International Journal of Environment, Agriculture and Biotechnology
Vol-10, Issue-4; Jul-Aug, 2025
Peer-Reviewed International Journal
Journal Home Page Available: https://ijeab.com/
Journal DOI: 10.22161/ijeab

Suspended Sediment Amount Estimation in River Using Anfis and Conventional Methods

Engin Mert Özer*, Bestami Taşar, Fatih Üneş, Mustafa Demirci

Department of Civil Engineering, Iskenderun Technical University, Türkiye
*Corresponding Author

Received: 13 Jun 2025; Received in revised form: 12 Jul 2025; Accepted: 19 Jul 2025; Available online: 26 Jul 2025

***Abstract*— *The amount of sediment in rivers is of great importance in terms of determining the amount of pollution, river transport, dam life and other factors. In this study, the sediment matter in river was estimated using Multiple Linear Regression (MLR), Sediment Rating Curve (SRC) and Adaptive Neuro-Fuzzy Inference System (ANFIS) methods. For sediment estimation, models were developed using the flow and precipitation data between the years 2014-2022 as input parameters from a station on the Yahara River near Catonsville, United States. The models were evaluated using both graphical and statistical analyses. Performance comparison between prediction models was conducted using three key criteria: coefficient of correlation (R), root mean square error (RMSE), and mean absolute error (MAE). Based on these evaluation metrics, the ANFIS model demonstrated superior performance compared to other models in predicting suspended sediment amount in the river.*

***Keywords*— Prediction, Sediment, Regression, SRC, ANFIS**

I. INTRODUCTION

When constructing structures on streams, accurately estimating the amount of sediments carried in rivers is of great importance in terms of engineering and water structure planning. It is known that the amount of sediments in streams not only damages agricultural areas and stream transportation, but also shortens the life of facilities built on streams. Especially in dam reservoirs, sediments accumulate in water storage facilities over time, reducing the capacity of the reservoirs and causing the reservoirs to become unable to fulfill its function. It is also of great importance to evaluate the erosion or accumulation that may occur at the bases of other structures such as bridges and viaducts in the stream for flood control purposes and to estimate the amount of sediments transport in order to take appropriate measures. Although it is the most accurate method, the measurement method of determining sediments directly from the stream is costly and time-consuming. Although many observation stations evaluate the flow of the river, sediments are not measured because it is expensive, time-consuming and difficult. It becomes impossible to observe and measure sediments in rivers due to disasters such as floods and inundations. In recent years, studies on estimating sediments in rivers have increased rapidly. Kisi (1) estimated the amount of sediments in the river using Artificial Neural Networks (ANN), Adaptive Neural Fuzzy Inference System (ANFIS), Sediment Rating Curve (SRC) and Multiple Linear Regression (MLR) models using daily flow data from 2 stations in the USA (1). Comparative studies have consistently demonstrated that advanced modeling techniques, particularly ANFIS and ANN models, outperform classical methods in sediment estimation applications. Rajaee et al. (2) investigated daily sediment concentration estimation using four different approaches: Fuzzy Logic (FL), Artificial Neural Networks (ANN), Multiple Linear Regression (MLR), and Sediment Rating Curve (SRC) methods. Their analysis utilized data from Little Black River and Salt River stations in the USA, incorporating river discharge, time-shifted river discharge, and sediment concentration measurements as input variables for the FL and ANN models. The results revealed that ANN-based estimates significantly outperformed classical methodologies. Mirbagheri et al. (3) conducted a comprehensive performance evaluation of Sediment Rating Curve (SRC), Artificial Neural Networks (ANN), and fuzzy rule-based models for river sediment concentration estimation. Using the coefficient of determination as the primary evaluation metric, their comparative analysis demonstrated that fuzzy rule-based models produced superior results for sediment concentration prediction. Fırat and Güngör (4) focused on sediment estimation using monthly flow rate measurements from three monitoring stations along the Büyük Menderes River in Turkey. Their study compared three methodologies: Artificial Neural Networks (ANN), Multiple Linear Regression (MLR), and Fuzzy Logic (FL). The evaluation results indicated that the FL approach achieved the highest performance among the tested methods. Mianaei and Keshavarzi (5) analyzed daily flow rate and sediment discharge measurements from the Escanaba River, comparing Fuzzy Logic algorithms against Sediment Rating Curve (SRC) methods. Their findings showed that sediment estimations generated through Fuzzy Logic algorithms consistently outperformed alternative approaches. Vafakhah (6) developed sediment estimation models using precipitation and flow rate data from the Kojor basin near the Caspian Sea in Iran. The study analyzed 776 days of data spanning from 2007 to 2010. Performance evaluation using Mean Square Error (MSE), Mean Absolute Error (MAE), and Correlation Coefficient (R) criteria demonstrated that the FL model achieved superior results compared to the ANN model. Olyaie et al. (7) conducted an extensive comparison of four methodologies for river sediment estimation in the USA: Artificial Neural Networks (ANN), Adaptive Neural Fuzzy Inference System (ANFIS), Sediment Rating Curve (SRC), and Coupled Wavelet Artificial Neural Networks (WANN). Their analysis revealed that while Sediment Rating Curve and ANN methods exhibited the lowest performance, WANN and ANFIS approaches delivered the most accurate results. Kisi and Zounemat (8) investigated sediment content estimation at two monitoring stations on the Muddy River in the USA using daily flow rate and sediment concentration data. Their comparative study evaluated four models: Artificial Neural Networks (ANN), Adaptive Neural Fuzzy Inference System (ANFIS), Sediment Rating Curve (SRC), and Clustered Neuro-Fuzzy (CNF). The results indicated that the CNF method provided superior performance in river sediment estimation. Shameei and Kaedi (9) employed Linear Genetic Programming for sediment quantity estimation using measurements from Rio Valenciano and Quebrada Blanca stations in the USA. In addition, artificial intelligence methods are widely used in studies in many fields of hydrology (10-11-12-13-14-15)

