Larhyss Journal, ISSN 1112-3680, n°64, Dec 2025, pp. 7-32

APPLICATION OF INTELLIGENT AUTOCORRELATED MODELS FOR RUNOFF SIMULATION A CASE STUDY OF THE IRANIAN SAMIYAN AND DOOSTBIGHLOU RIVERS

MOLAVI A.

Assistant Professor, Department of Water Sciences and Engineering, Ta.C., Islamic Azad University, Tabriz, Iran

ahad.molavi@iau.ac.ir

Research Article – Available at http://larhyss.net/ojs/index.php/larhyss/index
Received November 2, 2024, Received in revised form November 1, 2025, Accepted November 3, 2025

ABSTRACT

Auto-correlated simulation is one of the effective methods for predicting river flows in basins that do not have information and statistics on various climatic factors. In this study, the flow of Doostbighlou and Samiyan Rivers located in Ardabil province, Iran, was predicted with intelligent meta-exploration models of artificial neural networks, support vector machines, and their combination with wavelet transform, considering acceptable runoff delay times. The evaluation results showed that considering the wavelet transform in the individual methods of artificial neural networks and support vector machines can be an effective help in improving the performance of the above models. The t-test analysis indicated that at a 5% significance level, the null hypothesis (**H_0: $\mu_1 = \mu_2$**) was accepted only for the hybrid Wavelet-ANN and Wavelet-SVM models, while the alternative hypothesis (**H_1: $\mu_1 \neq \mu_2$**) was accepted for the standalone ANN and SVM models. Thus, at Samiyan station, hybridizing the individual artificial neural network model with the wavelet model has increased and decreased the R and RMSE parameters from 0.48 and 1.96 m^3s^{-1} to 0.82 and 1.02 m^3s^{-1}, respectively. Also, at Doostbighlou station, integrating artificial neural network and support vector machine models with wavelet transform analysis significantly improved their correlation coefficients, increasing them from 0.39 and 0.45 to 0.77 and 0.72, respectively.

Keywords: Hybrid wavelet, Hydrological variables, Time-delayed, Stream flow.

INTRODUCTION

Since the runoff series exhibits nonlinear characteristics, capturing the periodicity and regularity in the runoff series using a single model is challenging (Kalteh, 2013; Chibane and Ali-Rahmani, 2015; Atallah et al., 2024). The use of machine learning tools in hydrological predictions and simulations has made significant progress, especially in watersheds with little information and data (Mushtaq et al., 2024). Several models and methods are used in river flow simulation, and their performance and applications vary depending on the basin structure, the necessary and available data, and the purpose of the simulation. In order to simulate runoff in basins that do not have sufficient statistics and information, it will not be possible to use physical models that require more meteorological and hydrological variables. The autocorrelated stream flow simulation method, in which the only input data are the delayed stream flow discharges to the model, is a pioneering approach in such basins. In recent years, various meta-heuristic models such as support vector machines, artificial neural networks, combined wavelet-support vector machine models, and hybrid wavelet-artificial neural network models have found wide application in hydrological research and phenomena for which there is no specific algorithm.

Adamowski (2013) used support vector machine and artificial neural networks models to simulate rainfall-runoff in a mountainous basin with limited data in Uttaranchal, India. His study results showed the ability of the support vector machine model to predict direct runoff, base flow, and total flow in this mountainous and hilly region. Farajzadeh et al., (2014) used time series and artificial neural networks models to predict monthly flow and precipitation in the Urmia Lake basin in Iran. The results of flow prediction indicate that both models have good accuracy in estimating monthly flow. However, according to this study, there is not much difference between the two models compared in flow estimation. Cannas et al., (2006) investigated the effectiveness of preprocessing data in the application of artificial neural network models using continuous and discontinuous wavelet transforms. Their results showed that training the network with preprocessing data performed better than training the indecomposable network on chaotic immature signals. Andalib et al., (2020) simulated runoff at several stations using wavelet transform, self-organizing artificial neural network, and artificial intelligence methods in the Little River Watershed (LRW). The results showed that AI (Artificial Intelligence) models combined with wavelet transform, self-organizing artificial neural network, and shared information improve the ability to predict multi-station runoff by up to 23% compared to AI models that use the Markov Property. Kiani Asl et al., (2023) prepared the flood potential map of the Maroon Basin using the random forest and support vector machine learning methods. In this study, the required parameters were prepared and then converted into a readable format for the R software environment to run the random vector machine and random forest model. The results showed that the RF and SVM methods simulated the flood potential map of the Maroon Basin with an accuracy of 0.997, 0.947%, respectively. Dehghani et al., (2022) studied the performance of the hybrid wavelet-support vector machine model to estimate the river discharge of the Dez basin based on daily hydrometric statistics of stations located upstream of the dam and compared its results with the support vector model. The results of the study indicated the acceptability of hybrid structures in runoff modeling and the hybrid wavelet-support vector machine model had better performance in flow prediction. Samantaray et al., (2022) studied the hybrid SVM-SSA (Support Vector Machine with Salp Swarm Algorithm) model and conventional SVM and artificial neural network (ANN) models for runoff prediction in the Baitarani River Basin, Odisha, India. The test results showed that the hybrid models had better prediction accuracy compared to conventional methods and could be recommended in modeling for the precipitation-runoff process and runoff prediction. Mehta et al., (2023) employed GIS, remote sensing, and the Curve Number method to assess the effects of land use/land cover (LULC) changes on runoff dynamics in the Ambica River Basin from 1990 to 2010. Their findings revealed substantial LULC alterations, elevating runoff potential and flood vulnerability, underscoring the need for sustainable land use policies. The study emphasizes LULC management as a key strategy for reducing hydrological risks. Verma et al., (2023) performed a comparative evaluation of CMIP5 and CMIP6 models to examine their efficacy in simulating hydrological processes within a reservoir catchment located in Chhattisgarh, India. The study identified notable discrepancies in climate projections between the two model ensembles, emphasizing their critical implications for water resource management strategies. The findings highlight the necessity of meticulous model selection in regional hydrological assessments under evolving climatic conditions. Mehta et al., (2023) employed hierarchical clustering and Thiessen polygons to optimize rain gauge networks for flood forecasting in the Narmada River Basin, using data from 2010–2018. The study categorized stations into two clusters and applied the HEC-HMS model for rainfall-runoff analysis, validating results with observed dam inflows. Regression analysis confirmed strong correlations between observed and simulated runoff, demonstrating the model's reliability. The findings emphasize the adequacy of existing rain gauge networks and the effectiveness of the proposed approach for accurate flood prediction in similar basins. Kumar et al., (2023) evaluated multiple ML models, including CatBoost, XGBoost, and LGBM, for forecasting river inflow in the Garudeshwar watershed, integrating temporal lag and seasonal data. Their findings demonstrated CatBoost's superior performance in minimizing prediction errors (MAE, RMSE) and achieving high R^2 values, outperforming other models, particularly in handling categorical and continuous variables. The study underscores ML's effectiveness in hydrological modeling, providing valuable insights for flood control and water supply planning. Kantharia and Mehta (2024) developed an integrated rainfall-runoff model based on fuzzy logic, incorporating soil moisture dynamics to enhance hydrological simulations. Their study, conducted in the Damanganga Basin, demonstrated that fuzzy logic effectively addresses uncertainties inherent in rainfall-runoff processes. The proposed model exhibited improved predictive accuracy in runoff estimation by explicitly accounting for soil moisture variability. This research underscores the efficacy of fuzzy logic-based approaches in hydrological modeling, particularly in regions with limited data availability. Mehta and Yadav (2024) investigated rainfall-runoff modeling in the Purna River Basin utilizing the HEC-HMS hydrological model. The study adopted the Green-Ampt method for infiltration loss estimation and evaluated the performance of the SCS and Snyder unit hydrograph methods for runoff transformation. Comparative analysis revealed that the SCS unit hydrograph method exhibited superior predictive accuracy, with a coefficient of determination of 0.9680 and a Nash-Sutcliffe efficiency (NSE) of 0.928, outperforming the Snyder method. The findings suggest that the SCS-based approach offers greater reliability for runoff simulation in the Purna River Basin. Kantharia et al., (2024) utilized an Adaptive Neuro-Fuzzy Inference System (ANFIS) integrated with soil moisture data to improve rainfall–runoff modeling in the Damanganga River basin. Their results demonstrated enhanced predictive performance, emphasizing the critical role of soil moisture dynamics in hydrological simulations for effective water resource management. The study reinforces the applicability of ANFIS as a robust computational tool for capturing complex nonlinear relationships in catchment-scale hydrological processes. Panda et al., (2025) studied WE (Wavelet Ensemble) models for several estimation cases under various ground and flow conditions and data and compared them with individual ML models. The results indicated that WE models had better acceptability than individual and ensemble ML models. Given the demonstrated efficacy of Support Vector Machine (SVM) and Artificial Neural Network (ANN) methods, as well as their hybrid models incorporating wavelet transforms, in hydrological modeling, there remains a significant research gap regarding the application of autocorrelation analysis in runoff prediction. Specifically, a comprehensive evaluation of these models using both direct and lagged runoff data has yet to be conducted for the Doostbighlou and Samiyan Rivers in Ardabil Province, Iran. These rivers serve as critical water sources, supplying drinking water and supporting agricultural activities in the region, which is a key agricultural hub. To address this gap, the present study aims to assess the predictive performance of four modeling approaches: standalone Artificial Neural Network (ANN), standalone Support Vector Machine (SVM), Wavelet-Artificial Neural Network (W-ANN) hybrid model, and Wavelet-Support Vector Machine (W-SVM) hybrid model, in the autocorrelation-based simulation of monthly runoff in the aforementioned rivers.

