Water Science & Technology

Water Science & Technology Vol 87 No 3, 812 doi: 10.2166/wst.2023.014

Comparing machine-learning-based black box techniques and white box models to predict rainfall-runoff in a northern area of Iraq, the Little Khabur River

Biju Theruvil Sayed[a,*], Hani Kaid Al-Mohair[b], Ahmed Alkhayyat[c], Andrés Alexis Ramírez-Coronel[d] and Mohamed Elsahabi[e]

[a] Department of Computer Science, Dhofar University, Salalah, Oman
[b] Computer and Information Technology, Dammam Community College, King Fahd University of Petroleum & Minerals, Dhahran, Kingdom of Saudi Arabia
[c] College of Technical Engineering, Computer Technology Department, The Islamic University, Najaf, Iraq
[d] CES University, Medellín, Colombia
[e] Civil Engineering Department, Faculty of Engineering, Aswan University, Aswan 81542, Egypt
*Corresponding author. E-mail: b_sayed@du.edu.om

ABSTRACT

The rainfall-runoff process is one of the most complex hydrological phenomena. Estimating runoff in the basin is one of the main conditions for planning and optimal use of rainfall. Using machine learning models in various sciences to investigate phenomena for which statistical information is available is a helpful tool. This study investigates and compares the abilities of HEC-HMS and TOPMODEL as white box models and adaptive neural fuzzy inference system (ANFIS) and gene expression programming (GEP) as black box models in rainfall-runoff simulation using 5-year statistical data. Using the inputs of rainfall and temperature of the previous day and discharge in the steps of the previous 2 days reduced the prediction error of both models. Examining the role of different parameters in improving the accuracy of simulations showed that the temperature as an effective parameter in cold months reduces the amount of prediction error. A comparison of R^2, RMSE, and MBE showed that black box models are more effective forecasting tools. Among the black box models, the ANFIS model with $R^2 = 0.82$ has performed better than the GEP model with $R^2 = 0.76$. For white box models, the HEC-HMS and TOPMODEL had R^2 equal to 0.3 and 0.25, respectively.

Key words: ANFIS, GEP, HEC-HMS, rainfall-runoff, TOPMODEL

HIGHLIGHT

- In the current study, as a novel strategy, it was tried to compare machine-learning-based black box techniques and white box models to predict rainfall-runoff in the Northern area of Iraq (case study: The Little Khabur River)

1. INTRODUCTION

Runoff from rainfall is one of the most significant sources of water supply in most regions of the world, particularly in arid and low-rainfall regions. In areas with arid to semi-arid climatic conditions, scientific methods of collecting and controlling surface water to provide water and prevent damage become increasingly essential. Hence, a more accurate estimation of the runoff resulting from rainfall in catchment basins is crucial (Yavari *et al.* 2022). Knowledge of the natural power of runoff production in basins is one of the essentials for basic planning for the optimal use of runoff (Li *et al.* 2018; Liu & Lu 2022). In dry countries where environmental issues and water scarcity are increasing, it is necessary to know about water resources in order to analyze the relationship between food production, economic activities, and the ecosystem (Khan & Coulibaly 2006). Also, many studies have been conducted on the impacts of runoff on the water quality of reservoirs and rivers and the resulting problems. It is easier to predict floods by simulating the rainfall-runoff process. Approximately 40% of natural hazards are related to floods, and every year 20–300 million people are affected by problems caused by floods. It is predicted that by 2050, floods will cause a thousand billion dollars in annual damage (Młyński *et al.* 2018; Youssef *et al.* 2021). These issues show the importance of rainfall-runoff predicting.

Forecasting the rainfall-runoff process has been one of the complex issues in water resources management. Studies have shown that optimization algorithms like Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), neural networks, machine learning, and artificial intelligence can be suitable solutions to solve complex problems related to water resources like, simulation rainfall-runoff, coastal engineering, and safety in water supply networks (Kamranzad & Samadi 2013; Nourani *et al.* 2019b; Shahmirnoori *et al.* 2022). In recent years, various rainfall-runoff models have been developed and modified. Runoff rainfall models include artificial intelligence, conceptual, experimental, and physical models. Conceptual or white box models are based on physical regulations and can represent hydrological behaviors with practical expression. Compared to physical models, white box models require less data, and also, these models require less calculation volume and time (Kamranzad & Samadi 2013; Bartoletti *et al.* 2018). In modeling hydrological phenomena, there are two conceptual approaches: white box and black box models. White box models are presented based on governing mathematical relationships, equations, and physical parameters in the phenomenon. Models such as HEC-HMS, TOPMODEL, and HBV are conceptual models. These models can use fewer physical parameters than other models to estimate the appropriate response of the catchment area against precipitation, which is an advantage of these models (Abdessamed & Abderrazak 2019). White box models are presented based on governing mathematical relationships, equations, and physical parameters in the phenomenon. The purpose of white box models is to conduct scientific research on each component of the hydrological cycle's main operation to fully understand the mechanism and how the components work with each other. In black box models, it is impossible to present equations and mathematical relationships. In these models, the physical parameters affecting it cannot be estimated easily. This group aims to develop reliable connections between the measured parameters of the hydrological cycle, which are used to solve technical and scientific problems (Nourani *et al.* 2019a).

