

Suspended sediment concentration estimation using artificial neural networks and neural-fuzzy inference system case study: Karaj Dam

Abdolmajid Mohammadi*, Gholamhossein Akbari and Gholamreza Azizzian

Civil Engineering Department, Engineering Faculty, Sistan and Baluchestan University, Zahedan, Iran.
abdolmajidmohammadi.civil@gmail.com*

Abstract

Simulation and sediment assessment of the river are of the significant and practical issues in water resource management. To estimate the suspended sediment concentration of Karaj dam in this study, simultaneous water discharge data, base, water temperature and sediment density of Siraa Station located at Karaj dock entry have been used. Artificial Neural Fuzzy Inference System (ANFIS), Artificial Neural Network (ANN) and Sediment Rating Curve (SRC) modeling was used. Correlation coefficient (R) and Root Mean Square Error (RMSE) are considered the model's assessment criteria. The results show a higher accuracy of fuzzy model assessments in comparison with neural networks and sediment rating curve assessments.

Keywords: Sediment, Neural-fuzzy inference system, artificial neural network, Sediment rating curve

Introduction

Sediments are transferred by the rivers due to passing from sedimentary basins and cause erosion of the sides and bottom and also sedimentation throughout its course. The proper understanding of the features and river behaviors in order to apply the proper maintenance and management programs for river health is an essential and unavoidable issue. According to the reports delivered, many of our country's docks are faced with sediment problem and premature filling of reservoirs. It is evident that given the significance of the issue, paying more attention to this issue can play a major role in the country's economic development and lack of attention to its measurement and precise calculation is led to wasting away the national capitals (Muhammad Zamani, 2007). Artificial neural network in the field of water engineering is considered to solve various problems which satisfactory results are achieved, including the study performed by Ebrahimi Mohammadi *et al* (2010) compared in an article the efficiency of artificial neural network models, multivariate regression and sediment rating in assessing the daily suspended loading of Koreh Sang Station located at Haraz River based on daily precipitation and discharge as well as taking into consideration the RMSE and R criteria which indicate a better performance of artificial neural network models (Ebrahimi & Mohammadi, 2010). Nourani *et al* (2009) introduced a model based on fuzzy logic structure to estimate sediment suspended load of Khaivchay River located at Ardebil province which had a better results in comparison with classical methods and also artificial neural networks (Nourani, 2009).

Study area and analyzing data used

Geographical Position of the Study Area

The study area of this research is Karaj dock. Karaj dock is located at the permanent Karaj River which is considered the part of Karaj catchment. Considering geographical position, this basin is located between $35^{\circ}53'40''$ to $36^{\circ}10'45''$ latitudes and $51^{\circ}.2'3''$ to $51^{\circ}22'.6''$ in the north and north-west of Tehran. Also, the space of

this area has been calculated with Arc GIS 8.5 software which the extent of basin is hereby 832236736 square meters (Kordowani, 2006).

Data collection

To estimate sediment load, hydraulic and hydrological data of Siraa Station located at Karaj Dock Entry have been used. Data required for this study is related to the report on regional water organization in Tehran province. Data used with 40 years statistics associated with the years from 1967 to 2007 and 3037 measured data including Celsius temperature, base according to centimeter, course discharge based on Cubic meters per second and sediment density based on tons per day are selected and used in this study. The noteworthy point in this regard is lack of quantitative data which is naturally one of the problems faced with studying this kind.

Some of the statistical information related to data including mean, standard deviation, minimum and maximum data per each variable used are presented in Table 1.

