



A multi-step strategy for enhancing the rainfall-runoff modeling: combination of lumped and artificial intelligence-based hydrological models

Babak Mohammadi¹ · Mir Jafar Sadegh Safari^{2,3} · Babak Vaheddoost⁴ · Mustafa Utku Yilmaz⁵

Received: 31 December 2024 / Accepted: 31 July 2025
© The Author(s), under exclusive licence to Springer Nature B.V. 2025

Abstract

Accurate rainfall-runoff (RR) modeling holds significant importance in environmental management, playing a central role in understanding the dynamics of water cycle. In this respect, the precision in the determination of RR is crucial for mitigating the adverse effects of both water scarcity and excessive runoff, ensuring the sustainable management of ecosystems and water resources. As a primary hydrological variable, runoff engages in direct interactions with other hydrological variables. Due to the complexity of the RR process, two primary approaches are commonly used in modeling, namely conceptual (lumped) models and artificial intelligence (AI) models. Conceptual approaches are based on hydrological processes and use a larger number of hydrological variables, yet they often exhibit lower performance compared to AI models. In contrast, AI models rely on fewer parameters and lack physical interpretability, but demonstrate high performance. This study merges the advantages of both lumped and AI techniques to develop an advanced RR model. Hence, the applicability of several lumped and AI-based models in estimating the streamflow rates with the help of basic meteorological variables is investigated. The lumped hydrological models, namely the Modello Idrologico SemiDistribuito in continuo (MISD), Identification of Unit Hydrographs and Component Flows from Rainfall, Evaporation, and Streamflow (IHACRES), and Génie Rural à 4 paramètres Journalier (GR4J), are employed in conjunction with AI algorithms as Radial Basis Function (RBF) neural networks, Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multilayer Perceptron (MLP). An ensemble of conceptual models (MISD, IHACRES, and GR4J) and three AI models (MLP, RBF, and ANFIS) with various lag times are considered as effective variables, where Support Vector Machine (SVM) was utilized as a feature selection method with five different kernels in determining the best inputs. Afterward, the SVM-ANFIS model, as the best model, is hybridized with Ant Colony Optimization (ACO) to develop the SVM-ANFIS-ACO model. It is found that the coupling of lumped and AI methodologies considerably enhanced the accuracy of the RR models; and SVM-ANFIS-ACO outperformed other models in streamflow computation.

Keywords Artificial intelligence · Conceptual models · Optimization · Rainfall-runoff

Extended author information available on the last page of the article

1 Introduction

Rainfall-runoff (RR) modeling is an essential hydrological problem in flood control, agricultural planning, and water resources management. It provides insights for estimating both precipitation and streamflow in a river basin to mitigate adverse environmental effects and promote the effective use of water resources. Various approaches such as empirical, physical, and conceptual methodologies can be applied to hydrological models for better understanding of the hydrological process. It allows stakeholders to conduct certain policies for water allocation, protect ecosystems, and flood control (Chow, 1965; Singh, 2018; Kim et al., 2021; Mohammadi et al., 2022; Tao et al., 2023; Yilmaz, 2025).

For the sake of understanding the water cycle in a catchment, hydrological modeling can be performed. Among various hydrological modeling approaches, conceptual models present hydrological cycle components through some simplifications (Senatore et al., 2020; Tian et al., 2020) which provide a generic frame for the perception of hydrological characteristics in the studied area. Although conceptual hydrological methods are valuable for better understanding of the water cycle process in a region, their computational ability is not as high as artificial intelligence (AI) techniques. AI-based hydrological models not only provide better computational accuracy, but also are more effective in selection of variables affecting the hydrological process (e.g., topographic and atmospheric conditions), which are incorporated into the models (e.g., Aytek et al., 2008; Wang et al., 2009; Rasouli et al., 2012; Sanikhani et al., 2025).

Previous research has frequently employed conceptual hydrological models for streamflow estimation and water allocation worldwide (e.g., Brocca et al., 2011; Garavaglia et al., 2017; Masseroni et al., 2017; Mohammadi et al., 2023). Likewise, AI techniques are also widely used for modeling hydrological problems (e.g., Hung et al., 2009; Abbott & Marohasy, 2012; Fahad et al., 2023; Vaheddoost et al., 2023a; Yaseen et al., 2017; Kim et al., 2020; Shekar et al., 2025).

But, previous studies have identified several limitations when applying either standalone hydrological or AI models. For instance, conceptual models demonstrate reduced performance when dealing with extreme hydrological events (e.g., peak flows) or rapidly changing environmental conditions, as their underlying assumptions and parameter sets are usually calibrated based on historical norms (Tan et al., 2020; Peredo et al., 2022; Van Kempen et al., 2021; Stephens et al., 2019). So far, studies have shown that these models often underestimate peak flows during extreme rainfall events and struggle to capture the non-linear dynamics of rapid hydrological responses. Moreover, the increasing frequency of extreme weather events due to climate change further challenges these models (Duethmann et al., 2020). By utilizing standalone AI models, though computationally powerful, these studies lack physical interpretability and fail to incorporate domain knowledge of hydrological processes. Often behaving as ‘black boxes,’ the outcomes of these models are difficult to understand (Höge et al., 2022; Herath et al., 2021). Additionally, AI models’ performance is heavily dependent on the quality and quantity of training data, requiring extensive historical records that may not be available for every study area.

To be considered in this respect, coupling multiple approaches for modeling hydrological problems may enhance the computational ability of the developed model to gain the advantages of base methods (Fernando et al., 2012; Young & Liu, 2015; Hosseini and Mahjouri, 2016; Yilmaz et al., 2023; Fathi et al., 2025; Boodoo et al., 2025). It enables the incorpora-

tion of more variables into the model for better hydrological process modeling, where a single model fails to consider those variables from the beginning. For instance, Xiong and O'Connor (2002) coupled the lumped Soil Moisture Accounting and Routing (SMAR) and AI techniques. Zhang and Govindaraju (2003) and, Hosseini and Mahjouri (2016) implemented the Geomorphological Instantaneous Unit Hydrograph (GIUH) coupled with a support vector machine (SVM) and artificial neural network (ANN) for streamflow predictions. Shamseldin et al. (2007) merged the outcomes of five hydrological models for streamflow estimation. Young and Liu (2015) developed RR models by combining conceptual and AI algorithms.

The Coruh River, as a fast and relatively deep river, is one of the main water resources in northeastern Türkiye. These characteristics also make the river strategically significant for hydroelectric energy production and ecosystem conservation. Due to its important role in the region's water resources, several studies have been conducted to date to better understand its hydrological behavior. To this end, Yilmaz et al. (2023) developed a framework that combines statistical regionalization methods with an ANN to improve daily streamflow estimation, using the Coruh River Basin as a case study. This approach enhanced estimation accuracy and provided insights into the river's hydrological dynamics. Other studies have also investigated the efficiency of AI algorithms for streamflow estimation in the basin, such as Mehr et al. (2015) and Gunes et al. (2021), highlighting their potential for hydrological modeling.

Although the best algorithms require reliable and reasonable input(s), it is still not easy to take in the most relevant variables. As explained above, most studies conducted so far have examined the riverine streamflow process using process-based or black-box models. In the current study however, the performance of three lumped hydrologic models namely Modello Idrologico Semi-Distribuito in continuo (MISD), Identification of unit Hydrographs and Component flows from Rainfall, Evaporation, and Streamflow (IHACRES), and Génie Rural à 4 paramètres Journalier (GR4J) are compared with three AI models namely Radial Basis Function (RBF) neural network, Adaptive Neuro-Fuzzy Inference System (ANFIS) and Multilayer Perceptron (MLP) for streamflow estimation in semi-arid basins. To address the limitations associated with both process-based and AI approaches, models are augmented by integrating the insights from physics-based principles with the capabilities of AI techniques. This approach leverages a fusion of physical understanding and AI performance, offering a more robust and comprehensive solution. The study aims to (i) develop a novel framework for RR modeling that operates effectively with minimal input data, (ii) develop and compare the outcomes of the listed lumped and AI-based models, and (iii) perform hybridization and develop a strategy to promote the accuracy of the selected models that uses the advantages of both conceptual hydrological models and data-driven techniques.

2 Study area and data

The study area is located in northeastern Türkiye in the Coruh River Basin (see Fig. 1). This basin spans 39°52' ~ 41°32' N latitudes and 39°40' ~ 42°35' E longitudes. It's a heavily forested mountainous region characterized by steep terrain. The climate in the south is predominantly continental, while the north experiences a mild and rainy climate due to the

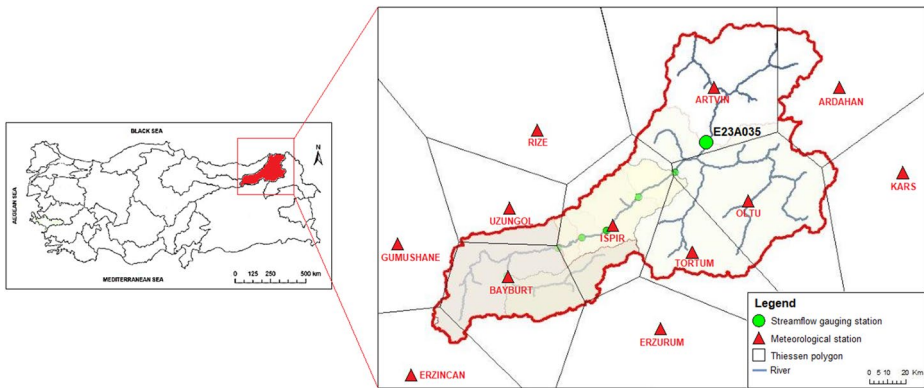


Fig. 1 Locations of the streamflow gauging station and meteorological stations in the Coruh River Basin

proximity to the Black Sea. Summers are hot and dry, winters are cool and rainy, and most of the precipitation occurs in spring. Winter temperatures often hover around or below 0 °C, leading to snow accumulation that gradually melts with rising spring temperatures.

The Coruh River emerges from the Mescit Mountains at an altitude of 3000 m above sea level (m ASL) within the borders of Erzurum Province and is one of the fastest and deepest rivers in the world. This river passes through Bayburt, Erzurum, and Artvin and flows into the Black Sea in Georgia. The river has a total length of 431 km, of which about 20 km are in Georgian territory. The river discharge varies seasonally, while water levels rise in spring and falls in summer, increasing the risk of flooding during extreme weather conditions. Despite the presence of dams and relatively high precipitation levels in the region, the persistent vulnerability to drought underscores the necessity for more proactive and integrated water resource management strategies (GDWM, 2020; Seleğ & Aksu, 2020; Vaheddost et al., 2023b).