In black box models, it is impossible to present equations and mathematical relationships. In these models, the physical parameters affecting it cannot be estimated easily. This group aims to develop reliable connections between the measured parameters of the hydrological cycle, which are used to solve technical and scientific problems. Black box models estimate the intended output by receiving input and performing mathematical operations. All black box models have parameters and coefficients estimated according to the input and output data of observations. Therefore, black box models depend on input and output data in quantity and quality (Andrews & Boyne 2010; Nourani *et al.* 2019b). Due to the nonlinear nature of most hydrological processes, nowadays, artificial neural networks (ANN) and machine learning as self-constructed estimator functions are widely used in modeling and forecasting nonlinear time series in hydrological processes. In general, the primary and most obvious advantages of black box models compared to other statistical and conceptual methods can be categorized in the form of things such as not requiring prior knowledge, applying a nonlinear filter under the title of stimulus function, and the ability to process multivariate inputs with different characteristics (Molajou *et al.* 2021).

In the study of Hamdan *et al.* (2021) which took place in northern Iraq, the researchers used the HEC model to assess rainfall-runoff. The results confirmed the performance of the model for rainfall-runoff simulation. Jaber *et al.* (2017) analyzed rainfall-runoff in a part of Iraq. They used Geographic Information System (GIS) and remote sensing to analyze the data and to produce the runoff depth map for the study area. Misra *et al.* (2009) compared the support vector machine model with an artificial neural network to forecast runoff and sediment amounts in India. The results showed that the support vector machine and neural network performed well predicting runoff and sediment load, but the support vector machine showed relatively better performance. Adamowski (2013) compared a support vector machine and an artificial neural network in rainfall-runoff modeling in India by calculating RMSE, R, Slop, intercept, and MBE. The results revealed that both models were suitable. Veintimilla-Reyes *et al.* (2016) used two neural network types in their study, and Anusree & Varghese (2016) used three neural network models to predict rainfall-runoff. The results showed the neural networks' positive performance in their research. Nourani (2017) employed a new type of neural network to forecast rainfall-runoff in their research. Humphrey *et al.* (2016) predicted rainfall-runoff by combining a conceptual model and a neural network. Miguélez *et al.* (2009) predicted urban runoff by combining ANN with the GA. This study used an artificial intelligence technique to model precipitation and surface runoff production for an urban area. The results of their work showed that runoff modeling using ANN had satisfactory results.

The literature review shows that in most research, only one conceptual model has been compared with the artificial intelligence model, and more models have not been used in the studies. However, examining more models and examining them with one another seems necessary because, based on the conditions in a region, different models can have different results in rainfall-runoff results, which should be considered. Also, considering that in most research, only two separate models have been compared, it is impossible to analyze the performance of two groups of conceptual models (white box) and black box with high accuracy. In fact, in a watershed, one conceptual model may be less accurate than a black box model but higher than another. In addition to the mentioned points, in previous research, ANN and support vector machines have been mainly used to investigate rainfall-runoff. The application of two ANFIS and GEP methods together has not been studied in past studies, which is considered in this study. Therefore, in this research, rainfall-runoff in Iraq is simulated with two conceptual models, HEC-HMS and TOPMODEL, and two black box models, Adaptive Neural Fuzzy Inference System (ANFIS) and Gene Expression Programming (GEP). In this study, statistical data of 5 years (2015–2020) have been used to estimate rainfall-runoff on a daily scale.

2. MATERIALS AND METHODS

2.1. Study area

The study area in this research is the Little Khabur River. This river is located in the north of Iraq, in the Kurdistan region and Dohuk province. The average rainfall in this area is 940 mm per year. This area has a warm Mediterranean climate with hot summers and cool winters with high rainfall. On average, the temperature reaches 40°C in the summer and −3°C in the coldest seasons. The summers of this region are hot and dry, and sometimes there is no rain in the summer. This river flows towards the southwest, and after passing through Zakho city, it flows into the Hezil Suyu river (Ewaid *et al.* 2020). Figure 1 shows the geographical location of the Little Khabur River in Iraq.

2.2. HEC-HMS model

HEC-HMS is one of the most reliable software in the field of hydrology, which uses several mathematical models to calculate the rainfall-runoff hydrological model, flood trends, infiltration, and other cases. This model is one of the mathematical models with several sub-models in the sections of surface runoff, base flow, and channel flow and is used to simulate the hydrological behavior of watersheds. HEC-HMS has the main sections, including the general outline of the watershed, characteristics of sub-basins such as the area and type of hydrograph and flows modeling, the essential characteristics of waterways such as loss rates and trending methods, characteristics of control sites, climate model flow; Including meteorological stations, the method of effecting the amount of precipitation of each station on the amount of water and control indicators (Halwatura & Najim 2013; Wildhaber *et al.* 2017). Lu *et al.* (2013) and Wildhaber *et al.* (2017) used this program in their research and reported its results as favorable. In general, HEC-HMS software is one of the most well-known programs for runoff simulation, and it is widely used in the United States. Different methods include the Clark hydrograph, the Schneider unit hydrograph, and the SCS unit hydrograph to convert rainfall into runoff. It is inexpensive to use and has the ability for automatic calibration (Wildhaber *et al.* 2017).

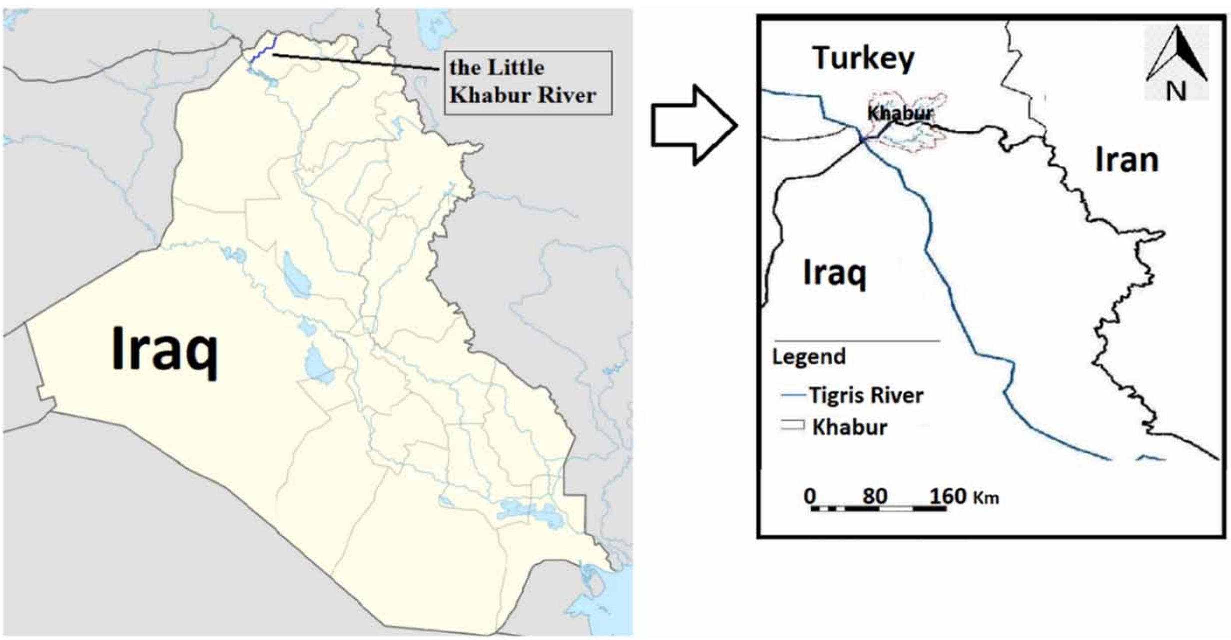

Figure 1 | The geographical location of the Little Khabur River in Iraq.

2.3. TOPMODEL model

TOPMODEL is a semi-distributive model in which the concept of a hydrological response unit based on topographic moisture index is used for basin homogenization. The topographical index usually introduces the topographical information used in this model. This index shows the tendency of flow accumulation and its movement in the direction of the downstream slope by the force of gravity. In this model, it is assumed that the hydraulic conductivity of the saturated zone can be estimated using the topographic slope of the earth's surface, and the relationship between the soil transfer ability and the depth is a power function proportional to the lack of saturation. It has been widely used to model precipitation and runoff in the past decades. This model's inputs include evaporation and transpiration, basin runoff height potential, precipitation, basin delay function, and topographic humidity index. In order to optimize the model parameters, the random jump method was used. TOPMODEL model is one of the standard models in hydrological analysis in many European countries (Campling *et al.* 2002; Khan & Coulibaly 2006).

2.4. Adaptive neural fuzzy inference system

ANFIS combines the advantages of ANN and fuzzy logic (FL) in a single framework. It provides rapid learning and adaptive interpretation capabilities for modeling complex patterns and understanding nonlinear relationships. ANFIS has been applied and practiced in various fields and has provided solutions to common iterative problems with improved temporal and spatial complexity. Therefore, it can effectively simulate rainfall-runoff, which is a complex problem. If the output of each neuro-fuzzy network layer is shown as $Q_{i,j}$, which is the output of the ith group in the first layer, then the performance of the different layers can be expressed as follows (Yaseen *et al.* 2018; Adnan *et al.* 2021):

Layer 1: Each node in this layer is equivalent to a fuzzy set (Equation (1)).

$$Q_i^1 = \mu_{Ai}(x) \tag{1}$$

Layer 2: In these layers, the input signals are multiplied and produce output (Equation (2)).

$$w_i = \mu_{Ai}(x) \times \mu_{Bi}(y) \quad i = 1, 2 \tag{2}$$

Layer 3: In this layer, calculate the ratio of the activity degree of the i-th rule to the sum of the activity degrees of all the rules (Equation (3)).

$$\bar{\bar{\bar{W}}}_i = \frac{w_i}{w_1 + w_2}, \qquad i = 1, 2 \tag{3}$$

Layer 4: The output of each node in this layer is as follows (Equation (4)).

$$Q_i^4 = \bar{W}_i f_i = \overline{W_i}(p_i x + q_i y + r_i) \tag{4}$$

where $\bar{W}_i$ is the output of the third layer and $\{p_i, q_i, r_i\}$ are the set of adaptive parameters of this layer. These parameters are called result parameters.

Layer 5: Every node in this layer, which is shown as $\sum$, calculates the final output value in the form of Equation (5).

$$Q_i^5 = \sum_i \bar{W}_i f_i = \frac{\sum_i W_i f_i}{\sum_i W_i} \tag{5}$$

2.5. Gene expression programming

Gene Expression Programming is a generalized GA that was first proposed based on Darwin's theory. In this way, in the direction of evolution, a population selectively abandons the unsuitable population and creates modified children. Tree structures are created from the set of functions (mathematical operators used in formulas) and terminals (problem variables and fixed numbers). Before the execution steps of GEP, the following preliminary steps must be determined by the user: 1 – The set of terminals (random fixed number problem variables), 2 – The set of mathematical operators used in the formulas, 3 – Selecting the fit function to measure the fit of the formulas, 4 – Determining Controlling parameters of the program execution, 5 – Criterion of the end and presentation of the results of the program execution, such as the number of new population production, determining a specific value for fitting the formulas, if the reasonable rate is equal to or greater than that value, the execution will be stopped. One of the positive points of the GEP method is that the search operators of GEP always generate functional structures and are fitted to genetic diversity. Another positive point of GEP is its unique, multigenic nature which permits the development of more complex programs composed of several subprograms. So GEP can be a suitable method for estimating rainfall-runoff simulation, considered a complex problem (Fernando *et al.* 2012; Molajou *et al.* 2021). In this research, in order to simulate the rainfall-runoff process with the GEP model, GeneXproTools 4.0 software is utilized.