The relevant study by Koua et al. (2019), investigates the relationship between rainfall and runoff in the Buyo Lake watershed in southwestern Côte d'Ivoire, particularly in the context of climate change. Through hydrological analysis and modeling, the authors assess how variations in precipitation patterns, driven by climate variability, affect surface runoff dynamics in the region. The research highlights increasing irregularities in rainfall distribution, leading to both flood risks and water resource challenges. Their findings underscore the importance of integrated watershed management strategies to adapt to evolving climatic conditions and ensure sustainable water resource use in the basin.

In their pertinent 2016 study, Faregh and Benkhaled apply a GIS-based implementation of the SCS-CN (Soil Conservation Service - Curve Number) method to estimate surface runoff in the Sigus watershed, located in northeastern Algeria. The research integrates geographic information systems (GIS) with hydrological modeling to analyze land use, soil types, and rainfall data for accurate runoff prediction. The study demonstrates the efficiency of combining spatial analysis tools with empirical hydrological methods to support watershed management, flood mitigation, and water resource planning, particularly in semi-arid regions sensitive to climatic variability (Argaz, 2018; Assemian et al., 2021; Chadee et al., 2023).

In their significant 2015 publication, Abdi and Meddi present a distributed rainfall–runoff modeling study applied to two watersheds in eastern Algeria. Their research focuses on evaluating hydrological responses under varying climatic and topographic conditions using a distributed modeling approach. By incorporating spatial heterogeneity in rainfall distribution, land use, and soil characteristics, the study aims to improve the accuracy of runoff simulations. The findings contribute to a better understanding of catchment-scale hydrological behavior, offering valuable insights for flood risk management, hydraulic infrastructure design, and sustainable water resource planning in semi-arid and mountainous environments (Benslimane et al., 2020; Rouissat and Smail, 2022; Baudhanwala et al., 2023; Ben Said et al., 2024; Ezz, 2025; Do et al., 2025).

On the other hand, in their important 2020 article, Riahi et al. examine runoff and soil erosion processes within homogeneous hydrological units of gentle slope located in a watershed in the middle valley of the Medjerda River in Tunisia. The study emphasizes the influence of land topography, soil structure, and rainfall characteristics on erosion dynamics in low-gradient terrains (Mfoutou and Diabangouaya, 2019; Chabokpour and Azamathulla, 2025). By analyzing field data and applying erosion modeling techniques, the authors assess sediment transport risks and quantify the runoff potential. Their contribution is particularly relevant for developing conservation strategies, land use planning, and watershed management policies in Mediterranean environments prone to soil degradation and water resource challenges (Niang et al., 2015; Zerkaoui et al., 2016; Morsli et al., 2017; Jelisavka and Goran, 2018; Pang and Tan, 2023).

In light of the demonstrated potential of intelligent and hybrid machine learning models for hydrological forecasting, the present study aims to rigorously evaluate and compare the performance of four modeling approaches, namely, standalone Artificial Neural Networks (ANN), standalone Support Vector Machines (SVM), Wavelet-ANN (W-ANN), and Wavelet-SVM (W-SVM), in the autocorrelation-based simulation of monthly runoff. Focusing on the Doostbighlou and Samiyan Rivers in Iran's Qarasu watershed, the research seeks to determine the extent to which wavelet-based hybridization enhances the predictive accuracy of machine learning models in data-scarce, agriculturally significant regions. This study addresses a critical methodological gap and contributes to the advancement of runoff simulation techniques under limited data conditions.

MATERIAL AND METHODS

Location of the Study Site

The study site was the hydrometric stations of Doostbighlou and Samiyan rivers located in the Qarasu watershed of Iran. The Qarasu basin is a sub-basin of the Aras basin, located in the central part of Ardabil province, and a small part of it is located in East Azerbaijan Province. This basin is located at the geographical coordinates of 47° 32′ to 48°41′ east longitude and 37° 47′ to 38° 52′ north latitude, and its area is 7706 km^2. The difference in altitude between the lowest point of the basin (774 m) and its highest point at the peak of Mount Sabalan (4786 m) is 4012 meters. The climate of the basin is in the Mediterranean and semi-arid range. About 42% of the basin's precipitation occurs in the spring. The lowest seasonal precipitation occurs in the summer with 10%. The flow regime, depending on the amount and type of rainfall, increases in frequency from the middle of the water year, i.e., late March to mid-spring, and then decreases until late summer. The reason for this is due to the melting of winter snow and heavy spring rainfall compared to low summer rainfall, as well as human exploitation of river water. The extensive network of surface and groundwater in the basin has made this basin an important place in terms of providing water resources for Ardabil province (Esfandiari and Qarachoorloo, 2023). In this study, monthly runoff data from Doostbighlou and Samiyan hydrometric stations during the statistical years 2000 to 2023 were used to simulate monthly runoff in an autocorrelated manner using intelligent and intelligent-hybrid models. The locations of the aforementioned stations in the Qarasu basin are presented in Fig. 1.

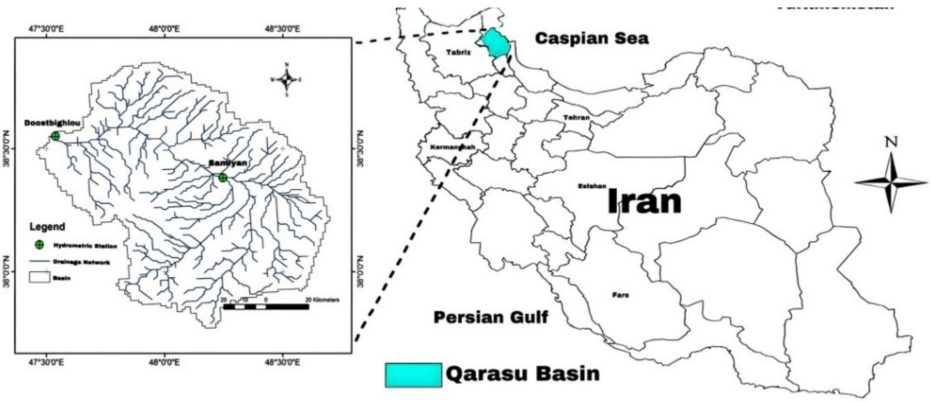


Figure 1: Location of study stations in the Qarasu watershed

Models used in the study

Support Vector Machine

In this model, the optimal separating plane in the nonlinear case will be in the form of Eq. (1).

$$f(x) = \varphi(x)^T w + w_0 \tag{1}$$

The dual Lagrange function is also in the form of the following Eq. (2).

$$L_d = \sum_{i=1}^{n} a_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j y_i y_j \varphi(x_i)\varphi(x_j) \tag{2}$$

In the above equations, $\varphi(x)$ is the kernel function. Radial Basis Function (RBF), polynomial function of degree d and sigmoid function (perceptron) are three kernel functions that are commonly used in SVM (Kavzoglu and Colkesen 2009). In order to estimate the flow using the support vector machine model, various types of kernel functions can be examined, in fact, choosing the appropriate function in using this model is very important and will bring different results. In runoff and rainfall-runoff simulation studies, the RBF kernel function model is mainly used (Eskandari and Nouri, 2010). The calculation process of this model has been done by coding in the MATLAB environment. The formulas of the radial basis function, the polynomial function of degree d, and the sigmoid function (perceptron) are presented in Eqs. (3) to (5), respectively. RBF parameters were logged for each station to ensure reproducibility.