As illustrated in Fig. 1, the streamflow gauging station numbered E23A035 is selected for hydrological simulation of the Coruh River Basin. This station is strategically situated close to the outlet of the basin along the Coruh River's mainstream, precisely at the 48th-kilometer point of the Artvin-Yusufeli highway. The drainage area of this station spans 15,687.1 km², and its elevation is 435 m ASL. To conduct analysis, we collected observed streamflow records from the General Directorate of State Hydraulic Works of Türkiye (abbreviated as DSI in Turkish) for 18 water years, totaling 6,574 days (October 1, 1992, to September 30, 2010). It's worth noting that no data reconstruction is applied due to the completeness of the time series. In the selected sub-basin, which corresponds to the drainage area of the streamflow gauging station, the temperature and precipitation values were calculated using Thiessen polygon method. Then, the selected data from seven meteorological stations whose Thiessen polygons overlapped with the selected sub-basin (as shown in Fig. 1) were used for analysis. The weights assigned to these meteorological stations were determined using Thiessen polygons; as Oltu (0.30), Ispir (0.24), Bayburt (0.21), Tortum (0.14), Artvin (0.07), Erzincan (0.03), and Uzungöl (0.01). To ensure consistency in the analysis, the acquired daily-scale precipitation and temperature from seven stations during the same time frame as the streamflow data were used. These meteorological data were acquired from the Turkish State Meteorological Service (abbreviated as MGM in Turkish). Table 1 provides

Table 1 Properties of the streamflow gauging station and meteorological stations

Station type	Station number	Station name	Location		Elevation (m)
			Latitude (°)	Longitude (°)	
Stream-flow gauging station	E23A035	Inanli	40.8888	41.7161	435
Meteorological stations	17045	Artvin	41.1752	41.8187	613
	17089	Bayburt	40.2547	40.2207	1584
	17094	Erzincan	39.7523	39.4868	1216
	17569	Uzungol	40.6193	40.4435	1114
	17666	Ispir	40.4868	40.9997	1223
	17668	Oltu	40.5497	41.9951	1312
	17688	Tortum	40.3013	41.5409	1576

Table 2 Statistical characteristics of the drainage area of the streamflow gauging station

Variable	Minimum	Maximum	Mean	S_x	C_v	C_s
Q (m ³ /s)	14.200	1145.000	122.083	133.567	1.094	2.322
P (mm)	0.000	31.397	1.330	2.736	2.056	3.558
T_{min} (°C)	-22.769	20.840	3.952	8.400	2.126	-0.391
T_{max} (°C)	-11.245	39.549	16.014	10.828	0.676	-0.117
T_{mean} (°C)	-17.113	29.214	9.579	9.515	0.993	-0.226

Q represents daily streamflow, P refers to daily total precipitation, and T_{mean} , T_{min} , and T_{max} are daily mean, minimum, and maximum temperatures, respectively. S_x is the standard deviation, C_v refers to the coefficient of variation, and C_s is the coefficient of skewness

information about the elevation and location of the stations, while Table 2 presents statistics related to the streamflow, precipitation, and temperature values for the drainage area of the streamflow gauging station.

3 Methods

3.1 Lumped hydrological models

3.1.1 Modello Idrologico Semi Distribuito in continuo (MISD)

The MISD is a conceptual hydrological model introduced by Brocca et al. (2011) for flood prediction in the Upper Tiber River, Italy. The primary emphasis of the MISD lies in the soil moisture module, which is demonstrated to alter the storage capacity and RR modeling accordingly. The MISD approach is implemented as a lumped hydrological model in the current research which used the daily temperature and rainfall information as inputs at the catchment scale for daily runoff simulation. Evaporation and transpiration processes lead to exit the water from the first layer, and this process can be computed via a linear function of the saturated soil and potential evaporation, while a non-linear relationship is employed for calculating the infiltration process of water movement from the soil surface to the root zone (Masseroni et al., 2017). In the MISD model, runoff production is caused by three

distinct components: the constituents of surface excess saturation, soil second layer, and the components related to subsurface runoff. The initial two are collected through the employment of the instantaneous geomorphological unit hydrograph (IGUH) and subsequently conveyed to the outflow, whereas the groundwater runoff is channeled to the outlet through the linear reservoir method. The Curve Number (CN) approach was used to evaluate losses, and the linear reservoirs and the IGUH methods are employed for tracking the precipitation in regions where runoff directly can be connected to the main waterway, and the main waterway routing can be calculated via linear broadcast method (Brocca et al., 2011). To summarize the setup, Table 3 details the parameters and corresponding optimal values in establishing the MISD model.

3.1.2 Identification of unit hydrographs and component flows from rainfall, evaporation, and streamflow data (IHACRES)

The IHACRES model represents a conceptual hydrological framework to conduct RR modeling using the minimum possible parameters in the catchment scale (Jakeman & Hornberger, 1993). Temperature and precipitation datasets are utilized by the IHACRES model as the input data for simulating runoff, which can also be employed as the measured runoff data for both the model calibration process and the examination of its performance. This model's foundation comprises two reduction models that operate non-linearly, in addition to a linear hydrograph model. The model for non-linear reduction transforms precipitation into rainfall that serves as effective rainfall by incorporating the ratio of evaporation and infiltration (Croke & Jakeman, 2004). The calculation of the basin saturation and basin moisture indices is carried out at each time step to calculate effective rainfall. Consequently, the parameters given in Table 4 are utilized as the best option for the application of the IHACRES model in the analysis.

3.1.3 Génie Rural à 4 paramètres Journalier (GR4J)

The GR4J is a conceptual RR model that has been employed for hydrological modeling tasks with acceptable accuracy in various regions (Sezen et al., 2019; Boumenni et al., 2017). Perrin et al. (2003) introduced the GR4J hydrological model as a lumped RR model specifically for daily-based simulations. The GR4J uses two phases as storage elements, which are the routing storage and production phase (Perrin et al., 2001). The routing stores

Table 3 The parameters of the MISD model used in this study

Parameter descriptions	Optimal value
Hydraulic conductivity for 1st layer	0.875
Hydraulic conductivity for 2nd layer	65
The exponent of drainage for 1st layer	10
Parameter of potential evapotranspiration	0.865
An exponent of drainage for 2nd layer	11.803
Snow module parameter degree-day	0.282
Fixed water capacity 1st layer	150
Total water capacity of 2nd layer	884.782
Initial conditions (fraction of W_{max})	0.477
Exponent runoff	15
Coefficient lag-time relationship	0.5

Table 4 Optimized parameters of the IHACRES model in this study

Parameter descriptions	Optimal value
Flow threshold	121.23
Temperature to potential evapotranspiration conversion factor	0.88
Plant stress threshold factor	0.91
The time constant governing the rate of recession of quick-flow	19.21
The time constant governing the rate of recession of slow-flow	92.39
The proportion of slow-flow to total flow	0.89

Table 5 The parameters of the GR4J model that were calibrated for this study

Parameter descriptions	Optimal value
One day ahead of the max capacity of the routing store	31.29
The time base of unit hydrograph UH1	0.87
Groundwater exchange coefficient	2.54
Maximum capacity of the production store	90.11

process can be conducted by two-unit hydrographs (UH) in the GR4J model; (1) the first UH generates (10% of) effective runoff, and (2) the other 90% of runoff can be generated by the unit routing store as the second UH. GR4J considers four physically-based parameters during its calibration phases, representing storage capacity (mm), groundwater contribution (mm), routing storage capacity (mm), and routing mechanism of the daily UH (Kumalioglu & Fistikoglu, 2019). Eventually, the parameters set given in Table 5 are used in the application of the GR4J model.

3.2 Artificial intelligence models

3.2.1 Multilayer perceptron (MLP)

The concept of the perceptron as an artificial neuron was first introduced by McCulloch and Pitts (1943). The connection between input and output vectors in an MLP network is non-linear. This is achieved by the connections between neurons in one layer and those in the adjacent layer (either the preceding or succeeding layer). The output of every neuron is subject to the multiplication with weight factors and subsequently presented as input to a non-linear excitation function. During the training stage, the perceptron is exposed to training data, and the weights in the network are adjusted to minimize the error between the desired output and the model's actual output, either until the error reaches a predefined goal or until a specified number of training iterations is completed. Similar to all modeling procedures, new inputs (absent during the training phase) are employed to validate the model, as is the case with all modeling processes.

3.2.2 Radial basis function network (RBF)

The RBF is extensively applied across various fields and typically yields satisfactory results. This is due to its capacity to effectively approximate functions, particularly where there exist non-linear relationships among the variables in the natural system (Kavaklioglu et al., 2018). The RBF architecture includes three layers: input, hidden, and output. For each neuron in the layers, the transfer function is determined through the definition of the input-target samples. Within the RBF modeling procedure, the output is obtained by a linear combination of the inputs. The RBF is reliable for variety of purposes in hydrological science such as time series modeling (Heddam, 2016).

3.2.3 Adaptive Neuro-Fuzzy Inference System (ANFIS)

The ANFIS method introduced by Jang et al. (1997) utilizes neural network and fuzzy logic features to find the hidden relationship among the variables involved in the problem. The applicability of ANFIS in hydrological science was demonstrated in several studies such as Kisi et al. (2017) and Aghelpour et al. (2020). Through the modeling procedure of ANFIS, the connections among inputs and output variables are detected by fuzzy rules to develop the input structure of the system through the implementation of neural networks. In this study, for the application of the Fuzzy Inference System (FIS) and fuzzy clustering approach, respectively the Sugeno and Fuzzy Cluster Means (FCM) techniques are utilized (Takagi & Sugeno, 1985; Bezdek et al., 1984).

3.2.4 Support Vector Machine (SVM)

As a supervised learning approach, SVM is widely implemented for regression and classification problems. Within the SVM algorithms, a set of n -dimensional vectors is classified into separate groups which is a technique commonly used in several investigations (Cortes & Vapnik, 1995). In order to classify non-linearly separable vectors, a kernel function is used to transform the multidimensional vectors into a higher-dimensional space (Mehr et al., 2020). In the current study, five kernel functions are considered in the modeling as: dot, radial, polynomial, Epanechnikov, and ANOVA kernels to remove less effective inputs by examining their normalized weight.

3.2.5 Ant Colony Optimization (ACO)

The ACO is a metaheuristic algorithm introduced by Dorigo (1992) for solving optimization problems (Dorigo et al., 1999). It is inspired by ants' behavior in searching for the optimal path to reach food from their initial colony location (Dorigo, 1992). Real ants use pheromones to guide each other to resources as they explore their area. Therefore, a higher pheromone concentration on a road means that more ants use that path, and the path (the quickest route to food) is right (Khosravi et al., 2018). Similar to other optimization algorithms, this method possesses optimizer operators. In this study, the ACO algorithm is implemented by the maximum number of iterations=1000, the population size=40, the intensification factor=0.5, and the deviation-distance ratio=1.