Table 1. Statistical information related to observational raw data.

| Maximum | Minimum | Standard deviation | Mean | Measurement variable |
|---------|---------|--------------------|-------|----------------------|
| 49 | -5.1 | 48.5 | 28 | Water temperature |
| 411 | 1.5 | 252.6 | 184 | Base |
| 30.2 | 0.150 | 23.8 | 14.6 | Flow discharge |
| 1500 | 0.05 | 1736.39 | 201.2 | Sediment density |

Input and output data of artificial neural network and neural-fuzzy inference system

The parameters of simultaneous water temperature, base and flow discharge are considered the inputs of artificial neural network and fuzzy inference system and sediment density parameter are considered the outputs. Of 3037 observational data, 2126 samples in the artificial

neural network and neural-fuzzy inference system are randomly selected for training and 911 samples have been selected at random to be tested. In other words, 70% of these data are applied for network training and the remained 30% are selected for the test. Experimental data are selected in such a way that they are appropriately distributed among the total data single points, maximum and minimum points are not included in the test collection.

Two stances are used for observational data for analyzing artificial neural network and neural-fuzzy inference system. These stances are raw data and normalized data, respectively. To normalize data, relation (1) has been used which make the entries standardized between 0.1 and 0.9.

$$x'_i = \left(\frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \right) \times 0.8 + 0.1 \tag{1}$$

In this relation, X'_i is the normalized input data, X_i is the real amount of input data, X_{\min} and X_{\max} are minimum and maximum entry data, respectively. Finally, the network output can be reverted to its initial stance by making standardized algorithm converted (Bardstani, 2010).

Assessment Indicators

To assess the performance of constructed prediction models in this study, three statistical indicators have been used:

1. Correlation coefficient according to relation (2)

$$CORR = \frac{\sum_{i=1}^n (x_i^o - \bar{x}^o)(x_i^p - \bar{x}^p)}{\sqrt{\sum_{i=1}^n (x_i^o - \bar{x}^o)^2 \sum_{i=1}^n (x_i^p - \bar{x}^p)^2}} \tag{2}$$

2. Error square mean root according to relation (3),

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i^o - x_i^p)^2}{n}} \tag{3}$$

3. Mean absolute error according to relation (4),

$$MAE = \frac{\sum_{i=1}^n |x_i^o - x_i^p|}{n} \tag{4}$$

In the above relation, X_i^o is the observed parameter, X_i^p predicted parameter and n is the number of data. Correlation coefficient which determines the relation

between the two variables changes ranged between -1 and +1. The closer the amount of this indicator to +1, the stronger will be the direct correlation between variables under study. Error squares mean root is an indicator mostly applied to measure the differences between predicted values by a model and the observed real amount of that variable. The above-mentioned differences also called remnants, and this indicator in fact adds up the differences together to present a unified amount from the model power for prediction. Therefore, this indicator is a suitable tool to measure the accuracy of this model. Mean absolute error is also a quantity dedicated to measure the closeness of the predicted values and real outputs of the models. Indicators of error squares mean root and mean absolute error are positive and the closer is it to zero, the less will be error (Houshmand, 2011).

Materials and methods

Artificial neural network

The proper structure used for modeling the issue of sediment density is the networks with progressive orders that multi-layered Perspstron (MLP) network has been used in this study. The general configuration of this network is three entries of water temperature, base and flow discharge and a sediment discharge output. The input layer includes three neurons that receive water temperature, base and flow discharge and is resulted into output layer of sediment discharge. There are also various neurons in the hidden layer. The function of these units' activities in this layer is hyperbolic tangent function. The output of this layer is limited to 1 and -1. Linear activation function has been used in the output layer for increasing the network speed. Not only dose using the linear function increase learning speed several times, but it can provides the output values without making any change in the network.

Learning methods of MLP network is based on the post-observance error algorithm. There are three learning methods for this regard which Markowart-Lonburgh method is the fastest of them and the present network has been trained by this algorithm.

