

## Artificial hummingbird algorithm-optimized boosted tree for improved rainfall-runoff modelling

Lyce Ndolo Umba<sup>a,b</sup>, Ilham Yahya Amir<sup>a</sup>, Gebre Gelete<sup>a,c</sup>, Hüseyin Gökçekuş<sup>a</sup> and Ikenna D. Uwanuakwa<sup>a,\*</sup>

<sup>a</sup> Department of Civil Engineering, Near East University, Nicosia, Mersin-10, Turkey

<sup>b</sup> Faculty of Sciences, University of Lubumbashi, Lubumbashi, Haut-Katanga, Democratic Republic of Congo

<sup>c</sup> Environmental and Atmospheric Sciences Research Group, Scientific Research Center, Al-Ayen University, Thi-Qar, Nasiriyah 64001, Iraq

\*Corresponding author. E-mail: ikeuwanuakwa@gmail.com; ikenna.uwanuakwa@neu.edu.tr

### ABSTRACT

Rainfall-runoff modelling is a critical component of hydrological studies, and its accuracy is essential for water resource management. Recent advances in machine learning have led to the development of more sophisticated rainfall-runoff models, but there is still room for improvement. This study proposes a novel approach to streamflow modelling that uses the artificial hummingbird algorithm (AHA) to optimize the boosted tree algorithm. The AHA-boosted tree algorithm model was compared against two established methods, the support vector machine (SVM) and the Gaussian process regression (GPR), using a variety of statistical and graphical performance measures. The results showed that the AHA-boosted tree algorithm model significantly outperformed the SVM and GPR models, with an  $R^2$  of 0.932, RMSE of 5.358 m<sup>3</sup>/s, MAE of 2.365 m<sup>3</sup>/s, and MSE of 28.705 m<sup>3</sup>/s. The SVM model followed while the GPR model had the least accurate performance. However, all models underperformed in capturing the peak flow of the hydrograph. Evaluations using both statistical and graphical performance measures, including time series plots, scatter plots, and Taylor diagrams, were critical in this assessment. The results suggest that the AHA-boosted tree algorithm could potentially be a superior alternative for enhancing the precision of rainfall-runoff modelling, despite certain challenges in predicting peak flow events.

**Key words:** artificial hummingbirds, boosted tree, optimization, rainfall-runoff

### HIGHLIGHTS

- Introduced artificial hummingbird algorithm (AHA) to optimize boosted tree algorithm in rainfall-runoff modelling.
- AHA-boosted model significantly outperforms SVM and GPR methods.
- Enhances precision in hydrological studies using advanced machine learning.
- Challenges in peak flow prediction were not adequately addressed by the models.
- Aligns with *Journal of Hydroinformatics*' focus on computational hydrosystem advancements.

### INTRODUCTION

According to Hadi & Tombul (2018), the science of hydrology examines the cycle and motions of water between the atmosphere, soil, and earth's surface. Rainfall-runoff processes, for example, are hydrological processes that are distinguished by their great complexity and degree of temporal and spatial variability. Determining these non-linear, complex, dynamic, and non-stationary interactions between the hydrological processes is thus a constant challenge for hydrologists (Nawaz *et al.* 2016). When designing, managing, and planning water resources and hydrology (Kisi *et al.* 2013; Anusree & Varghese 2016), mitigating drought, and planning flood control work (Nourani & Komasi 2013), as a result, over the past few years, modelling of the rainfall-runoff process has received considerable attention and has taken centre stage in hydrological research. Numerous models have been developed and are still being worked on to represent these intricate and unpredictable hydrological processes (Nourani *et al.* 2021; Gelete *et al.* 2023). These models can be generically categorized as black-box and physically based models.

Artificial intelligence (AI), which translates input values to output values, is now preferred over physically based models for accurate rainfall-runoff modelling as it becomes more challenging to take into account all the physical aspects of the modelling (Nourani & Komasi 2013). The use of AI models in modelling the complex non-linear hydrologic process has recently

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

been acknowledged as a useful technique (Kisi *et al.* 2012). AI models are capable of handling random and missing data in addition to long-term data prediction (Besaw *et al.* 2010).

Due to its capacity to capture intricate, non-linear interactions between input and output variables, machine learning (ML) approaches have been used more and more in hydrological modelling (Zhang *et al.* 2021). Support vector machines (SVM), random forests (RF), and artificial neural networks (ANN) are a few of the techniques that have demonstrated promising results in streamflow and rainfall-runoff process prediction (Zhang *et al.* 2021). The choice of input variables and calibrating of the algorithm hyperparameters can have a big impact on how well these models perform (Uwanuakwa & Akpinar 2020). As a result, finding the optimum hyperparameters is essential for increasing the precision of ML models.

Automatic calibration of ML models has been a critical aspect of hydrological modelling for nearly three decades. Early works by Babovic *et al.* (1994), Savic & Walters (1995), and others laid the foundation for calibrating hydrodynamic models using simulated evolution and genetic algorithm techniques. These approaches later expanded to perform multi-objective calibration, as highlighted by Soon & Madsen (2005). This study aims to build upon these foundational works by employing newer optimization techniques, specifically the artificial hummingbird algorithm (AHA), to calibrate boosted tree algorithms in rainfall-runoff modelling.

Metaheuristic algorithms are a type of optimization technique that has been increasingly employed in hydrological modelling due to their ability to find near-optimal solutions in a reasonable computational time. These algorithms, inspired by natural phenomena and using stochastic processes to explore the solution space, include genetic algorithms (GA), particle swarm optimization (PSO), and ant colony optimization (ACO).

Various advantages exist for employing metaheuristic algorithms in hydrological modelling. First, these algorithms can manage complex, non-linear, and multimodal objective functions, which are frequently encountered in hydrological models. Second, they are capable of handling a large set of model parameters, a common feature of physically based models. Third, they offer a set of nearly optimal solutions that are beneficial for conducting uncertainty analysis. Consequently, this study aims to merge ML with advanced optimization techniques to enhance the accuracy and reliability of hydrological modelling through the automatic calibration of model hyperparameters.

