Turkish J. Eng. Env. Sci.
26 (2002) , 15 – 25.


# Suspended Sediment Estimation and Forecasting using Artificial Neural Networks

**Hikmet Kerem CIĞIZOĞLU**
*İstanbul Technical University, Civil Engineering Faculty,
Maslak, 80626, İstanbul, TURKEY*

Received 02.10.2000

**Abstract**

The methods available in the literature for sediment concentration estimation are complicated and time consuming and necessitate cumbersome parameter estimation procedures. In this study, artificial neural networks (ANNs) are used to forecast and estimate sediment concentration values. The forecasting results obtained using previously observed sediment values were close to the real ones. The sediment concentration estimation, on the other hand, using only observed river flow values and the previous sediment value in a nearby river as input, provided realistic approximations in terms of mean squared error (MSE) and total sediment amount. The ANN estimates are compared also with corresponding classical regression ones and found to be significantly superior.

**Key Words:** Suspended sediment, Forecasting, River flow

## Introduction

Sediment yield is defined as the total sediment outflow from a watershed measurable at a point of reference during a specified period of time. The sediment outflow from the watershed is induced by processes of detachment, transportation, and deposition of soil materials by rainfall and runoff. Estimates of sediment yield are required in a wide spectrum of problems such as the design of reservoirs and dams, transport of sediment and pollutants in rivers, lakes and estuaries, design of stable channels, dams and debris basins, undertaking cleanup following floods, protection of fish and wildlife habitats, determination of the effects of watershed management, and environmental impact assessment (Singh et al., 1988). Fine sediment has long been identified as an important vector for the transport of nutrients and contaminants such as heavy metals and micro-organics. Suspended sediment is important in its own right, since its presence or absence exerts an important control on geomorphological and biological processes in rivers and estuaries.

Sediment yield Y(t) at a given point in space (say, a watershed outlet) can be represented as

$$Y(t) = \bar{Y}(t) \; + \; \varepsilon(t) \tag{1}$$

in which $\bar{Y}(t)$ is the mean value or deterministic component of Y(t), and $\varepsilon$(t) is the deviation from or fluctuation around the mean value or stochastic component of Y(t). The relative contribution $\bar{Y}(t)$ or $\varepsilon(t)$ to Y(t) depends on the watershed and space –time scales. Clearly, Y(t) encompasses the full range of variability from being entirely deterministic to being entirely stochastic. All sediment models are special cases of (1).

The deterministic models can be distinguished as empirical and conceptual. These models usually require long data records, so that average annual sediment yield can be determined. The conceptual models combine the mechanics of sediment transport with empirical relationships. Both the empirical and conceptual models approximate the physical processes controlling sediment yield (Singh et al., 1988).

Another way to represent the complex sediment behaviour is to interpret a sequence of sediment yield measurements as random. If the processes governing sediment yield are assumed to be stochastic, and thus governed by the laws of probability, the sediment yield can be described by a stochastic process and associated probability distributions (pdf).

The application of physics-based distributed process complex computer software programs is often problematic, due to the use of idealised sedimentation components, or the need for massive amounts of detailed spatial and temporal environmental data which is not available.

Many of the available techniques for time series analysis assume linear relationships among variables. In the real world, however, temporal variations in data do not exhibit simple regularities and are difficult to analyse and predict accurately. Linear recurrence and their combinations for describing the behaviour of such data are often found to be inadequate.

It seems necessary that nonlinear models such as artificial neural networks (ANNs), which are suited to complex nonlinear models, be used for the analysis of real world temporal data. ANN is a model inspired from the structure of the brain, is well suited to such tasks as pattern recognition, combinatorial optimisation, and discrimination. The ANN learns to solve a problem by developing a memory capable of associating a large number of input patterns with a resulting set of outputs or effects. The ANN develops a solution system by training on examples given to it. These tools contain no preconceived ideas about the manner in which a model ought to be structured or work. It also provides a flexible approach, with the power to provide different levels of generalisation, and can produce a reasonable solution from small data sets. The modeller has control over the data inputs and irrelevant variables can be identified or removed during the model building process.

There are numerous studies related to the application of ANNs to various problems frequently encountered in water resources. The nonlinear ANN approach was shown to provide a good representation of the rainfall-runoff relationship (Hsu et al., 1995, Minns and Hall, 1996). The application of the radial basis function type of ANNs to model the rainfall runoff process has also been examined (Fernando and Jayawardena, 1998, Mason et al., 1996). Tokar and Johnson (1999) employed neural network methodology to forecast daily runoff as a function of daily precipitation, temperature, and snowmelt for the Little Patuxent River in Maryland. Campolo et al. (1999 a,b) used ANNs to forecast river flows during heavy rainfall and low-flow periods. ANNs were also considered to be a powerful tool for use in various groundwater problems (Ranjithan et al., 1993, Rogers and Dowla, 1994). Raman and Sunilkumar (1995) investigated the use of ANNs in synthetic reservoir inflow series generation. Boogaard et al. (1998) introduced Auto-Regressive neural networks (ARNN) for the non-linear analysis and modelling of time series whereas See and Openshaw (1998) outlined a methodology incorporating the neural network and the fuzzy logic in forecasting problem. ANNs were also used in unit hydrograph derivation (Lange, 1998), regional flood frequency analysis (Hall and Minns, 1998), estimation of sanitary flows (Djebbar and Alila, 1998) and modelling hydraulic characteristics of severe contraction (Kheir El-Din, 1998). Abrahart (1998) presented an embedded solution for neural networks and the problem of accumulated error.

