



EXPLICIT NEURAL NETWORK IN SUSPENDED SEDIMENT LOAD ESTIMATION

Özgür Kişi*, Ali Aytek†

Abstract: Correct estimation of sediment volume carried by a river is very important for many water resources projects. Traditionally, artificial neural networks (ANNs) are used as black-box models without understanding what happens inside the box. The question is that, how anyone who may be unfamiliar with ANNs can apply this kind of models in any other study, while the model has not been formulated. This paper proposes an explicit neural network (ENN) formulation which is simple and can be used, by anyone who is even not familiar with ANNs, for modeling daily suspended sediment-discharge relationship. The daily streamflow and suspended sediment data from two stations on Tongue River in Montana are used as case studies. Two different sediment rating curves (SRC), multi-linear regression (MLR) and nonlinear regression (NLR) are also applied to the same data. The ENN estimates are compared with those of the SRC, MLR and NLR models. The root mean square errors (RMSE), mean absolute errors (MAE), correlation coefficient (R) and model efficiency (E) statistics are used to evaluate the performance of the models. The comparison results reveal that the suggested model performs better than the conventional SRC, MLR and NLR.

Key words: *Suspended sediment load, explicit neural networks, rating curves, modeling*

Received: November 23, 2012

Revised and accepted: November 27, 2013

1. Introduction

Modeling suspended sediment concentration is vital important for many water resource projects related with channel navigability, reservoir filling, hydroelectric-equipment longevity, river aesthetics, fish habitat and scientific interests (Kisi et al. 2006). The estimation of suspended sediment is enormously difficult because it is closely related to flow and the mechanism of their relationship is non-linear

*Özgür Kişi – Corresponding Author

Canik Basari University, Civil Engineering Department, Hydraulics Division, 55000 Samsun, Turkey, E-mail okisi@basari.edu.tr

†Ali Aytek

Gaziantep University, Civil Engineering Department, Hydraulics Division, 27310 Gaziantep, Turkey

and they have sophisticated interactions to each other (Sivakumar and Wallender, 2005).

Artificial neural networks (ANNs) have been successfully used in a number of diverse fields including water resources (Ozkan *et al.* 2011; Yalkin & Korkmaz, 2012; Balara *et al.* 2013). In the hydrological forecasting context, recent experiments have reported that the ANNs may offer a promising alternative for rainfall-runoff modeling (Shamseldin, 1997; Tokar & Johnson, 1999; Solomatine & Dulal, 2003; Giustolisi & Laucelli, 2005; Nourani *et al.* 2009; 2011), streamflow prediction (Kang *et al.*, 1993; Chang & Chen, 2001 ; Campolo & Soldati, 2003; Cigizoglu, 2003; Cigizoglu, 2005a, 2005b; Kisi, 2004a, 2005a; Hu *et al.*, 2005; Cigizoglu & Kisi, 2005; Jayawardena *et al.*, 2006; Kisi & Cigizoglu, 2007; Zhang *et al.*, 2011; Yilmaz *et al.* 2011), reservoir inflow forecasting (Saad *et al.*, 1996; Jain *et al.*, 1999; Bae *et al.*, 2007), and suspended sediment estimation (Jain, 2001; Tayfur, 2002; Cigizoglu, 2004; Kisi, 2004b; 2005b; Cigizoglu, 2006; Cigizoglu & Alp, 2006; Cigizoglu & Kisi, 2006; Kisi, 2008; Kisi *et al.*, 2012; Lafdani *et al.* 2013; Liu *et al.* 2013). Jain (2001) used a single ANN approach to establish sediment-discharge relationship and found that the ANN model could perform better than the rating curve. Tayfur (2002) developed an ANN model for sheet sediment transport and indicated that the ANN could perform as well as, in some cases better than, the physically-based models. Cigizoglu (2004) investigated the accuracy of a single ANN in estimation and forecasting of daily suspended sediment data. Kisi (2004b) used different ANN techniques for daily suspended sediment concentration prediction and estimation and he indicated that multi-layer perceptron could show better performance than the others. Kisi (2005b) developed an ANN model for modeling suspended sediment and compared the ANN results with those of the rating curve (RC) and multi-linear regression (MLR). He found that the ANN model performed better than the RC and MLR. Cigizoglu & Kisi (2006) developed some methods to improve ANN performance in suspended sediment estimation. Kisi (2008) proposed a data driven algorithm for obtaining ANN models in daily sediment estimation. Kisi *et al.* (2012) modeled discharge-sediment relationship using ANN with artificial bee colony algorithm. Lafdani *et al.* (2013) used ANN and support vector machine for prediction of daily suspended sediment load. Aytek and Kisi (2008) developed new formulae based on genetic programming for estimating suspended sediment. Recently, ANN models have been developed by using field data, experimental results, atmospheric variables and even numerical data. ANNs are traditionally used as black-box models and no one is interested in the fundamental hidden formulation. The output is obtained from black-box by introducing input data without understanding what happens inside the box. The accuracy of the models is then evaluated by comparing the models' output with the observed data. The question should be that, how anyone who may be unfamiliar with ANNs can apply this kind of models in a study, while the model's formulation does not exist. Recent studies such as Aytek *et al.* (2008), Guven *et al.* (2006) and Khorchani & Blanpain (2005) are observed in civil engineering literature to overcome this problem by providing the explicit formulation on which the neural network system is based.

The main objective of this study is to determine an explicit neural network (ENN) formulation that evaluates the amount of suspended sediment in a river

as a function of input variables. The ENN is applied to daily streamflow and suspended sediment data of two stations operated by the U.S. Geological Survey (USGS) and the results are compared with those of the sediment rating curves and multi-linear and nonlinear regression. To the best knowledge of the authors, no work has been reported in the literature that addresses the explicit neural network formulations for estimation of suspended sediment load.

2. Methodology

2.1 Explicit neural network (ENN)

In this study, the multilayer perceptron (MLP) ANN architecture (Rumelhart, 1986) with one single hidden layer is considered. MLP is layered feed forward networks typically trained with static backpropagation. Their main advantage is that they are easy to use, and that they can approximate any input/output map. The disadvantages are that they train slowly, and require lots of training data (typically three times more training samples than network weights). The basic element of a neural network is an artificial neuron as shown in Fig. 1, which consists of three main components; weights, bias, and an activation function. Each neuron receives inputs x_i ($i = 1, 2, \dots, n$) attached with a weight w_{ij} ($j \geq 1$) which shows the connection strength for a particular input for each connection. Every input is then multiplied by the corresponding weight of the neuron connection and summed as

$$W_i = \sum_{j=1}^n w_{ij}x_j \quad (1)$$

A bias b_i , is a type of correction weight with a constant non-zero value, is added to the summation in Equation (1) as

$$U_i = W_i + b_i = \sum_{j=1}^n w_{ij}x_j + b_i \quad (2)$$

In the architecture tangent-sigmoid transfer function is utilized as

$$y_i = f(U_i) = \frac{2}{1 + e^{-2U_i}} - 1 \quad (3)$$

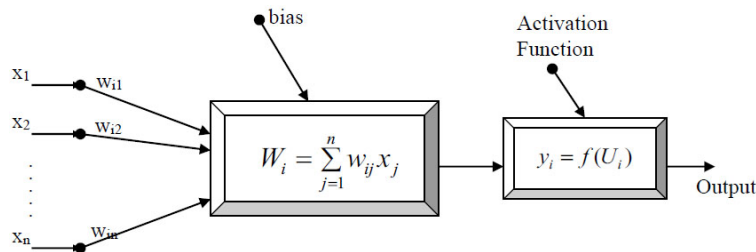


Fig. 1 The basic element of ANN neuron (Aytek et al. 2008).