## METHODOLOGY

The dataset used in the study was collected in the Katar catchment located in the Ethiopian Central Rift Valley basins, the sub-catchment covered a 3,298 km area and is situated along longitude 38.899°–39.41° E and latitude 7.359°–8.165° N as shown in Figure 1.

The study area climatically experiences semi-arid to sub-humid climatic conditions with an average annual precipitation of 980.85 mm and temperature of 18 °C. The Katar River and its tributaries which flow into the Lake Ziway consist of the main source in the catchment and support various activities such as irrigation, fishing, and other uses for water within the study area communities. The source of the meteorological data was six meteorological stations at Arata, Assela, Bekoji, Kalumsa, Ogolcho, and Sagure between 2008 and 2017. The discharge for the catchment is measured at a hydrometry station at Abura.

The input variables used are lagged runoff (Qt-1, Qt-2, Qt-3, Qt-4, and Qt-5) and rainfall (Pt). These variables have been found in the previous literature to have explained the variance in the measured rainfall-runoff (Gelete 2023b). Furthermore, the boosted tree algorithm was chosen because of its high accuracy predictive capabilities, just like other ML algorithms (Subramani *et al.* 2023), the boosted tree algorithm could suffer from overfitting and requires parameter fine-tuning. Therefore, the novel AHA was employed for parameter tuning. To compare the performance of the optimized algorithm, the traditional support vector machine (SVM), and Gaussian process regression (GPR) were used to compare the performance of the AHA-boosted tree method.

In order to ensure that the selected input variables are important in the model training, a feature importance process was performed using the easyGSA algorithm to measure the sensitivity of the input variables to output variance.

A total of 3,653 data points were collected spanning from 2008 to 2017, the dataset was partitioned into training, from 2008 to 2014 (70%), and testing, from 2015 to 2017 (30%). The training dataset was used to train the model, and the performance of the trained models was evaluated with the test dataset. It is worth noting here that the same data point for training and testing was used for each of the models to reduce bias that may occur due to data quality. Finally, the model evaluation parameters were used to measure each of the model performances on the test dataset.

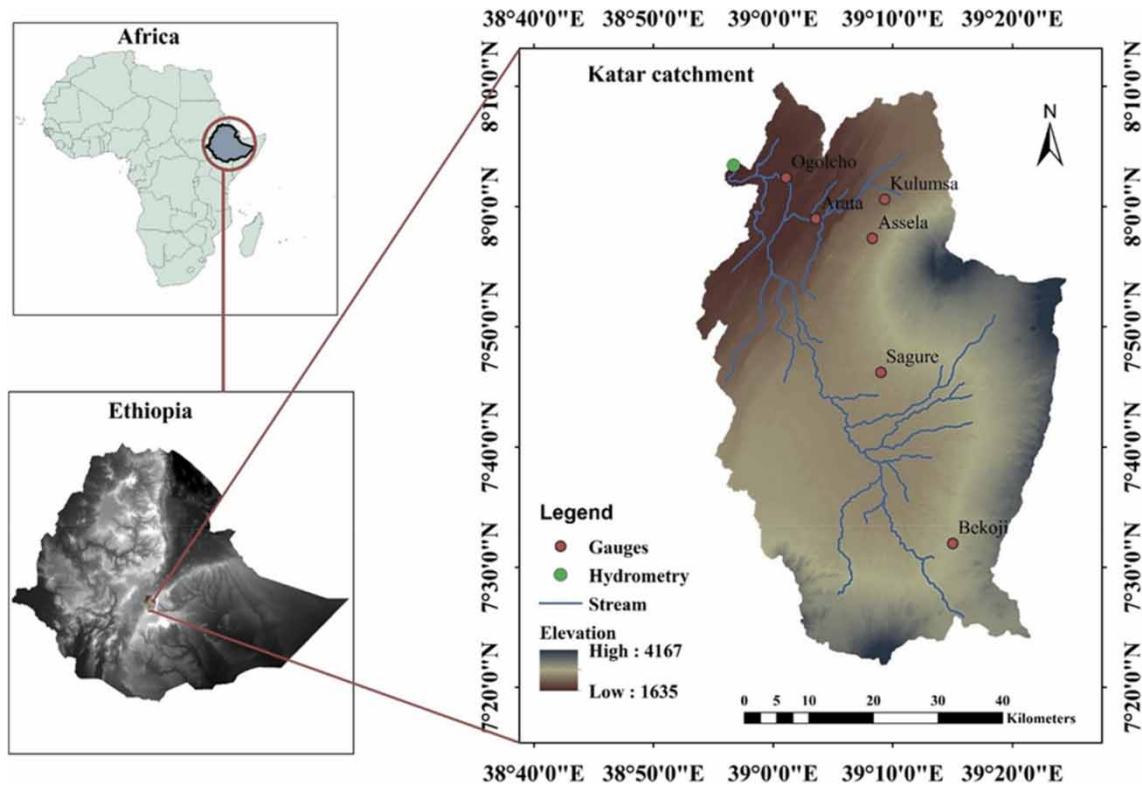


Figure 1 | Map of the study area.

### Artificial hummingbird algorithm

AHA is called the artificial hummingbird algorithm. It is a new bio-inspired, based on the intelligent foraging of hummingbirds and the special flight, is also a meta-heuristics (Zhao *et al.* 2022). Compared to other birds, hummingbirds possess a unique brain structure with a larger hippocampus, demonstrating their remarkable memory for foraging, making them the most intelligent birds studied (Ward *et al.* 2012). However, AHA imitates the directed, territorial, and migratory foraging behaviours of hummingbirds. It is modelled after the flight abilities and clever foraging techniques of hummingbirds using a visiting table (Zhao *et al.* 2022). AHA comprises three fundamental components: food sources, hummingbirds, and visit tables:

- Food sources examine nectar quality, nectar-refilling rate, and last visit to flowers.
- Hummingbirds are assigned to specific food sources, and they share information with other hummingbirds.
- Visit tables record visit levels for each food source, with high visit levels indicating priority visits. This helps hummingbirds obtain more nectar from food sources.

