ARTIFICIAL NEURAL NETWORK MODELS FOR ESTIMATION OF SEDIMENT LOAD IN AN ALLUVIAL RIVER IN INDIA

Archana Sarkar  
Rakesh Kumar  
Sanjay K. Jain  
R.D. Singh  
National Institute of Hydrology  
Roorkee, India.

The magnitude of sediment transport by rivers is a major concern for water resources planning and management. The methods available for sediment estimation are largely empirical, with sediment rating curves being the most widely used in India. In this study, sediment rating curve and artificial neural network (ANN) techniques have been applied to model the sediment-discharge relationship of an alluvial river. Daily data of sediment load and discharge of the Kosi River in India have been used. A comparison has been made between the results obtained using ANNs and sediment rating curves. The sediment load estimations in the river obtained by ANNs have been found to be significantly superior to the corresponding classical sediment rating curve ones. Also, an ANN approach can give information about the structure of events (e.g., hysteresis in the sediment-discharge relationship) which is not possible to achieve with sediment rating curves.
INTRODUCTION

The sediment outflow from the watershed is induced by processes of detachment, transportation and deposition of soil materials by rainfall and runoff. The assessment of the volume of sediments being transported by a river is 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; determination of the effects of watershed management; and environmental impact assessment. Fine sediment has long been identified as an important factor 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 (Morris and Fan, 1997).

Sediment rating curves are widely used to estimate the sediment load being transported by a river. Sediment load is defined as the sediment flow in a river measurable at a point of reference during a specified period of time. A sediment rating curve is a relation between the sediment and river discharges. Such a relationship is usually established by a regression analysis, and the curves are generally expressed in the form of a power equation. Rating curves are developed on the premise that a stable relationship between concentration and discharge can be developed which, although exhibiting scatter, will allow the mean sediment yield to be determined on the basis of the discharge history. A problem inherent in the rating curve technique is the high degree of scatter, which may be reduced but not eliminated. Concentration does not necessarily increase as a function of discharge.

Mathematically, a rating curve may be constructed by log-transforming all data and using a linear least squares regression to determine the line of best fit. The log-log relationship between load and discharge is of the form (Morris and Fan, 1997)

\[ S = aQ^b \]  

(1)

and the log-transformed form will plot as a straight line on log-log paper

\[ \log S = \log a + b \log (Q) \]  

(2)

where, \( S \) = sediment load, \( Q \) = discharge, \( \log a \) and \( b \) are regression constants.

A regression equation will minimize the sum of squared deviations from the log-transformed data, which is not the same as minimizing the sum of squared deviations from the original dataset and which introduces a bias that underestimates the load at any discharge. Ferguson (1986) reports that this bias may result in underestimation by as much as 50 percent. Ferguson and others have suggested bias correction factors, but their appropriateness is uncertain (Walling and Webb, 1988). The rating curve technique is not adequate in view of the complexity of the problem.

On the other hand, the application of physics-based distributed process computer simulation offers another possible method of sediment prediction. But the application of these complex software programs is often problematic, due to the use of idealized sedimentation components, or the need for massive amounts of detailed spatial and temporal data which are not available. Simpler approaches are therefore required in the form of ‘conceptual’ solutions or ‘black-box’ modelling techniques. Neurocomputing provides one possible answer to the problematic task of sediment transfer prediction. Recently, artificial neural networks (ANN) have emerged as powerful tools to model nonlinear processes. The artificial neural networks operate in a manner
Artificial Neural Network Estimation of Sediment Load  Sarkar, Kumar, Jain, and Singh

An ANN is a computing system made up of a highly interconnected set of simple information processing elements, analogous to a neuron, called units. The neuron collects inputs from both a single and multiple sources and produces output in accordance with a predetermined nonlinear function. An ANN model is created by interconnection of many of the neurons in a known configuration. The primary elements characterizing the neural network are the distributed representation of information, local operations and nonlinear processing. Figure 1 shows the general structure of a three layer back propagation ANN. The theory of ANN has been described in many textbooks such as Haykin (1994).

The main principle of neural computing is the decomposition of the input-output relationship into series of linearly separable steps using hidden layers (Haykin, 1994). Generally there are four distinct steps in developing an ANN-based solution. The first step is the data transformation or scaling. The second step is the network architecture definition, where the number of hidden layers, the number of neurons in each layer, and the connectivity between the neurons are set. In the third step, a learning algorithm is used to train the network to respond correctly to a given set of inputs. Lastly, comes the validation step in which the performance of the trained ANN model is tested through some selected statistical criteria.