The Levenberg-Marquardt algorithm that is more powerful than the standard back propagation algorithm (Kisi, 2004b; Cigizoglu & Kisi, 2005) is employed to minimize the RMSE of the network in this study. The back propagation with gradient descent technique is a steepest descent algorithm, while the Levenberg-Marquardt algorithm is an approximation to Newton's method (Marquardt, 1963). If we want to minimize a function $V(\underline{x})$ with respect to the parameter vector \underline{x} , then Newton's method would be

$$\Delta \underline{x} = -[\nabla^2 V(\underline{x})]^{-1} \nabla V(\underline{x}) \quad (4)$$

where $\nabla^2 V(\underline{x}) =$ the Hessian matrix and $\nabla V(\underline{x}) =$ the gradient. Let assume that $V(\underline{x})$ is a sum of squares function

$$V(\underline{x}) = \sum_{i=1}^N e_i^2(x) \quad (5)$$

then it can be shown that

$$\nabla V(x) = J^T(\underline{x}) \underline{e}(\underline{x}) \quad (6)$$

$$\nabla^2 V(x) = J^T(\underline{x}) J(\underline{x}) + S(\underline{x}) \quad (7)$$

where $J(x) =$ the Jacobean matrix and

$$S(\underline{x}) = \sum_{i=1}^N e_i \nabla^2 e_i(\underline{x}) \quad (8)$$

For the Gauss-Newton method it is assumed that $S(\underline{x}) \approx 0$, and the update (4) becomes

$$\Delta \underline{x} = [J^T(\underline{x}) J(\underline{x})]^{-1} J^T(\underline{x}) \underline{e}(\underline{x}) \quad (9)$$

The Levenberg-Marquardt modification to the Gauss-Newton method is

$$\Delta \underline{x} = [J^T(\underline{x}) J(\underline{x}) + \mu I]^{-1} J^T(\underline{x}) \underline{e}(\underline{x}) \quad (10)$$

The parameter μ is multiplied by some factor (β) when a step increases $V(\underline{x})$. When a step would result in a reduced $V(\underline{x})$, μ is divided by β . When μ is large the algorithm becomes steepest descent (with step $1/\mu$), while the algorithm becomes Gauss-Newton for small μ . The Levenberg-Marquardt algorithm can be considered a trust-region modification to Gauss-Newton. The computation of the Jacobean matrix is the key step in this algorithm. The terms in the Jacobean matrix can be computed by a simple modification to the back propagation algorithm for the neural network-mapping problem (Hagan & Menhaj, 1994).

2.2 Sediment Rating Curve (SRC)

The sediment discharge rating curve is sometimes called a suspended sediment transport graph or a suspended sediment transport relationship. It is a relationship between water and sediment discharge. A rating curve consists of a graph or equation, relating sediment discharge or concentration to stream discharge, which can be used to estimate sediment loads from the streamflow record. A sediment rating curve describes the average relation between discharge and suspended sediment concentration for a certain location. The sediment rating curve generally represents a functional relationship of the power form:

$$S = aQ^b \quad (11)$$

in which Q is water discharge (m^3/s) and S is suspended sediment load (mg/l or ton/day) (Sandy, 1990). Values of a and b for a particular stream are determined from data via a linear regression between $(\log S)$ and $(\log Q)$. After log-transformation to the arithmetic domain and application of the Ferguson (1986) correction factor, the sediment load occurring at a specific discharge can be estimated using the following expression:

$$S = CF.a.Q^b \quad (12)$$

where CF is the log-transformation bias correction factor. Specifically,

$$CF = e^{2.65s^2} \quad (13)$$

where e is the exponential function and s is the standard error of the regression equation. In the applications, first sediment rating curve (Eq. 11) is denoted as SRC1 and the second one with bias correction factor (Eq. 12) is denoted as SRC2.

2.3 Multiple Linear Regression (MLR)

If it is assumed that the dependent variable Y is effected by m independent variables X_1, X_2, \dots, X_m and a linear equation is selected for the relation among them, the regression equation of Y can be written as:

$$y = a + b_1x_1 + b_2x_2 + \dots + b_mx_m \quad (14)$$

y in this equation shows the expected value of the variable Y when the independent variables take the values $X_1 = x_1, X_2 = x_2, \dots, X_m = x_m$.

The regression coefficients a, b_1, b_2, \dots, b_m are evaluated, similar to simple regression, by minimizing the sum of the e_{yi} distances of observation points from the plane expressed by the regression equation (Bayazit and Oguz, 1998):

$$\sum_{i=1}^N e_{yi}^2 = \sum_{i=1}^N (y_i - a - b_1x_{1i} - b_2x_{2i} - b_mx_{mi})^2 \quad (15)$$

In this study, the coefficients a, b_1, b_2, \dots, b_m are determined using least squares method.

2.4 Nonlinear Regression (NLR)

Assume that the dependent variable Y is nonlinearly effected m independent variables X_1, X_2, \dots, X_m and following nonlinear equation (Eq. 16) can be written for the relation among them

$$y = a_0 x_1^{a_1} \cdot x_2^{a_2} \dots x_m^{a_m} \quad (16)$$

The Eq. (17) is simply obtained by applying logarithm transformation to the Eq. (16)

$$\log y = \log a_0 + a_1 \log x_1 + a_2 \log x_2 + \dots + a_m \log x_m \quad (17)$$

The coefficients $a_0, a_1, a_2, \dots, a_m$ can be simply determined using least squares method as explained in the MLR section.

3. Study Area

The daily streamflow and suspended sediment time series data from two stations on Tongue River in Montana are used. The location of the stations is shown in Fig. 2. The downstream station (station no: 6308500) at Miles City and the upstream station (station no: 6307830) below Brandenburg Bridge near Ashland are operated by USGS. The drainage areas at these sites are 13932 km² for the downstream station and 10521 km² for the upstream station. For these stations, the daily time series data were downloaded from web server of the USGS (<http://webserver.cr.usgs.gov/sediment>). Tongue River flows 286 miles from Big Horn Mountains in Wyoming to the confluence with the Yellowstone River near Miles City, Montana. The total catchment area is roughly 5,400 square miles. The river has naturally high suspended solids due to topography, geology and soils (MDEQ, 2003). Historic references (early 1800s) report that the river (Tongue) was very muddy and shallow, with shifting sand bars and quicksand present in the channel near Miles City (NRCS, 2002; USEPA, 2007).

After examining the data and noting the periods in which there are gaps in one or more of the two variables, the periods for training and testing are chosen. The data of October 01, 1977 to September 30, 1980 (75% of the whole data) are chosen for training, and data for October 01, 1980 to September 30, 1981 (25% of the whole data) are chosen for testing in all applications. It may be noted that the periods from which training and testing data are chosen span the same temporal seasons (October–September).