In this study, three models were developed and compared for sediment estimation on the Yahara River near Catonsville, United States: Multiple Linear Regression (MLR), Sediment Rating Curve (SRC), and Adaptive Neuro-Fuzzy Inference System (ANFIS). The models utilized flow data from the monitoring station spanning 2014-2022 as input parameters for predicting sediment transport, with their performance evaluated through comparative analysis.

II. MATERIAL AND METHODS

STUDY AREA

This study used 2923 daily meteorological data from station with number 0709000 of the United States Geological Survey (USGS) located at 43°12'32 North latitude and 89°21'09 " longitude on the Yahara River near Catonsville, United States, between the years 2014-2022 (Figure.1). The Yahara River is approximately 62 miles (100 km) long (including the distance between the lakes in between) and drains an area of 536 miles^2 (1,390 km^2). It is part of the Mississippi River watershed via the Rock River. The Yahara River connects the lakes around which the city of Madison is built. Figure 2 shows that Changes in Daily Average Sediment (mg/L) Amount of the Yahara River in the Study Station Between 2015 and 2022

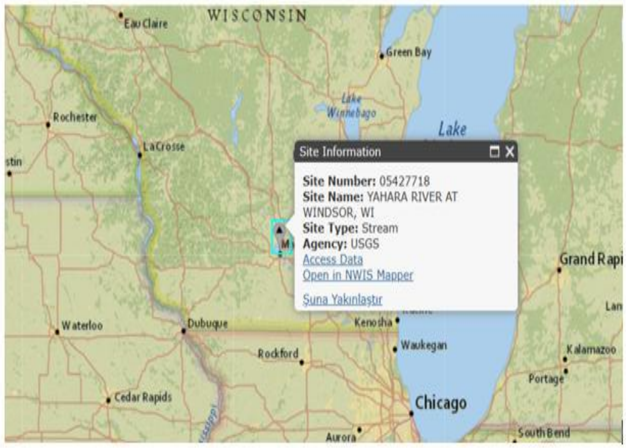

Fig.1. Study area location and Yahara river image (USGS)

