



Application of augmented bat algorithm with artificial neural network in forecasting river inflow in Malaysia

Wei Joe Wee¹ · Kai Lun Chong² · Ali Najah Ahmed¹ · Marlinda Binti Abdul Malek³ · Yuk Feng Huang² · Mohsen Sherif^{4,5} · Ahmed Elshafie⁶

Received: 16 October 2022 / Accepted: 16 November 2022 / Published online: 17 December 2022
© The Author(s) 2022

Abstract

Hydrologists rely extensively on anticipating river streamflow (SF) to monitor and regulate flood management and water demand for people. Only a few simulation systems, where previous techniques failed to anticipate SF data quickly, let alone cost-effectively, and took a long time to execute. The bat algorithm (BA), a meta-heuristic approach, was used in this study to optimize the weights and biases of the artificial neural network (ANN) model. The proposed hybrid work was validated in five different study areas in Malaysia. The statistical tests analysis of the preliminary results revealed that hybrid BA-ANN was superior to forecasting the SF at all five selected study areas, with average RMSE values of 0.103 m³/s for training and 0.143 m³/s for testing as compared to ANN standalone training and testing yielding 0.091 m³/s and 0.116 m³/s, respectively. This finding signifies that the implementation of BA into the ANN model resulted in a 20% improvement. In addition, with an R^2 score of 0.951, the proposed model showed a better correlation than the 0.937 value of R^2 of standard ANN. Nonetheless, while the proposed work outperformed the conventional ANN, the Taylor diagram, violin plot, relative error, and scatter plot findings confirmed the disparities in the proposed work's performance throughout the research regions. The findings of these evaluations highlighted that the adaptability of the proposed works would need detailed investigation because its performance differed from case to case.

Keywords Artificial neural network · Bat meta-heuristic algorithm · Streamflow forecasting · Uncertainty analysis

Introduction

Water, renewable but limited, is vital for all lives on the planet Earth. Sustained, rapid population and economic

growth add to the challenge of maintaining water availability, raising questions about the continuous usage of reservoir water at unsustainable levels. Therefore, the importance of assessing and anticipating changes in water availability has

✉ Ali Najah Ahmed
Mahfoodh@uniten.edu.my

Wei Joe Wee
weewejoe95@gmail.com

Kai Lun Chong
chongkailun1993@gmail.com

Marlinda Binti Abdul Malek
marlinda@iiu.edu.my

Yuk Feng Huang
huangyf@utar.edu.my

Mohsen Sherif
msherif@uaeu.ac.ae

Ahmed Elshafie
elshafie@um.edu.my

² Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, Jalan Sg. Long, Bandar Sg. Long, 43000 Kajang, Selangor, Malaysia

³ Cataclysmic Management and Sustainable Development Research Group (CAMSDE), Department of Civil Engineering, Kulliyah of Engineering, International Islamic University Malaysia, Selangor, Malaysia

⁴ Civil and Environmental Engineering Department, College of Engineering, United Arab Emirates University, P.O. Box 15551 Al Ain, United Arab Emirates

⁵ National Water and Energy Center, United Arab Emirates University, P.O. Box 15551 Al Ain, United Arab Emirates

⁶ Department of Civil Engineering, Faculty of Engineering, University of Malaya (UM), 50603 Kuala Lumpur, Malaysia

¹ Institute of Energy Infrastructure and Department of Civil Engineering, College of Engineering, Universiti Tenaga Nasional (UNITEN), 43000 Kajang, Selangor, Malaysia

grown in recent years due to population increase, industry, and urbanization, particularly in extreme conditions (Osman et al. 2022). Water shortage is one of the most challenging difficulties encountered in predominantly semiarid locations, and it is most likely influenced by several variables, one of which is accessibility (Kordani et al. 2022). Streamflow (SF) is one of the components of water transport from the land to waterbodies that is significant for management. The importance of accurate and reliable SF forecasting includes timely and efficient water resource management (Fathian et al. 2019), irrigation management verdicts (Tongal and Boojj 2018), flood risk assessment (Shafizadeh-Moghadam et al. 2018), and planning releases (Sapitang et al. 2020). Knowing the quantity and the duration in advance is therefore essential.

Nonetheless, characteristics of streamflow, such as non-linearity, stochasticity, and non-stationarity, pose complicated features that impair the effectiveness of contemporary hydrological models (Chabokpour et al. 2020). Also, given the complexities of modeling SF, particularly in the spatial and temporal domain, a physically based model may oversimplify or underestimate the actual SF process when the underlying pattern is not comprehended thoroughly (Chong et al. 2020; Martina et al. 2011). Evidenced such a problem was highlighted in the study by (Staudinger et al. 2011), where the minimum SF was poorly simulated using a hydrological model. Pellicer-Martínez et al. (2015) noticed that the hydrological modeling process conducted was often without considering the groundwater aspect. They are also often designed to handle maximum SF circumstances. Thus, the application in low-streamflow conditions relies on the relevance of presented hydrological processes (Nicolle et al. 2014). However, if a proper model calibration is conducted, a physically based model is remarkably excellent for obtaining reliable predictions (Samarasinghe et al. 2022). Another technical difficulty with SF forecasting is real-time forecasting and the resources required for continuous monitoring. Although most reservoirs have monitoring techniques to monitor river streamflow, they fail to use the data to produce a consistent database. Practically, machine learning techniques, which use fewer computational resources to execute the same task, have been the focus of study and science.

Data-driven modeling has gained popularity over the past 20 years in the hydrology and water resources industries. It is principally owing to their capacity to extract nonlinearity, stochasticity, and patterns from historical river flow data. For instance, Azamathulla et al. (2010) adopted support vector machine (SVM) to predict sediment transport. Apart from sediment transport prediction, a number of studies have utilized machine learning in the associated sectors of water-related fields, such as rainfall prediction and side weir discharge prediction. (Azamathulla et al. 2016; Chaplot 2021). Rashki Ghaleh Nou et al. (2019) show that traditional ANN and SVM are

inadequate in estimating scour depth near submerged weirs. They instead used a self-adaptive extreme learning machine (SAELM), which outperforms the classical machine learning model. Emadi et al. (2022) demonstrated that the integrated wavelet component in machine learning models can perform efficiently under water scarcity situations. The factors of climatic classes, river morphology, hydrological factor, and land-use characteristics all play a significant part in the model's accuracy. Despite the fact that these excellent studies have garnered widespread attention in the hydrological community, it is worth noting that the limits of AI-based models remain universal and a hot topic among academics. These drawbacks include problems such as over-fitting, model generalization, and local convergence rate (Nawi et al. 2017).

ANN, for example, is a traditional AI approach that employs gradient-based optimizers and is subject to the mentioned drawbacks. The time series complexity and the searchability of the training method are issues concerning the application of gradient-based optimizers. In a similar optimization process, researchers, on the other hand, have used meta-heuristic algorithms, a nature-inspired optimization strategy, to solve combinatorial optimization approaches. Meta-heuristics are primarily employed to generate near-optimal solutions since they lack the power to provide optimal solutions but maintain properties that allow them to be utilized independently or in conjunction with conventional approaches on a wide range of issues (Chong et al. 2021; Pei et al. 2019). Despite its mathematical shortcomings and excessive time consumption, GA, for example, may be quickly modified and utilized to tackle the concerns of complicated situations (Beg and Islam 2016). The cuckoo optimization algorithm (COA) performs well in global convergence, contributes to research capabilities at both the local and global levels, and employs Lévy flights for global searching (He et al. 2018), but has a poor performance, a low convergence rate, and a tendency to drop to the optimum local value (Qu and He 2015).

Yang (2010) developed a fast, adaptable, and user-friendly optimization model (BA) based on real-time challenges and reservoir systems, allowing them to move quickly from the exploration to the exploitation phase. BA theoretically combined the advantages of modern optimization algorithms with more realistic technique formulations to provide a better outcome. As of present, only a few researchers have used BA in forecasting or prediction tasks. In 2015, hybrid NN and BA were used to forecast stock prices (Golmaryami et al. 2015; Shahvaroughi Farahani and Razavi Hajiagha 2021). Aside from that, Banadkooki et al. (2020) used BA to evaluate and optimize the weight and structure of an ANN to forecast time series suspended sediment load. In a separate study, Xing et al. (2016) used BA to determine the parameters of SVM for monthly SF prediction at the Yichang station on the Yangtze River in China. They eventually claimed that the BA-based

SVM model outperformed the ANN and cross-validation-based SVM models in terms of accuracy. However, as stated earlier in the paragraph, the uses of BA in forecasting river SF remain relatively restricted.

The main contribution of this work is to provide an optimal approach through hybridization for forecasting upstream river SF in Malaysia. Various datasets were obtained from credible Malaysian authorities, and comprehensive statistical analysis and data pre-processing were performed throughout the model development. The performance of the ANN method and its hybridization with a new bio-inspired optimizer to supplement the regular ANN was validated using five selected study areas. Additional model validations were carried out employing various advanced analytics such as the Taylor diagram, violin plot, relative error, and scatter plot to analyze the model calculation inaccuracy. A multi-time scale streamflow forecasting research was also carried out to forecast future SF in all study areas from 2021 to 2025, a period of five years (60 months).

