Estimation of Suspended Sediment Concentration using Artificial Neural Networks

A. R. Senthil Kumar, Avinash Agarwal, R. D. Singh and R. K. Nema

National Institute of Hydrology, Roorkee-247667
e-mail : arsk@nih.ernet.in

ABSTRACT

The assessment of the sediment volume transported by river water is very important in the design and management of water resources project. Several methods have been proposed to predict suspended sediment concentration based on the properties of flow and sediment. The equations proposed by the investigators for the estimation of sediment concentration have the limitations due to the simplification of important parameters and boundary conditions. Recently, neural networks approach has been applied in many areas of water resources due to its capability in representing any nonlinear processes by given sufficient complexity of the trained networks. In this paper, development of an Artificial Neural Network (ANN) model for predicting the suspended sediment concentration at the upstream of Bhakra reservoir is presented. The results of the ANN model for calibration indicated that the all range of sediment concentration values were simulated fairly well. Whereas the high range values of sediment concentration were slightly deviated from the observed values during the validation of the model. The performance of ANN model was compared with multiple linear regression model (MLR) and was found better than the MLR model.

INTRODUCTION

Misuse and mismanagement of the catchment area contribute higher sediment load in the river reaches (Tejwani, 1984). Estimation of sediment concentration has become important in river engineering practices, river training, river management and the design of hydraulic structures mainly power plants. Another important use is in water quality monitoring (Nagy et al., 2002). Several methods have been proposed to predict suspended sediment concentration based on the properties of flow and sediment. The equations proposed by the investigators are based on empirical, dimensional approach and semitheoretical approach (Garde and Rangaraju, 1985; Graf, 1984; Rijn, 1984 a,b,c). The applicability of these equations to all rivers is limited due to the simplification of important parameters and boundary conditions considered (Nagy et al., 2002). Recently, neural networks approach has been applied in many areas of water resources due to its capability in representing any nonlinear processes by given sufficient complexity of the trained networks (Maier and Dandy, 2000). ANNs are proven to produce improved performance over other black box models in numerous hydrological studies (Hsu et al., 1995). The applications of artificial neural networks (ANN) in modelling the suspended sediment concentration are reported by Nagy et al., 2002 and Tayfur et al., 2003, Kisi,2004, Kisi, 2005. Nagy et al. selected the input neurons based on properties of flow and sediment. The properties considered are tractive shear stress, velocity ratio, suspension parameter, longitudinal slope, water depth ratio, Froude number, Reynolds number, stream width ratio, depth scale ratio and mobility number. The results of the ANNs model were compared with the results of the conventional models. Tayfur et al developed the ANNs model for predicting suspended sediment concentration by considering the rainfall intensity and slope as input neurons and compared with the performance of fuzzy model. It is concluded that ANNs model the sediment movement better than the other formulae including regression model. The main advantage of the ANN models over traditional models is that it does not require information about the complex nature of the underlying process under consideration to be explicitly described in mathematical form. This paper discusses the development of a model for the prediction of suspended sediment concentration using ANN at Kasol gauging site of river Sutlej.

THE STUDY AREA

The catchment of Sutlej river up to Kasol is considered for analysis. The study area is located upstream of Bhakra reservoir. Bhakra reservoir is very big reservoir located on the Sutlej river with a water spread area of 168.35 SQ. KM at full reservoir elevation of 515.11m. The original designed capacity of the reservoir at dead storage, live storage are 2431.806 million m^3 and 7436.034 million m^3 respectively. The capacity of reservoir in the year 1994 at dead storage, live storage are 1805.610 million m^3 and 6779.683 million m^3 respectively. The catchment area up to Kasol is 56980sq.km and is presented in Fig. 1. The impounding of water was started in the year 1959. The power production from the reservoir is 1050 MW. 35 lakh hectares of land is irrigated under this reservoir.