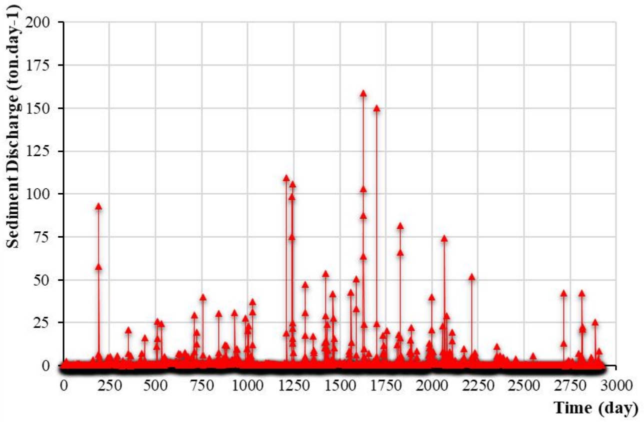

Fig.2. Changes in Daily Average Sediment ton/day) Amount of the Yahara River in the Study Station Between 2015 and 2022

III. METHOD

Multiple Linear Regression (MLR)

Multiple linear regression analysis is a statistical method used to model relationships between variables based on cause-and-effect principles. This technique allows researchers to examine how two or more variables interact and influence each other within a given system. The classification of regression analysis depends on the number of independent variables employed: Simple Regression: Uses a single independent variable to predict the dependent variable

Multiple Regression: Incorporates two or more independent variables to estimate the dependent variable. In engineering applications, variables typically exhibit interdependence rather than operating independently. This interconnected nature makes Multiple Linear Regression (MLR) particularly valuable for analyzing complex systems where multiple factors simultaneously influence outcomes.

The MLR method quantifies the impact of independent variables on dependent variables through regression coefficients within the mathematical equation. These coefficients represent the magnitude and direction of influence that each independent variable exerts on the dependent variable.

The mathematical formulation of Multiple Linear Regression is presented in Equation 1 below.

$$Y_i = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n) + \varepsilon_i \quad (1)$$

This equation contains linear expressions. In this equation, X_i (i = 1, …, n) represents the independent variables, Y_i represents the dependent variable, β represents the regression coefficients and ε represents the error.

Sediment Rating Curve (SRC)

Sediment Rating Curve (SRC) shows the relationship between the amount of sediment matter measured at any control section of the stream and the flow rate. If the sediment discharge is represented by S and the flow rate is represented by Q, the relationship between these two is shown in equations (2) and (3)

$$S=aQ^b \quad (2)$$

It is expressed with a functional relationship as follows. When we convert this relationship to logarithmic base;

$$\log S=a+b\log Q \quad (3)$$

The expression is obtained. In this equation, a and b obtained by establishing linear regression between logQ and logS show the coefficients depending on the river flow characteristics (16-17-18) The a and b values in the equation are found from the equations created according to the relationship between S and Q. With these curves used in sediment matter estimations, average monthly or annual sediment matter amounts can be estimated. In a large study examining the S-Q relationship given in the equation (19), it was shown that there is no uniform relationship. In some streams, the S-Q relationship follows a two-armed change. Due to hydrological reasons, the amount of sediment matter at different times of the flow may be different for the same flow rate value. Sometimes, accurate sediment matter estimation cannot be made using these curves and is insufficient. For the application area in the presented study, sediment rating curve for the training data (SRC) is shown in Fig. 3.

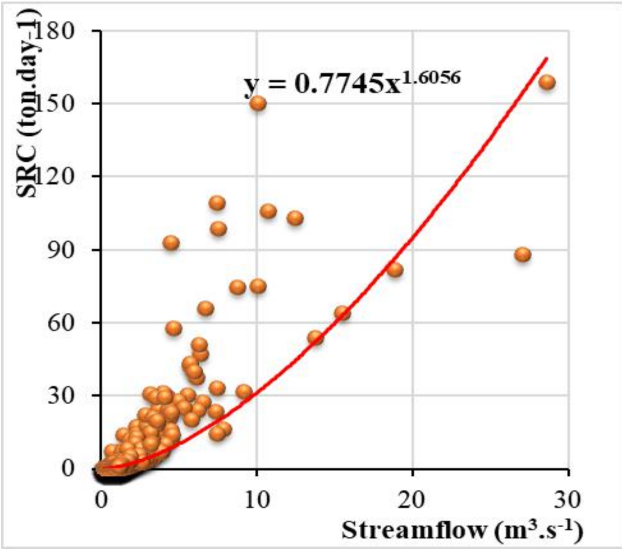

Fig.3. Scatter plot of the SRC model in the training phase

Adaptive Neural Fuzzy Inference System (ANFIS)