$$K(x, x_i) = \exp\left(-\|x-x_i\|^2/\sigma^2\right) \tag{3}$$

$$K(x, x_i) = \left[t+(x, x_i)\right]^d \tag{4}$$

$$K(x, x_i) = \tanh(k_1(x, x_i) + k_2 \tag{5}$$

Artificial Neural Network

Since normalized data should be used to increase the efficiency and speed of implementation of artificial neural networks (Zare Abyaneh et al., 2010), first the runoff data was placed in the range [-1,1] using Eq. (6), and normalized.

$$x_n = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{6}$$

To simulate the flow, an artificial neural network model of the type of multilayer perceptron network with one hidden layer with a large number of neurons was used, and the sigmoid tangent function was used to map information from the input layer to the hidden layer and the linear stimulus function was used to map information from the hidden layer to the output layer. The calibration of the multilayer perceptron networks was performed with the Levenberg-Marquardt error backpropagation training algorithm (fast and accurate convergence) and the maximum number of iterations in the network learning process was considered to be 1,000. The number of neurons in the hidden layer was determined by trial and error. The process started with a small number of neurons and continued by adding neurons up to a maximum of 20 neurons. Data split into 70% training, 15% validation, and 15% testing, repeated 5 times to assess stability.

Wavelet Analysis

A wave is defined as an oscillating function. A sinusoidal function is a wave. For these functions, Fourier analysis is used. In fact, Fourier analysis is a wave analysis in which functions or signals are expanded in terms of sine and cosine functions (Misiti, et al., 1996). The analyzed function is considered a wavelet function if the following dual conditions hold.

The wavelet has finite energy as follows:

$$E = \int_{-\infty}^{+\infty} \left|\varphi(t)\right|^2 dt \langle \infty \tag{7}$$

If $\varphi(f)$ is the Fourier transform of $\phi(t)$, then Eq. (8) must hold.

$$\int_{0}^{+\infty} \frac{|\varphi(f)|^2}{f} df \langle \infty \qquad (8)$$

Hybrid Wavelet - Support Vector Machine

In the hybrid wavelet and support vector machine model, after converting the data series into several subseries using wavelet analysis and selecting the Daubechies type 4 mother wavelet as the most widely used mother wavelet, the obtained subseries were considered as input to the support vector machine (Taie Semiromi et al., 2024). Considering the number of data, the series was analyzed at two levels and the operation continued with radial, polynomial and linear basis kernel functions.

Hybrid Wavelet - Artificial Neural Network

In this method, the input time series of training and testing data were analyzed into subseries using wavelet analysis. The first step of the series analysis is the selection of the mother wavelet, which was used due to the widespread use of the Daubechies type 4 wavelet. The number of analysis levels is usually considered to be an integer part of the logarithm of the series length (Mushtaq et al., 2024). After determining the subseries created by wavelet analysis as the input of the artificial neural network, data normalization was performed. To determine the optimal number of neurons in the hidden layer of the artificial neural network, the efficiency of 1 to 20 neurons in the hidden layer was evaluated. In this method, multilayer perceptron networks were trained using the Levenberg–Marquardt error backpropagation training algorithm with a maximum number of iterations in the network learning operation of 1,000. All calculations in this research were performed through coding and in the MATLAB software environment.

Time delay of inflow discharges

For runoff modeling, the main variable of discharge and its time delays were used as input to the model. To determine the acceptable delay, the correlation coefficient assumption test was used as follows.

The correlation is not significant: $H_0 : \rho = 0$

The correlation is significant: $H_1 : \rho \neq 0$

Which ρ is the correlation coefficient for the population. In this regard, the coefficient k according to Eq. (9) was used.

$$k = r\left(\frac{1-r^2}{n-2}\right)^{-1/2} \qquad (9)$$

In this equation, r is the correlation coefficient for the sample and n is the amount of data. If the absolute value of k calculated is greater than the value of k in the table related to the 95% confidence level with n-2 degrees of freedom, then the correlation coefficient will be significant. The value of k related to the 95% confidence level in the probability table is +1.98 and -1.98, so if Eq. (10) is established in each series, the correlation will be significant.

$$\left|k_{cal}\right| \geq 1.98 \tag{10}$$

Therefore, the correlation coefficient between runoff at time t and runoff with time lags of one month, two months, three months, etc. is calculated. The corresponding k value of each is compared with Eq. (10) to determine its significance. If the r of runoff with each of the considered lags is greater than 0.2 or smaller than -0.2, it indicates that the correlation is significant and can be used as an input variable.

Model Performance Evaluation Criteria

The runoff values obtained from four models of support vector machine, artificial neural network, combination of wavelet with artificial neural network, and combination of wavelet with support vector machine were evaluated under the statistical indices RMSE (root mean square error), MAE (mean absolute error), R (correlation coefficient), and also the t-test with observed values (Erich et. Al., 2022).

$$R = \frac{\Sigma_{i=1}^{N}(Q_{xi} - \overline{Q}_{x})(Q_{yi} - \overline{Q}_{y})}{\sqrt{\Sigma_{i=1}^{N}(Q_{xi} - \overline{Q}_{x})^{2}}\sqrt{\Sigma_{i=1}^{N}(Q_{yi} - \overline{Q}_{y})^{2}}} \tag{11}$$

$$RMSE = \left(\sum_{i=1}^{n}\left(Q_{y_i} - Q_{x_i}\right)^{2} / n\right)^{1/2} \tag{12}$$

$$MAE = \left(\sum_{i=1}^{n}\left|Q_{y_i} - Q_{x_i}\right|/n\right) \tag{13}$$

In these relations, Q_{xi} and Q_{yi} are the observed and calculated values of discharge at i-th time step, n is the number of data, $\overline{Q_x}$ and $\overline{Q_y}$ are the average of the observed and calculated values of discharge, respectively. The smallness of the RMSE and MAE indices will indicate the high accuracy of the model.

Flowchart: Hybrid Modeling Workflow for Runoff Simulation

The modeling workflow is summarized in the flowchart below.

```
START
 │
 ▼
[Data Preparation]
 ├─ Normalize data (Eq. 6: [-1,1] scaling)
 └─ Select time lags (Eq. 9-10: 1–3 months via correlation test)
 │
 ▼
[Wavelet Decomposition]
 ├─ Mother wavelet: Daubechies-4 (db4)
 ├─ Decompose into subseries:
 │   ├─ Approximation (low-frequency)
 │   └─ Detail (high-frequency)
 └─ 2-level decomposition
 │
 ▼
[Hybrid Model Construction]──────────────────────────────┐
 │                         │
 ├─ [Wavelet-ANN]                   ├─ [Wavelet-SVM]
 │   ├─ Input: Wavelet subseries         │   ├─ Input: Wavelet subseries
 │   ├─ Architecture: MLP (1 hidden layer)   │   ├─ Kernel: RBF/Polynomial/Sigmoid
 │   ├─ Activation: Tangent sigmoid        │   └─ Optimize kernel parameters
 │   ├─ Training: Levenberg-Marquardt       │
 │   └─ Optimize neurons (1–20)          │
 │                         │
 ▼                         ▼
[Model Training & Validation]          [Model Training & Validation]
 ├─ 70% training, 30% testing            ├─ 70% training, 30% testing
 └─ Iterations: 1000                └─ MATLAB implementation
 │
 ▼
[Evaluation Metrics]
 ├─ Statistical indices:
 │   ├─ RMSE (Eq. 11)
 │   ├─ MAE (Eq. 12)
 │   └─ R (Eq. 13)
 └─ Hypothesis testing:
    ├─ t-test ($H_0$: $\mu_1=\mu_2$ at $\alpha=0.05$)
    └─ Compare with standalone ANN/SVM
 │
 ▼
[Calibration]
 ├─ Apply calibration coefficients
 └─ Recalculate RMSE/MAE (Figs. 10–11)
 ▼
END
```

RESULTS AND DISCUSSION

Determining acceptable runoff delays

The input combination of artificial neural network, support vector machine, hybrid wavelet- artificial neural network and hybrid wavelet-support vector machine models was performed by considering the number of delays studied in the discharge at both stations. The results obtained are presented in Figs. 1 and 2.

