Conference Paper, Published Version

Asadiani Yekta, A. H.; Marsooli, R.; Soltani, F.

Suspended sediment estimation of Ekbatan reservoir sub basin using Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Artificial Neural Networks (ANN), and sediment rating curves (SRC)

River Flow 2010 - Dittrich, Koll, Aberle & Geisenhainer (eds) - © 2010 Bundesanstalt für Wasserbau ISBN 978-3-939230-00-7

Suspended sediment estimation of Ekbatan Reservoir Sub Basin using Adaptive Neuro-Fuzzy Inference Systems (ANFIS), Artificial Neural Networks (ANN), and Sediment Rating Curves (SRC)

Amir Hossein Asadiani Yekta
Islamic Azad University, Hamadan Branch, Hamadan, Iran

Reza Marsooli
The University of Mississippi, UNIVERSITY, MS, USA

Foad Soltani
Institute of Civilization and Development, Hamadan, Iran

ABSTRACT: In this study, the intelligent methods such as fuzzy logic and neural networks are used to predict the suspended sediment load as a function of water discharge data and the results compared with rating curve method. Two intelligent methods, Artificial Neuro-Fuzzy Inference Systems (ANFIS) and Artificial Neural Networks (ANN), are trained using measured data of Ekbatan reservoir sub basin, which is located in the west of Iran, between 1964 and 2006. To identify the influence of time period of water discharge data on suspended sediment load, four scenarios are examined and the best scenario, which simulates suspended sediment load at time""using water discharge data at the same time, is used for prediction of suspended sediment laod of Ekbatan sub basin.The Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Relative Error (RE) are used as error evaluation criteria to train, verify, and compare the results of developed models. Results show that the developed intelligent methods, Artificial Neuro-Fuzzy Inference Systems (ANFIS) and Artificial Neural Networks (ANN), are more accurate than rating curve. Study results also show that the ANFIS is more accurate than ANN for estimation of suspended sediment load as a function of discharge.

Keywords: ANFIS, ANN, Rating Curve, Suspended Sediment Load, EKbatan Reservoir

1 INTRODUCTION

Although suspended sediment load can be predicted using numerous developed equations their results often differ from each other and from measured data due to complexity of sediment transport nature. In recent years, simulation models for prediction of suspended sediment load of rivers have been popular among researchers because of Progress of computer models. Artificial Neural Networks (ANN) and Artificial Neuro-Fuzzy Inference Systems (ANFIS) are two well-known models for prediction of hydraulic and hydrology events. Many researchers have studied the application of Artificial Neural Networks in vital topics of hydrology and hydraulics such as prediction of sediment load, rainfall-runoff modeling, flow prediction etc. Cigizoglu (2002) made a comparison between ANNs and SRC for suspended sediment estimation and found that the estimations obtained by ANN's were significantly superior to the corresponding classical sediment rating curve ones. Agarwalet et al (2006) simulated the runoff and sediment yield using artificial neural networks as daily, weekly, ten-daily, and monthly monsoon runoff and sediment yield from an Indian catchment using back propagation artificial neural network (BPANN) technique, and compared the results with observed values obtained from using single- and multi-input linear transfer function models. They showed that the ANN model gives pretty reliable results. Kisi (2005) investigated the abilities of neuro-fuzzy (NF) and neural network (NN) approaches to model the stream flow–suspended sediment relationship for two stations—Quebrada Blanca station and Rio Valenciano station—operated by the US Geological Survey. He found that the NF model gives better estimates than the other technique.

The scope of this study is the suspended sediment estimation of Ekbatan dam using an intelligent method to get more accurate results compared to the rating curve. Two ANN and ANFIS algorithms are trained using measured water and sediment discharge data of Yalfan gauging station which is located at the entrance of Ekbatan dam in Iran. Dependency of suspended sediment load at time t to water discharge at different time periods, such as t and t-1, is evaluated using four scenarios for measured data of a US Geological Survey station where sufficient required data are available. The best scenario is applied for suspended sediment estimation of Ekbatan sub basin using ANFIS and ANN.

Four statistic parameters have been used to determine the accuracy of the models.