Hummingbirds visit food sources with high nectar-refilling rates, using the AHA algorithm for optimization. They simulate guided, territorial, and migrating foraging behaviours using three mathematical models.

#### 1. Guided foraging

In AHA, the hummingbirds tend to visit the food source with the highest visit level from the food sources with the most nectar volume during the guided foraging phase. The AHA algorithm models three flight skills: omnidirectional, diagonal, and axial:

$$\text{Axial: } D^{(i)} = \begin{cases} 1 & \text{if } i = \text{randi}([1, d]) \\ 0 & \text{else} \end{cases} \quad i = 1, \dots, d \quad (1)$$

$$\text{Diagonal: } D^{(i)} = \begin{cases} 1 & \text{if } i = P(j), j \in [1, k], P = \text{rand perm}(k), k \in [2, r_1 \cdot (d - 2) + 1] \\ 0 & \text{else} \end{cases} \quad i = 1, \dots, d \quad (2)$$

$$\text{Omnidirectional: } D^{(i)} = 1 \quad i = 1, \dots, d \quad (3)$$

Axial flight allows a hummingbird to fly along any coordinate axis in the search space, while diagonal flight allows it to fly from one corner to the opposite. Omnidirectional flight allows it to fly in a direction projected to each coordinate axis. Where  $r_1$  is a random number in  $(0, 1)$ ,  $randi([1, d])$  generates an integer at random from 1 to  $d$ , and  $randperm(k)$  generates an integer at random from 1 to  $k$ . The mathematical model of guided foraging is as follows:

$$v_i(t + 1) = x_{i,tar}(t) + a \cdot D \cdot (x_i(t) - x_{i,tar}(t)) \quad (4)$$

where  $x_i(t)$  represents the location of the  $i$ th food source at that time,  $x_{i,tar}(t)$  is the location of the  $i$ th hummingbird's intended target food source, and  $a$  is a directed factor that follows the normal distribution  $N(0, 1)$ .

## 2. Territorial foraging

Hummingbirds in the territorial foraging phase move to their own region and seek new food sources after eating flower nectar within their territory. The mathematic equation imitating the local search of hummingbirds in the territorial foraging strategy and a candidate food source is obtained as follows:

$$v_i(t + 1) = x_i(t) + D \cdot b \cdot x_i(t) \quad (5)$$

Consider  $b$  as a random variable that represents the territorial factor, and assume that  $b$  follows the standard normal distribution  $N(0,1)$ .

## 3. Migration foraging

Hummingbirds migrate to distant regions when there is not enough food in the areas where they frequently feed. The AHA defines a migration coefficient to determine the best number of iterations. If the coefficient is higher, the hummingbird will abandon the old source and stay at the new one. The mathematical model of migration foraging can be expressed as:

$$x_{wor}(t + 1) = Low + r \cdot (Up - Low) \quad (6)$$

where  $x_{wor}$  is the food source with the worst nectar-refilling rate in the population.

## Optimization of gradient boosting regressor

The AHA was used to optimize the hyperparameters of the Gradient Boosting Regressor, an ML algorithm for regression tasks. The goal of the optimization was to minimize the mean squared error (MSE) between the model's predictions and the actual values.

The AHA algorithm first initializes a population of solutions, each of which represents a set of hyperparameters. The algorithm then repeatedly updates these solutions over a specified number of iterations, guided by the objective function, which calculates the MSE of the gradient boosting regressor. The updating process involves three types of foraging behaviours: diagonal flight, omnidirectional flight, and axial flight. The algorithm also uses two types of foraging strategies: guided foraging and territorial foraging. Additionally, a migration foraging step is performed every  $2 \cdot npop$  iteration, which replaces the worst solution with a new, randomly generated one.

Each solution's fitness is evaluated using the objective function, and the best solution (i.e., the one that yields the lowest MSE) is recorded and returned after all iterations are completed. Once the best hyperparameters are found, a new Gradient Boosting Regressor is trained using these hyperparameters. The model's predictions on the training and testing sets are then compared with the actual values to calculate various performance metrics, including mean absolute error (MAE), mean square error (MSE), root-mean-squared error (RMSE), and  $R$ -squared.

## Sensitivity analysis

Nowadays, the incorporation of more factors for enhancing model prediction usually increases the complexity in the determination of the dependent variable effects on the output variance. To assess and quantify uncertainty of the model parameters, the researcher utilizes sensitivity analysis tools to understand model output variance. Saltelli (2002) defines 'the study of how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input'. There are two main kinds of sensitivity analysis global sensitivity analysis (GSA) and one-factor-at-a-time

(OAT). The GSA is better than the OAT because it determines the effects of parameter interactions on the model output (Saltelli *et al.* 2019). The impact of input parameter uncertainty on the measured rainfall-runoff variance was investigated in this study using the GSA. The easyGSA MATLAB application, which offers an easy method for users with fundamental MATLAB experience, was used to conduct the GSA. The meta-modelling analysis tool called easyGSA optimizes firstly the hyperparameters after fitting the GPR or ANN model with the model parameters. Secondly, a Monte Carlo simulation is used to calculate the Sobol indices. The easyGSA's first-order Sobol indices ( $S_i$ ) measure how different parameters influence the uncertainty of the rainfall-runoff. In contrast, the Total Sobol indices ( $S_{Ti}$ ) determine the interaction impact of model parameters on the uncertainty of the measured rainfall-runoff. The first order and total Sobol indices mathematically are defined as follows;

$$S_i = \frac{V_{x_i}(E_{x_{-i}}(y|x_i))}{V(y)} \quad (7)$$

$$S_{Ti} = 1 - \frac{V_{x_i}(E_{x_{-i}}(y|x_{-i}))}{V(y)} \quad (8)$$

where  $V(y)$  is the unconditional variance of  $y$ , found when all parameters  $x_i$  are allowed to vary, and  $E_{x_{-i}}(y|x_i)$  is the average of  $y$  when one parameter is fixed. Where all the variances of  $y$  is a function  $x_i$ , hence  $\sum S_i$  equals to 1 for perfectly linear models and for non-linear models  $\sum S_i$  is less than 1. Further,  $\sum S_{Ti}$  is equal or more than 1 for a perfectly linear model and a non-linear model, respectively.

### Model evaluation methods

In the field of ML, different methods have been used to measure the error and variance in a predicted dataset without any of the methods having superiority over the other. The commonly used statistical performance indices parameter include the mean-square error (MSE), RMSE, MAE, mean absolute percentage error (MAPE), and coefficient of correlation ( $R$ ) and coefficient of determination ( $R^2$ ) (Ardabili *et al.* 2018; Shamshirband *et al.* 2019). According to Wu & Chau (2013), the evaluation of model performances should include absolute and relative error measurements of the model. Furthermore, Chadalawada & Babovic (2019) noted that a comprehensive set of performance metrics should be employed in hydrological studies to evaluate model configurations. These metrics should include accuracy metrics, correlation metrics, statistical efficiency, complexity and variability metrics, and hydrological efficiency where applicable. Such a holistic approach ensures that the model's predictive accuracy, robustness, and established relationships with observed values are adequately assessed. This study adopted the traditional RMSE to measure the average error percentage, mean absolute error (MAE) for absolute error, and the coefficient of determination ( $R^2$ ) to measure the variance between the predicted and the measured values. According to Chai & Draxler (2014), RMSE is preferable when dealing with models expecting Gaussian error distribution. Additionally, The KGE statistic and volumetric efficiency (VE) were also used to account for model fit and statistical efficiency. Mathematical expressions of RMSE, MAE, and  $R^2$  are given thus;

$$R^2 = \frac{SSR}{SST} \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Y_{Predi} - Y_{Obi})^2}{N}} \quad (10)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |Y_{Obi} - Y_{Predi}| \quad (11)$$

$$MSE = \frac{\sum_{i=1}^N (Y_{Predi} - Y_{Obi})^2}{N} \quad (12)$$

$$KGE = 1 - \sqrt{(r-1)^2 + (\beta-1)^2 + (\gamma-1)^2} \quad (13)$$

$$\beta = \frac{\mu_{Predi}}{\mu_{Obi}} \quad (14)$$

$$\gamma = \frac{CV_{Predi}}{CV_{Obi}} = \frac{\sigma_{Predi}/\mu_{Predi}}{\sigma_{Obi}/\mu_{Obi}} \quad (15)$$

$$NSE = 1 - \frac{\sum_{i=1}^N (Y_{Obi} - Y_{Predi})^2}{\sum_{i=1}^N (Y_{Obi} - Y_{Obi-mean})^2} \quad (16)$$

$$VE = 1 - \frac{\sum_{i=1}^N |Y_{Predi} - Y_{Obi}|}{\sum_{i=1}^N Y_{Obi}} \quad (17)$$

where  $N$  is the number of observed values,  $Y_{Pred1}, Y_{Pred2}, \dots, Y_{PredN}$  are predicted values;  $Y_{Ob1}, Y_{Ob2}, \dots, Y_{ObN}$  are the observed values; SSR is the sum of squared regression derived by  $\sum (Y_{Predi} - \bar{Y})^2$ ; SST is the total variation contained in the dataset derived by  $\sum (Y_{Obi} - \bar{Y})^2$  and  $\bar{Y}$  is the mean of the  $Y$  value,  $r$  is the correlation coefficient between simulated and observed runoff,  $\beta$  is the bias ratio,  $\gamma$  is the variability ratio,  $\mu$  is the mean runoff in  $m^3/s$ , CV is the coefficient of variation,  $\sigma$  is the standard deviation (SD) of runoff in  $m^3/s$ , and the indices Predi and Obi represent predicted and observed runoff.

## RESULTS AND DISCUSSIONS

This study compared four AI-based models, GPR, SVM, AHA-boosted, and boosted tree, for modelling the rainfall-runoff process. The comparison was performed using different statistical and graphical performance measures and the results are discussed as follows.

### Input variable selection

The feature importance calculated by EasyGSA shows that the problem is non-linear, see details in Figure 2. The extract  $\sum S_i$  is equal to 0.56, while  $\sum S_{Ti}$  is 2.03 indicating a non-linear relationship (Saltelli *et al.* 2008). The lagged rainfall-runoff Qt-1 and Qt-4 have a strong effect on the variation of the measured rainfall-runoff as shown in the higher  $S_i$  and  $S_{Ti}$  values which measure the individual effects of parameters and the interactive effects of model parameters on the output variance, respectively. The lagged Qt-5 and Pt parameters showed the least effect on both the individual and interactive effects of model parameters. All selected inputs were observed to have a significant effect on the model output variance, indicating high reliability in modelling rainfall-runoff.

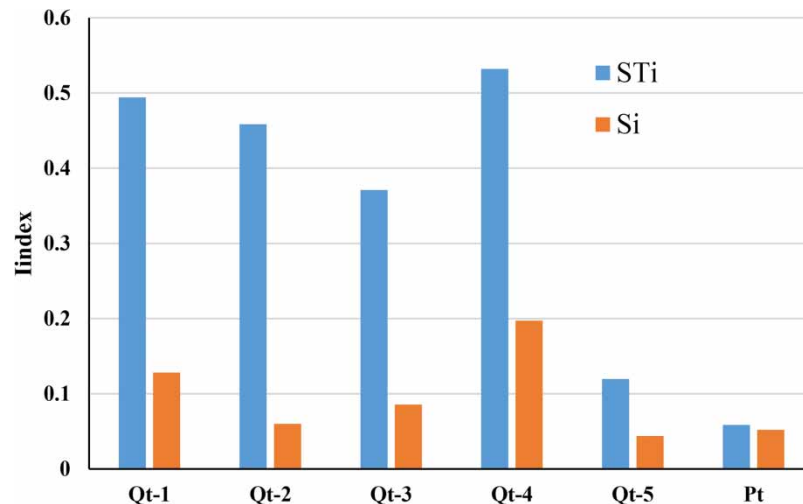


Figure 2 | Feature importance.