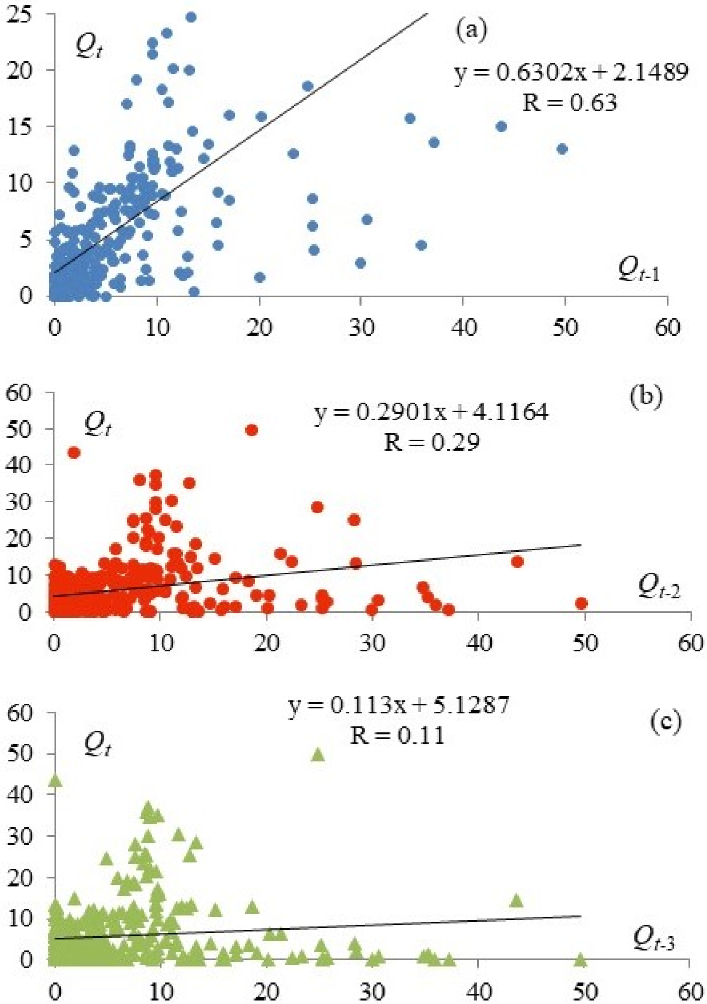


Figure 2: Comparison of current runoff with one-to-three-month lag times at Doostbighlou hydrometric station. (a) Runoff with a one-month delay, (b) Runoff with a two-month delay, (c) Runoff with a three-month delay.

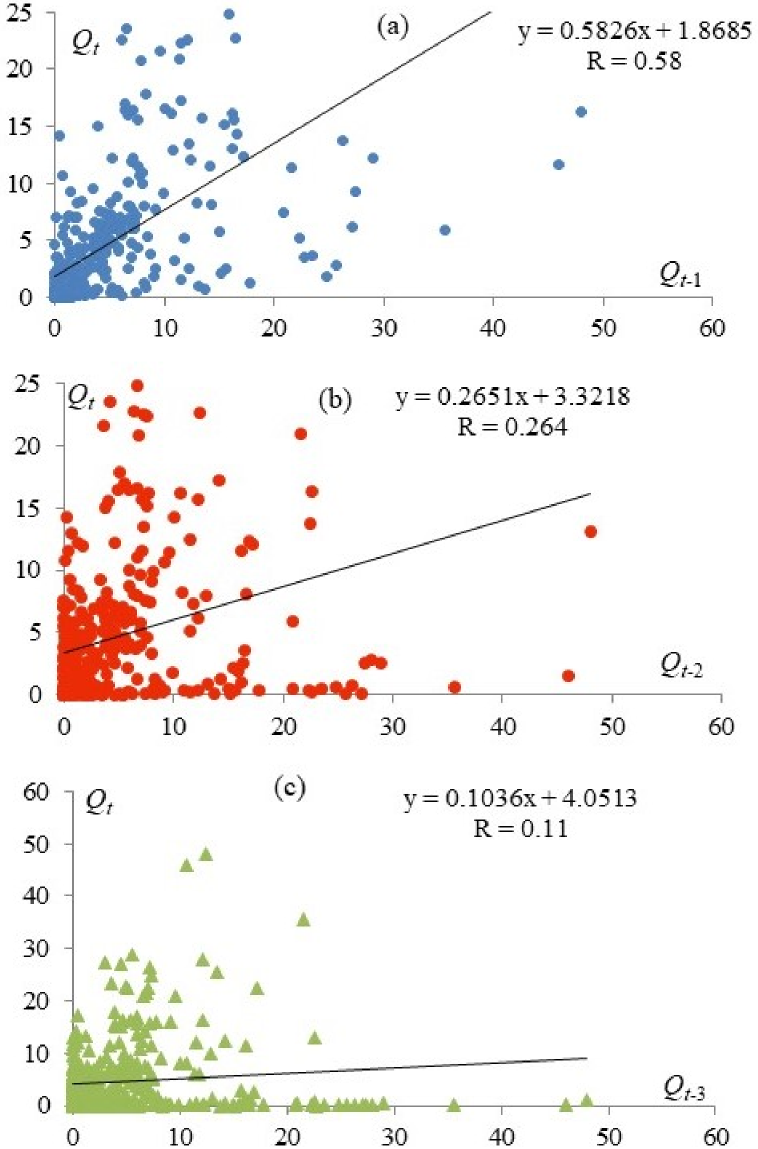


Figure 3: Comparison of current runoff with one-to-three-month lag times at Samiyan hydrometric station. (a) Runoff with a one-month delay, (b) Runoff with a two-month delay, (c) Runoff with a three-month delay

Review of the performance and efficiency of the models

After modeling, runoff prediction was performed with all four methods of artificial neural network and combination of wavelet transform - artificial neural network, support vector machine and combination of wavelet - support vector machine and the results were evaluated with observational data. In Table 1, the R and RMSE values of the artificial neural network and hybrid wavelet transform - artificial neural network models are presented for each of the neurons in the validation stage.

Table 1: Values of statistical indices R and RMSE of ANN and WAV - ANN models

neuron	Doostbighlou				Samiyan			
	WAV - ANN		ANN		WAV - ANN		ANN	
	R	RMSE	R	RMSE	R	RMSE	R	RMSE
1	0.73	1.60	0.42	1.72	0.83	1.90	0.45	2.09
2	0.73	1.59	0.38	1.54	0.81	1.29	0.50	2.08
3	0.62	1.07	0.35	1.48	0.84	1.14	0.46	2.49
4	0.66	1.09	0.42	1.68	0.82	0.97	0.47	2.02
5	0.71	1.13	0.43	1.45	0.81	1.34	0.44	1.88
6	0.69	1.47	0.40	4.77	0.81	1.77	0.46	2.16
7	0.62	1.22	0.41	1.73	0.80	0.98	0.49	2.38
8	0.77	0.97	0.32	1.69	0.76	1.27	0.38	2.52
9	0.74	1.57	0.34	1.70	0.80	1.18	0.48	1.93
10	0.77	1.26	0.41	2.07	0.77	1.28	0.50	2.04
11	0.71	1.32	0.39	1.18	0.74	1.09	0.36	2.08
12	0.70	2.59	0.35	1.70	0.78	1.20	0.45	3.42
13	0.61	1.09	0.34	1.58	0.81	1.06	0.37	1.97
14	0.63	1.49	0.33	1.98	0.71	1.58	0.49	1.97
15	0.69	1.29	0.36	5.36	0.65	1.45	0.35	2.43
16	0.69	1.13	0.40	2.02	0.77	0.99	0.43	2.25
17	0.67	1.30	0.42	1.90	0.81	0.94	0.44	2.37
18	0.71	1.27	0.21	5.15	0.74	1.23	0.42	2.25
19	0.69	0.96	0.39	1.40	0.78	1.40	0.48	1.89
20	0.65	1.08	0.40	1.97	0.78	1.19	0.44	2.60

According to Table 1, the following results can be expressed:

Doostbighlou Hydrometric Station

In using the artificial neural network model, the optimal state is in using 11 neurons; because in this case, the minimum RMSE and maximum R are obtained. In using the hybrid wavelet-artificial neural network model, the minimum RMSE and maximum R are obtained in using 8 neurons, therefore, the corresponding state was considered in the calculations. The hybrid wavelet-artificial neural network model had higher accuracy in simulating the runoff of Doostbighlou Station compared to the single artificial neural network model.

