GSC Advanced Research and Reviews

eISSN: 2582-4597 CODEN (USA): GARRC2

Cross Ref DOI: 10.30574/gscarr

Journal homepage: https://gsconlinepress.com/journals/gscarr/

GSC Advanced Research and Reviews

GSC Online Press INDIA

(Review Article)

Check for updates

An overview of streamflow prediction using random forest algorithm

M.M Jibril [1,*], Aliyu Bello [2], Ismail I Aminu [3], Awaisu Shafiu Ibrahim [4], Abba Bashir [4], Salim Idris Malami [1,5], Habibu M.A [6] and Mohammed Mukhtar Magaji [7]

[1] *Department of Civil Engineering, Faculty of Engineering, Kano University of science and technology, KUST, Wudil, Nigeria.*
[2] *Department of Civil Engineering, Faculty of Engineering Baze University, Abuja, Nigeria.*
[3] *Department of Civil Engineering, Faculty of Engineering Bayero University, Kano, Nigeria.*
[4] *Department of Civil Engineering, Federal University, Dutsin-ma, Katsina, Nigeria.*
[5] *Department of Civil Engineering, Institute for Infrastructure & Environment, Heriot-Watt University, Edinburgh, UK.*
[6] *Department of Electrical Engineering, Federal University Gusau, Nigeria.*
[7] *Department of Information and Communication Technology (ICT), Yusuf Maitama Sule University Kano, Nigeria.*

GSC Advanced Research and Reviews, 2022, 13(01), 050–057

Publication history: Received on 14 April 2022; revised on 06 October 2022; accepted on 09 October 2022

Article DOI: https://doi.org/10.30574/gscarr.2022.13.1.0112

Abstract

Since the first application of Artificial Intelligence in the field of hydrology, there has been a great deal of interest in exploring aspects of future enhancements to hydrology. This is evidenced by the increasing number of relevant publications published. Random forests (RF) are supervised machine learning algorithms that have lately gained popularity in water resource applications. It has been used in a variety of water resource research domains, including discharge simulation. Random forest could be an alternate approach to physical and conceptual hydrological models for large-scale hazard assessment in various catchments due to its inexpensive setup and operation costs. Existing applications, however, are usually limited to the implementation of Breiman's original algorithm for extrapolation and categorization issues, even though several developments could be useful in handling a variety of practical challenges in the water sector. In this section, we introduce RF and its variants for working water scientists, as well as examine related concepts and techniques that have earned less attention from the water science and hydrologic communities. In doing so, we examine RF applications in water resources, including streamflow prediction, emphasize the capability of the original algorithm and its extensions, and identify the level of RF exploitation in a variety of applications.

Keywords: Artificial intelligent; Random Forest; Streamflow; Machine learning

1. Introduction

Bregman's Random forest was introduced in (2001), is an ensemble machine learning approach that predicts using a large number of classification or regression trees (CART) [1] and Because of its high stability and generality, it has been widely used in a variety of fields. [2] In this scenario, the response variable, the number of flood reports per occurrence, is modelled using regression; thus, the Random Forest model is an ensemble of regression trees that may be used in a variety of fields., including land subsidence, invasive plant, groundwater, gully head susceptibility, and forest fire susceptibility[3]. During regression tree training, rules based on the response variable are constructed to partition observations until the resulting predictions have a minimum level of node impurity, and the majority of results collected from decision trees are regarded as the RF's final output. Breiman describes node impurity for regression trees as the total of the squared deviations between the expected and observed values.

* Corresponding author: MM Jibril
Department of Civil Engineering, Faculty of Engineering, Kano University of science and technology, KUST, Wudil, Nigeria.

During the last decades, a model called Artificial Intelligence (AI) particularly neural network (NN) or artificial neural networks (ANN) emerged and becomes well-known in the various field of science and engineering. ANN mimic how the human brain is functioning with regard to the complex interactions, pattern recognition, classification, and perception [4]–[8]. More research on machine learning can be found in [7], [9]–[15].