The performance comparison between the employed models is presented in Table 1. As shown in the table, all the models provide very good results based on their  $R^2$  value as per the guideline of Moriasi *et al.* (2015). Based on the result, the AHA-boosted model led the best prediction result with testing phase  $R^2$ , RMSE, MAE, and MSE values of 0.932, 5.358 m<sup>3</sup>/s, 2.365 m<sup>3</sup>/s, and 28.705 m<sup>3</sup>/s, respectively. The second-best prediction performance was obtained using the SVM model with  $R^2$ , RMSE, MAE, and MSE values of 0.929, 5.554 m<sup>3</sup>/s, 2.313 m<sup>3</sup>/s, and 17.262 m<sup>3</sup>/s, respectively. The least accurate performance was obtained in the GPR model.

In addition to the initial performance metrics discussed, it is worth examining other metrics that contribute to a more comprehensive understanding of the models' efficiency and robustness. Specifically, fitness metrics such as Kling–Gupta Efficiency (KGE) and Nash–Sutcliffe Efficiency (NSE), as well as the VE, which falls under the category of statistical efficiency.

From the fitness metrics, the AHA-boosted model showed a KGE of 0.922 and an NSE of 0.931, further solidifying its process in fitting the observed data effectively. The Boosted Tree model followed closely with a KGE of 0.905 and an NSE of 0.926. SVM and GPR were not far behind, registering KGE values of 0.884 and 0.943, and NSE values of 0.918 and 0.917, respectively. These fitness metrics are vital for gauging how well the model fits the observed data and add more depth to our understanding of each model's capabilities.

Furthermore, for the statistical efficiency, the VE was examined. Here, all models showed relatively similar performances. The AHA-boosted and boosted tree models both had a VE of 0.82, while the SVM slightly outperformed them with a VE of 0.83. GPR also showed a VE of 0.82. Statistical efficiency metrics like VE are crucial for evaluating the robustness of the models from a statistical viewpoint.

Taken together, these additional metrics align well with our initial findings. They not only reaffirm the AHA-boosted model's superior predictive capabilities but also offer a nuanced understanding of how each model fares across different categories of performance metrics. This multi-faceted evaluation makes a compelling case for the robustness and reliability of the AHA-boosted model in hydrological modelling.

Different studies have used various model performance measures such as statistical and graphical or a combination of both. Gelete *et al.* (2023) and Harmel *et al.* (2014) recommended the use of both graphical and statistical indices for more effective model performance evaluation. This is because some of the statistical performance measures can show good performance even when low values are poorly fitted (Moriasi *et al.* 2015). In situations like these, graphical measures can be instrumental in offering supplementary evidence, enabling precise identification of those areas where the model's performance lacks adequacy. Thus, in this study, the accuracy of the developed models was also evaluated using a timeseries plot, Taylor diagram, and scatter plot.

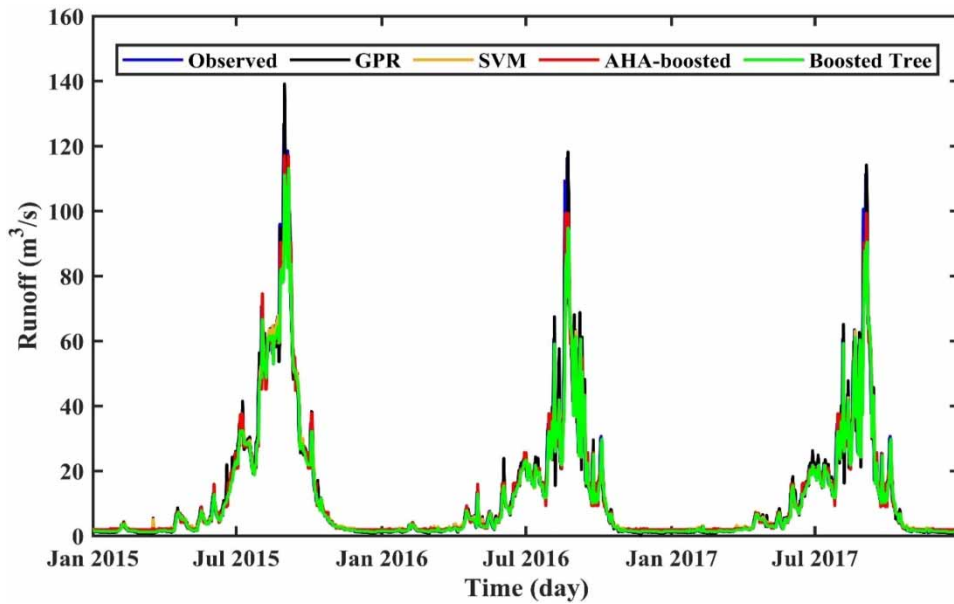
The time series plot of the developed models for rainfall-runoff modelling during the testing period is shown in Figure 3. As shown in the figure, all the models perform well in predicting the trend of the time series. However, all of the models failed to capture the peak flow of the hydrograph. For example, the observed peak discharge during the testing period that occurred on 30 August 2015, was 126.779 m<sup>3</sup>/s whereas the predicted peak runoff by GPR, SVM, AHA-boosted model, and boosted regression tree (BRT) was 92.23, 85.45, 90.2, and 79.41 m<sup>3</sup>/s, respectively. This indicates GPR, SVM, AHA-boosted model, and BRT models underestimated the peak discharge by 27.25, 32.6, 28.85, and 37.36%, respectively. Underestimation of the peak discharge by AI-based models has been reported in many studies (Hadi & Tombul 2018; Nourani *et al.* 2021; Tiban-gayuka *et al.* 2022; Gelete 2023a).

The performance of the developed models was also evaluated using a scatter plot as shown in Figure 4. As shown in the figure, in the AHA-optimized model, the data points were close to the line of equality indicating the best performance of the models. The data points were more scattered for the GPR model, especially in high-flow periods.