The scatter plots of the downstream and upstream stations data are given in Fig. 3. It can be seen that there is a nonlinear and scattered relationship between discharge and sediment data for both stations. Fig. 2 indicates the presence of outliers (also see Tab. I). In the downstream data set, a suspended sediment load value of 27200 ton/day is observed while the other values are below 20000 ton/day. In the upstream data set, the suspended sediment load values of 84400 ton/day and 81600 ton/day are observed while the other sediment values are below 50000 ton/day. These outliers are also used in the training period. These values give an additional difficulty to the models in estimation. The models calibrated using such outliers generally give overestimations of the low sediment values. The daily statistical parameters of the streamflow and sediment data for each station are given

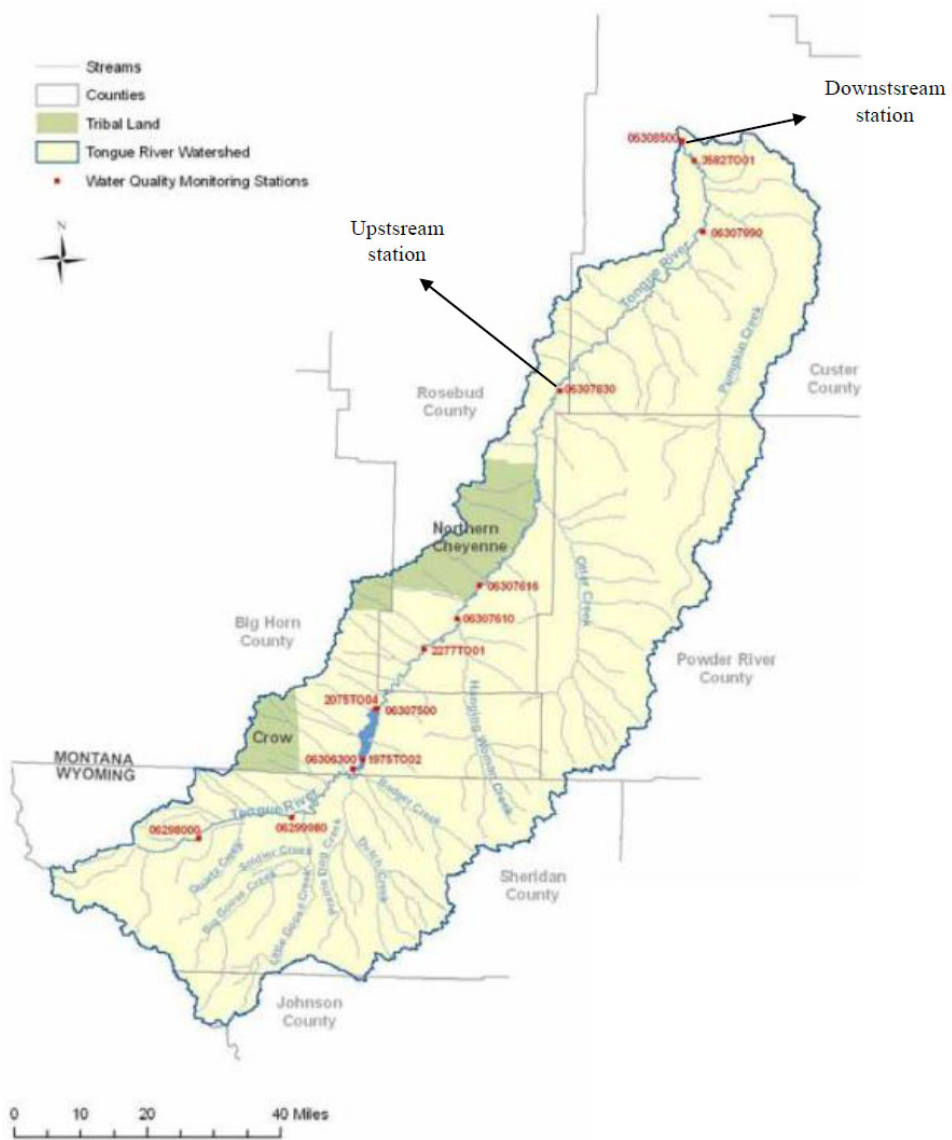


Fig. 2 The location of the downstream and upstream stations on Tongue River (USEPA, 2007).

in Tab. I. In the table, the x_{mean} , S_x , C_v , C_{sx} , x_{max} and x_{min} denote the mean, standard deviation, coefficient of variation, skewness, maximum and minimum, respectively. The skewness and coefficient of variation of flow and sediment data of the upstream and downstream stations are high, particularly for the training (calibration) data. In the calibration flow data, x_{min} and x_{max} values are 1.9 and 218 m³/s respectively for the downstream station. However, the testing flow data

set extremes are $x_{\min} = 0.1 \text{ m}^3/\text{s}$, $x_{\max} = 62.3 \text{ m}^3/\text{s}$. The value of x_{\min} for the calibration flow data is higher than that for the corresponding test set for the downstream station. This may cause extrapolation difficulties in estimation of low sediment values.

The auto and cross-correlation coefficients for the downstream and upstream station data are given in Tab. II. In this table, the S_t and Q_t represent the sediment load and discharge at time t , respectively. It may be seen that the auto-correlations of the sediment data are good for both stations. However, the correlations between the sediment load and discharges are normally not good and in fact are very poor for the calibration period data of the downstream and upstream stations.

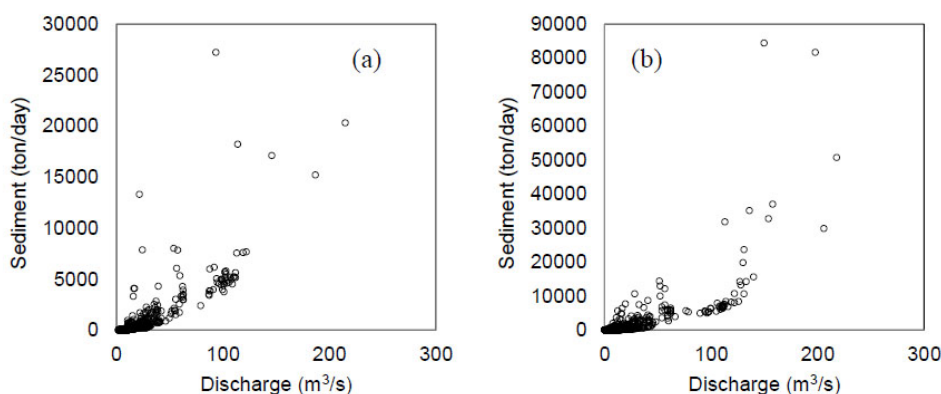


Fig. 3 Scatterplots of the (a) downstream and (b) upstream data.

4. Performance Measures

It is important to define the criteria by which the performance of the model and its prediction accuracy will be evaluated in model development process. The current study will use various statistical measures to assess the model performance, namely, the root mean square errors (RMSE), the mean absolute errors (MAE), the correlation coefficient (R) and the coefficient of efficiency (E). The RMSE describes the average difference between model results and observations. The MAE yields a more balanced perspective of the goodness-of-fit at moderate sediment values (Karunanithi *et al.*, 1994). The R measures the degree to which two variables are linearly related. The coefficient of efficiency, E , measures the differences between the observations and predictions relative to the variability in the observed data itself. A value of 0.5 and above indicates a satisfactory performance whereas a value below 0.5 indicates an unsatisfactory performance (Moriasi *et al.* 2007). The $RMSE$, MAE and E are defined as:

$$RMSE = \frac{1}{N} \sqrt{\sum_{i=1}^N ((Y_e)_i - (Y_o)_i)^2} \tag{18}$$

Data set	Station	Basin area (km ²)	Data type	x_{mean}	S_x	$C_v (S_x/x_{mean})$	C_{sx}	x_{max}	x_{min}
Training	Downstream (6308500)	13 932	Flow (m ³ /s)	15.6	24.9	1.60	4.1	218	1.9
			Sediment (ton/day)	996	4860	4.88	11.5	84400	3.4
Testing	Upstream (6307830)	10 521	Flow (m ³ /s)	15.1	20.6	1.36	4.4	215	1.8
			Sediment (ton/day)	444	1726	3.89	8.5	27200	1.5
Testing	Downstream (6308500)	13 932	Flow (m ³ /s)	9.0	12.9	1.43	3.1	62.3	0.1
			Sediment (ton/day)	369	1235	3.35	4.0	7400	0.1
	Upstream (6307830)	10 521	Flow (m ³ /s)	10.3	12.5	1.21	3.1	62.9	1.9
			Sediment (ton/day)	235	835	3.55	5.4	8000	3

Tab. I The daily statistical parameters of data set for the stations.

