Application of ANFIS and ANN based models for forecasting of Iyidere River from Riza catchment.

Chelang A.Arslan
Assis.prof
Kirkuk University /Engineering College/Civil Engineering Department
Email:dr.chelenkalparslan@gmail.com

Abstract: Two different techniques for forecasting monthly streamflow of Iyidere river at Riza were investigated in this study. First of all the adaptive neuro fuzzy interference ANFIS was applied to the monthly data of the iyidere stream then the artificial neural networks were applied also using radial basis function networks RBNN and generalized regression neural networks GRNN. A combination of the streamflow data with the most important climatic factors nearby the stream station was investigated for both techniques. Coefficient of determination $R^2$, Mean absolute error MAE, $E_{nash}$ and $R_{bias}$ statistics parameters are used to evaluate the models performance.

Index terms: ANFIS, ANN, GRNN, RBNN, $E_{nash}$, MAE, $R_{bias}$, Streamflow.

1. INTRODUCTION

Stream flow forecasting is very important for many areas of water engineering such as dam planning, flood prediction and estimation of domestic water supply. Using different forecasting methods which are not based on physics equations, such as neural networks and Fuzzy Logic are becoming very common in use in various water resources engineering field. Artificial neural networks have been shown to give successful results in many fields of hydrology and water resources. Chen et al.,( 2006). Markus et al. (1995) used ANNs with back propagation algorithm to predict monthly stream flows at a gauging station in southern Colorado by using snow water equivalent and temperature. Karunanithiet al. (1994) compared the Artificial neural networks ANNs with traditional methods for modeling qualitative and quantitative water resource variables. Other researches Maier and Dandy (1996); Shamseldin (1997); Olsson et al. (2004) also used this technique successfully. Zealand et al. (1999) proved the success of training the ANN with back-propagation algorithm for 1-week-ahead streamflow forecasting. Imrie et al. (2000) applied the cascade correlation and back-propagation algorithms for short term streamflow forecasting. Actually Neural networks tries to emulate the operation of human brain the same is true for fuzzy systems. A neuro fuzzy system include ANN and fuzzy system, Keskin et al., (2006). Neuro Fuzzy logic system is a hybrid intelligent system that combines neural network with a fuzzy system. Fuzzy logic and neural networks are natural complementary tools in building intelligent systems. Nath, (2007). A daptive Neuro Fuzzy based interference System (ANFIS) introduced in hydrologic forecasting by Jang (1993) for forecasting river flow. Zadeh (1965) produced Fuzzy set theory with relative membership concept and proposed fuzzy optimum theory. Tayfur et al. (2003). A model Based on simulating rainfall-stream flow using fuzzy logic and ANN was produced by Tayfur and Singh (2006). A neuro fuzzy network model for forecasting the inflow of Brazilian hydroelectric plants was developed by Ballini et. al (2001).

In this research a comprehensive comparison of ANFIS model and ANN two different types in Iyidere river forecasting from Riza basin at Turkey was produced. The monthly data for this river was from 1995 to 2000. The effect of most effective climatic factors on stream flow were also included in the applied simulating models.

2. ADAPTIVE NEURO-FUZZY INFERENCE SYSTEM (ANFIS)

An adaptive neural network can be described as a network which consists a number of nodes connected through directional links. Each of these node is characterized by a node function with fixed or adjustable parameters. Learning or training phase of neural networks is an operation which aims to determine parameter values to fit the training data. The basic learning rule is the well-known back propagation method which seeks to minimize some measure of error, usually sum of squared differences between network’s outputs and desired outputs (Kay et al., 2002), (Ab aboodi, 2014). Most fuzzy inference systems can be classified into three types due to the types of inference operations upon “if-then rules”. These types are Mamdani’s system, Sugeno’s system and Tsukamoto’s system. The most commonly used one is Mamdani’s system. Sugeno’s system is more compact and computationally efficient. The output is crisp, so, without the time consuming and mathematically intractable defuzzification operation, it is by far the most popular candidate for sample-data based fuzzy modeling and it lends itself to the use of adaptive techniques (Takagi and Sugeno, 1985). In first-order Sugeno’s system, a typical rule set with two fuzzy IF/THEN rules can be expressed as:

Rule 1: If x is A_1 and y is B_1, then $f_1 = px_1 + qy_1 + r_1$ (1).