The Adaptive Network-Based Fuzzy Inference System (ANFIS) is a hybrid artificial intelligence framework that integrates artificial neural networks with fuzzy inference systems. Originally developed by Jang in the early 1990s, this methodology has gained widespread recognition for its capability to model nonlinear functions and estimate chaotic time series effectively [20-21].

ANFIS functions through an interconnected network of nodes, with each node serving as an individual processing unit. This structural design allows the system to harness the advantages of both artificial neural networks and fuzzy logic inference techniques, creating a sophisticated hybrid learning algorithm that combines the best aspects of both approaches.

Fuzzy inference systems are built upon two fundamental approaches: Mamdani and Assilian approach and Takagi and Sugeno approach

These methodologies provide the theoretical foundation for ANFIS implementation and determine how fuzzy rules are constructed and applied. ANFIS implementation requires structured datasets containing both input and output variables. The system employs an error-minimization training principle to optimize the membership functions of fuzzy sets. Through this iterative process, ANFIS automatically generates fuzzy rules for the Fuzzy Inference System (FIS), ensuring optimal performance for the given problem domain.

The ANFIS architecture, illustrated in Figure 3, consists of several key components: x, y, z, t: Independent variables that serve as system inputs a_1, a_2, b_1, b_2, c_1, c_2, d_1, d_2: Input parameters that define the system's operational characteristics

∏ (pi): Membership functions that determine the degree of membership for input variables

N: Fuzzy rules that govern the system's decision-making process

w_i: Parameter weights that influence the relative importance of different system components

The hybrid nature of ANFIS provides several advantages:

Combines the learning capabilities of neural networks with the interpretability of fuzzy logic Automatically optimizes membership functions through training

Generates fuzzy rules without manual intervention

Effectively handles nonlinear and complex relationships in data

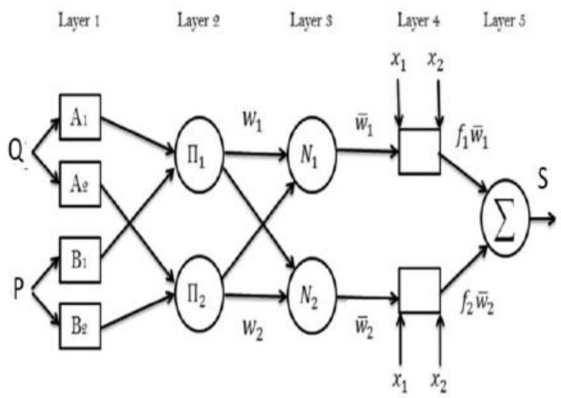

Fig.4 ANFIS structure for two inputs and one output

In the figure 4;

Layer 1: Membership Function Layer

The first layer focuses on membership function selection and determination of linguistic variable membership levels. In this study's ANFIS model, each independent variable utilizes two membership functions to define the fuzzy sets and their corresponding membership degrees.

Layer 2: Rule Layer

The second layer contains fixed nodes denoted by the symbol "∏" (product). These nodes calculate the products of outputs from the first layer, representing the resulting fuzzy rules. Each node in this layer corresponds to a specific fuzzy rule activation strength.

Layer 3: Normalization Layer

This layer consists of fixed nodes indicated by the symbol "N" (normalization). The ANFIS system normalizes the values within the network structure at this stage, ensuring that the fuzzy rule firing strengths are properly scaled. The normalized values serve as the layer's output.