3.3 Improvement of models via hybridization

The current study proposed stepwise hydrological modeling via hydrological conceptual- and AI-based algorithms (Fig. 2). This framework uses combination of the meteorological dataset, and the output of previous steps as the inputs of the next steps to use the benefits of both the meteorological dataset in time t and the persistence in the simulated runoff time series at $t-1$, $t-2$, $t-3$, $t-5$, $t-10$, and $t-30$.

Three main steps are followed in this respect: (1) RR modeling through conceptual hydrological modeling approaches (namely GR4J, IHACRES, and MISD) and three AI-based models (namely ANFIS, MLP, and RBF). To achieve a fair comparison between the selected models, the same variables are used as inputs in the models (given in step 1). In this respect, GR4J, IHACRES, and MISD use temperature and precipitation as input datasets. However, their structure calculates evapotranspiration for RR modeling; therefore, the precipitation, temperature, and evapotranspiration are simultaneously used as inputs to the

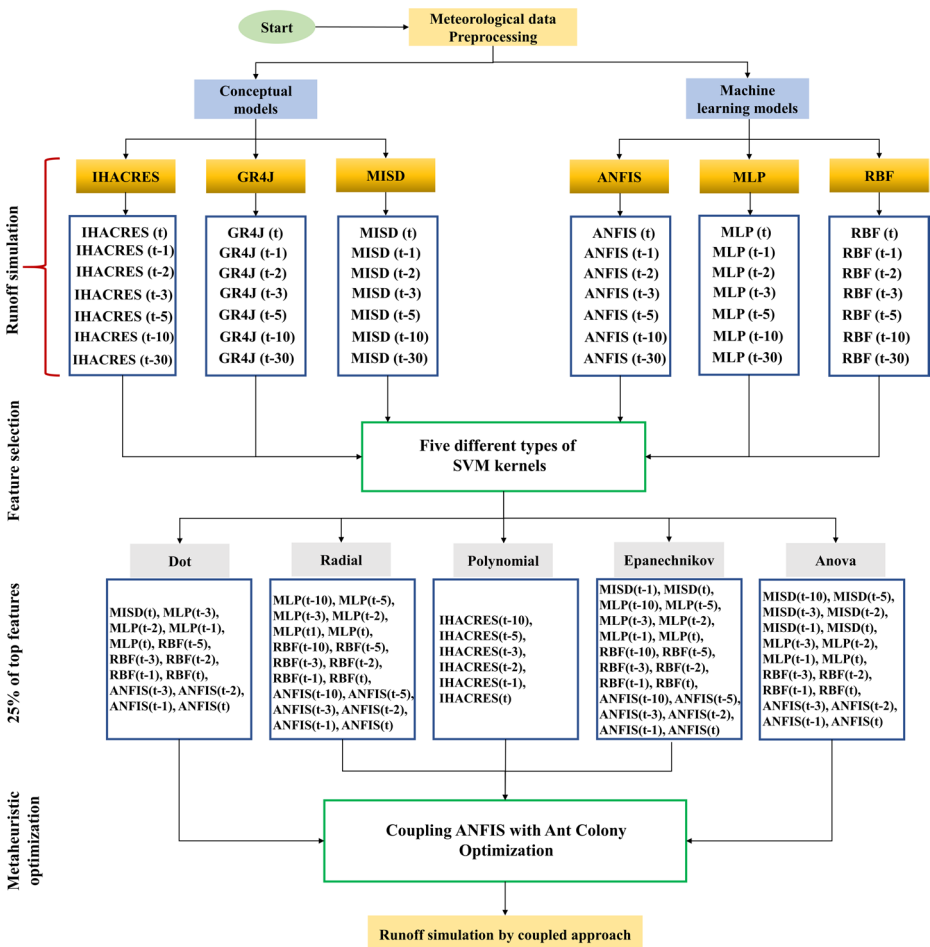


Fig. 2 Flowchart of the proposed framework in the current study

ANFIS, MLP, and RBF models. In step (2) the outputs of six applied models in step (1) are considered as potential variables in step (2). For this aim, an ensemble simulated runoff of six applied models in $t-1$, $t-2$, $t-3$, $t-5$, $t-10$, and $t-30$ lag time series is generated and fed to the SVM algorithm as a feature selection method. Five different SVM kernels, namely Dot, Radial, Polynomial, Epanechnikov, and ANOVA are applied to the feature selection procedure. Therefore, simulated runoff in $t-1$, $t-2$, $t-3$, $t-5$, $t-10$, and $t-30$ lags are analyzed by five different SVM kernels to reach the best time series based on each kernel's algorithms. In total, 25 of the top values at each output in step (2) are selected as outcomes to consider the most effective variables based on each kernel behavior. In step (3), a meta-heuristic technique which is inspired by ant colony behavior is considered for optimizing the ability of the ANFIS in hydrological simulation. The proposed framework (SVM-ANFIS-ACO) trains the ANFIS model via an optimal method, which leads to better training and therefore less error for the ANFIS model. In addition, step (3) plays a combination role for all outputs of the previous steps, and this step helps runoff modeling via the advantages of conceptual and AI techniques. To address potential overfitting issues, a strategy including optimal data splitting (80% training data), SVM-based feature selection, and early stopping during the training process is implemented. These measures ensured our models reliability and relatively maintained good generalization capability, as evidenced by the consistent performance across training and test sets.

In this respect, the proposed framework presents a novel multi-step strategy that distinctively integrates conceptual lumped models with advanced AI techniques to overcome limitations inherent in each approach used independently. Unlike traditional frameworks that typically employ either conceptual or AI-based models in isolation, this methodology uniquely combines both approaches through a sequential three-step architecture: first establishing baseline predictions with both model types, then enhancing variable selection through multiple SVM kernel functions applied to model outputs at various time lags, and finally implementing metaheuristic optimization via ACO to refine the ANFIS model performance. This integrated approach effectively bridges the gap between physically-based understanding and computational efficiency, capturing both the underlying hydrological processes through conceptual models and the complex non-linear relationships through AI techniques.

3.4 Comparison metrics

The present study utilized various statistical metrics to examine the efficacy of the implemented models for simulating RR. Hence, mean absolute error (MAE), Nash-Sutcliffe Efficiency (NSE) (Nash & Sutcliffe, 1970), and Kling-Gupta Efficiency (KGE) (Gupta et al., 2009) are utilized as

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

$$MAE = \frac{1}{N} \sum_{i=1}^N |(R_{i(measured)} - R_{i(predicted)})| \quad (2)$$

$$NSE = \left[1 - \frac{\sum_{i=1}^N (R_{i(measured)} - R_{i(predicted)})^2}{\sum_{i=1}^N (R_{i(measured)} - \bar{R}_{(measured)})^2} \right] \quad (3)$$

where N is the number of datasets, $R_{i(measured)}$ and $R_{i(predicted)}$ are the measured and predicted runoff values, respectively; and $\bar{R}_{(measured)}$ and $\bar{R}_{(predicted)}$ respectively are the mean measured and mean predicted runoff values; r is the linear correlation coefficient between predicted and simulated runoff values, α refers to a ratio for the standard deviation of the predicted runoff values to the standard deviation of the measured runoff values, and β refers to a ratio for the predicted runoff value means to the mean of the measured runoff.

4 Results

4.1 Comparison of conceptual and standalone AI models

As detailed previously, three conceptual models (i.e., GR4J, IHACRES, and MISD) and three AI models (i.e., ANFIS, MLP, and RBF) are used in step (1). The performance of six different models in predicting runoff are listed in Table 6. To assess the performance of these models, four different evaluation metrics namely: Mean Absolute Error (*MAE*), Nash-Sutcliffe Efficiency (*NSE*), and Kling-Gupta Efficiency (*KGE*) are used. Based on the conducted analysis, the MLP, RBF, and ANFIS showed reasonable performance based on the evaluation metrics. In this respect, low *MAE* indicated low bias in estimation, while high *NSE* and *KGE* values indicated favorable similarity between observation and estimation. Yet, the results of IHACRES and GR4J did not show favorable performance on the training dataset, considering high *MAE*, and low rate of *NSE* and *KGE*. On the other hand, based on the testing dataset, the so-called MLP, RBF, and ANFIS showed better performance. But, as expected the performance in the testing dataset is relatively lower than the training set for all of the models. MISD model represented the best result among hydrological-based models by *MAE*=66.44 (m³/s), *NSE*=0.46, and *KGE*=0.61 for the testing phase; while the MLP model represented the best result among AI-based models by *MAE*=36.04 (m³/s), *NSE*=0.75 and *KGE*=0.73 for the testing phase. In general, AI-based models performed more accurate than hydrological-based models in both the training and testing phases. As expected, the RR simulation in the training phase yielded more accurate results compared to those of testing phases for all applied models. Since, AI-based models could not be used

Table 6 Performance of three hydrological-based models (GR4J, IHACRES, and MISD) and three AI-based models (ANFIS, MLP, and RBF) for RR modeling in step 1

Dataset	Performance metric	Hydrological-based model			Intelligence-based model		
		MISD	IHACRES	GR4J	MLP	RBF	ANFIS
Training	<i>MAE</i>	37.29	52.78	58.78	31.92	30.41	31.28
	<i>NSE</i>	0.72	0.51	0.46	0.79	0.80	0.79
	<i>KGE</i>	0.86	0.61	0.52	0.85	0.87	0.86
Testing	<i>MAE</i>	66.44	72.63	76.7	36.04	38.1	36.98
	<i>NSE</i>	0.46	0.39	0.29	0.75	0.73	0.73
	<i>KGE</i>	0.61	0.40	0.24	0.73	0.72	0.71

in physical interpretation of the runoff process, it is evident that the performance of the AI-based models is dependent on the quality and quantity of the dataset and parameter tuning.

Figure 3 shows three hydrographs and scatter plots related to the measured and simulated runoff from three different hydrological-based models (GR4J, IHACRES, and MISD). According to the time series plots, it is obvious that the GR4J, IHACRES, and MISD models captured different patterns in the measured data. More precisely, the MISD showed a favorable temporal correlation with the measured data, indicating its ability to capture the RR patterns. The IHACRES, however, showed a moderate level of agreement, particularly during the peak flow periods. The GR4J model performed better in capturing the disparity, despite the limitations of displaying hydrological processes, especially on the base- and peak flows. In the next step, the scatter plots are used to display the distribution of data points and their correlation against the perfect fit line (i.e. dashed red line). The MISD model performed better than the remaining models and showed a high level of agreement with the observations.

In the following, Fig. 4 depicts the scatter plots and corresponding time series plots of AI-based models in replicating the observed streamflow data. In the time series plots, deviations or agreements between the measured and simulated runoff can be observed. Analyzing the time series plots, it becomes apparent that each AI-based model exhibits varying degrees of agreement with the measured runoff. All three AI-based models closely matched the measured runoff for the majority of the time period, indicating a strong performance in reproducing the observed runoff patterns. However, all three AI-based models showed some limitations in peak flow simulations. Models based on AI demonstrated a high level of agreement with the measured runoff, as indicated by a tight cluster of points close to the reference line (compared with IHACRES, GR4J, and MISD models).

