International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org Volume 3 Issue 4 ǁ April 2014ǁ PP.54-59

The Artificial Neural Network and Time Series Models Assessment (ARMA) In River Flow Simulation (Case Study: Jamash River)

Mehrdad Fereydooni [1], Mojtaba Khalilizadeh [2], Mohsen Tabasi[3]
[1]*(Asistant Professor,Department of Civil Eng.Larestan Branch Islamic Azad University, Larestan, Iran. Corresponding author, m.fereydooni@yahoo.com.)*
[2]*(Asistant Professor,Department of Civil Eng.Larestan Branch Islamic Azad University, Larestan, Iran.)*
[3]*(M.sc.candidate, Larestan Branch Islamic Azad University, Larestan, Iran.)*

ABSTRACT: *One of the basic tools of water resources management is to predict the water demand and supply. Predicting suitable river flows in civil tasks, rivers rearrangement, fluid warning systems, and particularly, planning to achieve optimal utilization from the reservoirs of the dams, is likely a vital issue. Lots of different methods have been developed in order to predict river flows over the recent years. In current study, the water flow of JamashRiver has been simulated in two stations with Neural Network and Time Series Models (ARMA) with different structures,and finallythe neural network MLP-GDX is considered as the most suitable model, in accordance with statistical indexes R2=95/89 and RMSE= 4/99 percent.*

I. INTRODUCTION

One of the basic tools in water resources management is to predict the water demand and supply.Predicting suitable river flows in civil tasks, rivers rearrangement, fluid warning systems, and particularly, planning to achieve optimal utilization from the reservoirs of the dams, is likely a vital issue. Lots of different methods have been developed in order to predict river flows over the recent years that can be classified into two models: a conceptual model and a "*based on the statistical data*" model.Considering the conceptual models' requirement to accurate and complete data and knowledge about the physical alternatives which can affect the river flow in a particular place, and this has been impossible so far, researchers appealed to use statistical models. During the recent four decades, Time series models have been widely used as a statistical model in predicting the river flows. It has been observed that the *artificial-intelligence*methods have been broadly utilized in recent years such as the artificial neural networks as well as the Time series methods in the fieldsin which the relationship between the input and output isnon-linear. These artificial-intelligence methods act like a proper flight recorder which is less restricted by physical matters and can model the non-linear and non-static processes of the river flow with no need to model the environmental and geometrical factors which are effective on the river flow.In this study the monthly flow of Jamash River is predicted in two stations of Sikhouran and Sarkhain a monthly scale by means of artificial neural networks and Time seriesmethods and their functions are compared with each other. Water-controlling structures such as dams, flood walls, …play an efficient role in reducing or eliminating the damages made by flood. However, in a lot of cases, the topographical or economic factors make it impractical to control the flood. Therefore, predicting the river flow provides alternative tools to reduce the damages of the flood. Warning prior to a progressive flood, allows evacuating people, domestic animals and machinery. Considering appropriate results gained from the models, their utilization in these fields is adequately acceptable. Some researches performed on the river water flow by means of Artificial Neural Networks and Time series are: Marcos (1995), Hall & Mainz (1998), Dahn (1998), Madsen (2000), Kishi&Kebanz (2009), John Adamofski et al (2010), Rahnama (2003), Dastourani& Wright (2004), Rezaei et al. (2007), Barani (2002), Sartaj et al. (2009), KarimiMasouleh et al (2010).

II. STUDIED REGION

With an area of about 1001 km^2, Jamash River watershed in North East of Bandar Abbas located between Eastern longitude 49 30 56 to 32 28 56 and Northern latitude 558 29 30 and 431 55 27 in the Persian Gulf basin (see fig. 1-2) and considering the divisions of the country, it is located in the region of Bandar Abbas city (Hydrological Report of Jamash).

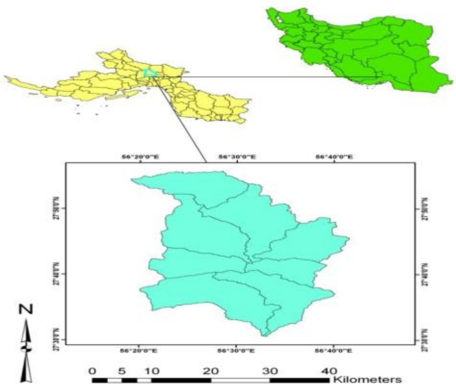


Figure (1): the location of Jamash River watershed

III. UTILIZED NETWORKS AND ALGORITHMS:

Based on how the nodes connection in the architecture of the neural networks, they are divided into the following types:

3-1- Progressive Networks

In a progressive network, the nodes are located in consecutive layers and their relationship is unilateral and when an input pattern is applied to the network, the first layer calculates its output values and delivers it to the next layer. The next layer receives these values as the input and transfers its output to the next one. Each node transfers the signal to the nodes of the next layer.

$$net_j = \sum_{j=1}^{L} w_{ij}x_j + w_0$$

Where,
$net_j$is the net input to the motor function,
w_{ij} is the matrix of the network weights
$\mathbf{x_j}$ is the input vector of the network
andw_0 is the bias vector of the network

3-2- Reverse Networks

They are the networks with at least one reverse flow from the output; that is, the layer received in this kind of network is affected not only by connecting to the input layer, but also bythe output vector. This network uses the output data as the new improved weights, therefore, allows the weights to return to the input. Figure (3-3) shows the subsets of these two networks each of which is used for certain purposes.

Various Training Algorithms:
All implemented stages were performed in a network using different mathematical algorithms.

4-1- Luneburg – Marquardt Algorithm

This is the modified Newton's classical algorithm which is used to find a suitable solution for the problems with minimization requirement. This method, like Newton's method, considers an approximation to Hazen's matrix of weights variations:

$$X_{k+1} = X_k - \left[J^T + \mu I\right]^{-1} J^T + e$$

Where,
X is the neural network weights,
K is the number of rehearsals,
T indicates the transposed matrix,
J is Jacobin's standard matrix of the network performance which should be minimized,
μ is the number which controls the training process,
and$e$ is the residual error vector.

4-2- Momentum Back-error Propagation Descent Gradient

In this method, the back-error propagation algorithm is used to estimate the network error and determine the weight vector and network bias (critical limit) so that it has the least error. The momentum parameter creates a kind of motional inertiato change the weights. This leads the system to get to the converge stage with fewer samples and in a shorter time interval. If the used data has fewer numerical errors, it is possible to use the above mentioned momentum, but in inaccurate systems lower amounts should be adopted.

4-3- Bayzyn Setting:

Bayzyn setting is an algorithm which automatically sets suitable values for the functional parameters. In this method, weights and biases of the network are considered as random variables with particular distributions. The statistical techniques are used to estimate the parameter with unknown changes. The advantage of this algorithm is that the size of the network has less effect on its results.

Time Series (ARMA)

ARMA model with ***p*** and ***q*** parameters is generally expressed as follow:

$$\mathbf{x}_t - \emptyset_1 \mathbf{x}_{t-1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots - \emptyset_p \mathbf{x}_{t-p} = \mathbf{z}_t + \theta_1 \mathbf{z}_{t-1} + \cdots + \theta_q \mathbf{z}_{t-q}$$

Where,

$\mathbf{X}_{t-i}$: observations of groundwater level at the time t-i,

$\mathbf{Z}_{t-i}$: Noise seed at the time t-i,

and$\mathbf{Ø}_i$, $\mathbf{Ø}_j$: the Model Coefficients.

There are three fundamental stages in modeling the time series including:

a) Specifications and model determination

In this stage, considering the statistical characteristics of the time series and comparing with the properties of different models, the type and the form of the model is already specified, which in fact, represents the manner of the time series. However, the form of the model is not a fixed and final one and may change in later stages.

b) Model Fitting

Model fitting over the Time series data means to evaluate the parameters of the model on the basis of data given together with reliable investigation on the evaluated parameters.

c) Recognizing the Model Credit

In this stage the quality of the adopted model will be checked, i.e. how well the model has been fitted on the sample data and furthermore, whether the existing assumptions of the adopted model are true for the Time series.

The criterion to Evaluatethe Workability and Errorsof the Model:

The evaluation criteria to assess the workability of different models have been represented by the International Meteorological Organization.

- The average square root of the error will be estimated as follow:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)^2}{N}}$$

In the above equation, y1 is the “observed results”, ŷ1 is the “estimated results” and N is the “frequency of the observations, and RMSE shows the difference between the observed and calculated values.

R^2presents the Network’s Efficiency and is expressed as follow:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

The most optimum answer of this model will be gained when **RMSE** tends toward zero and R^2 tends toward one, and in Time series the Residual Analysis must be checked as well as the above stages.