The application of ANNs to sediment concentration estimation is, however, not available in the literature. In this study, initially, ANNs are used to forecast the present or future sediment value using the past sediment values as input. Then the river flow values are used in the input layer to estimate the sediment value. In this part of the study, the training of the ANNs was carried out using the observed flow and sediment values in a nearby river, and during the testing stage the sediment values on other rivers were estimated using the observed flows. Finally the ANN estimation results were compared with the nonlinear regression outputs. The term forecasting is used in the study as in the case of having the same variable in both input and output layers. If the input layer contains variable(s) different than the output layer from the term estimation is preferred.

## The structure of ANNs

The learning process or training forms the interconnection between neurons. The strength of these interconnections is adjusted using an error convergence technique so that a desired output will be produced for a known input pattern.

Many training procedures are discussed in the literature. Error back propagation is one of the most commonly used procedures. The processing units are arranged in layers. The method is generally an iterative nonlinear optimisation approach using a gradient descent search method. Mason et al. (1996) and Fernando et al. (1998) have used radial basis function networks for training. They concluded that ANNs trained either using radial basis function or error back propagation provided comparable estimations.

Error back propagation provides a feed forward neural network, giving the capacity to capture and represent relationships between patterns in a given data sample (Eberhart and Dobbins, 1990). Detailed information about the structure of the error back propagation training procedure is provided in the appendix.

### Analysis of data

An extensive sediment monitoring network was established within the LOIS programme, involving the main tributaries of the River Humber (UK). One of the key objectives of the Land-Ocean Interaction Study (LOIS) established in 1992, was to quantify and characterise the flux of materials from river basins to oceans (Wass and Leeks, 1999). This provided an opportunity to deploy an extensive suspended sediment river network within the rivers of the study area, on a scale not previously attempted in the UK. A turbidity monitoring system was developed to provide a continuous record of suspended sediment concentration in the rivers, from which the fluxes were calculated. Linear relationships were established between suspended sediment concentration and turbidity to enable the conversion of nephelometric turbidity (NTU) to suspended sediment concentration (mg/l). The measurements were undertaken during the period October 1994-November 1997 (Wass and Leeks, 1999).

An overview of the river monitoring strategy is given by Leeks et al. (1997). The related information for the available flow and suspended data is presented in Tables 1 and 2. The suspended data is unfortunately not continuous throughout the observation period in all the five stations. The longest concurrent continuous observation periods for suspended sediment are from 07-11-1994, 21:15 to 22-01-1995, 11:00, totally 6,900 values, and from 24-12-1996, 12:45 to 20-02-1997, 08:00, totally 5,500 values, in Swale Thornton Manor and Ouse Skelton. In all the training and testing studies explained in the following sections, the data in this mentioned period has been used.

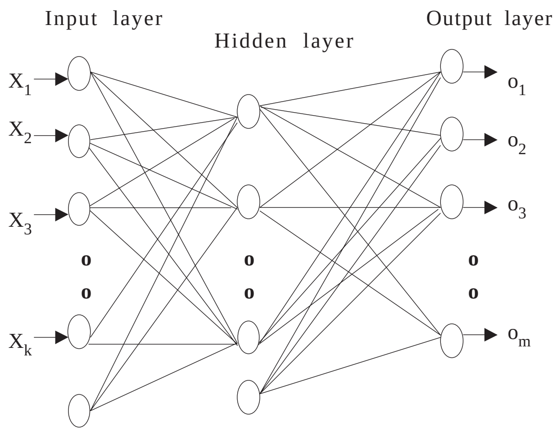


**Figure 1.** The structure of the artificial neural networks.

Since both sediment and flow values of the two rivers are skewed, the first order correlation coefficients for flow-sediment, flow-flow and sediment-sediment pairs are computed for logged values. In other words the logged sediment and flow values are assumed to be normally distributed. The results are presented in Table 3. The highest correlation, 0.98, is for the Ouse Skelton flow and the Swale Crakehill flow, whereas the lowest one, 0.64, is for the Ouse sediment and the Ouse flow.

Figures 2 to 5 illustrate the relation between suspended sediment and flow for the Ouse river and Swale river, respectively, both for natural and logged data. It is obvious that the relation between sediment and flow becomes relatively linear for the logged values. Double mass curves are obtained both for cumulative flow (in million $m^3$) and for cumulative sediment (in tonnes) comparing corresponding Swale River and Ouse River values (Figures 6 and 7). Both curves show clear homogeneity between both stations.

**Table 1.** River flow data

| River | Station | Observation period |
|---|---|---|
| Ouse | Skelton | from 01-09-93, 09:00 to 15-04-97, 03:45 |
| Swale | Crakehill | from 01-09-93, 09:00 to 14-04-97, 03:00 |
| Ure | Westwick Lock | from 01-09-93, 09:00 to 02-04-97, 08:00 |
| Nidd | Hunsingore Weir | from 01-09-93, 09:00 to 14-04-97, 03:00 |

**Table 2.** Suspended sediment concentration data

| River | Station | Observation period |
|---|---|---|
| Ouse | Skelton | from 03-10-94, 09:45 to 28-10-97, 16:00 |
| Swale | Thornton Manor | from 17-10-94, 21:15 to 11-09-97, 21:00 |
| Ure | Westwick Lock | from 09-06-94, 13:45 to 28-10-97, 14:00 |
| Nidd | Cowthorpe | from 10-08-94, 11:00 to 28-10-97, 13:00 |
| Swale | Catterick Bridge | from 13-10-94, 15:00 to 28-10-97, 15:00 |

**Table 3.** First order correlation coefficients ($r_1$) between stations.

| | Swale flow-Ouse flow | Swale sediment-Swale flow | Swale flow-Ouse sediment | Ouse sediment-Ouse flow | Swale sediment-Ouse sediment |
|---|---|---|---|---|---|
| $r_1$ | 0.98 | 0. 77 | 0.66 | 0.64 | 0.77 |