There are numerous studies related to the application of ANNs to various problems frequently encountered in water resources (ASCE Task Committee, 2000). But the application of the ANN approach for modelling sediment-discharge process is very recent, and has already produced very encouraging results. In a research project by Rosenbaum (2000), ANN technique has been used to predict sediment distribution in Swedish harbors. Jain (2001) used the ANN approach to establish an integrated stage-discharge-sediment concentration relation for two sites on the Mississippi River. Based on the comparison of results for two gauging sites, he has shown that the ANN results
are much closer to the observed values than the conventional technique. In a study by Nagy et al. (2002), an ANN is used to estimate the natural sediment discharge in rivers in terms of sediment concentration. They have addressed the importance of choosing an appropriate neural network structure and providing field data to that network for the training purpose and found that the ANN approach gives better results compared to several commonly used formulas of sediment discharge. Yitian and Gu (2003) applied the ANN technique to modelling daily discharge and annual sediment discharges in the Jingjiang reach of the Yangtze River and Dongting Lake, China. The authors demonstrated that the ANN technique is a powerful tool for real-time prediction of flow and sediment transport in a complex network of rivers. Agarwal et al. (2004) developed daily, weekly, ten-daily and monthly sediment yield ANN models for the Vamsadhara River basin in India. Raghuvanshi et al. (2006) also developed ANN models to predict runoff and sediment yield on a daily and weekly basis for a small agricultural watershed in the upper Siwane River in India. The authors found the performance of ANN models superior to regression models. Cigizoglu and Alp (2007) used the generalized regression neural network technique for river suspended sediment estimation in the Juniata Catchment in the USA. The authors found ANN estimations significantly superior to conventional method results.

In the present study, the ANN technique along with the conventional sediment rating curve technique has been applied to model the sediment-discharge relationship of the Kosi River in India using daily data of sediment load and discharge at the Birpur gauging site.

**PERFORMANCE EVALUATION CRITERIA**

The statistical and hydrological evaluation criteria used in the present study are root mean square error (RMSE), correlation coefficient (r), coefficient of efficiency (CE) or coefficient of determination (r²) and volumetric error (EV).

**Root Mean Square Error (RMSE)**

It yields the residual error in terms of the mean square error expressed as (Yu, 1994)

\[
RMSE = \sqrt{\frac{\text{residual variance}}{n}} = \left( \frac{\sum_{j=1}^{n} (Y_j - \hat{Y}_j)^2 / n}{n} \right)^{1/2}
\]

where, \( Y \) and \( \hat{Y} \) are the observed and estimated values respectively and \( n \) is the number of observations.

**Correlation Coefficient (r)**

It is expressed as (Haan, 1977)

\[
r = \frac{\sum_{j=1}^{n} (Y_j - \bar{Y})(\hat{Y}_j - \bar{Y})}{\sqrt{\sum_{j=1}^{n} (Y_j - \bar{Y})^2 \sum_{j=1}^{n} (\hat{Y}_j - \bar{Y})^2}} \times 100
\]

where, \( \bar{Y} \) and \( \bar{Y} \) are the means of observed and estimated values.
**Coefficient of Efficiency (CE)**

Based on the standardization of residual variance with initial variance, the coefficient of efficiency can be used to effectively compare the relative performance of the two approaches (Nash and Sutcliffe, 1970). It is expressed as

\[
CE = \left\{ 1- \frac{\text{residual variance}}{\text{initial variance}} \right\} \times 100 = \left\{ 1- \frac{\sum_{j=1}^{n} (Y_j - \hat{Y}_j)^2}{\sum_{j=1}^{n} (Y_j - \bar{Y})^2} \right\} \times 100 \tag{5}
\]

The coefficient of efficiency is also commonly known as the coefficient of determination \(r^2\) which may be written in a number of ways and represents the fraction of variance that is explained by regression. The closer this ratio is to unity, the better is the regression relation (Haan, 1977).