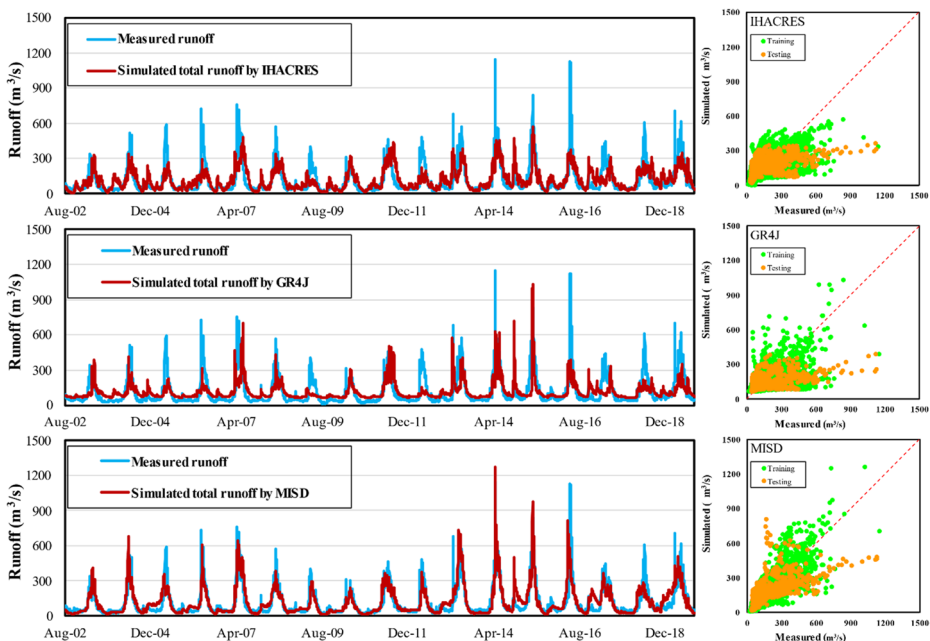


Fig. 3 Time Series and scatter plots of GR4J, IHACRES, and MISD models

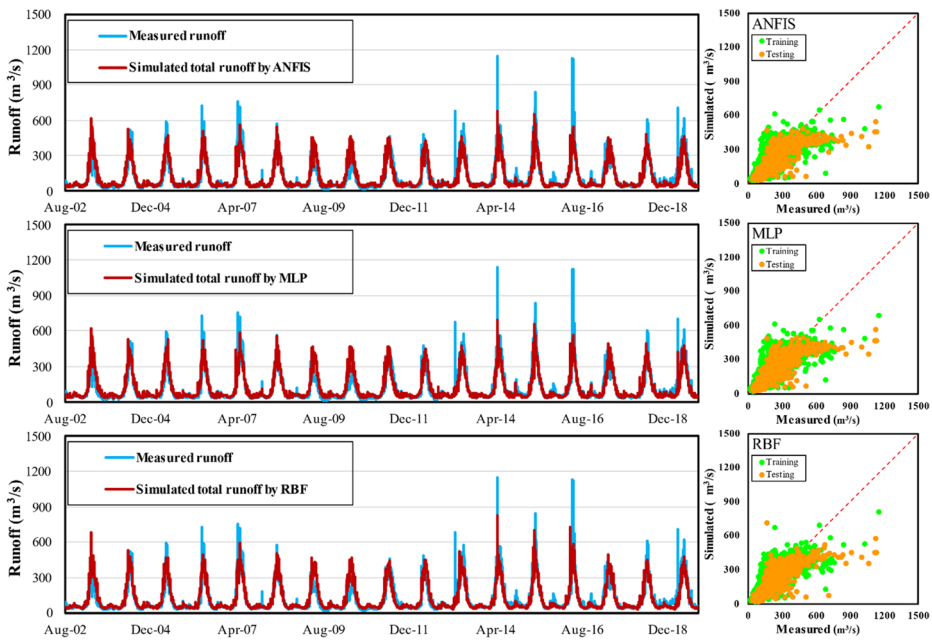


Fig. 4 Hydrographs and scatter plots of ANFIS, MLP, and RBF models

4.2 Feature selection via SVM

During step (2) of the modeling task, the normalized weights of each SVM kernel contributed to the input selection for RR modeling. Figure 5 presents the results of the SVM in feature selection, which is employed to identify the most effective inputs for the coupled models. In this analysis, five different SVM kernels are utilized: Dot, Radial, Polynomial, Epanechnikov, and ANOVA. The inputs for SVM consisted of the outputs from the hydrological models and the AI models. Additionally, lag times of the simulated runoff from these models are incorporated as inputs, considering t , $t-1$, $t-2$, $t-3$, $t-5$, $t-10$, and $t-30$ lag time series (where t represents the original time series, $t-1$ represents a one-day lag time, and $t-30$ represents a 30-day lag time). For each SVM kernel, a specific set of variables is suggested to be effective for runoff prediction in the coupled models. In this respect, the allocation of the variables takes place based on their respective weights assigned by the SVM kernels. Hence, weights ≤ 0.75 are considered as the threshold for the variable selection (Table 7).

These kernel functions significantly impact model performance by determining which combination of inputs (lagged outputs from previous models) are selected for the final ANFIS model. The subsequent ACO optimization enhances the ANFIS performance based on the selected features, creating five distinct model configurations with different input structures, with identically optimized parameters.

Afterward, weights equal or above 0.75 are selected in step (3), considering that each kernel function has advantages and disadvantages (Zhong-Da et al., 2014). In step (2), the inputs selected by the SVM-ANFIS-ACO model are determined. Considering the data provided in Table 7, the first column depicts the type of kernel function used in determination

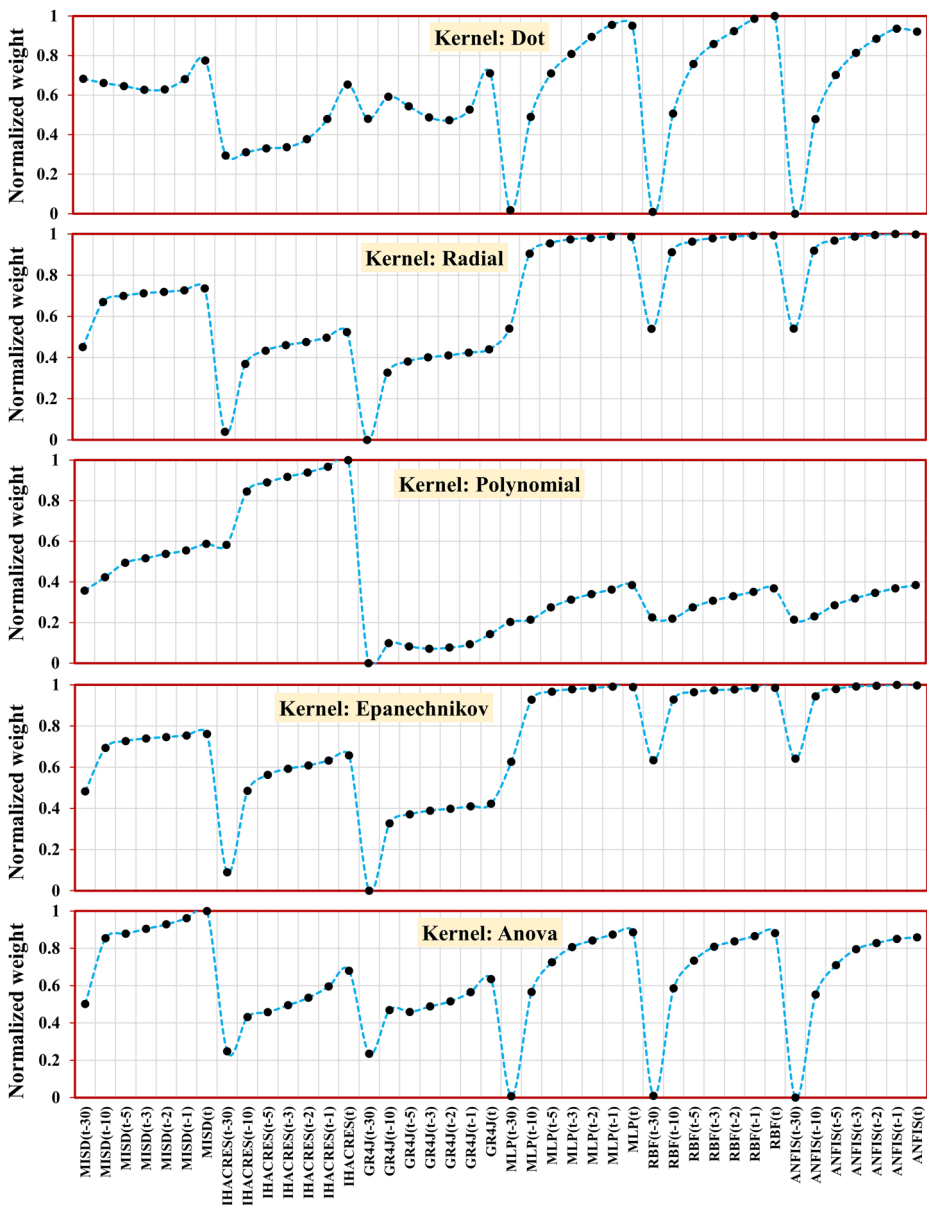


Fig. 5 Determination of the most effective features of the hydrological-based and AI-based models using SVM

of decision boundary of the models. The second column in Table 7, shows the variables used in estimations, and various combinations of time-series data for IHACRES, MISD, GR4J, MLP, RBF, and ANFIS models. The specific lag times used are defined, such as $t-n$ which are measured relative to the current time, t . The third column, on the other hand, indicates the name of the models together with the integrated kernels and inputs. For example, the

Table 7 Effective variables based on the different SVM kernels

Kernel	Selected input(s)	Model
Dot	MISD(t), MLP(t-3), MLP(t-2), MLP(t-1), MLP(t), RBF(t-5), RBF(t-3), RBF(t-2), RBF(t-1), RBF(t), ANFIS(t-3), ANFIS(t-2), ANFIS(t-1), ANFIS(t)	SVM-AN-FIS-ACO1
Radial	MLP(t-10), MLP(t-5), MLP(t-3), MLP(t-2), MLP(t-1), MLP(t), RBF(t-10), RBF(t-5), RBF(t-3), RBF(t-2), RBF(t-1), RBF(t), ANFIS(t-10), ANFIS(t-5), ANFIS(t-3), ANFIS(t-2), ANFIS(t-1), ANFIS(t)	SVM-AN-FIS-ACO2
Polynomial	IHACRES(t-10), IHACRES(t-5), IHACRES(t-3), IHACRES(t-2), IHACRES(t-1), IHACRES(t)	SVM-AN-FIS-ACO3
Epanechnikov	MISD(t-1), MISD(t), MLP(t-10), MLP(t-5), MLP(t-3), MLP(t-2), MLP(t-1), MLP(t), RBF(t-10), RBF(t-5), RBF(t-3), RBF(t-2), RBF(t-1), RBF(t), ANFIS(t-10), ANFIS(t-5), ANFIS(t-3), ANFIS(t-2), ANFIS(t-1), ANFIS(t)	SVM-AN-FIS-ACO4
Anova	MISD(t-10), MISD(t-5), MISD(t-3), MISD(t-2), MISD(t-1), MISD(t), MLP(t-3), MLP(t-2), MLP(t-1), MLP(t), RBF(t-3), RBF(t-2), RBF(t-1), RBF(t), ANFIS(t-3), ANFIS(t-2), ANFIS(t-1), ANFIS(t)	SVM-AN-FIS-ACO5

first row uses the Dot kernel and selected inputs of MISD(t), MLP(t-3), MLP(t-2), MLP(t-1), MLP(t), RBF(t-5), RBF(t-3), RBF(t-2), RBF(t-1), RBF(t), ANFIS(t-3), ANFIS(t-2), ANFIS(t-1), ANFIS(t), resulting in the model named as SVM-ANFIS-ACO1.