2.6. Combining the inputs of black box models

The data include the daily runoff, rainfall, and temperature during the 5-year period. The inputs of the black box models are based on previous studies and information from relevant authorities. Based on this, seven scenarios have been considered. Table 1 shows the composition of input data. P is rainfall, T is temperature, and Q is stream flow. 80% of the total data was used for training and the remaining 20% was used for testing.

2.7. Evaluation criteria

The statistical parameters of root mean square error, coefficient of determination, and the mean bias error were used to check the models. To calculate the root mean square error, coefficient of determination, and the mean bias error, Equations (6)–(8 were used, respectively (Adamowski 2013; Nourani *et al.* 2019b; Azizi & Nejatian 2022):

$$\mathrm{RMSE} = \sqrt{\frac{\sum (Q_e - Q_o)^2}{N}} \tag{6}$$

$$R^2 = 1 - \frac{\sum (Q_e - Q_o)^2}{\sum (Q_e - \bar{Q}_o)^2} \tag{7}$$

$$\mathrm{MBE} = \frac{\sum_{i=1}^{N} (Q_e - Q_0)}{N} \tag{8}$$

where N is the number of data, Q_e is the simulated flow and Q_o is the observed flow.

3. RESULTS AND DISCUSSION

3.1. White box models

The results of the evaluation criteria of HEC-HMS and TOPMODEL models are shown in Table 2. A 5-year data was used for training, and 1-year data for testing. The error values of both models obtained in the test phase were lower than in training, which shows that the models could not be trained well and could not simulate the runoff with the accuracy of the training phase. The error values of both HEC-HMS and TOPMODEL models are close to each other in the test phase, and both

Table 1 | Combination of inputs

Scenarios	Input	Output
1	$Q(t-1)$	$Q(t)$
2	$Q(t-1)Q(t-2)$	$Q(t)$
3	$Q(t-1)Q(t-2)P(t-1)$	$Q(t)$
4	$Q(t-1)Q(t-2)T(t-1)$	$Q(t)$
5	$Q(t-1)Q(t-2)P(t-1)P(t-2)$	$Q(t)$
6	$Q(t-1)Q(t-2)T(t-1)T(t-2)$	$Q(t)$
7	$Q(t-1)Q(t-2)P(t-1)T(t-1)$	$Q(t)$

Table 2 | Results of R^2, MBE, and RMSE for HEC-HMS and TOPMODEL models

	R^2		MBE		RMSE	
	HEC-HMS	TOPMODEL	HEC-HMS	TOPMODEL	HEC-HMS	TOPMODEL
Train	0.35	0.14	−0.18	−0.0015	1.72	1.83
Test	0.24	0.16	−0.70	0.11	2.55	2.49

models predict runoff with little accuracy. The RMSE of HEC-HMS is higher than TOPMODEL. The most significant difference is related to the MBE, which indicates that TOPMODEL estimated the rainfall-runoff much more than the actual value.

Figure 2 illustrates the simulation outcomes of two white box models in the test phase. Both models have estimated the runoff in cooler seasons (autumn and winter) as lower than the actual value. Both models assessed the peak discharge as much higher than the actual value in the summer season. In the drier seasons (spring and summer), TOPMODEL has simulated rainfall-runoff changes with a constant trend throughout most of the period and has shown the rainfall-runoff much more than the actual value. However, the HEC-HMS model shows a better performance in simulating rainfall-runoff in drier seasons.

3.2. Black Box models

The results of the performance of ANFIS and GEP with different scenarios are displayed in Table 3. The results in this table demonstrate that the discharge input one day before cannot predict the rainfall-runoff well, and with the addition of the discharge 2 days before, the ANFIS error value stays constant, and GEP shows a decrease. According to the results of scenarios 1 and 2, both rainfall models of the previous day were added to scenario 2, and scenario 3 was implemented. Comparing the error rate of scenario 3 with scenario 2 in the test phase shows that ANFIS and GEP models have reduced the error amount, which is better results, and rainfall and discharge significantly affected the accuracy of runoff estimation.

Scenario 4 was done by mixing the flow rate input up to two previous time steps and the previous day's temperature. The results revealed that the ANFIS model simulated the runoff with the lowest error value and the highest R in the test phase. Nevertheless, the accuracy of the GEP simulation with scenario 4 and the addition of the previous day's temperature were similar to the previous day's rainfall (scenario 3).