Downstream station		S_{t-1}	S_{t-2}	S_{t-3}	Q_t	Q_{t-1}	Q_{t-2}
Training period	S_t	0.749	0.511	0.435	0.444	0.416	0.377
Testing period	S_t	0.875	0.724	0.644	0.700	0.657	0.600
Upstream station		S_{t-1}	S_{t-2}	S_{t-3}	Q_t	Q_{t-1}	Q_{t-2}
Training period	S_t	0.707	0.387	0.308	0.565	0.490	0.426
Testing period	S_t	0.910	0.842	0.778	0.796	0.734	0.669

Tab. II The auto- and cross-correlations of the data of the downstream and upstream stations.

$$MAE = \frac{1}{N} \sum_{i=1}^N |(Y_e)_i - (Y_o)_i| \tag{19}$$

$$E = 1 - \frac{\sum_{i=1}^n ((Y_e)_i - (Y_o)_i)^2}{\sum_{i=1}^n ((Y_e)_i - Y_m)^2} \tag{20}$$

where N is the number of data set, Y_m , Y_{ei} and Y_{oi} are mean, estimated and observed suspended sediment loads, respectively.

5. Application and Results

The main aim of this study is to derive an ENN formulation for suspended sediment-discharge relation as a function of input parameters. Input parameters and weights of the trained ANN are extracted to form an explicit expression in the following manner.

Each input is multiplied by a connection weight (Eq. 1) and then biases are simply added to this multiplication (Eq. 2) and finally, the sum is transformed through a transfer function (sigmoid) (Eq. 3) to generate an output. In order to acquire accurate results from the ENN, before to the execution of the training process, input and output parameters are normalized in the range of $(-0.95; 0.95)$ by

$$\Gamma_{\text{normalized}} = a\Gamma + b \tag{21}$$

where Γ represents parameters used in the ENN training process, a and b are normalization coefficients of that particular parameter. The sediment-discharge relation for downstream station can be expressed as

$$Su_t = \left(\frac{30220.556}{1 + e^{\left(\frac{228.5046}{1 + e^{-2U_1}} + \frac{118.3855}{1 + e^{-2U_2}} - 119.5014 \right)}} - 1509.523 \right) \tag{22}$$

where $U_i = W_i + b_i$ and estimated as;

$$U_i = \begin{bmatrix} -0.02918 & 0.028144 & -0.00006 \\ 0.036927 & -0.03583 & 0.00007 \end{bmatrix} \times \begin{bmatrix} Qu_t \\ Qu_{t-1} \\ Su_{t-1} \end{bmatrix} + \begin{bmatrix} -1.180308 \\ 0.941279 \end{bmatrix} \tag{22b}$$

The sediment-discharge relation for upstream station can be expressed as

$$S_t = \left(\frac{93774}{1 + e^{\left(-\frac{32.1226}{1 + e^{-2U_1}} + \frac{53.9930}{1 + e^{-2U_2}} + \frac{1.1267}{1 + e^{-2U_3}} - \frac{2.7208}{1 + e^{-2U_4}} + \frac{19.6218}{1 + e^{-2U_5}} - 17.3568 \right)}} - 4685.3 \right) \tag{23}$$

where $U_i = W_i + b_i$ and estimated as;

$$U_i = \begin{bmatrix} 0.138485 & -0.153338 & -0.000636 \\ 0.068515 & -0.074774 & -0.000315 \\ -0.173993 & 0.195361 & 0.000034 \\ 0.010027 & -0.006283 & 0.000036 \\ -0.032954 & -0.015656 & -0.000076 \end{bmatrix} \times \begin{bmatrix} Qd_t \\ Qd_{t-1} \\ Sd_{t-1} \end{bmatrix} + \begin{bmatrix} -3.008234 \\ -2.256174 \\ 0.481494 \\ -0.590769 \\ 16.393527 \end{bmatrix} \quad (23b)$$

One of the most important tasks in ANN studies is to determine the optimal network architecture which is related to the number of neurons in the hidden layer. Generally, the trial and error approach is used. In this study, the best architecture of the network is obtained by trying different number of neurons. The trial starts from two until seven, and the performance of each network is checked by R^2 and Akaike Information Criterion (AIC) defined as (Akaike, 1974);

$$AIC = N \ln(MSE) + 2k \quad (24)$$

where N is the number of exemplars in the training set, k is the number of network weights and MSE is the mean squared error. AIC is used to measure the exchange between training performance and network size. The goal is to minimize AIC to obtain a network with the best generalization. Seven input combinations (Tab. III and Tab. V), each with 7 neurons (eg. Fig. 4), totally 49 alternatives are applied to reach the best architecture of the ENN formulation. The same procedure is used for both training and test periods. In these alternatives, the input combination (vi) with 5 neurons for the downstream station and 2 neurons for the upstream stations are selected as the optimum structure of the ENN. The relationship between the number of neurons ranging from 2 to 7 and the corresponding AIC and R values are presented in Fig. 3 for the testing stage. It is clearly seen in Fig. 4a that the AIC decreases and R increases till 5 neurons for the downstream station. The best hidden neuron number is 5 for the downstream station. From Fig. 4b, the optimum number of hidden neuron seems to be 2 for the upstream station ($R = 0.924$ and $AIC = -3022$). Based on these analyses, the optimal architecture of the ENN is constructed as 3-5-1 for the downstream and 3-2-1 for the upstream representing the number of input, hidden and output neurons, respectively.

For the downstream station, the RMSE, MAE and R of each ENN model in test period are given in Tab. III. Tab. III indicates that the ENN model whose inputs are the current discharge and the discharge and suspended sediment load of one previous period (input combination (vi)) has the smallest RMSE (317 ton/day) and MAE (34.2 ton/day). The ENN performance for the input combination (i) (only current discharge) is the worst since the hysteresis effect between sediment load and discharge. That is the say that the sediment loads for a given level of streamflow discharge in the rising stage of a streamflow hydrograph are greater than on the falling stage. This confirms that the practice of using sediment load versus discharge is misleading as said by McBean & Al-Nassri (1988). The critical issue in training ENN is avoiding overfitting as it reduces its capacity of generalization. If too many neurons are used, the network has too many parameters and may overfit the data. In contrast, if too few neurons are included in the network, it might not be possible to fully detect the signal and variance of a complex data set. Here, the

Input combinations	Nodes in hidden layer	RMSE (ton/day)	MAE (ton/day)	R
(i) Q_t	3	542	52.7	0.842
(ii) S_{t-1}	6	320	34.6	0.939
(iii) Q_t and S_{t-1}	6	525	58.0	0.903
(iv) S_{t-1} and S_{t-2}	5	319	38.7	0.935
(v) Q_t , S_{t-1} , and S_{t-2}	4	393	34.9	0.942
(vi) Q_t , Q_{t-1} , and S_{t-1}	5	317	34.2	0.940
(vii) Q_t , Q_{t-1} , S_{t-1} , and S_{t-2}	2	318	36.3	0.939

Tab. III The RMSE, MAE and R of ENN models in test period – Downstream station.