Various neurons have been used in the hidden layer and the number of their optimums for minimizing the error should be determined by trial and error. The network is first trained by little neurons in the hidden layer and the number of neurons gradually increases and the network will be re-trained. Because the initial selection of weight and threshold is randomly in each training stage and this issue causes the network is located at different initial conditions, so each structure has been tested by different initial structure in order to ensure not to falling into local minimums and gaining universal minimum (Akbari, 2011)

Artificial neural-fuzzy inference system

The coding system was performed in order to design Artificial Neural-Fuzzy Inference System using MATLAB program. In this code, genfis 2 (datin, datout, r) order was used to generate fuzzy inference system in that:

Table 2. Results obtained from ANN prediction capability

| MAE | RMSE | CORR | Stage | The model under study | Type of data |
|--------|--------|--------|----------|-----------------------|--------------|
| 1.1005 | 2.1099 | 85.27% | Training | ANN | Raw |
| 1.4607 | 2.5109 | 84.14% | Test | | |
| 0.1091 | 0.1346 | 91.78% | Training | ANN | Normalized |
| 0.814 | 0.1006 | 92.66% | Test | | |

Table 3. Results achieved from ANFIS prediction capability

| MAE | RMSE | CORR | Stage | The model under study | Type of data |
|--------|--------|--------|----------|-----------------------|--------------|
| 0.8964 | 1.0685 | 90.45% | Training | ANFIS | Raw |
| 0.7897 | 0.9029 | 91.69% | Test | | |
| 0.0096 | 0.0135 | 95.16% | Training | ANFIS | Normal ized |
| 0.0102 | 0.0257 | 94.61% | Test | | |

Table 4. Assessing the results of sediment estimation in sediment rating curve

| MAE | RMSE | CORR | Stage | Sediment rating curve model | Type of data |
|--------|---------|--------|----------|-----------------------------|--------------|
| 3.0537 | 3.3156 | 79.84% | Training | $Q_s = 15.22Q_w^{1.299}$ | Raw |
| 2.6183 | 2.9685 | 80.95% | Test | | |
| 1.937 | 2.16581 | 84.76% | Training | $Q_s = 0.8769Q_w^{0.8494}$ | Normali zed |
| 1.6052 | 1.9874 | 85.01% | Test | | |

Table 5. Results obtained from ANN and ANFIS prediction capability for observed raw data

| MAE | RMSE | CORR | Stage | Assessment method |
|--------|--------|--------|----------|-------------------|
| 0.8964 | 1.0685 | 90.45% | Training | ANFIS |
| 0.7897 | 0.9029 | 91.69% | Test | |
| 1.1005 | 2.1099 | 85.27% | Training | ANN |
| 1.4607 | 2.5109 | 84.14% | Test | |
| 3.0537 | 3.3156 | 79.84% | Training | SRC |
| 2.6183 | 2.9685 | 80.95% | Test | |

Table 6. Results obtained from ANN and ANFIS prediction capability for observed normalized data

| MAE | RMSE | CORR | Stage | Assessment method |
|--------|--------|--------|----------|-------------------|
| 0.0096 | 0.0135 | 95.16% | Training | ANFIS |
| 0.0102 | 0.0257 | 94.61% | Test | |
| 0.1091 | 0.1346 | 91.78% | Training | ANN |
| 0.0814 | 0.1006 | 92.66% | Test | |
| 1.037 | 2.1658 | 84.76% | Training | SRC |
| 1.6052 | 1.9874 | 85.01% | Test | |

Datin: input data

Datout: output data

R: clustering radius which varies between zero and 1.

Defining large amounts for r is resulted in producing a little cluster and defining small values for r will produce a large number of small clusters. Each cluster is a representative of a membership function in which each membership function shows a fuzzy set in data used for extracting the conditions (Jome Bidokhti, 2011).

Reducing clustering method has been used to extract the neural-fuzzy system role extraction and optimizing

membership functions parameters. Characteristics of ANFIS models used in this study are as follows:

- Type of fuzzy extraction system: Sogeno system
- Type of membership functions: Gousin functions
- Method of rule production: reducing clustering
- Educational algorithm: Hibrid algorithm
- Number of education courses (Epoch): 100 ones
- Method of optimum determination: trial and error

Sediment rating curve model (SRC)

Of the prevalent methods in estimating the load discharge of suspended sediment in Hydrometric stations is making relationship between sediment discharge data and their corresponding discharge data which is obtained from practicing data between amounts of flow discharge and its corresponding suspended sediment which is formulated as follows:

$$Q_s = a \times Q_w^b \quad (5)$$

In this equation, Q_s is suspended sediment discharge (based on tones per day) and Q_w is its corresponding flow discharge (based on cubic meters per second) and "a" and "b" are constant coefficients of the equation (Dehgahni, 2001).