Rule 2: If x is A_2 and y is B_2, then $f_2 = px_2 + qy_2 + r_2$ (2).

Figure(1) illustrates basic ANFIS structure.
Each node i in the first layer which is shown in Figure(1) is an adaptive node that represents membership functions which can be explained by generalized bell functions:

\[ Z_{1,i} = \mu_t(X) = \frac{1}{1 + |X-c_i|/a_i} \]  

where \( X \) is input to the node and \( a_i, b_i \) and \( c_i \) are adaptable variables known as premise parameters. The outputs of this layer are the membership values of the premise part. This product represents the firing strength of a rule. The second layer consists of the nodes which multiply incoming signals and sending the product out.

\[ Z_{2,1} = W_{1} = \mu_t(x) \mu_d(y) \]  

In the 3rd layer, the nodes calculate the ratio of the \( i \)th rules firing strength to the sum of all rules' firing strengths.

\[ z_{3,1} = W_{1} = \frac{w_i}{w_1 + w_2 + w_3 + w_4} \]  

where \( W \) is the output of Layer 3 and \( \{ p_1, q_1, r_1 \} \) are the parameter set. Parameters of this layer are referred to as consequent parameters.

The single node in the fifth layer computes the final output as the summation of all incoming signals

\[ f = \sum_{i=1}^{n} W_i f_i \]  

output that is forwarded to next node in the sequence. This parallel processing allows for great advantages in data analysis. ANNs are widely used in various branches of hydraulic engineering and their property to approximate complex and nonlinear equations makes it useful tools in econometric analysis. Each network consists of an input layer, an output layer and one or more hidden layers. Hassan AM et al., (2009).

### 3.1 RADIAL BASIS FUNCTION NEURAL NETWORKS (RBFNN)

Radial Basis Function (RBF) is a powerful, fast learning, and self-organized neural network. It is better than BP network in approximation, classification and learning speed, especially in processing highly nonlinear problems. RBF neural network was proposed by Moody and Darken (1980). It includes three layers: an input layer, a hidden radial basis neuron layer and a linear neuron output layer. Its structure is illustrated in Figure(2).

### 3.2 GENERALIZED REGRESSION NEURAL NETWORKS (GRNN)

A GRNN is a variation of the radial basis neural networks, which is based on kernel regression networks. A GRNN does not require an iterative training procedure as back propagation networks. It approximates any arbitrary function between input and output vectors, drawing the function estimate directly from the training data. In addition, it is consistent that as the training set size becomes large, the estimation error approaches zero, with only mild restrictions on the function. Kim B, et. Al., (2004). Figure (3) represents the basic structure of generalized regression networks.

**Figure(2) Structure of Radial Basis Function Neural Networks.**

**Figure(3) Structure of Generalized Regression Networks.**
4. THE STUDY AREA:

Iyidere stream is located between cities of Rize and Trabzon (Figure 4). The stream is about 160 km long and 20 m wide. Substrate of the stream consisted of rock and stone. The average flow speeds are about 2 and 4 m/sec in lower and upper basins, respectively.

5. APPLICATION AND RESULTS:

The Monthly flow data which were taken from station no 2218 of Iyidere stream at Riza state in Eastern black sea basin, Turkey with the climatic factors of Riza state around stream are used in this study. The location of the stream is shown in Figure (4). The climatic factors which were thought to be effective for describing the streamflow are: Monthly mean rainfall, Monthly maximum rainfall, monthly mean temperature and monthly humidity. The data sets for monthly flow values for Iyidere stream and other considered factors were taken from 1955-2000 as monthly mean values. Four statistic parameters were used to evaluate the models performance which are coefficient of determination $R^2$, Mean absolute error, MAE, Nash-Sutcliffe efficiency $E_{nash}$ and percent bias $R_{bias}$.