Samiyan Hydrometric Station

Considering that in the artificial neural network model in the Samiyan Hydrometric Station, the minimum RMSE and maximum R are obtained in using 19 neurons; the corresponding values were used in the calculations. In using the hybrid wavelet-artificial neural network model, the optimal state is in using 4 neurons. In using the support vector machine and hybrid wavelet-support vector machine models, considering that in both Doostbighlou and Samiyan hydrometric stations, runoff with one and two months of delay shows a significant correlation, accordingly, runoff modeling with the above methods with two months of delay was considered as the input of the models and the optimal modeling process was carried out. After obtaining the results, the comparison between the observed and simulated runoff values is presented in Figs. 4 to 7 for Doostbighlou and Samiyan stations, respectively.

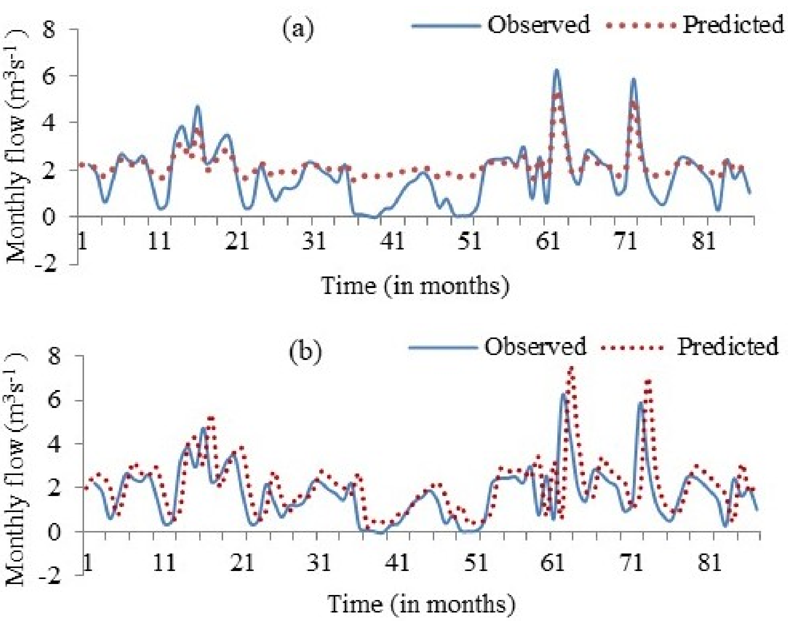


Figure 4: Comparison of observed and simulated runoff using Artificial Neural Network (ANN) and Support Vector Machine (SVM) models at the Doostbighlou Hydrometric Station, incorporating a two-month time lag. (a) ANN, (b) SVM

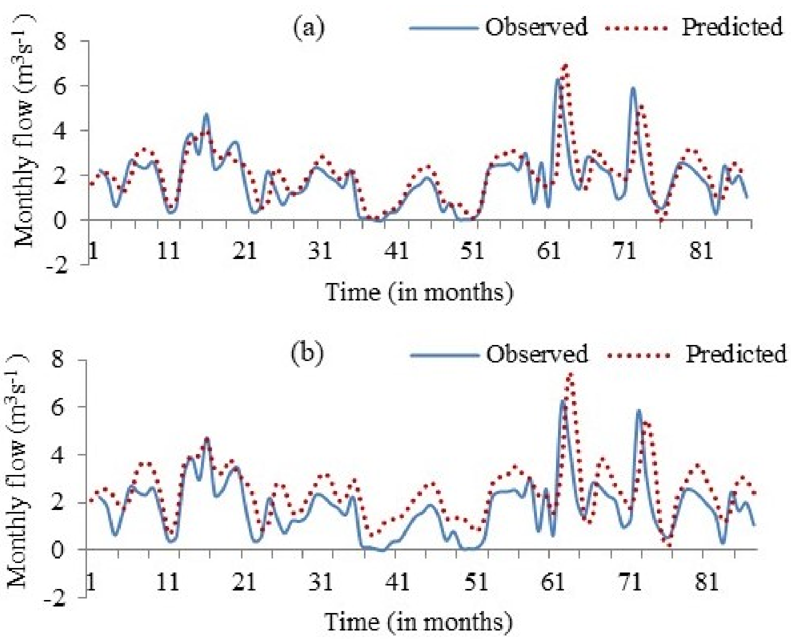


Figure 5: Comparison of Observed and Simulated Runoff Using Wavelet-Artificial Neural Network (WAV-ANN) and Wavelet-Support Vector Machine (WAV-SVM) Models at Doostbighlou Hydrometric Station, Considering a Two-Month Time Lag.". (a) WAV-ANN, (b) WAV-SVM

As can be seen from the graphs in Figs. 4 and 5 for the Doostbighlou hydrometric Station, the agreement between the observed runoff values and the simulated runoff values using the hybrid wavelet-artificial neural network and wavelet-support vector machine is very high compared to single models of artificial neural networks and support vector machines. This indicates the effectiveness of considering the wavelet transform in runoff prediction. Also, at this station, the agreement and consistency between the observed and simulated runoff obtained from the hybrid wavelet-artificial neural network model was better than the hybrid-support vector machine model.

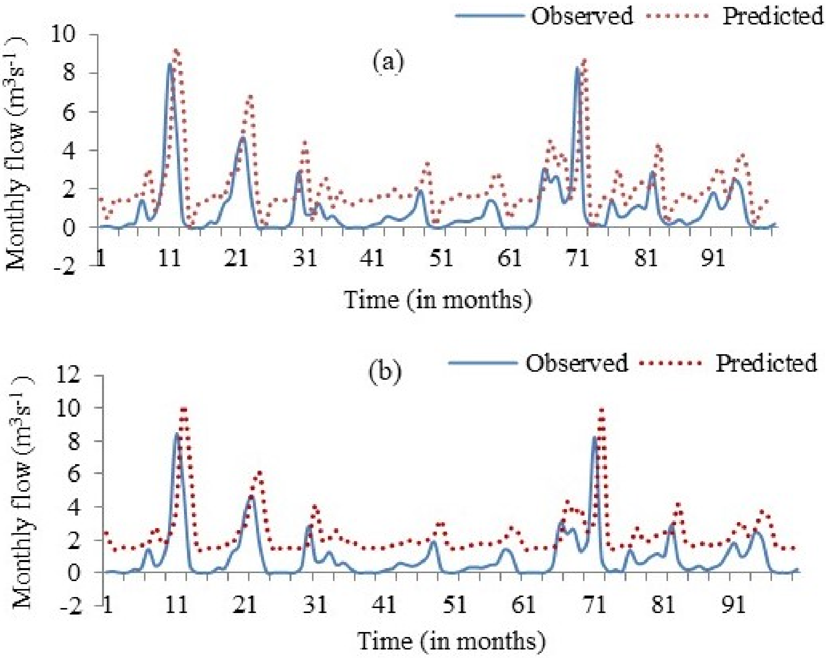
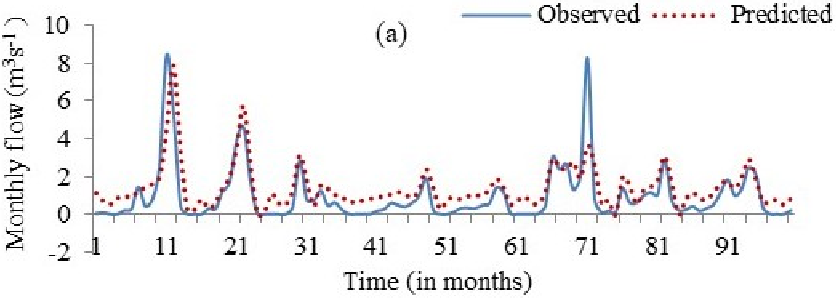


Figure 6: Comparison of observed and simulated runoff using ANN and SVM models at the Samiyan Hydrometric Station, incorporating a two-month time lag. (a) ANN, (b) SVM



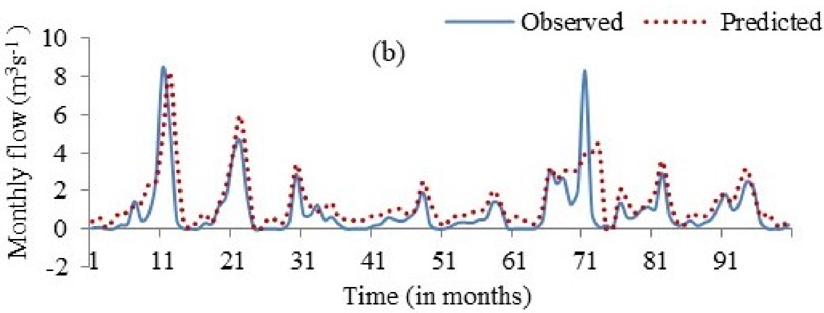


Figure 7: Comparison of Observed and Simulated Runoff Using WAV-ANN and WAV-SVM Models at Samiyan Hydrometric Station, Considering a Two-Month Time Lag. (a) WAV-ANN, (b) WAV-SVM

After simulating the runoff of both stations with the four studied methods, the models were calibrated and the results were evaluated using statistical indicators and t-test. The evaluation results for all models are presented in Table 2.

