THE SCIENCE WORLD
A Monthly e Magazine
ISSN:2583-2212
February 2024 Vol.4(2), 752-756

Popular Article

Machine learning models for daily river discharge prediction

S. B. Tarate[1], S. M. Raut[1], I. S. Singh[1], Manoj Kumar[1] and Vinod Kumar Padala[1]
[1]ICAR-National Research Centre for Makhana, Darbhanga – 846005, Bihar, India

https://doi.org/10.5281/zenodo.10702750

Introduction

Water resources are crucial for sustainable development, and their effective management is essential (Guldal and Tongal, 2010). Given the spatio-temporal variability of rainfall, it is imperative to focus on flood and drought management. River discharge prediction is vital for mitigating the risks associated with these hazards, enabling both structural and non-structural measures to be implemented. Accurate rainfall-discharge modeling is valuable for optimizing water utilization, enhancing flood forecasts, and facilitating drought management. Various models have been suggested to understand the rainfall-runoff dynamics within a watershed's fluvial system (Agarwal *et al.* 2006). These models fall into distinct categories such as physical, conceptual, stochastic, and data-driven models (Senaviratne *et al.* 2014). Data-driven models like Artificial Neural Network (ANN), Adaptive Neuro-Fuzzy Inference System (ANFIS) etc. demonstrate real-time capabilities in predicting hydrological phenomena.

Keywords: Machine learning, ANN, ANFIS, Wavelet transform, river discharge prediction

Artificial Neural Network (ANN)

ANN is a biologically inspired computational system; its behavior depends on a large number of processing elements (neurons), which are interconnected in such settings to pass some information (Yin *et al.* 2011). ANN simulates the biological neuron system of the human brain. ANNs generalization ability is the learning from input and associated output data; this is the basic idea behind the working of an ANN. Neural networks have enough capability to simulate non-

Published 25.02.2024

linear complex processes. ANN is a very inexpensive tool for prediction as it performs its function even with slow computers. The ANNs can learn from input-output pairs, self-organize and interactively adapt their structure. The ANN models can perform better without having prior knowledge of the system and can be applied to solve the problem of such a system which is not clearly understood. A mathematical model needs precise information of all contributing variables but trained ANN could simulate process behavior even in the presence of incomplete information. These models require less data for development of models as these models have more tolerance to incomplete and noisy data. The outcome is due to the collective behavior of the whole data and hence minimizes the effect caused due to disturbed data if any (Agarwal *et al.* 2006). With the advancement in the performance of computers, the ANN theory has been widely applied and studied during the past 30 years. ANN is considered a self-learning and self-adaptive function approximator with great potential to model and predict non-linear hydrological time series (Nourani and Komasi, 2013). Elsafi (2014) predicted the overflow of the Nile River in Sudan, resulting in floods in nearby areas, and concluded that ANN proved to be a dependable tool for forecasting flood hazards along the Nile River. Multi-layer perceptron (MLP) based ANN architecture with one hidden layer is shown in Fig. 1.