The proposed hybrid hydrological model

Artificial neural network (ANN)

As the name implies, ANN is built on the structural components of neural networks inherent in all biological organisms, which are then translated into analogous computing systems. Because it is a calculus-based algorithm, ANN is a heuristic search algorithm (it takes numerous approximations to find non-optimal but adequate results). The basic architecture of ANN comprises three layers: the input layer, the hidden layer, and the output layer. The capacity of ANN to foresee and predict both short-term and long-term data that is often too complicated to be appraised using standard modeling approaches has led to its widespread application. Furthermore, ANN has a high tolerance for errors, fast performance, and strong generalization skills when learning from historical data. Back propagation (BP) was employed to train the model. Figure 1 depicts the flowchart of the adopted ANN model.

Nodes are necessary for the input and hidden layers. Each node in its respective layer is associated with its weight. The data is subsequently routed through hidden layers that include one or more activation function components (depending on the number of nodes in hidden layers). The forward propagation of the neural network can be computed using Eq. (1).

$$Y_j = \sum_i^n X_i W_{ij} + b \tag{1}$$

$$O_j = f(Y_j) \tag{2}$$

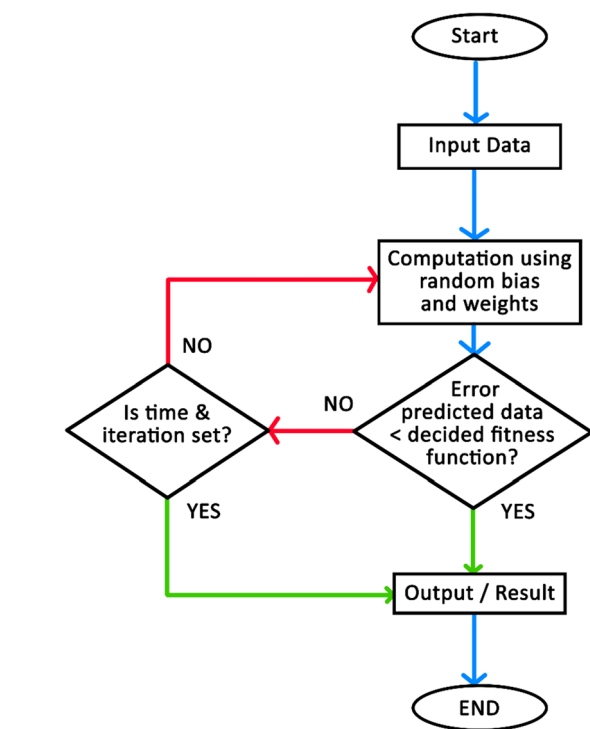


Fig. 1 The flow of the standard ANN model

where Y_j = hidden/final value derived from the summation of input data and its associated weights, X_i = data input/previous input value, and W_{ij} = weight associated between input/previous layer and output layer. O_j is referred to as the activation function, a vital component that introduces non-linearity property to the ANN model. Once the information reaches the output layer, they will compute an error between the predicted and observed value using Eq. (2).

$$\frac{\delta E}{\delta O_j} = (Y_{obs} - Y_j) \sum O_j^1 \tag{3}$$

$$\frac{\delta E}{\delta Y_j} = \frac{\delta E}{\delta O_j} \sum O_j^1 \tag{4}$$

$$\frac{\delta E}{\delta W_{ji}} = \frac{\delta E}{\delta Y_j} \frac{\delta Y_j}{\delta X_i} \tag{5}$$

where E_k = computed error between observed, Y_{obs} and predicted value, Y_k and O_j^1

$$W_{ji} = W_{ji} - n \frac{\delta E}{\delta W_{ji}} \tag{6}$$

= derivative of the activation function, n = learning rate.

Evolutionary artificial intelligence model (BA-ANN)

Yang (2010) was the first to develop BA in 2010, using the behavior of a bat’s echolocation to uncover its prey as the foundation of the algorithm. As previously stated, BA depends on microbat echolocation patterns to manage pulse emission rates and loudness while recognizing prey and avoiding obstacles in the absence of light. It is accomplished by employing echolocation mechanisms to convert information from sound waves emitted by prey and obstacles. In BA, the sole component will be used in echolocation, while time delay and loudness fluctuations will have no bearing on optimization difficulties. Figure 2 depicts the flow of the hybrid BA-ANN model, which was used in this investigation. BA is critical in the operations of random signal bandwidth adjustment via harmonics search. The theoretical fundamental of this algorithm is the assumption that the bats are soaring with velocity (v_i) from an initial position (s_i) with a minimum frequency (f_{min}), fluctuating wavelength (λ), and loudness (L_0) to seek out prey. They are equipped

with the capability of regulating the frequencies (f_i) as well as the rate of frequencies [$r \in (0, 1)$].

Case study

Study area

As illustrated in Fig. 3, for simplicity, the five major selected study areas in Malaysia were designated from A to E. The importance of such a study is that they are mainly in charge of the majority of energy-producing divisions in Malaysia. Besides, the benefits from flood control, irrigation supports, and clean drinking water are provided. Table 1 shows the specifics of the streamflow and rainfall study areas.

Data sources and pre-processing

According to previous research, the primary contributing characteristics of river flow in the water resources field are rainfall, flow rate, temperature, water level, streamflow, and water consumption. The parameters are vital as they greatly influence the state and condition of the water source. However, based on the literature review conducted, only rainfall and streamflow had shown consistent significance in affecting the output data, whereas the effects of temperature, water level, and water consumption were still quite vague. Therefore, with

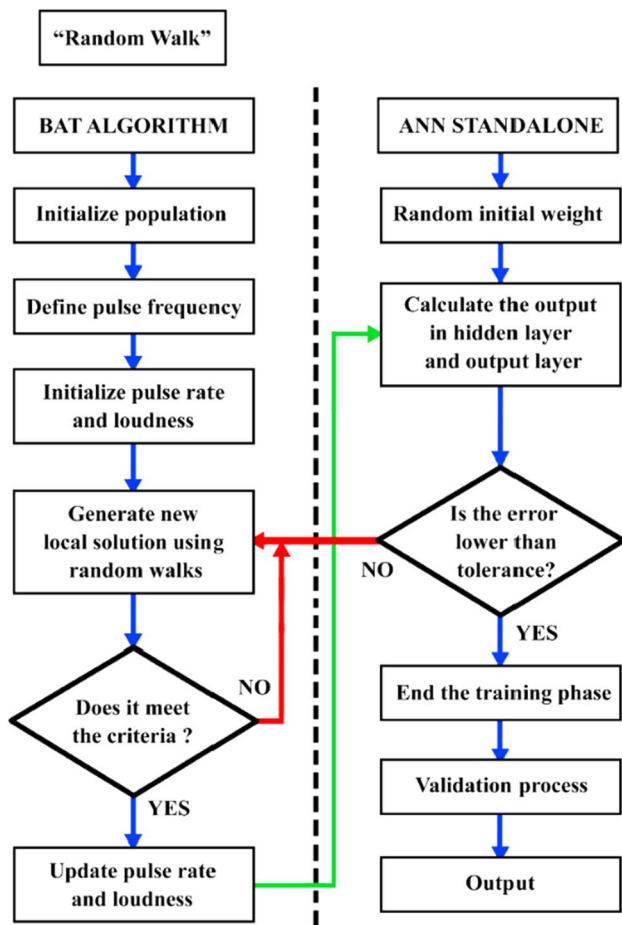


Fig. 2 The flowchart of the proposed hybrid BA-ANN model

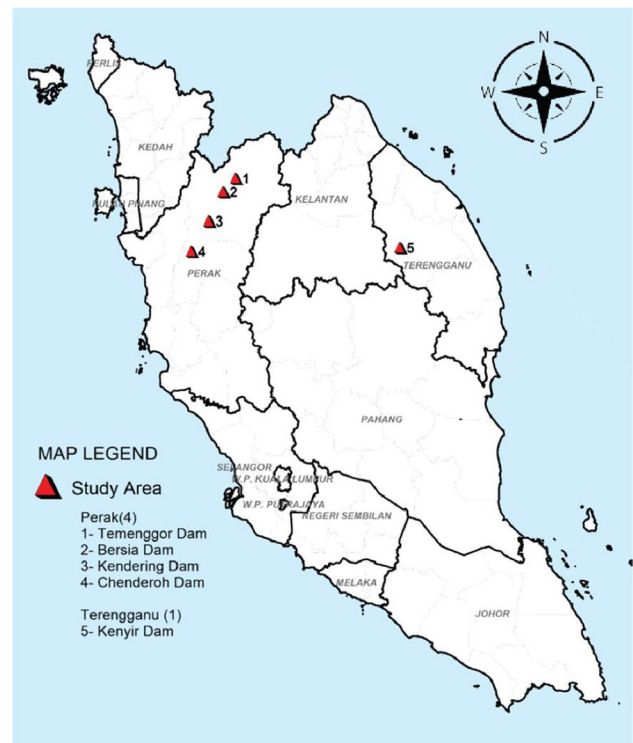


Fig. 3 Location of the five selected study areas in Malaysia

Table 1 Rainfall and streamflow data details of the five selected study areas in Malaysia