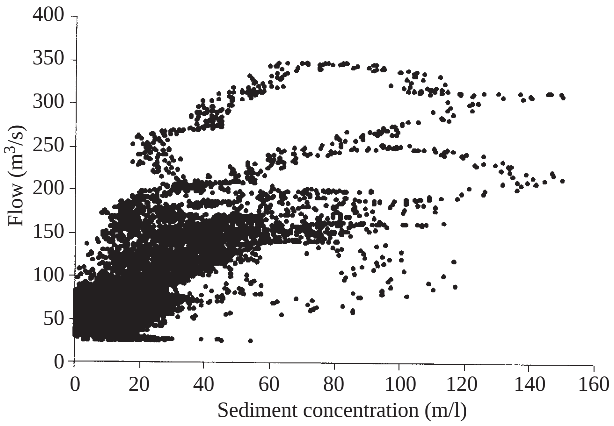


**Figure 2.** Suspended sediment versus logged flow data for Ouse river.

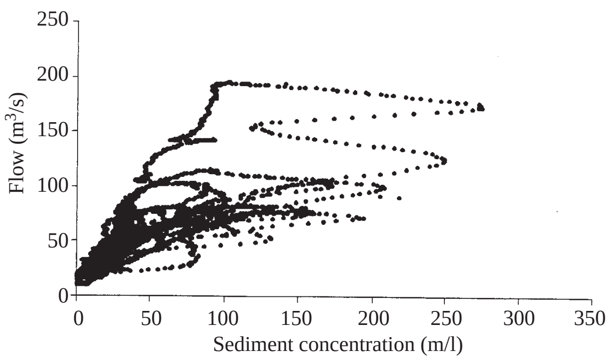


**Figure 3.** Suspended sediment versus logged flow data for Swale river.

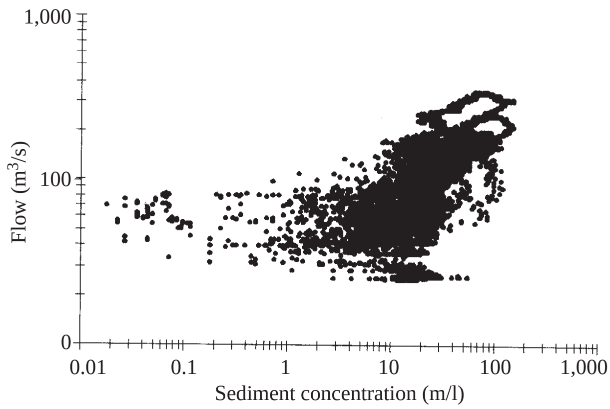


**Figure 4.** Logged suspended sediment versus logged flow data for Ouse river.

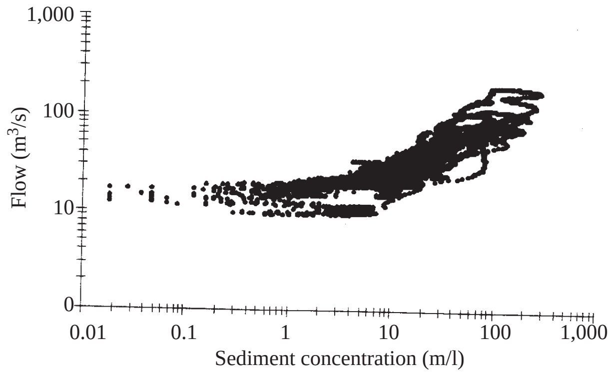


**Figure 5.** Logged suspended sediment versus logged flow data for Swale river.

## Application of ANNs to the data

A code in FORTRAN language was written following the steps explained in the appendix. The application of ANNs to forecast and estimate the suspended sediment concentration consisted of two steps. The first step was the training of the neural networks. The back propagation method was employed to train the ANNs. The determination of the number of hidden layers and the number of nodes both in the input layer in each hidden layer, which provided the best the training results, were the primary considerations in training procedure. The criteria for the evaluation of the training simulation was the final MSE value computed as presented in the appendix part. Initially randomly generated normal values, between –3 and 3, are assigned for correlation weights. The training input and output values are scaled between 0 and 1 simply using

$$[(x_i - x_{\min})/x_{\max} - x_{\min})]$$

where $x_{min}$ and $x_{max}$ represent the minimum and maximum of all the values in the input or the output layer. If the variable in each layer is different, e.g. if river flow is the input and the suspended sediment is the output, the scaling is realised for each layer separately. In the case of having the same variable in both input and output layers, i.e. in forecasting the suspended sediment using the previous sediment values as input, all the values in both layers are scaled together. During the training procedure the input layers are selected randomly. Once the training stage was completed, the testing stage began using the optimum values found for the number of nodes in each layer, the number of hidden layers, the learning rate, the momentum rate and the correlation weights.

The optimum learning, $\eta$, and the momentum, $\alpha$, rates were found after trying various values and observing the MSE produced at the end of the testing stage. It is seen that picking high values like 0.5 and 0.9 for $\eta$ and $\alpha$, as done by Raman and Sunilkumar (1995), throws the network into oscillations or saturates the node outputs. Saturation occurs when the net input to the function producing node output is a large value (either positive or negative) and variations in the input thus have little effect on the output (Eberhart and Dobbins, 1990). It is seen that $\eta$ and $\alpha$ should be decreased if the number of input and output layers are to be increased. The iteration number, on the other hand, increases by decreasing $\eta$ and $\alpha$ values. In this study, 0.02 and 0.01 values are found adequate for $\eta$ and $\alpha$, respectively. The testing stage results using these parameter values are presented in detail in the following part of the study.