Results and discussion

Results obtained from artificial neural network and neural-fuzzy inference system

Two stances have been used for observational data to analyze artificial neural network and neural-fuzzy inference system. These stances are raw data and normalized data, respectively. The resulted obtained from predicting ANN and ANFIS models for each raw and normalized stance have been shown in Tables 2 and 3. By comparing the network assessment parameters for each stance of observational data, it can be generally said that data normalization in the neural network is associated with improved network performance.

In Fig. 1 to 8, predicted density indicators by artificial neural network and neural-fuzzy inference system against observed measured density indicator for the set of training and test data have been presented.

Results of sediment rating curve model (SRC)

In the sediment rating curve, data have been divided into training data and test data the same as artificial neural system and neural-fuzzy inference system. Training data were used for power fitting curve and test data for analyzing model. Results of sediment rating curve with normal and real data have been shown in Table 4. There is no need for data normalization in sediment rating curve and the results are achieved according to input data ranges; however, sediment rating model was also formed by normal data to compare the results of neural network models and neural-fuzzy inference system. Also in Fig. 9 sediment ratings system can be observed in both raw and normalized stances corresponding to training data.



Fig. 1. Comparison of ANN results with raw observed values for training stage

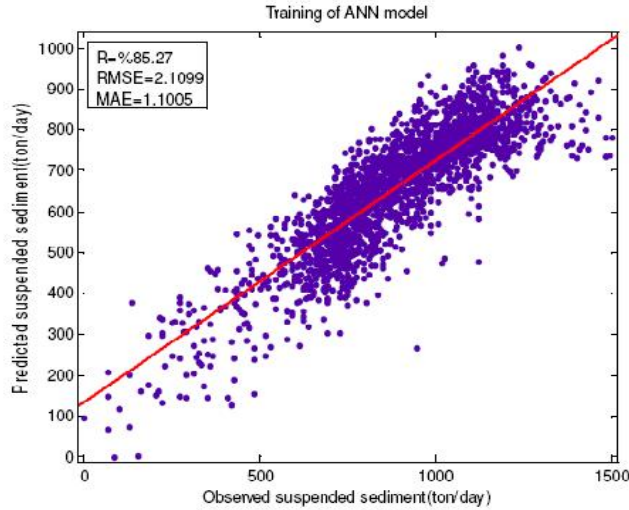


Fig 4. Comparison of ANN results with normalized observed values for test stage

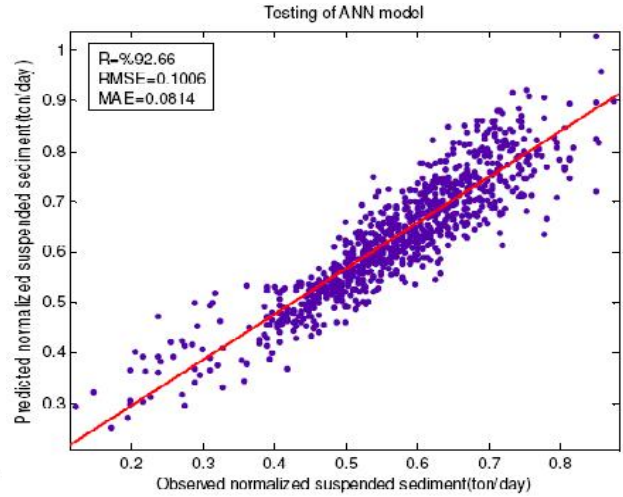


Fig. 2. Comparison of ANN results with raw observed values for test stage

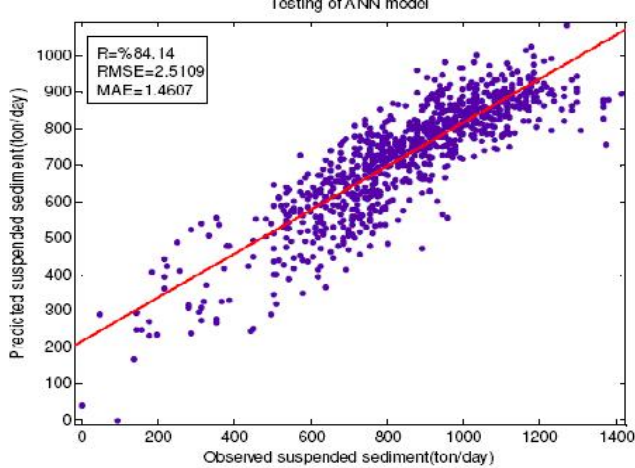


Fig 5. Comparison of ANFIS results with raw observed values for training stage

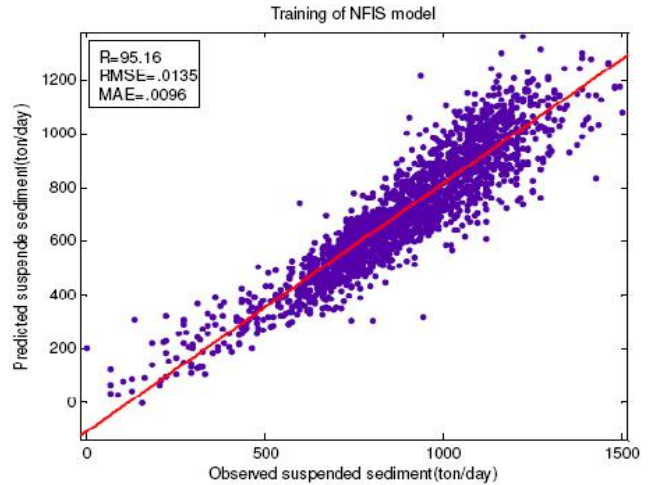


Fig 3. Comparison of ANN results with normalized observed values for training stage

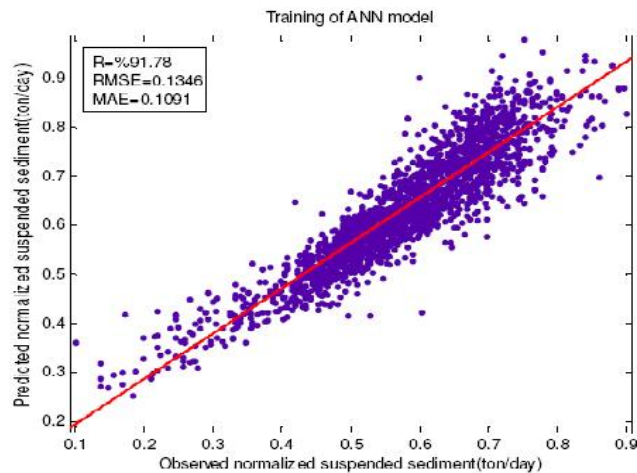


Fig 6. Comparison of ANFIS results with raw observed values for test stage

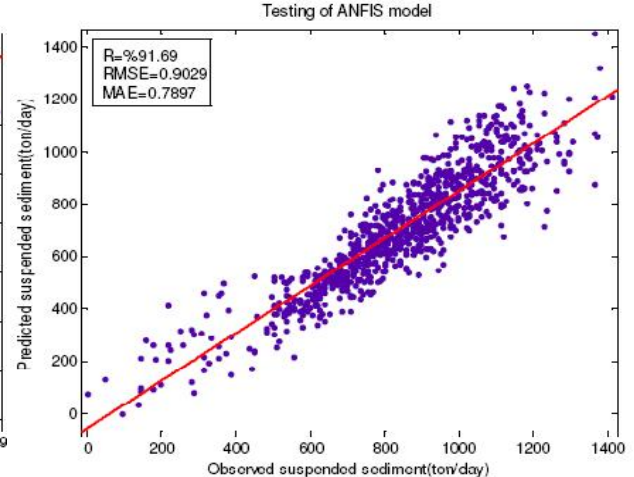


Fig 7. Comparison of ANFIS results with normalized observed values for training stage