Study area	ID	Location	State	Latitude (N)	Longitude (E)
Rainfall data details					
A	5610063	Sungai Rui, Hulu Perak	Perak	05° 36' 15"	101° 04' 50"
B	5007020	Larut & Matang, Sungai Kurau	Perak	05° 00' 43"	100° 43' 52"
C	5206032	Selama, Sungai Kerian	Perak	05° 13' 45"	100° 41' 20"
D	5128001	Hulu Terengganu	Terengganu	05° 08' 35"	102° 50' 40"
E	5710061	Sungai Tasek, Hulu Perak	Perak	05° 42' 30"	101° 00' 00"
Streamflow data details					
A	5610401	Sg. Rui, Hulu Perak	Perak	05° 36' 10"	101° 03' 45"
B	5007421	Larut & Matang, Sg. Kurau	Perak	05° 00' 45"	100° 43' 55"
C	5206432	Selama, Sg. Kerian	Perak	05° 13' 45"	100° 41' 20"
D	5129438	Sg. Telemong, Hulu Terengganu	Terengganu	05° 09' 00"	102° 55' 00"
E	5610401	Sg. Rui, Hulu Perak	Perak	05° 36.00' 10"	101° 03.00' 45"

41 years (492 months) of historical data from the year 1980 until 2020 retrieved from DID and DOE, the parameters to be used for analysis purposes of this study are historical rainfall (RF) data and historical streamflow (SF) data.

Data segmentation is critical in the ANN context for performance and to minimize data contamination. By referring to the empirical analysis conducted by Gholamy et al. (2018), it was deduced that the best results were attained in 70–80% of the data allocated for training and the remainder, 20–30% for testing. Thus, for both ANN standalone and hybrid BA-ANN models of this study, 70% of historical data are used for training (from January 1980 to August 2008), and the remaining 30% is used for testing the model (from the year September 2008 to December 2020). Apart from data partition, prior to training, data normalization was carried out to alter the existing raw data values to a range from 0 to 1. It is due to both forecasting models having an activation function that is bounded and susceptible to ranges that are not initially defined. The statistical tests were only then applied to the denormalized data, which was obtained through a reversion process to revert to the actual streamflow magnitude.

Model assessment and benchmark model comparison

In this work, ANN was used as a benchmarking model to BA-ANN to assess the accuracy of the forecast SF from the upstream river into the reservoirs of five selected study areas. Furthermore, several statistical tests were conducted using data from past and forecasted results to assess the validity of the forecasting model for all the selected study areas. The statistical tests comprise of root-mean-squared error (RMSE), coefficient of determination; often known as R-squared (R^2), and mean absolute error (MAE). In general, the greater the predicting skill, the lower the derived statistical values. In contrast, a higher R^2 value is preferable. RMSE calculates the average squared difference between estimated

and actual values. The mean of the residuals in the dataset is calculated by MAE. To assess the forecasting model’s accuracy, R^2 provides a fitted regression line. The employed statistical indices are defined below:-

$$RSME = \sqrt{\frac{1}{n} \sum_{i=0}^n (x_i - x_f)^2} \tag{7}$$

$$R^2 = 1 - \frac{\sum (x_i - x_f)^2}{\sum (x_i - \bar{x}_f)^2} \tag{8}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |x_i - x_f| \tag{9}$$

In addition to the four criteria listed above, an uncertainty analysis using the Sequential Uncertainty Fitting algorithm version 2 (SUFI-2) was performed to assess the model’s dependability and consistency. The analysis was presented in the form of a Taylor diagram, violin plot, relative error, and scatter plot.

Results and discussion

Solution representations (architecture and input selection)

The input selection process is one of the most critical aspects of constructing a forecasting model for assessing the significance of a model’s correlation between its inputs and outputs. The model-free technique was chosen to be applied to the model in this study, which leverages existing data for training, testing, and validation of the forecasting model. Table 2 displays the four streamflow vector input scenarios considered in this study as inputs for the AI-based model.

Aside from the input/output vector, a few parameters were considered when designing the architecture of the ANN standalone model. Different input scenarios and optimal network architecture for the proposed model are explored and discussed.

However, Scenario 1 is the baseline analysis using only R_t as the starting input value of historical RF data. In Scenario 2, the model will investigate the relationship between R_{t-1} (the starting input value of historical RF data with lag-time 1) and S_{t-1} (the initial input value of historical SF data with lag-time 1). Scenario 3 is an extension of Scenario 2, in which additional S_{t-2} and R_{t-2} , along with all the variables in Scenario 2, are analyzed. Finally, Scenario 4 took into account all possible input combinations. The preliminary step is conducted using only one study area.

The hyperbolic tangent function was used in this study through trial-and-error analysis to aid in comprehending the complicated large size of the network and reducing the time required during the learning process. This hyperbolic tangent function delivers output varying from -1 to 1 . A momentum (α) value of 0.3 and a learning rate (β) value of -0.3 were utilized during the development process. The scale of architecture networks was also examined, including the number of hidden nodes. Table 3 shows the simulation results of the ANN standalone forecasting model in four scenarios. Based on the overall consideration, an ANN architecture was utilized to predict the one-step ahead of monthly streamflow forecasting, with input variable based on scenario 3.

Similarly, the BA parameters utilized to improve the ANN were acquired. In addition to the parameters required in the ANN model, several parameters must be put into consideration in an effort to regulate the bat's echolocation wave that would allow it to find its prey. The optimal value of population size, iteration, pulse rate, and loudness was recognized by carrying out the trial-and-error method. Table 4 shows the simulation results of the hybrid BA-ANN forecasting model in four scenarios. Based on the overall assessment, a population size of 5 , a pulse rate of 0.35 , and a loudness of 0.2 were obtained. The best architecture for the proposed work is with an input variable based on scenario 3.

Table 2 Basis of the equations for all four scenarios of the forecasting models

Scenario	Input	Output
1	R_t	S_0
2	$R_{t-1} + S_{t-1}$	S_0
3	$R_{t-2} + S_{t-2} + R_{t-1} + S_{t-1}$	S_0
4	$R_{t-2} + S_{t-2} + R_{t-1} + S_{t-1} + R_t$	S_0

Table 3 Simulations of ANN standalone forecasting model in four scenarios

Scenario	1	2	3	4
Iteration	550	250	300	350
Hidden layer	2	2	2	2
Elapsed time (s)	54	29	33	47
RMSE (m^3/s)				
Train	0.129	0.116	0.093	0.091
Test	0.134	0.121	0.121	0.117
R				
Train	0.783	0.900	0.973	0.978
Test	0.642	0.820	0.962	0.951
R^2				
Train	0.782	0.882	0.955	0.953
Test	0.570	0.798	0.932	0.937
MAE (m^3/s)				
Train	0.035	0.014	0.003	0.005
Test	0.047	0.015	0.006	0.007

Statistical evaluation of proposed model at five selected study areas

Tables 5 and 6 summarize the reservoir performance indicators based on the best input scenario, scenario 3. Notably, the hybrid BA-ANN models outperform the standard ANN model across all metrics. One of the comparison aspects is the computational time. The shorter time needed for the proposed work indicates that BA improves the searchability rate of the ANN model toward the global optimum. The sped-up convergence rate is essential in reservoir streamflow

Table 4 Simulations of hybrid BA-ANN forecasting model in four scenarios

Scenario	1	2	3	4
Population	5	5	5	5
Iteration	450	450	200	400
Hidden layer	1	1	1	1
Elapsed time (s)	44	46	21	39
RMSE (m^3/s)				
Train	0.135	0.098	0.085	0.089
Test	0.138	0.117	0.113	0.115
R				
Train	0.780	0.856	0.980	0.977
Test	0.641	0.746	0.974	0.963
R^2				
Train	0.770	0.867	0.983	0.987
Test	0.583	0.744	0.948	0.970
MAE (m^3/s)				
Train	0.035	0.022	0.004	0.004

Table 5 Statistical test performance of the ANN standalone model at five selected study areas

Study area	Time taken (s)	RMSE (m ³ /s)		R ²		MAE (m ³ /s)	
		Train	Test	Train	Test	Train	Test
D	33.0	0.093	0.121	0.955	0.932	0.003	0.006
A	32.0	0.107	0.154	0.973	0.949	0.003	0.003
B	33.0	0.096	0.138	0.970	0.960	0.006	0.008
C	31.0	0.112	0.149	0.923	0.914	0.003	0.005
E	32.0	0.108	0.152	0.976	0.930	0.004	0.003
Average	32.2	0.103	0.143	0.959	0.937	0.004	0.005

Table 6 Statistical test performance of the hybrid BA-ANN model at five selected study areas

Study area	Time taken (s)	RMSE (m ³ /s)		R ²		MAE (m ³ /s)	
		Train	Test	Train	Test	Train	Test
D	21.0	0.085	0.113	0.983	0.948	0.004	0.006
A	22.0	0.094	0.115	0.993	0.967	0.003	0.007
B	21.0	0.096	0.116	0.986	0.926	0.002	0.003
C	20.0	0.087	0.119	0.959	0.949	0.001	0.002
E	23.0	0.095	0.117	0.955	0.966	0.003	0.005
Average	21.4	0.091	0.116	0.975	0.951	0.003	0.005

forecasting, particularly in real-time forecasting, as it can provide more time for the operator to make crucial decisions when an impending shortage/overflow occurs. The results also indicated that the ANN standalone model has several limitations, particularly during local convergence, influencing the time-consuming speed convergence. The results are consistent with the findings in the study by Zaini et al. (2018). They did, however, anticipate the streamflow using additional meteorological characteristics as input, which were not included in this study. The additional parameter could have affected the accuracy of the model. The current study demonstrates that the proposed work might still perform effectively without the extra meteorological factors.

