Artificial neural networks for estimating daily total suspended sediment in natural streams

Gokmen Tayfur[1] and Veysel Guldal[2]

[1]Corresponding author. Department of Civil Engineering, Izmir Institute of Technology, Urla, Izmir, Turkey 35340. E-mail: *gokmentayfur@iyte.edu.tr*

[2]Department of Civil Engineering, Suleyman Demirel Unversity, Isparta, Turkey

Received 18 June 2004; accepted in revised form 29 March 2005

Nordic Hydrology **Vol 37 No 1 pp 69–79**

Abstract Estimates of sediment loads in natural streams are required for a wide spectrum of water resources engineering problems from optimal reservoir design to water quality in lakes. Suspended sediment constitutes 75–95% of the total load. The nonlinear problem of suspended sediment estimation requires a nonlinear model. An artificial neural network (ANN) model has been developed to predict daily total suspended sediment (TSS) in rivers. The model is constructed as a three-layer feedforward network using the back-propagation algorithm as a training tool. The model predicts TSS rates using precipitation (P) data as input. For network training and testing 240 sets of data sets were used. The model successfully predicted daily TSS loads using the present and past 4 days precipitation data in the input vector with $R^2 = 0.91$ and $MAE = 34.22$ mg/L. The performance of the model was also tested against the most recently developed non-linear black box model based upon two-dimensional unit sediment graph theory (2D-USGT). The comparison of results revealed that the ANN has a significantly better performance than the 2D-USGT. Investigation results revealed that the ANN model requires a period of more than 75 d of measured P-TSS data for training the model for satisfactory TSS estimation. The statistical parameter *range* ($x_{\min} - x_{\max}$) plays a major role for optimal partitioning of data into training and testing sets. Both sets should have comparable values for the *range* parameter.

Keywords Artificial neural network; back-propagation; feedforward; parameter range; sediment graph theory; suspended sediment

Introduction

Estimates of sediment loads in natural streams are required for the optimal design and operation of water resources structures such as reservoir, dam and stable channel. Furthermore, it is known that sediment affects pollutant transport causing a water quality problem in surface water bodies. Sediment loads in rivers basically consist of bed load and suspended sediment. According to Yang (1996), suspended sediment constitutes 75–95% of the total load. Suspended sediment has long been identified for the transport of nutrients and contaminants such as heavy metals and micro-organics (Cigizoglu 2004).

Due to its importance, sediment transport has been experimentally and mathematically studied for years. Experimental studies have led to the development of numerous empirical equations to predict sediment rates (Yang 1996). These equations, when predicting measured data even for the same flow conditions, differ significantly from each other. This is because each equation is developed from a particular experiment and the resulting models do not have the universal ability to be applied to different situations (Tayfur 2003).

The mathematical studies have led to physics-based and black-box models. The physics-based models are expressed by partial differential equations (Tayfur 2001, 2002a; Guo and Jin 2002; Pittaluga and Seminara 2003). These are complicated models capable of providing

doi: 10.2166/nh.2005.031

spatial and temporal variations of the state variables. For these models to be effective, detailed information on topographic and soil characteristics of a watershed, and temporal and spatial variation of precipitation, should be available at a grid scale dictated by the numerical mesh. Such data is rarely available. Furthermore, the assumptions in the derivation of these equations significantly simplify the actual physical processes (Tayfur 2003). In addition, it is likely to have convergence and stability problems in the numerical solution of these highly non-linear equations. Hence, they, in the end, may not offer much advantage over the lumped (black-box) models.

In the black-box models, attention is concentrated on determining catchment response to inputs rather than on a detailed description and modelling of the catchment. In other words, the black-box analysis is applied as an investigation of the behaviour of the catchment system with no reference to its inner properties and the physical laws governing its processes. That means that all the information concerning the behaviour of the catchment of interest is represented in the input–output data set. In this direction, many studies have been carried out (Bruce 1975; Rendon-Herrero 1978; Williams 1978). Most of these models involve the estimation of excess precipitation and separation of base flow and are based on the assumption that concentration varies linearly with excess rainfall. As non-linear black-box models of catchment, functional series have been applied for runoff prediction (Amorocho 1967; Muftuoglu 1984, 1991; Xia *et al.* 1997). Only recently have Guldal and Muftuoglu (2001) developed a non-linear black box model based upon two-dimensional unit sediment graph theory for TSS prediction. The model is able to represent the overall erosive behaviour of a catchment without requiring excess rainfall and direct runoff. The model successfully predicted TSS using the current and antecedent precipitation. This model is proved to be superior to existing black-box (conventional second-order functional series) models. However, the model is based upon the assumptions that the catchment is time-invariant and precipitation has a uniform spatial distribution.