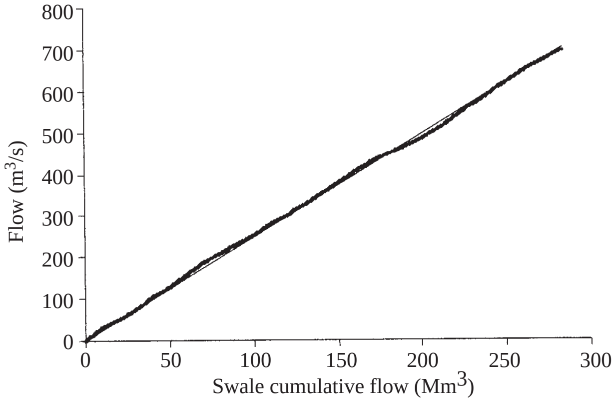


**Figure 6.** Double mass curve for river flow comparing Swale flow with Ouse flow for a common period of 75 days.

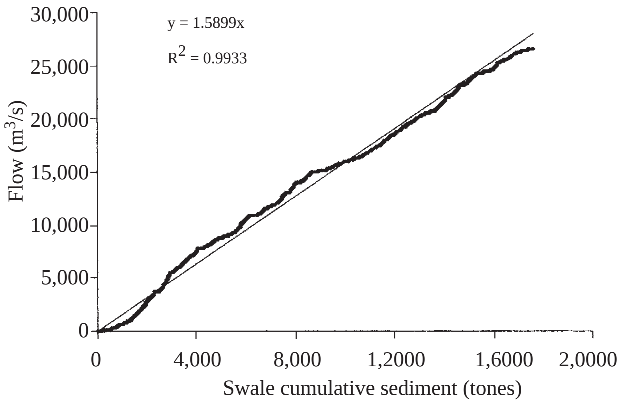


**Figure 7.** Double mass curve for sediment comparing Swale sediment with Ouse sediment for a common period of 75 days.

Initially, the applicability of the ANNs in suspended sediment concentration forecasting is investigated. The efficiency of ANNs in forecasting the suspended sediment concentration at present time t and at time steps ahead was the subject of the study. The first 5,900 suspended sediment values were considered for training the neural networks. The last 1,000 values (5,901th to 6,900th) covering a period of 10 and half days (from 07-01-1995 to 18-01-95) were then examined for testing. Different numbers of input values are considered in the input layer to forecast the unique sediment value at time t in the output layer. The last consecutive i (t-i, .., t-1) suspended sediment values consisted the input layer. The MSE for the testing period was 36.4, 33.33, 27.60, 43.47 and 32.40, for 4, 5, 6, 7 and 8 input values, respectively. Since 27.60 was the lowest MSE, the input layer node number was taken to equal 6 for the forecasting study. The optimum hidden layer node number is found 3, and increasing the hidden layer number did not provide any improvement in MSE during the testing stage. The forecasted values were very close to the observed ones (Fig. 8). The total sediment amount estimated by ANNs for this period was found to be 1,742 tonnes compared with the observed 1,776 tonnes. The performance of ANNs are compared with those of the classical Autoregressive Models. AR(1) and AR(**??**) models are used for forecasting sediment values. The parameters of these models are computed as mentioned by Box and Jenkins (1976) using the first 5,900 suspended sediment values. The forecasting was carried out for the last 1,000 values (5,901th to 6,900th) as done by ANNs. The obtained MSE values were 46.20 and 42.50 for AR(1) and Ar(**??**) models, respectively, being higher than the corresponding ANN results.

The same procedure is applied to forecast the suspended sediment concentration ahead at time t+24 using ANNs (Fig. 9 ). The MSE increased to 53.65 while forecasting the suspended sediment at t+24 and the total estimated sediment amount was 1,711 tonnes. It is obvious that sl increasing the time step ahead, the forecasting accuracy decreases.

Then ANNs are used to estimate the suspended sediment concentration values using the flow values as input. The first 5,900 flow and suspended sediment values observed in Ouse Skelton were used for the training stage. Again here, 6 input values in the input layer provided the minimum MSE during the testing stage. The six consecutive flow values (t-5, t-4, . . . , t) constituted the input layer to estimate the unique suspended sediment value at time t. The training was carried out for 13 to 5,900$^{th}$ value, whereas the testing covered the 5901$^{th}$ to 6,900$^{th}$ values (from 07-01-1995 to 18-01-95). Examining the results presented in Fig. 10 it can be said that the ANN estimated suspended sediment concentration time series does not capture the erratic behaviour of the observed one but instead provides an average approximation. The MSE and the ANN estimated sediment total at the end of the testing period were 89.70 and 1,856 tonnes, respectively. The difference between the estimated and observed sediment total (1,776 tonnes) was 4.5% . This shows that if ANNs are trained with river flow and suspended sediment data for the same time period and the same station, then the general behaviour of the suspended sediment time series in another time interval can be estimated using the available flow series as input. This

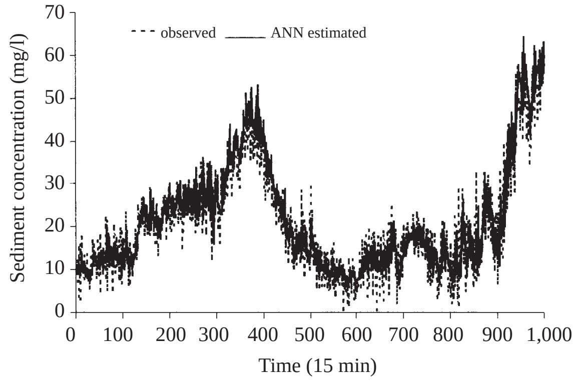


**Figure 8.** The forecasted Ouse Skelton suspended sediment values at time t for time period 07-01-1995 to 18-01-1995 using the last 6 (t-6, t-5, ...,t-1) values as input with a MSE=27.60.