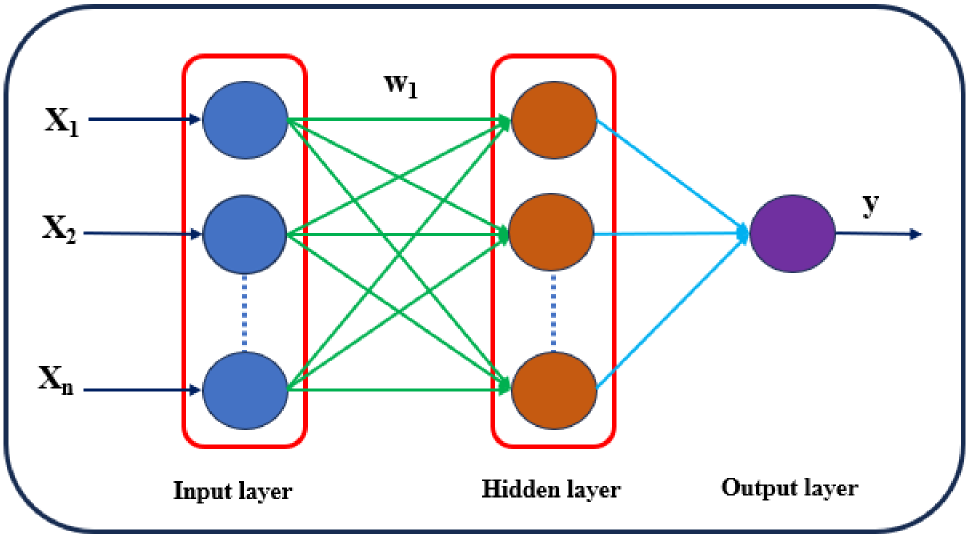

Fig. 1 ANN architecture with one hidden layer

Adaptive Neuro-Fuzzy Inference System (ANFIS)

When modeling river discharge using time-series data, uncertainties are common. To address this, fuzzy theory can be applied to handle uncertainties and solve real-world problems. The ANFIS is created by combining ANN and fuzzy logic, benefiting from both approaches (Seo *et al.* 2015). Unlike the ANN model, which requires a trial-and-error process to determine the optimal network architecture, the ANFIS model avoids this tedious procedure. ANFIS is a multilayer self-organizing network structure that adjusts the parameters of the fuzzy system to predict the system's output. The Fuzzy Inference System (FIS) is a popular computing framework based on fuzzy set theory. It relies on fuzzy operators, fuzzy sets, and a knowledge base as its foundational elements. FIS is capable of extracting fuzzy rules either from expert knowledge or numerical data, constructing a rule base adaptively. It can effectively handle the intricate translation of human intelligence into fuzzy systems. Fuzzy theory is valuable for modeling uncertainties in real-world problem-solving. In predicting peak river discharge flow, the ANFIS model outperforms the ANN model in terms of accuracy, prediction error, and computational efficiency (Nourani and Komasi, 2013). A general ANFIS structure with two rules, two inputs, and one output is shown in Fig. 2.

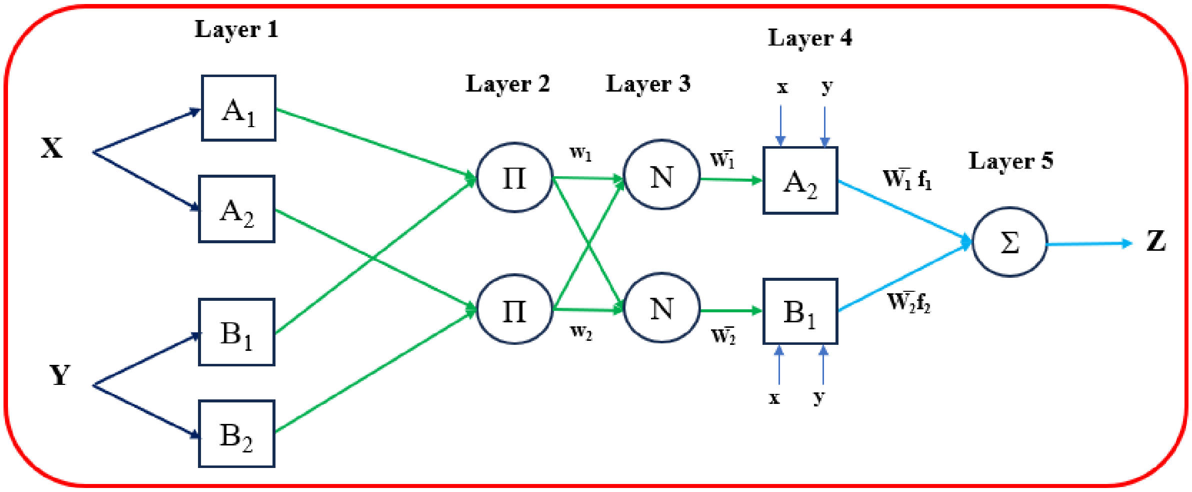

Fig. 2 General ANFIS structure with two rules, two inputs, and one output

Wavelet coupled ANN and ANFIS

Typically, the Fourier transform is used for examining static data, but hydrological time series signals are often non-stationary. The challenge with Fourier transforms in modeling non-stationary hydrological data is that they only focus on the central frequency feature of the time series. This makes them suitable only for static time series. To address this limitation, the wavelet transform can be employed. It proves effective when dealing with non-stationary data in modeling. The wavelet transform generates different series that highlight potential trends, seasonal variations, and internal correlations among various components. Unlike Fourier transform, wavelet transform breaks down the original time series data into distinct sub-components without losing any information. It extracts necessary details from historical data using only a few coefficients, unveiling hidden information within the dataset. This method provides a more condensed representation of the original time series data, offering a timescale view of relationships between processes. Consequently, data-driven models using wavelet-transformed sub-components enhance performance by capturing essential information present in the central time series hydrological data.

Ramadevi *et al.* (2019) developed data-driven models like autoregressive integrated moving average (ARIMA) (i.e. linear) and wavelet coupled ANN (i.e. non-linear) models for rainfall-runoff modeling of the Sarada River basin situated in the Andhra Pradesh state of India. The findings revealed that the wavelet-coupled ANN models outperformed both ARIMA and simple ANN models in terms of prediction accuracy. Solgi *et al.* (2017) predicted daily and monthly river flow of the Gamasiyab River in Iran by combining wavelet theory with ANN and ANFIS. The outcomes showed that the wavelet-coupled hybrid ANFIS and ANN models performed better than the simpler ANFIS and ANN models, both for daily and monthly scales.

Conclusion

Artificial intelligence-based data-driven machine learning models outperform traditional statistical models in predicting river discharge. These models excel in capturing the complex, non-linear patterns in time series data. Researchers are highly interested in these models due to their simplicity, minimal data needs, and efficient time performance when modeling non-linear aspects of time series data. The hybrid wavelet-coupled soft computing models, in particular, show improved predictive performance compared to simple soft computing models. This is because wavelet-coupled models can effectively identify sudden changes, periods, and trends in the original time series data. However, it's essential to validate the performance of these hybrid models under various climatic and geographic conditions for daily river discharge prediction.

References

Agarwal, A., Mishra, S. K., Ram, S. and Singh, J. K. (2006). Simulation of runoff and sediment yield using artificial neural networks. *Biosystems Engineering,* **94**(4):597-613

Elsafi, S. H. (2014) Artificial neural networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal*, **53**:655–662

Guldal, V. and Tongal, H. (2010). Comparison of recurrent neural network, adaptive neuro-fuzzy inference system and stochastic models in Egirdir lake level forecasting. *Water Resources Management*, **24**:105–128

Nourani, V. and Komasi, M. (2013). A geomorphology-based ANFIS model for multi-station modeling of rainfall-runoff process. *Journal of Hydrology,* **490**:41-55

Ramadevi, K. N. V., Ramana, R. V., Rao, Y. R. S. and Kumar, S. (2019). Development of data driven rainfall-runoff model for the Sarada River basin. *International Journal of Recent Technology and Engineering*, **7**(6C2):508-512

Senaviratne G. M. M. M. A., Udawatta, R. P., Anderson, S. H., Baffaut, C. and Thomp-son, A. (2014). Use of fuzzy rainfall–runoff predictions for claypan watersheds with conservation buffers in Northeast Missouri. *Journal of Hydrology*, **517**:1008–1018

Solgi, A., Zarei, H., Nourani, V. and Bahmani, R. (2017). A new approach to flow simulation using hybrid models. *Applied Water Science*, **7**:3691-3706

Seo, Y., Kim, S., Kisi, O. and Singh, V. P. (2015). Daily water level forecasting using wavelet decomposition and artificial intelligence techniques. *Journal of Hydrology*, **520**:224–243

Yin, D., Shu, L., Chen, X., Wang, Z. and Mohammed, M. E. (2011). Assessment of sustainable yield of Karst water in Huaibei, China. *Water resource management*, **25**:287-300