2 ANN MODEL

An ANN consists of a number of data processing elements called neurons or nodes that are grouped in layers. The input layer neurons receive input data or information and transmit the values to the next layer of processing elements via connections. This process is continued until the output layer is reached. This type of network in which data flows in one direction (forward) is known as a feed-forward network. The application of ANN models has been the topic of a large number of recent literatures, such as the book by Lingireddy and Brion (2005).

A model of a neuron has three basic parts: input weights, a summer, and an output function. The input weights scale values used as inputs to the neuron, the summer adds all the scaled values together, and the output function produces the final output of the neuron. Often, one additional input, known as the bias is added to the system. If a bias is used, it can be represented by a weight with a constant input of one. Figure 1 shows a simple ANN with three inputs and one output.

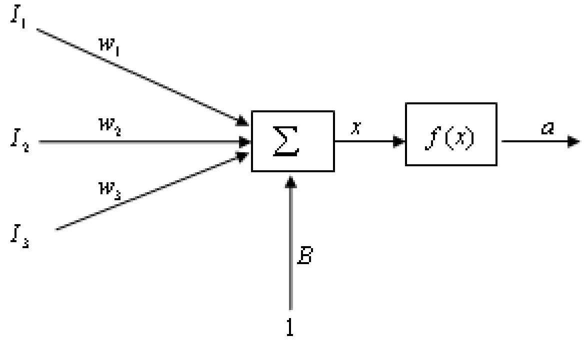

Figure 1. A schematic neuron model

I_1, I_2 and I_3 are the inputs, w_1, w_2 and w_3 are the weights, B is the bias, x is an intermediate output, and a is the final output. The equation for a is given by

$$a = f(w_1 I_1 + w_2 I_2 + w_3 I_3 + B) \qquad (1)$$

where f is a transfer function.

Detailed information can be found in related literature such as Lingireddy and Brion (2005).

3 ANFIS MODEL

Jang (1993) presented Adaptive Neuro-Fuzzy Inference Systems (ANFIS). As Figure 2 shows, the ANFIS model includes 5 layers that are summarized below.

Layer 1: Every node i in this layer is an adaptive node with node functions such as

$$O_{l,i} = \mu A_i(x) \qquad (2)$$

where x is the input to the ith node and A_i is a linguistic label associated with this node. $O_{l,i}$ is the membership grade of a fuzzy set A and it specifies the degree to which the given input x satisfies the quantifier A. The membership functions for A can be shown as

$$\mu_A(x) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2b_i}} \qquad (3)$$

where a, b, and c are the parameter sets. As the values of these parameters change, the bell-shaped function varies accordingly, thus exhibiting various forms of membership functions on linguistic label A_i. In fact, any continuous and piecewise differentiable functions, such as commonly used triangular-shaped membership functions, are also qualified candidates for node functions in this layer. Parameters in this layer are referred to as premise parameters. The outputs of this layer are the membership values of the premise part.

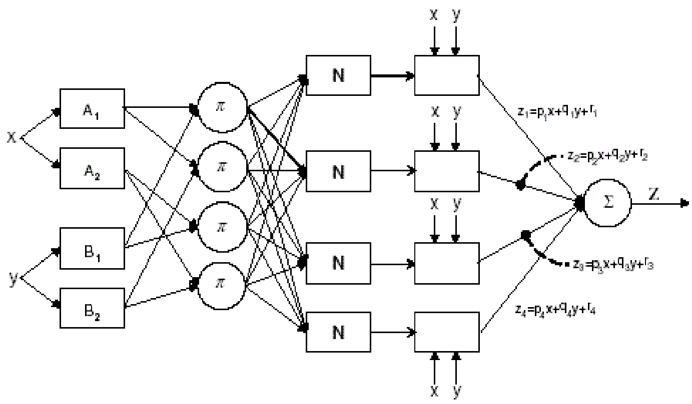

Figure 2. ANFIS architecture.