Scenarios 5, 6, and 7 with four input parameters could not provide more favorable results. In some cases, they provided similar or weaker results than scenarios 3 and 4 with fewer input parameters. It should be noted that with the increase in the number of inputs, the computing speed of the model decreased, so running the model with a smaller number of inputs

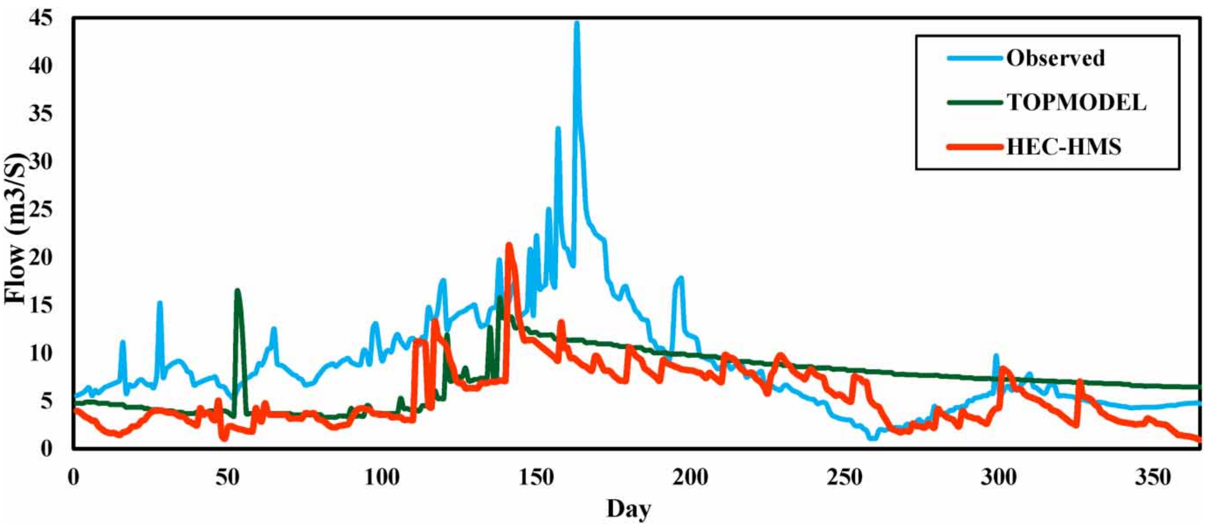

Figure 2 | Daily observed flow and flow simulated with TOPMODEL and HEC-HMS in the test phase.

Table 3 | Results of R^2, MBE, and RMSE for ANFIS and GEP models

Scenarios		R^2		MBE		RMSE	
		ANFIS	GEP	ANFIS	GEP	ANFIS	GEP
1	Train	0.55	0.53	0	−0.13	1.37	1.71
	Test	0.65	0.68	−0.14	−0.25	1.41	1.68
2	Train	0.56	0.64	0	−0.16	1.35	1.75
	Test	0.57	0.72	−0.16	−0.30	1.38	1.63
3	Train	0.70	0.66	0	−0.09	1.15	1.47
	Test	0.77	0.76	−0.01	−0.14	1.24	1.50
4	Train	0.65	0.65	0	−0.16	1.26	1.29
	Test	0.82	0.76	−0.06	−0.06	1.32	1.53
5	Train	0.75	0.78	0	−0.05	1.11	1.10
	Test	0.77	0.76	−0.01	−0.15	1.46	1.57
6	Train	0.71	0.67	0	−0.11	1.2	1.24
	Test	0.76	0.75	−0.02	−0.12	1.54	1.54
7	Train	0.68	0.61	0	−0.04	1.1	1.35
	Test	0.68	0.74	−0.11	−0.13	1.79	1.59

is desirable if the results are similar. Adding the rainfall 2 days before to the ANFIS model (scenario 5) improved the results, but its effect was less than the previous day's temperature. Also, by adding the temperature two days before (scenario 6), the accuracy of the simulation decreased. By adding the rainfall two days before to scenario 3, scenario 7 was created, which slightly reduced the accuracy of GEP and increased the accuracy of ANFIS.

The previous day's temperature and rainfall in scenarios (3 and 4) had a similar effect on the simulation of rainfall-runoff in the GEP model. However, in the ANFIS model, the effect of the previous day's temperature on the improvement of runoff simulation results was more significant than the previous day's rainfall. ANFIS in scenario 4 and GEP in scenarios 3 and 4 were selected as acceptable scenarios. Scenario 4 in ANFIS with RMSE 1.32, R^2 0.82, and MBE −0.06 provided the best result in the test stage. The RMSE of the best GEP scenario is about 16% higher than that of ANFIS, and ANFIS has been able to simulate the runoff more accurately.

Figure 3 shows the observed values and flow simulation related to scenarios 3 and 4 in the test phase. According to this figure, it can be seen that in scenario 3, the fitting line of ANFIS and GEP models did not differ much in the low flows.