[1]–[3]. A single decision tree is well known for having a high variance, being susceptible to noise, and demonstrating statistical instability. Bootstrap aggregation is used to generate numerous decision trees by randomly selecting the observed dataset with replacement to reduce over-fitting and statistical instability. The RF study can be applied bootstrap aggregation not only by leveraging a portion of input data but also by exploiting randomly chosen input variables or features for tree node splitting. [16], [17] Random forest predicts the mean output for new input data using the set of observed input–output data for training. [18]. RF cannot be overfit when compared to other conventional statistical techniques, and they are highly helpful when there are few sample sites [17] and several prospective forecasters. While dealing with a high number of predictor variables, random forest is useful. During the RF model simulation procedure, two critical parameters must be determined: the number of trees included in the forest (n_{tree}) and the number of predictors assessed at each node (m_{try}). shown below [2].

$$\begin{cases} m_{try} = log_2(M+1) \\ m_{try} = \sqrt{M} \\ m_{try} = \frac{M}{3} \end{cases} \quad \text{...} \quad (1)$$

Where M denotes the number of input variables specified in the original dataset

Figure 1 shows the process for developing random forest models and applying them to the stream network. [19]. For more understanding of data-driven approaches refer to [20], [21], [30]–[33], [22]–[29].

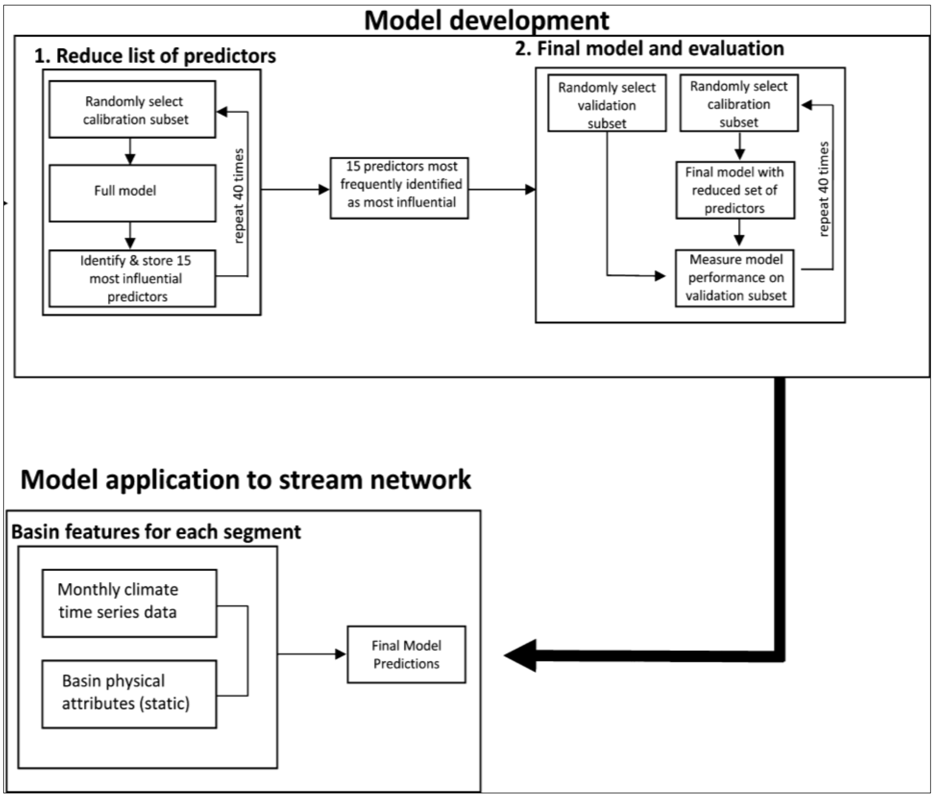

Figure 1 Methodology for random forest model development, and model application to the stream network

2. An overview of the related literature

Miller et al. [19] use a statistical machine learning technique— random forest modelling— to estimate natural flows at monthly time intervals from 1950 to 2015, In furthermore, he achieved an overall Nash–Sutcliffe efficiency of0.85; observed/expected ratio=0.94, indicating good consistency between anticipated and observed flows at approximately 2,000 stream gages. Sha et al.[2] Adopt five models to compare daily streamflow estimates, namely: extreme learning machine (basic ELM), extreme learning machine with kernels (ELM-kernel), random forest (RF), back-propagation neural network (BPNN), and support vector machine (SVR). The results reveal that the ELM-kernel model performed better than the other models, and the basic ELM model performed the worst. The RF model performed marginally better than the other models in predicting peak flows, while the ELM-kernel model performed the best in predicting low flows. Sadler et al.[1] Employs two data-driven models, Poisson regression and Random Forest regression, which have been trained to predict the frequency of flood reports per storm event as a proxy for flood severity given extensive environmental data (i.e., rainfall, tide, groundwater table level, and wind conditions).

Tongal et al.[16] Simulate and forecast streamflow using Support Vector Regression (SVR), Artificial Neural Networks (ANNs), and Random Forest (RF) as functions of precipitation (P), temperature (T), and potential evapotranspiration (PET). Schoppa et al.[18] use random forest to evaluate the performance of a large-scale flood discharge simulation, compare the predictive performance of random forest to the conceptual hydrological modelling package 'hydromad,' and measure the impact of catchment characteristics on model performance. His study discovered that random forest outperforms hydromad in the modelling of low and medium flood magnitudes. Cid et al. [17] identify crucial biological factors for identifying the aquatic state as flowing or unconnected pools using random forest and classification tree analysis.

Amare et al.[3] Used frequency ratio (FR) and random forest (RF) algorithms to predict gully susceptibility. The results showed that using the top four most important gully predictor factors: drainage density, elevation, land use, and groundwater table provided the best prediction accuracy using the FR and RF models. Peng et al.[34] Predict monthly streamflow using random forest (RF), BP neural network (BPnet), and traditional support vector machine (SVM) models, which optimize parameters that used a grid algorithm. His results show that the random forest model has higher prediction accuracy and necessitates less calculation while dealing with complex nonlinear hydrological time series. Fan et al.[35] Random Forest regression was used to analyze the relationships between air temperature, precipitation, and streamflow changes.

Li et al.[36] use elastic net regression (ENR), support vector regression (SVR), random forest (RF), and extreme Gradient Boosting (XGB) models and propose a modified multi-model integration method named a modified stacking ensemble strategy (MSES) for monthly streamflow prediction. The methods were applied to the Three Gorges Reservoir in the Yangtze River Basin, and the results show that RF and XGB provide better and more stable forecast performance than ENR.

3. Results and discussion

References	Case study	AI model	Data scale	Research remark	Performance metrics
Peng et al.[34]	Jinsha river, China	Random forest, BPNN, SVM	1954 to 1986 (monthly)	The prediction performance of the Bp neural network is the worst, while that of the support vector machine model and the random forest model is comparable.	RMSE, NSE, R^2 and MRE
Fan et al.[35]	Rivers Lancang–Mekong and Nu–Salween, China		From 1950 to 2010 (monthly)		

Li et al.[36]	Yangtze River, China	ENR, SVR, RF and XGB	1965–2016. (monthly)	The RF and XGB perform better in terms of forecasting and have higher and more stable accuracies than the ENR and SVR. .	RRMSE, MAPE, QR1 and QR2
Shortridge et al.[37]	Gilgel Abbay River.in northwest Ethiopia	GLM, GAM, MARS, ANN, RF and M5	1961 to 2004 (monthly)	Other approaches, particularly GAMs and random forests, are capable of effectively capturing non-linear interactions and lend themselves to simplified display of model structure.	MAE and NSE
Pasupa et al.[38]	Chao Phraya River in Thailand	LR, kernel regression, SVR, RBF kernel, kNN, and RF.	hourly	SVR with RBF kernel function in conjunction with 72-hour lag feature was discovered to be the best competitor.	RMSE
Li et al.[39]	Chile, China	RF, SVM, BRT, and SLM	1956 to 2000 (monthly)	This study reveals that the RF is an effective approach for reconstructing streamflow and a useful tool for analysing past hydrological change; it also has huge potential for recreating temperature and precipitation.	R, R^2, NSE and ST
Mohr et al.[40]	Chile, China	RF	Daily	Peak ground velocity and elevation extremes are identified as the most essential for predicting streamflow response using random forest classification.	
Piniewski,[41]	Poland	RF	Daily	The proposed random forest model has a mean predicted accuracy of 79 percent, which is high when compared to other models.	
Latif et al.[42]	Knowung river at Cedar ford, Australia	LSTM, RF and TB	Daily	It is suggested that future studies use the LSTM model to predict hydrological characteristics in various locations.	RMSE, NSE
Nhu,[43]	Zrebar Lake, iran	RF and M5	Daily	When compared to other developed algorithms, the M5P model predicts maximum lake water level succinctly.	R^2, RMSE, MAE, NSE PBIAS and PSR
Rezaie-Balfet al.[44]	Siira, Bilghan, and Gachsar,	GEP, RFR, EEMD-VMD-GEP	Daily	The feature of the AI methods, especially the EEMD-VMD-RFR algorithm, is to be able to be further	NSE, RMSE, MAE and RSD

	in Karaj basin, Iran	and EEMD-VMD-RFR		employed in the study region and provide more flexibility by adding desired decision variables for reservoir management.	
Khosravi,[45]	Taleghan catchment in northern Iran	M5P, RF and M5R	1979–2012(Daily)	AI algorithms gave a higher performance than the physically based models IHACRES, SWAT and HSPF for the Taleghan catchment.	R^2, RMSE. MAE, NSE, KGE and BIAS

4. Conclusion

Ongoing global climatic change is predicted to improve the global hydrologic cycle, affecting streamflow and water availability and perhaps disrupting river discharge regimes. Rivers serve as the primary supply of water for human activities, and mastering the streamflow variation of a river is critical for water resource planning, management, and consumption. Short-term projections of streamflow time series that are accurate and dependable are critical for water resource management. However, in the last decade, substantial research into new approaches for data-driven streamflow prediction has been motivated by the development of increasingly complex machine learning algorithms, paired with rapid gains in processing capabilities. In the last decade, many popular models such as support vector machines (SVMs), random forest (RF), and artificial neural networks (ANNs) have been employed for water resource research. This research was conducted based on streamflow prediction using Random Forest.

Recommendation

Other studies recommend using more than three distinct artificial intelligence models to produce the most accurate results. These studies used Random Forest to review streamflow prediction, and it is recommended in other studies to use more than three distinct artificial intelligence models to produce the most accurate results. In many cases, there is no visible variation in performance between models, so the more models available, the better and the comparison.

Compliance with ethical standards

Acknowledgements

The authors acknowledged this research to Humans and Machines LTD and Civil Lake LTD for their funding assistance.

Disclosure of conflict of interest

No conflict of interest.

References

[1] J. M. Sadler, J. L. Goodall, M. M. Morsy, and K. Spencer, "Modeling urban coastal flood severity from crowd-sourced flood reports using Poisson regression and Random Forest," *J. Hydrol.*, vol. 559, pp. 43–55, 2018, doi: 10.1016/j.jhydrol.2018.01.044.

[2] X. Li, J. Sha, and Z. L. Wang, "Comparison of daily streamflow forecasts using extreme learning machines and the random forest method," *Hydrol. Sci. J.*, vol. 64, no. 15, pp. 1857–1866, 2019, doi: 10.1080/02626667.2019.1680846.

[3] S. Amare *et al.*, "Susceptibility to gully erosion: Applying random forest (RF) and frequency ratio (FR) approaches to a small catchment in Ethiopia," *Water (Switzerland)*, vol. 13, no. 2, 2021, doi: 10.3390/w13020216.

[4] S. I. Abba, A. G. Usman, and S. Is, "Simulation for response surface in the HPLC optimization method development using arti fi cial intelligence models : A data-driven approach," vol. 201, no. March, 2020, doi: 10.1016/j.chemolab.2020.104007.

[5] A. G. Usman, M. H. Ahmad, R. N. Danraka, and S. I. Abba, "The effect of ethanolic leaves extract of Hymenodictyon floribundun on inflammatory biomarkers: a data-driven approach," *Bull. Natl. Res. Cent.*, vol. 45, no. 1, 2021, doi: 10.1186/s42269-021-00586-y.

[6] U. M. Ghali *et al.*, "Advanced chromatographic technique for performance simulation of anti-Alzheimer agent: an ensemble machine learning approach," *SN Appl. Sci.*, vol. 2, no. 11, 2020, doi: 10.1007/s42452-020-03690-2.

[7] S. I. Abba *et al.*, "Comparative implementation between neuro-emotional genetic algorithm and novel ensemble computing techniques for modelling dissolved oxygen concentration," *Hydrol. Sci. J.*, vol. 66, no. 10, pp. 1584–1596, 2021, doi: 10.1080/02626667.2021.1937179.

[8] S. Idris, M. A. A. Musa, S. I. Haruna, U. U. A. A. G. Usman, and M. I. A. Abba, "Implementation of soft - computing models for prediction of flexural strength of pervious concrete hybridized with rice husk ash and calcium carbide waste," *Model. Earth Syst. Environ.*, 2021, doi: 10.1007/s40808-021-01195-4.

[9] A. Alamrouni, F. Aslanova, S. Mati, H. S. Maccido, and A. A. Jibril, "Multi-Regional Modeling of Cumulative COVID-19 Cases Integrated with Environmental Forest Knowledge Estimation : A Deep Learning Ensemble Approach," *Int. J. Environ. Res. Public Heal.*, vol. 19(2), pp. 1–23, 2022, doi: https://doi.org/10.3390/ijerph19020738.

[10] S. I. Malami, F. H. Anwar, S. Abdulrahman, S. I. Haruna, S. I. A. Ali, and S. I. Abba, "Implementation of hybrid neuro-fuzzy and self-turning predictive model for the prediction of concrete carbonation depth: A soft computing technique," *Results Eng.*, vol. 10, no. May, p. 100228, 2021, doi: 10.1016/j.rineng.2021.100228.

[11] K. Mahmoud *et al.*, "Prediction of the effects of environmental factors towards COVID-19 outbreak using AI-based models," *IAES Int. J. Artif. Intell.*, vol. 10, no. 1, pp. 35–42, 2021, doi: 10.11591/ijai.v10.i1.pp35-42.

[12] U. M. Ghali, M. Alhosen, A. Degm, A. N. Alsharksi, Q. Hoti, and A. G. Usman, "Development Of Computational Intelligence Algorithms For Modelling The Performance Of Humanin And Its Derivatives In HPLC Optimization Method Development," vol. 9, no. 08, pp. 110–117, 2020.

[13] A. G. USMAN, S. IŞIK, S. I. ABBA, and F. MERİÇLİ, "Artificial intelligence–based models for the qualitative and quantitative prediction of a phytochemical compound using HPLC method," *Turkish J. Chem.*, vol. 44, no. 5, pp. 1339–1351, 2020, doi: 10.3906/kim-2003-6.

[14] G. M. Khalid and A. G. Usman, "Application of data-intelligence algorithms for modeling the compaction performance of new pharmaceutical excipients," 2021.

[15] S. I. Abba *et al.*, "Comparative implementation between neuro-emotional genetic algorithm and novel ensemble computing techniques for modelling dissolved oxygen concentration," *Hydrol. Sci. J.*, vol. 0, no. 0, 2021, doi: 10.1080/02626667.2021.1937179.

[16] H. Tongal and M. J. Booij, "Simulation and forecasting of streamflows using machine learning models coupled with base flow separation," *J. Hydrol.*, vol. 564, pp. 266–282, 2018, doi: 10.1016/j.jhydrol.2018.07.004.

[17] N. Cid *et al.*, "A biological tool to assess flow connectivity in reference temporary streams from the Mediterranean Basin," *Sci. Total Environ.*, vol. 540, pp. 178–190, 2016, doi: 10.1016/j.scitotenv.2015.06.086.

[18] L. Schoppa, M. Disse, and S. Bachmair, "Evaluating the performance of random forest for large-scale flood discharge simulation," *J. Hydrol.*, vol. 590, no. September, p. 125531, 2020, doi: 10.1016/j.jhydrol.2020.125531.