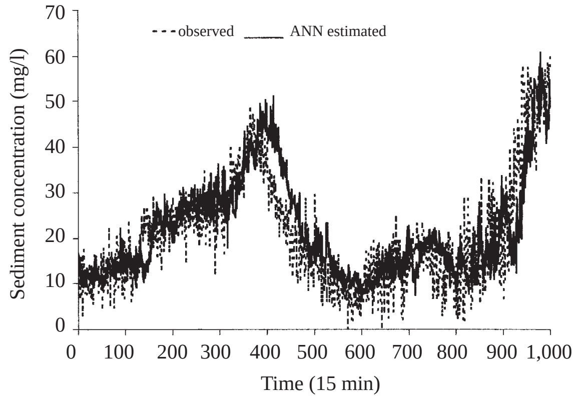


**Figure 9.** The forecasted Ouse Skelton suspended sediment values at time t+24 for time period 07-01-1995 to 18-01-1995 using the last 6 (t-6, t-5, ...,t-1) values as input with a MSE=53.65.

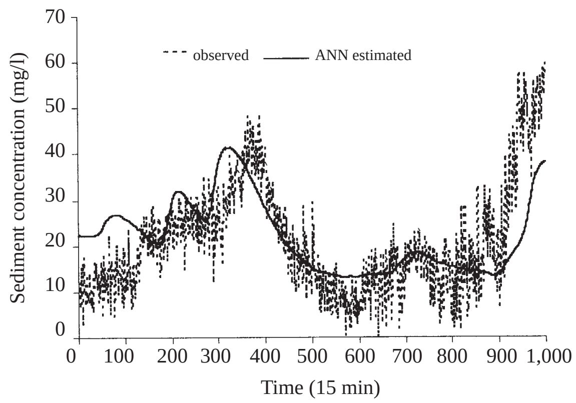


**Figure 10.** The estimated Ouse Skelton suspended sediment values for time period 07-01-1995 to 18-01-1995 using the last 6 (t-6, t-5, ...,t-1) Ouse Skelton flow values as input with a MSE=89.70.

is a significant conclusion since the sediment data may be absent for some time intervals and a rough estimation would really be of great significance.

It should be kept in mind that for the annual sediment budget computations for the related reservoirs' annual sediment totals are considered. Therefore, the obtained results carry significance since the ANN estimated suspended sediment values represent the average behaviour providing a total sum close to the original one.

The next step of the study was the estimation of the suspended sediment concentration in a river using the river flow and suspended sediment concentration data in another river for training the neural networks. This approach can be justified by the fact that the catchments of two rivers show similar physical characteristics. The training period covered 70 days from 09-11-1994 to 18-01-1995 and 55 days from 25-12-1996 to 19-02-1997. Since the input layers were selected randomly, the discontinuity between two periods did not cause a problem. After trying various numbers of nodes in the input layer, it is seen that 6 input nodes produced the minimum MSE during the testing stage. The sediment values in the Swale River were used for training purposes with six consecutive flows at times t-5,.., t in the input layer to forecast the sediment value at time t. The trained neural networks were used to estimate the suspended sediment value at Ouse Skelton at time t using the Ouse River flows at time t-5, . . . , t as input. The testing period covered 70 days from 09-11-1994 to

18-01-1995. The testing stage results for Ouse Skelton suspended sediment values are presented in Fig. 11. The MSE was 298 and the estimated sediment total, 22,319 tonnes, differed from the observed one, 24,276 tonnes, by 8.0% . It can be said that the ANN estimated the time series, underestimated the observed peaks, but again provide an average representation for the observed series. This is expected because as long as the ANNs are not adaptive or dynamic they all yield average trends. The same study has been carried out by adding a Swale sediment concentration value at time t to the input layer and decreasing the number of flow inputs from 6 to 5 in both training and testing stages. Thus, training stage input layer consisted of five Swale flows at times t-4, t-3, . . . , t and a Swale sediment value at time t-1 making a total of 6 inputs in the input layer whereas the testing stage input layer included the Ouse flows at t-4, t-3, . . . , t and a Swale sediment value at time t-1. The resulting ANN-estimated Ouse sediment series are closer to the original one, with a MSE equal to 249, compared with the previous one, where only flows were used in the input layer (Fig. 12). The estimated sediment amount was 21,567 tonnes being 11% less than the observed 24,276 tonnes. The addition of a sediment value to the input layer provided closer estimates, especially for the peaks, whereas the total estimated sediment amount was not significantly different from the previous case. The addition of further Swale sediment values to the input layer did not provide an improvement in estimation.

The same data set is used afterwards to estimate Swale River sediment. In this case during the training stage, which covered again periods from 09-11-1994 to 18-01-1995 and from 25-12-1996 to 19-02-1997, Ouse sediment at time t is estimated using Ouse flows at times t-4, t-3, . . . , t and an Ouse sediment value at time t-1. During the testing stage (from 09-11-1994 to 18-01-1995) Swale sediment is estimated using Swale flows at times t-4, t-3, . . . , t and an Ouse sediment value at time t-1. The MSE was 1,230 and the estimated sediment sum, 6,745 tonnes, differed from the observed 7,000 tonnes by 3.6%.