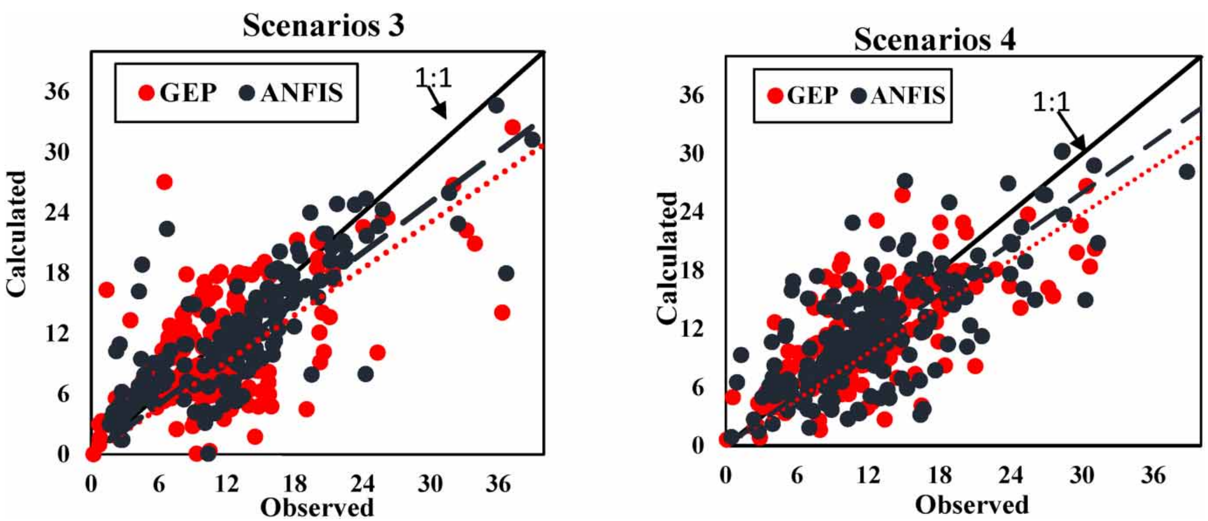

Figure 3 | The computational and observational fit line at the test phase in scenarios 3 and 4.

They had the same performance in simulating low flows, but the fitting line of the ANFIS model for flows higher than 18 m^3/s is closer to the 1:1 line. This is due to the better prediction of high discharges by the ANFIS model. In general, both models predict the runoff as less than the actual value, and the negative value of the MBE error also confirms this issue. However, the underestimation in GEP is more than ANFIS.

3.3. Comparing results of black box and white box models

Table 4 illustrates the performance of four study models in simulating the rainfall-runoff process. The error values of rainfall-runoff process simulation in black box models are lower than in white box models. Also, black box models had lower R^2 values than white box models. Figure 3 demonstrates that ANFIS and GEP could estimate the rainfall-runoff process more accurately, showing the power of black box models in analyzing the rainfall-runoff process and more accurate simulation. While the RMSE of ANFIS was lower than the RMSE of GEP, the accuracy of both ANFIS and GEP methods was acceptable. According to the results of this research, it can be concluded that the two white box models (HEC_HMS and TOPMODEL) could not simulate the dominant processes of the basin due to their structure. According to the available data, the black box models' flexibility was generally higher than the white box models. One of the reasons for the good performance of black box models can be considered as their inputs, which are the flow rates in the previous time step, which significantly impact the rainfall-runoff simulation's accuracy. Due to the fact that in black box models, only input and output data are examined, the presence of noisy data significantly affects the model's performance.

In general, the results of this research showed that black box models could work well in rainfall-runoff estimation because it was shown by Adamowski (2013) and Veintimilla-Reyes *et al.* (2016) that machine learning is a suitable tool for predicting rainfall and runoff. This study also demonstrated that the ANFIS model is suitable for rainfall-runoff estimation. The positive performance of the ANFIS model in estimating runoff precipitation was also reported in the research of Nourani & Komasi 2013. The better performance of the black box model compared to the white box model in Khan & Coulibaly (2006) study has the same result. They showed that the ANFIS model could more accurately estimate the rainfall-runoff process than the HBV model.

To investigate the effect of rainfall and temperature in different months on the accuracy of rainfall-runoff simulation, scenarios 3 and 4 were used in GEP and ANFIS models, which had the highest accuracy. Tables 5–7 show the evaluation criteria of scenarios 3 and 4 in different months for the two models. The highest R^2-value of ANFIS and GEP models with scenario 4 is for March and June, respectively, which are almost similar (Table 5). The lowest R^2 also corresponds to August and September, and the results of both models are almost similar.

The RMSE results in Table 6 demonstrate no significant difference between GEP and ANFIS scenarios and models in October, and the rainfall-runoff prediction error was almost the same. In December, ANFIS-3 and GEP-4 provided the same results. In January, scenario 4, both models, and in February and March, the GEP-4 scenario had lower RMSE. Because in the winter and cold months of the year, sometimes the weather was snowy, the temperature parameter reduced the amount of model error.

When the temperature increased in April and May, there was no difference between the scenarios and models, and the results were almost identical. In June, in scenario 4, both models were slightly better than scenario 3. In the hot months of July, August, and September, models and scenarios did not differ much; their RMSE was almost the same. In general, from December to March, when the weather becomes cold, the temperature parameter increases the accuracy of the forecast. In spring and summer, except June, no difference was observed between scenarios 3 and 4. Table 7 shows that the models predicted rainfall-runoff less than the actual amount in March and April.

Table 4 | Comparison of R^2, MBE, and RMSE of withe box and black box models in the test phase

	HEC-HMS	TOPMODEL	GEP	ANFIS
R^2	0.3	0.25	0.76	0.82
MBE	−0.32	0.1	−0.06	−0.06
RMSE	1.8	1.78	1.52	1.32

Table 5 | Comparing the R^2-value from the GEP and ANFIS models for scenarios 3 and 4

	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep
GEP-3	0.59	0.68	0.82	0.64	0.55	0.94	0.98	0.81	0.99	0.78	0.43	0.22
GEP-4	0.54	0.27	0.84	0.68	0.71	1.00	0.92	0.80	0.98	0.80	0.65	0.18
ANFIS-3	0.73	0.35	0.81	0.67	0.57	0.65	0.94	0.66	0.78	0.54	0.76	0.33
ANFIS-4	0.38	0.54	0.75	0.94	0.93	0.98	0.92	0.71	0.81	0.57	0.32	0.05

