Research Article **GJAI (2019) 1:2**

Global Journal of Artificial Intelligence
(DOI:10.28933/GJAI)

Neuro-Fuzzy Approach to River Sediment Yield Prediction

Balogun O. [1], Akinboro S. A[2], Ogunseye, A. A[3]

[1,2] College of Information and Communication Technology, Bells University of Technology, Otta, Nigeria; [3]Department of Electronic & Electrical Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

ABSTRACT

This work is motivated by the critical role that sediment yield prediction plays in preventing natural and economic disasters. Methods based on regression techniques have been used to solve the problem but they are generally inadequate in predicting river sediment yield because of the inherent complexity of the problem. This work uses the Adaptive Neuro-Fuzzy Inference System (ANFIS) to solve the problem. The ANFIS model accepts four input data namely temperature, rainfall, water stage and water discharge and gives on output data that represents the sediment yield. The ANFIS model was developed and simulated with MATLAB 7.0 using the Levenberg-Marquardt optimization method and trained with a maximum of 1500 epochs at a learning rate of 0.5. the results obtained was compared with the ones obtained with the Artificial Neural Network (ANN) model and it was found that the ANFIS model performs better than the ANN model.

Keywords: ANFIS, ANN, Neuro-Fuzzy, Backpropagation algorithm, Sediment yield.

***Correspondence to Author:**
Ogunseye, A. A
Department of Electronic & Electrical Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria

How to cite this article:
Balogun O., Akinboro S. A, Ogunseye, A. A. Neuro-Fuzzy Approach to River Sediment Yield Prediction. Global Journal of Artificial Intelligence, 2019; 1:2.

eSciPub LLC, Houston, TX USA.
Website: http://escipub.com/

1. Introduction

Sediments are objects that exists within an aquatic environment. They supply nutrients and provide ecological balance in the environment they exist. River sediments include clay particles, silt, sand, gravel, rock, etc. They originate from river bed load transport, beach bank erosion, and land runoff and are naturally sorted by size through prevalent hydrodynamic conditions. In general, fast moving water will contain coarse-grained sediments and slow-moving water contain fine grained sediments (Pollution issues, 2016).

The assessment of the volume of sediments being transported by a river is important to the quality of water. Water quality refers to the physical, chemical, biological, and radiological characteristics of water. It is a measure of the condition of water relative to the requirements of one or more biotic species and/or to any human need or purpose (Johnson et al, 1997). Some other importance of the assessment of volume of sediment transportation include estimation of sediment transport in rivers, design of dams, reservoirs andchannels, environmental impact assessment, and determination of the efficacy of watershed management and other catchment treatment (Lohani et al, 2009).

Sediment yield is therefore the total volume of sediments/sand particles that leaves a drainage basin in a given period. Some variables like temperature and rainfall that determine the observed sediment yield are all controlled by nature and human activities. They are, in most cases, changing and in some cases, similar to certain degrees thereby making the underlying physical laws unknown or not precisely known. This makes it rather difficult to model the phenomenon adequately.

For instance from data gathered, it was shown that the temperature obtained was not sequential in readings. There were periods that it was similar to those of previous years while its overall effect on sediment yield was minimal and years where it was a far value from what was expected. In the course of the sediment yield data gathering, it was observed that at most times, none of the variables could be truly said to be linear in nature. Ogundoyin I.K. (2009) posits that the rating relationships based on regression techniques are generally not adequate in view of the inherent complexity with regard to sediment yield. The non-linear nature of sediment yield has led to the development of intelligent machines that can model the characteristic variables that determine sediment yield. This aspect of Information and Communication Technology (ICT) is known as soft computing.

In this research, a Neuro-Fuzzy model is used to model sediment yield. The Fuzzy aspect of the model classifies data as a member of classes while ANN on the other hand solves complex problems by decomposing them into smaller simpler elements and inputting the data into a layered network. A suitable algorithm is then introduced to train the network. The interaction of nodes (which may contain another network) contributes to the overall behavior of the network. The combination of these two methods has allowed the strengths of the individual models to be harnessed to develop a more efficient model.