Layer 2: This layer consists of the nodes labeled π which multiplies incoming signals and sends the product out. For instance, in node 1

$$w_i = \mu_{A_i}(x_1)\mu_{B_i}(x_2) \quad (4)$$

Layer 3: In this layer, the nodes labeled N calculate the ratio of the ith rule's firing strength to the sum of all rules' firing strengths

$$\overline{w} = \frac{w_i}{\sum_1^n w_i} \quad , i = 1,2,...,n \quad (5)$$

Layer 4: This layer's nodes are adaptive with node functions

$$Z_i = \overline{w}_i f_i = \overline{w}_i (p_i x_1 + q_i x_2 + r_i) \quad (6)$$

where p_i, q_i, and r_i are named consequent parameters.

Layer 5: This layer's single fixed node labeled Σ computes the final output as the summation of all incoming signals

$$\sum_{i=1}^{n} \overline{w}_i f_i \quad (7)$$

More information on ANFIS can be found in Jang (1993).

4 SEDIMENT RATING CURVE

Regression approach is an alternative to estimate the suspended sediment load concentration. Sediment-discharge rating curve is a regression approach that connects the river hydrology to sediment transport. The rating curve estimates sediment load using flow discharge in the following form

$$Q_s = aQ_w^b \quad (8)$$

where Q_s is the suspended sediment discharge, Q_w is flow discharge, and a and b are the coefficients determined by regression analysis.

5 EVALUATION OF INPUT AND OUTPUT DATA

To evaluate the importance of choosing proper sediment and water discharge in the case of time period, the measured data of a US Geological Survey gauging station, which contains a wide range of required data, is used for test of different scenarios. The measured data of station number 01442750 (DELAWARE R AT DUNNFIELD, NJ) between 1969 and 1973 is used to train developed ANFIS and ANN models. Of course, the relation of sediment and water discharge differs in Delaware and the gauging station used in Ekbatan case study. Because of lacke of measured data at Ekbatan reservoir, the measured data of Delaware river was taken to compare the differen scenarios.

The models are verified using data between 1974 and 1975. Table 1 shows the statistic parameters of data used for training and testing the models. Since the available data include a wide range of data, which can be seen in Table 1, the following equation is used to normalize measured data.

$$X_{Normal} = X_i / X_{Max.} \quad (9)$$

where X_{Normal} is normalized data, X_i is ith data, and $X_{Max.}$ is maximum value.

Table 1. Statistic parameters of USGS data

	Data	Average	Max.	Min.	Standard deviation
Training data	Discharge (m^3/s)	190.8	2710	44.7	186.2
	Sediment (mg/l)	11.1	640	0	29.5
Verification data	Discharge (m^3/s)	205.8	2030	44.2	37178
	Sediment (mg/l)	7.8	169	1	135.6

5.1 *Input and Output Scenarios*

If C_t shows sediment concentration at time t, which its unit is day, and water discharge at the same time is Q_t, four following scenarios are evaluated using ANFIS and ANN to identify the best relation between sediment and water discharge.

1- C_t and Q_t
2- C_t and C_{t-1}
3- C_t, Q_t, and Q_{t-1}
4- C_t, C_{t-1}, Q_t, and Q_{t-1}

5.2 *Error Evaluation Criteria*

The Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and Relative Error (RE) are used to estimate the quality of results with measured data.

$$MAE = \frac{1}{n}\sum_{i=1}^{n} \left| y_{actual_i} - y_{forecast_i} \right| \quad (10)$$

$$RE = \frac{(y_{actual} - y_{forecast})}{y_{actual}} \times 100 \quad (11)$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n} \left(y_{actual_i} - y_{forecast_i} \right)^2 \quad (12)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{actual_i} - y_{forecast_i}\right)^2} \quad (13)$$

5.3 *ANFIS and ANN Development*

Two MATLAB codes were prepared for ANFIS and ANN models. These codes and measured data of DELAWARE station are used to train and verify the models. After try-and-error tests, the best ANFIS and ANN network found in this study is the one with 3 layers including one input layer, one middle layer, and one output layer. The best membership function in ANFIS model is the Gaussian function which input parameters are classified as low, average, and high. The architecture of the ANFIS and ANN models are given in Table 2. The third column indicates which input parameters are classified as low, average, and high in the ANFIS model. The fourth colum indicates which the ANN models have 3 layers. The first digit stands for the number of input data, the second one for the number of neurons in the middle layer, and the last one for the number of output data.

