ABSTRACT

Building a numerical model in not an easy task. In addition to its need for verification, it requires special skills and experiences in mathematics, computer programming and may take long period to be ready for operation and prediction. On the other hand, artificial neural network (ANN) prediction models are more efficient in predictions once they are trained from examples or patterns. These types of ANN models need large amount of data which should be at hand before thinking to develop such models. In this paper, the capability of ANN model to predict suspended sediment in 2-D flow field is investigated. The data used for training the network are generated from a pre-verified 2-D hydrodynamic and a 2-D suspended sediment models which were recently developed by the authors. About two-thirds of the data are used for training the network while the rest of the data are used for validating and testing the developed ANN model. Field data measured by hydraulic research Institute are used to compare the results of the ANN model. A tanh activation function is used at the hidden layer which consisted of 5-10-1. The conjugate gradient learning algorithm is adopted. The results of the developed ANN model proved that the technique is reliable in such field compared to both the results of the previously developed models and the field data provided that the trained network is used to generate prediction within the range of training data.

Keywords: Sediment transport, Suspension sediment, Artificial neural networks, numerical modeling, River hydraulics, Nile River, Hydrodynamic modeling

1. INTRODUCTION

The subject of sediment transport in alluvial streams gains its importance with the increasing of water resources utilization. Extensive researches had been done in the sediment transportation and many problems related to the alluvial streams in the past four decades had been investigated. The total load of the transporting sediment is the summation of the bed load and the suspended load. The knowledge of the distribution of sediment in the vertical is eventually used to compute the total suspended load carried by the stream. Observations in laboratory flumes and natural streams have shown that the concentration of suspended load in the vertical direction decreases with increase in distance from the bed. However, laboratory investigations to predict sediment transport are generally very time-consuming, costly and even not possible for many practical engineering problems. Therefore, mathematical models for predicting sediment transport were developed using different techniques. Several one dimensional models were developed, see for example Thomas and Prasuhn\[1\], Bhallamudi and Chaudhry\[2\] and Guo and Jin\[3\]. Examples of 2-D models include those developed by Lin and Shen\[4\], Van Rijn\[5\], Celik and Rodi\[6\], Van Rijn et al.\[7\] and Elfiky et al.\[8\].

Dawdy and Vanoni\[9\] evaluated several developed numerical models and pointed out that "none of the movable-bed models evaluated was found to yield wholly satisfactory results". Instead of mathematical model, a relatively new computational tool, ANN, is used in the present paper to predict the suspended sediment load in terms of the flow depth, the velocities components in x and y directions and the sediment carrying capacity. Since, the method learns from examples, a large set of data should be available. Practically, field data of different rivers should be used to train and validate the ANN but it is not available at the time being at the author hands Therefore, the developed SED-2 numerical model by Elfiky et al.\[8\] is used to generate the required data for training and verification of the ANN to test and prove its capability to predict the suspended sediment concentration once it gets trained..
2. OVERVIEW OF ANN

Artificial Neural Network (ANN), is a structure composed of a number of interconnected units (artificial neurons) Each neuron has an input/output characteristics and implements a local computations or function, Schalkoff\(^{10}\). Hence, the overall ANN of interconnected neurons displays a corresponding functionality. Usually, the application determines the required functionality via specifications. The designer of the ANN has to specify the parameters that satisfy these specifications. The output of any neuron is determined by its input/output characteristics, its interconnection to other neurons and possible external inputs. A neural system should be capable of storing information through training. Thus the objective of training the ANN is to develop an internal structure enabling the ANN to correctly identify or classify new similar patterns. Thus, neural network is a dynamic system, its state changes over time in response to external inputs or an initial unstable state. Various types of ANN are in use and could be reviewed from Schalkoff\(^{10}\). Most of the applications of ANNs in fields of water Engineering were reviewed in Negm\(^{11}\). In this paper, the multilayer feedforward network or the multilayer perceptrons is used in modeling suspended sediment concentration in river flow.

