Amirkabir Journal of Civil Engineering

Amirkabir J. Civil Eng., 54(11) (2023) 891-892
DOI: 10.22060/ceej.2022.21208.7653

Uncertainty Analysis of Artificial Intelligence Models in Forecasting River Flow (Case Study: Karun River)

Y. Mehdizadeh Zare-Anari [1], M. Najafzadeh [1], S. Anvari [2*]

[1] Graduate Student, Water Engineering Department, Faculty of Civil and Surveying Engineering, Graduate University of Advanced Technology, Kerman, Iran

[2] Assistant Professor, Department of Ecology, Institute of Science and High Technology and Environmental Sciences, Graduate University of Advanced Technology, Kerman, Iran

ABSTRACT: An accurate estimation of the discharge flow of natural streams plays a key role in irrigation planning, the design of bridges embedded in waterways, the management of reservoirs and dams, and the design of flood-warning systems. In recent decades, several studies have been conducted using artificial intelligence (AI) to accurately estimate the flow. Despite the proven accuracy level of AI methods, in many cases, there are uncertainties that occur for a variety of reasons. Insufficient knowledge of these uncertainties in the flow modeling process can have irreversible effects. In this research, monthly flow data, measured from Armand hydrometric station, located in Karun River basin, during a-28 year period (from 1980 to 2008) were used. First, the Thomas and Fiering method was used to generate flow series data and consider them as input variables. Then, flow forecast modeling was performed by three AI methods, namely Model Tree (MT), Gene Expression Programming (GEP), and Multivariate Adaptive Regression Spline (MARS). Statistical indicators such as correlation coefficient (R) and Root Mean Square Error (RMSE) were used to evaluate the accuracy of the models. In terms of the training stage, the MARS model with R=0.839 and RMSE=28.624 m3/s performed better than the other two models. Additionally, in the testing stage, the MT model with R=0.784 and RMSE=34.441 m3/s showed a more appropriate performance than other models. The Monte Carlo simulation method was used to calculate the uncertainty of the models, so the results showed that the r-factor parameter, which was the average width of the confidence band in the MT model, was equal to 1.67, indicating a lower and more optimal number compared to MARS (1.92) and GEP (2.025) models. Moreover, the usability of the statistical criterion for quantifying uncertainty, (known as 95PPU) indicated that the GEP model with 95PPU of 64% was selected as a more appropriate percentage than MARS (61%) and MT-M5 (55%).

Review History:

Received: Mar. 12, 2022
Revised: Jun. 27, 2022
Accepted: Jul. 11, 2022
Available Online: Aug. 09, 2022

Keywords:

Flow prediction
Artificial intelligence models
Uncertainty analysis
Monte-Carlo technique

1- Introduction

Population growth and increasing demand for water and food products, on the one hand, as well as the limitation of available freshwater resources, have made water resources management even more necessary. In this regard, improving the accuracy of river flow forecasts and taking into account its uncertainties play a key role in managing water resources [1-3]. With the advent of data-driven artificial intelligence models, a wide range of artificial intelligence (AI) techniques has been used to predict river flows. Artificial intelligence methods have shown the sufficient capacity to identify hidden patterns in time series and have proven high reliability compared to statistical methods [1-4]

The mentioned research as well as other studies carried out in the field of river flow forecasts show that although AI models have been successfully used in the field of flow forecasting in recent decades, however, like other hydrological models of basic physics, has many uncertainties, which mainly originate from the climatic-hydrological factors of the catchment area. Therefore, it is necessary to consider and quantify the uncertainties associated with these factors to improve the prediction results and ultimately manage water resources as best as possible. In this regard, Abbasi et al. [1] used Support Vector Regression (SVR) and Multiple Linear Regression (MLR) models to predict the inflow to Bukan Dam. They also used three approaches: random forest (RF), deep automatic encoder (DAE) and principal component analysis (PCA) to preprocess the input variables. The results showed that the DAE-SVR and RF-SVR models compared to the PCA-MLR, RF-MLR, DAE-MLR, PCA-SVR hybrid models had a significant improvement in the accuracy of forecasting one month later. Due to the superior performance of three robust AI models (namely multivariate adaptive regression spline [MARS], gene-expression programming [GEP], and model tree [MT]) in comparison with other AI models, no research has been done by these models to predict the flow

*Corresponding author's email: s.anvari@kgut.ac.ir

Copyrights for this article are retained by the author(s) with publishing rights granted to Amirkabir University Press. The content of this article is subject to the terms and conditions of the Creative Commons Attribution 4.0 International (CC-BY-NC 4.0) License. For more information, please visit https://www.creativecommons.org/licenses/by-nc/4.0/legalcode.

of the Karun River. Therefore, the main goal of the current research is to predict the flow of this river (at the Armand hydrometric station), and to investigate the uncertainty of these predictions using the Monte Carlo method.

