Irrigation Sciences and Engineering Vol. 41, No. 2, Summer 2018, p183-195

Shahid Chamran University of Ahvaz

EXTENDED ABSTRACT

Application of Shannon Entropy for Selecting the Optimum input Variables in River Flow Simulation using Intelligent Models (Case Study: SofyChay)

F. Akhoni Pourhosseini[1*] and M.A. Ghorbani[2]

1*- Corresponding Author, Ms.c student of Water Resources Engineering, University of Tabriz, Iran*(fateme.pourhosseini@yahoo.com).*

2- Associate Professor, University of Tabriz, Iran.

Received:1 August 2016 Accepted:1 March 2017

Keywords: Artificial Neural Network, Bayesian Network, Entropy, river flow, Sofychay River.

Introduction

Accurate prediction of the river flow is an important element in the management of surface water resources, dam reservoir operation, flood control and drought. Selecting appropriate inputs for intelligent models is vital to increase the accuracy and efficiency of the models. Since river flow prediction is of great importance in water resources, researchers have been exploring different approaches over the past several decades. Various methods have been devised to predict the flow of the river over the past years. In general, we can classify conceptual models and data-driven methods. Over the past four decades, time series models have been widely used in river flow prediction (Dawson et al., 2008). Intelligent systems are used to predict nonlinear phenomena. The Bayesian Network and the Artificial Neural Network are among these methods. Ahmadi et al. (2014) studied the comparison of performance of support vector machine and network methods in forecasting daily flow of the Barandozachay River. The results showed that, both methods are close to each other and are suitable for river flow simulation. But in mid-range forecasting and the minimum backup car model, it's much better than the business network model. Shannon entropy theory was first developed by Shannon and then widely used in various scientific issues.

The purpose of this study is to use the Shannon Entropy Theory to find the best combination of input variables for artificial neural network and Bayesian network models to predict the flow. Therefore, for this purpose, the Sufi River of the studied area was selected.

Methodolog

The Sufychy River is a 120 km long river from the southeastern area of Sahand Mount Sahand, East Azarbaijan Province. In this research, the Tazekan station information is used. In order to estimate the flow of the next month, the flow of the previous months is used. Using the time series correlation matrix with a delay of six months, input to the Shannon entropy was introduced. Shannon (1948) showed that events with high probability of occurrence shows less information. Conversely, the more likely it is to occur, the more information it receives. New information reduces uncertainties. Entropy theory is used as an indicator for quantifying relevant items (Masoumi, 2007, Kerachiyan, and others).

Bayesian network, a probabilistic graph pattern, consistedof a set of variables and a corresponding probability for each one. A direct and non-circular graph in which nodes represent the variables of the problem and in the Bayesian network, in addition to the quality of the relationship between the variables of the problem,An artificial neural network receives signals

from input units, and these signals are embedded over the network and eventually move to the output neuron. The main controlling parameters of artificial neural network models are connective resistances between neurons, which are referred to as weight and inertia.
In order to compare the observational and predictive values, the R^2, RMSE and SI scores were used (Ghorbani and Dehgani 2015).

$$R^2 = \left[\frac{\sum_N^{i=1}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_N^{i=1}(x_i-\bar{x})^2 \cdot \sum_N^{i=1}(y_i-\bar{y})^2}}\right]^2 \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(x_i-y_i)^2}{N}} \quad (2)$$

$$SI = \frac{\frac{1}{n}\sum_{i=1}^{n}|x_i-y_i|}{\bar{x}} \quad (3)$$

Results and Discussion

In the present study, according to Fig. 1, it can be concluded that the effects of $Q_{(t-1)}$, $Q_{(t-2)}$ and $Q_{(t-3)}$ on the output $Q_{(t-6)}$, relative to other parameters, respectively, and can be introduced as an optimal combination of input parameters to models.