1- Analysis Results:

Data preparation of these models will be usually done in two stages:

a) Data Normalization:

The first step to prepare the data for modeling by neural networks is data normalization. There are several methods of data normalization. In this study, equations 1-4 are used to normalize the data.

b) Data Classification:

$$x_{normalize} = \frac{(x - x_{\min})}{(x_{\max} - x_{\min})}$$

The most important effort taken after data normalization is to determine the set of training, examining and validationdata. The examining data set which examines the model generalization capability must explain and represent the whole data sets. In this study, 70% of the total data is used for training, while 30% is used for examining as well. The results gained via different structures are presented with six input patterns by means of error evaluation criteria***R²*** and **RMSE** in table (1-7). Considering the results, it has been clarified that the best input pattern for prediction is the fourth one, that is, the waterflow with two delays and precipitation, evaporation and temperature, each one with a delay, same as figure (1-7). In this study, the best model is determined with 5 nodes in the first layer and 6 nodes in the middle one. Thus, this type of model is used to simulate and predict the water flow rate.

Input Pattern	Network	Network Structure	R^2	RMSE
1	FNN-LM	6-8-1	0.5904	0.012718
	FNN-GDX	6-4-1	0.5937	0.0103704
	FNN-BR	5-7-1	0.5882	1.1352167
2	FNN-LM	6-8-1	0.7622	0.0175108
	FNN-GDX	5-8-1	0.7155	0.0140149
	FNN-BR	4-4-1	0.6787	0.0123548
3	FNN-LM	5-7-1	0.8784	0.0086025
	FNN-GDX	4-5-1	0.787	0.0094812
	FNN-BR	4-4-1	0.8645	0.0145221
4	FNN-LM	5-6-1	0.9713	0.0044997
	FNN-GDX	4-8-1	0.671	0.051875
	FNN-BR	4-7-1	-0.197	0.0127306
5	FNN-LM	5-6-1	0.8002	0.0175013
	FNN-GDX	5-4-1	0.4026	0.0184363
	FNN-BR	4-8-1	0.7568	0.0129763
6	FNN-LM	4-7-1	0.7996	0.011954
	FNN-GDX	4-4-1	0.5441	0.013388
	FNN-BR	5-6-1	0.8109	0.0248813