**Volumetric Error (EV)**

This is also called relative absolute error in volume (Yu, 1994) and is estimated as

\[
EV = \left\{ \frac{\sum_{j=1}^{n} (\hat{Y}_j - Y_j)}{\sum_{j=1}^{n} (Y_j)} \right\} \times 100 \tag{6}
\]

**THE STUDY AREA AND DATA AVAILABILITY**

In the present study, time series data of sediment load for one gauging station on the Kosi River in the state of Bihar in India has been used. Kosi is a large alluvial river with low gradients and wide flood plains. Meandering or lateral shifting of alluvial rivers produces cutoff meanders, oxbow lakes and distinctive landforms. Tectonic and environmental changes can cause aggradations and degradation in alluvial rivers and lead to high soil erosion and meandering. The Kosi River carries a mean annual discharge of 1.6x10³ m³/sec, with monsoon discharge 10 times the lean period discharge. The river carries a very high concentration of suspended sediment load during the monsoon months. The normal flood discharge of the Kosi usually varies from 1.5 to 2.0 million cusecs. About 75 to 84 percent of the total runoff occurs in the monsoon months of June to October. On an average total sediments are 0.20 percent of the total runoff. About 95 percent of the silt load comes down the river during the monsoon floods and only 5 percent of the sediments come down in the remaining non-monsoon months. The total runoff during the non-monsoon months, however is on an average about 19 percent of the total annual runoff.

For the present study, discharge and sediment load data for a period of five years (2000-2004) have been used. However, during this period some of the data are missing and therefore data of 1502 days have been used. These data are available at the Birpur gauging site. There is a barrage at Indo-Nepal border near Birpur. The unit of discharge data is cusec while sediment load data is in cubic feet (cft) per day.

**METHODOLOGY**

**Rating Curve Analysis**

In this study, a regression analysis has been carried out to develop the sediment rating equation for the Kosi river at the Birpur gauging site. For this purpose, the available data have been
considered in two parts. The first part is used to calibrate the mathematical equation for sediment rating and the second, to validate it. Out of 1502 days of data, 1002 days data were used for the calibration of the rating curve and the remaining 500 days data for validation.

The sediment rating equation between sediment load and discharge for Kosi River at Birpur site is developed in the form of Equation (1)

\[ S = 0.3368Q^{2.2373} \]  

where

\[ S = \text{Sediment load in the River Kosi at the Birpur site in } 10^3\text{cft.} \]
\[ Q = \text{Discharge in the River Kosi at the Birpur site in } 10^3\text{cusec.} \]

Artificial Neural Network Analysis

For the development of ANN models the total data have been divided into training and testing periods of 1002 days and 500 days respectively. The models provide the sediment load at time step \( t \), with \( S_t \) as output. It has been shown by many investigators that the current sediment load can be better mapped by considering, in addition to the current value of discharge, the sediment and discharge at previous times. Therefore, in addition to \( Q_t \), i.e., discharge at time step \( t \), other variables such as \( Q_{t-1}, Q_{t-2}, \ldots \) etc. and \( S_{t-1}, S_{t-2}, \ldots \) etc were considered as the input. Various combinations of input variables considered for training of ANN in the present study are given in Table 1 together with the output variables.

From the many different types of ANNs that have been developed with different objectives, the multilayer perceptron was chosen for application in this research, as it is particularly suited to regression problems and is the most common type of network applied to modelling of various hydrological problems (Hsu et al., 1995). These ANNs model complex multivariate nonlinear functions, such as the sigmoid function. The composite function is fitted to the data by modifying the shape-defining parameters of the component nonlinear functions in an iterative training process, which minimizes the error between the estimated outputs and the target outputs.

In the course of this investigation, neural network analysis was conducted using the Neural Power-2.5 ANN package (NeuralPower, 2003). A back-propagation ANN with the generalized delta rule as the training algorithm has been employed in this study. The structure for all simulation models are three layer BPANN which utilizes a nonlinear sigmoid activation function uniformly between the layers. Nodes in the input layer are equal to number of input variables, nodes in hidden layer are varied from the default value by the NP package (2 to 7) for various ANN models to approximately double of input nodes (Zhu et al., 1994) and the nodes in the output layer is one as the models provide single output.

The modelling of ANN started with the normalization (re-scaling) of all inputs and output with the maximum value of respective variable reducing the data in the range 0 to 1 to avoid any

<table>
<thead>
<tr>
<th>ANN Model</th>
<th>Input Variables</th>
<th>Output Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN - 1</td>
<td>( Q_t )</td>
<td>( S_t )</td>
</tr>
<tr>
<td>ANN - 2</td>
<td>( Q_t, Q_{t-1}, S_{t-1} )</td>
<td>( S_t )</td>
</tr>
<tr>
<td>ANN - 3</td>
<td>( Q_t, Q_{t-1}, Q_{t-2}, S_{t-1}, S_{t-2} )</td>
<td>( S_t )</td>
</tr>
<tr>
<td>ANN - 4</td>
<td>( Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, S_{t-1}, S_{t-2}, S_{t-3} )</td>
<td>( S_t )</td>
</tr>
<tr>
<td>ANN - 5</td>
<td>( Q_t, Q_{t-1}, Q_{t-2}, Q_{t-3}, Q_{t-4}, S_{t-1}, S_{t-2}, S_{t-3}, S_{t-4} )</td>
<td>( S_t )</td>
</tr>
</tbody>
</table>
Artificial Neural Network Estimation of Sediment Load  Sarkar, Kumar, Jain, and Singh