Table 6 | Comparing the RMSE value from the GEP and ANFIS models for scenarios 3 and 4

	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep
GEP-3	0.27	0.31	0.52	1.06	0.67	1.62	0.50	0.35	0.49	0.42	1.41	0.35
GEP-4	0.30	0.26	0.38	0.49	0.51	1.08	0.53	0.35	0.44	0.40	1.39	0.33
ANFIS-3	0.27	0.26	0.38	0.84	0.70	2.16	0.51	0.37	0.49	0.41	1.40	0.33
ANFIS-4	0.31	0.56	0.43	0.35	0.68	1.62	0.52	0.35	0.44	0.41	1.39	0.36

Table 7 | Comparing the MBE value from the GEP and ANFIS models for scenarios 3 and 4

MBE	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep
GEP-3	0.19	0.28	0.31	0.20	−0.19	−0.74	−0.81	0.09	0.23	0.20	0.19	0.37
GEP-4	0.37	0.28	0.31	0.20	0.09	−0.46	−0.52	0.08	0.19	0.17	0.11	0.30
ANFIS-3	−0.19	0.29	−0.28	−0.06	−0.11	−1.11	−1.24	0.00	0.22	0.21	0.17	0.42
ANFIS-4	0.09	0.28	0.35	0.07	−0.19	−0.69	−0.83	0.10	0.14	0.11	0.15	0.19

4. CONCLUSION

This study used HEC-HMS and TOPMODEL as white box models and ANFIS and GEP as black box models to simulate rainfall-runoff during a 5-year statistical period. A combination of runoff, rainfall, and temperature data on a daily scale was selected as input parameters. Then, seven events were defined for rainfall-runoff simulation. Using rainfall and temperature inputs of the previous day and discharge in the previous two days' steps reduced the prediction error of both models. In this research, by investigating the rainfall and temperature factors in improving the simulations, it was found that the temperature factor as an effective parameter in the cold months reduced the amount of prediction error. Also, it was found that the accuracy of both black and white box models in the simulation of the hot months of the year is low, and the accuracy of the runoff prediction has not improved by changing the scenario. Comparing R^2, RMSE, and MBE demonstrated that black box models act as more effective tools in predictions, and using white box models as a suitable tool for the initial estimation was acceptable. Both black box models (ANFIS and GEP) could simulate daily rainfall-runoff. Although the accuracy of ANFIS was higher than the GEP model, using the GEP model is easier than the ANFIS model and can be used as a practical and competitive method with other methods in rainfall-runoff simulation. It should be noted that although black box models do not need equations with a physical basis, understanding the processes and the effect of the chosen factors in the input data on the output of the models is very important. In the absence of understanding, it cannot be Expected good results. If the black box models are run for a range that has not been trained, they usually use extrapolations that are not reliable. Also, some black box models, such as ANFIS, require much time to run, which makes using these models complicated and time-consuming.

Due to the importance of rainfall-runoff, it is suggested that researchers use other black box models, such as Support Vector Machine and Bayesian Network, in future studies. It is also recommended to use and examine other white box models, like HBV and HYMOD, to compare with black box models in future studies.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

Abdessamed, D. & Abderrazak, B. 2019 Coupling HEC-RAS and HEC-HMS in rainfall–runoff modeling and evaluating floodplain inundation maps in arid environments: case study of Ain Sefra city, Ksour Mountain. SW of Algeria. *Environmental Earth Sciences* **78** (19), 1–17.

Adamowski, J. 2013 Using support vector regression to predict direct runoff, base flow and total flow in a mountainous watershed with limited data in Uttaranchal, India. Annals of Warsaw University of Life Sciences-SGGW. *Land Reclamation* **45** (1), 71–83.

Adnan, R. M., Petroselli, A., Heddam, S., Santos, C. A. G. & Kisi, O. 2021 Short term rainfall-runoff modelling using several machine learning methods and a conceptual event-based model. *Stochastic Environmental Research and Risk Assessment* **35** (3), 597–616.

Andrews, R. & Boyne, G. A. 2010 Capacity, leadership, and organizational performance: testing the black box model of public management. *Public Administration Review* **70** (3), 443–454.

Anusree, K. & Varghese, K. O. 2016 Streamflow prediction of Karuvannur River Basin using ANFIS, ANN and MNLR models. *Procedia Technology* **24**, 101–108.

Azizi, H. & Nejatian, N. 2022 Evaluation of the climate change impact on the intensity and return period for drought indices of SPI and SPEI (study area: Varamin plain). *Water Supply* **22** (4), 4373–4386.

Bartoletti, N., Casagli, F., Marsili-Libelli, S., Nardi, A. & Palandri, L. 2018 Data-driven rainfall/runoff modelling based on a neuro-fuzzy inference system. *Environmental Modelling & Software* **106**, 35–47.

Campling, P., Gobin, A., Beven, K. & Feyen, J. 2002 Rainfall-runoff modelling of a humid tropical catchment: the TOPMODEL approach. *Hydrological Processes* **16** (2), 231–253.

Ewaid, S. H., Abed, S. A., Al-Ansari, N. & Salih, R. M. 2020 Development and evaluation of a water quality index for the Iraqi rivers. *Hydrology* **7** (3), 67.

Fernando, A. K., Shamseldin, A. Y. & Abrahart, R. J. 2012 Use of gene expression programming for multimodel combination of rainfall-runoff models. *Journal of Hydrologic Engineering* **17** (9), 975–985.