Table (1-7): Results gained via different structures with 6 input patterns

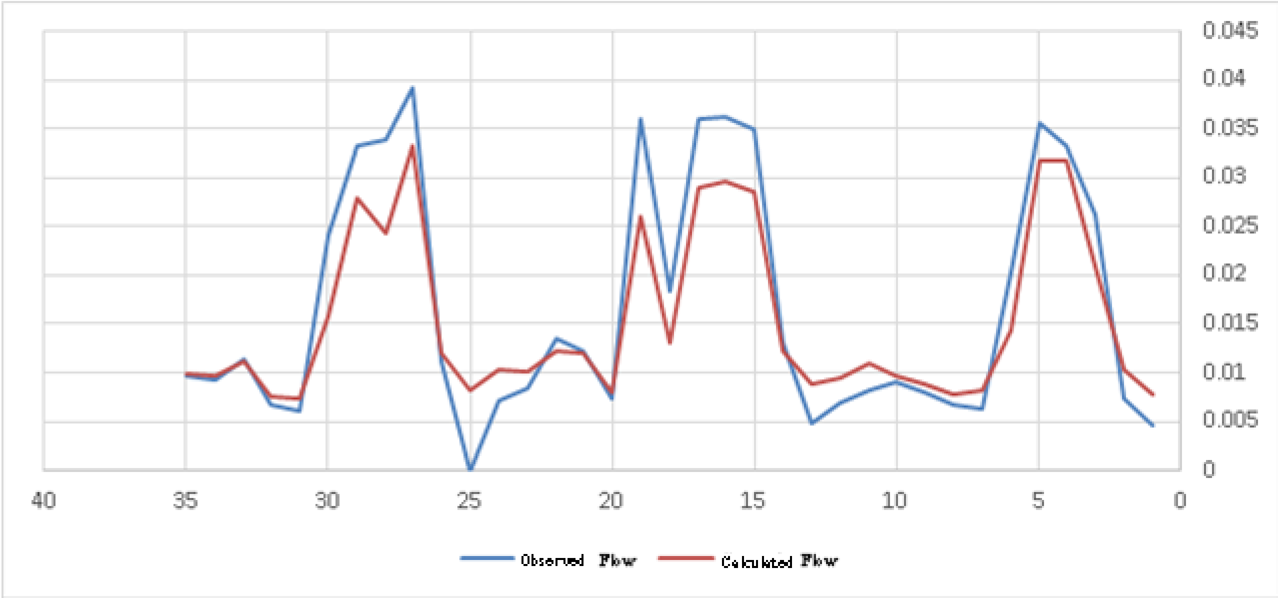
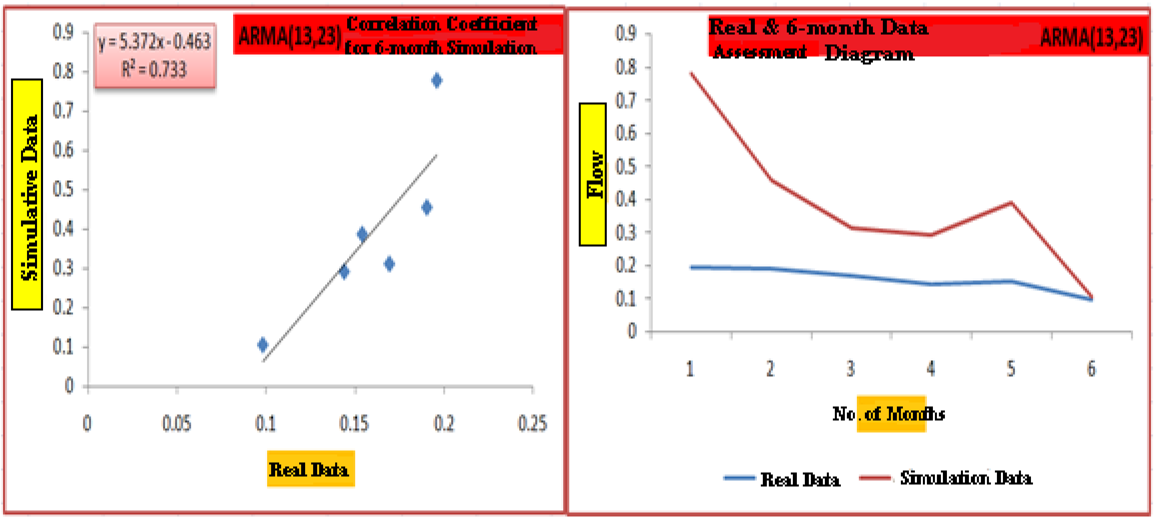


Figure (1-7): Observed and calculated water flow diagram of the fourth pattern with network type FNN-LM model



The results of the Stoke Stykymodel (ARMA) in simulation did not meet the requirements at all for 36 months due to the lack of data and the 6-month simulation is related to ARMA (p=13,q=21) with correlation coefficients equal to R^2=0.733 and **RMSE**=0.2604 (see fig. 2-7).

Figure (2-7): Dotted diagram of R^2 correlation coefficient and the true and simulated data assessment diagram

Briefly, it can be stated that considering the results achieved through different models the artificial neural networks capability is beyond the Time series models.

REFERENCES:

[1] Marcos (1995). They used neural networks during their researches to predict monthly water flow of a river in Western Colorado.

[2] Hall and Maness (1998) in two southern and western regions of England, evaluated the peak flow rate of the watersheds lacked the hydrometer station by means of different methods.

[3] Dahn (1998) in a study in the central highlands of Vietnam in two watershed regions of Da Nhimn and Nga, predicted the daily water flow of the river using artificial neural network method with the method of training after error propagation and physical method of the Tank prediction model.

[4] KarimiMasouleh (1388 / 2009) Application of neural networks in predicting Karoon River's water flow rate based on the statistical data of the last six months precipitation of the upstream stations – the Eighth International Seminar on River Engineering

[5] Rahnama (2003) Studied Qareghajbolson flood prediction using artificial neural networks method.
[6] Rezaei and colleagues (2007), using physiographical features of Sefidrood River watershed regions and one and five-day precipitations before the example flood, tried to model the peak water flow rate regionally.
[7] Sartaj, Majid et al (1387, 2008) predicting the River flow by means of an artificial neural network – The 3rd Iranian Water Resources Management Conference – October 2008
[8] Mohammadi, Mahmood, et al (1385, 2006) predicting average weekly water flow rate entering Doroodzan Dam Reservoir by artificial neural network –National Conference on Irrigation and Drainage Networks Management- May 2006
[9] KarimiMasouleh et al (2009)investigated the neural networks application in predicting Karoon River's water flow rate on the basis of the statistical data of the lastsix months precipitation of the upstream stations.