saturation effect that may be caused by the use of sigmoid function (accomplished through the Neural Power package). All interconnecting links between nodes of successive layers were assigned random values called weights. A constant value of 0.15 and 0.8 respectively has been considered for the learning rate $\mu$ and the momentum term $\alpha$ which were selected after hit and miss trials. The quick propagation (QP) learning algorithm has been adopted for the training of all the ANN models. QP is a heuristic modification of the standard back propagation and is very fast (NeuralPower, 2003). The network weights were updated after presenting each pattern from the learning data set, rather than once per iteration.

The criterion selected to avoid over training was generalization of ANN through cross-validation (Haykin, 1994). For this purpose, the training data set for all the models was further divided in two subsets. First subset (800 datasets) for estimation of weights of the ANN model and second subset (202 dataset) for evaluation of the performance of ANN model with data of first subset. Training was stopped when the error for the second subset started increasing. In this way, the training set has been used to assess the performance of various candidate model structures, and thereby choose the best one. The particular ANN model with the best performing parameter values was then trained on the full training data set (1502 daily data), and the generalized performance of the resulting network has been measured on the test data set (500 daily data) to which it has never before been exposed. The performance of the model was tested through the statistical criterion discussed earlier.

**DISCUSSION OF RESULTS**

The values of the performance criteria from various ANN models as well a rating curve for both training (calibration) and testing (validation) data sets are presented in Table 2. The training and testing results are discussed separately.

**Calibration (Training) Results**

It can be seen from Table 2 that all the ANN models have outperformed the conventional sediment rating curve technique in terms of various performance criteria. In terms of root mean

**Table 2. Comparative performance of various ANN models and conventional sediment rating curve.**

<table>
<thead>
<tr>
<th>Model</th>
<th>Calibration (Training)</th>
<th>Validation (Testing)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RMSE</td>
<td>$r$</td>
</tr>
<tr>
<td>ANN Model (Input, Hidden, Output)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN – 1</td>
<td>4745.0</td>
<td>94.5</td>
</tr>
<tr>
<td>(1,2,1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN – 2</td>
<td>3666.9</td>
<td>96.7</td>
</tr>
<tr>
<td>(3,3,1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN – 3</td>
<td>3001.6</td>
<td>97.8</td>
</tr>
<tr>
<td>(5,4,1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN – 4</td>
<td>2802.1</td>
<td>98.1</td>
</tr>
<tr>
<td>(7,6,1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ANN – 5</td>
<td>3408.1</td>
<td>97.1</td>
</tr>
<tr>
<td>(9,7,1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conventional Procedure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sediment Rating Curve</td>
<td>8038.7</td>
<td>92.6</td>
</tr>
</tbody>
</table>
square error (RMSE), rating curve model performed the worst (8038.7), whereas, ANN-4 model performed the best (2802.1). The correlation coefficient (r) and coefficient of determination ($R^2$) for various ANN models have been higher than the conventional sediment rating curve technique. However, both the values have been highest for ANN-4 model ($r = 98.1\%, \; \text{and} \; r^2 = 96.2\%$), whereas the values for rating curve are only 92.6\% and 85.7\% respectively. In terms of volumetric error (EV), the performance of ANN-4 model has been the best with the least error (19.24), whereas the rating curve model performed the worst with highest value of error 40.68) which is almost double the ANN error.

**Validation (Testing) Results**

All the ANN models have outperformed the conventional sediment rating curve technique in terms of various performance criteria during testing/validation also. In terms of RMSE, again the rating curve model performed the worst (5252.2), whereas, ANN-4 model performed the best (2965.6). In terms of r and $r^2$, ANN-4 model performed the best (97.6\% and 95.2\% respectively). The rating curve model performed the worst for r and $R^2$ also (95.7\% and 91.6\% respectively). The volumetric error (EV) has been lowest for ANN-2 model (21.81), whereas, it has been highest for the rating curve model (39.14). The best performing ANN model in other criteria, i.e., ANN-4 model has also got a low value of EV (23.13).