Model	Training				Testing			
	RMSE (ton/day)	MAE (ton/day)	R	E	RMSE (ton/day)	MAE (ton/day)	R	E
ENN	615	53.0	0.953	0.98	317	34.2	0.969	0.94
SRC1	2822	149	0.818	0.31	650	48.5	0.961	0.61
SRC2	2954	196	0.818	0.66	449	36.7	0.961	0.87
MLR	1603	135	0.847	0.88	508	73.6	0.917	0.90
NLR	2540	128	0.853	0.63	346	34.3	0.956	0.72

Tab. IV The RMSE, MAE, R and E of each model in test period – Downstream station.

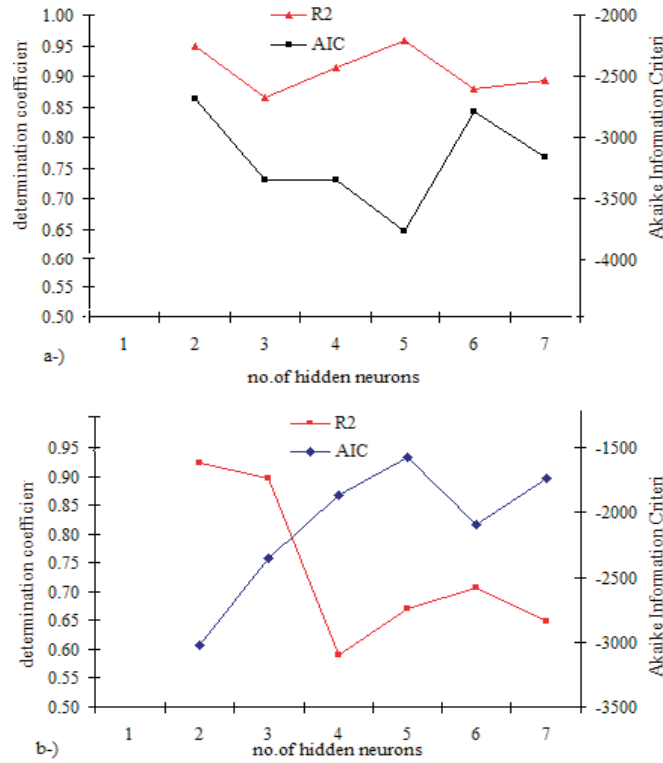


Fig. 4 The effect of number of hidden neurons on the ENN performance in a) testing period for downstream (6308500) station and b) testing period for upstream (6307830) station.

hidden node numbers of the ENN are determined using the trial and error method. For the downstream station, optimum hidden node numbers of the ENN models are found to vary between 2 and 6. The best way to avoid overfitting is to use lots of training data. For noise-free data, if we have at least 5 times as many training cases as there are weights in the network, we are unlikely to suffer from overfitting. The other way to avoid overfitting problem is to use of different ENN structures (Sudheer *et al.*, 2002). In this study, different ENN structures are tried and 3-year of flow data (1096 days) are used for training of the ENN models. The 20 weights ($3 \times 5 + 5 = 20$) are used for the most complex ENN(3,5,1) models comprising 3 inputs, 5 hidden and 1 output nodes (input combination (v)). The training data seem to be enough to avoid overfitting.

SRC1, SRC2, MLR and NLR formulas obtained for the downstream station are:

$$S = 0.7066.Q^{2.0589} \quad (25)$$

$$S = 0.7066CF.Q^{2.0589} \quad (26)$$

$$S = 401.11.Q_t - 362.32.Q_{t-1} + 0.661.S_{t-1} - 272.8 \quad (27)$$

$$S = 0.518.Q_t^{1.3215}Q_{t-1}^0S_{t-1}^{0.4158} \quad (28)$$

CF is calculated as 1.496 in Eq. (26).

The training and test results of the optimal ENN model whose inputs are the Q_t , Q_{t-1} , and S_{t-1} compared with those of the SRC1, SRC2, MLR and NLR models in Tab. IV for the downstream station. It can be obviously seen from this table that the ENN model performs better than the rating curves, linear and nonlinear regression techniques. The RMSE, MAE and E results imply that the bias term CF increases the rating curve accuracy. The suspended sediment estimates of ENN, SRC1, SRC2, MLR and NLR models are shown in Fig. 5 in the form of hydrograph and scatterplot (the latter plotted as double logarithmic for better representation). As seen from the figure (especially from the scatterplots) that the ENN model approximates the corresponding observed suspended sediment values better than the rating curve and regression techniques. The MLR also performs better than the SRC models. Significantly underestimations are clearly seen for the NLR model. Among the SRC models, the SRC2 performs better than the SRC1 model.

The estimation of total sediment load is also considered for comparison due to its importance in reservoir management. The ENN, SRC1, SRC2 and NLR models respectively estimate the observed total sediment load of 134689 t as 129348, 79169, 118423 and 95297 t with underestimations of 4, 41.2, 12.1 and 29.2% while the MLR method computes the total sediment load as 162903 t, with overestimation of 20.9%. The ENN model provides better total sediment load estimate.

For the upstream station, RMSE, MAE, R^2 and E of ENN models are given in Tab. V. Here, also the ENN model provides best accuracy for the input combination (vi).

SRC1, SRC2, MLR and NLR formulas obtained for the upstream station are:

$$S = 0.4296.Q_t^{2.1022} \quad (29)$$

$$S = 0.4296.CF.Q_t^{2.1022} \quad (30)$$

$$S = 131.9.Q_t - 110.52.Q_{t-1} + 0.641.S_{t-1} - 163.1 \quad (31)$$

$$S = 0.991.Q_t^{0.6086}Q_{t-1}^{0.4549}S_{t-1}^{0.4326} \quad (32)$$

CF is calculated as 1.389 in Eq. (30).

The comparison of the ENN, SRC, MLR and NLR models were presented in Tab. VI. For the upstream station, the ENN model also shows better accuracy than the rating curve and regression techniques. Here also the SRC2 performs better than the SRC1 model with respect to RMSE and MAE. The MLR performs much better than the SRC models from the RMSE, R^2 and E viewpoints. The suspended sediment estimates of each model and observed values are compared in Fig. 6. It can be seen from scatterplots that the ENN model predictions are much closer to the exact fit line than those of the SRC1, SRC2, MLR and NLR models, especially for the high values (>100 ton/day). The MLR seems to perform better than the SRC models. Here also the NLR gives significantly underestimations. The ENN and MLR predictions of the total sediment load, 85731 t are 97470 and 89403 t, with overestimations error of 13.7 and 4.3%, while the SRC1, SRC2 and

