Journal of Crop and Weed 7(2): 51-58 (2011)

Hydrological processes modeling using RBNN - a neural computing technique

A. SINGH, [1]M. IMTIYAZ, [1]R. K. ISAAC AND [1]D. M. DENIS

Regional Research Station (OAZ), Uttar Banga Krishi Viswavidyalaya, Majhian-733133, West Bengal
[1]Vauge School of Agricultural Engineering and Technology, Sam Higginbottom Institute of Agriculture, Technology and Sciences (Formerly AAIDI), Allahabad -211 007, Uttar Pradesh
Received:10.04.2011, Revised: 19.11.2011, Accepted : 25.11.2011

ABSTRACT

Severe erosion in the watersheds under Damodar Valley Corporations (DVC), Hazaribagh, Jharkhand, India has been taking place for a long time and several soil and water conservation measures are being adopted by the Soil Conservation Department under DVC. For effective planning of soil conservation programs, hydrologic models can always be of help. Radial basis neural network (RBNN) model is neural network model which requires lesser data. In the present study, one of the watersheds under DVC named Nagwa was selected for simulating surface runoff and sediment yield. Maximum and minimum daily temperature and rainfall were used as input for RBNN model training and validation for surface runoff and runoff was included when simulating for sediment yield. The RBNN model was trained for the year 1991-2000 and validated for the year 2005-2007. Results indicate that coefficient of determination (R^2), Nash-Sutcliffe simulation efficiency (NSE) and root mean square error (RMSE) values for SWAT model were found to be 0.74, 0.74 and 0.41mm during training and 0.65, 0.63 and 4.15 mm during validation period respectively. The model performed quite well for simulation of sediment yield with R^2, NSE and RMSE values of 0.77, 0.69 and 0.24t ha^{-1} during training period and 0.88, 0.82 and 0.65t ha^{-1} during validation period, respectively. It could be stated that RBNN model based on simple input could be used for estimation of monthly runoff, sediment yield, missing data, and testing the accuracy of other models.

Key words: ANN, radial basis neural network, runoff, sediment yield, simulation, modeling

Mathematical models have been applied to simulate hydrological processes during the last couple of decades. Hydrological models based on certain governing equations which define the various physical processes affecting the hydrologic behaviors of the watershed. These models are certainly good and have been validated throughout the world on various watersheds for simulating stream flow and sediment yield. They require considerably huge data comprising of physical nature of the watershed and which may be subjective also. On the other hand, artificial neural network (ANN) is black box model which does not need the information regarding the physical characteristics of the watershed. ANN models have gained considerable popularity due to its simplicity and robustness and nowadays been used in simulating hydrological processes. These are also called system theoretic models which establish a relationship between input and the output functions without considering the complex physical laws governing the natural process such as rainfall-runoff transformation. ANNs are basically based on the functioning of the biological processes of a human brain. As we learn the things in our day to day life and our actions are normally based on our past experiences. Similarly, ANNs are also trained with known input and output and then employed to predict the output with known input. For various complex nonlinear environmental problems, ANNs have an advantage over distributed parameter models due to the lesser data requirements and they are more suited for long-term forecasting (Mutlu *et al.*, 2008). Park and Sandberg (1993) proved that Radial Basis Function (RBF) networks with one hidden layer are capable of universal approximation. Fernando and Jayawardena (1998) reported that the RBF type network is found to perform better than feed forward network trained with back propagation algorithm. However, the application of RBF type neural networks to hydrological problems is still rare, but recently it is getting more attention due to its advantages over feed forward networks. There are numbers of applications of ANN modeling of hydrological processes due to its simplicity and flexible means (Maier and Dandy 1996; Coulibaly *et al.*, 2000; Persson *et al.*, 2001; Rajurkar *et al.*, 2002). American Society of Civil Engineers (ASCE) Task Committee on the Application of Artificial Neural Networks in Hydrology (ASCE, 2000a, b) has also made an exhaustive review investigating the role of ANNs in different fields of hydrology. Campolo *et al.* (1999) applied the distributed rainfall data for the prediction of water levels at the catchment outlet. The model performed better when the water levels observed in the recent past were also used as input along with the rainfall data. Zhang and Govindaraju (2000) used a modular neural network for the prediction of catchment runoff, and utilized Bayesian concepts in deriving the training algorithm. The type of data that should be included as input for modeling purpose using ANN, largely depends upon the availability and response of the model in terms of error. Kisi *et al.* (2008) found radial basis neural network better than other two ANN models such as feed-forward neural network and generalized regression neural network in an investigation to improve the accuracy of the stream flow-suspended sediment rating curve for daily suspended sediment estimation. Agrawal *et al.* (2009) used the back propagation artificial neural network modeling technique to forecast the runoff and sediment yield from an Indian watershed during the monsoon period.