Tayfur (2003) summarised the basic reasons behind the difficulties in all these aforementioned modelling efforts of sediment transport. Sediment transport is a complicated problem. There is still a lack of well defined strong correlation between sediment concentration and a dominant variable. Due to the stochastic nature of sediment movement, it is difficult to precisely define at what flow condition a sediment particle will begin to move. In fluvial hydraulics, the boundary is movable and the resistance to flow is variable and there is a lack of a reliable and consistent method for the prediction of the variation of roughness coefficient. There is an inability to predict bed forms on a sound theoretical basis—even if the bed form is given the form roughness still varies significantly. Sediment discharge depends on the gradation and shape of sediment, percent of bed surface covered by coarse material, availability of bed material for transport, variations of hydrological cycle, rate of supply of fine material or wash load, water temperature, channel pattern and bed configuration and strength of turbulence. Tremendous uncertainties are involved in estimating sediment loads at different flow and sediment conditions under different hydrologic, geologic and climatic constraints.

For that reason, as an alternative to the existing models, an ANN model to estimate TSS in natural rivers is proposed. The ANN is a black-box model capable of solving highly non-linear complex problems. Mathematically ANNs can be viewed as a universal approximator. They have the ability to identify a relationship from given patterns and this enables them to solve large-scale complex problems, such as pattern recognition, non-linear modelling, classification, association, and control. In recent years, ANNs have been applied to solve many hydrological problems (ASCE Task Committee 2000b). With regard to sediment transport, Tayfur (2002b) applied ANNs to predict sheet sediment bed loads from plots having varying gradients under varying rainfall intensities. Cigizoglu (2002, 2004) applied ANNs to estimate TSS in natural streams using flow discharge data in the input vector.

Since Cigizoglu (2002, 2004) undertook estimation of TSS using flow discharge data this study is limited to using only precipitation data in the input vector of the ANN model to estimate TSS. To predict TSS via precipitation data is important. According to Sivapalan *et al.* (2003), one of the main issues for hydrologists today is the prediction of the hydrologic variables in ungauged or poorly gauged watersheds. Drainage basins in many parts of the world are ungauged or poorly gauged and in some cases existing measurement networks are declining (Sivapalan *et al.* 2003). There is an extensive network of precipitation gauges in most parts of the world. To predict hydrologic variables through precipitation data would therefore be very beneficial for hydrologists.

Jain and Indurthy (2003), in their comparative study, have shown that ANNs outperformed the unit hydrograph model and statistical models of linear multiple regression (LMR) and non-linear multiple regression (NLMR) in rainfall–runoff modelling. Similarly, Cigizoglu (2002, 2004) has shown the superiority of ANN over LMR, NLMR, autoregressive (AR) and sediment rating curve (SRC) models in TSS modelling. Since the performance of ANN was tested against deterministic, statistical and stochastic models, in this study the performance of ANNs will be tested against the non-linear black-box model based on two-dimensional sediment graph theory of Guldal and Muftuoglu (2001).

Artificial neural network (ANN)

The hydraulic and/or hydrologic applications of ANN generally consider a three- layer feedforward network, as shown in Figure 1. In a feedforward ANN, the input quantities (x_i) are fed into the input layer neurons that, in turn, pass them on to the hidden layer neurons (z_i) after multiplication by connection weights (v_{ij}) (Figure 1). A hidden layer neuron adds up the weighted input received from each input neuron ($x_i v_{ij}$) and associates it with a bias (b_j) (i.e. $net_j = \sum x_i v_{ij} - b_j$). The result ($net_j$) is then passed on through a non-linear transfer function to produce an output.

The learning of ANNs is generally accomplished by the most commonly used supervised training algorithm of the back-propagation algorithm. The objective of the back-propagation algorithm is to find the optimal weights that would generate an output vector $\boldsymbol{Y} = (y_1, y_2, \ldots, y_p)$ as close as possible to the target values of the output vector $\boldsymbol{T} = (t_1, t_2, \ldots, t_p)$ with the selected accuracy. The optimal weights are found by minimising a predetermined error function (E) of the following form (ASCE Task Committee 2000a):

$$E = \sum_P \sum_p (y_i - t_i)^2 \qquad (1)$$

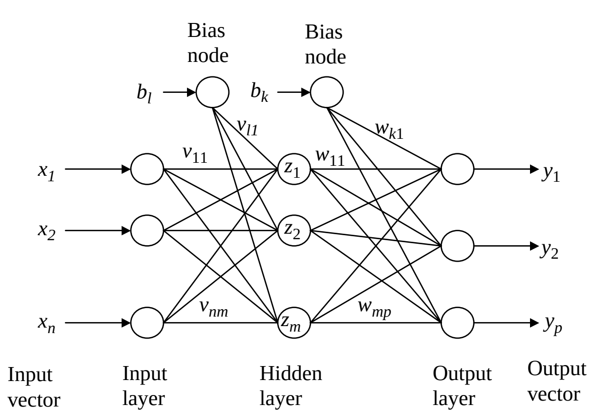

Figure 1 Schematic representation of a three-layer feed-forward artificial neural network

where y_i = the component of an ANN output vector $\boldsymbol{Y}$; t_i = the component of a target output vector $\boldsymbol{T}$; p = the number of output neurons; and P = the number of training patterns.