Layer 4: Defuzzification Layer

All nodes in the fourth layer are adaptive nodes that perform defuzzification operations. The normalized weight values (w) from the third layer are multiplied by first-degree polynomial equations. The output of this layer is expressed as "$w_1 \times f_1$", where w_1 represents the normalized firing strength and f_1 represents the polynomial function.

Layer 5: Output Layer

The final layer contains a single fixed node that performs summation operations. This node calculates the total result of all preceding operations, represented by the symbol "Σ" (summation), producing the final ANFIS output.

IV. RESULTS

A total of 2,923 data points from an 8-year period (2014-2022) were collected from the studied station. The dataset was divided into training and testing sets, with 75% of the data (2,192 measurements) allocated for model training and 25% (731 measurements) reserved for testing.

The model performance was evaluated using three statistical metrics: the correlation coefficient (R), root mean square error (RMSE), and mean absolute error (MSE), as defined in Equations 4-6.

$$R^2 = \left(\frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{\left(n(\Sigma x^2) - (\Sigma x)^2\right)}\sqrt{\left(n(\Sigma y^2) - (\Sigma y)^2\right)}} \right) \quad (4)$$

$$RMSE = \left((1/N) \textstyle\sum_{i=1}^{N} |S_{measurment} - S_{Predict}|^2 \right)^{1/2} \quad (5)$$

$$MAE = \frac{1}{N} \left(\textstyle\sum_{i=1}^{N} |S_{measurment} - S_{Predict}| \right) \quad (6)$$

Table 1. Error rates and correlation changes of the models

Model	Model Inputs	MAE (ton/d)	RMSE (ton/d)	R
MLR	Q(t), P(t)	1.38	1.758	0.787
SRC	Q(t), P(t)	0.33	1.296	0.829
ANFIS	Q(t), P(t)	0.34	0.846	0.888

MAE: Mean absolute error, **RMSE:** Root mean square error, **R**: Correlation coefficient Q(t,): daily streamflow (m^3/s), P(t): daily precipitation (mm)

When Table 1 is examined, the ANFIS model showed the best performance with the highest coefficient of correlation and the lowest error rates (MAE: 0.34, RMSE 0.846, R: 0.788). The MLR model showed the worst performance with the low coefficient of determination and high error rates (MAE: 1.38, RMSE 1.758, R: 0.620).

Multiple Linear Regression (MLR) Results

In this study, the daily river flow (Q), Daily precipitation (P) were used in the MLR model applications to estimate the sediment concentrations (S). The MLR equation used in the sediments estimation is shown in equation (7). For the MLR results, the scatter and distribution graphs of the training data and the scatter and distribution graphs of the test data are shown in Figure 5 and Figure 6, respectively.