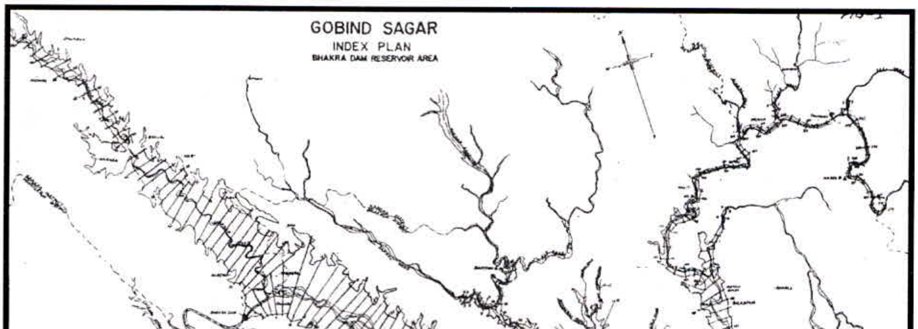

Fig. 1. Index map of Bhakra Reservoir

ANN – AN OVERVIEW

ANNs are a form of computing inspired by the functioning of the brain and nervous system and are discussed in detail in a number of hydrologic papers. For example, Portugal, 1995; Minns and Hall, 1996; See et al, 1997; Danh et al, 1998; Zealand et al, 1999; ASCE, 2000a,b; Maier and Dandy, 2000; Elshorbagy, 2000. The architecture of a feed forward ANN can have many layers where a layer represents a set of parallel neurons. The basic structure of ANN 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 outputs are produced. The neurons in the layers are interconnected by strength called weights. A typical three-layered feed forward ANN is shown in Fig. 2

In general, a neuron can have n inputs, labeled from 1 through *n*. For example neuron 3 in the hidden layer shown in Fig. 2, n= 2. In addition, each neuron has an input that is equal to 1.0, called *bias*. Each neuron *j* receives information from every node *i* in the pervious layer. A weight (w_{ji}) is associated with each input (x_i) to node *j*. The effective

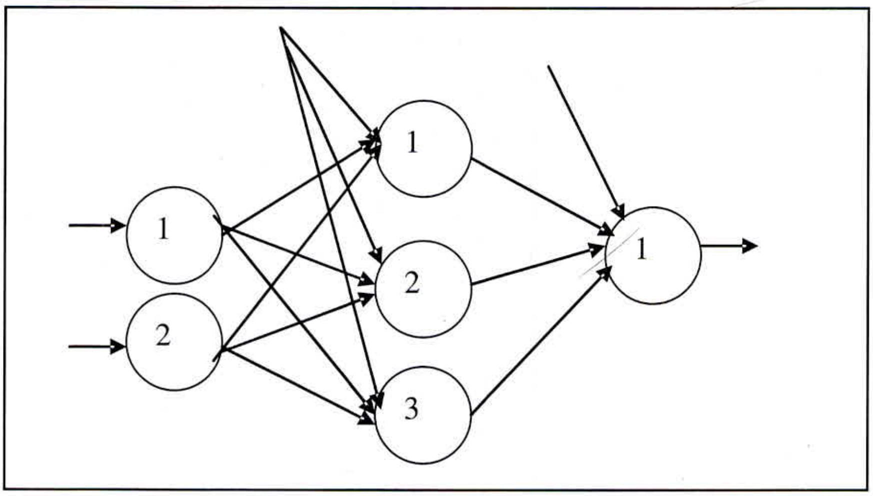

Fig. 2 : A Typical Three-Layer Feed Forward ANN (ASCE, 2000a)

incoming information (NET_j) to node *j* is the weighted sum of all incoming information, otherwise known as the net input, and is computed as:

$$NET_j = \sum_{i=0}^{n} w_{ji} x_i \qquad (1)$$

where x_0 and w_{j0} are called as the bias term ($x_0 = 1.0$) and the bias respectively. Equation 1 applies to the nodes in the output layer and hidden layer(s). The weighted sum of input information is passed through an activation function, called transfer function, to produce the output from the neuron. The transfer function introduces some nonlinearity in the network, which helps in capturing the nonlinearity present in the function being mapped. The commonly employed transfer function is the sigmoid function (ASCE, 2000a) and is given as follows:

$$OUT_j = \frac{1}{1 + e^{-NET_j}} \qquad (2)$$

The interconnected weights are adjusted using a learning algorithm such that the output from the ANN model is very close to the observed values by minimizing the error through a mathematically formulated procedure. This procedure is called training of network.

Using a set of examples from a given problem domain, comprising inputs and their corresponding outputs, an ANN model can be trained to learn the relationship between the input-output pairs. The feed forward ANN is generally adapted in all studies because of its applicability to a variety of different problems (Hsu et al., 1995). However, there are no guidelines in developing an effective ANN architecture, though some researchers have reported suggestions that can be implemented while developing an ANN model. For instance, Maier and Dandy (2000) report that not more than one hidden layer is required in feed forward networks because a three-layer network can generate arbitrarily complex decision regions. Also, the appropriate input vector to the ANN model can be identified according to the procedure of Sudheer et al. (2000).

The input values should be normalised to the range between 0 and 1 before passing into a neural network since the output of sigmoidal function is bound between 0 and 1. Minns and Hall, (1996), Dawson and Wilby (1998), Sajikumar and Thandaveswara (1999), and Burian et al (2001) emphasised the importance of the normalisation of data and gave the procedure to normalise. The output from the ANN should be denormalised to provide meaningful results. In this study, following equation is used to normalize the data set:

$$N_i = \frac{R_i - Min_i}{Max_i - Min_i} \qquad (3)$$

where R_i is the real value applied to neuron *i*; N_i is the subsequent normalized value calculated for neuron *i*; Min_i is the minimum value of all values applied to neuron *i*; Max_i is the maximum value of all values applied to neuron *i*.

Training a network is a procedure during which an ANN processes training set (input-output data pairs) repeatedly, changing the values of its weights, according to a predetermined algorithm and the environment in which the network is embedded. The main objective of training (calibrating) a neural network is to produce an output vector $Y = (y_1, y_2, \ldots, y_p)$ that is as close as possible to the target vector (variable of interest or forecast variable) $T = (t_1, t_2, \ldots, t_p)$ when an input vector $X = (x_1, x_2, \ldots, x_p)$ is fed to the ANN. In this process, weight matrices W and bias vectors V are determined by minimizing a predetermined error function as explained as follows:

$$E = \sum_P \sum_p (y_i - t_i)^2 \qquad (4)$$

where t_i is a component of the desired output T; y_i is the corresponding ANN output; p is the number of output nodes; and P is the number of training patterns.

Back propagation is the most popular algorithm used for the training of the feed forward ANNs (Hsu et al, 1995; Dawson and Wilby, 1998; Thirumalaiah and Deo, 1998; Sajikumar and Thandaveswara, 1999; Tokar and Jhonson, 1999; Zealand et al, 1999; Thirumalaiah and Deo, 2000; ASCE, 2000a; Elshorbagy et al, 2000; Maier and Dandy, 2000; Burian et al, 2001). Each input pattern of the training data set is passed through the network from the input layer to output layer. The network output is compared with the desired target output, and an error is computed based the equation 4. This error is propagated backward through the network to each neuron, and the connection weights are adjusted based on the equation

$$\Delta w_{ij}(n) = -\varepsilon * \frac{\partial E}{\partial w_{ij}} + \alpha * \Delta w_{ij}(n-1) \qquad (5)$$

where $\Delta w_{ij}(n)$ and $\Delta w_{ij}(n-1)$ are weight increments between node *i* and *j* during *n*th and *(n-1)*th pass, or epoch (ASCE, 2000a). A similar equation is written for correction of bias values. In the equation 5, e and a are called learning rate and momentum respectively. The momentum factor can speed up training in very flat regions of the error surface and help prevent oscillations in the weights. A learning rate is used to increase the chance of avoiding the training process being trapped in a local minima instead of global minima. The literature by Rumelhart et al, 1986 can be referred for the details of the algorithm.

PERFORMANCE EVALUATION OF ANN MODEL

The whole data length is divided into two, one for calibration (training) and another for validation of artificial neural network model. The performance during calibration and validation is evaluated by performance indices such as root mean square error (RMSE), model efficiency (Nash and Sutcliffe, 1970) and coefficient of correlation (R). They are defined as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^{K} (t-y)^2}{K}} \tag{6}$$

$$\text{Efficiency} = 1 - \frac{\sum (t-y)^2}{\sum (t-\bar{t})^2} \tag{7}$$

$$\text{Coefficient of Correlation} = \frac{\sum TY}{\sqrt{\sum T^2 \sum Y^2}} \tag{8}$$

where K is the number of observations; t is the observed data; y is computed data; $T = t - \bar{t}$ in which $\bar{t}$ is the mean of the observed data; and $Y = y - \bar{y}$ in which $\bar{y}$ is the mean of the computed data.