In the back-propagation algorithm, the effect of the input is first passed forward through the network to reach the output layer. After the error is computed, it is then propagated back towards the input layer with the weights being modified. The gradient-descent method, along with the chain rule of differentiation, is employed to modify the network weights as (ASCE Task Committee 2000a)

$$\Delta v_{ij}(n) = -\delta \frac{\partial E}{\partial v_{ij}} + \alpha \Delta v_{ij}(n-1) \tag{2}$$

where $\Delta v_{ij}(n)$ and $\Delta v_{ij}(n-1)$ = the weight increments between node i and j during the nth and $(n-1)$th pass or epoch; δ = the learning rate; and α = the momentum factor.

The details of the theory of ANNs are given in ASCE Task Committee (2000a).

Non-linear black-box model (2D-USGT)

Guldal and Muftuoglu (2001) developed a non-linear black-box model based upon two-dimensional unit sediment graph theory (2D-USGT) on the basis of the qualitative description of the catchment. In their model, the output (the system's response) is related to input by functional equations whose parameters—called response functions—are determined by a calibration procedure based upon the historical data. The discrete form of the model is given as (Guldal and Muftuoglu 2001)

$$y_n = \sum_{i=1}^{n} \sum_{j=i}^{n} h_{i,j} x_i x_j \tag{3}$$

where y_n = sediment concentration; n = number of intervals in the memory, x_i = effective precipitation in the ith interval; x_j = effective precipitation in the jth interval; and $h_{i,j}$ = coefficient representing the contribution rate from the ith interval under the influence of the unit effective precipitation in the jth interval.

The set of coefficients $h_{i,j}$, called the '*response function*', can be interpreted as a family of sediment graphs, each resulting from the unit effective precipitation in an interval under the influence of the unit effective precipitation in an antecedent interval. This is referred to as the 2D unit sediment graph.

The modified form of Equation (3) is the special second-order functional series. The discrete form of the modified model is expressed as (Guldal and Muftuoglu 2001)

$$y_n = \sum_{i=1}^{l} h_i x_i + \sum_{i=1}^{k} \sum_{j=i}^{k} h_{i,j} x_{i+l} x_{j+l} \tag{4}$$

where l and k = the numbers of intervals in the delayed and immediate response periods respectively; $n = l + k$; and h_i, $h_{i,j}$ = ordinary finite-period and 2D finite-period unit sediment graph, respectively (i.e. h_i represent the catchment response to the unit effective rainfall in the ith interval and $h_{i,j}$ represent the response to the unit effective rainfall in the $(l+j)$th interval under the influence of the unit effective rainfall in the $(l+i)$th interval).

The model given by Equation (4) can only be used for data generation. Nevertheless, if there is a lag L between the effect in the final unit interval of the memory and the response, the models representing the prediction lead time can then be applied as a predictor. For this purpose, the models are applied repeatedly, for successive predictions or generations of TSS, by shifting the time origin. Thus, with a new parameter of a shift counter (c), Equation (4) can be expressed by the following form that has a generation and prediction capability

(Guldal and Muftuoglu 2001):

$$y_{n+L+c} = \sum_{i=1}^{l} h_i x_{i+c} + \sum_{i=1}^{k} \sum_{j=i}^{k} h_{i,j} x_{i+l+c} x_{j+l+c} \quad (5)$$

The details of the 2D-USGT model can be obtained from Guldal (1997) and Guldal and Muftuoglu (2001).

Precipitation and TSS data

The data used in this study were directly obtained from the illustrations given by Brown and Choate (1989). The same data were used by Guldal and Muftuoglu (2001) to calibrate and validate their model of 2D-USGT. A technical report written by Carey *et al.* (1988) provided a hint that the data belong to a medium-sized catchment in the Tennessee Basin. After a few futile attempts, it was decided not to carry out further investigation on the identity and properties of the catchment. This is because the ANN model developed in this study does not require the properties of the catchment.

The data contained 240 d of measured precipitation (P) and TSS data (Figures 2(a, b)). Table 1 presents cross-correlation values between the two variables. According to Table 1, after time lag of 4 d, the cross-correlation is close to zero. That means, after a time lag of 4 d, there is no significant effect of P on TSS. This fact was preserved in the application of the ANN model. For example, the input vector contained, in one of the applications of the ANN

