ciência e natura

River Flow Forecasting using artificial neural network (Shoor Ghaen)

Mohsen Rezaei [1], Ahmad Ali Akbari Motlaq [2], Ali Rezvani Mahmouei [3], Seyed Hojjatollah Mousavi [4]

1. Faculty member, Department of Civil Engineering, University of Zabol (m59_rezaee@yahoo.com)
2. Student in Civil Engineering Water, University of Zabol (emehr915575@gmail.com)
3. Faculty member, Department of Civil Engineering, Bozorgmehr University of Qaenat (rezvani@buqaen.ac.ir)
4. Faculty member, Department of Civil Engineering, University of Zabol (h.mousavi@uoz.ac.ir)

Abstract

In our country, most of the rivers located in dry and warm climate areas are seasonal, and many of them have experienced floods. That, along with concerns about scarcity of water resources and the need to control surface water, makes identification, modeling, and simulation of rivers' behavior, necessary for to long-term planning and proper and rational use of river flows potential. Rainfall phenomenon and the resulting runoff in watersheds, as well as predicting them are of nonlinear system types. Artificial neural networks are able to analyze and simulate phenomena in nonlinear and uncertain system where the relationship between the components and system parameters are not well known or describable. Shoor Ghayen River, with 100 km length is the biggest seasonal river of Qaenat city and the main source of water in Farrokhi storage dam. Therefore, in this study according to the rainfall and runoff statistic of Khonik Olya hydrometric and Ghayen synoptic stations between 1976-1977 and 2010-2011 water years, precipitation phenomena and river runoff was predicted. MATLAB software is used to perform calculations. For modeling artificial neural network, 85 percent of data were used for training the proposed method, the remaining 15% were used for validating the method using 10 neurons, and a network with an error of less than 5% was developed for each month. The maximum correlation in evaluation phase was for April with the value of 0.99, and the minimum was for June and August with a value of 0.92. Overall results indicate optimum performance of artificial neural networks in predicting runoff caused by rainfall. It is also found that better results can be achieved by standardizing the data.

1. Introduction

The most important issue when dealing with water crisis is the management of these valuable resources. Therefore, due to the limited water resources, proper and optimized management is the most important task of policy-makers and engineers in this field. One of the most important elements of water resources management is predicting the volume of these resources, especially predicting the flow of the rivers. In our country, most of the rivers in different geographical areas are seasonal, and many of them have experienced floods. This, as well as scarcity of water resources and the necessity of harnessing surface waters indicate the importance of identification and simulation of river behavior for long-term planning and utilization of river flows potential.

Due to the multiplicity of hydrological factors that govern rainfall phenomena in watershed basins, the reaction of most of the basins against precipitation is complex [1]. Among the methods that have attracted attention of many researchers in recent years in analysis of complex and nonlinear phenomena, is the use of artificial neural networks. Consequently, it has found many applications in water engineering problems including hydrology [2].

Wu and Chau (2011) simulated flows of Daning and Luzhou Rivers in China using artificial neural network model and singular spectrum analysis, and achieved acceptable results [14]. Misaghi and Mohammadi used artificial neural network with different input vectors in simulating rainfall – runoff phenomena and river flow routing and argued that by reducing the number of neurons in the output layer, the network's ability to achieve the desired results will improve. In addition, The Multi-Layer Perceptron (MPL) network with delta method was more efficient in the process of evaluation and validation, and by using various activation functions instead of a fixed activation function, more satisfactory results are provided [8]. Dawson and Wilby (1998) suggested precipitation- runoff models for forecasting river flows in two flood basins of England (Amber and Mole) using artificial neural network models [9]. Zare and Bayat compared neural intelligent models and empirical methods in a study to estimate Zayandehroud River annual runoff in Isfahan province, and concluded that performance of neural intelligent models is able to provide better results compared to empirical methods [3].

2. The Basin Under Study

The basin of the river under study (Shoor Ghayen River) includes Ghayenat region with an area of 18783 square kilometers at 59 degrees 11 minutes 38 seconds longitude and 33 degrees 43 minutes and 47 seconds latitude. The basin under study with an average elevation of 1420 meters above sea level, the average annual rainfall of 173 mm and maximum and minimum monthly temperature of 42°C and 23°C, respectively, have a dry climate. This river with 100 Km length and an average gradient of 0.8 percent originates from West and Southwest highlands, and following a northwest – southeast path, enters Afghanistan in the eastern regions of Iran. In recent years, Farrokhi Dam has been constructed on the river for the storage of flood flows with a reservoir volume of 5 Million cubic meters. Figure 1 shows map of the study area.

Figure 1: Map of the study area and Shoor Ghayen river basin

3. Materials and methods Structure and algorithm of artificial neural networks

Artificial Neural Network is one of computational methods that tries to provide a mapping between input space (the input layer) and optimal space (the output layer) by understanding the inherent relationships between data with the help of training procedures and using processors called neurons.