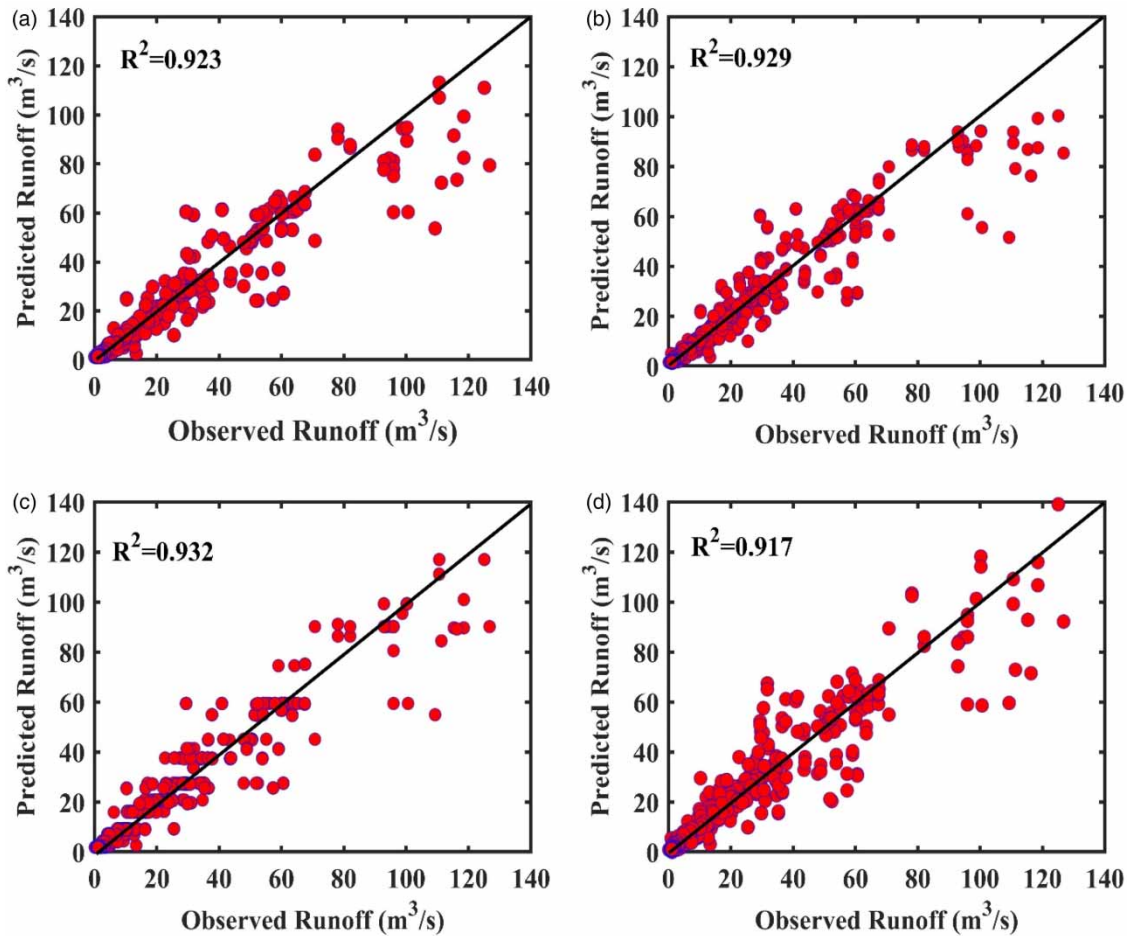
**Table 1** | Performance of the test dataset with the models

Models	$R^2$	RMSE	MAE	MSE	KGE	NSE	VE
Boosted Tree model	0.923	5.839	2.165	34.12	0.905	0.926	0.82
AHA-boosted	0.932	5.36	2.365	28.72	0.922	0.931	0.82
SVM	0.929	5.554	2.313	30.85	0.884	0.918	0.83
GPR	0.917	5.884	2.277	34.62	0.943	0.917	0.82

RMSE, MAE, and MSE are in m<sup>3</sup>/s.



**Figure 3** | The time series plot of observed and predicted runoff during the testing period.



**Figure 4** | Scatter plot of observed runoff versus predicted runoff using (a) BRT, (b) SVM, (c) AHA-boosted, and (d) GPR during the testing period.



A comparison of the model accuracy was also performed in the testing phase based on a two-dimensional Taylor's diagram, as shown in Figure 5. This diagram summarizes and highlights different statistical performance indices such as correlation coefficient ( $r$ ) and SD in a single graph to compare the deviation of predicted values from the actual value (Taylor 2001). From Figure 5, it can be seen that the better goodness of fit of runoff modelling was obtained from the AHA-boosted model value of  $r=0.9657$  and SD of  $19.03 \text{ m}^3/\text{s}$ , followed by SVM ( $r=0.9637$  and  $\text{SD}=18.67 \text{ m}^3/\text{s}$ ), BRT ( $r=0.96$  and  $\text{SD}=18.51 \text{ m}^3/\text{s}$ ) and GPR ( $r=0.958$  and  $\text{SD}=19.71 \text{ m}^3/\text{s}$ ).

## CONCLUSIONS

In this research, we examined the efficacy of four AI models – GPR, BRT, SVM, and AHA-boosted models – in modelling the rainfall-runoff process for the Katar catchment in Ethiopia. The study was executed in two phases. The first phase involved selecting dominant inputs, while the second phase entailed a comparative performance assessment of the developed models using statistical and graphical measures. While all models demonstrated strong predictive performance, the AHA-boosted model emerged as the most accurate, showcasing superior  $R^2$ ,  $R$ , RMSE, MAE, and MSE values, and also in terms of fitness and statistical efficiency using NSE, KGE, and VE.

However, it is crucial to highlight the role of automatic calibration in this study, as it is a vital component in hydrological modelling that has been under scrutiny for nearly three decades. Our AHA-boosted model leveraged advanced optimization techniques for automatic calibration, enhancing both the model's accuracy and reliability. This contributes to the existing body of work by offering a new calibration technique that aligns closely with the current needs and complexities of hydrological modelling.