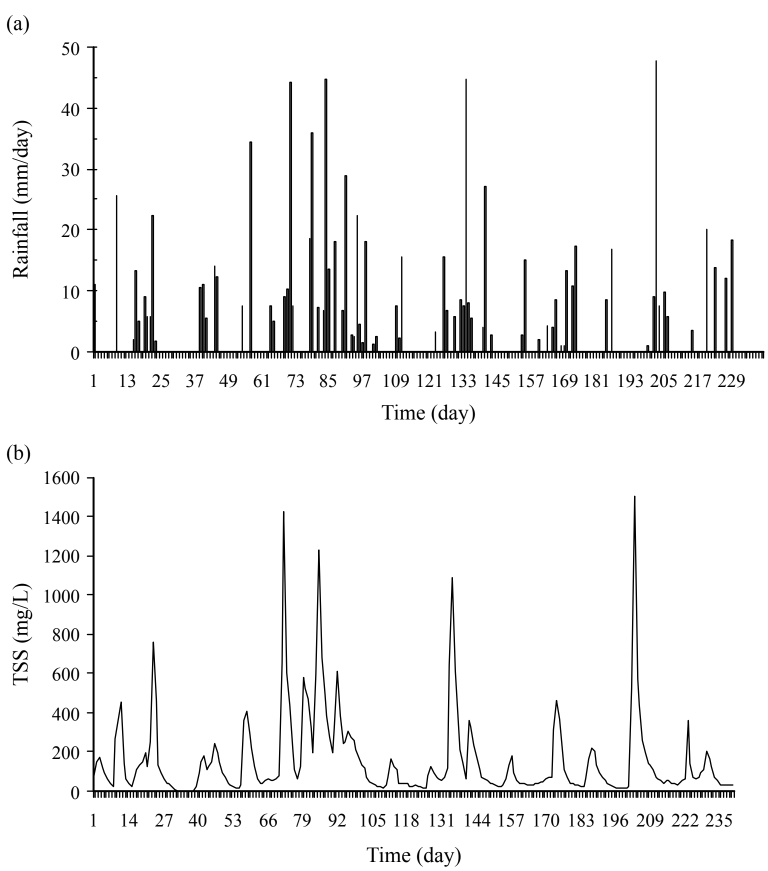

Figure 2 Measured data: (a) precipitation and (b) TSS

Table 1 Cross-correlations between precipitation (P) and TSS

Time lag zero-day	(t)	$r_{P,TSS,0}$	0.3664
Time lag one-day	($t-1$)	$r_{P,TSS,1}$	0.7209
Time lag two-days	($t-2$)	$r_{P,TSS,2}$	0.4735
Time lag three-days	($t-3$)	$r_{P,TSS,3}$	0.2709
Time lag four-days	($t-4$)	$r_{P,TSS,4}$	0.1104
Time lag five-days	($t-5$)	$r_{P,TSS,5}$	0.0505

model, P data at present (t) and at time lags $t-1$, $t-2$, $t-3$ and $t-4$ (Table 1). Table 2 presents the statistical parameters (mean,$\bar{x}$; standard deviation, s_x; skewness coefficient, c_{sx}; coefficient of variation, c_{vx}; overall minimum, $x_{\min}$ and overall maximum, $x_{\max}$) of the P-TSS data. As seen in Table 2, both data sets have comparable low values of coefficients of variation and skewness.

Application of ANN model

The developed ANN model was applied to estimate TSS from P data. For this purpose, many different cases of input vectors were tried. In all these cases, the constructed network employed a single hidden layer (Figure 1) and the sigmoid activation function (Tayfur 2002b). Due to the nature of the sigmoid function, all external input and output values before passing them into the network were standardised by Equation (6) (Tayfur *et al.* 2005) as

$$x_i = 0.1 + \frac{0.8(x_i - x_{\min_i})}{(x_{\max_i} - x_{\min_i})} \tag{6}$$

where $x_{\max_i}$ and $x_{\min_i}$ are the maximum and minimum values of the ith neuron in the input layer for all the feed data vectors, respectively.

Equation (6) compresses all the data into the range of 0.1–0.9 to overcome problems associated with upper-limit and lower-limit saturation. Note that, without standardisation, large values input into an ANN would require extremely small weighting factors to be applied and this could cause a number of problems (Dawson and Wilby 1998).

Before starting the training procedure, small random values of −1.5–1.5 for network connection weights and 1.0 for biases were assigned. The network was successfully trained with 20 000 iterations, and 0.04 and 0.1 values for the learning rate and momentum factor terms, respectively.

Several cases of input vector were tried so as to find the optimal number of neurons in the input layer. Table 3 summarises the properties of network structure for each case. The optimal number of neurons in the hidden layer of the network (Figure 1) was obtained by the trial and error procedure for each case. According to Table 3, the best result was obtained for the third case where present and past 4-d values of the P data were used in the input vector. This is consistent with the information in Table 1 stating that after a time lag of 4 d, the cross-correlation is close to zero. In this particular case, the network had 5 input neurons, 8 hidden neurons and 1 output neuron (Table 3).