Halwatura, D. & Najim, M. M. M. 2013 Application of the HEC-HMS model for runoff simulation in a tropical catchment. *Environmental Modelling & Software* **46**, 155–162.

Hamdan, A. N. A., Almuktar, S. & Scholz, M. 2021 Rainfall-runoff modeling using the HEC-HMS model for the Al-Adhaim river catchment, northern Iraq. *Hydrology* **8** (2), 58.

Humphrey, G. B., Gibbs, M. S., Dandy, G. C. & Maier, H. R. 2016 A hybrid approach to monthly streamflow forecasting: integrating hydrological model outputs into a Bayesian artificial neural network. *Journal of Hydrology* **540**, 623–640.

Jaber, H. S., Mansor, S., Pradhan, B. & Ahmad, N. 2017 Rainfall–runoff modelling and water balance analysis for Al-Hindiyah barrage, Iraq using remote sensing and GIS. *Geocarto International* **32** (12), 1407–1420.

Kamranzad, B. & Samadi, M. 2013 Assessment of soft computing models to estimate wave heights in Anzali port. *Journal of Marine Engineering* **9** (17), 27–36.

Khan, M. S. & Coulibaly, P. 2006 Bayesian neural network for rainfall-runoff modeling. *Water Resources Research* **42** (7), W07409.

Li, J., Li, L., Dong, W. & Li, H. 2018 Purification effects of amended bioretention columns on phosphorus in urban rainfall runoff. *Water Science & Technology* **78** (9), 1937–1945.

Liu, W. & Lu, Z. 2022 Investigating the influences of concave depths on stormwater runoff and pollution retention of urban grasslands. *Water Science & Technology* **86** (9), 2441–2453.

Lu, J., Li, Z., Tang, R., Tang, B., Wu, H., Yang, F., Labed, J. & Zhou, G. 2013 Evaluating the SEBS-estimated evaporative fraction from MODIS data for a complex underlying surface. *Hydrological Processes* **27** (22), 3139–3149.

Miguélez, M., Puertas, J. & Rabuñal, J. R. 2009 Artificial neural networks in urban runoff forecast. In *International Work-Conference on Artificial Neural Networks*. Springer, pp. 1192–1199.

Misra, D., Oommen, T., Agarwal, A., Mishra, S. K. & Thompson, A. M. 2009 Application and analysis of support vector machine based simulation for runoff and sediment yield. *Biosystems Engineering* **103** (4), 527–535.

Młyński, D., Petroselli, A. & Walega, A. 2018 Flood frequency analysis by an event-based rainfall-runoff model in selected catchments of southern Poland. *Soil and Water Research* **13** (3), 170–176.

Molajou, A., Nourani, V., Afshar, A., Khosravi, M. & Brysiewicz, A. 2021 Optimal design and feature selection by genetic algorithm for emotional artificial neural network (EANN) in rainfall-runoff modeling. *Water Resources Management* **35** (8), 2369–2384.

Nourani, V. 2017 An emotional ANN (EANN) approach to modeling rainfall-runoff process. *Journal of Hydrology* **544**, 267–277.

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

Nourani, V., Davanlou Tajbakhsh, A., Molajou, A. & Gokcekus, H. 2019a Hybrid wavelet-M5 model tree for rainfall-runoff modeling. *Journal of Hydrologic Engineering* **24** (5), 4019012.

Nourani, V., Molajou, A., Najafi, H. & Danandeh Mehr, A. 2019b Emotional ANN (EANN): a new generation of neural networks for hydrological modeling in IoT. In *Artificial Intelligence in IoT*. Springer, pp. 45–61.

Shahmirnoori, A., Saadatpour, M. & Rasekh, A. 2022 Using mobile and fixed sensors for optimal monitoring of water distribution network under dynamic water quality simulations. *Sustainable Cities and Society* **82**, 103875.

Veintimilla-Reyes, J., Cisneros, F. & Vanegas, P. 2016 Artificial neural networks applied to flow prediction: a use case for the Tomebamba river. *Procedia Engineering* **162**, 153–161.

Wildhaber, M. L., Wikle, C. K., Moran, E. H., Anderson, C. J., Franz, K. J. & Dey, R. 2017 Hierarchical stochastic modelling of large river ecosystems and fish growth across spatio-temporal scales and climate models: the Missouri River endangered pallid sturgeon example. Geological Society, London, Special Publications, **408**(1), 119–145.

Yaseen, Z. M., Ghareb, M. I., Ebtehaj, I., Bonakdari, H., Siddique, R., Heddam, S., Yusif, A. A. & Deo, R. 2018 Rainfall pattern forecasting using novel hybrid intelligent model based ANFIS-FFA. *Water Resources Management* **32** (1), 105–122.

Yavari, F., Salehi Neyshabouri, S. A., Yazdi, J., Molajou, A. & Brysiewicz, A. 2022 A novel framework for urban flood damage assessment. *Water Resources Management* **36** (6), 1991–2011.

Youssef, A. M., Abu-Abdullah, M. M., AlFadail, E. A., Skilodimou, H. D. & Bathrellos, G. D. 2021 The devastating flood in the arid region a consequence of rainfall and dam failure: Case study, Al-Lith flood on 23th November, Kingdom of Saudi Arabia. Zeitschrift für Geomorphologie, 115–136.

First received 1 December 2022; accepted in revised form 8 January 2023. Available online 24 January 2023