A typical ANN consists of three layers (4-10-1) is shown in Figure 1. Each layer is in turn consists of a number of neurons and links connecting them. The input variables determine the number of neurons in the input layer and the input data vectors are applied to the input layers from an external source. A bias neuron is normally used with input of unity to shift or scale the activation function. The output layer is where the output are processed and are sent to an external source for further analysis or extra treatments or plotting, etc. The layers between the input and the output are hidden where the entire processing are not accessible. The number of neurons in the hidden layer is determined by solving the application several times using networks of different sizes i.e. by a trial and error procedure until the error is minimized. The most common nonlinear transfer functions are the sigmoidal functions including the logistic and the hyperbolic tangent. The latter function is given by Equation (2)

\[
O_{h_j} = \frac{\exp(I_{h_j}) - \exp(-I_{h_j})}{\exp(I_{h_j}) + \exp(-I_{h_j})}
\]

where \(I_{h_j}\) is the input to the neuron \(j\) of the hidden layer and given by

\[
I_{h_j} = \sum_{i=1}^{m} x_i w_{ij} + b_j
\]

where \(x_i\) is the input of the neuron \(i\) in the input layer with \(m\) is the number of neurons in the input layer and \(b_j\) is the bias of the unit. The \(w_{ij}\) is the weights vector of the connections between the neurons of the input layer and the neurons of the hidden layer.

The outgoing signal from the hidden neuron is then combined with the weights of the connections between the neurons of the hidden layers and those of the output layer yielding the output of the output layer, \(O_{o_k}\) using a linear combine function defined by Eq.(3).

\[
O_{o_k} = \sum_{j=1}^{n} O_{h_j} C_{jk} + b_k
\]

in which \(C_{jk}\) is weight of the connection between neuron \(j\) of the hidden layer and neuron \(k\) of the output layer and \(b_k\) is the bias to the neuron \(k\).
Training the network involves the determination of the weight vectors \( w_{ij} \) and \( c_{jk} \) such that the sum of squares of the error between the actual value of the output and the desired value of the output is minimal. The network weights are randomly assumed within a particular range. Then they are updated through training of the network in the direction of minimizing the errors. The range within which weights are assumed should be selected carefully by trial and error. The training algorithm is explained in Negm et al.\(^{12}\) while the systematic steps for building the network was presented in Negm et al.\(^{13}\).

### 3. APPLICATIONS OF ANN IN THE FIELD OF SEDIMENT TRANSPORT

Although many applications in the field of Hydraulic Engineering are available such as Karunanith et al.\(^{14}\), Grubert\(^{15}\), Tawfik et al.\(^{16}\), Thirumalaiah and Doe\(^{17}\), Dibike et al.\(^{18}\), Dibike and Abbott\(^{19}\), Abdeen\(^{20}\), Kashefipour et al.\(^{21}\) and Wright et al.\(^{22}\), very few applications in the field of sediment transport were published. Nagy\(^{23}\) used ANN to estimate the natural sediment discharge in rivers in terms of sediment concentration. This was achieved by training the network to extrapolate data collected from reliable sources. Selecting an appropriate neural networks structure and a training algorithm as well as providing data to the network were addressed using the back-propagation algorithm. It was concluded that the predicted sediment concentrations agreed well with the measured ones. Jain\(^{24}\) used the ANN approach to establish an integrated stage-discharge-sediment concentration relation. The data of two gauging sites was used to compare the performance of ANN and the conventional curve-fitting approach. It was concluded that for both data sets, the ANN results are much closer to the observed values than the conventional techniques. The ANN approach can successfully model the hysteresis effect that is associated with unsteady flow in open channels. Nagy et al.\(^{25}\) used the ANN approach to estimate the natural sediment discharge in rivers in terms of sediment concentration. Their network was trained based on collected field data for different natural streams. They stated that the ANN approach produces more better results compared to several commonly used formulae of sediment discharge. In the present application the training data was generated using previously developed numerical models and the prediction of the ANN (for test data set) were compared to both field data and the numerical model results.

### 4. COLLECTION OF DATA FOR TRAINING, VALIDATION AND TEST

The numerical sediment transport model (SED-2) developed by Elfiky et al.\(^{8}\) was used to generate the suspended sediment load ((kg/m.sec) for a canal reach of 830 m long. The two basic
inputs of the SED-2 model (velocities) was obtained by running the HYD-2 model that was developed
by Elfiky et al.\textsuperscript{26}. The inputs to the SED-2 model include the velocities in x and in y directions, the
flow depth and the sediment transport capacity. The output of the model is the suspended sediment
concentration. Figure 2 shows a definition sketch for the reach where the model was applied. The
reach is confined between cross section (1) at KM 57.000 and cross section (2) at KM 57.830 on El-
Noubaria canal. Also, El-Nasser canal was included in the simulation using SED-2. Since the site
where the numerical model was applied to generate the data, consisted of two canals, each canal was
simulated separately using the ANN because the change in the flow direction at the canal junction was
misunderstood by the network leading to very poor neural network model. The effective total number
of generated data points are 963 for El-Nubraia canal reach.