The proposed models can achieve higher accuracy of streamflow forecasting value than the standard ANN model while maintaining a faster convergence rate. For instance, from the result of the hybrid model simulation in Table 6, it can be deduced that at 200 iterations, the average testing value for RMSE of all five study areas is 0.116 m³/s, as opposed to 0.103 m³/s value of that standard ANN model. Study area D has the lowest RMSE for training among the study areas based on the proposed model, at 0.085 m³/s, while study area B has the highest, at 0.096 m³/s. Study area D maintains the lowest RMSE testing result of 0.113 m³/s, whereas study area C has the highest at 0.119 m³/s, trailing study area B. While the proposed hybrid approach decreases RMSE by 20%, the R² measure only results in a minor marginal increase, revealing that the R² metric might not be sufficient to assess their effectiveness. As presented in Tables 5 and 6, superior exploitation and exploration elements of BA improve the searchability of ANN by minimizing the risks

of being caught in local optimums, resulting in increased streamflow forecasting accuracy.

To visualize the capability of ANN, whose works without explicit knowledge of the underlying physical process, in forecasting the streamflow, the observed and forecasted streamflow for all five study areas were plotted and are shown in Fig. 4. The typical ANN model is unable to reliably estimate low streamflow in all five locations, worsening the situation when there is water scarcity. Peak flow is regularly overestimated or underestimated by the model, causing the reservoir to overflow, especially during periods of severe rain. With the addition of BA to ANN, the proposed model can accurately estimate whether the flow is at its peak or its minimum.

Supplementary analysis (all scenarios in each study area)

The preceding section emphasizes a single scenario, scenario 3, as it was proven to be the best input scenario for ANN and hybrid models. However, as stated in Sect. 4.1, it was only applied to study area D as a preliminary stage for testing the various derived scenarios and model architecture. With the optimized parameters tabulated in Sect. 4.1, the models were developed for other study areas, but only the 3rd scenario was adopted. The models were developed for additional study areas using the optimized parameters reported in Sect. 4.1. However, as the only adopted scenario was the 3rd scenario, it poses the issue of whether the proposed paradigm applies to other cases. Therefore, it is

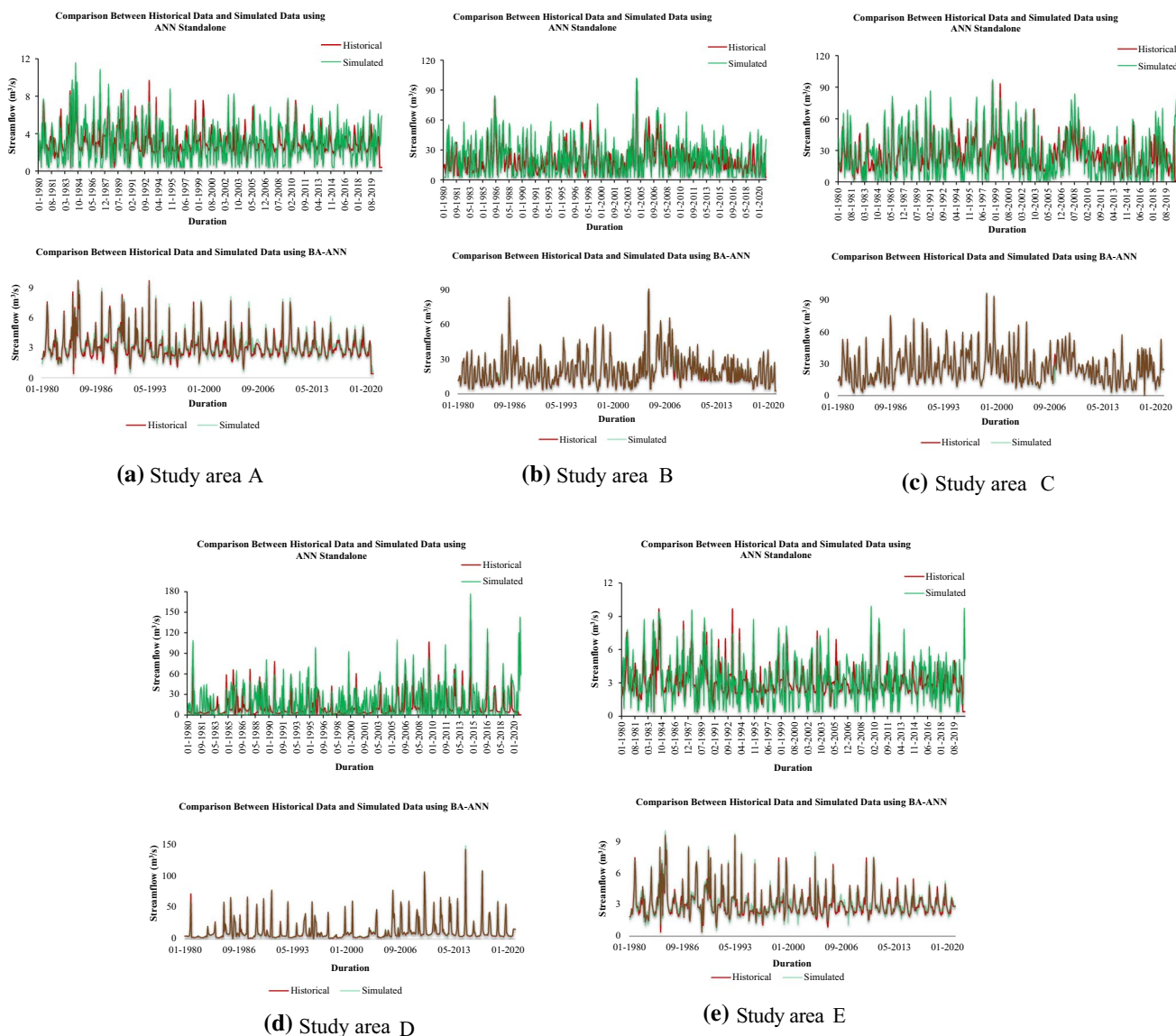


Fig. 4 Streamflow simulation and forecast for all study areas

crucial to validate the findings further by applying them to the remaining study areas adopted in this study.

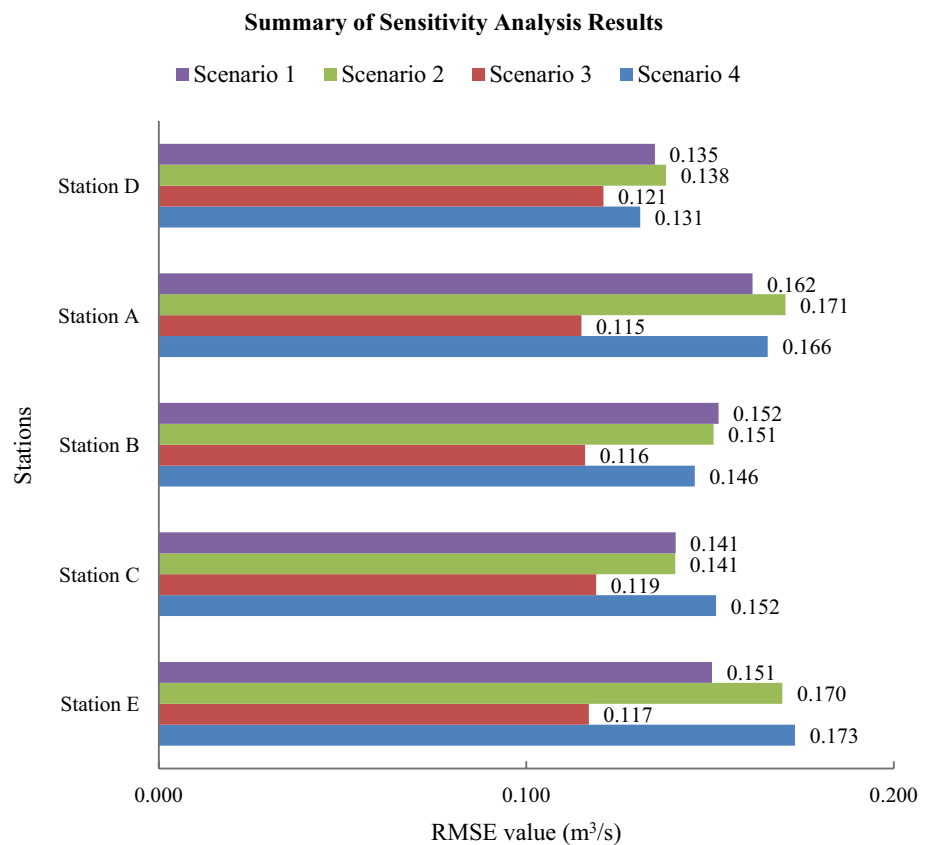
The findings tabulated in Table 7 and Fig. 5 demonstrate that scenario 3 remains the best scenario even when applied

to the other four study areas, as evidenced by the lowest RMSE value compared to another scenario. Given that all of the analyses imply scenario 3, it is plausible to assume that the elements to be considered for the equation in SF