The system consists of a large number of fully interconnected processing elements called neurons, which act in coordination to solve a problem. Each neural network is made of three layers including input, intermediate, and output. Transmission functions typically used in artificial neural network are sigmoid and hyperbolic tangent functions. Figure 2 shows an overview of a neural network structure.
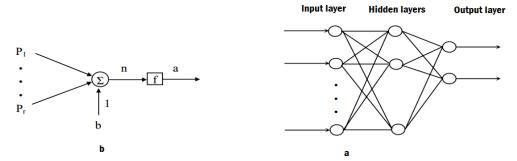

Figure 2: (a) The general structure of artificial neural networks - (b) - Neuron with r input

These networks are mathematical structures capable of showing processes and non-linear components for communication between the input and output of each system. The most interesting feature of artificial neural network is its ability to learn based on training patterns that through the process of learning, neural networks will find optimal weights. Different algorithms have been provided for training neural networks among which the back-propagation learning method so widely used in hydrological models.

4. The model theory

Data normalization: data standardization means scaling them to fit within a certain range, which is done between 0.5 and 1 according to the following formula:

$$X_{normalized} = 0.5 + 0.5(\frac{X_{real} - X_{min}}{X_{max} - X_{min}})$$

2. Finding the best network structure: The best network structure is to create the right combination of layers and the number of their neurons, the type, and the number of membership functions, to achieve minimum

error and the highest correlation, which is done by test and trial method.

3. Network training: About 85% of data were used for network training, and the training algorithm is presented.

4. Network Test: The rest of the data (15%) were allocated for review and testing; the test is necessary to make sure of trained network accuracy.

In the present study, Backpropagation with Feedforward architecture was used for network training in artificial neural network modeling, and among different algorithms of this learning method, Levenberg Marquardt algorithm with Learngdm weight learning function with 10 neurons were used. The Log-sigmoid (Logsig) and tangent sigmoid (Tansig) transfer functions were used in intermediate layer and linear function (Pureline) in the output layer.

5.Evaluation Criteria

The following common statistics validation criteria are used to evaluate the used methods (functions):

$$R = \frac{\sum\limits_{i=1}^{N} t_i p_i}{\sqrt{\sum\limits_{i=1}^{N} t_i^2}\sqrt{\sum\limits_{i=1}^{N} p_i^2}}$$

1. Regression coefficient :

$$R = \frac{\sum\limits_{i=1}^{N} t_i p_i}{\sqrt{\sum\limits_{i=1}^{N} t_i^2}\sqrt{\sum\limits_{i=1}^{N} p_i^2}}$$

2.Calculate the root mean square error (RMSE):

3.And for the average error (MAE):

Artificial neural networks algorithm was used in predicting runoff from rainfall is shown in Figure 3:

$$MAE = \frac{\sum\limits_{i=1}^{N}(T_i - P_i)}{N}$$

```
┌─────────────────────────────────────┐
│            Collecting data           │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│    Categorization and creating a Matrix    │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│    Input and output data normalization    │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│      Determining the structure      │
│           of the network            │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│ Selecting the type of activation function, values of training parameter, number of │
│ hidden layers, the maximum number of hidden neurons, and the type of function │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│ Network training with LM algorithm starting from the lowest hidden neurons │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│            Evaluate and             │
│            network test             │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│ Network regression analysis, determining the correlation between │
│ the observed and predicted and determination of error percent │
└─────────────────────────────────────┘

┌─────────────────────────────────────┐
│       Does the network have the error below       │
│ five percent and a correlation coefficient close to one? │
└─────────────────────────────────────┘

                  Yes

           End and save program
```
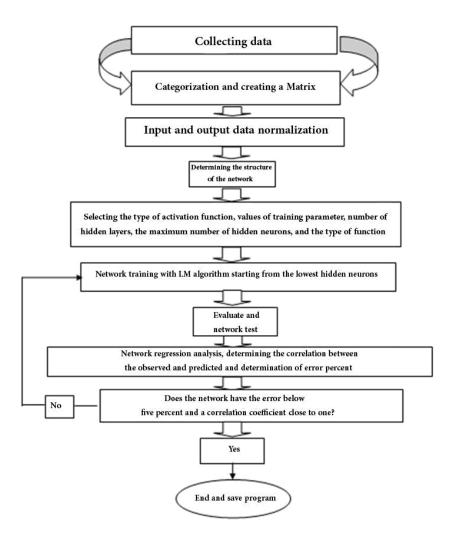
Figure 3: Flowchart of the project

6 Neural network design

In this study, the neural network model was assessed using different inputs. Usually, rainfall is the first option to simulate runoff and to assume a base discharge, the discharge of one or two months ago or more can be used as an input. The other inputs such as initial moisture index can be added later.