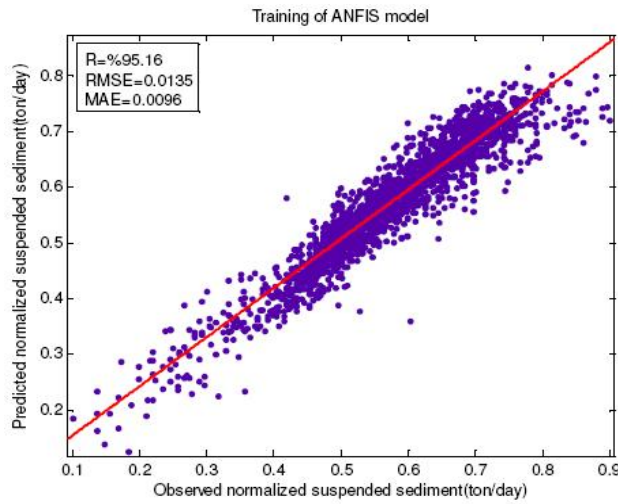


Fig 8. Comparison of ANFIS results with normalized observed values for test stage

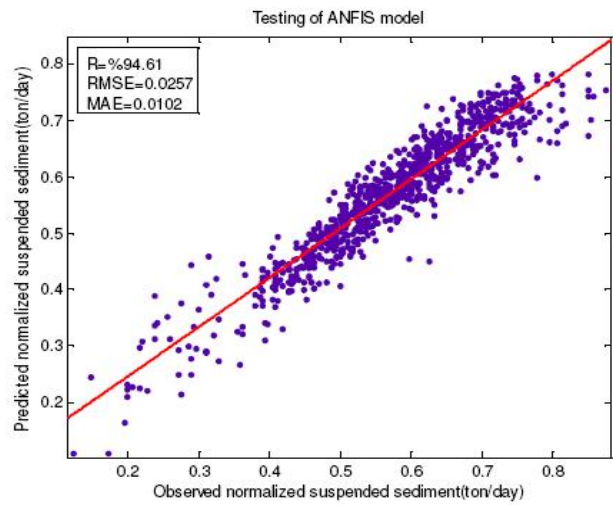


Fig. 9. Sediment rating curve for (a) raw and (b) normalized values.