$$S = 15{,}57 + 0{,}91 * B_t + 5{,}96 * Q_t \quad (7)$$

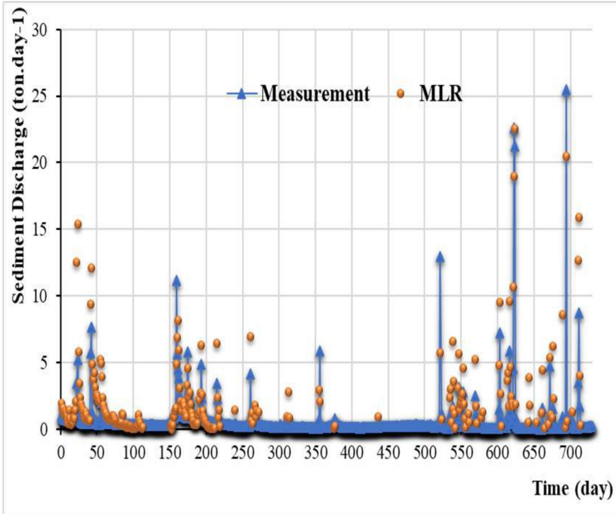

Fig.5. Distribution plot of the MLR model in the testing phase

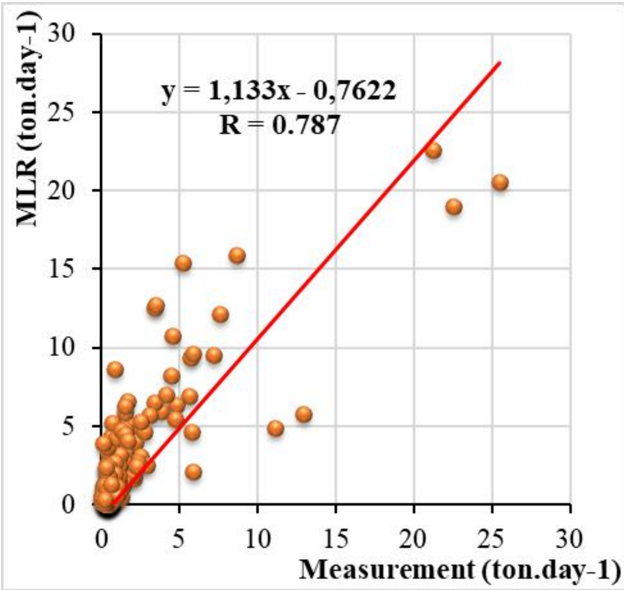

Fig.6. Scatter plot of the MLR model in the testing phase

When the distribution graph drawn for the CDR is examined, a large degree of compatibility is seen between the estimated values and the measured values. The compatible relationship between the estimated and measured values is also seen with the high determination value (R=0.787) in the scatter graph.

Sediment Rating Curve (SRC) Results

For the sediment rating curve results, the scatter and distribution graphs of the training data and the scatter and distribution graphs of the test data are shown in Figure 7 and Figure 8, respectively.

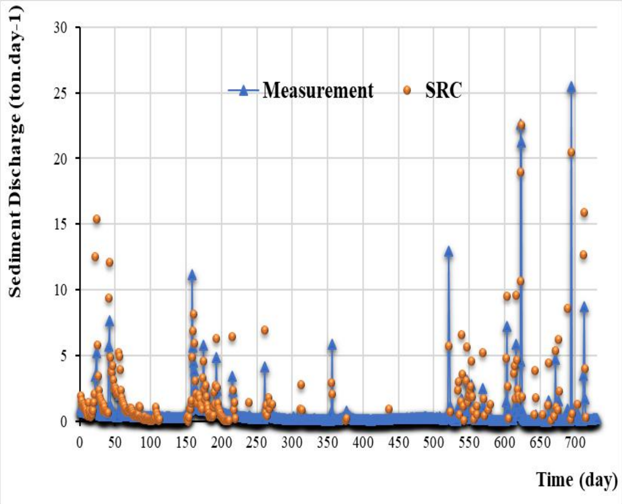

Fig.7. Distribution plot of the SRC model in the testing phase

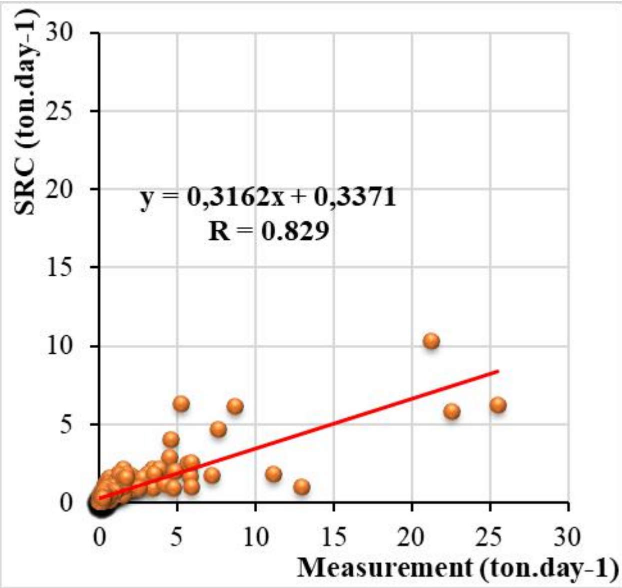

Fig.8. Scatter plot of the SRC model in the testing phase

When the distribution graphs drawn for the Sediment rating curve are examined, the predicted values do not seem to be very compatible with the measurement values. When the shape scatter graph for the test data is examined, a low correlation coefficient o is observed (R=0.829).