Table 2. Developed ANFIS and ANN models

Scenario	Input data	ANFIS	ANN
1	Q_t	3	(1,1,1)
2	C_{t-1}	3	(1,1,1)
3	Q_t and Q_{t-1}	3	(2,1,1)
4	C_{t-1}, Q_t, and Q_{t-1}	3	(3,1,1)

All four scenarios have been modeled using developed ANFIS and ANN models and their results compared using the error evaluation criteria. Comparison of results reveals that the first scenario has the most accurate estimation. Therefore, estimation of sediment load at time *t*, C_t, using water discharge at the same time, Q_t, has been chosen as the best scenario. The estimation errors of ANFIS and ANN models have been compared with sediment rating curve, SRC, in Table 3. The second scenario, which relates C_t to C_{t-1} shows a big error which is because of the less effect of these two sediment parameters in prediction of the sediment discharge compared to affect of water discharge. Also, scenario 4 consist of C_{t-1} which results to big error. This comparison indicates that the estimation of the sediment load at time *t*, C_t, is not affected by C_{t-1}.

Table 3. Comparison of accuracy of models

Model	Error	Scenario 1	2	3	4
ANFIS	MAE	3.18	4.38	6.03	3.83
	MSE	24.3	100.7	110.6	124.2
	RMSE	4.9	10.0	10.5	11.1
	RE (%)	3.83	124.2	11.1	19.2
ANN	MAE	8.1	8.8	11.1	7.7
	MSE	179	403.7	418.5	496.7
	RMSE	13.5	20.1	20.4	22.3
	RE (%)	21.1	24.0	49.4	32.2
SRC	MAE	9.0	-	-	-
	MSE	196.4	-	-	-
	RMSE	14.0	-	-	-
	RE (%)	37.4	-	-	-

Also, the efficiency of models to estimate the peak value of sediment concentration has been tested and the results are summarized in Table 4. This table shows that while ANN and SRC models have significant error to estimate peak sediment concentration, ANFIS model has closer results to measured data.

Table 4. Accuracy of models for peak sediment estimation

Measured peak data	Estimated data ANFIS	ANN	SRC	Relative error % (RE) ANFIS	ANN	SRC
227	250	297	97	10.3	30.8	57.2
180	200	240	59	11.1	33.4	67.2
140	120	81	43	14.1	42.2	69.2
168	151	116	119	10.3	30.8	29.3
78	85	99	28	9.0	27.0	64.2
112	101	79	94	9.8	29.5	15.8
99	112	138	21	13.1	39.4	78.3
200	190	170	63	5.0	15.1	68.4

In general, the results of ANFIS and ANN models which were trained and verified using measured data of a US Geological Survey gauging station revealed that the estimation of suspended sediment load at time t is more accurate when water discharge at the same time is used as input data. Therefore, the first scenario is applied for estimation of suspended sediment load of Ekbatan sub basin.

6 SUSPENDED SEDIMENT ESTIMATION OF EKBATAN RESERVOIR

6.1 *Ekbatan Sub Basin*

The sub basin of Ekbatan reservoir extends between $48^\circ\,28'$ and $48^\circ\,42'$ latitude and between $34^\circ\,35'$ and $34^\circ\,45'$ longitude. This sub basin is located at the southeast of Hamadan which is one of the biggest cities in the west of Iran. The area and perimeter of Ekbatan sub basin are about 221.55 km^2 and 62.2 km, respectively. The maximum elevation of this basin is 3584 and the minimum one is 1920 meter above the sea level. The Ekbatan basin includes two rivers called Abro and Abshineh with two gauging stations, Abro and Yalfan, which are located in these rivers, respectively. Water discharge data has been measured from 1955 to 2006 at the Yalfan station. Also, measured suspended sediment concentration data are available between 1964 and 1995 and another time period between 2003 and 2006.

6.2 *Results*

ANFIS and ANN models have been trained using the best scenario of the previous sections, which relates sediment load at time *t*, to water discharge at the same time. Also, for raising the accuracy of the results, logarithmic scale of data have been used to train and verify the developed models.