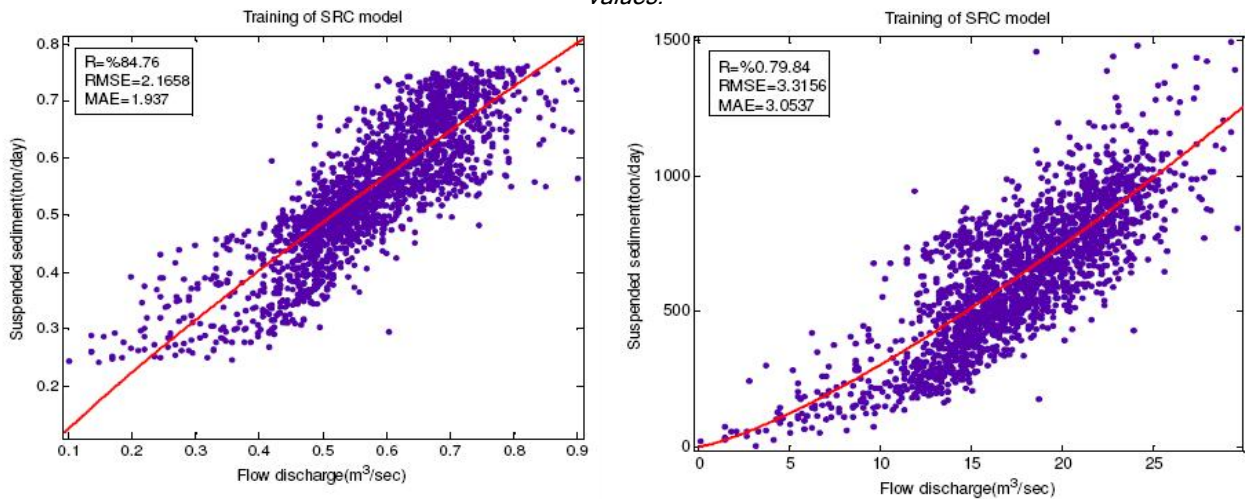


Fig. 10. Observed and predicted values by ANN and ANFIS for raw observed data

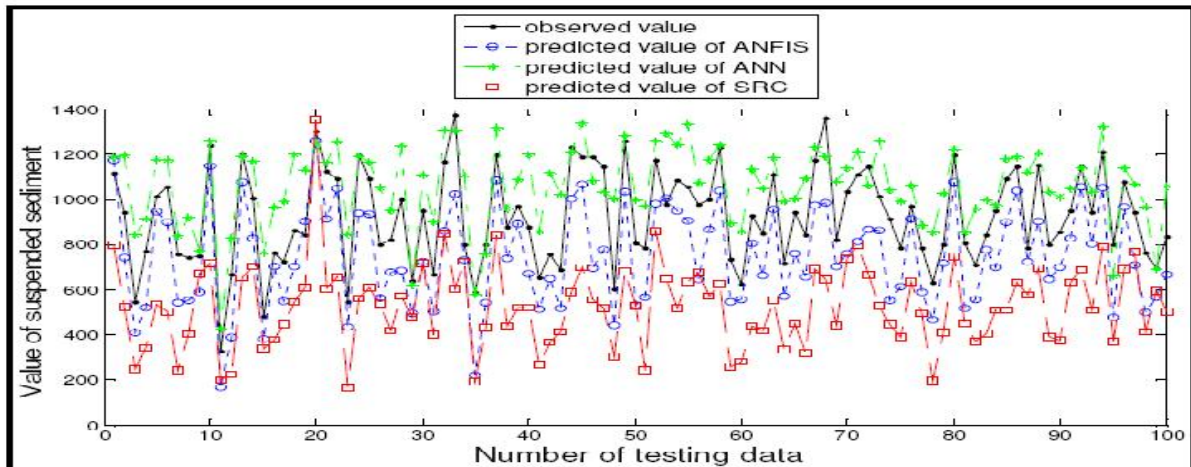
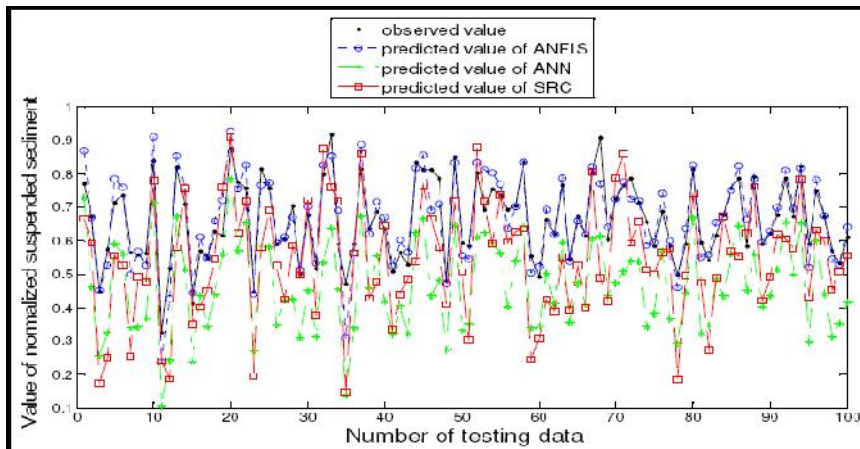


Fig. 11. Observed and predicted values by ANN and ANFIS for normalized observed data



Comparison the results achieved from ANN, ANFIS & SRC

The correlation coefficient and the criteria for error analysis for three utilized model corresponding to total raw and normalized data have been presented in Tables 5 and 6. The results of this tables show that neural-fuzzy inference system with 90.45% and 95.16% correlation coefficient is considered the best model for raw and observed normalized data, respectively and 91.69% and 94.61% correlation coefficient for raw and observed normalized data, respectively at the test stage. For this method can be led to a more accurate result in comparison with artificial neural network and sediment rating curve methods in non-linear modeling of load assessment of suspended sediments.