Email: ajai_jpo@yahoo.com

Cobaner *et al.* (2009) compared an adaptive neuro-fuzzy approach with three different ANN techniques, namely, the generalized regression neural networks, radial basis neural networks and multi-layer perceptron and two different sediment rating curves. Many researchers have applied hydrological models to simulate the rainfall-runoff processes and sediment yield at the outlet of the Nagwa watershed, application of RBNN model does not exist in the peer reviewed journal. As such the present work on simulation of the hydrological processes at single outlet of the watershed using neural computing technique such as RBNN was undertaken.

MATERIALS AND METHODS

Study area

Nagwa watershed is located in the upper Damodar valley under the Damodar Valley Corporation, Hazaribagh, Jharkhand, India. The area of the watershed is approximately 92.46 km^2 of which about 30-40% is under shrubs & forest and the remaining under cultivation and other uses. The average elevation of the command is 540 m above mean sea level. The topography of the watershed is undulating with flat land in major parts. The slope of the watershed ranges from 1 to 16 % with an average of 2 %. The average annual rainfall of the area is 1200 mm of which more than 80% occurs during the monsoon months from June to October and the rest in the winter months (December and January). The daily temperature ranges from a maximum of 42.5^0 C (1st May, 1999) to a minimum of 2.5^0 C (18th January, 1999). The daily mean relative humidity varies from a minimum of 21.72% in the month of April to a maximum of 90.36% in the month of September. The overall climate of the area is classified as sub-humid sub-tropical. The study area comes under the Sub-humid tract of Eastern Plateau in the 12th agro-ecological zone of India. The texture of red and yellow soils of the watershed ranges from sandy loam to clay loam. Soil structures vary from moderately fine sub angular blocky to coarse sub angular blocky. The overall soils of the watershed are neutral to slightly acidic with medium organic matter and low salt content. Bulk density of the soils varies around 1.4 to 1.5 gcc^{-1} with moderately low saturated hydraulic conductivity ranging from 40 to 170mm day^{-1}. It is bounded by latitudes of 23^{0}59'08"N to 24^0 05'41"N and longitudes of 85^{0}16'35"E to 85^{0}23'45"E.

Radial basis neural networks (RBNN)

RBNNs are general purpose networks which can be used for a variety of problems including system modeling, prediction and classification. In general, an RBNN is any network which makes use of radially symmetric and radially bounded transfer functions in its hidden layer. The Euclidean distance is determined from the point being evaluated to the center of each neuron and a radial basis function is applied to the distance to compute the weight for each neuron. RBNN has an input layer, a hidden layer and an output layer. The neurons in the hidden layer contain Gaussian transfer functions whose outputs are inversely proportional to the distance from the center of the neuron. RBNNs are similar to K-Means clustering and have a variable number of neurons that is usually much lesser than the number of training points. The output of the RBNN is calculated as given hereunder (Mutlu *et al.,* 2008)

$$Y = \sum W \theta \left(\|X - C\|\right) \qquad \ldots(1)$$

where, X = input value, Y = output value, $\theta(\ )$ = radial basis function, W = weights connecting the hidden and output nodes, C = centre of hidden node, which depends on the observed input data and $\|X - C\|$ = Euclidean distance between the input and hidden nodes.

The network was optimized for number of nodes at hidden layer 2 for a given parameter of momentum rate, linear coefficient, numbers of nodes at prototype layer, learning rule and transfer function for RBNN model. The number of hidden layer neurons significantly influences the performance of a network. If this number is small, the network may not achieve a desired level of accuracy, while with too many nodes it will take a long time to get trained and may sometimes over fit the data. The change of hidden 2 nodes showed that 25 numbers of nodes at hidden layer yielded the lowest RMSE and highest correlation coefficient (R) values. However, there was a trend of decreasing R and increasing RMSE with increase of number of nodes. In case of RBNN, epoch level was set from the number of training data set and later it was varied from 8 to 64. There was no impact of changing the epoch level on the performance of the ANN model. It was fixed at 64 arbitrarily. Momentum rate was varied to determine the optimum rate while keeping other parameter as constant. RMSE and R values reached at peak when momentum rate was 0.40. Further, optimization was carried out with optimized number of nodes at hidden layer 2 as 25 and momentum rate as 0.40. Number of prototype layer nodes has a significant effect on performance of neural network in RBNN. Lowest RMSE and highest R were obtained at number of prototype layer nodes of 200. Similarly, linear coefficient of 0.60 yielded better performance. The highest R value was obtained in case of Delta-Rule.