These parameters are defined as:

$$R^2 = \frac{\sum_{i=1}^{n}(A_t-F_{\text{mean}})(F_t-F_{\text{mean}})^2}{\sum_{i=1}^{n}(A_t-A_{\text{mean}})^2 \sum_{i=1}^{n}(F_t-F_{\text{mean}})^2}$$

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^{n} |A_t - F_t|$$

$$E_{\text{nash}} = 1 - \frac{\sum_{t=1}^{n}(A_t-F_t)^2}{\sum_{t=1}^{n}(A_t-\overline{A})^2}$$

$$R_{\text{bias}} = 100 \times \frac{\sum_{t=1}^{n}(F_t-A_t)}{\sum_{t=1}^{n}A_t}$$

where $A_t$ is the observed value and $F_t$ is the forecasted value and $F_{\text{mean}}, A_{\text{mean}}$ are the mean value of the series (Chokmani et al., 2008; Tiyaki S et al., 2014. The percent bias expressed the degree to which the estimation is under or over estimation. A positive percent bias indicates overestimation, whereas a negative percent bias indicates underestimation (Salas et al. 2000). The optimum value of $R_{bias}$ is 0.0. A brief description of $R_{bias}$ and $E_{\text{nash}}$ was given also by Moriasi et al. (2007), this description can be summarized as:

- **VG**: $0.75 < E_{\text{nash}} \leq 1.00$; $R_{bias} < -10$
- **G**: $0.65 < E_{\text{nash}} \leq 0.75$; $10 \leq R_{bias} \leq -15$
- **Satisfactory (S)**: $0.50 < E_{\text{nash}} \leq 0.65$; $15 \leq R_{bias} \leq 25$
- **Unsatisfactory (US)**: $E_{\text{nash}} \leq 0.50$; $R_{bias} \geq 25$

The analysis suggested that 6 antecedent flow values are adequate with the selected climatic factors as input vectors to the ANFIS and ANN models with different combinations. Different ANFIS architectures were tried to find optimum one. Using 2 Gaussian membership function gave the best performance. Different ANN structures were also tried for both RBNN and GRNN. The data sets for all input variables were divided for training and test periods. Table (1) illustrates the results for ANFIS model with investigated architectures. The best result was...
fond to be by using input combinations as \( Q_{t-1}, Q_{t-3}, Q_{t-4}, Q_{t-5} \), monthly mean Rainfall \( R_m \), monthly maximum rainfall \( R_{mx} \), monthly mean temperature \( T_{mean} \), monthly mean humidity \( H \). The best comparison can be represented by following Figure (5) which shows the monthly flow values which is resulted from the applied models with the observed data.

By using radial basis function networks RBFNN following results which are illustrated in Table (2) were found. The Results represents the best found one for each tried input combination after applying different numbers of neurons at the hidden layer and different spread values. The values of all statistical parameters were better than ANFIS models. The best RBFNN model was by applying input combination which is consisting of flow values of the stream for lag 1,2 and 3 months with the values of monthly mean Rainfall \( R_m \), monthly maximum rainfall \( R_{mx} \), monthly mean temperature \( T_{mean} \), monthly humidity \( H \) by using spread value =0.04 and 200 neurons at the hidden layer. Since the value of coefficient of determination was found to be 0.9866, the same value was found for \( E_{nash} \), \( R \) bias value was = 0.1576 and with mean absolute error =1.544.

Another ANN model was investigated which is generalized regression neural networks GRNN with the same tried input combinations. The results of this trial are illustrated in the Table(3). It is clear from this table that the performance of the RBFNN was better than GRNN performance. By comparing the applied methods on the Iyidere stream the high performance of RBFNN models was concluded. As it is seen from the results that the RBFNN is slightly superior than GRNN and ANFIS models.

The detailed comparison can be also shown with scatter plot for each applied model as shown in following figures.
It is important to say that for all applied models the performance was clearly improved by including the climatic factors with streamflow values therefore the results of the models which are consisted from streamflow values only were not produced.

6. CONCLUSIONS

The potential of ANFIS, RBGNN and GRNN for prediction of future values of Iyidere streamflow at Riza region northern Turkey has been presented in this research. Different input combinations with different architectures were applied. It was found that the RBNN models showed better accuracy. ANFIS model performs better than GRNN models while the RBFNN models were better than ANFIS models. The performance of all the applied models were evaluated by statistical indices such as R², MAE, EMAE, and R bias. All the applied models were performing better with including the rainfall and humidity and temperature factors, Comparatively, RBNN models performance was found to be superior as compared to GRNN and ANIS models.

REFERENCES:


