



# Evaluating machine learning models in predicting dam inflow and hydroelectric power production in multi-purpose dams (case study: Mahabad Dam, Iran)

Seyed Mohammad Enayati<sup>1</sup> · Mohsen Najarchi<sup>2</sup> · Osman Mohammadpour<sup>3</sup> · Seyed Mohammad Mirhosseini<sup>1</sup>

Received: 6 March 2024 / Accepted: 28 July 2024 / Published online: 20 August 2024  
© The Author(s) 2024

## Abstract

This study aimed to forecast dam inflows and subsequently predict its capability in producing HEPP using machine learning and evolutionary optimization techniques. Mahabad Dam, located in the northwest of Iran and recognized as one of the nation's key dams, served as a case study. First, artificial neural networks (ANN) and support vector regression (SVR) were employed to predict dam inflows, with optimization of parameters achieved through Harris hawks optimization (HHO), a robust optimization technique. The data of temperature, precipitation, and dam inflow over a 24-year period on a monthly basis, incorporating various lag times, were used to train these machines. Then, HEPP from the dam was predicted using temperature, precipitation, dam inflow, and dam evaporation as input variables. The models were applied to data covering the years 2000 to 2020. The results of the first part indicated both hybrid models (HHO-ANFIS and HHO-SVR) improved the prediction performance compared to the single models. Based on the results of Taylor's diagram and the error evaluation criteria, the HHO-ANFIS hybrid model (RMSE, MAE, and NSE of 3.90, 2.41, and 0.86, respectively) exerted better performance than HHO-SVR (RMSE, MAE, and NSE of 4.39, 2.70, and 0.86, respectively). The results of the second part showed that using the HHO algorithm to optimize single models (RMSE, MAE, and NSE of 0.2, 10, and 0.90, respectively) predicted HEPP better than single models (RMSE, MAE, and NSE of 0.2, 10, and 0.90, respectively). The results of Taylor's diagram also showed that the HHO-ANFIS model exerted better performance. The findings of this study indicated the promising performance of machine learning models optimized by metaheuristic algorithms in the simultaneous prediction of dam inflows and HEPP in multi-purpose dams for better management and allocation of surface water resources.

**Keywords** Machine learning · Optimization algorithms · River flow · Hydropower · Mahabad Dam

## Introduction

Rivers are essential water resources for domestic, industrial, and agricultural use (Tabbusum and Dar 2021; Wu et al. 2005; Ahmed et al. 2014). Various characteristics of rivers and their tributaries add complexity to river basin management (Jimeno-Sáez et al. 2018; Khosravi et al. 2021). Therefore, engineers need to understand the effects of structural measures such as dams and reservoirs as well as non-structural strategies such as disaster prevention plans and flood control methods (Mohammadi 2021; Musarat et al. 2021). A basic flood prevention plan can be used to predict the flow rate of a river at a specific period (Nguyen et al. 2021). Accurate estimation of river flow rate in watersheds and water resources systems plays a fundamental role in the efficient and timely management of water projects (Rahimzad et al. 2021; Samadianfard et al. 2019). Also, it is important

✉ Mohsen Najarchi  
mohsennajarchi@yahoo.com

Seyed Mohammad Enayati  
SM.Enayati@iau.ac.ir

Osman Mohammadpour  
Osman.Mohammadpour@iau.ac.ir

Seyed Mohammad Mirhosseini  
Mo.mirhosseini@iau.ac.ir

<sup>1</sup> Department of Civil Engineering, Arak Branch, Islamic Azad University, Arak, Iran

<sup>2</sup> Department of Water Engineering, Arak Branch, Islamic Azad University, Arak, Iran

<sup>3</sup> Department of Water Engineering, Mahabad Branch, Islamic Azad University, Mahabad, Iran

to have access to information about river flow to achieve optimal performance in river management, design flood warning systems, and make appropriate planning in this regard (Fathian et al. 2019; Poul et al. 2019; Puttinaovarat and Horkaew 2020; Sharma and Goel 2024).

Dams are reliable places to collect runoff so that they can be used to meet water needs in critical situations, prevent floods, and generate hydroelectric energy (Milan et al., 2018). Dams also play a significant role for other purposes such as fish breeding and attracting tourists. Therefore, the construction of multi-purpose dams is placed on the agenda of governments in the water resources management sector. In general, the prediction of reservoir inflow is important for water supply management and planning. However, the input values in the future are uncertain and therefore it is necessary to obtain as much information as possible about the possible changes of this variable in the future. This is crucial for decision-making and optimal exploitation of water-related infrastructure (Noorbeh et al. 2020). Greater performance in predicting reservoir inflows enhances the reliability of water supply from the dam, being particularly crucial during drought periods (Kim et al. 2022). Determining the actual and optimal amounts of water resources, as well as the release of water from reservoirs, provides valuable information to planners to help manage and optimally allocate water resources (Babaei et al. 2019; Meshram et al. 2019; Obahoundje et al. 2024). It is also important to predict the dam inflow in terms of reducing the possible risks of floods and droughts (Turner et al. 2020; Latif and Ahmed 2024).

Although non-renewable energies such as fossil fuels are used in some underdeveloped or developing countries, due to their negative effects on human health and the environment, as well as higher costs than some renewable energies such as hydroelectricity, wind, and solar energies, they are less used nowadays. The increasing demand for electricity led to the rapid application of renewable energies, which aimed to achieve sustainable energy and progress in the economy (Peng et al. 2023). The hydroelectricity

produced by the dams is of great importance because most governments understand the demand for more production of electricity. Hydropower is one of the main sources of renewable energy due to its low cost and near-zero pollution and the potential to quickly respond to high electricity demands (Dmitrieva 2015; Banadkooki et al. 2020; Sahin and Ozbay Karakus, 2024). Aside from mitigating greenhouse gas emissions, hydropower plants play a crucial role in climate change adaptation by managing water resources. Presently, approximately 20% of global electricity is generated by hydropower plants, with nearly 60 countries relying on them for over half of their electricity production (Clerici and Alimonti 2015; Bilgili et al 2015). Accurate modeling of hydroelectric power production (HEPP) facilitates improved planning and management of water and energy resources. This, in turn, enhances resource productivity, boosts the prospects of water and energy sustainability, mitigates environmental challenges, and helps prevent the depletion of natural resources (Zolfaghari and Golabi 2021).

In the field of HEPP, researchers have used machine learning to achieve favorable results (Choubin et al. 2019; Dehghani et al. 2019; Shu et al. 2024). Some of them are summarized in Table 1. The adaptive neuro-fuzzy inference system (ANFIS) is a machine learning model that combines artificial neural networks (ANN) and fuzzy inference systems (FIS) (Esmaili et al. 2021), providing a robust modeling approach (Firat and Güngör, 2007). However, despite its advantages, ANFIS may encounter performance issues due to potential trapping in local minima during training (Hashemi et al. 2014). To address this challenge, researchers have turned to evolutionary optimization algorithms to enhance ANFIS prediction performance (Milan et al. 2021; Kayhomayoon et al. 2021a, b; Seifi and Riahi-Madvar 2019). Among these algorithms, the Harris hawks optimization algorithm (HHO) stands out as particularly effective (Kayhomayoon et al. 2021a, b; Milan et al. 2021). HHO, inspired by swarm intelligence, mimics the collective behavior and hunting mechanisms of Harris hawks (Abbasi et al. 2021; Moayedi et al. 2019).

**Table 1** Selected related works of hydropower production forecasting

Author (year)	Prediction variable	ML models	Time step
Hannon et al. (2023)	Hydroelectric power	ANN, ARIMA, SVR*	Daily, monthly, seasonally
Guo et al. (2011)	Streamflow forecasting	SVR	Monthly
Wang et al. (2011)	Hydroelectric power	SVR	Seasonally
Zhang et al. (2020)	Reservoir inflow prediction	ANN, SVR, ANFIS	Daily
Stefenon et al. (2022)	Hydroelectric power	LSTM, SVR	–
Barzola-Monteses et al. (2022)	Hydropower production	ANN, LSTM	Monthly
Dehghani et al. (2019)	Power production	ANFIS-GWO	Monthly

\*ANN, Artificial Neural Network; ARIMA, autoregressive integrated moving average; SVM, Support vector regression; ANFIS, Adaptive Neuro-fuzzy Inference System; LSTM, Long short-term memory

The HEPP, the allocation of water from the dam for the downstream areas, and flood control depend on the water volume of the dam and its inflow. It is crucial to predict the amount of dam inflow and the amount of hydroelectricity produced from it during different months of the year, especially the hot months of the year when the most electrical energy is used. Knowing the amount of dam inflow can help in better management and more optimal allocation of the dam's water resources. Mahabad Dam is one of the important dams in the northwest of Iran, which is of special importance due to supplying a large part of the drinking and agricultural needs of the Mahabad area, producing hydroelectric energy, and also reviving Lake Urmia. This research aimed to predict the dam inflow using machine learning techniques optimized by evolutionary algorithms. Then, the amount of hydroelectric energy produced by the dam is predicted.

The main novelty of this study involves the simultaneous usage of machine learning models in predicting the inflow and then estimating the HEPP. Also, utilizing recently introduced hybrid models in the estimation of HEPP is among the novelty of this research. It has been tried to use the minimum available data for predictions to enable the implementation of this approach in any range. For this purpose, precipitation and temperature variables have been used to predict the dam inflow. The dam inflow along with precipitation, evaporation, and the outflow from the dam are also used in the form of input patterns to predict HEPP in the dam. The HHO algorithm has been used to improve the prediction performance of ANFIS and SVR models. Therefore, the purpose of this research is to investigate the performance of optimization algorithms in improving machine learning models to achieve a suitable forecasting system for the dam inflow and then the hydroelectric energy produced, which has been addressed in a limited number of previous research.

## Materials and methods

### Study area and data

The Mahabad River catchment basin ( $45^{\circ} 25' - 46^{\circ} 45' E$ ,  $36^{\circ} 23' - 36^{\circ} 45' N$ ) lies to the south of Lake Urmia, covering an area of  $1524.53 \text{ km}^2$ , which represents approximately 3% of Lake Urmia's catchment area. The Mahabad River originates from the convergence of the Bitas tributary in the east and the Kotar tributary in the west (Fig. 1). In 1967, the Mahabad reservoir dam was constructed at the confluence of these tributaries, alongside another dam downstream aimed at diverting water for irrigation and flood control to prevent downstream flooding events. The Mahabad Dam stands as one of the country's ten largest water dams, boasting a structure height of 47.5 m and a length of 700 m, built on a pebble and clay bed. Annually, the Mahabad dam receives an inflow

of 339 million cubic meters (MCM) of water. Adjacent to Mahabad city, the dam forms a permanent wetland utilized by the Regional Water Company to provide drinking and agricultural water to the city and surrounding villages. The lake created by the dam spans an area of 360 ha (Enayati et al. 2022; Nematollahi and Sanayei 2023).

A dam is constructed with different purposes, one of the most important of which is the production of hydroelectric power. Hydropower, a form of renewable energy, harnesses the energy from water stored in dams or flowing in rivers to generate electricity. The water flow entering the turbine rotates its blades and causes the rotation of the generator, followed by the production of electrical energy (Fig. 2).

The amount of outflow and the energy produced from the dam relies on the amount of dam inflow. The long-term monthly average (2000–2020) of inflow to the Mahabad Dam reservoir shows that in the first five months of the year (January to May), more inflow has entered the dam reservoir than in the following seven months (Fig. 3). At the same time, the amount of outflow from the dam was the highest in April to September. Also, the amount of HEPP was the highest from April to September. As a result, the highest volume of outflow from the dam occurs in the months in which the minimum inflow to the dam occurred. Therefore, proper planning is needed for better and more stable management of the dam.

To predict the dam inflow ( $Q_{in}$ ), several input variables including temperature ( $T$ ), precipitation ( $P$ ), and inflow with different monthly delays (one to three months denoted by  $Q_{in(t-1)}$ ,  $Q_{in(t-2)}$ , and  $Q_{in(t-3)}$ , respectively) were considered (Table 2). To determine the most appropriate combination of input variables, several scenarios were defined to consider as the machine learning inputs. The correlation coefficient of each input variable with the output was computed, allowing for the formulation of various input scenarios based on these coefficients. Analysis revealed that the dam inflow one month earlier and precipitation exhibited the strongest correlation with the output. Consequently, the first and second scenarios incorporated the dam inflow one month earlier and precipitation individually. Furthermore, based on the correlation coefficient, the third scenario encompassed a combination of these two variables. In total, six scenarios were defined utilizing the results of the correlation coefficients.

Several variables including  $P$ ,  $Q_{in}$ , outflow from the dam ( $Q_{out}$ ), dam evaporation ( $E$ ), and hydroelectric power produced one month earlier ( $HEPP_{(t-1)}$ ) were used to predict the HEPP in the Mahabad Dam. Similar to the method of defining input scenarios to predict the dam inflow, the results of the correlation coefficient were used to define scenarios for the prediction of HEPP. A total of five input scenarios were developed, in which the HEPP in one month earlier was considered the first input scenario, while the last input scenario included all input variables. Each scenario in the first

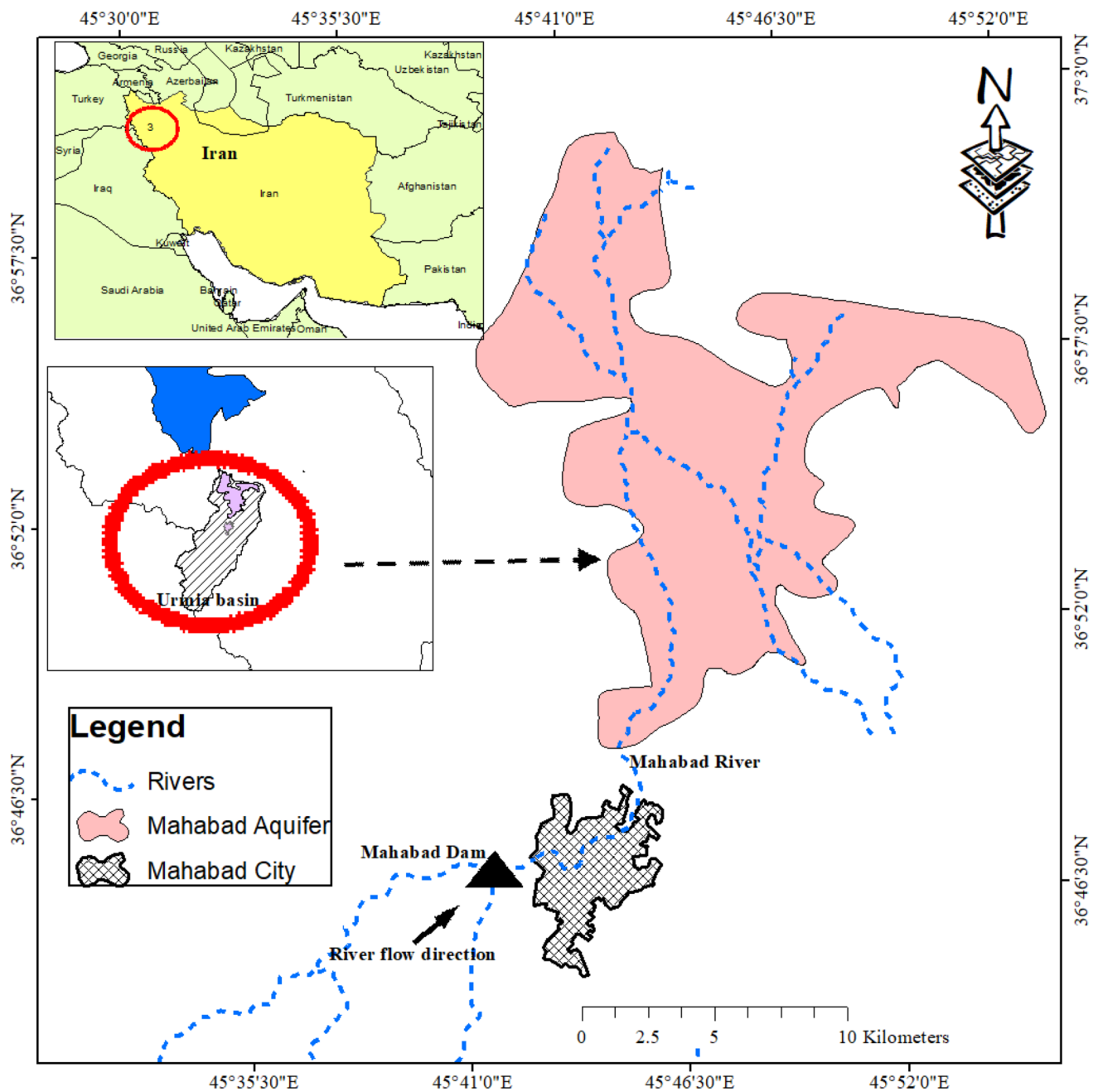


Fig. 1 Location of the Mahabad Dam catchment

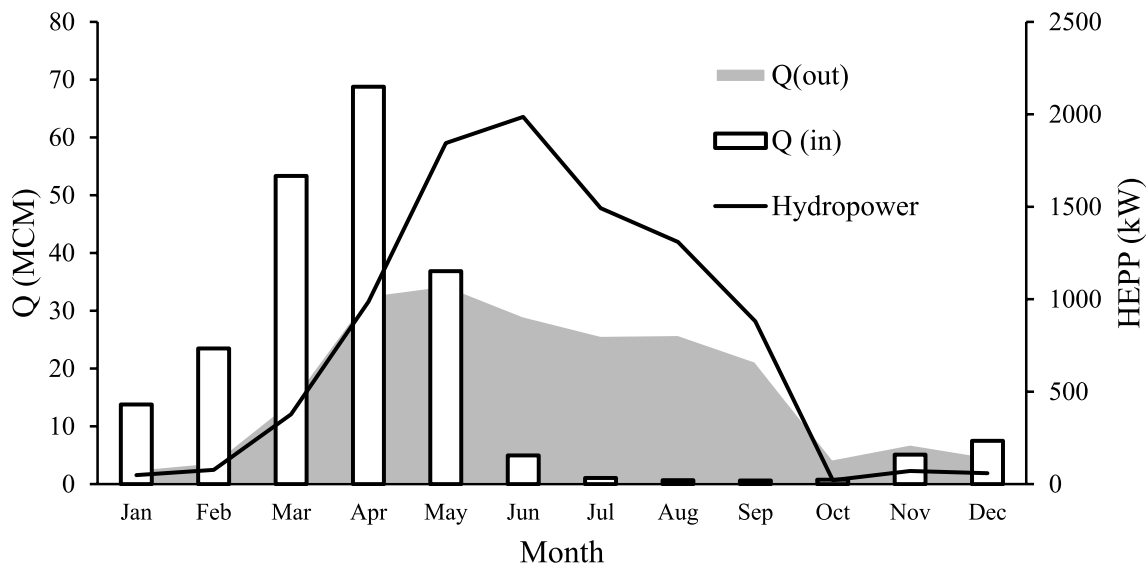
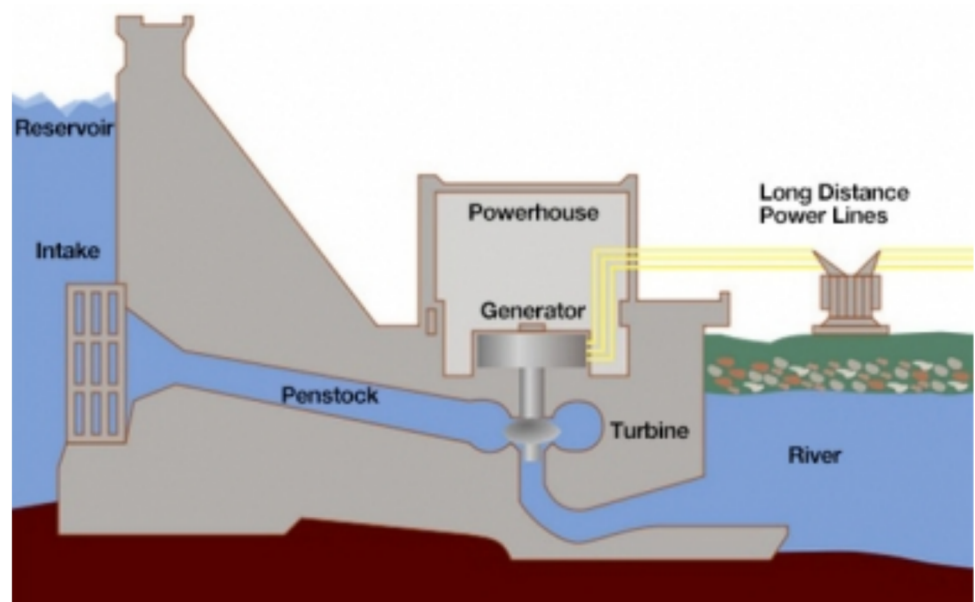
and second parts was implemented by the machine learning models optimized by evolutionary algorithms to predict the output variable.

### Methodology

Figure 4 depicts the process of modeling the amount of dam inflow. According to the scenarios defined as the input of machine learning methods, temperature, precipitation, and inflow to the reservoir with different delays have

been used. Modeling of the dam inflow was conducted by two base models, ANFIS and SVR. In the following, an efficient optimization method called HHO was used to improve the prediction performance, so the two hybrid models of HHO-ANFIS and HHO-SVR were developed. Each scenario was used as the input of the machine learning models, and the best input combination was selected for each model. The results of the models along with each scenario were analyzed using error evaluation criteria such as RMSE, MAE, and NSE and visual methods such as

**Fig. 2** The HEPP mechanism in the hydropower plant (blog. wika.us)



**Fig. 3** Time series of inflow and outflow of the dam reservoir and HEPP

Taylor’s diagram. In the following, the HEPP of the dam reservoir was predicted from the results of the first part, together with the parameters of the outflow from the dam, evaporation, and precipitation. In this section, the SVR model was used for prediction, which was optimized by the HHO algorithm. Similar to the first part, in this part, different input scenarios defined by the combination of input variables were considered. Each scenario was evaluated by both models and the results were compared. About 70% of the data were considered for training, while the rest was used to test the method.

### Adaptive neuro-fuzzy inference system

The ANFIS model, created by integrating ANN and FIS according to Jang (1993), offers superior performance compared to standalone ANN and FIS models. ANFIS mitigates common limitations observed in ANN and FIS, such as overfitting and sensitivity to membership function definitions. A typical ANFIS structure comprises five layers. In the first layer, the generalized Gaussian membership function  $\mu$  generates a new output,  $Out_{i1}$ , based on the inputs  $x$  and  $y$  (Eq. 1)

**Table 2** Input scenarios for machine learning to predict  $Q_{in}$  and HEPP. R values show the correlation coefficients

Scenario	Input variables	Output variable
Part1	S <sub>1</sub> $Q_{in(t-1)}$ (R=0.56)	$Q_{in}$
	S <sub>2</sub> P (R=0.55)	
	S <sub>3</sub> $Q_{in(t-1)} + P$	
	S <sub>4</sub> $Q_{in(t-1)} + P + Q_{in(t-2)}$ (R = -0.0.06)	
	S <sub>5</sub> $Q_{in(t-1)} + P + Q_{in(t-2)} + T$ (R = -0.28)	
	S <sub>6</sub> $Q_{in(t-1)} + P + Q_{in(t-2)} + T + Q_{in(t-3)}$	
Part2	S <sub>1</sub> HEPP <sub>(t-1)</sub> (R=0.71)	HEPP
	S <sub>2</sub> HEPP <sub>(t-1)} + E (0.63)</sub>	
	S <sub>3</sub> HEPP <sub>(t-1)} + E, Q<sub>out</sub> (0.62)</sub>	
	S <sub>4</sub> HEPP <sub>(t-1)} + E, Q<sub>out</sub>, Q<sub>in</sub> (R=0.10)</sub>	
	S <sub>5</sub> HEPP <sub>(t-1)} + E, Q<sub>out</sub>, Q<sub>in</sub>, P (R=-0.12)</sub>	

$$Out_{1i} = \mu A_i(x), i = 1, 2 \tag{1}$$

$$Out_{1i} = \mu B_{i-2}(y), i = 3, 4 \tag{1}$$

where

$$\mu(x) = e^{-\left(\frac{x - P_i}{\sigma_i}\right)^2} \tag{2}$$

and  $A_i$  and  $B_i$  are the membership values of  $\mu$ , while  $P_i$  and  $\sigma_i$  are the equation parameters. The output of each node is obtained in the second layer using Eq. (3)

$$Out_{2i} = \mu A_i(x) * \mu B_{i-2}(y) \tag{3}$$

Afterward, the output of layer 2 is normalized in layer 3 (Eq. 4)

$$Out_{3i} = \bar{W}_i = \frac{\omega_i}{\sum_{i=1}^2 \omega_i} \tag{4}$$

The output is then used in a linear combination equation

$$Out_{4,i} = \bar{w}_i f_i = \bar{w}_i (p_i x + q_i y + r_i) \tag{5}$$

where  $p$ ,  $q$ , and  $r$  are parameters defined for the  $i$ -th node. The model's output is obtained using Eq. (6).

$$Out_{4,i} = \sum_i \bar{w}_i f_i \tag{6}$$

### Support vector regression

Introduced by Vapnik (1998), SVR is a supervised learning machine specifically designed for regression tasks. SVR employs a structured risk minimization approach, utilizing a kernel function to create an optimal separating hyperplane. This hyperplane aims to maximize the geometric margin

while minimizing the upper bound of the generalization error, as outlined by:

$$d = w^T x + b \tag{7}$$

where  $w$  is the weight vector,  $b$  is bias, and  $d$  and  $x$  belong to the training sample  $J = \{x_i, d_i\}_{i=1}^N$ . Considering the  $\epsilon$ -insensitive loss function, the coefficients  $y$  and  $d$  are obtained by minimizing the risk function:

$$\frac{1}{2} \|w\|^2 + C \sum_{i=1}^N |y_i - d_i|_\epsilon \tag{8}$$

which is subjected to the constraints brought in Eqs. (9) to (12) for  $i = 1, 2, \dots, N$

$$d_i - y_i \leq \epsilon + \zeta_i \tag{9}$$

$$y_i - d_i \leq \epsilon + \zeta'_i \tag{10}$$

$$0 \leq \zeta_i \tag{11}$$

$$0 \leq \zeta'_i \tag{12}$$

where  $C$  represents a constant determining the trade-off between the training error and the penalization term  $\|w\|^2$ , and  $y_i$  denotes the estimator output, and  $\zeta$  along with  $\zeta'$  are nonnegative parameters. To address the optimization problem of Eq. (8), Lagrange multipliers are introduced, facilitating the expression of the minimization formula. Additionally, a kernel function is introduced to transform the problem into a nonlinear regression problem. Various kernel functions, such as linear, polynomial, radial basis function (RBF), and sigmoid kernels, have been introduced for use in SVR structures. For this study, the RBF kernel, widely utilized, was selected during the SVR modeling (Kayhomayoon et al. 2023).

### Harris hawks optimization (HHO)

HHO draws inspiration from the hunting behavior of Harris hawks, as detailed in the work by Heidari et al. (2019). This approach comprises two distinct phases known as soft and hard sieges. During the soft siege, when the prey retains ample energy, it attempts to evade capture through unpredictable and misleading jumps. The hawks employ a gentle surrounding strategy to gradually wear down the prey. On the other hand, in the hard siege phase, the prey is thoroughly fatigued and possesses minimal escape capability. Harris hawks employ a more intense surrounding approach, culminating in a surprise pounce to capture the prey. The algorithm involves the random movement of

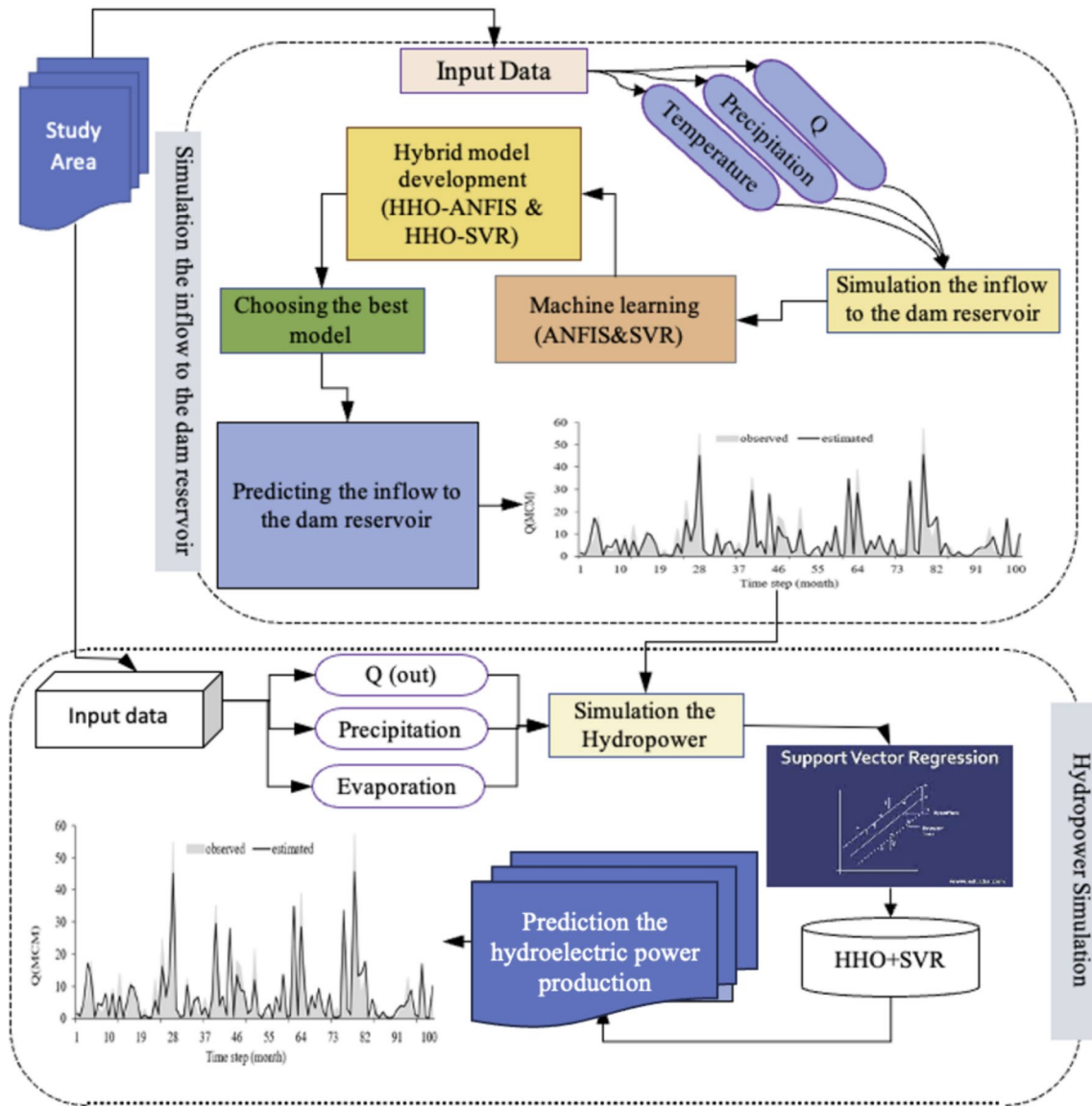


Fig. 4 The flowchart of the current study

Harris hawks as they seek out prey, with their positions determined by Eq. (13) (Fig. 5).

$$X(t + 1) = \begin{cases} X_{rand}(t) - r_1 |X_{rand}(t) - 2r_2 X(t)| & q \geq 0.5 \\ (X_{rabbit}(t) - X_m(t)) - r_3 (LB + r_4(UB - LB)) & q < 0.5 \end{cases} \quad (13)$$

where  $X(t)$  and  $X(t + 1)$  are the positions of hawks at iterations  $t$  and  $t + 1$ , respectively,  $X_{rabbit}(t)$  is the rabbit's position,  $r_1, r_2, r_3, r_4$ , and  $q$  are random numbers and are updated in each iteration,  $UB$  and  $LB$  and the lower and upper limits of variables,  $X_{rand}(t)$  is the position of an arbitrary hawk, and  $X_m$  is the average position of the population, which is obtained using Eq. (14) (Fig. 5).

$$X_m(t) = \frac{1}{N} \sum_{i=1}^N X_i(t) \quad (14)$$

where  $N$  is the total number of hawks, and  $X_i(t)$  is the position of each hawk in iteration  $t$ . The prey's energy during the escape is defined by Eq. (15)

$$E = 2E_0 \left(1 - \frac{t}{T}\right) \quad (15)$$

where  $T$  is the maximum iteration number,  $E$  is the prey's energy, and  $E_0$  is the initial energy (Fig. 5).

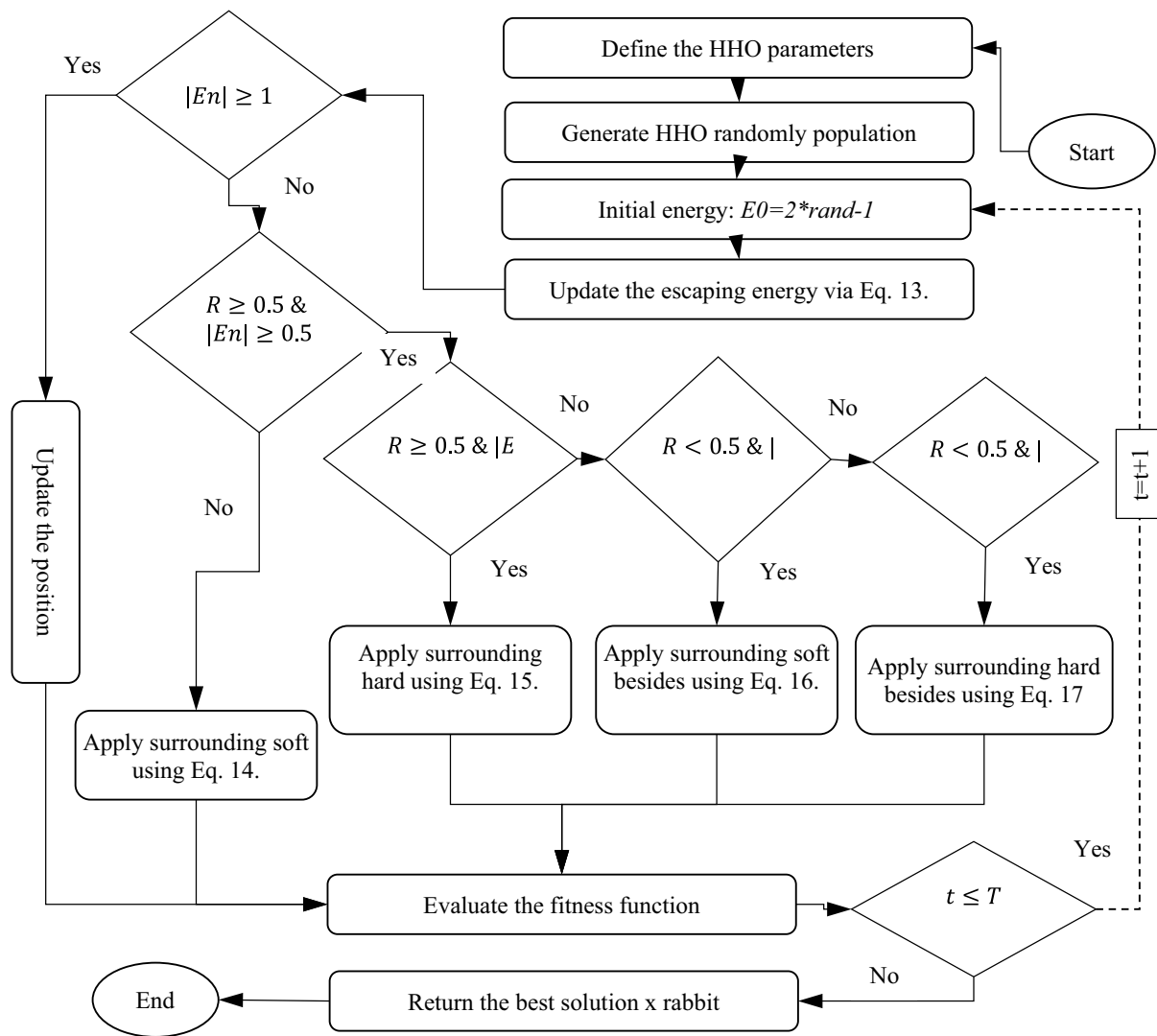


Fig. 5 Flowchart of the HHO algorithm

Table 3 Parameters used in the structure of algorithms

Algorithm	Parameter	Value
ANFIS	Fuzzy structure	Sugeno-type
	Initial FIS for training	Genfis3
	The type of membership functions	Gaussian
	The membership function of the output	Linear
	Optimization method	Hybrid
	Number of fuzzy rules	10
	The maximum number of epochs	2000
HHO	Number of search agent	40
	Iteration number	2000
	$\beta$	1.5
	Range partitions (weights and biases)	$[-3, +3]$
	Population size	30

Table 3 shows the values of the parameters used in the structure of ANFIS and HHO. HHO was used to improve the performance of SVR and ANFIS by creating hybrid methods of HHO-SVR and HHO-ANFIS. According to the table, the Sugeno structure was considered for ANFIS, where the output was in the form of a linear function. A total of 10 if-then fuzzy rules have been formulated for simulation in ANFIS, where they were replaced with HHO in the hybrid models. The maximum number of iterations during the optimization process was 2000.

### Performance evaluation criteria

Root mean square error (RMSE) (Eq. 16), Nash Sutcliffe index (NSE) (Eq. 17), and mean absolute error (MAE) (Eq. 18) were calculated to evaluate the performance



of scenarios and machine learning models used in this study (Najafzadeh et al. 2021):

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_o - x_p)^2}{n}} \quad 0 \leq RMSE \leq \infty \quad (16)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (x_o - x_p)^2}{\sum_{i=1}^n (x_o - \bar{x}_o)^2} \quad -1 \leq NSE \leq 1 \quad (17)$$

$$MAE = \frac{\sum_{i=1}^n |x_o - x_p|}{n} \quad 0 \leq MAE \leq \infty \quad (18)$$

where  $x_o$  is the observed (measured) value,  $x_p$  is the predicted (estimated) value, and  $n$  is the number of samples.  $\bar{x}_o$  and  $\bar{x}_p$  represent the average of the observed and predicted data, respectively. Also, some graphic diagrams such as Taylor’s diagram, scatterplots, and time series plots were used for the comparison of the models and scenarios.

## Results

### Simulation of the dam inflow

The prediction of current dam inflow was based on precipitation, temperature, and dam inflow one month earlier. Six input scenarios were generated from the combination of these variables, and their performance as inputs to machine learning models is summarized in Table 4. Evaluation criteria in the table indicate that scenarios incorporating more input variables exhibit superior modeling performance compared to those with fewer variables. Consequently, the sixth scenario, encompassing all input variables, demonstrated the highest modeling performance across all machine learning models utilized. Within this scenario, the HHO-ANFIS hybrid model performed the best, yielding RMSE, MAE, and NSE values for the test data of 3.9 MCM per month, 2.41 MCM per month, and 0.86, respectively. Conversely, for the ANFIS model, in the same scenario, RMSE, MAE, and NSE values were 16.24 MCM per month, 6.75 MCM per month, and -0.26, respectively, indicating instances where the ANFIS model

**Table 4** The performance of scenarios and machine learning models in the prediction of dam inflow

Scenario	Model	NSE		MAE (MCM)		RMSE (MCM)	
		Test	Train	Test	Train	Test	Train
S1	ANFIS	-0.08	0.55	6.10	4.48	11.80	8.65
	HHO-ANFIS	0.62	0.59	4.70	4.54	8.00	7.75
	SVR	0.50	0.56	5.10	4.80	8.50	7.89
	HHO-SVR	0.60	0.59	4.10	4.76	6.80	8.30
S2	ANFIS	0.34	0.57	6.00	4.79	10.46	8.31
	HHO-ANFIS	0.72	0.72	4.47	4.38	6.96	6.83
	SVR	0.60	0.79	4.60	4.56	6.94	6.78
	HHO-SVR	0.66	0.73	4.49	4.47	6.68	6.67
S3	ANFIS	0.50	0.77	4.87	3.10	12.5	5.96
	HHO-ANFIS	0.60	0.80	2.86	3.14	6.60	5.88
	SVR	0.58	0.70	3.20	3.35	6.70	6.12
	HHO-SVR	0.77	0.78	3.60	2.80	6.00	5.70
S4	ANFIS	0.55	0.85	5.42	2.61	8/59	4.62
	HHO-ANFIS	0.74	0.76	3.77	3.25	5.88	5.90
	SVR	0.72	0.74	3.80	3.30	5.92	5.95
	HHO-SVR	0.79	0.88	3.13	2.18	5.01	4.60
S5	ANFIS	-0.49	0.97	8.00	1.26	15.0	2.15
	HHO-ANFIS	0.71	0.87	3.60	2.27	6.25	4.54
	SVR	0.75	0.78	3.62	2.80	6.29	4.60
	HHO-SVR	0.70	0.83	3.12	3.20	6.30	5.20
<b>S6</b>	ANFIS	-0.26	0.92	6.75	1.40	16.24	3.04
	<b>HHO-ANFIS</b>	<b>0.86</b>	<b>0.90</b>	<b>2.41</b>	<b>2.00</b>	<b>3.90</b>	<b>4.17</b>
	<b>SVR</b>	<b>0.73</b>	<b>0.79</b>	<b>2.90</b>	<b>2.40</b>	<b>5.90</b>	<b>4.80</b>
	<b>HHO-SVR</b>	<b>0.86</b>	<b>0.89</b>	<b>2.70</b>	<b>2.51</b>	<b>4.39</b>	<b>4.19</b>

The superior ML model and scenario are shown in bold

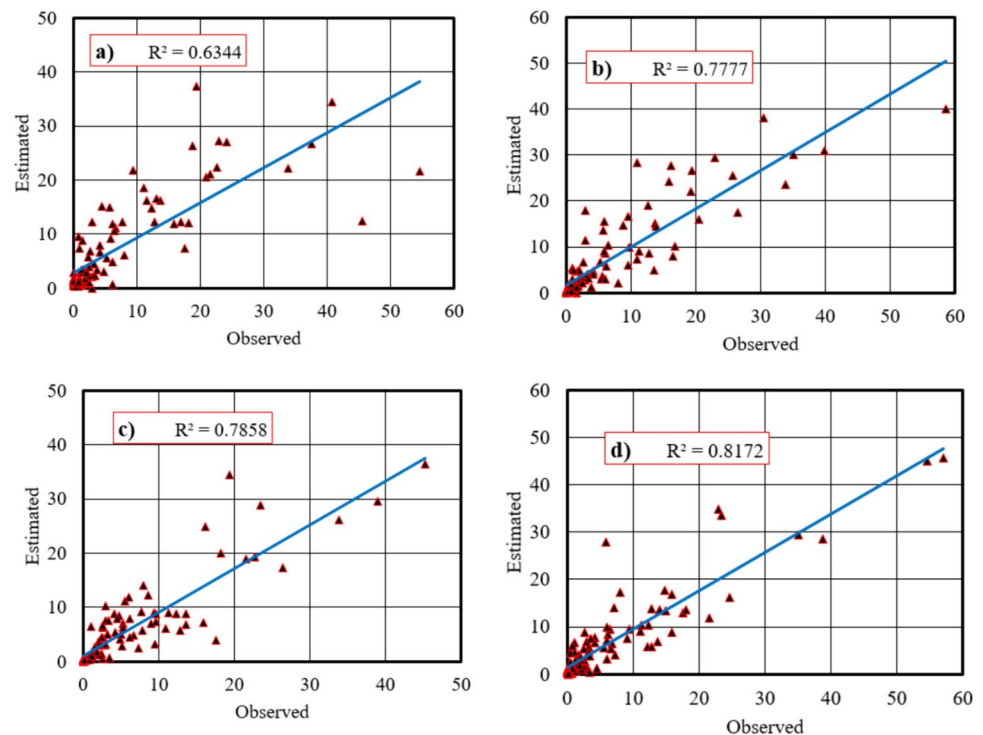
was trapped in local optimal points, resulting in unreliable predictions and a significant disparity in modeling accuracy between training and test data. However, noteworthy accuracy in predicting dam inflow for the ANFIS model was observed in the fourth scenario. This scenario incorporated inflow to the reservoir from one and two months earlier, along with monthly precipitation as input variables. For this scenario, the RMSE, MAE, and NSE error evaluation criteria for the test data were 8.59 MCM per month, 5.42 MCM per month, and 0.55, respectively.