2- Case study

This basin has an approximate area of 67,257 square kilometers and includes the permanent Karun River with a length of 890 kilometers and an average annual discharge of 18,700 million cubic meters. In this research, the Karon river basin has been investigated up to the location of the Armand hydrometric station. This station is located near the Karon 4 dam with an approximate height of 1082 meters, and its latitude and longitude are 46-50 and It is 31-40. The upstream basin of this dam has an area of 11,000 square kilometers and its annual average water discharge is 110 cubic meters per second. The average annual rainfall of this basin is approximately 750 mm, and its maximum and minimum height above the open water level is 4141 and 1044 meters, respectively. It should be mentioned that the data used in this research is the 28-year time series of the daily flow of Armand hydrometric station for the years 1981 to 2008. The highest amount of river flow is related to the months of May (478.86), March (403.39) and Farvardin (356 cubic meters per second), which is due to the rains of the previous months as well as runoff from snow melting. On the other hand, the lowest amount of discharge belongs to Shahrivar (21.21), Mehr (21.33) and March (23.05). Also, 70 percent of the data (years 1981 to 2000) were used for training and the rest, i.e. the statistics of the years 2001 to 2008, were used to test MARS, GEP and MT models. The input variables include river flow values for the previous one to three months [Q(t-1), Q(t-2), Q(t-3)].

3- Results and discussion

This section evaluated the performance of AI models by correlation coefficient (R) and root mean square error (RMSE). The statistical measures for the training stage of the MT and MARS models indicated approximately equal performance (R=0.841 and RMSE=36.789 m^3/s for MT and R=0.841 and RMSE=36.765 m^3/s for MARS). The GEP model stood at the third level of efficiency. Additionally, in the testing stage, statistical indices by MT were equal to R=0.87 and RMSE=44.253 m^3/s in compared with GEP (R=0.865 and RMSE=46.243 m^3/s) and MARS (R=0.848 and RMSE=47.43 m^3/s).

Additionally, the uncertainty results of MARS, GEP and MT models for predicting the flow of the Karun River have been calculated and quantified. Generally, smaller values of R-factor, which represents the uncertainty bandwidth, indicate good performance of the model. From this point of view, the MT model has the smallest R-factor (1.67) value compared to MARS (R-factor=1.92) and GEP MARS (R-factor=2.025). On the other hand, the 95PPU values that represent the percentage of definitive data (historical observations) are in the 95% uncertainty estimation band, for the mentioned models, it is between 59.5% and 64%. Finally, to judge the performance of the best model under conditions of uncertainty, both the minimum R-factor index and the maximum 95 PPU should be taken into consideration at the same time, and according to the mentioned cases, it can be said that the performance of the MT model under conditions of uncertainty is better than other models.

4- Conclusions

The present research is an attempt to predict the flow of the Karun River one month later using MARS, GEP and MT artificial intelligence models. Also, calculating the uncertainty of the river flow values by the Monte Carlo simulation method (MCS) and its effect on the uncertainty bands of the flow predictions by the mentioned models are other goals of this research. Results of AI techniques demonstrated that MT had the best performance in the prediction of flow discharge. Moreover, the performance of uncertainty analysis showed that MT had a promomising evaluation than GEP and MARS.

References

[1] M. Abbasi, A. Farokhnia, M, Bahreinimotlagh, R, Roozbahani. A Hybrid of Random Forest and Deep Auto-Encoder with Support Vector Re-gression Methods for Accuracy improvement and uncertainty reduction of long-term streamflow prediction. Journal of Hydrology. 597 (2021) 125717.

[2] F. Azarpira, S. Shahabi. Evaluating the capability of hybrid data-driven approaches to forecast monthly streamflow using hydrometric and meteorological variables. Journal of Hydroinformatics. 23(6) (2021) 1165-1181.

[3] Z.M. Yaseen, A. El-Shafie, O. Jaafar, H.A. Afan, H. A., K. N. Sayl. Artificial intelligence based models for stream-flow forecasting: 2000–2015. Journal of Hydrology. 530 (2015) 829-844.

[4] R. Noori, A.R. Karbassi, A. Moghaddamnia, D. Han, M.H. Zokaei-Ashtiani, A. Farokhnia, M. Ghafari Gousheh. Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction. Journal of Hydrology. 401 (2011) 177-189.

HOW TO CITE THIS ARTICLE
Y. Mehdizadeh Zare-Anari, M. Najafzadeh, S. Anvari, Uncertainty Analysis of Artificial Intelligence Models in Forecasting River Flow (Case Study: Karun River), Amirkabir J. Civil Eng., 54(11) (2023) 891-892.

DOI: 10.22060/ceej.2022.21208.7653