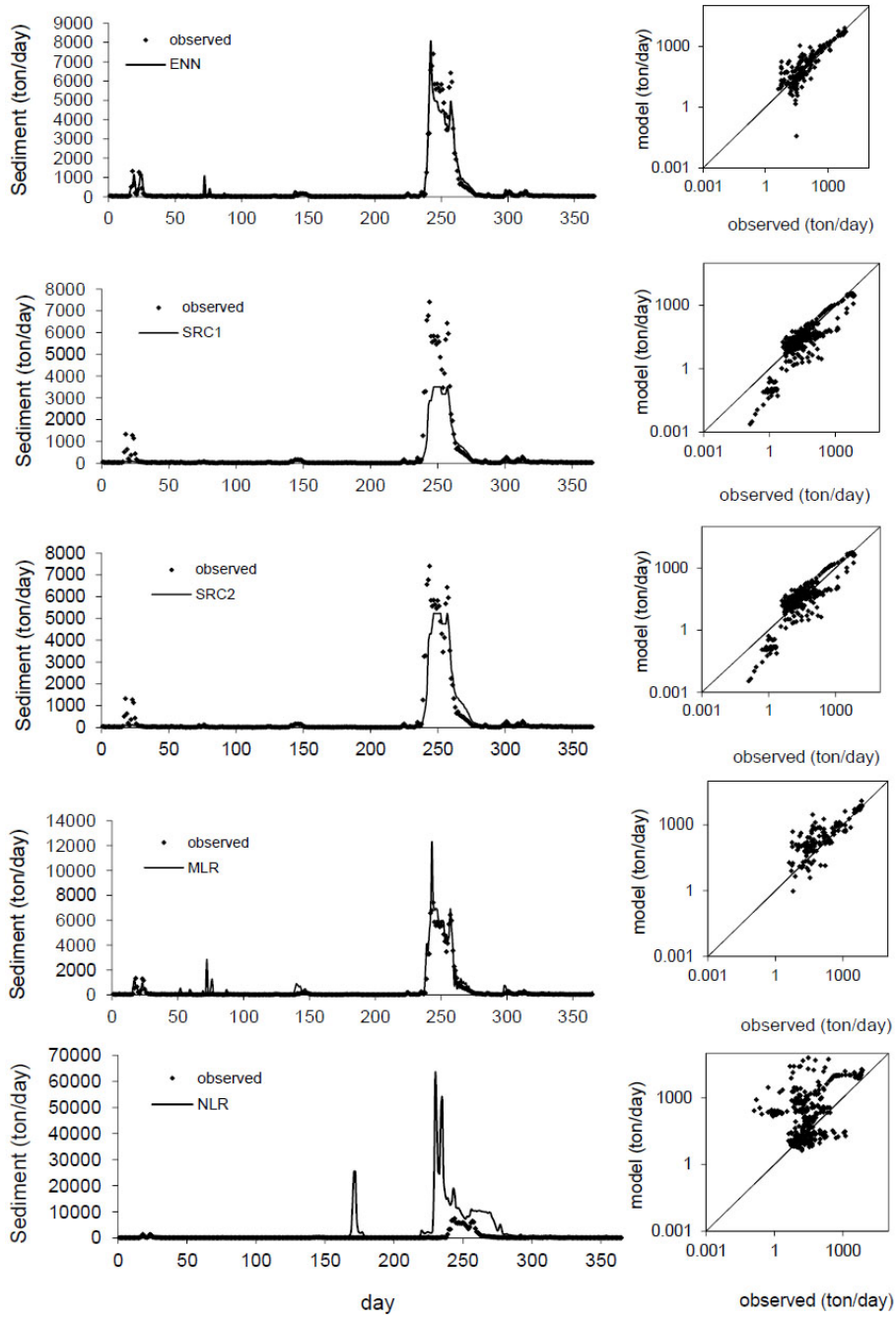


Fig. 5 Suspended sediment estimates of ENN, SRC1, SRC2, MLR and NLR models in test period – Downstream station.

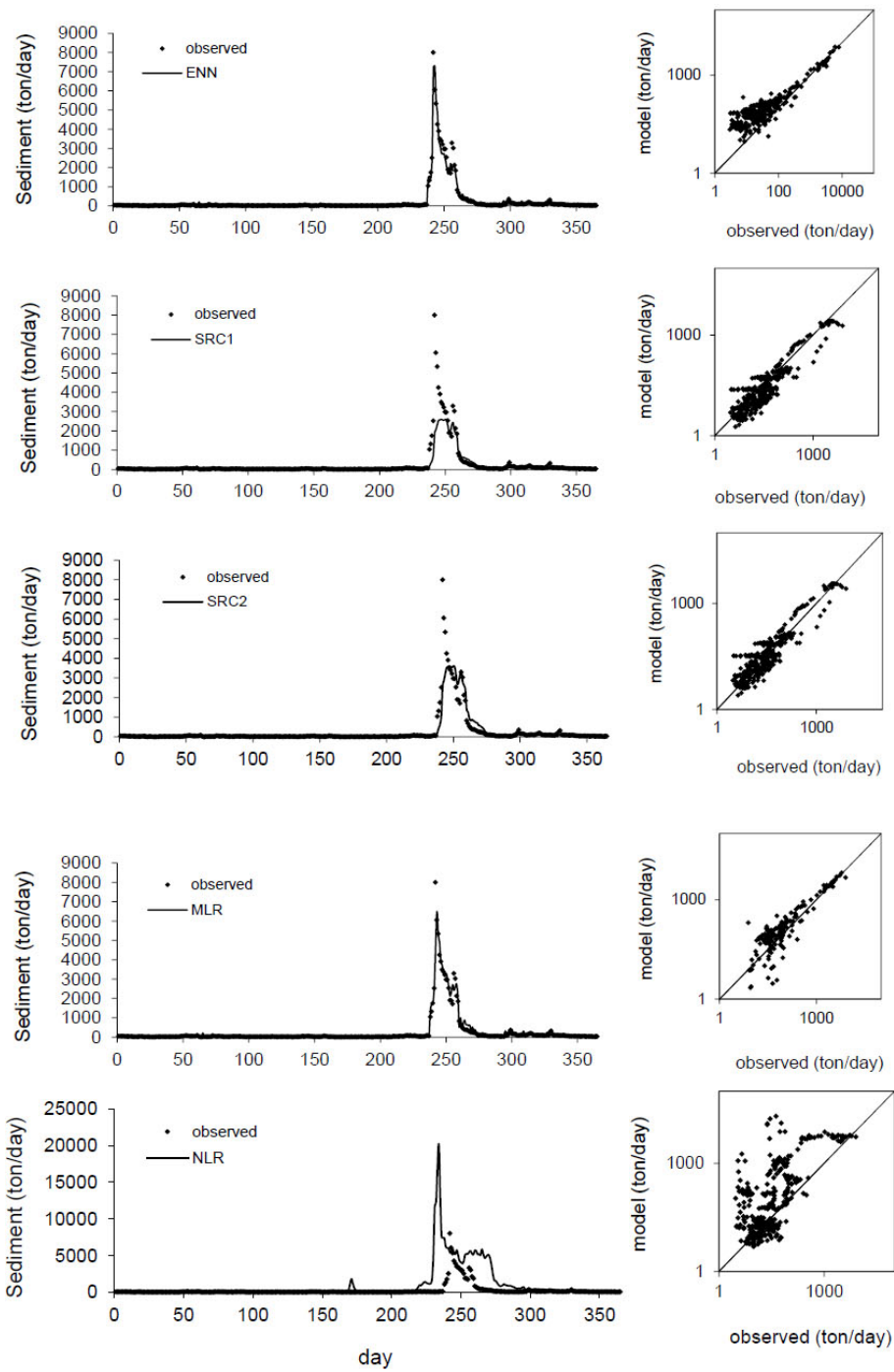


Fig. 6 Suspended sediment estimates of ENN, SRC1, SRC2, MLR and NLR models in test period – Upstream station.

Input combinations	Nodes in hidden layer	RMSE (ton/day)	MAE (ton/day)	R
(i) Q_t	3	390	33.2	0.782
(ii) S_{t-1}	6	393	22.2	0.814
(iii) Q_t and S_{t-1}	2	374	17.1	0.805
(iv) S_{t-1} and S_{t-2}	2	540	31.6	0.763
(v) Q_t, S_{t-1} , and S_{t-2}	5	331	24.1	0.849
(vi) Q_t, Q_{t-1} , and S_{t-1}	2	159	21.4	0.966
(vii) Q_t, Q_{t-1}, S_{t-1} , and S_{t-2}	4	292	18.1	0.890

Tab. V The RMSE, MAE and R of ENN models in test period – Upstream station.