Training of radial basis neural networks

The process of determining ANN weights is called learning or training and it is similar to calibration of a mathematical model. The RBNNs were trained with a training set of input of daily temperature and rainfall data and known output data of surface runoff. Later the daily output was aggregated to monthly data. At the beginning of training, the weights were initialized with a set of random values. The weights were systematically changed by the learning algorithm such that, for a given input, the difference between the ANN output and the actual output is small. Many learning examples are repeatedly presented to the network, and the process is terminated when this difference is less than a specified value. The root mean square error over the training samples was the objective function to be minimized. The first phase of training the network was a clustering phase. In this phase, the incoming weights to the prototype layer learnt to become the centers of clusters of input vectors. This clustering is done in NeuralWare Professional II/PLUS using a Dynamic K-Means algorithm. When the clustering phase finishes, the radii of the Gaussian functions at the cluster centers are set using a nearest neighbor procedure. The radius of a given Gaussian is set to the average distance to the two nearest cluster centers.

Model evaluation

The most widely used statistics reported for hydrologic calibration and validation are the coefficient of determination (R^2) and the Nash-Sutcliffe model efficiency (NSE) coefficient. Moriasi *et al.* (2007) recommended both graphical techniques and quantitative statistics to be used in model evaluation. Root mean square error (RMSE) is one of the error indices which are used in model evaluation. R^2 values range from 0 to 1. Higher values indicate less error variance and normally values greater than 0.5 are considered acceptable (Santhi *et al.,* 2001; Van Liew *et al.,* 2003). NSE ranges between $-\infty$ and 1.0. Values between 0.0 and 1.0 are generally considered as good. If it is found lesser than zero, it indicates that the mean observed value is a better predictor than the simulated value. The R^2, NSE and RMSE can be determined with the following equation.

$$R^2 = \left[\frac{\sum_{i=1}^{n} \left(Y^{obs} - \bar{Y}^{obs} \right)\left(Y^{sim} - \bar{Y}^{sim} \right)}{\sqrt{\sum \left(Y^{obs} - \bar{Y}^{obs} \right)^2} \sqrt{\sum \left(Y^{sim} - \bar{Y}^{sim} \right)^2}} \right] \quad ...(2)$$

where, where Y^{sim} are the simulated values , Y^{obs} are the observed values, $\bar{Y}^{obs}$ is the mean of observed values, and $\bar{Y}^{sim}$ is the mean of n simulated values.

...(3)

where Y_i^{obs} is the i^{th} observation for the constituent being evaluated, Y_i^{sim} is the i^{th} simulated value for the constituent being evaluated, Y^{mean} is the mean of observed data for the constituent being evaluated, and n is the total number of observations.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(Y^{sim} - Y^{obs} \right)^2}{n}} \quad ....(4)$$

RESULTS AND DISCUSSION

Surface runoff

The RBNN model was first optimized to achieve the lowest RMSE and greater correlation coefficient. The optimized neural network was applied to study the prediction of surface runoff. Several trial and error iterations were made to achieve the goal to be within the range of statistical parameters. Monthly observed and RBNN simulated values of runoff was plotted for the calibration period as shown in fig. 1 and R^2 value of 0.74 in fig. 2 shows that the RBNN model predicted accurately when compared to observed counterparts. NSE and RMSE values were obtained as 0.74 and 0.41 mm which demonstrate a very close match in almost all the training years between observed and simulated values of surface runoff (Table 1).

Table 1: Performance of RBNN model for monthly surface runoff and sediment yield during calibration and validation period

Calibration	Surface run off	Sediment yield
R^2	0.74	0.77
NSE	0.74	0.69
RMSE	0.41 mm	0.24 t ha^{-1}
Validation		
R^2	0.65	0.88
NSE	0.63	0.82
RMSE	4.15 mm	0.65 tha^{-1}

The trained RBNN neural network was applied to validate the model for a small watershed in Eastern India. The model was validated for June to October, 2005 to 2007. The scatter plot for the validation period has been shown in fig. 4 and indicates the coefficient of determination, R^2 as 0.65 which further demonstrates that model predicted closely with the observed values of surface runoff. Monthly observed and RBNN simulated values of runoff was plotted for the validation period and model simulated values follow the trend of observed values as shown in fig. 3. This demonstrates that the RBNN model predicted accurately when compared to observed counterparts. Thus, in an RBNN, for any point in the input space the response of the closest basis function plays a major role in the output of the network. RBNNs do not have any associated connections between input hidden nodes, which give a weighted input to each hidden node before the nonlinear transformation takes place. Consequently, a trained RBNN network's output will be accurate only if the input pattern falls close to the centre of the basis function.