Table 2 Statistical parameters of measured data (mean, $\bar{x}$; standard deviation, s_x; skewness coefficient, c_{sx}; coefficient of variation, c_{vx}; overall minimum, $x_{\min}$ and overall maximum, $x_{\max}$)

	$x_{\min}$	$x_{\max}$	$\bar{x}$	s_x	c_{vx}	c_{sx}
P (mm/d)	0	48.2	6.20	9.36	2.14	3.05
TSS (mg/L)	1	1500	156.03	216.46	1.39	3.24

Table 3 Different P input vectors to ANN model (P: precipitation; t: present time, $t-1$: time lag 1 d, ..., $t-5$: time lag 5 d)

	Number of neurons in layers				
Input vector	**Input**	**Hidden**	**Output**	R^2	***MAE*** **(mg/L)**
P_t, P_{t-1}, P_{t-2}	3	5	1	0.83	56.46
P_t, P_{t-1}, P_{t-2}, P_{t-3}	4	6	1	0.90	38.42
P_t, P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4}	**5**	**8**	**1**	**0.91**	**34.22**
P_t, P_{t-1}, P_{t-2}, P_{t-3}, P_{t-4}, P_{t-5}	6	10	1	0.89	46.79

Several cases of data partitioning into training and testing periods were tried in order to obtain the optimal period for training the ANN model. Table 4 presents 7 different partitioning cases and the related error measures for each case. Tables 5 and 6 present the statistical parameter values for the training and testing periods for each case in Table 4. According to Table 4, minimum error is obtained for Case 6 where 150 d of data were used in the network training. Employing more than 75 d in the training period results in satisfactory TSS estimations (Table 4). Case 1 and Case 2 in Table 4 produced poor results. This is because of two main reasons. First, 60 d, especially 30 d, of period is too short for the ANN model to capture the relation between the input (P) and output (TSS) variables. Second, though the other statistical parameter values are comparable for the training and testing periods, the values of the parameter *range* ($x_{\min} - x_{\max}$) for the training periods in Case 1 and Case 2 are quite a bit lower than the corresponding values for the testing periods (Tables 5 and 6). According to Table 4, using 120 d of data (Case 5) or 180 d of data (Case 7) in the training periods does not affect the performance of ANN model significantly.

Table 4 Different data partitioning into training and testing periods

	Training period (d)	**Testing period (d)**	R^2	***MAE*** **(mg/L)**
Case 1	30	210	0.696	77.58
Case 2	60	180	0.708	70.00
Case 3	75	165	0.888	52.69
Case 4	90	150	0.904	40.30
Case 5	**120**	**120**	**0.910**	**41.52**
Case 6	150	90	0.907	34.22
Case 7	180	60	0.919	38.50

Table 5 Statistical parameters for precipitation (mm/day) for the cases in Table 4

	Training period						Testing period					
	$x_{\min}$	$x_{\max}$	$\bar{x}$	s_x	c_{vx}	c_{sx}	$x_{\min}$	$x_{\max}$	$\bar{x}$	s_x	c_{vx}	c_{sx}
Case 1	**0.0**	**25.9**	3.35	7.12	2.13	2.53	**0**	**48.2**	4.45	9.61	2.16	3.08
Case 2	**0.0**	**34.7**	**3.58**	**7.59**	**2.12**	**2.74**	**0**	**48.2**	**4.60**	**9.84**	**2.14**	**3.06**
Case 3	**0.0**	**44.6**	**4.75**	**9.81**	**2.07**	**2.77**	**0**	**48.2**	**4.45**	**9.57**	**2.15**	**3.05**
Case 4	0.0	45.0	5.56	10.77	1.94	2.56	0	48.2	3.76	8.50	2.26	3.38
Case 5	0.0	45.0	4.93	9.94	2.02	2.96	0	48.2	3.68	8.57	2.33	3.65
Case 6	**0.0**	**45.0**	**5.09**	**10.18**	**2.00**	**2.73**	**0**	**48.2**	**3.03**	**7.57**	**2.50**	**3.99**
Case 7	0.0	45.0	4.74	9.59	2.02	2.85	0	48.2	3.05	8.38	2.75	4.13

Table 6 Statistical parameters for TSS (mg/L) for the cases in Table 4

	Training period						Testing period					
	x_{min}	x_{max}	$\bar{x}$	s_x	c_{vx}	c_{sx}	x_{min}	x_{max}	$\bar{x}$	s_x	c_{vx}	c_{sx}
Case 1	**1**	**760**	139.1	170.9	1.23	2.18	**1**	**1500**	159.1	224.5	1.41	3.24
Case 2	**1**	**760**	**123.5**	**142.6**	**1.15**	**2.18**	**10**	**1500**	**167.8**	**237.9**	**1.42**	**3.10**
Case 3	**1**	**1420**	**159.5**	**221.9**	**1.39**	**3.21**	**10**	**1500**	**155.2**	**217.1**	**1.40**	**3.24**
Case 4	1	1420	209.5	253.9	1.21	2.35	10	1500	123.9	186.5	1.51	4.32
Case 5	1	1420	179.8	229.8	1.28	2.68	10	1500	132.5	203.7	1.54	4.07
Case 6	**1**	**1420**	**183.4**	**232.0**	**1.26**	**2.59**	**10**	**1500**	**116.9**	**190.2**	**1.63**	**5.07**
Case 7	1	1420	167.9	218.4	1.30	2.77	10	1500	123.1	215.8	1.75	4.86