ANFIS Results

For ANFIS results, scatter and distribution graphs of training data and test data are shown in Figure 9 and Figure 10, respectively

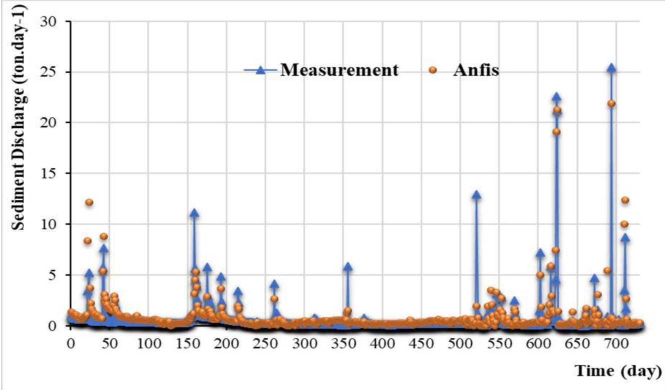

Fig.9. Distribution plot of the ANFIS model in the testing phase

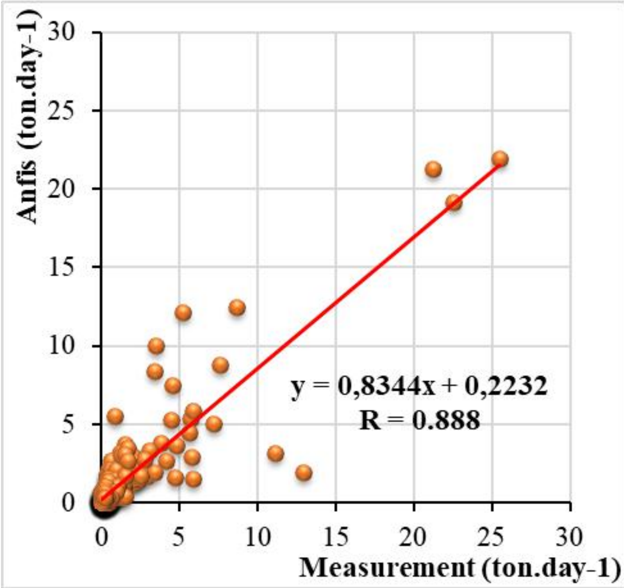

Fig.10. Scatter plot of the ANFIS model in the testing phase

When the distribution and scatter graphs of the test data for ANFIS are examined, it is seen that the predicted values and the measured values are compatible. When the shape distribution and scatter graphs for the test data are examined, it is seen that the coefficient of determination of the training data is higher than the correlation coefficient of the test data. (R=0.888)

V. CONCLUSION

A total of 2923 data from 8 years between 2014-2022 were used at the studied station. In the study, 75% of all data were used for training, 25% for training, 2192 data, and 731 measurement data for testing. The results found from the equations for the ANFIS, SRC, MLR prediction model of these data were calculated.

In the study, it was seen that the, ANFIS, SRC and MLR models can be used in daily sediment estimation on a regional basis. However, when the error calculations and determination values made within the scope of this study are taken into consideration, it was understood that the ANFIS method gives more accurate estimation results than the SRC and MLR models. It was understood that all three methods used in the study were at an acceptable level in terms of estimation results.

ACKNOWLEDGEMENTS

The data used in this study were downloaded from the web server of the USGS. The authors wish to thank the staff of the USGS who are associated with data observation, processing, and management of USGS Web sites.

REFERENCES

[1] Kisi, Ö. (2005). Daily river flow forecasting using artificial neural networks and auto-regressive models. Turkish Journal of Engineering and Environmental Sciences, 29(1), 9-20.

[2] Rajaee, T., Mirbagheri, S. A., Zounemat-Kermani, M. ve Nourani, V. (2009). Daily suspended sediment concentration simulation using ANN and neuro-fuzzy models. Science of the total environment, 407(17), 4916-4927.

[3] Mirbagheri, S. A., Nourani, V., Rajaee, T., & Alikhani, A. (2010). Neuro-fuzzy models employing wavelet analysis for suspended sediment concentration prediction in rivers. Hydrological Sciences Journal–Journal des Sciences Hydrologiques, 55(7), 1175-1189.

[4] Firat, M., & Güngör, M. (2010). Monthly total sediment forecasting using adaptive neuro fuzzy inference system. Stochastic Environmental Research and Risk Assessment, 24(2), 259-270.

[5] Mianaei, S. J., & Keshavarzi, A. R. (2010). Prediction of riverine suspended sediment discharge using fuzzy logic algorithms, and some implications for estuarine settings. Geo-Marine Letters, 30(1), 35-45.

[6] Vafakhah, M. (2013). Comparison of cokriging and adaptive neuro-fuzzy inference system models for suspended sediment load forecasting. Arabian Journal of Geosciences, 6(8), 3003-3018.

[7] Olyaie, E., Banejad, H., Chau, K. W., & Melesse, A. M. (2015). A comparison of various artificial intelligence approaches performance for estimating suspended sediment load of river systems: a case study in United States. Environmental Monitoring and Assessment, 187(4), 1-22.

[8] Kisi, O. & Zounemat-Kermani, M. (2016). Suspended sediment modeling using neuro-fuzzy embedded fuzzy c-means clustering technique. Water resources management, 30 (11), 3979-3994.

[9] Shamaei, E., & Kaedi, M. (2016). Suspended sediment concentration estimation by stacking the genetic programming and neuro-fuzzy predictions. Applied Soft Computing, 45, 187-196.

[10] Ünes, F. (2010). Dam reservoir level modeling by neural network approach: A case study Neural Network World. 20 (4), 461- 23.

[11] Üneş, F., Demirci, M. (2015). Generalized Regression Neural Networks for Reservoir Level. Modeling Technology and Applied Science (ICETA), Berlin, Germany.

[12] Kaya, YZ., Mamak, M., Üneş, F. (2016). Evapotranspiration prediction using M5T data mining method. International Journal of Advanced Engineering Research and Science, 3-12.

[13] Üneş, F., Gumuscan, F., Demirci, M. (2017). Prediction of Dam Reservoir Volume Fluctuations Using Adaptive Neuro Fuzzy Approach. European Journal of Engineering and Natural Sciences, 2 (Issue 1), pp. 144.

[14] Üneş, F., Demirci, M., Tasar, B., Kaya, YZ., Varçin, H. (2019). Modeling of dam reservoir volume using generalized regression neural network, support vector machines and M5 decision tree models. Applied Ecology and Environmental Research 17 (3), 7043-7055 24-2019

[15] M Yilmaz, F Tosunoğlu, NH Kaplan, F Üneş, YS Hanay Modeling Earth Systems and Environment 8 (4) Predicting monthly streamflow using artificial neural networks and wavelet neural networks models, 5547-556323-2022

[16] Vanscikle, J., Beschta, R. L., (1983). Supply-based models of suspended sediment transport in streams. *Water Resources Research*, 19, 3.

[17] Demiröz, E.. (1989). Sediment Sampling Activities in Turkey. *Fourth International Syposium on River Sedimentation*, Beijing, China, Nov. 1-5.

[18] Salas, J. D., Shin, H. S., (1999). Uncertainty analysis of reservoir sedimentation. *Journal of Hydraulic Engineering*, ASCE, 125, 4, April.

[19] Williams, J. R., (1978). A sediment graph model based on an instantaneus unit sediment graph. *Water Reseurces Research*, 14, 659-664.

[20] Jang JSR. Fuzzy modeling using generalized neural networks and kalman filter algorithm. Proceedings of the 9th National Conference on Artificial Intelligence, Anaheim, CA, USA, 1991 July 14–19. p.762–767.

[21] Jang JSR. ANFIS: Adaptive-Network-Based Fuzzy Inference System. IEEE Transactions on Systems, Man and Cybernetics 1993;23(3):665-685.