2. Related Work

Rughuwnshi et al (2006) developed Artificial Neural Network (ANN) models to predict both runoff and sediment yield of a small agricultural watershed on a daily and weekly basis. A total of five models were developed for predicting runoff and sediment yield, of which three models were based on a daily interval and the other two were based on a weekly interval. All five models were developed both with one and two hidden layers. Each model was developed with five different network architectures by selecting a different number of hidden neurons. The ANN models with a double hidden layer were observed to be better than those with single hidden layer. Further, the ANN model prediction performance improved with increased number of hidden neurons and input variables. As a result, models that consider both rainfall and temperature as input performed better than those considering rainfall alone as input.

Nagy et al (2002) used an artificial neural model to estimate the natural sediment discharge in rivers in terms of sediment concentration. This was achieved by training the network to extrapolate several natural streams of data collected from reliable sources. The selection of water and sediment variables used in the model was based on the prior knowledge of the conventional analyses, based on the dynamic laws of flow and sediment. Firat (2008) compared various Artificial Intelligence techniques for both Seyhan and Cine rivers. For river flow forecasting using the Seyhan River, the forecasting models were established using combinations of antecedent daily river flow records. On the other hand, for the Cine River, daily river flow and rainfall records were used in the input layer. For both stations, the data sets are divided into three subsets, training, testing, and verification data set. The results demonstrated that ANFIS model is superior to the GRNN (General Regression Neural Network) and FFNN (Feedforward Neural Network) forecasting models, and ANFIS can be successfully applied and provides high accuracy and reliability for daily river flow forecasting.

Pankaj Kumar (2011) used Adaptive Neuro-Fuzzy inference system to forecast crop yield by using it to map out the relationship between climatic data and its effects on crop yield. ANFIS was used to predict rice yield based on time series data of 27 years. (1981-1982, 2007-08).

ANFIS has also been used to predict stock market momentum. In their paper Agrawal et al (2010) devised an innovative approach for indicating stock market decisions that the investor should take for minimizing the risk. The system used Adaptive Neuro-Fuzzy Inference System (ANFIS) for taking decisions based on the values of technical indicators. . They used weighted moving averages, divergence and RSI (Relative Strength Index).

Bisht and Jangid (2011) made use of the Neuro-Fuzzy inference system to model river stage discharge. In their paper, river stage discharge models using Adaptive Neuro- Fuzzy Inference System and Linear Multiple Regression (LMR) methods were developed. They investigated the best model to forecast river discharge. Ten ANFIS models were developed out of which the best five ANFIS models were selected. The developed models were trained, tested and validated on the data of Godavari River at Rajahmundry, Dhawalaishwaram Barrage site in Andhra Pradesh. Comparing observed data and the estimated data through developed ANFIS models, it was proven that the developed ANFIS models predicted better results than the traditional models, like LMR.

Ozgur Kisi et al (2008) in their paper investigated the accuracy of an adaptive Neuro-Fuzzy computing technique in suspended sediment estimation. They collected the monthly streamflow and suspended sediment data from two stations, Kuylus and Salur Koprusu, in Kizilirmak Basin in Turkey. The data were used as case studies. The estimation results obtained by using the Neuro-Fuzzy technique were tested and compared with those of the artificial neural networks and sediment rating curves. Root mean squared errors, mean absolute errors and correlation coefficient statistics were used as comparing criteria for the evaluation of the models' performances. The comparison results reveal that the Neuro-Fuzzy models can be employed successfully in monthly suspended sediment estimation.

All scenarios of comparison suggest that ANFIS model is better suited and a more efficient model to predict nonlinear variable conditions.

3. The ANFIS model

The ANFIS model used in this work, shown in Figure 1, is a multilayered network made up of one input layer, three hidden layer and one output layer. The input layer takes in a maximum of four neurons and a minimum of two neurons. These neurons represent the temperature, rainfall, water level and water discharge variables. The neuron combination can be in any order but the chosen order must remain consistent throughout the simulation process. These input neurons are the data set used in training and validating the sediment yield prediction.

The second layer represent the layer where the Fuzzy rules are applied to the input data.