As explained above, Fig. 5 depicts the selected variables based on their assigned weights. According to the results and the weights derived from the SVM kernels, the allocation provided by SVM strengthened the predictions and improved both the understanding of crucial input variables and the overall accuracy of the coupled models.

4.3 SVM-ANFIS-ACO model

Table 8, depicts the results acquired for SVM-ANFIS-ACO models. According to the results, the SVM-ANFIS-ACO1, SVM-ANFIS-ACO2, SVM-ANFIS-ACO4, and SVM-ANFIS-ACO5 exhibit favorable performance metrics. The favorable results showed low MAE and high NSE and KGE, suggesting accuracy in prediction. Among the models evaluated on the testing dataset, SVM-ANFIS-ACO2 performed the best, achieving $MAE=35.17$ (m^3/s), $NSE=0.78$, and $KGE=0.72$. The worst-performing model is SVM-ANFIS-ACO3 with $MAE=75.38$ (m^3/s), $NSE=0.33$ and $KGE=0.41$.

Table 8 Performance of simulated runoff by SVM-ANFIS-ACO models via the most effective input datasets

Dataset	Per- for- mance metric	SVM-ANFIS-ACO1	SVM-ANFIS-ACO2	SVM-ANFIS- ACO3	SVM- AN- FIS- ACO4	SVM- AN- FIS- ACO5
Training	MAE	21.54	26.94	48.57	21.37	21.73
	NSE	0.90	0.85	0.55	0.90	0.89
	KGE	0.92	0.89	0.64	0.93	0.92
Testing	MAE	34.97	35.17	75.38	35.99	35.26
	NSE	0.77	0.78	0.33	0.76	0.76
	KGE	0.71	0.72	0.41	0.73	0.72

Figure 6 depicts the measured and predicted runoff for SVM-ANFIS-ACO models through scatter and time series plots. The evaluation of the SVM-ANFIS-ACO models in simulating observed runoff dynamics can be conducted through the analysis of time series graphs. Time series plots allow for the observation of deviations or agreements between measured and simulated runoff. It becomes clear from the analysis that the SVM-ANFIS-ACO models demonstrate varied levels of agreement with the measured runoff. Except for SVM-ANFIS-ACO3, all SVM-ANFIS-ACO models performed well in recreating the measured runoff patterns, showing the closest match to the measured runoff for the majority of the time period. In peak flow simulations, all SVM-ANFIS-ACO models, however, had some drawbacks, most of which were in the form of underestimated values. In terms of capturing the peaks, SVM-ANFIS-ACO2 and SVM-ANFIS-ACO4 outperformed others. This alignment shows that the models successfully simulate the dynamics of runoff and perform consistently. It is evident how the points are distributed and how they relate to the reference line in scatter plots. A tight cluster of locations around the reference line showed that SVM-ANFIS-ACO2 and SVM-ANFIS-ACO4 had a high level of agreement with the measured runoff. The simulated daily runoff values, however, considerably departed from the perfect fit line (i.e., $y=x$) and demonstrated a significant underestimate for all SVM-ANFIS-ACO models. For SVM-ANFIS-ACO2, matching points were more distributed around the 1:1 perfect fit line, while for SVM-ANFIS-ACO3, points deviated more from the perfect fit line.

5 Discussion

RR is a complex problem that is evaluated mostly with the help of either conceptual or AI-based methods. Conceptual models are valuable for better understanding the hydrological process (Kim et al., 2018) but have less computational accuracy. The AI models, on the other hand, have higher computational capability (Adnan et al., 2021; Khodakhah et al., 2022) although physical interpretation of the results is irrelevant in most cases. The main goal of the current study is to integrate the advantages of both conceptual and AI methods in a unique modeling approach to develop not only a robust but also a reliable hydrological model.

In total, 11 models are developed in the current study, including three standalone conceptual models (IHACRES, GR4J, and MISD), three standalone AI models (ANFIS, MLP, and RBF), and five SVM-ANFIS-ACO models based on five SVM kernel functions attributed to five different scenarios. The Taylor diagram of all applied models related to the test

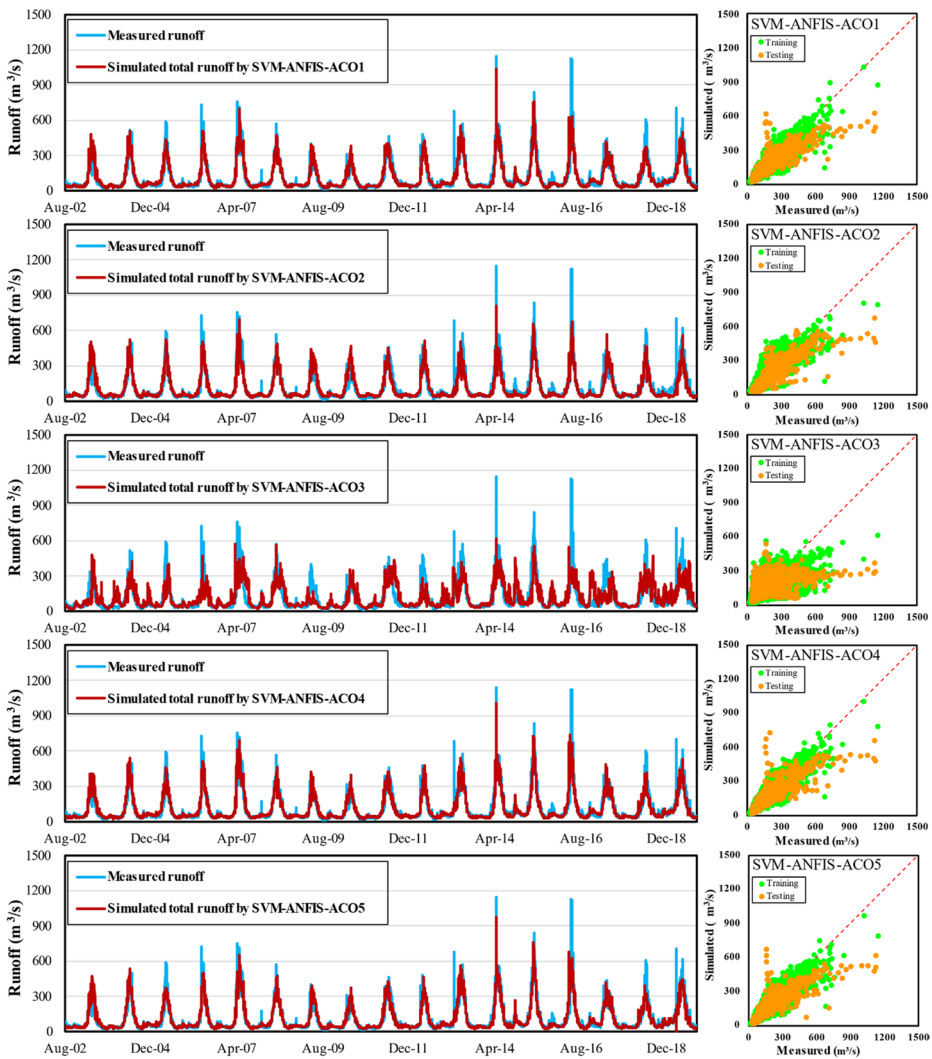


Fig. 6 Hydrographs and scatter plots of hybrid SVM-ANFIS-ACO models based on five different input datasets

section is shown in Fig. 7. The Taylor diagram visually represents the models' accuracy in terms of their proximity to the observed points, that are located at the reference position in the diagram. The radial distance from the observation point to a model estimation depicts the standard deviation of model(s), that examines the dispersion from the predicted data. The azimuthal angle shows the correlation coefficients, indicating the ability of the model to simulate the observed deviations. Hence, the Taylor diagram provides a comprehensive and multi-dimensional tool for the evaluation of the three hydrological models (IHACRES, GR4J, and MISD), three AI-based models (ANFIS, MLP, and RBF), and five coupled models (hybridization of AI with SVM-ANFIS-ACO1, SVM-ANFIS-ACO2, SVM-ANFIS-ACO3, SVM-ANFIS-ACO4, SVM-ANFIS-ACO5 models). Results showed