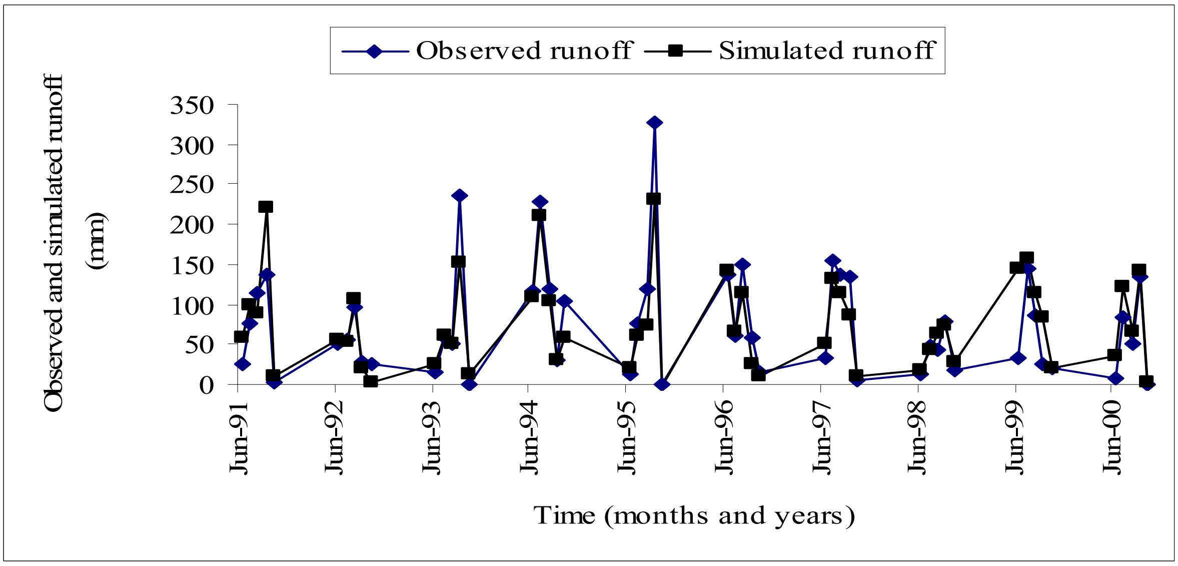

Fig. 1. Performance of radial basis neural network model for observed and simulated runoff (mm) during calibration period (1991-2000)

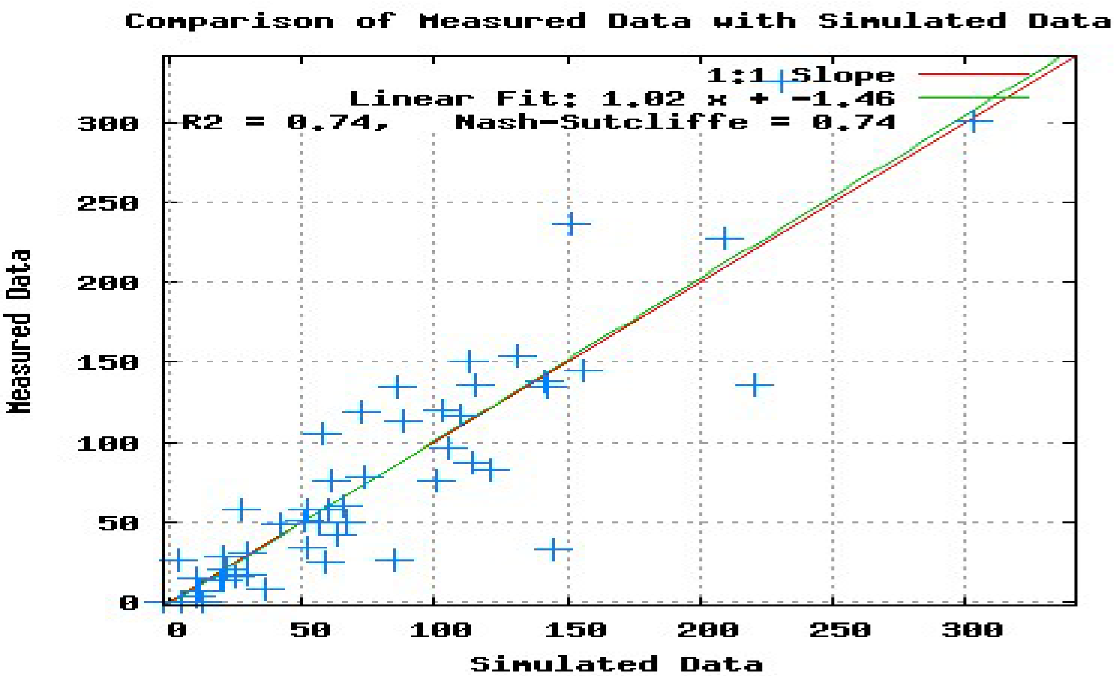

Fig. 2. Scattergram of observed and RBNN simulated monthly surface runoff during calibration period

Sediment yield

The major inputs for predicting sediment yield were the maximum and minimum temperature daily rainfall and surface runoff which were fed to develop neural network The RBNN model was trained during years 1991-2000 for simulating sediment yield. The monthly observed and simulated sediment yields have been compared graphically in fig 5. The trend of simulated monthly sediment follows quite well to the observed sediment yield during training period.