The ANN model has been trained using 300 sets of data at the Yalfan station which include measured sediment concentration and water discharge data. Also, 50 sets of measured data have been used for verification. After try-and-error tests, the best ANN model found for Ekbatan case study is the one with 3 layers including one input layer, one middle layer, and one output layer. The best model consists of one input, 2 middle, and one output nodes which should have a Sigmoid transfer function. The accuracy of the developed ANN model for three different types of inputs and targets, raw data, normalized data, and logarithmic scale of data, is shown in Table 5. This table shows that the ANN model with logarithmic normalized input and target data gives more accurate results. Data normalization decreases the data changes amplitude which results decrease in error of a model.

Table 5. Accuracy of ANN model of Ekbatan case study

Error	Raw data	Normalized data	Logarithmic normalized data
MAE	118.7	0.049	0.133
MSE	38640	0.0131	0.0311
RMSE	196.57	0.114	0.176
RE (%)	29.8	26.87	14.01

The ANFIS classifiers were designed by using bell shaped membership function, Gaussian membership function, and triangular membership function.

The comparison of results showed that the bell shaped Gaussian membership function has the most accurate results among other tested membership functions. Also the model classifies input data to 3 classes. Table 6 compares the results of ANFIS model with different number of membership function and different type of input and output data. As this table shows, the best model has three membership functions and input data must be in logarithmic normalized scale.

Table 6. Accuracy of ANFIS model of Ekbatan case study with Gaussian membership function

Type of input data	Error	Number of membership functions 2	3	5
Raw data	MAE	109.6	91.8	93.4
	MSE	42731	36123	33007
	RMSE	206.7	190.0	181.6
	RE (%)	1.97	7.83	10.3
Normalized data	MAE	0.051	0.059	0.057
	MSE	0.012	0.012	0.012
	RMSE	0.112	0.111	0.110
	RE (%)	23.3	14.7	13.3
Logarithmic normalized data	MAE	0.13	0.14	0.14
	MSE	0.030	0.033	0.033
	RMSE	0.175	0.183	0.184
	RE (%)	1.55	1.03	1.85

The comparison of results of ANFIS and ANN models for Ekbatann basin are denoted in figures 3 and 4. In these figures, gb means the bell-shaped Gaussian membership function and the followed digits indicates the Number of membership functions.

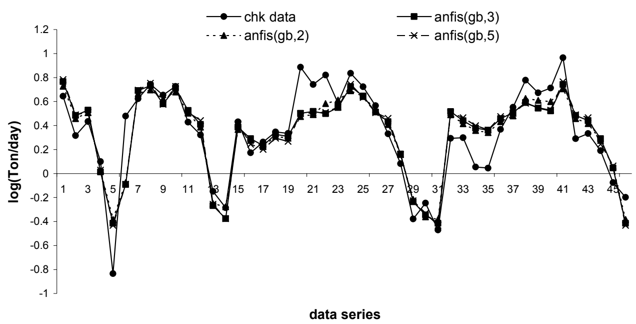

Figure 3. Comparison of ANFIS results with measured data

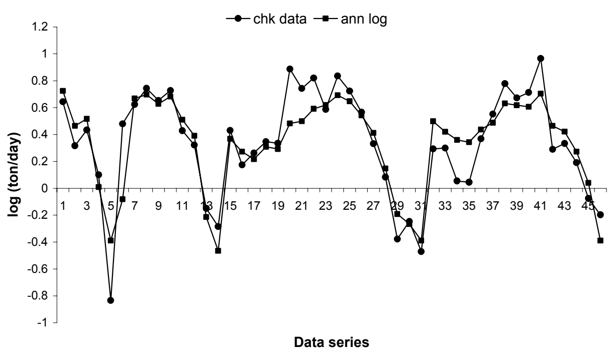

Figure 4. Comparison of ANN result with measured data

The developed sediment rating curve of Yalfan station has been used to compare the results of ANFIS and ANN models with this regression approach. Ganji (2006) classified the measured sediment data to 10 statistical classes and used data classification method to show the relation between incoming sediment discharge and water discharge by the following relation.

$$Q_s = 21.421 Q_w^{1.6602} \tag{14}$$

Where Q_s is the sediment discharge (ton/day), and Q_w is the water discharge (m^3/s).