MODEL DEVELOPMENT

For the current application, the daily rainfall values for Kalpa, Rackchham, Kaza, Namagia, Rampur, Suni, Berthin, Kahu, Kasol and Bhakra were available from 1987 to 2000. The discharge and suspended sediment concentration values at Kasol for the same period were also available. The ANN and MLR models had been developed for predicting the suspended sediment concentration at Kasol using the available data. The details of the model development are described in the following sections.

SELECTION OF INPUT

The ANN model for the prediction of suspended sediment concentration is normally developed using the antecedent rainfall, sediment concentration and discharge values of upstream stations as input vector. Determining the number of antecedent rainfall, sediment concentration and discharge values involves finding the lags of rainfall, sediment concentration and discharge values that have significant influence on the predicted sediment concentration. These influencing values corresponding to different lags can be very well established through statistical analysis of the data series. The input vector is selected generally by trial and error method; however, Sudheer et al. (2002) have presented a statistical procedure that avoids the trial and error procedure. They reported that the statistical parameters such as auto correlation function (ACF), partial auto correlation function (PACF) and cross correlation function (CCF) can be used for this purpose. The PACF of the sediment concentration series at Kasol with 95 % confidence levels and CCF of sediment concentration series at Kasol between rainfall of all stations upto kasol and discharge series at Kasol suggest the input vector to the ANN model. Based on PACF and CCF of the data series, the following input vector was selected for neural network training.

$$SC_{Kasol,t} = f(SC_{Kasol,t-1}, INFL_{Kasol,t-1}, INFL_{Kasol,t}, R_{Kasol,t-1}, R_{Rampur,t-1,} R_{Suni,t-1}) \quad (9)$$

In which SC, INFL and R are suspended Sediment concentration, discharge and rainfall value respectively.

MODEL TRAINING

The ANN models had been trained using back propagation algorithm. The whole data set were divided into two sets for the training and validation purpose of the ANN model. The data from 1991 to 2000 were considered for the training of the model since it contained the extreme values of sediment concentration. The data from 1987 to 1990 were considered for the validation of the model. The software used for the training of the model was MATLAB (The Mathworks, Inc., 2001). The number of the hidden neurons in the hidden layer was found by a trail and error procedure, and a number of 5 neurons is found to be optimum. Multiple Linear Regression (MLR) model had been developed for the prediction of suspended sediment concentration using the data considered in the development of ANN model. The performance of ANN and MLR models were evaluated based on the performance indices.

RESULTS AND DISCUSSIONS

The performance of best ANN model for the prediction of suspended sediment concentration during calibration and validation is presented in Figures 3 and 4 respectively along with the corresponding observed suspended sediment concentration. The visual inspection of the plots clearly demonstrates the potential of the developed ANN model in prediction of the suspended sediment concentration at Kasol. The results were further analyzed using statistical indices too. The results of the calibration and validation of the ANN and MLR models in terms of various statistical indices are presented in the Table 1.

The performance of MLR model for the prediction of suspended sediment concentration during calibration and validation is presented in Figures 5 and 6 respectively along with the corresponding observed suspended sediment concentration

The high coefficient of correlation during the calibration and validation for the prediction of suspended sediment concentration indicates that explained variance was high and the developed ANN model is good to estimate suspended sediment concentration with less error. The coefficient of correlation, RMSE, Model efficiency of

MLR model during calibration were lower than the values of ANN model. The performance of MLR model during validation was only slightly lower than the performance of ANN model. The RMSE of ANN, which is a measure of the residual variance, during calibration was very low compared to the value during the validation. The same trend was not observed in the performance of MLR model. The model efficiency of MLR was deteriorated during calibration as well as validation. The relative error in peak, which is a measure of model's ability to predict peak suspended sediment concentration, was very low during calibration of the ANN model. It was observed high during the validation of ANN model. The percentage error in peak suspended sediment concentration of MLR was high during calibration as well as validation. But overall performance of ANN model was better than MLR Model.

CONCLUSIONS

In this study, an ANN Model had been developed for predicting the suspended sediment concentration at Kasol gauging site, Sutlej River Basin. The ANN model had been developed using daily rainfall for basin up to Kasol and daily discharge and suspended sediment concentration data at Kasol from 1987 to 2000. The statistical parameters ACF, PACF and CCF had been used for selection of Input vector. The model

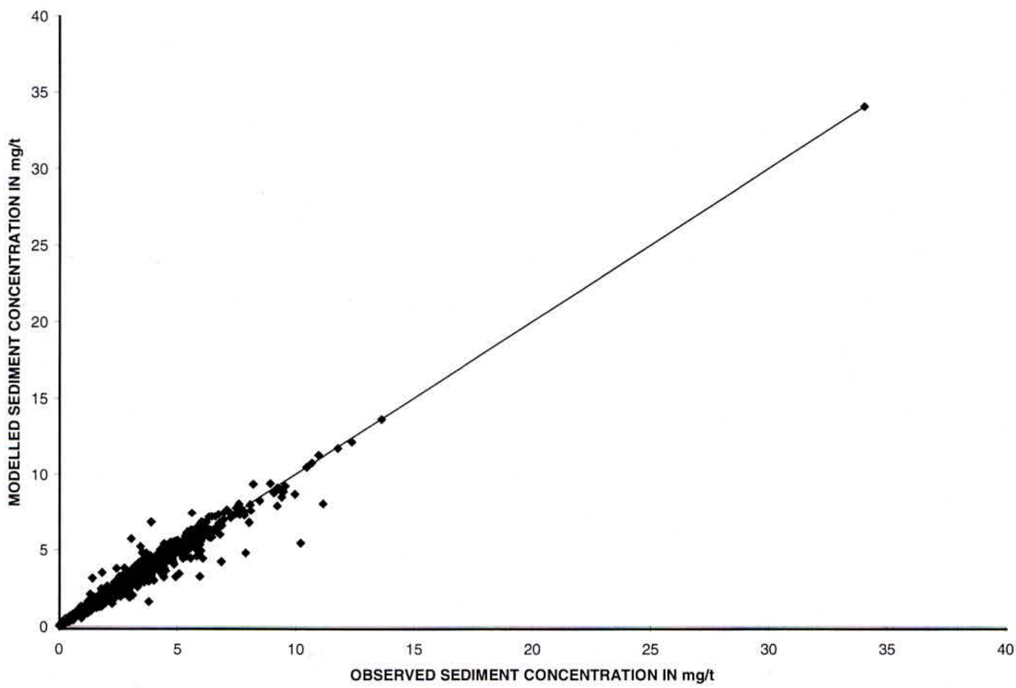

Fig. 1 : Scatter plot of observed Vs modeled sediment concentration for ANN calibration