The model under predicted in the year receiving high rainfall. The overall prediction of the monthly sediment yield during the whole calibration period was in close agreement with its observed values. Scatter plot between observed and predicted sediment yield data serve a strong basis to assess a model's accuracy. The closer the scatter points are to the line of the best fit, the better the model. The scatter plot between the observed and simulated monthly sediment yield along with the regression line is presented in fig. 6. The figure shows an even distribution of the simulated values about regression line for both lower and higher measured values. The R^2, NSE and RMSE values were found to be as 0.77, 0.69 and 0.24 tha^{-1} during training period. The trained model was validated for 2005 to 2007 for sediment yield loss from a small watershed. Monthly observed and RBNN simulated values of sediment yield were plotted for the validation period and model simulated values follow the trend of observed values as shown in fig. 7. This demonstrates that the RBNN model predicted satisfactorily when compared to observed sediment yield from the outlet of the watershed. Statistical analysis of observed and RBNN simulated monthly sediment yield (tha^{-1}) for the validation period has been presented in table 1. The scatter plot for the validation period (2005-2007) has been shown in fig. 8 and indicates the coefficient of determination, R^2 as 0.88 which demonstrate that model predicted closely with the observed values of sediment yield.

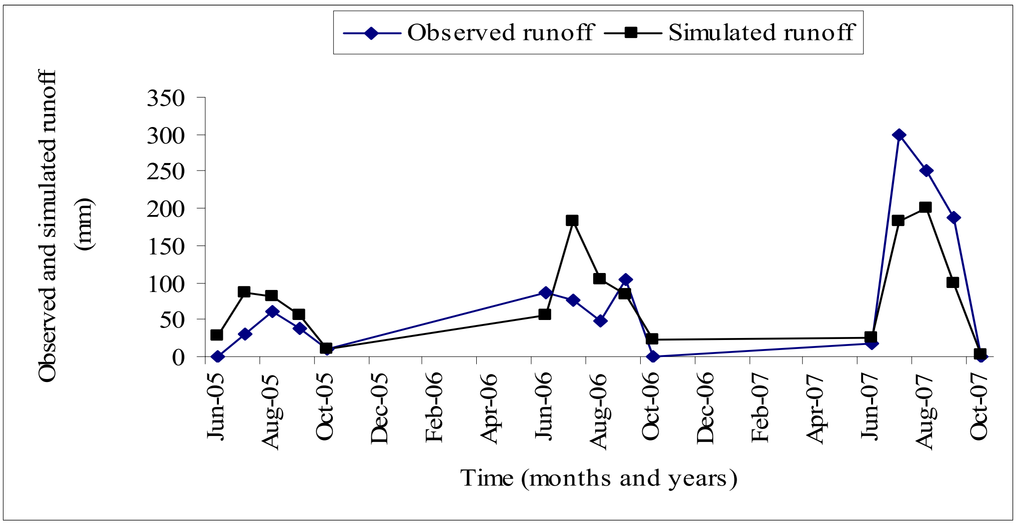

Fig. 3. Performance of radial basis function neural network for observed and simulated monthly runoff (mm) for model validation period (2005-2007)