Model	Training				Testing			
	RMSE (ton/day)	MAE (ton/day)	R	E	RMSE (ton/day)	MAE (ton/day)	R	E
ENN	554	42.0	0.956	0.88	159	21.4	0.976	0.96
SRC1	1133	63.7	0.840	0.65	465	26.6	0.904	0.34
SRC2	1621	95.7	0.840	0.15	396	25.9	0.904	0.28
MLR	678	46.3	0.948	0.82	226	27.3	0.964	0.93
NLR	937	41.5	0.841	0.56	413	21.5	0.901	0.45

Tab. VI The RMSE, MAE, R and E of each model in test period – Upstream station.

NLR yield as 59754, 82972 and 67254 t, with underestimations of 30.3, 3.2 and 21.6%, respectively. The SRC2 model estimate is closest to the observed one.

Conventionally, the current suspended sediment load is only related to the current discharge in the rating curves. However, the current suspended sediment load is not only depended on the discharge at the current time but also the discharges and suspended sediments at the previous times. For that reason, if the suspended sediment load is related to the current discharge and antecedent discharge and suspended sediment values through a MLR, better results are obtained. However, the MLR is incapable of setting up river discharge-sediment load relationship in comparison to the ENN models, since the MLR technique assume linear relationship amongst the variables. Such models require that the variables have normal distribution. The ENN models seem to be more adequate than the SRC1, SRC2, MLR and NLR for the process of establishing a rating relationship between river discharge and suspended sediment load. Such problems arise behave in a nonlinear manner. The main advantages of using ENNs are their flexibility and ability to model nonlinear relationships. Mathematically, an ENN may be treated as a universal approximator (ASCE Task Committee, 2000). This technique has already become a prospective research area with great potential due to the ease of application and simple formulation.

6. Concluding Remarks

Previously given ANN-based formulations are all implicit. Therefore, explicit neural network formulations have been proposed for modeling daily suspended sediment-discharge relationship in the present study. The ENN is very simple and it can be used by anyone who may be not familiar with ANNs, in a spreadsheet on a very simple PC or in a hand calculator. The ENN models were tested applying to different input combinations of daily streamflow and suspended sediment load data of two stations on Tongue River in Montana. The suspended sediment estimates based on ENN models were compared with two different sediment rating curves and multi-linear regression. The results obtained with ENN models were found to be better than those obtained using the conventional rating curve, MLR and NLR techniques. The comparison results suggested that the ENN approach may provide a superior alternative to the sediment rating curve techniques and regression techniques used in the current study. The difficulties in estimation of suspended sediment load using only current discharge, resulting from the hysteresis effect, were also indicated. The MLR technique performed better than the rating curves. The SRC2 model with a bias term is found to be much better than the unbiased SRC1 model. The study only used data from two areas and further work using more data from various areas may be required to strengthen these conclusions.

Acknowledgements

The second author would like to thank the Research Foundation of Gaziantep University for the support provided during this study. The data used in this study were downloaded from the web server of the USGS. The authors wish to thank the staffs of the USGS who are associated with data observation, processing, and

management of USGS Web sites. The authors also thank to the reviewers for their constructive comments and support on our paper.

References

- [1] Akaike, H.: A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, AC-19: 716–723, 1974.
- [2] ASCE Task Committee.: Artificial neural networks in hydrology. II: Hydrological applications, *J. Hydrol. Eng. ASCE*, 5(2), 124–137, 2000.
- [3] Aytek A., Kisi O.: A genetic programming approach to suspended sediment modelling, *Journal of Hydrology*, 351, 288–298, 2008.
- [4] Aytek A., Guven A., Yuce M. I., Aksoy H.: An explicit neural network formulation for evapotranspiration, *Hydrological Sciences Journal*, 53(4), 893–904, 2008.
- [5] Bae, D-H., Jeong, D. M., Kim, G.: Monthly dam inflow forecasts using weather forecasting information and neuro-fuzzy technique, *Hydrol. Sci. J.*, 52(1), 99–113, 2007.
- [6] Balara, D., Timko, J., Zilkova, J.: Application of neural network model for parameters identification of non-linear dynamic system. *Neural Network World*, 22(6), 495-510, 2012.
- [7] Yalkin, C., Korkmaz, E. E.: Neural network world: A neural network based selection method for genetic algorithms. *Neural Network World*, 23(2), 103–116, 2013.
- [8] Bayazit, M., Oguz, B.: *Probability and Statistics for Engineers*, Birsen Publishing House, Istanbul, Turkey, p. 159, 1998.
- [9] Campolo, M., Soldati, A.: Artificial neural network approach to flood forecasting in the River Arno, *Hydrol. Sci. J.* 48(3), 381–398, 2003.
- [10] Cigizoglu, H. K.: Estimation, forecasting and extrapolation of river flows by artificial neural Networks, *Hydrol. Sci. J.*, 48(3), 349–361, 2003.
- [11] Cigizoglu, H. K.: Estimation and forecasting of daily suspended sediment data by multi layer perceptrons, *Advances in Water Resources* **27**, 185–195, 2004.
- [12] Cigizoglu, H. K.: Application of the Generalized Regression Neural Networks to Intermittent Flow Forecasting and Estimation, *Journal of Hydrologic Engin.*, 10(4), 336–341, 2005a.
- [13] Cigizoglu, H. K.: Generalized Regression Neural Network in Monthly Flow Forecasting, *Civil Engineering and Environmental Systems*, 22(2), 71–84, 2005b.
- [14] Cigizoglu, H. K., Alp, M.: Generalized Regression Neural Network in modelling river sediment yield, *Advances in Engineering Software*, 37(2), 63–68, 2006.
- [15] Cigizoglu, H. K., Kisi, O.: Flow prediction by three back propagation techniques using k-fold partitioning of neural network training data, *Nordic Hydrology*, 36(1), 49–64, 2005.
- [16] Cigizoglu, H. K., Kisi, O.: Methods to improve the neural network performance in suspended sediment estimation, *J. of Hydrology*. 317, 221–238, 2006.
- [17] Chang, F-J., Chen, Y-C.: A counterpropagation fuzzy-neural network modeling approach to real time streamflow prediction, *J. of Hydrology*, 245, 153–164, 2001.
- [18] Ferguson, R. I.: River loads underestimated by rating curves, *Water Resour. Res.* 22(1), 74–76, 1986.
- [19] Giustolisi, O., Laucelli, D.: Improving generalization of artificial neural networks in rainfall–runoff modelling, *Hydrol. Sci. J.* 50(3), 439–457, 2005.
- [20] Guven, A., Gunal, M., Cevik, A. K.: Prediction of pressure fluctuations on sloping stilling basins. *Canadian Journal of Civil Engineering*, 33(11), 1379–1388, 2006.
- [21] Hagan M. T., Menhaj M. B.: Training feed forward networks with the Marquardt algorithm. *IEEE Trans Neural Networks*; 6:861–7, 1994.
- [22] Hu, T. S., Lam, K. C., Thomas, N. G. S.: A modified neural network for improving river flow prediction, *Hydrol. Sci. J.* 50(2), 299–318, 2005.