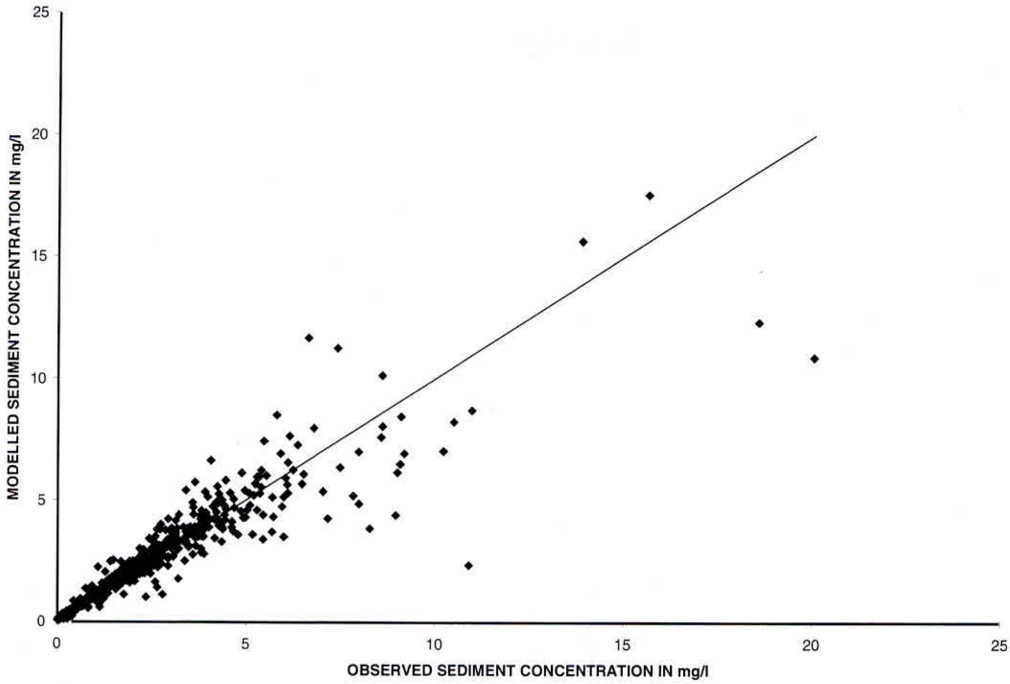

Fig 2 : Scatter plot of observed Vs modeled sediment concentration for ANN validation

Table 1 : Comparison of results between ANN model and MLR

	ANN model		MLR model	
	Calibration	Validation	Calibration	Validation
Coefficient of Correlation	0.9915	0.9500	0.9562	0.9450
RMSE	0.2344	0.5821	0.5293	0.5926
Model efficiency	0.9825	0.9025	0.9410	0.8915
% error in peak suspended sediment concentration	0.0084	-45.59	-59.46	-46.86

performance evaluation criteria used were coefficient of correlation, RMSE, model efficiency and percentage error of peak suspended sediment concentration. The best ANN model had been compared with MLR model. It is concluded from the validation and calibration results of the models that the ANN model performed better than MLR model.

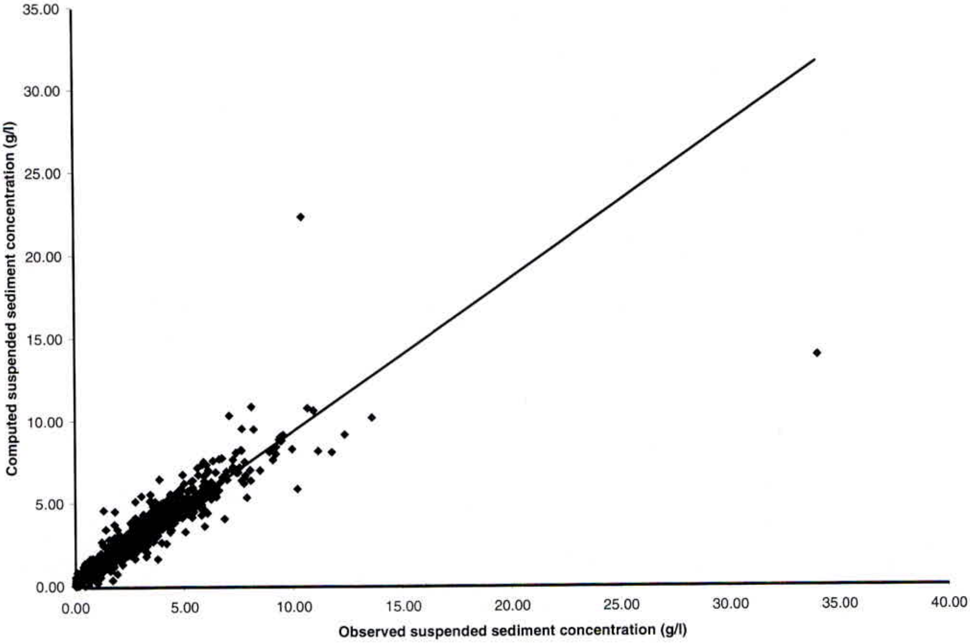

Fig. 3 : Scatter plot of observed Vs modeled sediment concentration for MLR calibration