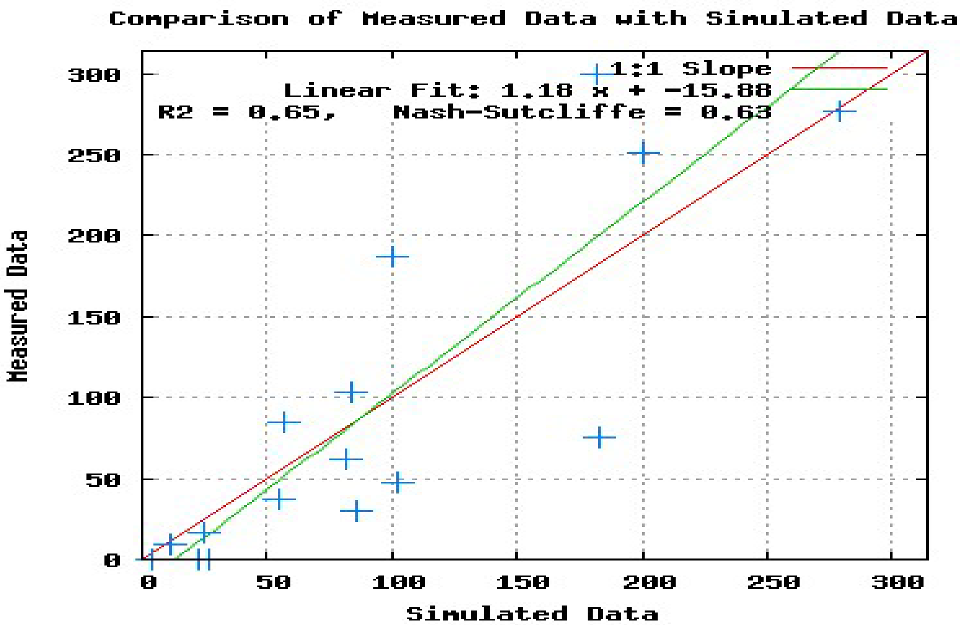

Fig. 4. Scattergram of observed and RBNN simulated monthly surface runoff during validation period

In the present study, an attempt was made to calibrate and validate the RBNN model, a data driven neural network model for surface runoff and sediment yield for a watershed where erosion and water quality problems exists. Monthly simulations for surface runoff and sediment yield show good agreement between measured and simulated counterparts. As RBNNs have the ability to continuously learn from the previous data and can result quick model development, they are particularly suited for modeling nonlinear systems where traditional parameter estimation techniques are not convenient or physical characteristics of watershed are not available. It could be stated that RBNN model based on simple input could be used for estimation of monthly runoff and sediment yield, missing data, and testing the accuracy of other models. RBNN model being the artificial neural network model lacks the spatial characteristics. For studying simulation at only single outlet of any watershed, RBNN model could be employed as an alternative model.

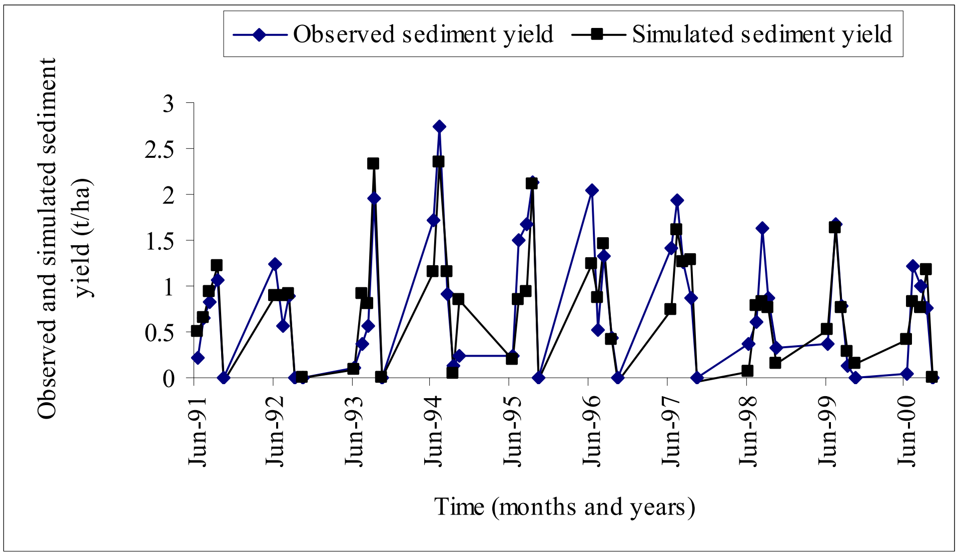

Fig. 5. Performance of radial basis function neural network for observed and simulated monthly sediment yield (tha^{-1}) for model calibration period (1991-2000)

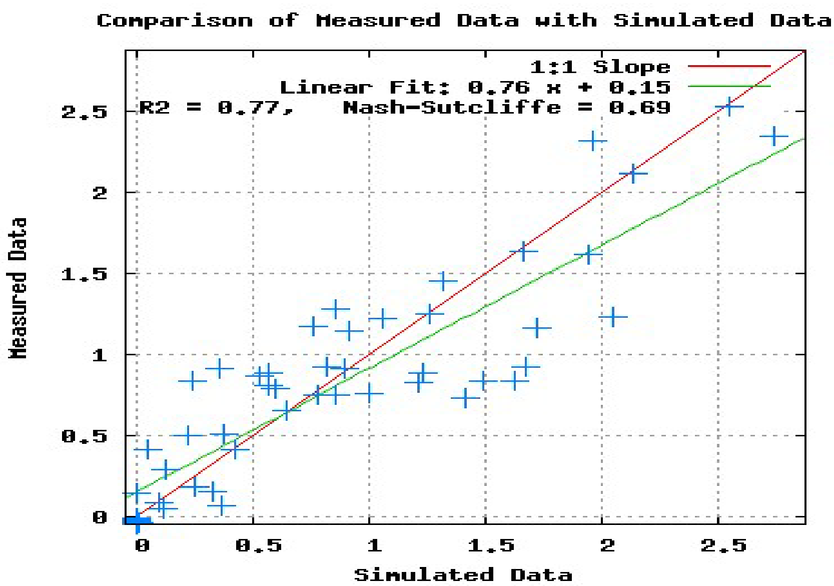

Fig. 6. Scattergram of observed and RBNN simulated monthly sediment yield during calibration period

