Hydrogeomorphology, Vol.6, No.22, Spring 2020, pp (30-33)

Received: 2019.12.27
Accepted: 2020.05.06

Investigating the Application of Hybrid Support Vector Machine Models in Predicting River Flow of Karkhe Basin

Reza Dehghani[1], Hasan Torabi Poodeh[*2], Hojjatollah Younesi[3], Babak Shahinejad[4]

1- Ph.D student in water structures, Lorestan University
2- Associate Professor, Department of Water Engineering, Lorestan University
3- Assistant Professor, Department of Water Engineering, Lorestan University
4- Assistant Professor, Department of Water Engineering, Lorestan University

Abstract

1- Introduction

River flow forecasting is one of the most important issues in water resources management and planning, especially in making the right decisions in the event of floods and droughts. Various approaches to hydrology have been introduced to predict river flow rates, among which, intelligent models are the most important ones. In this study, daily data of Karkheh catchment was used to evaluate the accuracy of models in river flow prediction. Daily flow modeling of Karkheh catchment rivers including Cham Anjir, Madianrud, Afrineh, Kashkan, Pol Zal, Jologogir and support vector-wavelet and back-vector-Bayesian models were used and the results were compared to verify the studied models . Few studies have investigated each of the mentioned models in predicting daily flow discharge, but the purpose of this study was to investigate these models simultaneously in a basin to predict daily river flow.

2- Methodology

In this study, the rivers of Karkheh Abriz Basin were selected as the study area and daily observational flow of this basin was used for calibration and validation of the models at Cham Anjir, Madianrud, Afrineh, Kashkan, Paul Zal, Jologir upstream stations. For this purpose,

* (Corresponding Author), **E-mail:**torabi.ha@lu.ac.ir

80% of daily flow data (2008-2016) were selected for the calibration of models and 20% data (2016-2018) was utilized for model validation. Backup vector machine is an efficient learning system based on bound theory of optimization. The Bayesian network is also a meaningful representation of the uncertain relationships between parameters in a process and is a non-polarized directed graph of nodes to represent random variables and bows to represent possible relationships between variables. Correlation coefficient, root mean square error, absolute mean error for evaluation and comparison of model performance were used in this study. Moreover, the Basin network is a meaningful representation of our uncertain relationships between parameters in a process, and a non-circular directional graph of nodes for displaying random variables and arcs to represent potential relationships between variables. The correlation coefficients, root mean square error, mean absolute error was used for evaluation and also comparison of the performance of models in this research.

Support Vector Machine used for Classification is called SVC and has been successfully used for many applications concerning the separation of data into two or several classes. The aim of using SVC is to find a classification criterion (i.e., a decision function) which can properly separate unseen data with a good generalization ability at the testing stage. This criterion, for a two-class data classification, can be a linear straight line with a maximum distance (margin) from the data of each class. This linear classifier is also known as an optimal hyperplane in SVC related discussions.

The wavelet transform has been proposed as an alternative to short-time Fourier transform and its purpose is to overcome the problems related to the frequency resolution power in short-time Fourier transform. In the wavelet transform, as in the short-time Fourier transform, the signal is divided into windows and the wavelet transform is performed on each of these windows, separately (Vapnik, 1998). A wavelet means a small wave, part or window of the main signal, whose energy is concentrated in time. Using a wavelet transform or analysis, a mother signal or time series can be broken down into wavelets with different levels and scales.

Bayesian networks are graphical models that are used to argue when there is complexity and uncertainty, or they are utilized in a graph that represents random variables and their dependencies (Kevin and Nicholson, 2010). In this graph, the nodes represent discrete or continuous random variables, and the orientation arcs that connect each pair of nodes to each other to represent the dependency between the variables. In fact, this grid is a graph of orientation in which there is no distance (Heckerman, 1997).

3- Results and discussion

The results showed that all three models had better results in the structures of 1 to 4 delay times than other specified structures. Moreover, comparison of the models showed that the hybrid model of support-wavelet vector machine had a better performance in flow forecasting. Overall, the results showed that using a hybrid backup vector machine model can be useful in predicting daily discharge.

4- conclusion

The results showed that an increase in the number of effective parameters in different models for simulation resulted in better performance in the discharge estimation. In addition, the results showed that the hybrid Support Vector Machine model had a better performance

Keywords: Bayesian Network, Support Vector Machine, Wavelet, Karkhe Basin.

5- References

Hekerman, D. (1997). Bayesian Networks for Data Mining, *Data Mining and Knowledge Discovery*, 1(1), 79-119.

Korb, K.B., Nicholson, A.E. (2010). *Bayesian artificial intelligence*, second edition. CRC Press, 360-452.

Shin, S., Kyung, D., Lee, S., Taik & Kim, J., and Hyun, J. (2005). An application of support vector machines in bankruptcy prediction model, *Expert Systems with Applications*, 28(4), 127-135.

Vapnik, VN. (1995). The Nature of Statistical Learning Theory. Springer, New York

Vapnik, VN. (1998). *Statistical learning theory*. Wiley, Ne.