[19] M. P. Miller, D. M. Carlisle, D. M. Wolock, and M. Wieczorek, "A Database of Natural Monthly Streamflow Estimates from 1950 to 2015 for the Conterminous United States," *J. Am. Water Resour. Assoc.*, vol. 54, no. 6, pp. 1258–1269, 2018, doi: 10.1111/1752-1688.12685.

[20] Q. B. Pham, S. I. Abba, A. G. Usman, N. Thi, and T. Linh, "Potential of Hybrid Data-Intelligence Algorithms for Multi-Station Modelling of Rainfall," 2019.

[21] M. Benaafi, M. A. Yassin, A. G. Usman, and S. I. Abba, "Neurocomputing Modelling of Hydrochemical and Physical Properties of Groundwater Coupled with Spatial Clustering , GIS , and Statistical Techniques," 2022.

[22] A. G. Usman *et al.*, "Simulation of liver function enzymes as determinants of thyroidism: a novel ensemble machine learning approach," *Bull. Natl. Res. Cent.*, vol. 46, no. 1, 2022, doi: 10.1186/s42269-022-00756-6.

[23] A. G. Usman, S. Işik, and S. I. Abba, "Hybrid data-intelligence algorithms for the simulation of thymoquinone in HPLC method development," *J. Iran. Chem. Soc.*, vol. 18, no. 7, pp. 1537–1549, 2021, doi: 10.1007/s13738-020-02124-5.

[24] S. I. Abba and A. G. Usman, "Modeling of Water Treatment Plant Performance using Artificial Neural Network: Case Study Tamburawa Kano-Nigeria," *Dutse J. Pure Appl. Sci. (DUJOPAS), Vol.*, vol. 6, no. 3, pp. 135–14, 2020.

[25] A. G. U. Selin and I. S. I. Abba, "A Novel Multi - model Data - Driven Ensemble Technique for the Prediction of Retention Factor in HPLC Method Development," *Chromatographia*, no. 0123456789, 2020, doi: 10.1007/s10337-020-03912-0.

[26] S. I. Abba *et al.*, "Emerging evolutionary algorithm integrated with kernel principal component analysis for modeling the performance of a water treatment plant," *J. Water Process Eng.*, vol. 33, no. December 2019, p. 101081, 2020, doi: 10.1016/j.jwpe.2019.101081.

[27] U. M. Ghali *et al.*, "Applications of Artificial Intelligence-Based Models and Multi- Linear Regression for the Prediction of Thyroid Stimulating Hormone Level in the Human Body," vol. 29, no. 4, pp. 3690–3699, 2020.

[28] A. G. Usman, S. Işik, S. I. Abba, and F. Meriçli, "Chemometrics-based models hyphenated with ensemble machine learning for retention time simulation of isoquercitrin in Coriander sativum L. using high-performance liquid chromatography," *J. Sep. Sci.*, vol. 44, no. 4, pp. 843–849, 2021, doi: 10.1002/jssc.202000890.

[29] S. I. Abba, A. G. Usman, and S. Işik, "Simulation for response surface in the HPLC optimization method development using artificial intelligence models: A data-driven approach," *Chemom. Intell. Lab. Syst.*, p. 104007, 2020, doi: 10.1016/j.chemolab.2020.104007.

[30] H. U. Abdullahi, A. G. Usman, and S. I. Abba, "Modelling the Absorbance of a Bioactive Compound in HPLC Method using Artificial Neural Network and Multilinear Regression Methods," vol. 6, no. 2, pp. 362–371, 2020.

[31] M. H. Ahmad, A. G. Usman, and S. I. Abba, "Comparative performance of extreme learning machine and Hammerstein–Weiner models for modelling the intestinal hyper-motility and secretory inhibitory effects of methanolic leaf extract of Combretumhypopilinum Diels (Combretaceae)," *Silico Pharmacol.*, vol. 9, no. 1, 2021, doi: 10.1007/s40203-021-00090-1.

[32] S. Ismail, R. A. A. A. G. Usman, and S. I. Abba, "Development of chemometrics - based neurocomputing paradigm for simulation of manganese extraction using solid - phase tea waste," *Model. Earth Syst. Environ.*, 2022, doi: 10.1007/s40808-022-01369-8.