Table 7 Results of trial and error for all scenarios in each study area

Scenario	Parameter (testing)	Study area				
		D	A	B	C	E
		RMSE	RMSE	RMSE	RMSE	RMSE
1	$R_t = S_0$	0.135	0.162	0.152	0.141	0.151
2	$R_{t-1} + S_{t-1} = S_0$	0.138	0.171	0.151	0.141	0.170
3	$R_{t-2} + S_{t-2} + R_{t-1} + S_{t-1} = S_0$	0.113	0.115	0.116	0.119	0.117
4	$R_{t-2} + S_{t-2} + R_{t-1} + S_{t-1} + R_t = S_0$	0.131	0.166	0.146	0.152	0.173

Fig. 5 Summary of sensitivity analysis (SA) results for all scenarios in each study area



forecasting should include: (1) R_{t-1} (initial input value of historical RF data with lag-time 1), and (2) R_{t-2} (initial input value of historical RF data with lag-time 2), (3) S_{t-1} (initial input value of historical SF data with lag-time 1), (4) S_{t-2} (initial input value of historical SF data with lag-time 2), and (5) S_0 (initial output value of forecasted SF data). In conclusion, identifying the best scenario equation to apply to the model is critical since insufficient input data frequently underperformed (scenario 1), while too much input has a detrimental impact (scenario 4).

Multi-step ahead forecasting analysis

In real-time circumstances, being able to foresee each conceivable event utilizing time series forecasting would be a great benefit. It would be highly motivating to be able to acquire exact, accurate time series, especially in the field of engineering. An iterative forecasting approach was used to forecast future SF in all research areas from 2021 to 2025, a total of five years, in order to assess the developed model’s capabilities in multi-step-ahead forecasting (60 months). However, it is worth noting that the cumulative error in the prediction is carried into the subsequent forecast in the iterative forecasting process, resulting in an increased error in streamflow accuracy. The results of the multi-step forward

forecasting are best represented graphically, as shown in Fig. 6.

Figure 7 depicts the likely maximum and lowest forecasted streamflow, which provides a discussion to clarify such results from Fig. 6. The value of SF for study area D, which served as the preliminary, ranges from 10.00 to 70.00 m³/s. The highest value, 77.87 m³/s, was predicted to occur in May 2022, while the lowest value, 0.88 m³/s, was forecasted to occur in January 2021. The SF value in study area A ranged from 2.00 m³/s to 6.00 m³/s. May 2023 was anticipated to have the highest value, at 7.68 m³/s, while July 2023 was predicted to have the lowest value, at 1.60 m³/s. The value of SF varied in study area B from 5.00 to 45.00 m³/s. The anticipated values were for the maximum value to occur in August 2021 at 53.10 m³/s and the lowest value to occur in April 2022 at 3.68 m³/s. Next, in study area C, the value of SF fluctuated from as low as 20.00 m³/s to as high as 65.00 m³/s. The estimated peak value for May 2023, 65.81 m³/s, is expected to fall to the anticipated minimum value for February 2024, 2.15 m³/s, around nine months later. Finally, the value of SF fluctuated from 2.00 to 5.00 m³/s for study area E. Forecasted SF data reveal that they will be at their highest in May 2023, as they will reach a value of 7.68 m³/s and at their lowest, in February 2025 as they will be at 1.57 m³/s.

Fig. 6 Forecasted streamflow data of all five study areas for 2021–2025

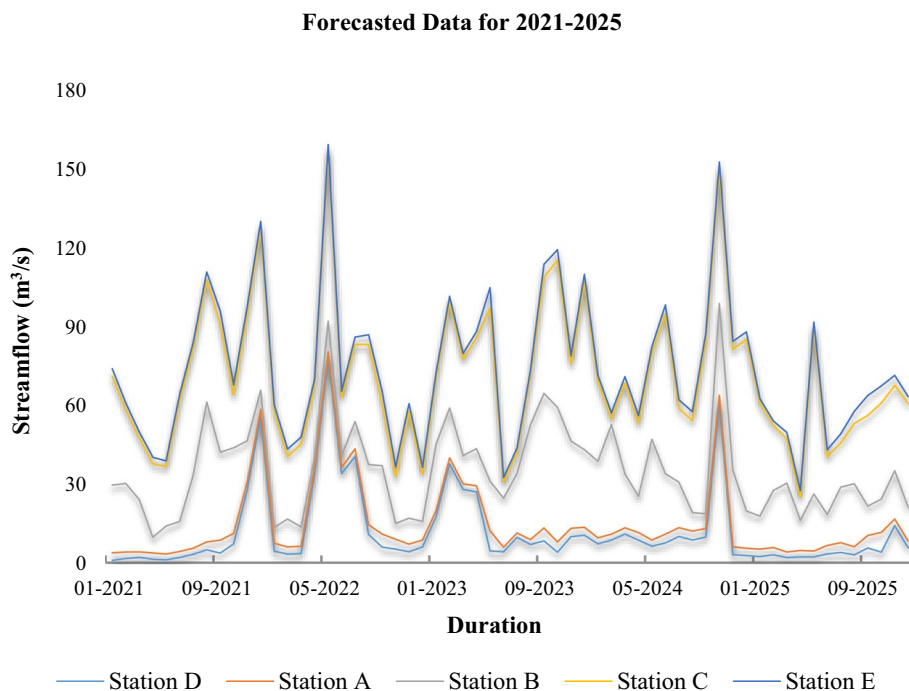
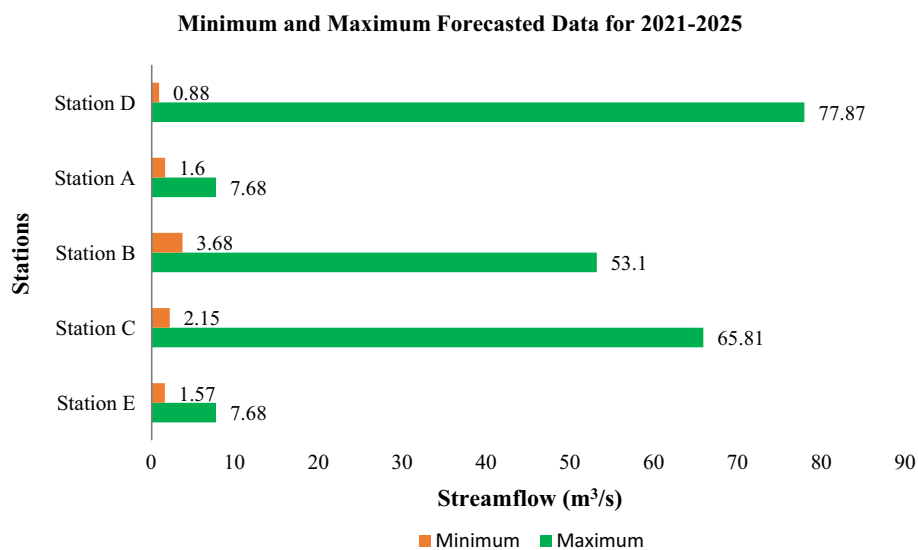


Fig. 7 Minimum and maximum forecasted data of all five study areas from 2021 till 2025