The general function of discharge in river for each month is defined as the following:

$Qt. = F (Q_{t-1}, Q_{t-2}, ..., P_t, P_{t-1}, ..., P_{t-n}, T_t, T_{t-1}, ...T_{t-n}, E_t, E_{t-1}, ..., E_{t-n})$

Where: Q is the flow discharge, P is precipitation, T is Temperature, E is evaporation, t is month number and n is the last month's numbers.

As it can be seen, the above equation is a big equation and it has many parameters. In general, discharge in every month depends on the discharge of the past few months and rainfall in the same month and past few months. The effect of temperature and evaporation is lower compared to other parameters. Therefore, only the effect of temperature in each month was used in the analysis.

In this research, input variables and discharge was modeled so that the flow in April, May, June, July, and August can be predicted using the total amount of precipitation, mean

temperature an discharge through consecutive months, because seasonal rivers often flow in autumn and winter. Input variables including total monthly rainfall and average temperature and discharge in several consecutive months and output flow (discharge in the coming months) were formed on a monthly basis and with different time delays (one, two and up to 6 months), and the effects and correlation of past time sequences of each variables on future discharge rate, and the most appropriate model for each of the input variables was determined using R^2. Overall, the proposed input models are given in Table 1.

Table (1): Suggested input models and their correlation coefficient

Discharge forecast	Average discharge	Average temperature	Total Precipitation
April	(October -September) $R^2 = 0.13$	(September) $R^2 = 0.09$	(December -September) $R^2 = 0.02$
May	(April – January) $R^2 = 0.14$	(April) $R^2 = 0.05$	(November-April) $R^2 = 0.06$
June	(May) $R^2 = 0.93$	(May) $R^2 = 0.05$	(December-May) $R^2 = 0.05$
July	(June) $R^2 = 0.78$	(June) $R^2 = 0.14$	(February- June) $R^2 = 0.06$
August	(July) $R^2 = 0.77$	(July) $R^2 = 0.08$	(February- July) $R^2 = 0.058$
September	(May- August) $R^2 = 0.52$	(August) $R^2 = 0.14$	(September-August) $R^2 = 0.05$

7.Discussion and Results

In this study, using data on water years of 1976-77 to 2010-11 of Khonik Olya Hydrometer Station and Ghayen Synoptic Station, the relationship between the discharge of the river and its discharge in recent months, as well as rainfall and temperature for every month and the last month were evaluated. Of available data, 85 percent was used for training and the other 15% was used to evaluate the method that was used in this study. As shown in Table 2, the maximum correlation coefficient in the training phase was 0.99 for April, and the minimum was 0.92 corresponding to June and August.

Table 2: Summary of statistical results of forecasting models in training phase

Parameter	RMSE(m3/s)	MAE(m3/s)	R	R2
April	0.01	0.002	0.99	0.98
May	0.01	0.0002	0.98	0.97
June	0.04	0.003	0.92	0.85
July	0.04	0.006	0.94	0.89
August	0.05	0.004	0.92	0.85
September	0.03	0.006	0.96	0.92

Table 3: Summary of statistical results of forecasting models in test phase

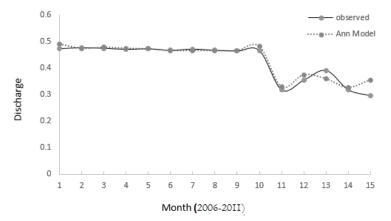| Parameter	RMSE(m3/s)	MAE(m3/s)	R	R2
April	0.21	0.09	0.99	0.98
May	0.007	0.002	0.96	0.92
June	0.16	0.12	0.98	0.96
July	0.03	0.014	0.97	0.94
August	0.04	0.04	0.93	0.86
September	0.12	0.08	0.99	0.98

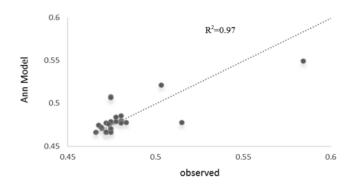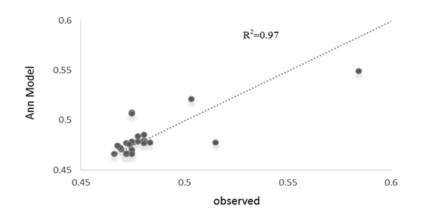Figure 4: The actual data and simulated test phase for spring

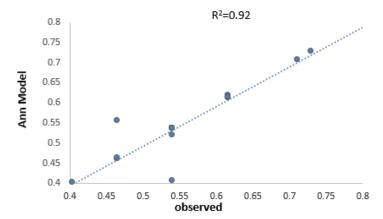Figure 5: Discharge correlation values observed and estimated in test phase in spring