The third layer of the ANFIS model represents the layer where normalization takes place after applying the Fuzzy rule. This is achieved using the result which is the output from the second layer as input to the third layer.

The normalized data is sent to the fourth layer where defuzzification takes place and then to the output layer. All of these processes occur on all of the input neurons and the output from each of the input Neuron is summed up to give the predicted sediment yield

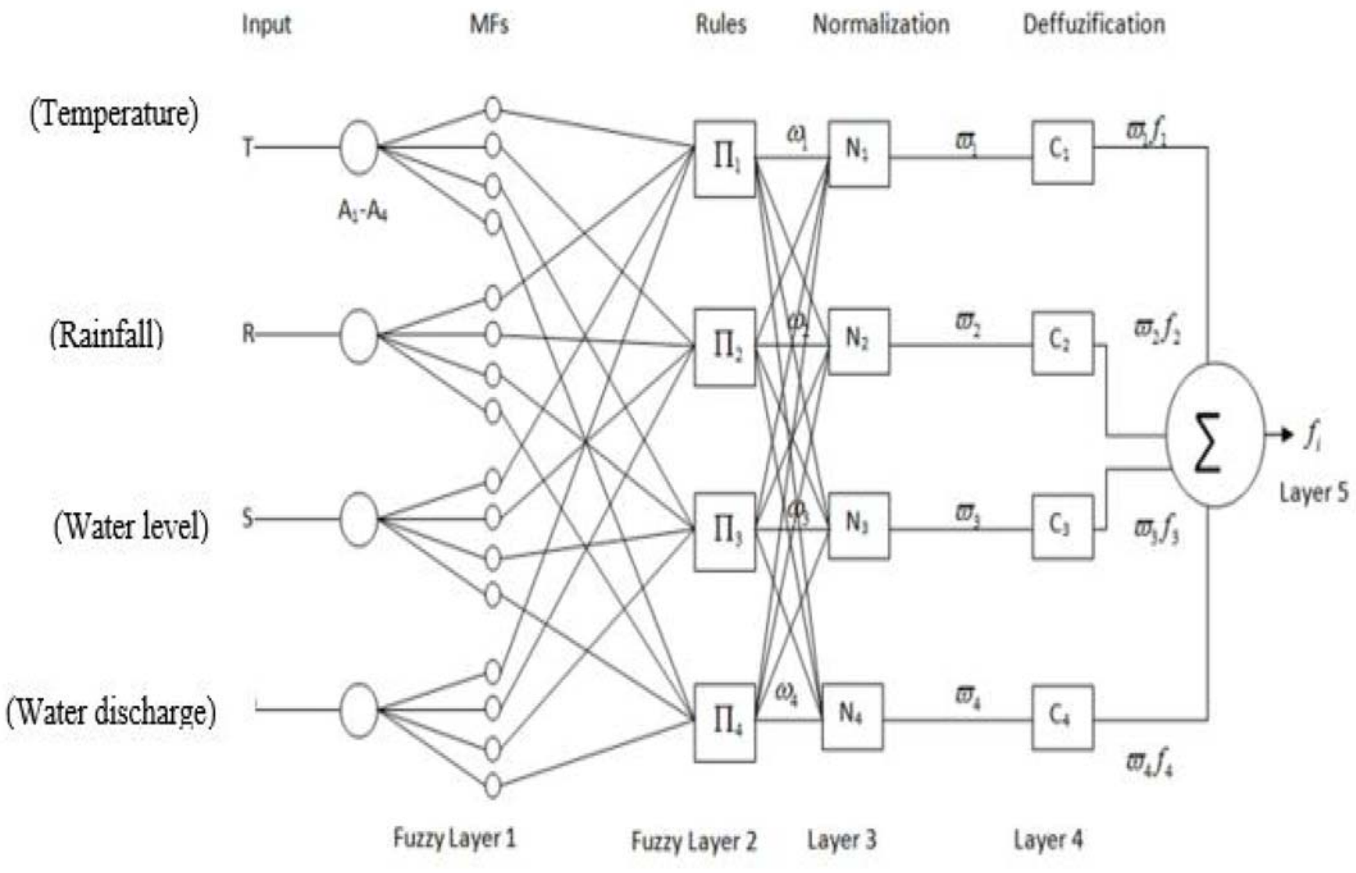

Figure 1: ANFIS Model

3.1 ANFIS FuzzyRules

The following Fuzzy logic rules are used in implementing the ANFIS system:

Rule1

If A1 is T and B1 is R and C1 is S and D1 is Q then *Y=f1=*$\overline{w}$*1f1*

Rule 2

If A2 is T and B2 is R and C2 is S and D2 is Q then *Y=f2=*$\overline{w}$*2f2*

Rule 3

If A3 is T and B3 is R and C3 is S and D3 is Q then *Y=f3=*$\overline{w}$*3f3*

Rule 4

If A4 is T and B4 is R and C4 is S and D4 is Q then *Y=f4=*$\overline{w}$*4f4*

3.2 Algorithm for Sediment yield Computation

Let the inputs be represented as follows:

Water temperature=T

Rainfall=R

Water level/Stage=S

Water Discharge=D

Sediment yield=*f*

i. Input values for T, R, S and D

ii. Calculate thenormalizedweight$\overline{w}$ −

*wi=μAi(T)*μAi(R)*μAi(S)*μAi(D)*......................................eq(1)

where μ is the membership function

$\overline{w}=\frac{w_i}{\sum_{i=1}^{4} w_i}$...eq (2)

for*i*=1…4

iii. To calculate the sediment yield

f= $\frac{T*\overline{w}_t+R*\overline{w}_r+S*\overline{w}_s+D*\overline{w}_d}{\overline{w}_t+\overline{w}_r+\overline{w}_s+\overline{w}_d}$.......................................eq(3)

where w_t = normalized weight for temperature w_r = normalized weight for rainfall

w_s = normalized weight for water level

w_d= normalized weight for water discharge

3.3 The Data

The data used for this research was obtained from the Ogun-Osun River that rises in the Iganran Hills (503 m) east of Shaki in the north-western part of Oyo state now Oyo and Osun states. The river flows southwards for approximately 410 km before discharging into the Lagos lagoon. Its main tributary is the Oyan which rises to the west of Shaki and incorporates the Ofiki River. The totalareaisabout23,700sq.km, occupying parts of Oyo, Osun, Ogun, and Lagos States in Nigeria. The data collected spans a period of four years and consist of a 385 row by 5 column data.

The data represents the values for five entities which are Water temperature (°c), rainfall (mm), Water level (mm), Water Discharge (m^3/s) and sediment (mg/l). The data is divided into two categories. A category is used in training the network and another category is used in testing the network. A total of 216 data sets was used for training the network while the rest was used for testing the network. The Neuro-Fuzzy algorithm was fed with three different set of inputs for each observed sediment level. The inputs are:

i. Water temperature andrainfall,
ii. Water temperature, rainfall and Waterlevel,
iii. Water temperature, rainfall, Water level and Waterdischarge.

3.4 Parameters for Performance Evaluation

Root Mean Square Error: The root mean square error is a frequently used measure of the differences between values (sample and population values) predicted by a model or an estimator and the values actually observed.

Root mean square error serves to aggregate the magnitudes of the error in predictions for various times into a single measure of predictive power. It's a good measure of accuracy and is majorly used to compare forecasting errors of different models for a particular variable. The performance of the models is accessed by satisfying the defined objective function of the model. The applicability and suitability of these models to hydrologic problem are measured using statistical criteria as follow:

$$RMSE=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(s_{oi}-s_{ci})^2} \text{ ....eq (4)}$$

where

RMSE is the root mean square error,

N is the number of observations, and

S_{oi} and S_{ci} are the ordinates of the observed and computed sediment yield respectively.