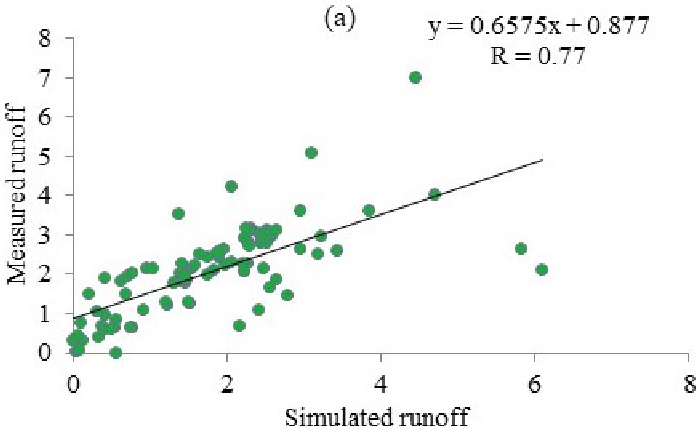
Table 2: Index values and statistical tests of the models used in Doostbighlou and Samiyan hydrometric stations

Method	RMSE m^3s^{-1}	MAE m^3s^{-1}	R	Variance	Mean difference	*t* statistic
			ANN			
Doostbighlou	1.18	0.95	0.39	1.28	0.41	-2.72
Samiyan	1.96	1.43	0.48	2.42	-1.62	-5.05
Mean	1.57	1.19	0.435	1.85	-0.605	
			SVM			
Doostbighlou	1.38	0.97	0.45	1.71	0.36	-2.84
Samiyan	2.15	1.71	0.45	2.15	-1.45	-6.28
Mean	1.765	1.34	0.45	1.93	-0.545	
			WAV ANN			
Doostbighlou	0.97	0.70	0.77	1.32	0.26	-1.44^{*}
Samiyan	1.02	0.73	0.82	1.37	-0.46	-1.56^{*}
Mean	0.995	0.715	0.785	1.345	-0.1	
			WAV SVM			
Doostbighlou	1.17	0.98	0.72	1.38	0.73	-1.93^{*}
Samiyan	1.12	0.70	0.78	1.97	-0.47	-1.83^{*}
Mean	1.145	0.84	0.75	1.675	0.13	

In Table 2, the * sign above some numbers indicates that at a significance level of 5%, the null hypothesis (H_0: $\mu_1=\mu_2$) is confirmed. For statistical comparison, the obtained *t*-statistics were evaluated using the critical value of 1.97 (*t*-critical) from the reference Table. From the aforementioned table, which presents the results of comparing observational and simulated runoff data at the two stations of Doostbighlou and Samiyan with four models: artificial neural network, support vector machine, wavelet-artificial neural network combination, and wavelet-support vector machine combination under the statistical indices RMSE, MAE, and R, as well as the t-test, the following can be noted.

The results of the t-test analysis at the two hydrometric stations of Samiyan and Doostbighlou showed that at a significance level of 5%, the first hypothesis (H_1: $\mu_1 \neq \mu_2$) was rejected only in the hybrid wavelet-artificial neural network and wavelet-support vector machine models and the null hypothesis (H_0: $\mu_1 = \mu_2$) was valid. Also, at a significance level of 5%, the first hypothesis (H_1: $\mu_1 \neq \mu_2$) was valid in the artificial neural network and support vector machine models and the null hypothesis (H_0: $\mu_1 = \mu_2$) was rejected. At Doostbighlou station, the combined wavelet-artificial neural network model had the highest R value (correlation coefficient) with a value equal to 0.77 among all the models used. At Samiyan station, the combined wavelet-artificial neural network model had the highest R value, with a value equal to about 0.82 among all the models used. Hybridizing the single artificial neural network model with the wavelet model at Doostbighlou station has increased and also decreased the R and RMSE parameters from 0.39 and 1.18 m^3s^{-1} to 0.77 and 0.97 m^3s^{-1}, respectively. Hybridizing the single artificial neural network model with the wavelet model at Samiyan station has increased and also decreased the R and RMSE parameters from 0.48 and 1.96 m^3s^{-1} to 0.82 and 1.02 m^3s^{-1}, respectively. Hybridizing the single support vector machine model with the wavelet model at Doostbighlou station has increased and decreased the R and RMSE parameters from 0.45 and 1.38 m^3s^{-1} to 0.72 and 1.17 m^3s^{-1}, respectively. Hybridizing the single support vector machine model with the wavelet model at Samiyan station has increased and decreased the R and RMSE parameters from 0.45 and 2.15 m^3s^{-1} to 0.78 and 1.12 m^3s^{-1}, respectively. Considering the acceptability of the wavelet-artificial neural network and wavelet-support vector machine models in the runoff simulation of both Doostbighlou and Samiyan stations compared to the single models, the calibration coefficients of the aforementioned models are presented in graphs in Figs. 8 and 9.

The graphs in Fig. 8 show the high correlation and agreement between the observed runoff values and the values obtained from the WAV-ANN and WAV-SVM models at Doostdighlou Hydrometric Station. As in this station, the correlation coefficient in the WAV-ANN model (Graph a) is equal to 0.77.



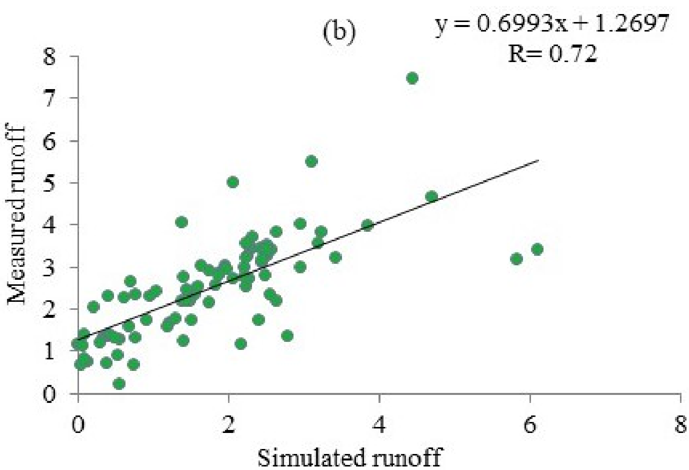


Figure 8: Calibration coefficients of Doostbighlou hydrometric station. (a) WAV-ANN, (b) WAV-SVM

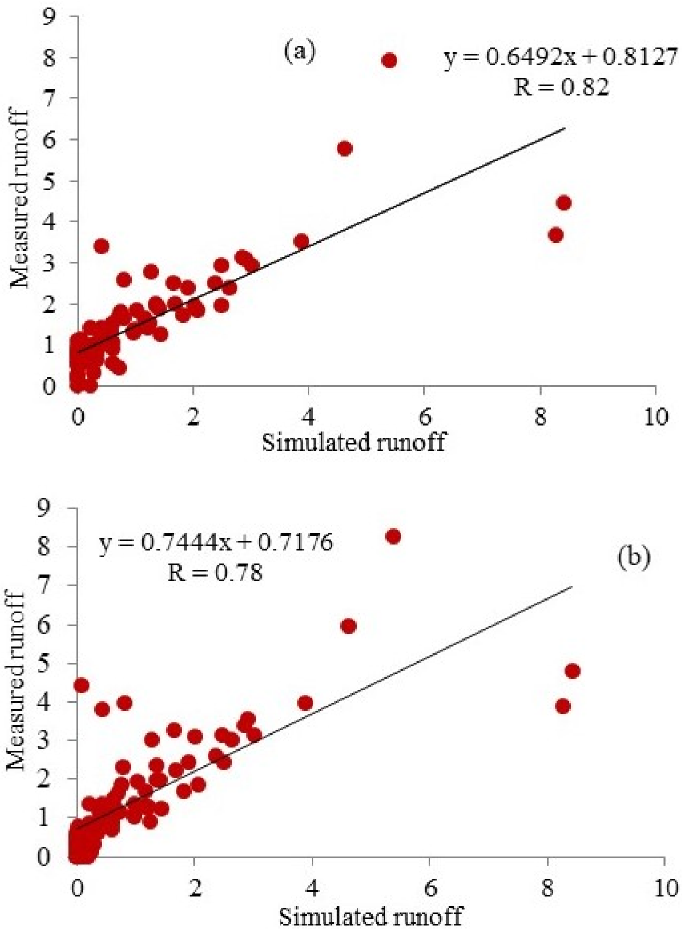


Figure 9: Calibration coefficients of Samiyan hydrometric station. (a) WAV-ANN, (b) WAV-SVM

The graphs (a) and (b) in Fig. 9 with correlation values of 0.82 and 0.78, which correspond to the WAV-ANN and WAV-SVM models, respectively, indicate the high agreement of these models in the runoff simulation of Samiyan station. Dehghani et al., (2022) also found similar results in their acceptance compared to the use of single intelligent models in their study of the performance of hybrid intelligent models with wavelet transforms that they used to estimate the runoff of the Dez River Basin in Iran.