Figure 3(a) presents measured data versus ANN-predicted data for the case where present and past 4-d values of P data (Table 3) were used in the input vector. In this case (Case 6 in Tables 4, 5 and 6), 150 data sets were used for the model training and the remaining 90-d TSS data were predicted by the ANN model. As seen in Figure 3(a), the slope of the regression line is close to 1 and the intercept is close to 0. The coefficient of determination (R^2) is 0.91 and the mean absolute error (*MAE*) is 34.22 mg/L. It can be concluded that the ANN successfully estimated the measured TSS data. Figure 3(b) also shows the trend of the

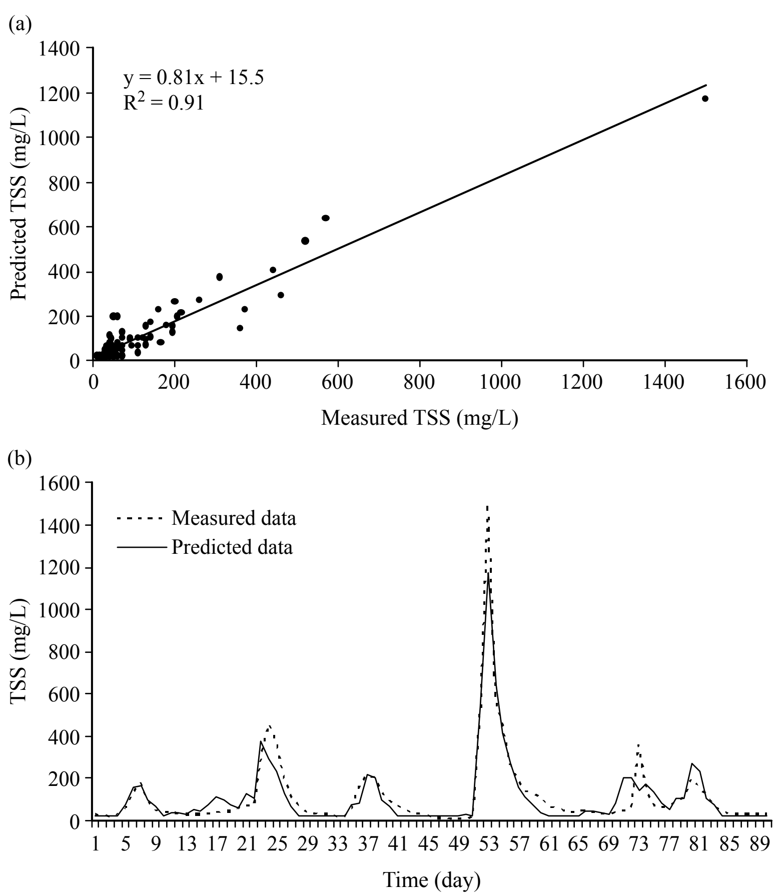

Figure 3 (a) Measured TSS data versus ANN-predicted TSS. (b) Simulating trend of measured TSS data by the ANN model (ANN model used present and past 4 d of precipitation in the input vector)

simulation. As seen, the model closely captured the trend of the measured data. Only in one observation, namely the observation on day 53 having an extreme value, did it slightly underestimate the measured data.

Several scenarios were tried to estimate TSS from past TSS values by ANNs. The best result was obtained when the past 6 measured TSS were used in the input vector with $R^2 = 0.67$. Also, in one scenario, TSS at $(t-1)$ (one day lag) was added to the input vector of the present and past 4 d of P data to estimate TSS at the present time (TSS_t). It was seen that incorporating TSS at $(t-1)$ did not improve the results significantly ($R^2 = 0.92$).

Application of 2D-USGT model

The 2D-USGT model was also applied to estimate TSS data from P data. In order to be consistent with the ANN model, the 2D-USGT model was calibrated with the first 150 sets of data. Then it was tested with the remaining data. Cross-correlation data in Table 1 were used to determine the memory parameters of the model.

Figures 4(a, b) present the prediction of TSS data by the 2D-USGT model having memories of 4 d and 1 d (linear and non-linear parts of total memory, i.e. $l = 4$, $k = 1$, that is $n = 5$ d of total memory). In other words, in this case, the model employed present and past 4 d of precipitation in the input vector. As seen in Figure 4(a), $R^2 = 0.73$ and $MAE = 68.2$, as opposed to $R^2 = 0.91$ and $MAE = 34.22$ in the case of the ANN model (Figure 3(a)). Also, as seen in Figure 4(b), the 2D-USGT model could not sufficiently capture the trend of the measured data—underpredicting extreme values and overpredicting low values.