[33] W. Asnake Metekia, A. Garba Usman, B. Hatice Ulusoy, S. Isah Abba, and K. Chirkena Bali, "Artificial intelligence-based approaches for modeling the effects of spirulina growth mediums on total phenolic compounds," *Saudi J. Biol. Sci.*, no. xxxx, 2021, doi: 10.1016/j.sjbs.2021.09.055.

[34] F. Peng, J. Wen, Y. Zhang, and J. Jin, "Monthly Streamflow Prediction Based on Random Forest Algorithm and Phase Space Reconstruction Theory," *J. Phys. Conf. Ser.*, vol. 1637, no. 1, pp. 0–6, 2020, doi: 10.1088/1742-6596/1637/1/012091.

[35] H. Fan and D. He, "Temperature and precipitation variability and its effects on streamflow in the upstream regions of the Lancang-Mekong and Nu-Salween Rivers," *J. Hydrometeorol.*, vol. 16, no. 5, pp. 2248–2263, 2015, doi: 10.1175/JHM-D-14-0238.1.

[36] Y. Li, Z. Liang, Y. Hu, B. Li, B. Xu, and D. Wang, "A multi-model integration method for monthly streamflow prediction: Modified stacking ensemble strategy," *J. Hydroinformatics*, vol. 22, no. 2, pp. 310–326, 2020, doi: 10.2166/hydro.2019.066.

[37] J. E. Shortridge, S. D. Guikema, and B. F. Zaitchik, "Empirical streamflow simulation for water resource management in data-scarce seasonal watersheds," pp. 11083–11127, 2015, doi: 10.5194/hessd-12-11083-2015.

[38] K. Pasupa and S. Jungjareantrat, "Water levels forecast in Thailand: A case study of Chao Phraya river," *2016 14th Int. Conf. Control. Autom. Robot. Vision, ICARCV 2016*, vol. 2016, no. November, pp. 13–15, 2017, doi: 10.1109/ICARCV.2016.7838716.

[39] J. Li, Z. Wang, C. Lai, and Z. Zhang, "Tree-ring-width based streamflow reconstruction based on the random forest algorithm for the source region of the Yangtze River, China," *Catena*, vol. 183, no. August, p. 104216, 2019, doi: 10.1016/j.catena.2019.104216.

[40] C. H. Mohr, M. Manga, C. Y. Wang, and O. Korup, "Regional changes in streamflow after a megathrust earthquake," *Earth Planet. Sci. Lett.*, vol. 458, pp. 418–428, 2017, doi: 10.1016/j.epsl.2016.11.013.

[41] M. Piniewski, "Classification of natural flow regimes in Poland," *River Res. Appl.*, vol. 33, no. 7, pp. 1205–1218, 2017, doi: 10.1002/rra.3153.

[42] S. D. Latif and A. N. Ahmed, “Application of deep learning method for daily streamflow time-series prediction: A case study of the kowmung river at Cedar Ford, Australia,” *Int. J. Sustain. Dev. Plan.*, vol. 16, no. 3, pp. 497–501, 2021, doi: 10.18280/IJSDP.160310.

[43] V. H. Nhu *et al.*, “Daily water level prediction of zrebar lake (Iran): A comparison between m5p, random forest, random tree and reduced error pruning trees algorithms,” *ISPRS Int. J. Geo-Information*, vol. 9, no. 8, 2020, doi: 10.3390/ijgi9080479.

[44] M. Rezaie-Balf, S. F. Nowbandegani, S. Z. Samadi, H. Fallah, and S. Alaghmand, “An ensemble decomposition-based artificial intelligence approach for daily streamflow prediction,” *Water (Switzerland)*, vol. 11, no. 4, pp. 1–31, 2019, doi: 10.3390/w11040709.

[45] K. Khosravi *et al.*, “Improving daily stochastic stream flow prediction: comparison of novel hybrid data-mining algorithms,” *Hydrol. Sci. J.*, vol. 66, no. 9, pp. 1457–1474, 2021, doi: 10.1080/02626667.2021.1928673