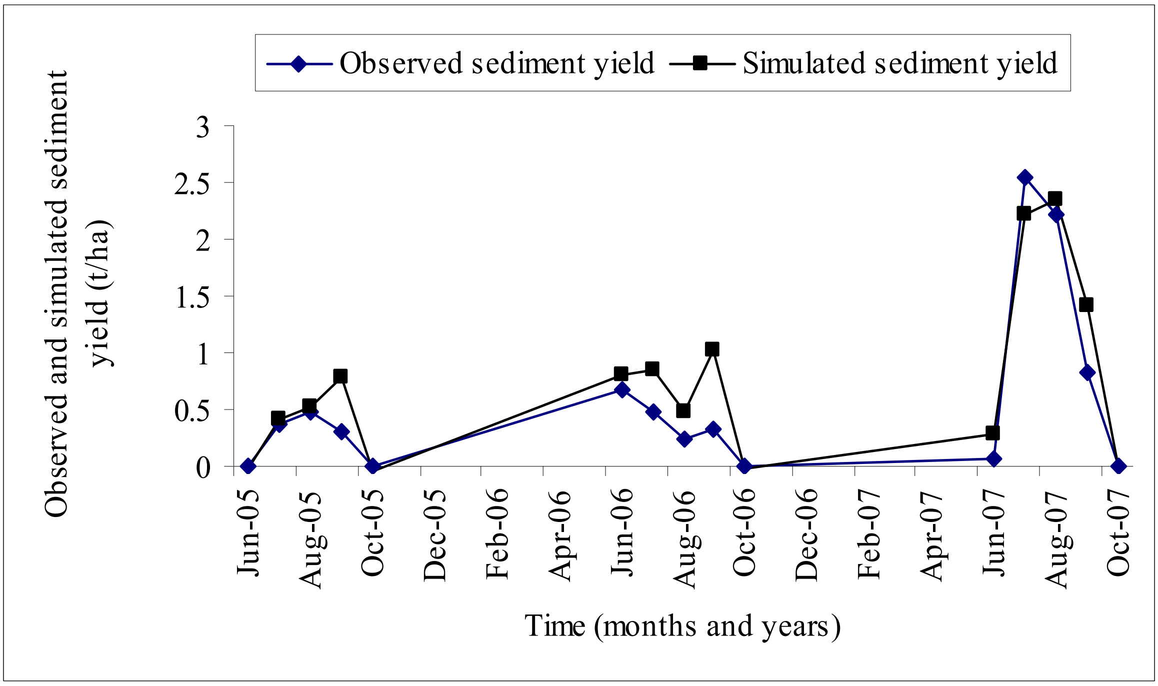

Fig. 7. Performance of radial basis function neural network for observed and simulated sediment yield for model validation period (2005-2007).