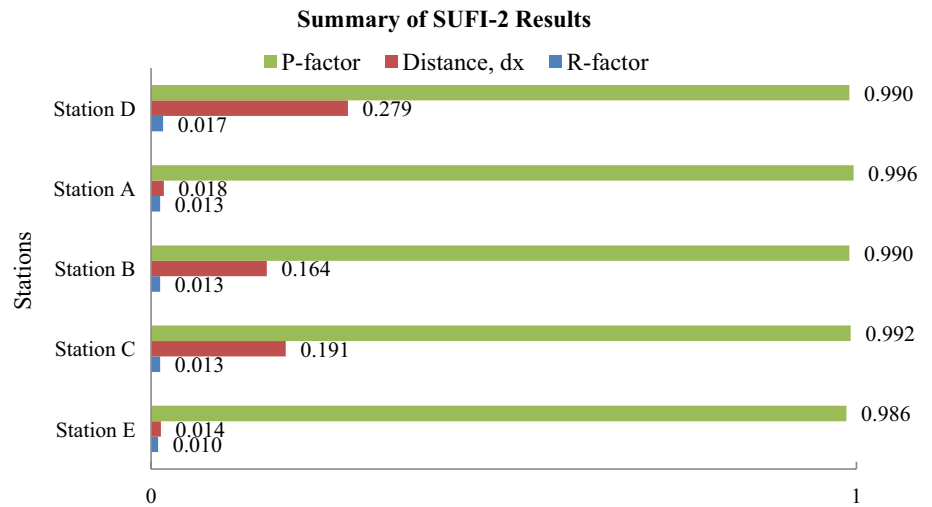


Uncertainty assessment

A 95% confidence interval (95 CI) was established to examine the performance uncertainty in the developed model. The *p*-factor and *r*-factor, two essential statistical parameters, were computed to show the degree of uncertainty (Fig. 8). These two components have distinct interpretations of outcomes, with a better-trained model having a *p*-factor around one and an *r*-factor near zero. Numerically, study area A scores the highest *p*-factor value, at 0.996 (99.6%), followed by study area C, which scores 0.992 (99.2%). Study area E got the lowest value, with 0.986 (98.6%), while study

areas D and B both have 0.990 (99.0%). The *r*-factors for the five reservoir study areas are 0.01, 0.017, and 0.013, with study area E being the best, study area D being the worst, and study areas A, B, and C being comparable. In general, *p*-factors larger than 0.7 and *r*-factors greater than 1.5 are considered statistically significant. Despite the worst *p*-factor value of 0.986 and *r*-factor value of 0.017, the models surpass the critical value significantly. The goodness-of-fit indices, such as the optimum average (*dx*) value being close to or equal to zero, may be used to quantify the uncertainty in streamflow forecasting. The lowest *dx* value in study areas A and E might indicate that the model captured a majority

Fig. 8 Summary of SUFI-2 results for hybrid BA-ANN model in all five study areas



of the underlying information in that streamflow time series. Meanwhile, study areas B, C, and D have dx values greater than 0.15. A notable observation here is that the *p*-factor and *r*-factor may have modest changes between study areas, necessitating the use of the dx as the variances in value were rather considerable and obvious.

Graphical representation of the proposed model proficiency

The inevitable concern is whether the same level of achievement will be maintained in subsequent years. To further compare the model performance, some graphical

representations, including relative error, scatter plots, and Taylor and violin diagrams, were utilized to show the findings gained from the best input combination.

Taylor diagrams based model performance The Taylor diagram is a diagnostic metric through a simple visual representation of model performances versus available data. The Taylor diagram for each of the five study areas is shown in Fig. 9. The terms R, RMSE, and SD demonstrate how comparable the observation records and prediction models are.

First and foremost, the placement of the developed model at each study area in the Taylor diagram shows a vital discovery. As every model falls inside the black dashed line,

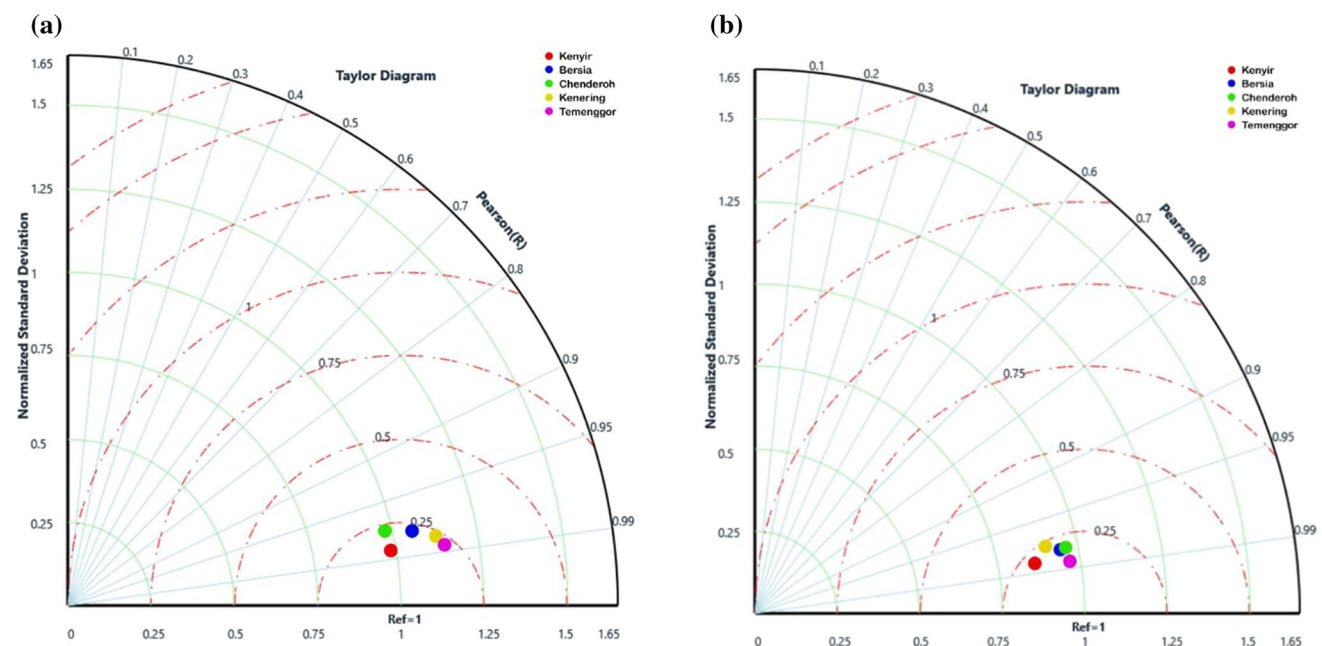


Fig. 9 Taylor diagram of hybrid BA-ANN at all five study areas: **a** training phase, **b** testing phase

they all tend to underestimate the variability of streamflow at their respective study area. Despite this, study areas B and E are relatively closer to their measured streamflow compared to the others. The fact that streamflow variability is generally underestimated may be due to the models' inadequacy in capturing streamflow during peak flows. Study area E typically performs the poorest in all three statistics throughout the validation period. Study area D did the best with an RMSE value of $0.113 \text{ m}^3/\text{s}$ and the highest value of R at 0.974. It should be observed that the simulated pattern that is closest to the reference point fits the data well. The remaining study areas performed averagely, with approximately an identical R correlation. Among the study areas with identical R correlation values, study area B is significantly closer to the observations and has a comparable standard deviation. In the meanwhile, study area A is less susceptible to spatial variability than study area B.

Violin diagrams based model performance A violin plot, as explained in the previous section, is a way of plotting numeric data to analyze the percentile ranking of accuracies, which includes the maximum and minimum value of the data, the median, and the interquartile range, together with the entire distribution of the simulated streamflow. Violin plot is another method used to compare the distribu-

tion of output models with reality. Violin plot is similar to a box plot with a more visible display and a more detailed description of the difference between the underlying distributions. Each violin plot's denser section demonstrates that the streamflow distribution is at its greatest in that range of values. The violin plots' tips display the range of values without filtering outliers and background noise.

The findings are recorded in Table 8 and presented in Fig. 10. The shape of the density mass function is highly comparable to observations in all five study areas. Nevertheless, a deeper look revealed increased dispersion and skewness in the study area D streamflow, as indicated by the highest $23.86 \text{ m}^3/\text{s}$ generated. Accordingly, upstream regions of study area B models have the highest likelihood of streamflow values (higher density), with an error of $0.42 \text{ (m}^3/\text{s)}$ as opposed to $0.03 \text{ (m}^3/\text{s)}$ in study area A. Despite having an average error of $0.15 \text{ m}^3/\text{s}$, there is a minor difference among such study areas; study areas D, A, and E all have comparable relative tails to recorded values at the upstream discharge, but B and C have longer tails. Except for the study area C model, which underestimated streamflow in the 25th percentile, other study area models overestimated streamflow in the 25th percentile. However, such variances are negligible, with a range difference between $0.41 \text{ m}^3/\text{s}$ at study area B and $0.12 \text{ m}^3/\text{s}$ at study area D. The 75th