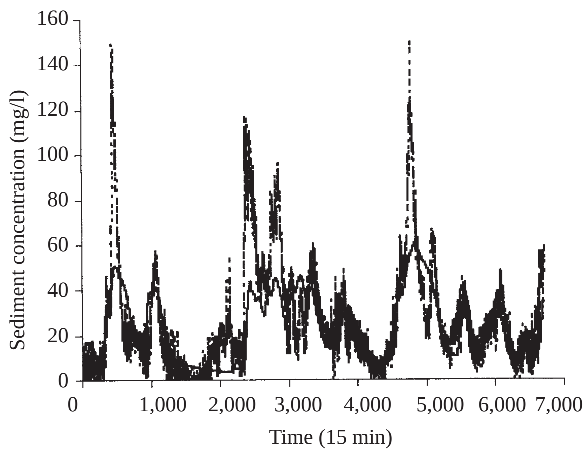


**Figure 11.** The estimated Ouse Skelton suspended sediment values for time period 09-11-1994 to 18-01-1995 using the last 6 (t-6, t-5, . . . , t-1) Ouse Skelton flow values as input with a MSE=298.34 following the training using Swale Thornton sediment as output and Swale flow as input for time periods 09-11-1994 to 18-01-1995 and 25-12-1996 to 19-02-1997 (the input layers are selected randomly).

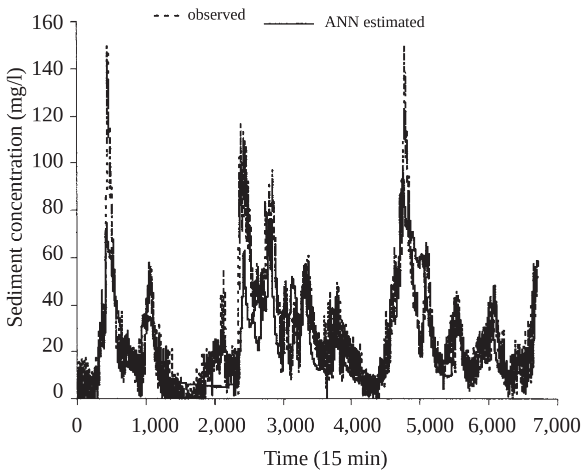


**Figure 12.** The estimated Ouse suspended sediment values for time period 09-11-1994 to 18-01-1995 using the last 5 (t-6, t-5, ...,t-1) Ouse flow values and 1 previous Swale sediment value as input with an MSE=249.00 following the training using Swale sediment as output and last 5 Swale flows and the previous Swale sediment as input for time periods 09-11-1994 to 18-01-1995 and 25-12-1996 to 19-02-1997 (the input layers are selected randomly).

This result is significant since it shows that a rough approximation of the suspended sediment concentration values can be obtained using the flow and suspended sediment data in another river in training ANNs, provided that the flow values for the station without sediment concentration data are available. Since the installation of a sediment measurement equipment is costly and, in general, the annual sediment sum is desired, the ANN method can be considered as a powerful tool in suspended sediment estimation. It should be also taken into account that in this study the rainfall data was not incorporated in the input layer since it was not available. The inclusion of rainfall data would certainly increase the accuracy of the estimation.

The next step of the study was to compare the ANN estimation results with those obtained with regression. The relation between sediment concentration and river flow is represented generally in the literature with the sediment rating curve as mentioned by Clarke (1994),

$$C = aQ^b + \varepsilon \tag{2}$$

where C and Q represent sediment concentration and river flow, respectively, whereas a and b are coefficients and $\varepsilon$ is the error term. This nonlinear relationship can be converted into a regression form taking logarithms of both parts:

$$x = d + fy + \varepsilon \tag{3}$$

where x and y represent the logged values of sediment concentration and river flow, respectively, whereas d and f are coefficients and $\varepsilon^*$ is the error term.

The data period from 25-12-1996 to 19-02-1997 is used to obtain the regression equation with Ouse sediment as a dependent variable and Swale sediment, Swale flow and Ouse flow as independent variables. Considering the non linear relation between variables, a nonlinear regression has been found adequate. The equation is as follows,

$$x = -0.94 - 0.02y_1 + 0.33y_2 + 0.75y_3; R^2 = 0.78 \tag{4}$$

where x, $y_1$, $y_2$ and $y_3$ represent the logged values of Ouse sediment, Swale sediment, Swale flow and Ouse flow, respectively. The obtained regression is applied to the data period 09-11-1994 to 18-01-1995 to estimate Ouse sediment using observed Swale sediment, Swale flow and Ouse flow values. The results are presented in Fig. 14. The MSE was 674 and the regression estimated sediment sum had a value of 39,319 tonnes being 62% higher than the observed 24,276 tonnes. The results show that ANN estimations for sediment concentration are far superior to classical regression estimates in terms of both MSE and total sediment amount. It should be also kept in mind that, in the regression analysis the observed Ouse sediment values were used as dependent variables to obtain the regression equation, whereas during the training stage of the ANNs only Swale sediment and flow were used.

## Discussion and conclusions

In the presented study ANNs have been used for forecasting and estimating suspended sediment concentration. It has been shown that even in the absence of observed sediment data it was possible to obtain reliable corresponding estimates by training ANNs using the sediment and flow data in a nearby river. The difference between the ANN estimated total sediment amount and the corresponding observed amount was either less than or close to 10% . The ANNs are more versatile than the regression-based models because of the freedom available with the choice of the number of hidden layers and the nodes associated with each of these layers. The ANN structure allows information to be processed along multiple paths simultaneously, thereby offering opportunities for parallel implementation. The results obtained are especially significant considering the expense of installing sediment measurement equipment and the importance of providing realistic future estimates for a river's potential sediment yield.