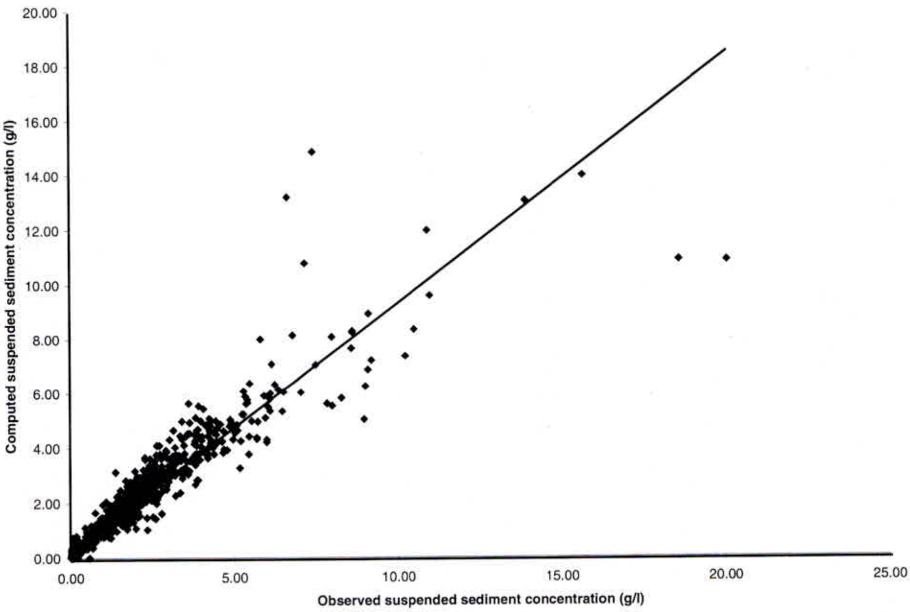

Fig 4 : Scatter plot of observed Vs modeled sediment concentration for MLR validation

REFERENCES

1. ASCE 2000a. Artificial neural networks in hydrology-I: Preliminary concepts, Journal of Hydrologic Engineering, ASCE, 5(2), pp. 115-123.
2. ASCE 2000b. Artificial neural networks in hydrology-II: Hydrologic applications, Journal of Hydrologic Engineering, ASCE, 5(2), pp. 124-137.
3. Atiya, A.F., M.E. Suzan, I.S. Samir. and S.E.Mohamed. 1999. A comparison between neural-network forecasting techniques-case study: river flow forecasting, IEEE Transactions on Neural Networks, 10(2), pp. 402-409.
4. Burian, S.J., S.R. Durrans, S.J. Nix. and R.E. Pitt. 2001. Training artificial neural networks to perform rainfall disaggregation, Journal of Hydrologic Engineering, ASCE, 6(1), pp. 43-50.
5. Central Water Commission. 1989. Manual on Flood Forecasting, River Management Wing, R. K. Puram, New Delhi.
6. Danh, N.T., H.N. Phien. and A.D. Gupta. 1999. Neural network models for river flow forecasting, Water SA, 25(1), pp. 33-39.
7. Dawson,C.W. and R. Wilby. 1998. An artificial neural network approach to rainfall-runoff modeling, Hydrological Sciences Journal, 43(1), pp. 47-66.
8. Elshorbagy, A., S.P. Simonovic. and U.S. Panu. 2000. Performance evaluation of artificial neural networks for runoff prediction, Journal of Hydrologic Engineering, 5(4), pp. 424-433.
9. Fernando, A.K. and A.W. Jayawardena. 1998. Runoff forecasting using RBF networks with OLS algorithm, Journal of Hydrologic Engineering, ASCE, 3(3), pp. 203-209.
10. Garde, R. J., and K. G. Rangaraju. 1985. Mechanics of sediment transportation and alluvial stream problems, Wiley Eastern Limited, New Delhi.
11. Graf, W. H. 1984. Hydraulics of sediment transport, Water Resources Publications, P.O. Box 2481, Littleton, Colorado 80161, U.S.A.
12. Hsu, K-L., H.V. Gupta. and S. Sorooshian. 1995. Artificial neural network modeling of the rainfall-runoff process, Water Resources Research, 31(10), pp. 2517-2530.
13. Kisi, O. 2004. Multi-Layer perceptrons with Levenberg-Marquardt training algorithm for suspended sediment concentration prediction and estimation, Hydrological Sciences Journal, 49(6), pp. 1025-1040.
14. Kisi, O. 2005. Suspended sediment estimation using neuro-fuzzy and neural network approaches, Hydrological Sciences Journal, 50(4), pp. 683-696.
15. Maier, H.R. and G.C. Dandy. 2000. Neural networks for the prediction and forecasting of water resources variables: A review of modelling Issues and applications, Environmental Modelling & Software, 15, pp. 101-124.
16. Minns, A.W. and M.J. Hall. 1996. Artificial neural networks as rainfall runoff models, Hydrological Sciences Journal, 41(3), pp. 399-418.
17. Nagy, H. M., K. Watanabe. and M. Hirano. 2002. Prediction of sediment load concentration in rivers using artificial neural networks, Journal of Hydraulic Engineering, 128(6), pp. 588-595.