In this study, the regression coefficient was modified using FAO modification method. Revised rating curve in USBR and data classification methods are shown by equations (15) and (16), respectively.

$$Q_s = 17.76 Q_w^{1.46} \tag{15}$$

$$Q_s = 54.39 Q_w^{1.66} \tag{16}$$

The results of these rating curves, which are compared in Table 7, reveal that the rating curve developed by data classification method is more accurate compared to other methods.

Table 7. Accuracy of rating curve of Ekbatan case study developed by different methods

Error	USBR	Data classification	FAO	
			USBR	Data classification
MAE	85.79	97.76	80.66	237.92
MSE	49366	37080	39556	161674
RMSE	222.18	192.56	198.88	402.08
RE (%)	62.67	-10.03	31.83	-179.30

Table 8 shows the comparison of ANFIS, ANN, and SRC models that have been developed using logarithmic normalized data. The comparison reveals that the ANFIS and ANN models are more accurate methods than sediment rating curve to estimate the suspended sediment concentration of Ekbatan basin. Also the ANFIS model has the best results among all three evaluated methods. The error evaluation parameters MAE MSE, and RMSE compare the time series data one by one and the error would be calculated while RE computes the total error of the time series. Because the total sediment load is the main issue in the reservoir sedimentation problems, and also because of the insignificant difference among MAE, MSE, and RMSE, as can be seen in Table 8, RE was taken as the main evaluation criterion.

Table 8. Comparison of ANFIS, ANN, SRC models

Error	ANFIS	ANN	SRC
MAE	0.14	0.133	97.76
MSE	0.033	0.0311	37080
RMSE	0.183	0.176	192.56
RE (%)	1.03	14.01	-10.03

7 CONCLUSION

The best scenario for estimation of suspended sediment load using water discharge and an intelligent model such as ANFIS and ANN is training these models with measured input and output data at the same time. In other words, the best combination of required data is the measured suspended sediment load data at time t as input and measured water discharge at the same time, t, as target or output.

The results of developed intelligent models revealed that the ANFIS model is more accurate than ANN and rating curve to predict the concentration of suspended sediment load and also rating curve is not as accurate as ANFIS and even ANN. Also it is concluded that normalized and logarithmic normalized input data produce more accurate results than raw input data.

REFERENCES

Agarwal, A. h., Mishra, S. K., and Singh, J. K. 2006. Simulation of runoff and sediment yield using Artifitial Neural Networks. Biosystems Engineering, 97(4), 597-613.

Ariffin, J., Abdul Ghani, A., Zakaria, N., Shukri Yahya, A. 2003. Sediment prediction using ANN and regression approach, 1st International Conference on Managing rivers in the 21st Century: Issues & Challenges

Asadiani. A.H., Slotani. F., Suspended sediment estimation using artificial neutral network (ANN) and adaptive nero-fazzy inference system (ANFIS),case study Ekbatan dam watershed, research report, 2008, Islamic Azad University, Hamedan branch, in Farsi

Bojadziv, G., Bojadziv, M. 1995. Fuzzy Sets, Fuzzy Logic, Applications, World Scientific, New Jersey, pp. 283.

Cigizoglu, H. 2002. Suspended sediment estimation for rivers using artificial neural networks and sediment rating curves. Turkish journal Eng.Env.Sci, 26, 27-36.

Ganji, N. 2006. Evaluation of Ekbatan reservoir sedimentation using sediment rating curve, Master thesis, Tarbiat Modares University, Tehran, Iran

Gorzalezany, M.B., Clusezek, A. 2000. Neuro-Fuzzy Systems for Rule Based Modeling of Dynamic Processes, Proceedings of ESIT, Aachen, Germany, pp 416-422.

Jang, J. S. R. 1993. ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Sys. Manage. and Cybernetics 23(3), 665–685.

Kisi, O. 2005. Suspended sediment estimation using neuro-fazzy and neural network approaches. Hydrological Sciences Journal, IAHS Press, 50(4), 683–696.

Li, H., Chen, G. L. P., Huang, H. P. 2001. Fuzzy Neural Intelligent Systems. Mathematical Foundation and the Applications in Engineering, CRC Press, pp. 371.

Lingireddy, S., Brion, G. M. 2005. Artificial Neural Networks in Water Supply Engineering. Published by the American Society of Civil Engineers