Error indicators of RMSE and MAE after calibration

The indices RMSE and MAE were recalculated after applying the calibration coefficients to the values obtained from the model simulations. Figs. 10 and 11 present a comparison between the indices RMSE and MAE before and after applying the calibration coefficients.

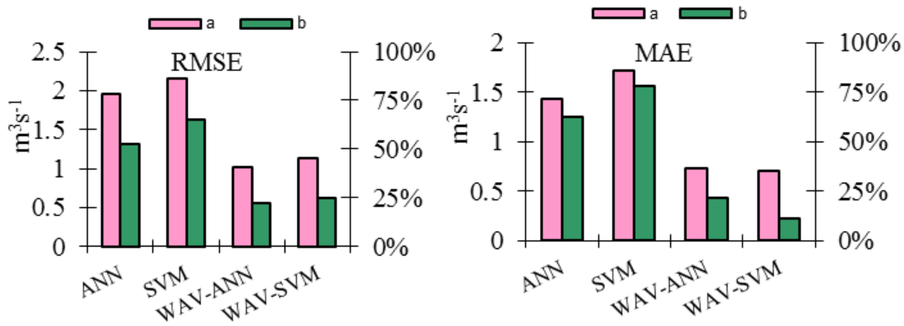


Figure 10: Comparison of statistical indices of RMSE and MAE at Samiyan Hydrometric Station before and after applying the fitting coefficients. (a) Before applying the fitting coefficients, (b) After applying the fitting coefficients.

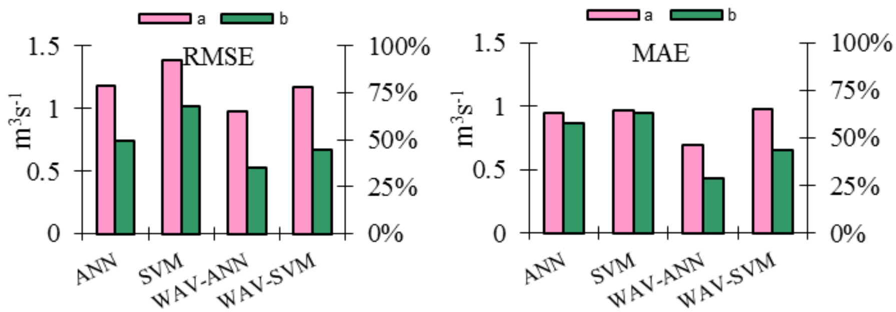


Figure 11: Comparison of statistical indices of RMSE and MAE at Doostbighlou Hydrometric Station before and after applying the fitting coefficients. (a) Before applying the fitting coefficients, (b) After applying the fitting coefficients

As is clear from the graphs in Figs. 10 and 11, applying calibration coefficients to the simulated values has significantly reduced the error indicators.

CONCLUSION

In this study, the performance and efficiency of intelligent models of artificial neural network, support vector machine and intelligent-hybrid models of wavelet-artificial neural network and wavelet-support vector machine were evaluated in the simulation of the next month's runoff of Doostbighlou and Samiyan hydrometric stations. In this regard, for each of the stations, the appropriate combination of model inputs (runoff with an acceptable number of delays) was determined and the flow was simulated in an autocorrelated manner, in which the following results can be expressed.

After considering the wavelet model in the individual models of artificial neural network and support vector machine in the runoff simulation, the values of the error indices were significantly reduced, indicating its positive effect on improving and increasing the efficiency of the aforementioned models in both Doostbighlou and Samiyan hydrometric stations. The highest agreement in runoff simulation with a correlation coefficient of 0.82 was related to Samiyan hydrometric station using the hybrid wavelet-artificial neural network model. In general, the RMSE error index of the support vector machine model in runoff simulation for both stations was high compared to the artificial neural network model. After the hybrid wavelet-artificial neural network model, the hybrid wavelet-support vector machine model had better accuracy and efficiency than other models used in both stations and can be recommended for predicting runoff in the aforementioned rivers. This study demonstrates the effectiveness of hybrid wavelet-based machine learning models (WAV-ANN and WAV-SVM) for simulating autocorrelated runoff in data-poor basins. However, several promising research directions can further improve hydrological modeling in similar regions. Including the integration of meteorological and climate variables into meteorological inputs: Future studies could consider precipitation, temperature, evapotranspiration, and snowmelt data to assess their impact on runoff prediction accuracy. Transfer learning: Regional and inter-basin generalization, applying pre-trained models to neighboring basins with limited data. Comparative studies: Evaluate model performance in different climatic and topographic regions. Impact of land use and climate change on LULC dynamics: Evaluate how land cover changes affect runoff patterns using GIS and remote sensing.

Declaration of competing interest

The author declares that he has no known competing financial interests or personal relationships that could influence the work reported in this article.

REFERENCES

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, Versita, Vol. 45, Issue 5, pp. 71-83.

ANDALIB G., NOURANI V., MONIRIFAR H., SHARGHI E. (2020). Development of hybrid wavelet-artificial neural network and support vector machine approach for multi-station rainfall-runoff modeling using clustering and mutual information tools, New approaches in civil engineering, Vol. 3, Issue 2, pp. 49-62.

ARGAZ A. (2018). 1D model application for integrated water resources planning and evaluation: case study of Souss river basin, Morocco, Larhyss Journal, No 36, pp. 217-229.

ASSEMIAN A.E., DJE BI DOUTIN S., SAMAKÉ Y. (2021). Consequences of the effects of climate change on water resources in a humic tropical zone of central eastern Côte d'Ivoire, Larhyss Journal, No 45, pp. 95-105.

ATALLAH M., DJELLOULI F., HAZZEB A. (2024). Rainfall-runoff modeling using the HEC-HMS model for the Mekerra wadi watershed (N-W Algeria), Larhyss Journal, No 57, pp. 187-208.

BAUDHANWALA D., KANTHARIA V., PATEL D., MEHTA D., WAIKHOM S. (2023). Applicability of SWMM for urban flood forecasting a case study of the western zone of Surat city, Larhyss Journal, No 54, pp. 71-83.

BEN SAID M., HAFNAOUI M.A., HACHEMI A., MADI M., BENMALEK A. (2024). Evaluating the effectiveness of the existing flood risk protection measures along wadi Deffa in El-Bayadh city, Algeria, Larhyss Journal, No 59, pp. 7-28.

BENSLIMANE M., BERREKSI A., BENMAMAR S., BOUACH A. (2020). Flood risk numerical simulation of Bejaia city urban zone (Algeria), Larhyss Journal, No 42, pp. 167-178.

CANNAS B., FANNI A., SEE L., SIAS G. (2006). Data preprocessing for river flow forecasting using neural network: wavelet transforms and data partitioning, Physics and Chemistry of the Earth, Vol. 31, Issue 18, pp. 1164-1171.

CHABOKPOUR J., AZAMATHULLA H.M.D. (2025). Multi-scale CFD analysis of erosion dynamics in heterogeneous rockfill structures implications for sustainable hydraulic engineering design, Larhyss Journal, No 61, pp. 217-239.

CHADEE A., NARRA M., MEHTA D., ANDREW J., AZAMATHULLA H. (2023). Impact of climate change on water resource engineering in Trinidad and Tobago, Larhyss Journal, No 55, pp. 215-229.

CHIBANE B., ALI-RAHMANI S.E. (2015). Hydrological based model to estimate groundwater recharge, real- evapotranspiration and runoff in semi-arid area, Larhyss Journal, No 23, pp. 231-242.