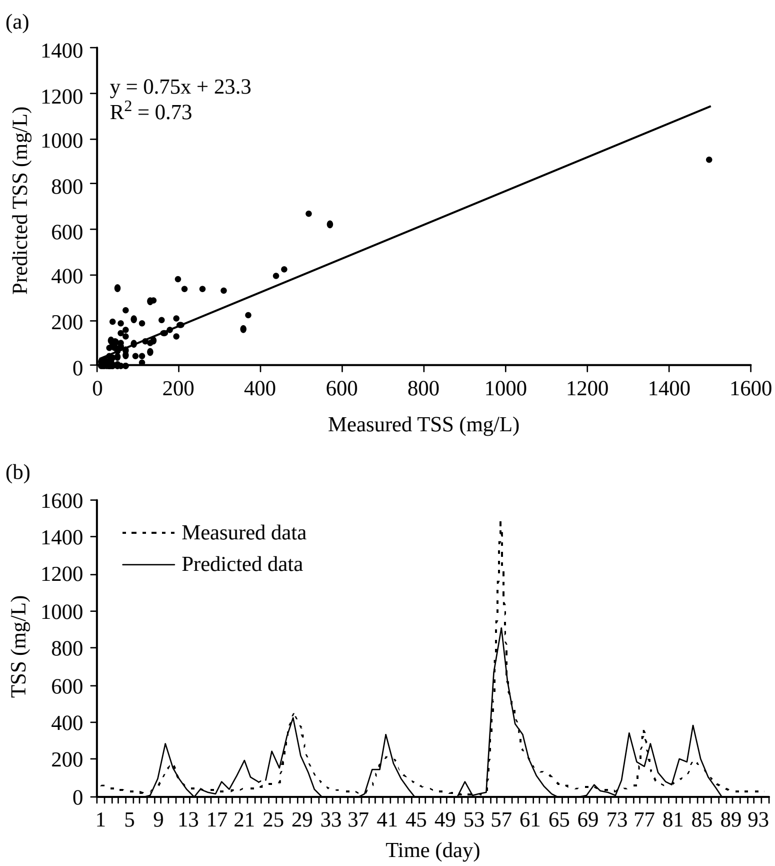

Figure 4 (a) Measured TSS data versus 2D-USGT-predicted TSS. (b) Simulating trend of measured TSS data by the 2D-USGT model (2D-USGT model used present and past 4 d of precipitation in the input vector)

Concluding remarks

This study presented satisfactory predictions of daily total suspended sediment (TSS) in natural rivers by artificial neural networks (ANNs) from precipitation (P) data. This result has an important implication for a basin where there is a partial gauging station for hydrological variables. One of the main issues for hydrologists today is the prediction of the hydrologic variables in ungauged, or poorly gauged, watersheds. Drainage basins in many parts of the world are ungauged or poorly gauged and in some cases existing measurement networks are declining. On the other hand, there is an extensive network of precipitation gauges in most parts of the world. Therefore, prediction of TSS through P data becomes very beneficial for hydrologists.

Investigation results revealed that at least present and past 3 d P data should be used in the input vector of ANN model for satisfactory TSS estimation. Cross-correlation values between the input and output variables can shed light on finding the optimal number of input neurons in the input layer of the ANN model.

Investigation results also revealed that, in order to obtain satisfactory TSS estimations by the ANN model, there should be a sufficient length of data record for the training period so that the ANN model can capture the relation between the input and output variables. In this study, this period is found to be more than 75 d for P-TSS study. Furthermore, it is found out that the statistical parameter values for the training and testing periods should be comparable for successful ANN applications. In particular, in partitioning data into two periods of training and testing, special attention should be given to the parameter *range*. The values of this parameter should be comparable for the two periods. Otherwise, as is presented in this study, ANNs would perform poorly in estimating TSS. This is consistent with the fact that ANNs are not good extrapolators (ASCE Task Committee 2000a).

ANNs showed a greater performance in the prediction of TSS using P data in the input vector than the non-linear black-box model based upon two-dimensional unit sediment graph theory (2D-USGT). The constructed three-layer feedforward network makes no assumption with regard to the hydrological processes. On the other hand, the 2D-USGT model makes assumptions that the catchment is time-invariant and precipitation has a uniform spatial distribution. That may be the reason for the poorer performance of the 2D-USGT model in predicting TSS.

ANNs are black-box models that can solve non-linear complex problems, such as the suspended sediment transport in natural streams, when provided with sufficient historical data of the process. The ANN does not make any assumption, as opposed to deterministic, statistical and stochastic models, on the physics of the process so as to simplify the process. However, ANN does not reveal any explicit mathematical relation between the input and output variables of the physical process. Hence, one is not able to gain much insight into understanding the physics of the process. Furthermore, though ANN has a very strong interpolation capacity, it lacks the extrapolation capability, especially for the cases for which it is not trained. Nevertheless, as opposed to the other models, it is quite simple to construct and train the model.

This study showed that ANNs can be successfully employed to estimate the complex non-linear river suspended sediment process in situtations where explicit knowledge of internal sub-process is not required. Predicting TSS loads that are required for the planing and operation of a wide spectrum of water resources structures from precipitation data makes ANNs a very promising planing and management modelling tool for the hydrologists.

References

Amorocho, J. (1967). The nonlinear prediction problem in the study of runoff cycle. *Wat. Res. Res*, **3**(3), 861–880.