5. COLLECTION OF FIELD DATA

The collected field measurements at two cross-sections on EL-Nubaria main canal are used to
compare the model results. The data were collected by the Hydraulic Research Institute (HRI),
National Water Research Center (NWRC), Delta Barrages, Egypt on November, 1998, Saad et al.\textsuperscript{27}.
The average velocity and the suspended sediment load were measured at two stations along both of
Sec. (1) at KM 57 and Sec. (2) at KM 57.83 on EL-Nubaria canal, see Figure 2.

6. BUILDING THE NETWORK

Many factors affect the accuracy of the network. Among these factors, the following ones
important.

6.1 Normalization of Data

Normalization ensures that each input contributes equally to the decision or the prediction
made by the network. If the input values were not normalized, an input data, which have large
numbers, will be more significant than that which has small numbers. Several methods could be used
for normalization. One of these methods the zero-mean unit-standard deviation normalization method
in which the mean and the standard deviation for each field is determined. Each field is then
normalized such that the mean value for the field becomes zero and the values at plus and minus one
standard deviation are mapped onto plus and minus one.

According to the Neural Connection\textsuperscript{28}, the normalized input data, which are provided to the neural
network, are classified into three sets, i.e. training, validation and test data sets. The training data is
used to train the proposed ANN and is taken as 70\% of the total records (2/3 of the data may be
enough for large set of data). Validation data is used to monitor neural network performance during
training phase and it represents 15\% of total input data. Test data is used to test the performance of a
trained ANN in generating the required prediction. The test data set is unseen data to the ANN model
and represents 15\% of the total utilized records by the present application.
6.2 The Initial Values of the Connections Weights
The choice of the connections weights have a large effect on the performance of the network. The best initial values of the connections weight are found by trial and errors by conducting many computer experiments and the correlation coefficient, R, between the target and the output of the proposed network is computed for each experiment. Also, the root mean square error, rmse, is computed. The values of the weights that generate output with maximum R and minimum rmse are chosen. In the present application, the best initial weights was assumed to be in the range ± 1.

6.3 The Number of Neurons in Each Hidden Layer.
Generally, increasing the number of neurons in the hidden layer improves the performance of network on the training data, but not necessarily on the validation data. If so many hidden neurons are used in a network, the network will have enough weights to exactly represent all the training patterns. Such network will be poor network because it will be able to generalize the solution. This mean means that the network is overtrained. Either overtrained or undertrained networks is not desirable. The correct number of neurons in the hidden layer is assessed by looking at the performance of the network on the validation data. As the total number of hidden units is increased from one, the network performance on the validation data increases rapidly. This is because each new hidden unit starts to represent one of the underlying features in the data set. As more units are added, performance levels off. At that point, the training should stop. However, adding further units may then cause a decrease in performance because the power of generalization is lost and the network begins to learn the noise present in the data. It is always better to use as few neurons as possible to achieve the desired result. Generally, the number of neurons depends on the complexity of the data and on both the number of input and output variables. From experience, a rough initial estimation to the number of neurons in the hidden layer may be the geometric mean of the neurons in both input and output layers. The procedure is achieved by conducting many computer experiments. In the present application, the best number of neurons in the hidden layer is 10. The results of the conducted experiments are presented in Figure 2 in terms of R and rmse. The best value of R and the minimum value of rmse are when the network has a size of 4-10-1. It should be noted a similar figure to figure 3 is prepared to select the optimal value of each of the important factors affecting the ANN performance but not presented here to avoid repetition.

6.4 The Number of Hidden Layers.
In most of the application one hidden layer will produce enough accuracy. However, more than one hidden layers can be used based on the complexity of the data structures. This can be achieved by conducting several computer experiments using single and multiple hidden layers and then a decision is taken based on the performance of the network. In this application, one hidden layer is found to be enough.

6.5 Activation Function at the Hidden Neurons
The type of activation functions used in the hidden layer is chosen by trials. In this application the tanh activation function is found to be the best one compared to the linear or the sigmoid.