DEHGHANI R., TORABIPOODEH H., YOUNESI H., SHAHINJAD B. (2022). Application of hybrid support vector machine wavelet (WSVM) model in predicting river flow, case study: Dez watershed, Watershed Engineering and Management, Vol. 13, Issue 1, pp. 98-110.

DO T.V.H., PHAM H.G., KIEU Q.L., TRAN T.N.H. (2025). The typical mechanisms and factors leading to flash floods in small watersheds in the mountainous region of Vietnam, a case study in the CHU VA stream watershed, Larhyss Journal, No 61, pp. 141-168.

ERICH C.F., GILMOUR J., MACHIN T., HENDRY L. (2022). Statistics for Research Students, 1st edition, University of Southern Queensland, Toowoomba, 101p.

ESFANDIYARI F., QARACHOORLOO M. (2023). Regional estimation of sediment yield using monthly sediment rating curve in Qarasu watershed, Journal of Applied Research in Geographical Sciences, Vol. 23, Issue 69, pp. 33-52.

ESKANDARI A., NOURI R. (2010). Developing a Proper Model for Online Estimation of the 5-Day Biochemical Oxygen Demand Based on Artificial Neural Network and Support Vector Machine, Ecology, Vol. 38, Issue 61, pp. 76-84.

EZZ H. (2025). Unexpected flooding in Mersa Matruh, Egypt - Investigating causes, hydrological analysis, and flood risk assessment, Larhyss Journal, No 61, pp. 371-399.

FARAJZADEH J., FAKHERI FARD A., LOTFI S. (2014). Modeling of monthly rainfall and runoff of Urmia Lake basin using feed-forward neural network and time series analysis model, Water Resources and Industry, Vol. 7, Issue 8, pp. 38-48.

http://cda.psych.uiuc.edu/matlab_pdf/wavelet_ug.pdf

FAREGH W., BENKHALED A. (2016). GIS based SCS-CN method for estimating runoff in Sigus watershed, Larhyss Journal, No 27, pp. 257-276. (In French)

JELISAVKA B., GORAN R. (2018). Irrigation of agricultural land in Montenegro: overview, Larhyss Journal, No 33, pp. 7-24.

KALTEH A.M. (2013). Monthly River flow forecasting using artificial neural network and support vector regression models coupled with wavelet transform, Computers and Geosciences, Vol. 54, Issue 2, pp. 1-8.

KANTHARIA V., MEHTA D., KUMAR V., SHAIKH M.P., JHA S. (2024). Rainfall–runoff modeling using an Adaptive Neuro-Fuzzy Inference System considering soil moisture for the Damanganga basin, Journal of Water and Climate Change, Vol. 15, Issue 5, pp. 2518-2531.

KANTHARIA V.C., MEHTA D.J. (2024). Integrated rainfall-runoff modelling using fuzzy logic considering soil moisture: case study of Damanganga Basin, International Journal of Hydrology Science and Technology, Vol. 18, Issue 3, pp. 300-318. DOI:10.1504/IJHST.2024.140855

KAVZOGLU T., COLKESEN I. (2009). A kernel functions analysis for support vector machines for land cover classification, International Journal of Applied Earth Observation and Geoinformation, Vol. 11, Issue 5, pp. 352-359.

KIANI ASL M.A., MOTESHAFFEH B., ROSHAAN S.H. (2023). Evaluation of machine earning algorithms (RF and SVM) in producing flood susceptibility mapping Maroon watershed, Iran-Watershed Management Science and Engineering, Vol. 17, Issue 61, pp. 51- 52.

KOUA T., ANOH K., EBLIN S.., KOUASSI K., KOUAME K. (2019). Rainfall and runoff study in climate change context in the buyo lake watershed (southwest Côte d'Ivoire), Larhyss Journal, No 39, pp. 229-258. (In French)

KUMAR V., KEDAM N., SHARMA K.V., MEHTA D., CALOIERO T. (2023). Advanced machine learning techniques to improve hydrological prediction: A comparative analysis of streamflow prediction models, Water, Vol. 15, Issue 14, pp. 1-24.
DOI:10.3390/w15142572

MEHTA D., DHABUWALA J., YADAV S.M., KUMAR V., AZAMATHULLA H.M. (2023). Improving flood forecasting in Narmada River basin using hierarchical clustering and hydrological modelling, Results in Engineering, Vol. 20, Issue 13, pp. 1-13.

MEHTA D., HADVANI J., KANTHARIYA D., SONAWALA P. (2023). Effect of land use land cover change on runoff characteristics using curve number: A GIS and remote sensing approach, International Journal of Hydrology Science and Technology, Vol. 16, Issue 1, pp. 1-16.

MEHTA D., YADAV S. (2024). Rainfall runoff modelling using HEC-HMS model: case study of Purna river basin, Larhyss Journal, Vol. 59, No 59, pp. 101-118.

MFOUTOU W., DIABANGOUAYA D.B. (2019). Bank erosion on the Mfilou river on Brazzaville, Larhyss Journal, No 39, pp. 299-311. (In French)

MISITI M., MISITI Y., OPPENHEIM G., POGGI J.M. (1996). Wavelet Toolbox, The MathWorks, Inc., Version 3, pp. 2-36.

MORSLI B., HABI M., BOUCHEKARA B. (2017). Study of marine intrusion in coastal aquifers and its repercussions on the land degradation by the use of a multidisciplinary approach, Larhyss Journal, No 30, pp. 225-237. (In French)

MUSHTAQ H., AKHTAR T., HASHMI M.Z., MASOOD A., SAEED F. (2024). Hydrologic interpretation of machine learning models for 10-daily stream flow simulation in climate sensitive upper Indus catchments, Theoretical and Applied Climatology, Vol. 155, Issue 6, pp. 1-18.

NIANG D., YACOUBA H., DOTO C.V, KARAMBIRI H., LAHMAR R. (2015). Land degradation impact on water transfer of soil-plant-atmosphere continuum in the Burkinabe Sahel, Larhyss Journal, No 24, pp. 109-127.

PANDA C., PANDA K.C., SINGH R.M., SINGH R., SINGH V.P. (2025). A generalized hydrological model for stream flow prediction using wavelet ensembling, Journal of Hydrology, Vol. 655, Issue 2.

PANG J.M., TAN K.W. (2023). Development of regional climate model (RCM) for Cameron highlands based on representative concentration pathways (RCP) 4.5 and 8.5, Larhyss Journal, No 54, pp. 55-70.

RIAHI R., BELAID H., HATIRA A., BACCOUCHE S. (2020). Contribution to the study of runoff and erosion of low slope homogeneous hydrological units of a watershed of the middle valley of Medjerda Tunisia, Larhyss Journal, No 43, pp. 119-137.

ROUISSAT B., SMAIL N. (2022). Contribution of water resource systems analysis for the dynamics of territorial rebalancing, case of Tafna system, Algeria, Larhyss Journal, No 50, pp. 69-94.

SAMANTARAY S., SAWAN S., SAHOO A., SATAPATHY P.S. (2022). Monthly runoff prediction at Baitarani River basin by support vector machine based on Salp swarm algorithm, Ain Shams Engineering Journal, Vol. 13, Issue 5, pp. 1-18.

TAIE SEMIROMI M., KOCH M. (2024). Statistical downscaling of precipitation in northwestern Iran using a hybrid model of discrete wavelet transform, artificial neural networks, and quantile mapping, Theoretical and Applied Climatology, Vol. 155, Issue 7, pp. 6591-6691.

VERMA S., KUMAR K., VERMA M.K., PRASAD A.D., MEHTA D., RATHNAYAKE U. (2023). Comparative analysis of CMIP5 and CMIP6 in conjunction with the hydrological processes of reservoir catchment, Chhattisgarh, India, Journal of Hydrology: Regional Studies, Vol. 50, pp. 1-42.
DOI: 10.1016/j.ejrh.2023.101533

ZARE ABYANEH H., BAYAT VARKSHI M., AKHAVAN S., MOHAMMADI M. (2010). Estimation of nitrate in Hamedan-Bahar plain groundwater using ANN and the effect of data resolution on prediction accuracy, Journal of Environment Studies, Vol. 37, Issue 58, pp. 129-140.

ZERKAOUI L., BENSLIMANE M., HAMIMED A., KHALDI A. (2016). Land development issues and water use conflict in mountain areas, case of the Beni Chougrane mountains (northwestern Algeria), Larhyss Journal, No 26, pp. 209-223. (In French).