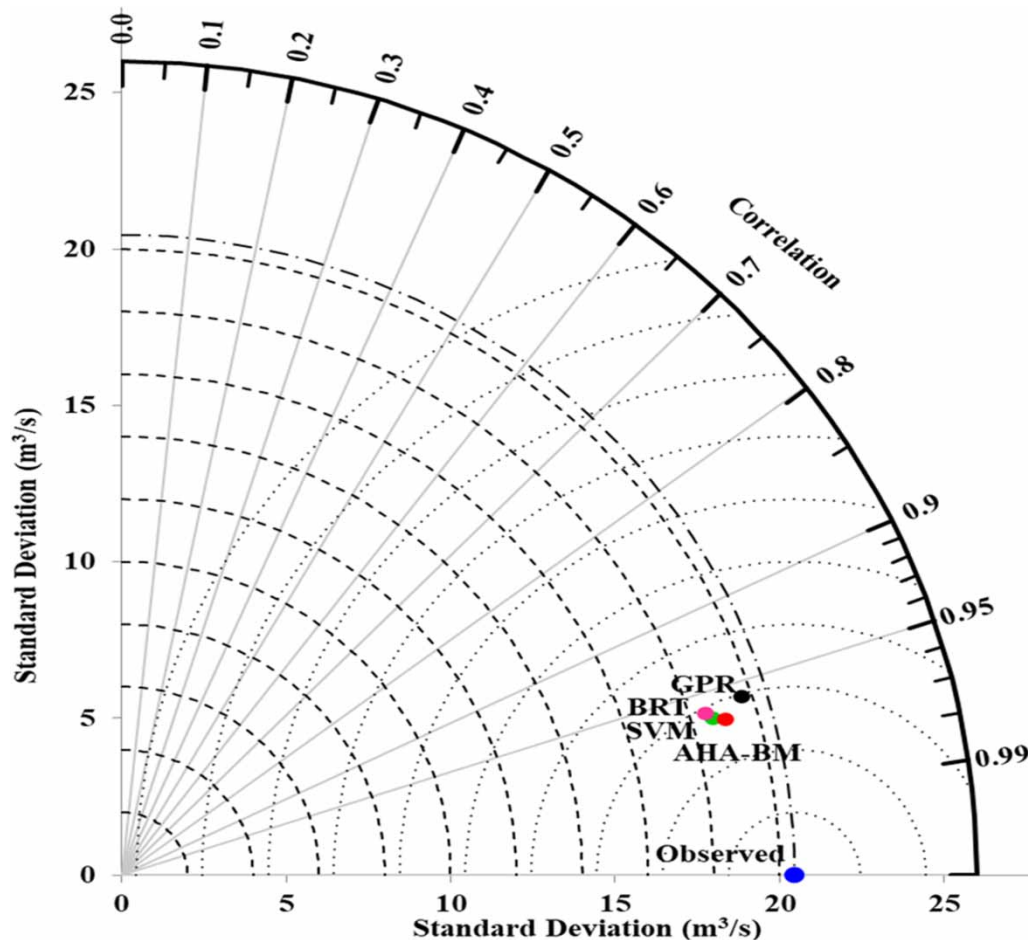


Figure 5 | Taylor's diagram for the model test performance.

The second most accurate model was SVM, followed by BRT and GPR. These results accentuate the potential advantages of optimization techniques, particularly the AHA-boosted model, in predicting rainfall-runoff processes.

For future work, we recommend extending this research to include more traditional models for comparative analysis. Given the strong performance of the AHA-boosted model in this study, it would be worthwhile to apply it to other hydrological processes. Furthermore, we suggest exploring other optimization algorithms in conjunction with emerging AI models to further improve modelling accuracy.

In summary, this research underlines the significance of automatic calibration in hydrological models and offers a new perspective on employing advanced optimization techniques for this purpose.

## DATA AVAILABILITY STATEMENT

Data cannot be made publicly available; readers should contact the corresponding author for details.

## CONFLICT OF INTEREST

The authors declare there is no conflict.

## REFERENCES

- Anusree, K. & Varghese, K. O. 2016 Streamflow prediction of Karuvannur River Basin using ANFIS, ANN and MNLN models. *Procedia Technol.* **24**, 101–108. Elsevier. <https://doi.org/10.1016/J.PROTCY.2016.05.015>.
- Ardabili, S. F., Najafi, B., Shamshirband, S., Bidgoli, B. M., Deo, R. C. & Chau, K. W. 2018 Computational intelligence approach for modeling hydrogen production: A review. *Eng. Appl. Comput. Fluid Mech.* **12** (1), 438–458. Taylor and Francis Ltd. <https://doi.org/10.1080/19942060.2018.1452296>.
- Babovic, V., Wu, Z. & Larsen, L. C. 1994 Calibrating hydrodynamic models by means of simulated evolution. In: *Hydroinformatics '94. Proc. 1st Int. Conf. Delft, 1994*. Vol. 1.
- Besaw, L. E., Rizzo, D. M., Bierman, P. R. & Hackett, W. R. 2010 Advances in ungauged streamflow prediction using artificial neural networks. *J. Hydrol.* **386** (1–4), 27–37. Elsevier. <https://doi.org/10.1016/J.JHYDROL.2010.02.037>.
- Chadalawada, J. & Babovic, V. 2019 Review and comparison of performance indices for automatic model induction. *J. Hydroinf.* **21** (1), 13–31. IWA Publishing. <https://doi.org/10.2166/HYDRO.2017.078>.
- Chai, T. & Draxler, R. R. 2014 Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature. *Geosci. Model Dev.* <https://doi.org/10.5194/gmd-7-1247-2014>.
- Gelete, G. 2023a Application of hybrid machine learning-based ensemble techniques for rainfall–runoff modeling. *Earth Sci. Inf.* (0123456789). Springer Berlin Heidelberg. <https://doi.org/10.1007/s12145-023-01041-4>.
- Gelete, G. 2023b Application of hybrid machine learning-based ensemble techniques for rainfall–runoff modeling. *Earth Sci. Inf.* **1**, 1–21. Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/S12145-023-01041-4/FIGURES/14>.
- Gelete, G., Nourani, V., Gokcekus, H. & Gichamo, T. 2023 Ensemble physically based semi-distributed models for the rainfall-runoff process modeling in the data-scarce Katar catchment, Ethiopia. *J. Hydroinf.* **25** (2), 567–592. IWA Publishing. <https://doi.org/10.2166/HYDRO.2023.197/1182440/JH2023197.PDF>.
- Hadi, S. J. & Tombul, M. 2018 Forecasting daily streamflow for basins with different physical characteristics through data-driven methods. *Water Resour. Manage.* **32** (10), 3405–3422. Springer Netherlands. <https://doi.org/10.1007/S11269-018-1998-1/FIGURES/7>.
- Harmel, R. D., Smith, P. K., Migliaccio, K. W., Chaubey, I., Douglas-Mankin, K. R., Benham, B., Shukla, S., Muñoz-Carpena, R. & Robson, B. J. 2014 Evaluating, interpreting, and communicating performance of hydrologic/water quality models considering intended use: A review and recommendations. *Environ. Model. Softw.* **57**, 40–51. Elsevier Ltd. <https://doi.org/10.1016/j.envsoft.2014.02.013>.
- Kisi, O., Shiri, J. & Nikoofar, B. 2012 Forecasting daily lake levels using artificial intelligence approaches. *Comput. Geosci.* **41**, 169–180. Pergamon. <https://doi.org/10.1016/J.CAGEO.2011.08.027>.
- Kisi, O., Shiri, J. & Tombul, M. 2013 Modeling rainfall-runoff process using soft computing techniques. *Comput. Geosci.* **51**, 108–117. Pergamon. <https://doi.org/10.1016/J.CAGEO.2012.07.001>.
- Moriasi, D. N., Gitau, M. W., Pai, N. & Daggupati, P. 2015 Hydrologic and water quality models: Performance measures and evaluation criteria. *Trans. ASABE* **58** (6), 1763–1785. <https://doi.org/10.13031/trans.58.10715>.
- Nawaz, N., Harun, S., Talei, A. & Chang, T. K. 2016 Event-based rainfall-runoff modeling using adaptive network-based fuzzy inference system. *J. Teknol.* **78** (9–4), 41–46. Penerbit UTM Press. <https://doi.org/10.11113/JT.V78.9693>.
- Nourani, V. & Komasi, M. 2013 A geomorphology-based ANFIS model for multi-station modeling of rainfall–runoff process. *J. Hydrol.* **490**, 41–55. Elsevier. <https://doi.org/10.1016/J.JHYDROL.2013.03.024>.
- Nourani, V., Gökçekuş, H. & Gichamo, T. 2021 Ensemble data-driven rainfall-runoff modeling using multi-source satellite and gauge rainfall data input fusion. *Earth Sci. Inf.* **14** (4), 1787–1808. Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1007/S12145-021-00615-4/FIGURES/12>.