18. Nash, J.E. and J.V. Sutcliffe. 1970. River flow forecasting through conceptual models:1. A discussion of principles, Journal of Hydrology, 10, pp. 282-290.
19. Portugal, M.S. 1995. Neural networks versus time series methods: a forecasting experience, 14th International Symposium on forecasting, 12-15 June 1995, Stockholm School of Economics, Stockholm, Sweden.
20. Rijn, L. C. 1984. Sediment transport, Part I: Bed load transport, Journal of Hydraulic Engineering, 110(10), pp. 1431-1456.
21. Rijn, L. C. 1984. Sediment transport, Part II: Suspended load transport, Journal of Hydraulic Engineering, 110(11), pp. 1613-1641.
22. Rijn, L. C. 1984. Sediment transport, Part III: Bed forms and alluvial roughness, Journal of Hydraulic Engineering, 110(12), pp. 1733-1754.
23. Rumelhart, D. E., E. Hinton. and J. Williams. 1986. Learning internal representation by error propagation, Parallel Distributed Processing, Vol. 1, MIT Press, Cambridge, Mass., pp. 318-362.
24. Sajikumar, S. and B.S. Thandaveswara. 1999. A non-linear rainfall-runoff model using an artificial neural network, Journal of Hydrology, pp. 216, 32-55.
25. See, L., M. Dougherty. and S. Openshaw. 1997. Some initial experiments with neural network models of flood forecasting on the river Ouse, Proceedings of Gecomputation'97 & SIRC'97, University of Otago, New Zealand, pp. 15-22.
26. Sudheer, K. P., A.K. Gosain. and K.S. Ramasastri. 2002. A data-driven algorithm for constructing artificial neural network rainfall-runoff models, Hydrological Processes,16, pp. 1325-1330.
27. Sudheer, K. P., P. C. Nayak. and K.S. Ramasastri. 2003. Improving Peak Flow Estimates in Artificial Neural Network River Flow Models, Hydrological Processes, 17(1), pp. 677-686.
28. The MathWorks, Inc. 2001. ANN Toolbox User's Guide, 3 Apple Hill Drive, Natick, MA 01760-2098.
29. Tayfur, G., S. Ozdemir. and V.P. Singh. 2003. Fuzzy logic algorithm for runoff-induced sediment transport from bare soil surfaces, Advances in Water Resources, 26, pp. 1249-1256.
30. Tejwani, K.G. 1984. Reservoir sedimentation in India-Its causes, Control, and Future course of action, Water International, 9, pp. 150-154.
31. Thirumalaiah, K. and M.C. Deo. 1998. River stage forecasting using artificial neural networks, Journal of Hydrologic Engineering, ASCE, 3 (1), pp. 26-32.
32. Thirumalaiah, K. and M.C. Deo. 2000. Hydrological forecasting using neural networks, Journal of Hydrologic Engineering, ASCE, 5 (2), pp. 180-189.
33. Tokar, A.S. and A. Johnson. 1999. Rainfall-runoff modeling using artificial neural networks, Journal of Hydrologic Engineering, ASCE, 4(3), pp. 232-239.
34. Zealand, C.M., D.H. Burn. and S.P. Simonovic. 1999. Short term streamflow forecasting using artificial neural networks, Journal of Hydrology, 214, pp. 32-48.