Correlation coefficient: Correlation coefficient is a coefficient that illustrates a quantitative measure of some type of correlation and dependence, meaning statistical relationships between two or more random variables or observed data values. It is usually between the ranges of +1 to -

1. It is used to evaluate the degree of relationship between two variables. When the correlated coefficient value carries a positive sign, it implies that as one variable increase/decrease, the other variables increase/decrease. The higher the value is towards any of the extremes (-1 or +1), the stronger the relationship between the variables involved.

$$cc=\frac{\sum_{i=1}^{N}(s_{oi}-s_o)\bar{X}(s_{ci}-\bar{s}_c)}{\sum_{i=1}^{N}(s_{oi}-\bar{s})^2 X \sum_{i=1}^{N}(s_{ci}-\bar{s}_c)^2} \text{........................ (5)}$$

CC is the correlation

$\bar{s}_0$ and $\bar{s}_c$ are the mean values of the observed computed values of the sediment yield coefficient of determination.

$$R^2 = (CC)^2 \text{...............................(6)}$$

where R is the mean square error

The RMSE objective is to be minimized. It helps get the difference or error between the output generated by the network model and the desired result. The network was presented with training examples, which consist of a pattern of activities that allow the system to predict to a high level of correctness the value we are lookingfor.

4. Results and Discussion

The performances of the fuzzy ANFIS based sediment yield developed for this study as compared to ANN in terms of RMSE and R^2 during the training and validation periods are shown in tables 1 and 2 respectively. For this study, four input combinations were used as inputs to the model during training and validation periods. The inputs are also shown in Tables 1 and 2. Figure 2 shows the graphical representation of the various data combination during the test and validation stages of the simulation.

During the simulation, the ANFIS model fitted the observed sediment data better than the ANN (simulated for the purpose of this research)in the validation stages except when fed with three variables where the ANN model had a RMSE of 23.2996 as compared to that of ANFIS model which had a RMSE of 23.3443. When four input data sets are applied to the validated ANFIS model, the Correlation Coefficient of 0.3860 was obtained, compared to a value of 0.378074 for ANN which implies the ANFIS model correlation coefficient tends greater to +1 i.e. there is a greater dependence between the observeddata.

Table 1: The comparison of ANN and Neuro-Fuzzy models in sediment Estimation (Training set): Oyan gauging station of Ogun-Osun RiverBasin

Model Input	Neuro-Fuzzy performance		ANN Performance	
	RMSE	R^2	RMSE	R^2
TR	8.666	0.6491	8.849	0.0408
TRL	8.769	0.5771	8.669	0.0279
TRLD	8.820	0.1850	8.744	0.0739

T=Temperatrue, R=Rainfall, L=Water level, D=Discharge

Table 2:The comparison of ANN and Neuro-Fuzzy models in sediment Estimation (Validation set): Oyan gauging station of Ogun-Osun RiverBasin

Model Input	Neuro-Fuzzy performance		ANN Performance	
	RMSE	R^2	RMSE	R^2
TR	23.3134	0.15234	23.4762	0.37479
TRL	23.3443	0.3385	23.2996	0.1674
TRLD	23.2129	0.386	23.53045	0.378074