- Saltelli, A. 2002 Sensitivity analysis for importance assessment. *Risk Anal.* **22** (3), 579–590. John Wiley & Sons, Ltd. <https://doi.org/10.1111/0272-4332.00040>.
- Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M. & Tarantola, S. 2008 *Global Sensitivity Analysis. The Primer*. John Wiley and Sons, Chichester, UK.
- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S. & Wu, Q. 2019 Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environ. Model. Softw.* **114**, 29–39. Elsevier Ltd. <https://doi.org/10.1016/j.envsoft.2019.01.012>.
- Savic, D. A. & Walters, G. A. 1995 *Genetic Algorithm Techniques for Calibrating Network Models. Report*.
- Shamshirband, S., Rabczuk, T. & Chau, K. W. 2019 A survey of deep learning techniques: Application in wind and solar energy resources. *IEEE Access* **7**, 164650–164666. Institute of Electrical and Electronics Engineers Inc. <https://doi.org/10.1109/ACCESS.2019.2951750>.
- Soon, T. K. & Madsen, H. 2005 Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration. *Water Resour. Res.* **41** (3). <https://doi.org/10.1029/2004WR003041>.
- Subramani, N., Easwaramoorthy, S. V., Mohan, P., Subramanian, M. & Sambath, V. 2023 A gradient boosted decision tree-based influencer prediction in social network analysis. *Big Data Cognit. Comput.* **7** (1), 6. Multidisciplinary Digital Publishing Institute. <https://doi.org/10.3390/BDCC7010006>.
- Taylor, K. E. 2001 Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* **106** (D7), 7183–7192. <https://doi.org/10.1029/2000JD900719>.
- Tibangayuka, N., Mulungu, D. M. M. & Izdori, F. 2022 Evaluating the performance of HBV, HEC-HMS and ANN models in simulating streamflow for a data scarce high-humid tropical catchment in Tanzania. *Hydrol. Sci. J.* **67** (14), 1–14. Taylor & Francis. <https://doi.org/10.1080/02626667.2022.2137417>.
- Uwanuakwa, I. D. & Akpınar, P. 2020 Investigations on the influence of variations in hidden neurons and training data percentage on the efficiency of concrete carbonation depth prediction with ann. *Adv. Intell. Syst. Comput.*, 958–965. Springer.
- Ward, B. J., Day, L. B., Wilkening, S. R., Wylie, D. R., Saucier, D. M. & Iwaniuk, A. N. 2012 Hummingbirds have a greatly enlarged hippocampal formation. *Biol. Lett.* **8** (4), 657–659. The Royal Society. <https://doi.org/10.1098/RSBL.2011.1180>.
- Wu, C. L. & Chau, K. W. 2013 Prediction of rainfall time series using modular soft computing methods. *Eng. Appl. Artif. Intell.* **26** (3), 997–1007. Pergamon. <https://doi.org/10.1016/j.engappai.2012.05.023>.
- Zhang, J., Chen, X., Khan, A., Zhang, Y.-k., Kuang, X., Liang, X., Taccari, M. L. & Nuttal, J. 2021 Daily runoff forecasting by deep recursive neural network. *Journal of Hydrology* **596**, 126067. <https://doi.org/10.1016/J.JHYDROL.2021.126067>
- Zhao, W., Wang, L. & Mirjalili, S. 2022 Artificial hummingbird algorithm: A new bio-inspired optimizer with its engineering applications. *Comput. Methods Appl. Mech. Eng.* **388**, 114194. North-Holland. <https://doi.org/10.1016/j.jhydrol.2021.126067>.

First received 31 July 2023; accepted in revised form 30 November 2023. Available online 13 December 2023