Therefore, it can be observed from the results of calibration as well as validation that the performance of ANN-4 model has been the best except in terms of volumetric error during validation. However, in terms of EV also, its performance has been very close to the best one. It can be noted that by the inclusion of input variables of previous time steps, the model performance improves up to the previous three days inputs. Beyond that, i.e., when the inputs of previous fourth day are included in the model, the performance starts deteriorating. It implies that beyond the previous three days values, no new information is actually supplied to the ANN model for training. And with higher number of input variables, the network becomes more complicated and may over-fit the data.

The plots between observed sediment load and estimated sediment load (using the conventional and ANN approach) for the calibration as well as validation period have been illustrated in Figure 2 (a), (b), (c) and (d) respectively. It is observed that the ANN (ANN-4) estimates show a better match with the observed values. It is also seen from the graph that there is a large variation in the conventional approach estimates (for both calibration and validation period). However, the ANN estimates (for both calibration & validation period) show a fairly good agreement with the observed values even at the high extremes.

The temporal variation of observed sediment load and the estimate using the conventional technique and ANN (ANN–4 model) for the calibration period is plotted in Figure 3. It is seen from the comparison of the graphs that the ANN estimates of sediment load very closely follow the observed curve, whereas the conventional approach has significant mismatch, particularly near the peaks.

Using the weights obtained in the training phase for each combination, the performance or in other words, the generalization capability of the ANN was tested using the validation period data. Figure 4 contains the graphical presentation of the results of the validation phase. Comparison of the sediment load estimates with the observed ones using the conventional approach and ANN approach are shown in the figure. Again here the ANN estimates are closer to the corresponding observed ones. The conventional approach estimates again show significant deviations from the corresponding observed ones.
that the hysteresis estimated by ANNs is almost the same as the observed one. It can be seen the hysteresis in the sediment load obtained by ANNs and the sediment rating relation are compared with the corresponding observed ones in Figure 5 for all the 1502 data sets. It can be seen that the hysteresis estimated by ANNs is almost the same as the observed one.

Figure 2. Plot between observed and estimated sediment load.

Figure 3. Comparison of observed and estimated sediment load during calibration.

To be able to test whether the ANN approach can give information about the structure of events, the hysteresis in the sediment load obtained by ANNs and the sediment rating relation are compared with the corresponding observed ones in Figure 5 for all the 1502 data sets. It can be seen that the hysteresis estimated by ANNs is almost the same as the observed one.
Based on these results, it is clear that the sediment load estimations obtained by the ANN technique are significantly superior to the corresponding classical sediment rating curve ones. It may also be added that as a result of training, a set of weights that represents the ‘knowledge’ of ANN is obtained and one does not get an explicit equation to work with. The weight distributions for the trained ANN-4 model are shown in Figure 6.

**CONCLUSIONS**

The ANN technique has been utilized for modelling the sediment-discharge process in an alluvial river. The data of the Birpur gauging site of the River Kosi in India have been used for the analysis. The results of the ANN have been compared with those of the conventional sediment rating curve approach. The ANN results have been found to be much closer to the observed values than the conventional technique. The study shows that the ANN technique can be successfully applied for the development of reliable relationships between sediment and discharge in a river.
when other approaches cannot succeed due to the high non-linearity in the relationship, especially in alluvial rivers. A significant advantage of using the ANN approach is that it can successfully model the hysteresis effect in the sediment-discharge relationship.

Moreover, the ANN technique has preference over the conventional methods as ANNs can accept any number of effective variables as input parameters without omission or simplification as is commonly done in the conventional methods. The presented ANN models have been developed by using only field river data, and they have no boundary conditions in application. The only restriction is that the models cannot estimate accurately the sediment load for data out of the range of the training pattern data. Such a problem can easily be overcome by feeding the training patterns with a wide range of data. Site engineers can calculate sediment load using the ANN without prior knowledge of the sediment transport theories, provided they know the bounds of the data used to generate the ANN.

ACKNOWLEDGMENTS

The authors are thankful to the two reviewers, Dr. Nayan Sharma, Professor, Dept of Water Resources Development and Management, Indian Institute of Technology, Roorkee, India and Dr. S. P. Agarwal, Scientist E, Water Resources Division, Indian Institute of Remote Sensing, Dehradun, India. Their valuable suggestions helped to improve the quality of paper.

REFERENCES


ADDRESS FOR CORRESPONDENCE
Archana Sarkar
National Institute of Hydrology
Roorkee-247667
India

Email: archana_nih@yahoo.co.in