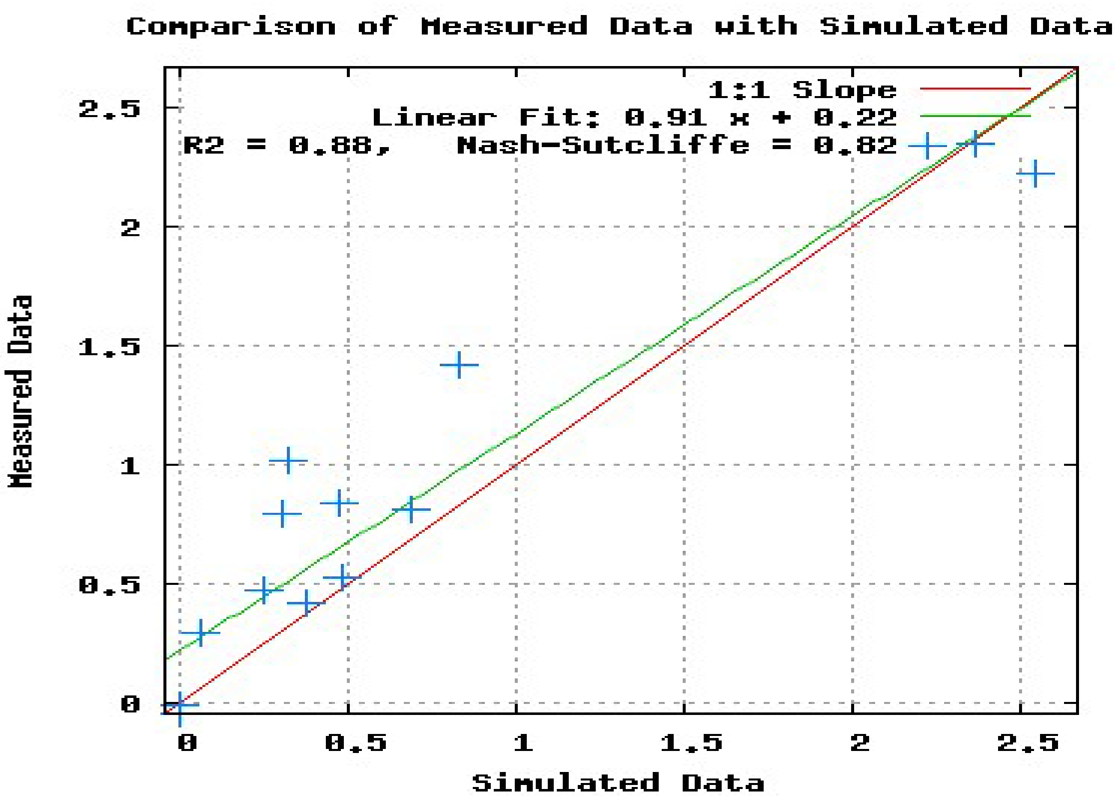

Fig. 8. Scattergram of observed and RBNN simulated monthly sediment yield during validation period

REFERENCES

Agrawal, A., Rai, R. K. and Upadhyay, A. 2009. Forecasting of runoff and sediment yield using artificial neural networks, *J. Water Resource and Protection*, **1**:368-75.

American Society of Civil Engineers (ASCE). 2000a. Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial neural networks in hydrology. I: preliminary concepts, *J.Hydrol. Eng. ASCE*, **5**: 115-23.

American Society of Civil Engineers (ASCE). 2000b.Task Committee on Application of Artificial Neural Networks in Hydrology. Artificial neural networks in hydrology. II: hydrologic applications, *J. Hydrol. Eng. ASCE*, **5**: 124-37.

Campolo, M., Andreussi, P. and Soldati, A. 1999. River flood forecasting with a neural network model. *Water Resour. Res.,* **35**: 1191-97.

Cobaner, M., Unal, B. and Kisi, O. 2009. Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *J. Hydrol.,* **367**: 52-61.

Coulibaly, P., Anctil, F., Bobee, B. 2000. Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. *J. Hydrol*., **230**: 244-57.

Fernando, A. K. and Jayawardena, A. W.1998. Runoff forecasting using RBF networks with OLS algorithm. *Journal of Hydrologic Engineering, ASCE*, **3**: 203–09.

Kisi, O., Ibrahim, Y. and Emrah, D. 2008. Modelling daily suspended sediment of rivers in Turkey using several data-driven techniques. *Hydrological Sci. J.*, **53**: 1270 – 85.

Maier, H. R. and Dandy, G. C. 1996. The use of artificial neural networks for prediction of water qualitv parameters. *Wat. Resour. Res*., **32**: 1013-22.

Moriasi, D. N., Arnold, J. G, Van Liew, M. W., Binger, R. L., Harmel, R. D. and Veith, T. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Trans. ASABE*, **50**:885-00.

Mutlu, E., Chaubey, I., Hexmoor, H. and Bajwa, S. G. 2008. Comparison of artificial neural network models for hydrologic predictions at multiple gauging stations in an agricultural watershed. *Hydrol. Proc.,* published online in at www.interscience.wiley.com, DOI: 10.1002/hyp.7136.

Park, J. and Sandberg, I. W. 1993. Universal approximations using radial basis function networks. *Neural computation*, **3**:246-57.

Persson, M, Berndtsson, R. and Sivakumar, B. 2001. Using neural networks for calibration of time domain reflectometry measurements. *Hydrol. Sci. J.,* **46**: 389-98.

Rajurkar M P, Kothyari UC and Chaube UC. 2002. Artificial neural networks for daily rainfall-runoff modelling. *Hydrol. Sci. J.,* **47**: 865-77.

Santhi, C., Arnold , J. G., Williams, J. R., Dugas, W. A. and Hauck L. 2001. Validation of the SWAT model on a large river basin with point and non point sources, *J. American Water Resour. Assoc.*, **37**:1169–88.

Van Liew, M. W. and Garbrecht, J. 2003. Hydrologic simulation of the Little Washita River experimental watershed using SWAT. *J. Amer. Water Resour. Assoc.*, **39**: 413-26.

Zhang, B. and Govindaraju, R. S. 2000. Prediction of watershed runoff using Bayesian concepts and modular neural networks. *Wat. Resour. Res.*, **36**: 753-62.