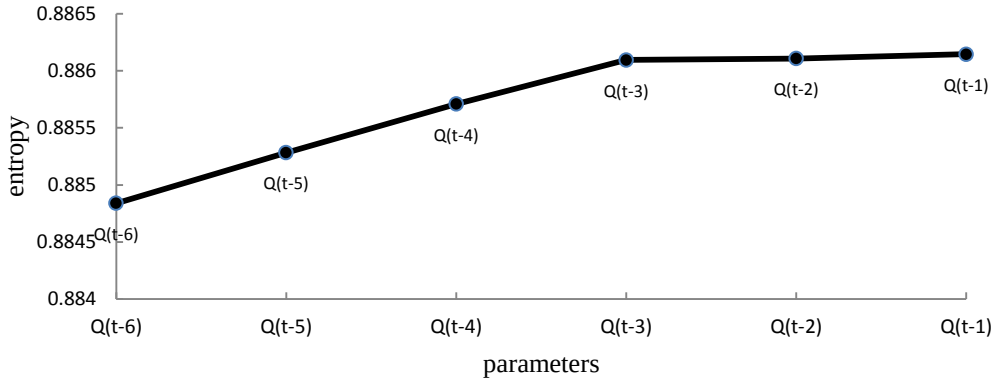

Figure 1- Entropy values of the parameters used at the station

The level of R^2 in the training sector is the third largest combination of the Bayesian network model, equivalent to 0.597, which is larger than the previous ones. Also, the dispersion index is SI = 5.485 and the root mean square error is RMSE = 3.421(m3 / s0.
The results of this study suggest an artificial neural network as a precise method for predicting the flow of the SofyChay River. Bohrani, Fatehi(2008) In predicting the flow of Nazluchay watershed, one of the sub-basins of Lake Urmia, the artificial neural network with climate indicators one year ago was able to predict the flow of one year with acceptable accuracy.

Discussion

River flow simulation is one of the important components in water resources management and flood forecasting. On the other hand, the proper combination of simulation is one of the factors influencing modeling. In this study, Shannon entropy was used to select the optimum input variables. Monthly river flow of SofyChay at the period of 1973-2012 was forecasted using Artificial Neural Networks and Bayesian Neural Network in Tazekand hydrometric station. Information entropy is used to quantify the amount of information in a data set. Shannon entropy used as a measure for quantifying the absence of knowledge about the characteristics of a system.

The monthly river flow data of Sofychay River with different antecedent are considered as input to the Shannon entropy. Shannon entropy results showed that time series with three antecedents provides better results to predict. Comparing the performance of ANN and BN models indicated that the performance of ANN method was better than BN method in predicting. In addition, the results showed that the Shannon entropy has better performance in selection of Optimum Combination of Variables into intelligent models.

References

1- Ahmadi, F., Dinpazhoh, Y. Fakherifard, A. and Darbandi, S., 2013. Comparing nonlinear time series models and genetic programming for daily river flow forecasting (Case study: Barandouz-Chai River), *Journal of Water and Soil Conservation, 22(1),* pp.151-169 (In Persian).

2- Bohrani, A. and Fatehi, A., 2008. Application of Artificial Neural Network in Stream Flow Forecasting using Climatic Indices (Case Study: Nazloochay River Basin), *Journal of Civil and Environmental Engineering University of Tabriz, 35(3),*pp. 51-62.(In Persian).

3- Dawson, C.W. Abrahart, R.J. Shamseldin, A.Y. Wibly, R.L. 2006. Flood estimation at ungauged sites uzingartifitial neural networks. *Journal of Hydrology. 319*,pp. 391-409.

4- Ghorbani, M.A.,and Dehgani, R.,2015. Application of Bayesian Neural Networks, Support Vector Machines and Gene Expression Programming Analysis of Rainfall - Runoff Monthly (Case Study: Kakarza River). *Journal of Irrigation Sciences and Engineering, 20. 39(2),* pp.125-138.(In Persian).

5- Masoumi, F. and Kerachiyan, R., 2007.Optimal Design of Groundwater Groundwater Quality Monitoring Systems Using Discrete Entropy Theory Case Study: Tehran Aquife, *Third National Congress of Civil Engineering, Shahid Beheshti,Tehran.*

6- Shannon, CE. 1948. A mathematical theory of communications, I and II. Bell Sys. Technol. *J. 27: 379-423.*