Figure 6: Discharge correlation values observed and estimated in training phase in May

Figure 7: Discharge correlation values observed and estimated in training phase in September

In addition, to predict the discharge of April, the average discharge from six time steps earlier, and total precipitation from four time steps earlier and the average temperature in one time step earlier was used relative to April discharge. This trend has the highest correlation coefficient for the other months, and R^2 is above 85% for April to September (Results are shown in Table 2). Overall results indicate optimum performance of artificial neural networks to predict the rainfall-runoff. Of other results of this research is that with standardizing the data better results can be achieved for prediction. The results of error level in neural network method indicate performance of neural network model in predicting rainfall-runoff.

8. Conclusion

Precipitation-runoff interactions and their correct evaluation is an inseparable part of every hydrological study. Complex, fuzzy, and non-linear factors effective on precipitation-runoff relationships cause mathematical modeling to be difficult. Rainfall-runoff models are important tools in water resources projects, and the performance of rainfall- runoff models depends on the right choice of model parameters. Discharge changes in drainage

basins, other than independent variables examined in this study, has a strong relationship with other climatic elements of the basin, therefore it is recommended that in these models, the role of other factors affecting the flow of rivers in drainage basins also to be taken into consideration.

Although by reducing the period of the input data, the model accuracy is also reduced, but in the absence of longer data, the method presented in Table 1 provides acceptable results, and the minimum suitable period with appropriate accuracy can be determined for the input data. Overall results indicate optimum performance of artificial neural networks in predicting rainfall-runoff. It is suggested that rainfall-runoff simulation with the impact of other factors and various algorithms for optimum evaluation of artificial neural network to be conducted for determination of the best algorithm.

9. References

Nourani, V., Salehi, K. 1998 Rainfall-runoff modeling using fuzzy neural network approach in comparison with fuzzy neural network approaches, The fourth National

Congress of Civil Engineering, Tehran University, pp. 1-8.

Keshvari, F., Shams Nia, S. A., 2014, Performance evaluation of artificial intelligence models (artificial neural networks and Adaptive Neural Fuzzy Inference System (ANFIS) In predicting monthly discharge of river flow (Case Study: Bahman Dam - Ghareaghaj River).

Zare Abyaneh, H., Bayat Varkeshi, M., 2011, Evaluation of intelligent and empirical models in estimating annual runoff, Iranian Journal of Soil and Water (Agricultural Science and technology), Vol. 25 (2): 365-379.

Salahi, B., Sarmasti, T., 2013, Simulation of rainfall-runoff in Gharehsou river southern sub-basin by using artificial neural networks (ANNs), Journal of Geography and Environmental Planning, Vol. 24, 52nd edition, Issue 4.

Ghaffari, G., Vafakhah, M., 2013, Rainfall - runoff process simulation using artificial neural network and adaptive fuzzy system in Haji Qushan Drainage Basin, Journal of Drainage Basin Management, Vol. 4, 8th edition.

Borhani Darian, E. R., Fatehi Marj, E. 2008, Application of artificial neural network in predicting river flow using Nazloochaei river drainage basin climate indexes, Journal of Technical College, Vol. 35, Issue 3.

Eskandarinia, E. R., et al., 2011, Evaluation of the effect of previous rainfall in estimation of river flow using artificial neural network in Bakhtiari River, Journal of watershed management, Vol. 2, Issue 3.

Misaghi, F. and K. Mohammadi. 2003. Simulation of rainfall runoff and river routing using ANN. Proceeding of the 6th International river congress. Ahvaz. pp: 455-462.

Dawson, W., Wilby, R. (1998). An artificial neural network approach to rainfall-runoff modeling Hydrological Sciences Journal, Volume, Issue 1, pp. 47-66.

Abbot, J., Marohasy, J., 2012. Application of artificial neural networks to rain forecasting in Queensland, Australia. Adv. Atmos. Sci. 29, 717–73.

S. Afshin, H. Fahmi, A. Alizadeh, H. Sedghi, F. Kaveh, Long term rainfall forecasting by integrated artificial neural network fuzzy logic-wavelet model in Karoon basin, Sci. Res. Essays 6 (6) (2011) 1200–1208.

Nastos, P.T., et al., 2013. Rain intensity forecast using artificial neural networks in Athens, Greece.Atmos. Res. 119, 153–160.

Singh, P., Borah, B., 2013. Indian summer monsoon rainfall prediction using artificial neural network. Stoch. Environ. Res. Risk Assess. 27, 1585–1599.

Wu, C.L., Chau, K.W. (2011).Rainfall–runoff modeling using artificial neural network coupled with singular spectrum analysis, Journal of Hydrology, vol, 399, pp. 394–409.