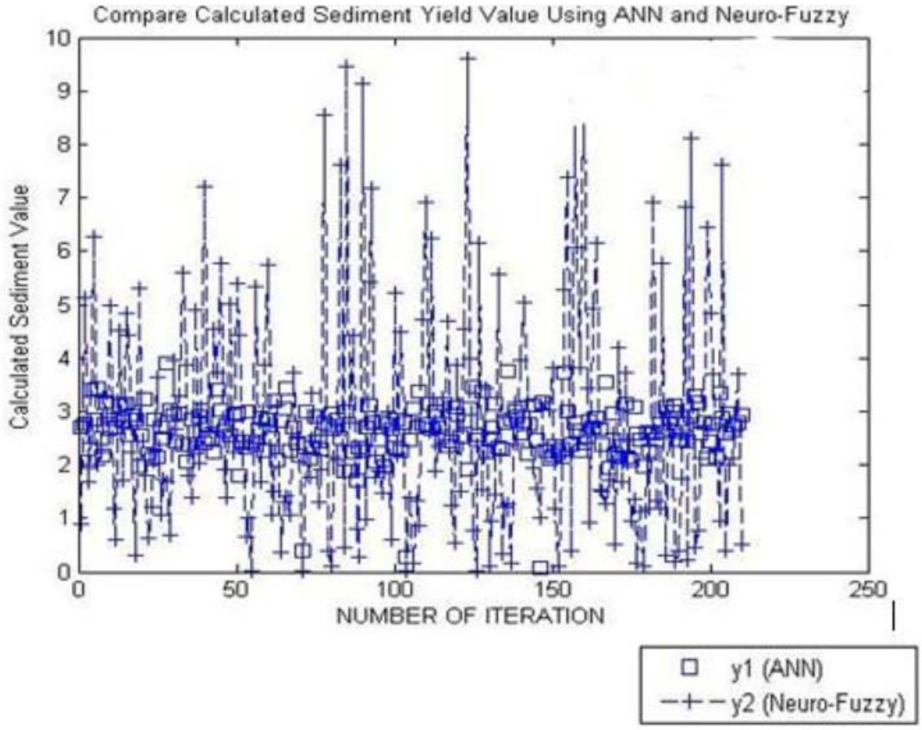

Figure 2: Graphical view of ANFIS and ANN using four variable (TRLD) (training set)

5. Conclusion

This work looked into the predictive strengths of both ANN and Neuro-fuzzy with focus on river sediment yield prediction. Data of a period of four years was used for both t the simulation and validation phases and at the end of the validation stage of the simulation using four input variables, the ANFIS model outperformed the ANN model. It had a lower RMSE of 23.2129 as against that of the of the ANN which had 23.53045. Their correlation coefficient was 0.3860 and 0.3780 respectively meaning that there is quantitative measure of some type of correlation and dependence between the observed and random variables of both models.

The ANN outperformed the ANFIS model only when they were fed with three variables. The ANN model had a RMSE of 23.299 as compared to that of ANFIS model which had a RMSE of 23.3443. This research has helped to improve the capability of ANN using the ANFIS model on the same data set that ANN has been used on.

REFERENCES

1. Agrawal Samarth, Manoj Jindal, Pillai, G.N (2010). Momentum Analysis based Stock Market
2. Bisht Dinesh C.S., and Ashok Jangid (2011) Discharge modelling using Adaptive Neuro-Fuzzy Inference System. International Journal of Advance Science and Technology Vol. 31, 1-5
3. Firat M. (2008). Comparison of Artificial Intelligence Techniques for river flow Forecasting. Hydrological and Earth system Sciences Journal.,12, 123-139.
4. Johnson, D.L., Ambrose, S.H., Bassett, T.J., Bowen, M.L., Crummey, D.E., Isaacson, J.S., Johnson, D.N., Lamb, P., Saul, M., and Winter-Nelson, A.E. (1997). Meaning of Environmental terms. Journal of environmental Quality. 26: 581-589.
5. Lohani, A.K., Goel, N.K., Bhatia, K.K.S (2009). Deriving stage discharge sediment concentration relationships using fuzzy logic. Hydrological Science journal 52(4), 793-807.
6. Nagy H., Watanabe, K., and Hirano, M. (2002). Prediction of Sediment Load Concentration in Rivers using Artificial Neural Network Model. *Journal of Hydrological Engineering*, 128(6), 588–595.
7. Ozgur Kisi, TefarukHaktanir, Mehmet Ardiclioglu, Ozgur Ozturk, EkremYalsin, and Salih Uludag (2008). Adaptive Neuro-fuzzy computing technique for suspended sediment estimation. Advances in Engineering Software, Volume 40,

issue 6, 438-444

8. Pankaj Kumar (2011) Crop yield forecasting by ANFIS. Mathematical theory and modeling journal Vol. 1 No. 3, ISSN 2225-0522, 56

9. Pollution issues (2016). Sedimentation. www.pollutionissues .com/Re-Sy/sedimentation.html. Date of last visit: 14/05/16

10. Raghuwanshi, N., Singh, R., and Reddy, L. (2006). Runoff and Sediment Yield Modeling Using Artificial Neural Networks: Upper Siwane River, India. J. Hydrol. Eng., 11(1), 71–79.

11. Ogundoyin I.K et al.,(2009). Artificial Neural Network Approach to River sediment yield prediction: A case study of Ogun-Osun River Basin: A case sudy of Ogun-Osun River Basin. Joural of Computer science and its applications, pp.8.