Also, to compare the results obtained from ANN, ANFIS and SRC models and the real amounts of sediment discharge for the first one hundred data test for raw and normalized data has been shown in Fig. 10 and 11. These figures show that the accuracy of neural-fuzzy inference system is much more than the accuracy of other methods. Therefore, using this method to estimate the load density of suspended sediment in the study area is recommended.

General Conclusion

Study of the obtained results show that the fuzzy method in identifying the sediment transition patterns in the vicinity under study has provided more acceptable results of artificial neural network methods and sediment rating curve. Fuzzy method is capable of identifying the time range of occurring sedimentary regimen change in the river.

In modeling with fuzzy inference system in comparison with sediment rating curve model, the relations among inputs (flow discharge, base and water temperature) and output (sediment density) in a smaller amount of time and with a relatively more accuracy. Another benefit of this method is that it is not sensitive to a small number of errors in statistical data which is resulted in better assessment of fuzzy model in comparison with sediment rating curve model. The results of this study show that fuzzy model can be used for adopting proper management methods catchment areas, especially the sedimentary issues and erosion.

Of the most important disadvantages of various sediments assessment is that there is no universal relationship which can be well responsible for all states and conditions.

Taking into consideration the above-mentioned method, due to the fact that different basins have different characteristics including slope of the area, amount and type of vegetation, soil material and are in different levels in terms of management, it is better the above mentioned factors are defined as an entry to neural models and neural-fuzzy models in addition to applied entries through collecting the statistics of various areas which have significant difference with each other in terms of the above-mentioned factors (Ahmadi, 2008).

References

1. Ahmadi H (2008) Using fuzzy inference system in assessing suspended sediment (case study: Talghan Watershed, 2nd year. No. 5.
2. Akbari GH (2011) Study of the quality of erosion and sedimentary transition using artificial neural networks and current methods of sediment" Case Study: 4th Conference on Iran's water resource management.
3. Bardstani, S (2010) Comparison of artificial neural network and sediment rating curve method to assess sediment load in Shapour River, 4th National Congress on Erosion and Sediment.
4. Dehgahni AA (2001) Comparison of the estimation of suspended load related to two sediment rating curve methods and artificial neural network (case study: Doug River in Golestan Province. *J. Agr. Sci. Natural Resources*. 16th edition, Journal 1-A.
5. Ebrahimi, Mohammadi, Sh (2010) Comparison of the efficiency of artificial neural networks, multivariate linear regression and sedimentary rating in assessing Haraz River suspended load", 4th National Conference on Erosion and Sediment.
6. Houshmand MM (2011) Estimating maximum depth of scour Bakh pier base using neural-fuzzy adoptive inference system. MA thesis on Electrical Engineering, University of Sistan and Baluchestan, Civil Faculty, Zahedan.
7. Jome Bidokhti (2011) Flow numerical modeling on Jami Launching and determining the maximum dynamic pressure using neural-fuzzy model", MA thesis on Civil Engineering, University of Sistan and Baluchestan, Zahedan.
8. Kordowani P (2006) Study of impacts of climate change (temperature and precipitation) on water changes of Karaj and Letian Docks, Applied Research Design of Water Resources Management Corporation in Iran.
9. Muhammad Zamani M (2007) Estimating the transition of river's sedimentary load using artificial neural network MA thesis on Civil Engineering, University of Tehran.
10. Nourani (2009) introducing a model to estimate sedimentary suspended load using fuzzy area and its comparison with classical methods and neural network", Case Study: Gharesou River located in Ardebil Province. 8th International Congress on Civil Engineering.