- [23] Jain, S.K.: Development of integrated sediment rating curves using ANNs, *J. of Hydraulic Eng., ASCE*, **127**(1), 30–37, 2001.
- [24] Jain, S. K., Das, D., Srivastava, D. K.: Application of ANN for reservoir inflow prediction and operation, *J. Water Res. Planning Mgmt ASCE* **125**(5), 263–271, 1999.
- [25] Jayawardena, A. W., Xu, P. C., Tsang, F. L., Li, W. K.: Determining the structure of a radial basis function network for prediction of nonlinear hydrological time series, *Hydrol. Sci. J.* **51**(1), 21–44, 2006.
- [26] Kang, K. W., Park, C. Y., Kim, J. H.: Neural network and its application to rainfall–runoff forecasting, *Korean J. Hydrosoci.* **4**, 1–9, 1993.
- [27] Karunanithi, N., Grenney, W. J., Whitley, D., Bovee, K.: Neural networks for river flow prediction, *J. Comp. Civil Eng. ASCE*, **8**(2), 201–220, 1994.
- [28] Kisi, O.: River flow modeling using artificial neural networks, *ASCE Journal of Hydrologic Engineering*, **9**(1), 60–63, 2004a.
- [29] Kisi, O.: Multi-layer perceptrons with Levenberg-Marquardt optimization algorithm for suspended sediment concentration prediction and estimation, *Hydrol. Sci. J.* **49**(6), 1025–1040, 2004b.
- [30] Kisi, O.: Daily river flow forecasting using artificial neural networks and auto-regressive models, *Turk. J. Eng. Environ. Sci.*, **29**, 9–20, 2005a.
- [31] Kisi, O.: Suspended sediment estimation using neuro-fuzzy and neural network approaches, *Hydrol. Sci. J.*, **50**(4), 683–696, 2005b.
- [32] Kisi, O.: Constructing neural network sediment estimation models using a data-driven algorithm, *Mathematics and Computers in Simulation*, **79**(1), 94–103, 2008.
- [33] Kisi, O., Cigizoglu, H. K.: Comparison of different ANN techniques in continuous and intermittent river flow prediction, *Civ. Eng. & Env. Sys.*, **24**(3), 211–231, 2007.
- [34] Kisi, O., Karahan, M. E., Sen, Z.: River suspended sediment modelling using a fuzzy logic approach. *Hydrol. Process.* **20**, 4351–4362, 2006.
- [35] Kisi, O., Ozkan, C., Akay, B.: Modeling discharge-sediment relationship using neural networks with artificial bee colony algorithm, *J. of Hydrology*, **428–429**, 2012, 94–103.
- [36] Khorchani, M., Blanpain, O.: Development of a discharge equation for side weirs using artificial neural networks. *Journal of Hydroinformatics*, **7**(1), 31–39, 2005.
- [37] Lafdani, E.K., Nia, A.M., Ahmadi, A.: Daily suspended sediment load prediction using artificial neural networks and support vector machines. *J. of Hydrology*, **478**, 50–62, 2013.
- [38] Liu, Q. J., Shi, Z. H., Fang, N. F., Zhu, H. D., AI, L.: Modeling the daily suspended sediment concentration in a hyperconcentrated river on the Loess Plateau, China, using the Wavelet–ANN approach. *Geomorphology*, **186**, 181–190, 2013.
- [39] Marquardt, D.: An algorithm for least squares estimation of non-linear parameters, *J. Soc. Ind. Appl. Math.*, **431–441**, 1963.
- [40] McBean, E. A., Al-Nassri, S.: Uncertainty in suspended sediment transport curves, *J. Hydrol. Engng ASCE*, **114**(1), 63–74, 1998.
- [41] MDEQ.: Water Quality Assessment for the Tongue River Planning Area, Montana. Prepared by USEPA and Tetra Tech, Inc. for the Montana Department of Environmental Quality, Water Quality Planning Bureau, 2003. Helena, Montana. Available online at <http://www.deq.state.mt.us/wqinfo/TMDL/TonguePowderRosebudTMDL.asp>
- [42] Moriasi, D. N., Arnold, J. G., Liew, M. W. V., Bingner, R. L., Harmel, R. D., Veith, T. L.: Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the American Society of Agriculture and Biological Engineers*, **50**(3), 885–900, 2007.
- [43] Nourani, V., Komasi, M., Mano, A.: A multivariate ANN–wavelet approach for rainfall–runoff modeling. *Water Resources Manage.* **23**, 2877–2894, 2009.
- [44] Nourani, V., Kisi, O., Komasi, M.: Two hybrid artificial intelligence approaches for modeling rainfall–runoff process. *J. Hydrol.*, **402**, 41–59, 2011.

Kişi O., Aytek A.: Explicit neural network in suspended sediment load estimation

- [45] NRCS.: Tongue River Stream Corridor Assessment. Montana Reaches. Phase II – Physical Habitat Assessment. Natural Resources Conservation Service. U.S. Department of Agriculture. Bozeman, Montana, 2002.
- [46] Ozkan, C., Ozturk, C., Sunar, F., Karaboga, D.: The Artificial Bee Colony algorithm in training Artificial Neural Network for oil spill detection. *Neural Network World*, 21(6), 473–492, 2011.
- [47] Rumelhart D. E.: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, Vol. 1, Cambridge, MA: MIT Press, 1986.
- [48] Saad, M., Bigras, P., Turgeon, A., Duquette, R.: Fuzzy learning decomposition for the scheduling of hydroelectric power systems, *Water Resour. Res.*, 32(1), 179–186, 1996.
- [49] Sandy, R.: *Statistics for Business and Economics*, McGraw-Hill Publishing, New York, pp. 1117, 1990.
- [50] Shamseldin, A.Y.: Application of a neural network technique to rainfall–runoff modelling, *J. Hydrol.*, 199, 272–294, 1997.
- [51] Sivakumar, B., Wallender, W., 2005. Predictability of river flow and suspended sediment transport in the Mississippi River basin: a non-linear deterministic approach. *Earth Surf. Process. Landforms Earth Surf. Process. Landforms*, 30, 665–677.
- [52] Solomatine, D. P., Dulal, K. N.: Model trees as an alternative to neural networks in rainfall–runoff modelling, *Hydrol. Sci. J.*, 48(3), 399–411, 2003.
- [53] Sudheer, K. P., Gosain, A. K., Rangan D. M., Saheb S. M.: Modelling evaporation using an artificial neural network algorithm, *Hydrological Process*, 16, 3189–3202, 2002.
- [54] Tayfur, G.: Artificial neural networks for sheet sediment transport, *Hydrological Sciences Journal, IAHS*, 47(6), 879–892, 2002.
- [55] Tokar, A. S., Johnson, P. A.: Rainfall–runoff modeling using artificial neural networks. *J. Hydrol. Eng., ASCE*, 4(3), 232–239, 1999.
- [56] USEPA: *Water Quality Assessment for the Tongue River Watershed, Montana*. U.S. Environmental Protection Agency Montana Operations Office and Tetra Tech, Inc, Montana, 2007.
- [57] Yalkin, C., Korkmaz, E. E.: Neural network world: A neural network based selection method for genetic algorithms. *Neural Network World*, 22(6), 495–510, 2012.
- [58] Yilmaz, A. G., Imteaz, M. A., Jenkins, G.: Catchment flow estimation using Artificial Neural Networks in the mountainous Euphrates Basin, *J. Hydrol.* 410, 134–140, 2011.
- [59] Zhang, Q., Wang, B. D., He, B., Peng, Y., Ren, M. L., 2011.: Singular spectrum analysis and ARIMA hybrid model for annual runoff forecasting. *Water Resources Manage.*, 25, 2683–2703, 2011.