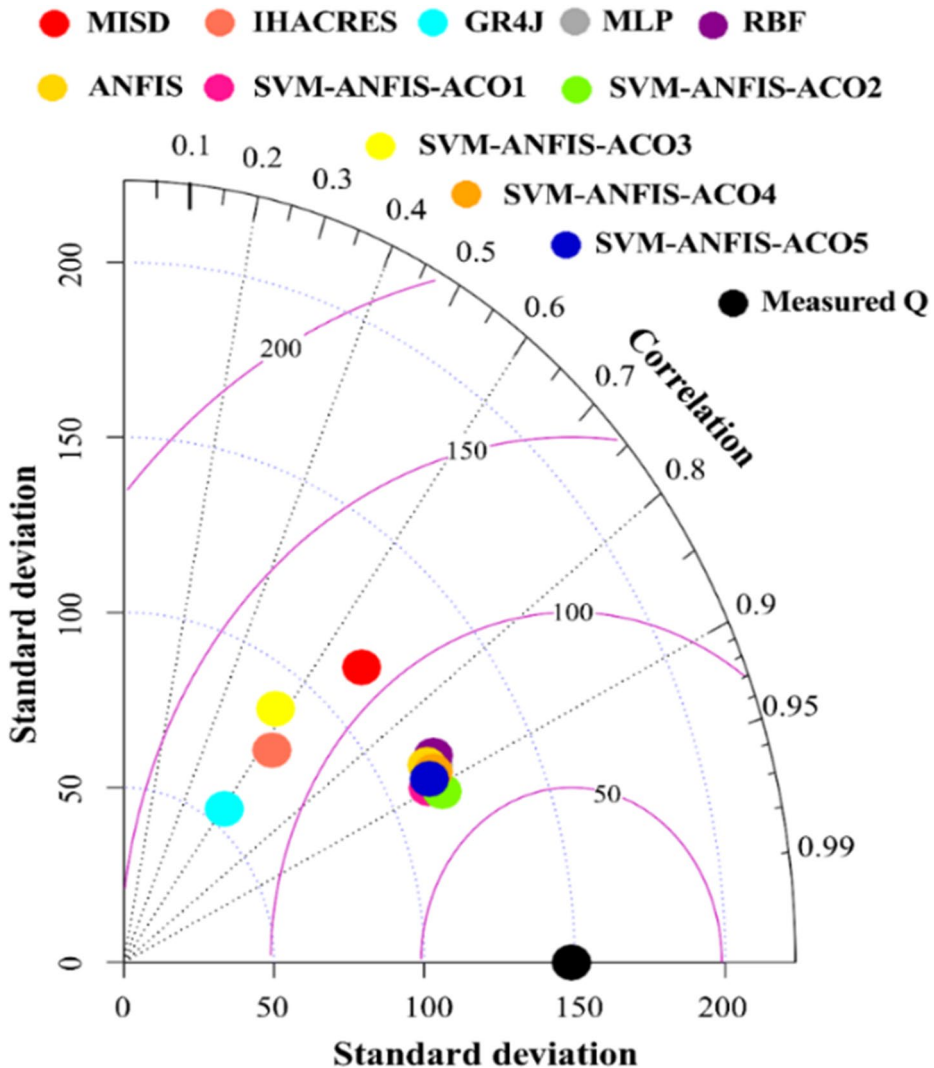


Fig. 7 Taylor diagram related to the simulated runoff by all applied models vs. measured runoff (for testing stages)

that the coupled models performed better (except SVM-ANFIS-ACO3), and therefore it seems the hybridization of AI with hydrological models improves the runoff estimations.

The positions of results of the models on the Taylor diagram reveal several key patterns that inform our understanding of model strengths and weaknesses. The conceptual models (GR4J, IHACRES, and MISD) are able to capture the general pattern of streamflow variations but systematically underestimate flow variability, particularly during peak flow events. A clear trade-off between correlation and accurate representation of variability is evident across model types. The hybrid SVM-ANFIS-ACO2 model is positioned closest to this ideal location. This indicates that the hybrid approach successfully addresses the limitations

of individual models by simultaneously improving correlation and accurately representing flow variability. The convergence of these models toward the reference point demonstrates how the hybrid framework effectively balances different aspects of statistical performance. The positions of different SVM-ANFIS-ACO variants reveal the influence of kernel function selection on model behavior. The SVM-ANFIS-ACO2 model achieved the best balance between correlation and variability representation, while the SVM-ANFIS-ACO3 model shows a slightly lower correlation compared to other SVM-ANFIS-ACO-based models. This pattern highlights how different kernel functions emphasize different aspects of the statistical relationship between predictors and streamflow.

A significant advantage of the proposed multi-step modeling framework is the minimal data requirements compared to traditional physically-based distributed models. The framework operates with basic hydrometeorological (precipitation, temperature, and evapotranspiration) time series as primary inputs, making it particularly suitable for implementation in regions with limited data availability. The minimal data requirements, however, introduce certain considerations regarding the framework's transferability to diverse geographic and climatic conditions. In regions dominated by processes not explicitly represented in our input variables, such as complex snowmelt dynamics in high alpine areas, extensive ground-water-surface water interactions in karst landscapes, or heavily regulated river systems, additional input variables or model adjustments may be beneficial.

The framework's hybrid nature partially mitigates these limitations by allowing the AI components to capture relationships that are not explicitly parameterized in the conceptual models. This compensatory mechanism is evidenced in our study by the substantially improved performance of the hybrid approaches compared to individual conceptual models. Future applications of this framework across diverse watersheds would benefit from a preliminary assessment of dominant hydrological processes and potential limitations of the minimal input approach for the specific context.

The current study exhibits several remarkable advancements in the field of RR modeling when compared to models documented in the hydrological literature. As already highlighted by Zhang et al. (2018), Zhou et al. (2022), and Mohammadi et al. (2022), conventional techniques, mostly conceptual hydrological models, are extensively employed to model RR processes based on the physical principles that govern hydrological phenomena. Although these models provide invaluable contributions to the underlying hydrological mechanisms, they often necessitate a variety of hydrological variables and may have computational efficiency limitations.

AI-based models, namely MLP, RBF, and ANFIS, have become increasingly popular as they possess the ability to deal with complex relationships within datasets through a reduced set of variables. However, the physical justification that is usually connected with conceptual hydrological models is absent in such models. The current study aims to fill the gap between these two modeling paradigms through the amalgamation of the strengths of both hydrological models and AI techniques. Then, we compared some of the common lumped hydrological models, including MISD, IHACRES, and GR4J, with our newly developed approach (SVM-ANFIS-ACO2 model), to increase the accuracy of hydrological modeling and capture the runoff time series pattern.

The hybrid SVM-ANFIS-ACO strategy offers distinct advantages over standalone models in terms of superior prediction accuracy, which can be attributed to three key factors: (1) complementary process representation. By integrating conceptual and AI approaches,

the hybrid strategy leverages both process-based understanding and data-driven pattern recognition, capturing hydrological dynamics that neither approach could fully represent alone. (2) Enhanced feature selection: the SVM-based feature selection process identified the most informative combinations of model outputs and time lags, effectively filtering out noise while preserving critical information about basin behavior. (3) Performance optimization: the ACO algorithm provided more effective optimization of the ANFIS performance compared to ordinary ANFIS, resulting in better convergence to optimal parameter values. These comparative advantages demonstrate that while the hybrid approach requires greater initial investment in model development, it offers substantial benefits in terms of prediction accuracy.

The enhanced accuracy and reliability of our hybrid RR modeling framework offer significant practical implications for water resource managers and decision-makers across multiple sectors. For flood management authorities, the improved prediction accuracy during peak flow events enables more precise flood forecasting, which can directly translate to more effective early warning systems. The framework's ability to capture both the timing and magnitude of flow events enables operators to maintain optimal reservoir levels while meeting multiple operational objectives. Urban water supply managers can utilize more accurate streamflow predictions to better assess water availability, particularly during seasonal transitions and drought onset periods. This enhanced predictive capability supports more informed decisions regarding water conservation measures, supply alternatives, and infrastructure investment planning. Environmental agencies can apply the model results to more accurately assess ecological flow requirements and evaluate the impacts of water abstractions on aquatic ecosystems. The improved prediction of flow regimes enables better protection of aquatic habitats through more precise environmental flow allocations and more effective monitoring of compliance with water quality standards.

6 Conclusions

6.1 Findings

Accurate simulation of runoff components is crucial for water resources management. The sophisticated dynamics of runoff generation and its interactions with hydro-meteorological factors are addressed in the conducted study. To achieve this, a multi-step strategy is used to enhance the estimation of streamflow using lumped and AI-based models. AI techniques are used to ease the simulation of underlying hydrological processes and reduce the bias in the prediction. Afterward, different configurations and scenarios are used to introduce the most suitable runoff model with the help of basic hydro-meteorological variables. The performance of lumped and AI-based models is examined to conclude that the ensemble approach performed better in the reduction of uncertainties when compared to the individual models. In this respect, the SVM-ANFIS-ACO2 model outperformed the other models and was recommended for further applications. The proposed multi-step strategy successfully leveraged the complementary strengths of different modeling approaches. Conceptual models provided physically-based process representations, while AI components captured complex non-linear relationships of RR. Also, the SVM-based feature selection with multiple kernel functions proved highly effective in identifying optimal input configurations. The Radial

kernel (used in SVM-ANFIS-ACO2) demonstrated superior performance in capturing the complex temporal dynamics of the watershed.

6.2 Research limitations

Each modeling approach employed in this study has inherent limitations that can affect performance in RR applications. The conceptual models use mathematical approximations of physical processes that may not fully capture complex watershed dynamics, particularly during extreme events. The standalone AI models offer limited physical interpretability, making it difficult to connect their operations with actual hydrological processes, and these models often perform poorly when predicting conditions outside their training data range. The applied feature selection methods also are sensitive to kernel selection, the performance of SVM for feature selection is highly dependent on appropriate kernel selection, with different kernels identifying different variable relationships. By combining conceptual and AI approaches, we leveraged their complementary strengths. The conceptual models provided physically-based process representation, while the AI components captured complex non-linear relationships not explicitly parameterized in the conceptual structure. This integration reduced the impact of simplified process representation in conceptual models and enhanced the physical relevance of AI approaches. Evidence of this complementary benefit is seen in the improved peak flow predictions, where the hybrid SVM-ANFIS-ACO2 model reduced peak flow error compared to the best standalone conceptual model. The SVM-based feature selection process with multiple kernel functions addressed the limitations of both conceptual and AI models by: (1) Identifying the most informative model outputs and lag times, effectively filtering out noise while preserving critical information. (2) Reducing the dimensionality of inputs to subsequent models, minimizing overfitting risk. (3) Creating a more robust input structure that improved extrapolation performance.

6.3 Recommendations for future research

Future research directions can explore the integration of other advanced AI architectures into this framework, particularly models such as Long Short-Term Memory (LSTM) networks and Transformer-based models. These architectures could potentially enhance the framework's ability to capture complexities in the RR relationship. Enhanced physical interpretability should be explored through process-constrained AI approaches that maintain superior predictive performance while improving transparency. This could involve developing visualization techniques for internal model states or implementing physics-informed neural networks that incorporate hydrological constraints. Multi-watershed validation studies across diverse geographical and climatic conditions would provide valuable insights into the framework's transferability limits and adaptation requirements. Such studies should systematically evaluate performance across watersheds with different characteristics, including size, climate, land use, and geological settings. Future research should focus on adding soil moisture indices as inputs, which would improve model sensitivity to changing watershed conditions, particularly during transition periods between dry and wet seasons. Incorporate attention-based neural network architectures that can dynamically focus on different input variables and time lags based on current conditions. Incorporate signal processing approaches (e.g., wavelet transforms) to better capture multi-scale temporal patterns in

both inputs and outputs, improving the representation of periodic features. These targeted improvements would address the specific error patterns observed in our current modeling framework while maintaining the advantages of our hybrid approach.

Acknowledgements We utilized Claude 3.7 Sonnet (developed by Anthropic) to assist with improving language quality throughout the manuscript. All content was thoroughly reviewed, verified, and edited by the authors to ensure scientific accuracy.