Table 8 Summary of violin plot at all five study areas

Study area	Data	Max. (m^3/s)	Upper (m^3/s)	Quad. 3 (m^3/s)	Med. (m^3/s)	Quad. 1 (m^3/s)	Min. (m^3/s)
D	His	172.74	19.82	10.25	5.96	3.82	0.40
	Sim	148.88	19.75	10.40	6.33	3.94	0.40
	Error	23.86	0.07	0.15	0.37	0.12	0.00
A	His	9.68	5.52	3.59	2.80	2.25	0.40
	Sim	9.75	5.55	3.80	2.97	2.41	0.52
	Error	0.07	0.03	0.21	0.17	0.16	0.12
B	His	90.52	48.52	26.39	17.05	11.11	2.52
	Sim	89.99	48.94	26.96	17.52	11.52	2.42
	Error	0.53	0.42	0.57	0.47	0.41	0.10
C	His	95.94	63.26	35.34	22.68	15.03	0.98
	Sim	96.96	63.10	35.34	22.98	14.98	1.38
	Error	1.02	0.16	0.00	0.30	0.05	0.40
E	His	9.68	5.52	3.59	2.80	2.27	0.40
	Sim	10.11	5.46	3.64	3.01	2.40	0.40
	Error	0.43	0.06	0.05	0.21	0.13	0.00
Avg. Error		5.18	0.15	0.20	0.30	0.17	0.12

where His = observed streamflow, Sim = predicted streamflow

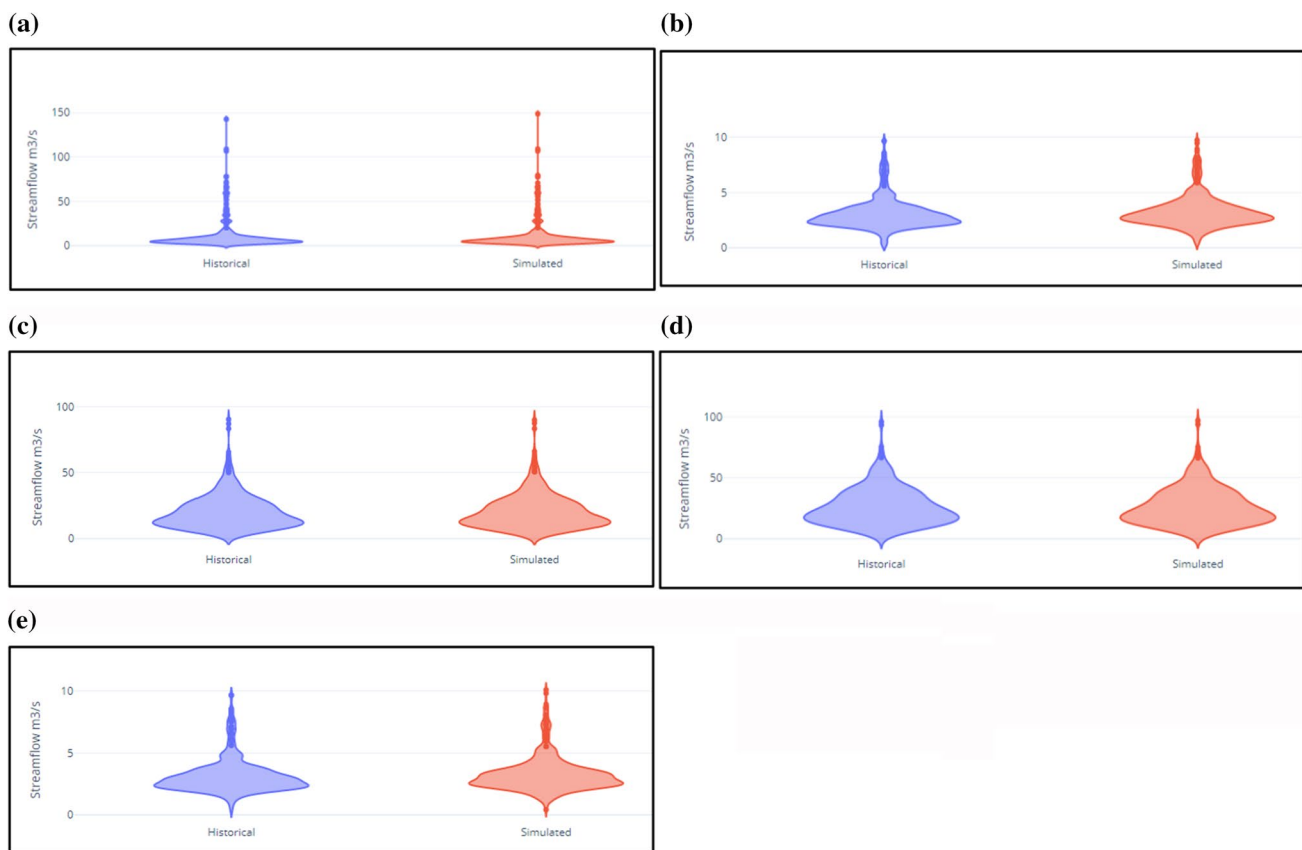


Fig. 10 Violin plot for all study areas: **a** study area D, **b** study area A, **c** study area B, **d** study area C, and **e** study area E

percentile found a similar result in all study areas as the 25th percentile, albeit with minor variance. For instance, a similar trend can be found in study areas D, A, and B, although with a worse outcome than the 25th percentile error. Besides, study area C is performing better at the 75th percentile than the 25th percentile, indicating that the model can forecast the peak flow relatively better than the low flow.

Scatter plots and relative errors based model performance The scatter plot of the estimated and observed streamflow was generated in this section. As seen in Fig. 11, shifting to the right raised both the X and Y values. As a result, all five sites showed a favorable association between historical and simulated data, with an R^2 value of more than 0.9. Furthermore, the computed relative error (RE) of the simulated data versus historical data for all five study areas revealed that the RE ranged from 1.39 to 2.51%. Study area A has the highest RE, at 2.51%, followed by E, at 2.49%. The RE is 1.57% at study area D and 1.44% at study area B, respectively. Finally, study area C scores the lowest, and hence the best, RE of 1.39%. The findings are shown in Fig. 12.

Conclusion and recommendation

In this study, an evolutionary AI-based model is developed and assessed for multi-time scale streamflow forecasting, with bat algorithm (BA) adopted, as an optimization approach, to search for the optimal set of weights and biases of the artificial neural network (ANN). The superior exploration and exploitation of BA can significantly improve the performance of ANN in achieving better forecasting accuracy. The effects of a few elements, including historical streamflow (SF) and rainfall (RF) data, were assumed to be significant contributors to the river's upstream entry into the reservoir.

The experimental verification on five selected study areas offered substantial evidence of the proposed work's applicability and dependability in all the adopted performance indicators. It is validated using uncertainty analysis using SUFI-2 that the results produced from the proposed hybrid model exceeded 95% for the p -factor and the average distance values for all study regions were near zero. Besides, a further implication of the research is that BA as an

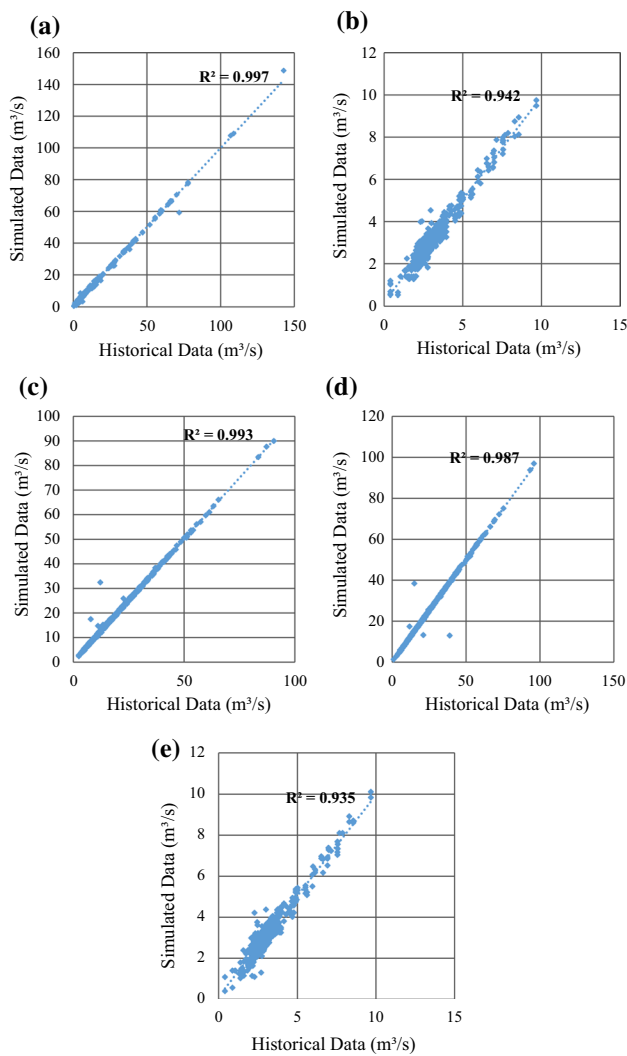
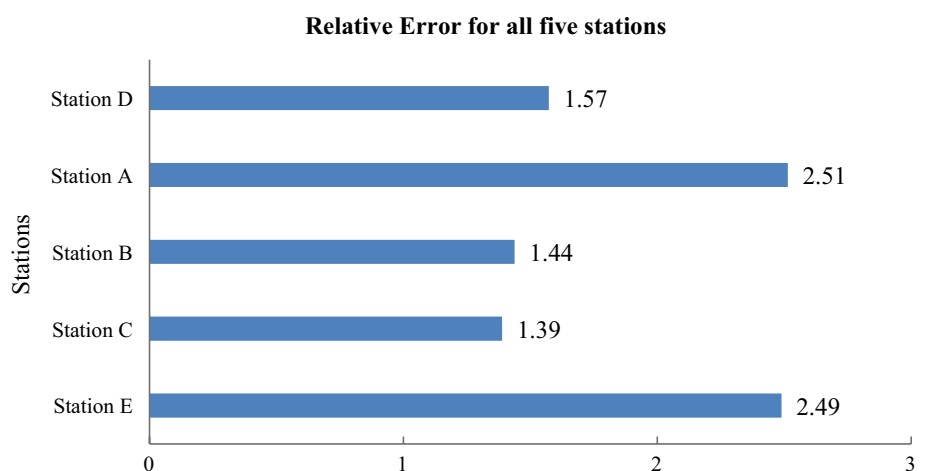