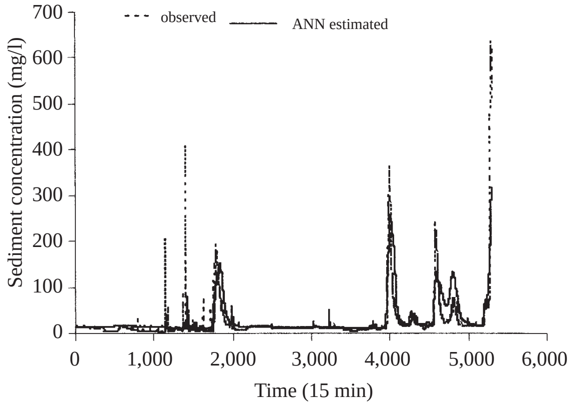


**Figure 13.** The estimated Swale suspended sediment values for time period 25-12-1996 to 19-02-1997 using the last 5 (t-6, t-5, ...,t-1) Swale flow values and 1 previous Ouse sediment value as input with an MSE=1,230 following the training using Ouse sediment as output and last 5 Ouse flows and the previous Ouse sediment as input for time periods 09-11-1994 to 18-01-1995 and 25-12-1996 to 19-02-1997 (the input layers are selected randomly).

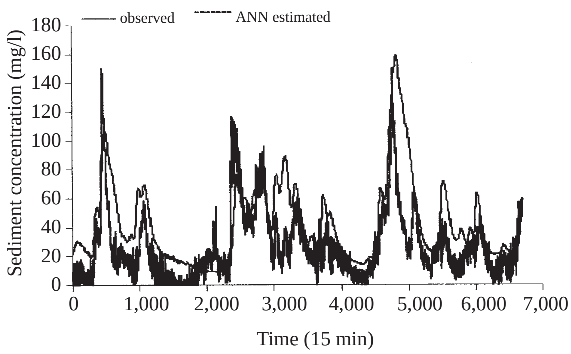


**Figure 14.** The regression estimated Ouse suspended sediment values for time period 09-11-1994 to 18-01-1995 using the Ouse flow, Swale sediment and Swale flow values as independent variables (MSE=674). The regression equation is obtained using the values for Ouse Skelton sediment, Ouse flow, Swale Thornton sediment and Swale flow for time period 25-12-1996 to 19-02-1997.

## Appendix

A neural network has an input layer, a hidden layer and an output layer. Each layer is made up of several nodes, and layers are interconnected by sets of weights. The pattern of connectivity and the number of neurons in each layer may vary within some constraints. No communication is permitted between the nodes within a layer, but the nodes in each layer may send their output to the nodes in the succeeding layers. The nodes receive input either from the initial inputs or from the interconnections.

Error back propagation involves two phases: a feed forward phase, in which the external input information at the input nodes is propagated forward to compute the output information signal at the output unit, and a backward phase, in which modifications to the connection strengths are made based on the differences between the computed and observed information signals at the output units (Eberhart and Dobbins, 1990).

At the beginning of a training process, the connection strengths are assigned random values. The learning algorithm modifies the strength in each iteration until the successful completion of the training. When the iterative process has converged, the collection of connection strengths captures and stores the knowledge and the information present in the examples used in the training process.

When presented with a new input pattern, a feed forward network computation results in an output pattern which is the result of the generalisation and synthesis of what ANN has learned and stored in its connection strengths.

The neural network employed in this study possessed a three-layer learning network consisting of an input layer, a hidden layer and an output layer, as shown in Fig. 1.

There are N data input patterns, each having a set of input values, $x_i$, i=1, . . . ., k at the input nodes with output values, $T_n$, n =1, . . . .., m at the output nodes. The input values are multiplied by the first interconnection weights, $w_{ij}$, j=1, . . . ,h at the hidden nodes, and the products are summed over the index, i, and become the inputs of hidden layers i.e.,

$$H_j = \sum_{i=1}^{k} w_{ij} x_i \quad j = 1, \ldots h$$

where $H_j$ is the input to the jth hidden node, $w_{ij}$ is the connection weight from the ith neuron to the jth neuron. Each hidden node is transformed through a sigmoid function to produce a hidden node output, $HO_j$, defined as

$$HO_j = f(H_j) = \frac{1}{1 + \exp\left[-\ (H_j + \theta_j)\right]}$$

where $H_j$ is the input to the node, $f(H_j)$ is the node output, and $\theta_j$ is a threshold or bias. The threshold, $\theta_j$, will be learned in the same manner as the weights. The output, $HO_j$ serves as the input to the succeeding layer and its procedure is continued until the output layer is reached. This is referred to as forward activation flow. The input to the m output nodes, $IO_n$, is expressed as

$$IO_n = \sum_{j=1}^{h} w_{jn} HO_{jn} \quad n = 1, \ldots m$$

These input values are processed through the sigmoidal function defined earlier to give the neural network output values, $O_n$. The subsequent weight adoption or learning process is accomplished by the back propagation learning algorithm.

The $O_n$ at the output layer will not be the same as the target value, $T_n$. For each input pattern, the sum of the squares of error, $e_p$, for the pth input pattern is

$$e_p = \sum_{n=1}^{m} (T_n - O_n)^2$$

and the average system error or mean square error (MSE), E, for all input patterns is

$$E = \frac{1}{2N} \sum_{p=1}^{N} \sum_{n=1}^{m} (T_{pn} - O_{pn})^2$$

where $T_{pn}$is the target value, $T_n$, for the pth pattern and $O_{pn}$ is the neural network output value, $O_n$, for the $p^{th}$ pattern.