6.6 The Learning Algorithm
The learning algorithm affects highly the performance of the networks. In the present application, the conjugate gradient is used to prevent the network from being trapped in a local minima. Unlike back-propagation, the conjugate gradient method does not proceed along the direction of the error gradient, but in a direction orthogonal to the one in the previous step. This prevents future steps from influencing the minimization achieved during the current step. The most common learning algorithms for multi-layers feedforward networks was introduced in the paper of ASCE task committee on application of ANN. In addition to the above factors, the maximum number of updates is important which is fixed when the validation error reaches to each minimum during training process.

Keeping in mind the above discussed factors, building the network for the present application is well represented by the flowchart shown by Figure 4.
Figure 3. Typical performance of the proposed network in terms of (a) R and (b) rmse while determining the best number of neurons in the hidden layer of the present application

### 7. RESULTS OF THE DEVELOPED NETWORK

The developed ANN model 4-10-1 with a tansh activation function and 1000 iterations (maximum updates) is used to predict the suspended sediment load for the El-Nubaria canal through the reach from KM 57.000 to KM 57.830. The results of the network are presented in three figures. Figure 5 presents the comparison between the ANN estimation and the values predicted from the previously developed numerical model (SED-2) for training data set. The correlation coefficient, R, for this set of data is 0.9993. Clearly, perfect agreement is obtained for this set and this expected because the generated data from the numerical model was used to train the network. The very few data points which seem to deviate from the line of perfect agreement are those points where the velocities are affected by the entering flow to the El-Nasser canal and hence the suspended sediment is also affected because a remarkable portion of suspended sediment flow to El-Nasser canal. It should be noted that El-Nasser canal was not included in the simulation using the neural network, in spite of its inclusion in the numerical model, because its inclusion interrupts the performance of the network. Figure 6 presents the results of ANN model for validation and test data sets versus those of the numerical model. The correlation coefficient is \( R = 0.9993 \) for validation data and equals \( R = 0.9992 \) for test data. The correlation coefficient for all data set is 0.9993. Figure 7 represents the variation of the residuals for all the three data sets versus the network predictions. The residuals seem to be distributed around the line of zero error, uncorrelated with the ANN outputs (estimated and predicted) and of very small values. The correlation coefficient of the residuals with the network prediction is very small and equals -0.0272. In this figure, the high values of the residuals are for the points that affected by El-Nasser canal where the velocity component in x direction is suddenly affected (because it changes its direction and becomes in y direction as the flow enters El-Nasser canal) and the suspended sediment load is in turn reduced compared to the upstream sections.
1. Iterations to find the best weights
2. Iterations to find best activation function
3. Iterations to find no. of hidden neurons
4. Iterations to find the number of hidden layers
5. Iterations to find max number of updates

Calculation of the R and rmse

Check stability of the network

Yes

Save the output of the best networks as i.e. training, validation and test data sets

Post processing of output and further analysis

Stop

Figure 4. Flowchart showing the basic steps of building a neural network for an application
Figure 5. Comparison between predictions of ANN and those of SED-2 numerical model for training data set (70% of whole data)

Figure 6. Comparison between predictions of ANN and those of SED-2 numerical model for both validation and test data sets (30% of whole data)
COMPARISONS

The collected field data at the two cross sections, 1 and 2 (Figure 2) are compared to both the predicted values using the numerical model and the neural network model in Figures 8 and 9 for sections (2) and (1). Clearly, good matching is observed between the field data and the models results at section (2). The results of both models are very close to each other. At Sec. (1), there are great agreement at the left and right stations while a gap is noticed between the models results and the field data in the middle station, perhaps due to the infl ow boundary effect, Elfiky et al.\textsuperscript{8)} Comparisons between results of ANN model and the numerical model at other sections as (3) and (4) indicate very great agreement (not presented here to reserve space). Comparisons between results of ANN model and those of numerical model at the longitudinal sections as L.S.1, L.S.2 and L.S.3 show also very close agreement between both results. Indicated in Figure 10, the comparison for L.S.1. representing the worst case among the three sections.

Fig. (8) Comparison between predictions of ANN, SED-2 numerical model and field data for Test data sets at section (1)
9. CONCLUSION

A multilayer feedforward artificial neural network (4-10-1) is used to estimate the suspended sediment concentration efficiently based on four inputs including the depth of flow, the components of flow velocities in x and y directions and the sediment carrying capacity. Since, the field data are very limited, a 2-D numerical model (SED-2) was used to generate the required training and validation data for the developed neural network. The present paper proved that the ANNs are a powerful computational tool for computing the suspended sediment concentration in rivers provided that the trained and verified network should be used to predict values within the training range otherwise, poor predictions are obtained.

10. REFERENCES