Author contributions

Term	Author(s)
Conceptualization	Babak Mohammadi, Mir Jafar Sadegh Safari
Methodology	Babak Mohammadi, Mustafa Utku Yilmaz, Babak Vaheddoost
Implementation of the computer code and algorithms	Babak Mohammadi
Validation	Mir Jafar Sadegh Safari, Babak Vaheddoost
Formal analysis	Babak Mohammadi, Mir Jafar Sadegh Safari
Investigation	Babak Mohammadi, Mir Jafar Sadegh Safari, Babak Vaheddoost, Mustafa Utku Yilmaz
Resources	Mustafa Utku Yilmaz
Data Curation	Mustafa Utku Yilmaz
Writing - Original Draft	Babak Mohammadi, Mir Jafar Sadegh Safari, Babak Vaheddoost, Mustafa Utku Yilmaz
Writing - Review & Editing	Mir Jafar Sadegh Safari, Babak Vaheddoost, Mustafa Utku Yilmaz
Supervision	Mir Jafar Sadegh Safari
Visualization	Babak Mohammadi, Mir Jafar Sadegh Safari, Babak Vaheddoost

Funding Not applicable.

Data availability The data are available by the corresponding author upon request.

Code availability The codes are available by the corresponding author upon request.

Declarations

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Conflict of interest The authors declare no competing interests.

References

- Abbot, J., & Marohasy, J. (2012). Application of artificial neural networks to rainfall forecasting in Queensland, Australia. *Advances in Atmospheric Sciences*, 29(4), 717–730. <https://doi.org/10.1007/s00376-012-1259-9>
- Adnan, R. M., Mostafa, R. R., Kisi, O., Yaseen, Z. M., Shahid, S., & Zounemat-Kermani, M. (2021). Improving streamflow prediction using a new hybrid ELM model combined with hybrid particle swarm optimization and grey wolf optimization. *Knowledge-Based Systems*, 230, Article 107379. <https://doi.org/10.1016/j.knsys.2021.107379>
- Aghelpour, P., Bahrami-Pichaghchi, H., & Kisi, O. (2020). Comparison of three different bio-inspired algorithms to improve ability of neuro fuzzy approach in prediction of agricultural drought, based on three different indexes. *Computers and Electronics in Agriculture*, 170, Article 105279. <https://doi.org/10.1016/j.compag.2020.105279>
- Aytek, A., Asce, M., & Alp, M. (2008). An application of artificial intelligence for rainfall-runoff modeling. *Journal of Earth System Science*, 117(2), 145–155. <https://doi.org/10.1007/s12040-008-0005-2>
- Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, 10(2–3), 191–203. [https://doi.org/10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)
- Boodoo, F., Hostache, R., Skifa, N., Guerin, J., & Delenne, C. (2025). Are LSTM and conceptual rainfall-runoff models able to cope with limited training datasets under diverse hydrometeorological conditions? *Modeling Earth Systems and Environment*, 11(2), Article 128. <https://doi.org/10.1007/s40808-025-02316-z>
- Boumenni, H., Bachnou, A., & Alaa, N. E. (2017). The rainfall-runoff model GR4J optimization of parameter by genetic algorithms and Gauss-Newton method: Application for the watershed Ourika (High Atlas, Morocco). *Arabian Journal of Geosciences*, 10, Article 343. <https://doi.org/10.1007/s12517-017-3086-x>
- Brocca, L., Melone, F., & Moramarco, T. (2011). Distributed rainfall-runoff modelling for flood frequency estimation and flood forecasting. *Hydrological Processes*, 25(18), 2801–2813. <https://doi.org/10.1002/hyp.8042>
- Chow, V. T. (1965). *Handbook of applied hydrology*. International Association of Scientific Hydrology. Bulletin. 10. <https://doi.org/10.1080/02626666509493376>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Croke, B. F. W., & Jakeman, A. J. (2004). A catchment moisture deficit module for the IHACRES rainfall-runoff model. *Environmental Modelling & Software*, 19(1), 1–5. <https://doi.org/10.1016/j.envsoft.2003.09.001>
- Dorigo, M. (1992). Optimization, learning and natural algorithms. PhD Thesis. Politecnico di Milano.
- Dorigo, M., Di Caro, G., & Gambardella, L. M. (1999). Ant algorithms for discrete optimization. *Artificial Life*, 5(2), 137–172. <https://doi.org/10.1162/106454699568728>
- Duethmann, D., Blöschl, G., & Parajka, J. (2020). Why does a conceptual hydrological model fail to correctly predict discharge changes in response to climate change? *Hydrology and Earth System Sciences*, 24(7), 3493–3511. <https://doi.org/10.5194/hess-24-3493-2020>
- Fahad, S., Su, F., Khan, S. U., Naeem, M. R., & Wei, K. (2023). Implementing a novel deep learning technique for rainfall forecasting via climatic variables: An approach via hierarchical clustering analysis. *Science of the Total Environment*, 854, Article 158760. <https://doi.org/10.1016/j.scitotenv.2022.158760>
- Fathi, M. M., Mehedi, A., Abdullah, M., Smith, V., Fernandes, A. M., Hren, M. T., & Terry, D. O. (2025). Evaluation of LSTM vs. conceptual models for hourly rainfall runoff simulations with varied training period lengths. *Scientific Reports*, 15, 15820. <https://doi.org/10.1038/s41598-025-96577-4>
- 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. [https://doi.org/10.1061/\(asce\)jhe.1943-5584.0000533](https://doi.org/10.1061/(asce)jhe.1943-5584.0000533)
- Garavaglia, F., et al. (2017). Impact of model structure on flow simulation and hydrological realism: From a lumped to a semi-distributed approach. *Hydrology and Earth System Sciences*, 21(8), 3937–3952. <https://doi.org/10.5194/hess-21-3937-2017>
- GDWM (General Directorate of Water Management). (2020). Coruh Basin flood management plan. Ankara: The Republic of Türkiye Ministry of Agriculture and Forestry. <https://www.tarimorman.gov.tr/SYGM/Sayfalar/Detay.aspx?SayfaId=53>

- Gunes, M., Parim, C., Yildiz, D., & Buyuklu, A. (2021). Predicting monthly streamflow using a hybrid wavelet neural network: Case study of the Coruh River Basin. *Polish Journal of Environmental Studies*, 30(4). <https://doi.org/10.15244/pjoes/130767>
- Gupta, H. V., Kling, H., Yilmaz, K. K., & Martinez, G. F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. <https://doi.org/10.1016/j.jhydrol.2009.08.003>
- Heddad, S. (2016). New modelling strategy based on radial basis function neural network (RBFNN) for predicting dissolved oxygen concentration using the components of the Gregorian calendar as inputs: Case study of Clackamas River, Oregon, USA. *Modeling Earth Systems and Environment*, 2(4), 1–5. <https://doi.org/10.1007/s40808-016-0232-5>
- Herath, H. M. V. V., Chadalawada, J., & Babovic, V. (2021). Hydrologically informed machine learning for rainfall–runoff modelling: Towards distributed modelling. *Hydrology and Earth System Sciences*, 25(8), 4373–4401. <https://doi.org/10.5194/hess-25-4373-2021>
- Höge, M., Scheidegger, A., Baity-Jesi, M., Albert, C., & Fenicia, F. (2022). Improving hydrologic models for predictions and process understanding using neural ODEs. *Hydrology and Earth System Sciences*, 26(19), 5085–5102. <https://doi.org/10.5194/hess-26-5085-2022>
- Hosseini, S. M., & Mahjouri, N. (2016). Integrating support vector regression and a geomorphologic artificial neural network for daily rainfall-runoff modeling. *Applied Soft Computing*, 38, 329–345. <https://doi.org/10.1016/j.asoc.2015.09.049>
- Hung, N. Q., et al. (2009). An artificial neural network model for rainfall forecasting in Bangkok, Thailand. *Hydrology and Earth System Sciences*, 13(8), 1413–1425. <https://doi.org/10.5194/hess-13-1413-2009>
- Jakeman, A. J., & Hornberger, G. M. (1993). How much complexity is warranted in a rainfall-runoff model? *Water Resources Research*, 29(8), 2637–2649. <https://doi.org/10.1029/93WR00877>
- Jang, J. S. R., Sun, C. T., & Mizutani, E. (1997). Neuro-fuzzy and soft computing—a computational approach to learning and machine intelligence [Book review]. *IEEE Transactions on Automatic Control*, 42(10), 1482–1484.
- Kavaklioglu, K., Koseoglu, M. F., & Caliskan, O. (2018). Experimental investigation and radial basis function network modeling of direct evaporative cooling systems. *International Journal of Heat and Mass Transfer*, 126, 139–150. <https://doi.org/10.1016/j.ijheatmasstransfer.2018.05.022>
- Khodakhah, H., Aghelpour, P., & Hamed, Z. (2022). Comparing linear and non-linear data-driven approaches in monthly river flow prediction, based on the models SARIMA, LSSVM, ANFIS, and GMDH. *Environmental Science and Pollution Research*, 29(15), 21935–21954. <https://doi.org/10.1007/s11356-021-17443-0>
- Khosravi, A., Nunes, R. O., Assad, M. E. H., & Machado, L. (2018). Comparison of artificial intelligence methods in estimation of daily global solar radiation. *Journal of Cleaner Production*, 194, 342–358. <https://doi.org/10.1016/j.jclepro.2018.05.147>
- Kim, K. B., Kwon, H. H., & Han, D. (2018). Exploration of warm-up period in conceptual hydrological modelling. *Journal of Hydrology*, 556, 194–210. <https://doi.org/10.1016/j.jhydrol.2017.11.015>
- Kim, T., Shin, J. Y., Kim, H., & Heo, J. H. (2020). Ensemble-based neural network modeling for hydrologic forecasts: Addressing uncertainty in the model structure and input variable selection. *Water Resources Research*, 56(6), Article e2019WR026262. <https://doi.org/10.1029/2019WR026262>
- Kim, T., Yang, T., Gao, S., Zhang, L., Ding, Z., Wen, X., Courly, J. J., & Hong, Y. (2021). Can artificial intelligence and data-driven machine learning models match or even replace process-driven hydrologic models for streamflow simulation?: A case study of four watersheds with different hydro-climatic regions across the CONUS. *Journal of Hydrology*, 598, 126423. <https://doi.org/10.1016/j.jhydrol.2021.126423>
- Kisi, O., Demir, V., & Kim, S. (2017). Estimation of long-term monthly temperatures by three different adaptive neuro-fuzzy approaches using geographical inputs. *Journal of Irrigation and Drainage Engineering*, 143(12), Article 04017052. [https://doi.org/10.1061/\(ASCE\)IR.1943-4774.0001242](https://doi.org/10.1061/(ASCE)IR.1943-4774.0001242)
- Kumanlioglu, A. A., & Fistikoglu, O. (2019). Performance enhancement of a conceptual hydrological model by integrating artificial intelligence. *Journal of Hydrologic Engineering*, 24(11), Article 04019047. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001850](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001850)
- Masseroni, D., Cislighi, A., Camici, S., Massari, C., & Brocca, L. (2017). A reliable rainfall–runoff model for flood forecasting: Review and application to a semi-urbanized watershed at high flood risk in Italy. *Hydrology Research*, 48(3), 726–740. <https://doi.org/10.2166/nh.2016.037>
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5, 115–133. <https://doi.org/10.1007/BF02478259>
- Mehr, A. D., Vaheddoost, B., & Mohammadi, B. (2020). ENN-SA: A novel neuro-annealing model for multi-station drought prediction. *Computers & Geosciences*, 145, Article 104622. <https://doi.org/10.1016/j.cageo.2020.104622>
- Mehr, A. D., Kahya, E., Sahin, A., & Nazemosadat, M. J. (2015). Successive-station monthly streamflow prediction using different artificial neural network algorithms. *International Journal of Environmental Science and Technology*, 12, 2191–2200. <https://doi.org/10.1007/s13762-014-0613-0>