ASCE Task Committee (2000a). Artificial neural networks in hydrology. I: Preliminary concepts. *J. Hydrol. Engng., ASCE*, **5**(2), 115–123.

ASCE Task Committee (2000b). Artificial neural networks in hydrology, II: Hydrologic applications. *J. Hydrol. Engng. ASCE*, **5**(2), 124–137.

Brown, R.T. and Choate, K.D. (1989). Daily suspended sediment model for water resources management. In *Sediment Transport Modelling*, S.S. Wang (ed.), ASCE, New York, pp. 642–647.

Bruce, R.R. (1975). Model for runoff of pesticides in small upland watersjed. *J. Environ. Qual.*, **4**(4), 541–548.

Carey, W.P., Brown, R.T. and Chathan, G.C. (1988). History of suspended sediment data collection and inventory of available data for the Tennessee and Cumberland river basins. *USGS Open File Rep.*, 88–497, Tennessee Technical University, Cookeville, TN.

Cigizoglu, H.K. (2002). Suspended sediment estimation and forecasting using artificial neural networks. *Turkish J. Engng. Environ. Sci., TUBITAK*, **26**, 15–25.

Cigizoglu, H.K. (2004). Estimating and forcasting of daily suspended sediment data by multi-layer perceptrons. *Adv. Wat. Res*, **27**(2), 185–195.

Dawson, W.C. and Wilby, R. (1998). An artificial neural network approach to rainfall-runoff modelling. *Hydrol. Sci. J.*, **43**(1), 47–66.

Guldal, V. (1997). A nonlinear system model for prediction of suspended sediment load of rivers. PhD thesis, Istanbul Technical University, Turkey.

Guldal, V. and Muftuoglu, R.F. (2001). 2D unit sediment graph theory. *J. Hydrol. Engng., ASCE*, **6**(2), 132–140.

Guo, Q.-C. and Jin, Y.-C. (2002). Modelling nonuniform suspended sediment transport in alluvial rivers. *J. Hydraul. Engng., ASCE*, **128**(9), 839–847.

Jain, A. and Indurthy, K.V.P. (2003). Comparative analysis of event-based rainfall-runoff modeling techniques—deterministic, statistical, and artificial neural networks. *J. Hydrol. Engng., ASCE*, **8**(2), 93–98.

Muftuoglu, R.F. (1984). New models for nonlinear catchment analysis. *J. Hydrol.*, **73**, 335–357.

Muftuoglu, R.F. (1991). Monthly runoff generation by non-linear mpodels. *J. Hydrol.*, **125**, 277–291.

Pittaluga, M.B. and Seminara, G. (2003). Depth-integrated modelling of suspended sediment transport. *Wat. Res. Res.*, **39**(5).

Rendom-Herrero, O. (1978). Unit sediment graph. *Wat. Res. Res.*, **14**(5), 889–901.

Sivapalan, M., Takeuchi, K., Franks, S.W., Gupta, V.K., Karambiri, H., Lakshmi, V., Liang, X., McDonnell, J.J., Mendiodo, E.M., O'Connell, P.E., Oki, T., Pomery, J.V., Schertzer, D., Uhlenbrook, S. and Zehe, E. (2003). *IAHS Decade on Prediction in Ungauged Basins (PUB), 2003–2012: Shaping An Exciting Future for the Hydrological Sciences*, **48**(6), 857–880.

Tayfur, G. (2001). Modelling two dimensional erosion process over infiltrating surfaces. *J. Hydrol. Engng.*, **6**(3), 259–262.

Tayfur, G. (2002a). Applicability of sediment transport capacity models for non-steady state erosion from steep slopes. *J. Hydrol. Engng., ASCE*, **7**(3), 252–259.

Tayfur, G. (2002b). Artificial neural networks for event-based non-steady state sheet erosion. *Hydrol. Sci, J., IAHS*, **47**(6), 879–892.

Tayfur, G. (2003). Modelling sediment transport. In *Watershed Hydrology (Int. Conf. on Water and Environment, 15–18 December, Bhopal, India)*, V.P. Singh and R.N. Yadava (eds), Allied Publishers PVT Limited, New Delhi, pp. 353–375.

Tayfur, G., Swiatek, D., Wita, A. and Singh, V.P. (2005). Case study: Finite element method and artificial neural network models for flow through Jeziorsko earthfill dam in Poland. *J. Hydraul. Engrg. ASCE*, **131**(6), 431–440.

Williams, J.R. (1978). A sediment graph model based on an instantaneous unit sediment graph. *Wat. Res. Res.*, **14**(4), 659–664.

Xia, J., O'Connor, K.M., Kachroo, R.K. and Liang, G.C. (1997). A non-linear perturbation model considering catchment wetness and its application in river flow forecasting. *J. Hydrol.*, **200**, 164–178.

Yang, C.T. (1996). *Sediment Transport Theory and Practice*, McGraw-Hill, New York.