Fig. 11 Scatter plot all study areas: violin plot for all study areas: **a** study area D, **b** study area A, **c** study area B, **d** study area C, and **e** study area E

Fig. 12 Results of relative error (RE) of hybrid BA-ANN for all five study areas



optimization tool is well compatible with ANN in addressing nonlinear and complex time series. BA performs well in pure optimization, although it differs somewhat from how the ANN model learns. This research suggests that BA can train an ANN model with adequate generalizability, as indicated by the reduced RMSE in the testing dataset (unforeseen). As this study may have indicated, the execution of the proposed work differed from case to case. It could provide valuable insight to other academic academics in their studies on regulating the discharge of hydroelectric water under different conditions using an effective streamflow forecasting technique. Last but not least, the proposed work has the potential to be further extended to deep learning models, which are possibly more complex with the additional parameters to be optimized.

The work conducted for this research unquestionably necessitates additional research to be carried out in the future. For future work, the outcomes of this research deliver the insights as follows:-

- (a) The association between existing data on climate, hydrology, and land use must be examined in more depth to investigate the components of SF forecasting.
- (b) The proposed forecasting model of this study could be extended to additional study areas in diverse geographical locations, climate zones, and land-use activities.
- (c) The hybrid BA-ANN model can be integrated with other deterministic models (i.e., Info Works RS) to enhance the management of SF.
- (d) The hybrid BA-ANN model can be used with another expert system (for example, fuzzy approaches) to develop a more comprehensive SF short-term and long-term monitoring strategy.

Acknowledgements The authors are grateful to the Department of Irrigation and Drainage (DID) Malaysia for providing data to conduct this study.

Data availability Data are available upon reasonable request from the corresponding author.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare no conflict of interest.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Azamathulla HM, Ghani AA, Chang CK, Hasan ZA, Zakaria NA (2010) Machine learning approach to predict sediment load—a case study. *CLEAN Soil Air Water* 38(10):969–976
- Azamathulla HM, Haghiabi AH, Parsaie A (2016) Prediction of side weir discharge coefficient by support vector machine technique. *Water Supply* 16(4):1002–1016
- Banadkooki FB, Ehteram M, Ahmed AN, Teo FY, Ebrahimi M, Fai CM, Huang YF, El-Shafie A (2020) Suspended sediment load prediction using artificial neural network and ant lion optimization algorithm. *Environ Sci Pollut Res* 27(30):38094–38116
- Beg AH, Islam MZ (2016) Advantages and limitations of genetic algorithms for clustering records. pp 2478–2483
- Chabokpour J, Chaplot B, Dasineh M, Ghaderi A, Azamathulla HM (2020) Functioning of the multilinear lag-cascade flood routing model as a means of transporting pollutants in the river. *Water Supply* 20(7):2845–2857
- Chaplot B (2021) Prediction of rainfall time series using soft computing techniques. *Environ Monit Assess* 193(11):1–11
- Chong KL, Lai SH, Ahmed AN, Zaafer WZW, Rao RV, Sherif M, Sefelnasr A, El-Shafie A (2021) Review on dam and reservoir optimal operation for irrigation and hydropower energy generation utilizing meta-heuristic algorithms. *IEEE Access* 9:19488–19505
- Chong KL, Lai SH, Yao Y, Ahmed AN, Jaafar WZW, El-Shafie A (2020) Performance enhancement model for rainfall forecasting utilizing integrated wavelet-convolutional neural network. *Water Resour Manag* 34(8):2371–2387
- Emadi A, Sobhani R, Ahmadi H, Boroomandnia A, Zamanzad-Ghavidel S, Azamathulla HM (2022) Multivariate modeling of agricultural river water abstraction via novel integrated-wavelet methods in various climatic conditions. *Environ Dev Sustain* 24(4):4845–4871
- Fathian F, Mehdizadeh S, Sales AK, Safari MJS (2019) Hybrid models to improve the monthly river flow prediction: Integrating artificial intelligence and non-linear time series models. *J Hydrol* 575:1200–1213
- Gholamy A, Kreinovich V, Kosheleva O (2018) Why 70/30 or 80/20 relation between training and testing sets: a pedagogical explanation
- Golmaryami M, Behzadi M, Ahmadzadeh M (2015) A hybrid method based on neural networks and a meta-heuristic bat algorithm for stock price prediction. *IEEE*, pp 269–275
- He X-S, Wang F, Wang Y, Yang X-S (2018) Nature-inspired algorithms and applied optimization. Springer, pp 53–67
- Kordani H, Chaplot B, Dehkharghani PR, Azamathulla HM (2022) People's participation in using treated wastewater as an approach for sustainability of ecosystem services, green spaces, and farmlands in peri-urban areas: the case study of Kalak-e Bala, Karaj, Iran. *Water Supply* 22(4):4571–4583
- Martina M, Todini E, Liu Z (2011) Preserving the dominant physical processes in a lumped hydrological model. *J Hydrol* 399(1–2):121–131
- Nawi NM, Hamzah F, Hamid NA, Rehman MZ, Aamir M, Ramli AA (2017) An optimized back propagation learning algorithm with adaptive learning rate. *Learning* 500:2
- Nicolle P, Pushpalatha R, Perrin C, François D, Thiéry D, Mathevet T, Le Lay M, Besson F, Soubeyroux J-M, Viel C (2014) Benchmarking hydrological models for low-flow simulation and forecasting on French catchments. *Hydrol Earth Syst Sci* 18(8):2829–2857
- Osman AIA, Ahmed AN, Huang YF, Kumar P, Birima AH, Sherif M, Sefelnasr A, Ebraheem AA, El-Shafie A (2022) Past, present and perspective methodology for groundwater modeling-based machine learning approaches. *Arch Comput Methods Eng* 29:1–17
- Pei W, Huayu G, Zheqi Z, Meibo L (2019) A novel hybrid firefly algorithm for global optimization. *IEEE*, pp 164–168
- Pellicer-Martínez F, González-Soto I, Martínez-Paz JM (2015) Analysis of incorporating groundwater exchanges in hydrological models. *Hydrol Process* 29(19):4361–4366
- Qu C, He W (2015) A double mutation cuckoo search algorithm for solving systems of nonlinear equations. *Int J Hybrid Inf Technol* 8(12):433–448
- Rashki Ghaleh Nou M, Azhdary Moghaddam M, Shafai Bajestan M, Azamathulla HM (2019) Estimation of scour depth around submerged weirs using self-adaptive extreme learning machine. *J Hydroinform* 21(6):1082–1101
- Samarasinghe JT, Basnayaka V, Gunathilake MB, Azamathulla HM, Rathnayake U (2022) Comparing combined 1D/2D and 2D hydraulic simulations using high-resolution topographic data: examples from Sri Lanka—Lower Kelani River Basin. *Hydrology* 9(2):39
- Sapitang M, Ridwan WM, Faizal Kushiar K, Najah Ahmed A, El-Shafie A (2020) Machine learning application in reservoir water level forecasting for sustainable hydropower generation strategy. *Sustainability* 12(15):6121
- Shafizadeh-Moghadam H, Valavi R, Shahabi H, Chapi K, Shirzadi A (2018) Novel forecasting approaches using combination of machine learning and statistical models for flood susceptibility mapping. *J Environ Manag* 217:1–11
- Shahvaroughi Farahani M, Razavi Hajiagha SH (2021) Forecasting stock price using integrated artificial neural network and metaheuristic algorithms compared to time series models. *Soft Comput* 25(13):8483–8513
- Staudinger M, Stahl K, Seibert J, Clark M, Tallaksen L (2011) Comparison of hydrological model structures based on recession and low flow simulations. *Hydrol Earth Syst Sci* 15(11):3447–3459

- Tongal H, Booi MJ (2018) Simulation and forecasting of streamflows using machine learning models coupled with base flow separation. *J Hydrol* 564:266–282
- Xing B, Gan R, Liu G, Liu Z, Zhang J, Ren Y (2016) Monthly mean streamflow prediction based on bat algorithm-support vector machine. *J Hydrol Eng* 21(2):04015057
- Yang X-S (2010) Nature inspired cooperative strategies for optimization (NICSO 2010). Springer, pp 65–74
- Zaini N, Malek MA, Yusoff M, Osmi SFC, Mardi NH, Norhisham S (2018) Bat algorithm and neural network for monthly streamflow prediction. AIP Publishing LLC, p 020260

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.