The aim of the back propagation algorithm is to minimise iteratively the averaged square error. This is accomplished by first computing the gradient ($\delta_n$) for each node on the output layer:

$$\delta_n = O_n(1 - O_n)(T_n - O_n)$$

The error gradient $\delta_j$ is then recursively determined for the hidden layers by computing the weighted sum of the errors at the previous layer:

$$\delta_j = HO_j(1 - HO_j) \sum_{n=1}^{m} \delta_n w_{jn}$$

The error gradients are then used to update the network weights:

$$\Delta w_{ij}(r) = \eta \delta_j x_i$$

$$w_{ij}(r+1) = w_{ji}(r) + \Delta w_{ji}(r)$$

Generally, to assure rapid convergence, large step sizes, which do not lead to oscillations, are used. The weight change after the nth data presentation is:

$$\Delta w_{ji}(r) = \eta \delta_j x_i + \alpha \Delta w_{ji}(r-1)$$

where $\alpha$ the a momentum rate term is used to improve convergence, $\eta$ is the learning rate which provides the step size during the gradient descent and r is the iteration number.

## References


Abrahart, R. J., “Neural Networks and the Problem of Accumulated Error: An Embedded Solution that Offers New Opportunities for Modelling and Testing”, Hydroinformatics Conference, Copenhagen, 1998.

Boogaard, H., Gautam, D.K., and Mynett, A.E., “Auto-Regressive Neural Networks for the Modelling of Time Series”, Hydroinformatics Conference, Copenhagen, 1998.

Box, G.E., and Jenkins, G.M., “Time Series Analysis: Forecasting and Control”, Holden-Day, San Francisco, 1976.

Campolo, M., Andreussi, P., and Soldati, A., “River Flood Forecasting with a Neural Network Model”, Water Resources Research, 35, 1191-1197, 1999a.

Campolo, M., Soldati, A., and Andreussi, P., “Forecasting River Flow Rate during Low-Flow Periods using Neural Networks”, Water Resources Research, 35, 11, 3547-3552, 1999b.

Clarke, R.T., “Statistical Modelling in Hydrology”, John Wiley & Sons, 412p, 1994.

Djebbar, Y., and Alila, Y., “Neural Network Estimation of Sanitary Flows”, Hydroinformatics Conference, Poster Presentations, Copenhagen, 1998.

Eberhart, R.C., and Dobbins, R.W., “Neural Network PC Tools: A Practical Guide”, 1990.

Fernando, D., and Jayawardena, A.W., “Runoff Forecasting using RBF Networks with OLS Algorithm”, Journal of Hydrologic Engineering, 3, 3, 203-209, 1998.

Hall, M. J., and Minns, A., “Regional Flood Frequency Analysis using Artificial Neural Networks”, Hydroinformatics Conference, Copenhagen, 1998.

Hsu, K., Gupta, H., and Sorooshian, S., “Artificial Neural Network Modelling of the Rainfall-Runoff Process”, Water Resources Research, 31, 10, 2517-2530, 1995.

Kheir El-Din , K.A., “Neural Network Application for Modelling Hydraulic Characteristics of Severe Contraction”, Hydroinformatics Conference, Copenhagen, 1998.

Lange,N., “Advantages of Unit Hydrograph Derivation by Neural Networks”, Hydroinformatics Conference, Copenhagen, 1998.

Leeks, G.J.L., Neal, C., Jarvie, H.P., Casey, H., and Leach, D.V., “The LOIS River Monitoring Network: Strategy and Implementation”, The Science of the Total Environment, 194/195, 101-109, 1997.

Lingireddy, S., “Aquifer-Parameter Estimation using Genetic Algorithms and Neural Networks”, Civil Engineering and Environmental Systems, 15, 2, 125-144, 1988.

Mason, J.C., Price, R.K., and Tem‘me, A., “A Neural Network Model of Rainfall- Runoff using Radial Basis Functions”, Journal of Hydraulic Research, 34, 4, 537- 548, 1996.

Minns, A.W., and Hall, M.J., “Artificial Neural Networks as Rainfall Runoff Models”, Hydrological Sciences Journal, 41, 3, 399-417, 1996.

Raman, H., and Sunilkumar, N., “Multivariate Modelling of Water Resources Time Series using Artificial Neural Networks”, Hydrological Sciences Journal, 40, 2, 145-163, 1995.

Ranjithan, S., Eheart, J.W., and Garrett, J.H., “Neural Network-Based Screening for Groundwater Reclamation under Uncertainity”, Water Resources Research, 29, 3, 563-574, 1993.

Rogers, L., and Dowla, F.U., “Optimization of Groundwater Remediation using Artificial Neural Networks with Parallel Solute Transport Modelling”, Water Resources Research, Vol.30, No.2, 457-481, 1994.

See, L., and Openshaw, S., “Using Soft Computing Techniques to Enhance Flood Forecasting on the River Ouse”, Hydroinformatics Conference, Copenhagen, 1998.

Singh, V.P., Krstanovic, P.F., and Lane, L.J., “Stochastic Models of Sediment Yield”, Modeling Geomorphological Systems, Chapter 9, Edited by M.G.Anderson, John Wiley & Sons Ltd., 1988.

Tokar, A.S., and Johnson, P.A., “Rainfall-Runoff Modelling using Artificial Neural Networks”, Journal of Hydrologic Engineering, 4, 3, 232-239, 1999.

Wass, P.D., Marks, S.D., Finch, J.W., Leeks, G.J.L., and Ingram, J.K., “Monitoring and Preliminary Interpretation of in-River Turbidity and Remote Sensed Imagery for Suspended Sediment Transport Studies in Humber catchment”, The Science of the Total Environment, 194/195, 263-283, 1997.

Wass, P.D., and Leeks, G.J.L., “Suspended Sediment Fluxes in the Humber catchment”, UK, Hydrological Processes, 13, 935-953, 1999.