- Mohammadi, B., Safari, M. J. S., & Vazifekkhah, S. (2022). IHACRES, GR4J and MISD-based multi conceptual-machine learning approach for rainfall-runoff modeling. *Scientific Reports*, 12(1), Article 12096. <https://doi.org/10.1038/s41598-022-16215-1>
- Mohammadi, B., Gao, H., Feng, Z., Pilesjö, P., Cheraghalizadeh, M., & Duan, Z. (2023). Simulating glacier mass balance and its contribution to runoff in Northern Sweden. *Journal of Hydrology*, Article 129404. <https://doi.org/10.1016/j.jhydrol.2023.129404>
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, 10(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Peredo, D., Ramos, M. H., Andréassian, V., & Oudin, L. (2022). Investigating hydrological model versatility to simulate extreme flood events. *Hydrological Sciences Journal*, 67(4), 628–645. <https://doi.org/10.1080/02626667.2022.2030864>
- Perrin, C., Michel, C., & Andréassian, V. (2001). Does a large number of parameters enhance model performance? Comparative assessment of common catchment model structures on 429 catchments. *Journal of Hydrology*, 242(3–4), 275–301.
- Perrin, C., Michel, C., & Andréassian, V. (2003). Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, 279(1–4), 275–289. [https://doi.org/10.1016/S0022-1694\(03\)00225-7](https://doi.org/10.1016/S0022-1694(03)00225-7)
- Rasouli, K., Hsieh, W. W., & Cannon, A. J. (2012). Daily streamflow forecasting by machine learning methods with weather and climate inputs. *Journal of Hydrology*, 414, 284–293. <https://doi.org/10.1016/j.jhydrol.2011.10.039>
- Sanikhani, H., Nikpour, M. R., & Jamshidi, F. (2025). Advanced framework for predicting rainfall-runoff: Comparative evaluation of AI models for enhanced forecasting accuracy. *Water Resources Management*, 39, 3273–3294. <https://doi.org/10.1007/s11269-025-04106-9>
- Selek, B., & Aksu, H. (2020). Water resources potential of Turkey. In N. Harmancioglu, & D. Altinbilek (Eds.), *Water resources of Turkey. World water resources* (Vol. 2). Springer. https://doi.org/10.1007/978-3-030-11729-0_8
- Senatore, A., Furnari, L., & Mendicino, G. (2020). Impact of high-resolution sea surface temperature representation on the forecast of small Mediterranean catchments' hydrological responses to heavy precipitation. *Hydrology and Earth System Sciences*, 24(1), 269–291. <https://doi.org/10.5194/hess-24-269-2020>
- Sezen, C., Bezak, N., Bai, Y., & Šraj, M. (2019). Hydrological modelling of karst catchment using lumped conceptual and data mining models. *Journal of Hydrology*, 576, 98–110. <https://doi.org/10.1016/j.jhydrol.2019.06.036>
- Shamseldin, A. Y., O'Connor, K. M., & Nasr, A. E. (2007). A comparative study of three neural network forecast combination methods for simulated river flows of different rainfall-runoff models. *Hydrological Sciences Journal*, 52(5), 896–916. <https://doi.org/10.1623/hysj.52.5.896>
- Shekar, P. R., Mathew, A., & Sharma, K. V. (2025). A hybrid CNN–RNN model for rainfall–runoff modeling in the Potteruvagu watershed of India. *CLEAN–Soil, Air, Water*, Article 2300341. <https://doi.org/10.1002/clen.202300341>
- Singh, V. P. (2018). Hydrologic modeling: Progress and future directions. *Geoscience Letters*, 5(1), Article 15. <https://doi.org/10.1186/s40562-018-0113-z>
- Stephens, C. M., Marshall, L. A., & Johnson, F. M. (2019). Investigating strategies to improve hydrologic model performance in a changing climate. *Journal of Hydrology*, 579, Article 124219. <https://doi.org/10.1016/j.jhydrol.2019.124219>
- Takagi, T., & Sugeno, M. (1985). Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 2(1), 116–132. <https://doi.org/10.1109/TSMC.1985.6313399>
- Tan, M. L., Gassman, P. W., Yang, X., & Haywood, J. (2020). A review of SWAT applications, performance and future needs for simulation of hydro-climatic extremes. *Advances in Water Resources*, 143, Article 103662. <https://doi.org/10.1016/j.advwatres.2020.103662>
- Tao, W., Shao, F., Su, L., Wang, Q., Zhou, B., & Sun, Y. (2023). An analytical model for simulating the rainfall-interception-infiltration-runoff process with non-uniform rainfall. *Journal of Environmental Management*, 344, Article 118490. <https://doi.org/10.1016/j.jenvman.2023.118490>
- Tian, J., Liu, J., Wang, Y., Wang, W., Li, C., & Hu, C. (2020). A coupled atmospheric–hydrologic modeling system with variable grid sizes for rainfall–runoff simulation in semi-humid and semi-arid watersheds: How does the coupling scale affects the results? *Hydrology and Earth System Sciences*, 24(8), 3933–3949. <https://doi.org/10.5194/hess-24-3933-2020>
- Vaheddoost, B., Safari, M. J. S., & Yilmaz, M. U. (2023a). Rainfall-runoff simulation in ungauged tributary streams using drainage area ratio-based multivariate adaptive regression spline and random forest hybrid models. *Pure and Applied Geophysics*, 180(1), 365–382. <https://doi.org/10.1007/s00024-022-03209-3>
- Vaheddoost, B., Yilmaz, M. U., & Safari, M. J. S. (2023b). Estimation of flow duration and mass flow curves in ungauged tributary streams. *Journal of Cleaner Production*, 409, Article 137246. <https://doi.org/10.1016/j.jclepro.2023.137246>

- Van Kempen, G., Van Der Wiel, K., & Melsen, L. A. (2021). The impact of hydrological model structure on the simulation of extreme runoff events. *Natural Hazards and Earth System Sciences*, 21(3), 961–976. <https://doi.org/10.5194/nhess-21-961-2021>
- Wang, W. C., et al. (2009). A comparison of performance of several artificial intelligence methods for forecasting monthly discharge time series. *Journal of Hydrology*, 374(3–4), 294–306. <https://doi.org/10.1016/j.jhydrol.2009.06.019>
- Xiong, L., & O'connor, K. M. (2002). Comparison of four updating models for real-time river flow forecasting. *Hydrological Sciences Journal*, 47(4), 621–639. <https://doi.org/10.1080/02626660209492964>
- Yaseen, Z. M., et al. (2017). Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model. *Journal of Hydrology*, 554, 263–276. <https://doi.org/10.1016/j.jhydrol.2017.09.007>
- Yilmaz, M. U. (2025). Enhancing sustainable flood risk management through a novel fuzzy multi-stakeholder framework. *Water Resources Management*, 39(3), 1453–1480. <https://doi.org/10.1007/s11269-024-04012-6>
- Yilmaz, M. U., Aksu, H., Onoz, B., & Selek, B. (2023). An effective framework for improving performance of daily streamflow estimation using statistical methods coupled with artificial neural network. *Pure and Applied Geophysics*, 180, 3639–3654. <https://doi.org/10.1007/s00024-023-03344-5>
- Young, C. C., & Liu, W. C. (2015). Prediction and modelling of rainfall–runoff during typhoon events using a physically-based and artificial neural network hybrid model. *Hydrological Sciences Journal*, 60(12), 2102–2116. <https://doi.org/10.1080/02626667.2014.959446>
- Zhang, B., & Govindaraju, R. S. (2003). Geomorphology-based artificial neural networks (GANNs) for estimation of direct runoff over watersheds. *Journal of Hydrology*, 273(1–4), 18–34. [https://doi.org/10.1016/S0022-1694\(02\)00313-X](https://doi.org/10.1016/S0022-1694(02)00313-X)
- Zhang, Z., Zhang, Q., Singh, V. P., & Shi, P. (2018). River flow modelling: Comparison of performance and evaluation of uncertainty using data-driven models and conceptual hydrological model. *Stochastic Environmental Research and Risk Assessment*, 32, 2667–2682. <https://doi.org/10.1007/s00477-018-1536-y>
- Zhong-Da, T., Xian-Wen, G., & Tong, S. (2014). Combination kernel function least squares support vector machine for chaotic time series prediction. *Acta Physica Sinica*, 63(16), 160508. <https://doi.org/10.7498/aps.63.160508>
- Zhou, Y., Cui, Z., Lin, K., Sheng, S., Chen, H., Guo, S., & Xu, C. Y. (2022). Short-term flood probability density forecasting using a conceptual hydrological model with machine learning techniques. *Journal of Hydrology*, 604, Article 127255. <https://doi.org/10.1016/j.jhydrol.2021.127255>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Babak Mohammadi¹  · **Mir Jafar Sadegh Safari**^{2,3}  · **Babak Vaheddoost**⁴  ·
Mustafa Utku Yilmaz⁵ 

✉ Mir Jafar Sadegh Safari
jafar.safari@torontomu.ca

Babak Mohammadi
babak.mohammadi@smhi.se

Babak Vaheddoost
babak.vaheddoost@btu.edu.tr

Mustafa Utku Yilmaz
utkuyilmaz@klu.edu.tr

- ¹ Hydrology Research Unit, Swedish Meteorological and Hydrological Institute, Norrköping, Sweden
- ² Department of Geography and Environmental Studies, Toronto Metropolitan University, Toronto, Ontario, Canada
- ³ Department of Civil Engineering, Yasar University, Izmir, Türkiye
- ⁴ Department of Civil Engineering, Bursa Technical University, Bursa, Türkiye
- ⁵ Department of Civil Engineering, Kırklareli University, Kırklareli, Türkiye