The findings indicate that overall, the HHO-SVR model outperformed the SVR model, consistently providing more accurate predictions of dam inflow across most scenarios. Particularly, the sixth scenario demonstrated the highest accuracy for the HHO-SVR model, with RMSE, MAE, and NSE of 4.39, 2.70, and 0.86, respectively. In comparison, the SVR model exhibited a prediction with RMSE, MAE, and NSE of 5.90, 2.90, and 0.73, respectively, in the same scenario, indicating higher accuracy compared to other scenarios. A comparison of the sixth scenario results between the SVR and HHO-SVR models reveals that the hybrid model significantly reduces model error and enhances the NSE value by approximately 0.13. Although the HHO-ANFIS hybrid model generally exhibits superior accuracy in estimating dam inflow compared to the HHO-SVR hybrid model, the performance of the single SVR model surpasses ANFIS, particularly for the test data. The lowest accuracy among the models was observed in scenarios  $S_1$  and  $S_2$ , each of which included only one input variable.

Figure 6 depicts scatterplots of observed and predicted values along with fitted regression lines. The proximity of values to the regression line indicates the model's performance, with points scattered relative to the line suggesting poorer performance. Upon examination, all models exhibit relatively good density around the regression line. The  $R^2$  values of the fitted models range from 0.63 to 0.82, with the lowest value attributed to the ANFIS model and the highest ( $R^2 = 0.82$ ) achieved by the HHO-ANFIS model. Notably, there's a significant disparity between the figures for HHO-ANFIS and HHO-SVR on one side, and SVR and ANFIS on the other. Some points in the figure exhibit considerable distance from the regression line, indicating significant discrepancies between observed and predicted values. This weakness is also evident in the SVR model. Conversely, other models demonstrate suitable density around the regression line, with  $R^2$  values closely aligned.

Taylor's diagram, depicted in Fig. 7, was utilized for further evaluation to examine the correlation and standard deviation values between predicted and observed values. The correlation coefficient for all four models fell within the range of 0.79 to 0.91, indicating the efficacy of all models in estimating dam inflow. Interestingly, the HHO-ANFIS model exhibited a slightly higher correlation coefficient compared to the other three models. The Root Mean Square Deviation (RMSD) value for HHO-ANFIS lies on the arc of 4.3, while for ANFIS, it slightly exceeds 6. Moreover, the diagram illustrates no significant difference between HHO-SVR and HHO-ANFIS modeling.

**Fig. 6** The scatterplot of the models, **a** ANFIS, **b** SVR, **c** HHO-SVR, **d** HHO-ANFIS



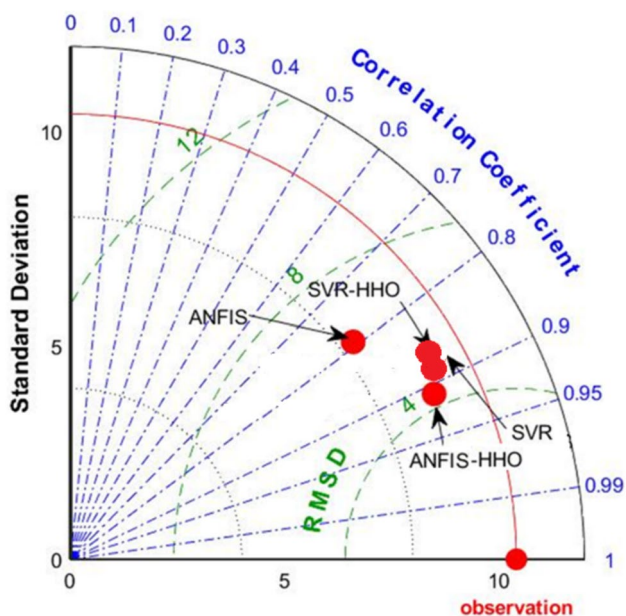


Fig. 7 Taylor's diagram of the models in predicting the dam inflow

Figure 8 illustrates that all models successfully predicted the minimum values (base inflow) with acceptable performance. However, ANFIS exhibited an overestimation of inflow across several months, including steps 37, 64, 73, 78, and 97. Similarly, the HHO-SVR model displayed overestimation for steps 28 and 80. Upon comparison, it is evident that the HHO-ANFIS model outperformed the other models

in predicting maximum values, with its data showing greater consistency with observational data.

### Simulation of hydroelectric power production

According to Table 5, which presents the error evaluation criteria of the models and scenarios, the HHO-SVR model demonstrated satisfactory performance in predicting the HEPP. Evaluation of the SVR model revealed that the  $S_3$  input scenario yielded superior prediction results compared to other scenarios. In this scenario, which included variables such as HEPP from one month earlier, evaporation, and outflow from the dam, the RMSE, MAE, NSE, and  $R^2$  values for the test data were 423.7 kW, 188 kW, 0.88, and 0.86, respectively. Conversely,  $S_2$  exhibited the lowest prediction performance, with RMSE, MAE, NSE, and  $R^2$  values of 762.7 kW, 42 kW, 0.72, and 0.49, respectively. This suggests that evaporation significantly contributed to improving prediction performance in the third scenario compared to the second. In the fifth scenario, where all input variables were included, predictions yielded error evaluation criteria of RMSE = 392.7 kW, MAE = 177 kW, NSE = 0.91, and  $R^2 = 0.88$  for the test data, demonstrating appropriate prediction accuracy using the HHO-SVR hybrid model. Among the scenarios utilizing the hybrid model, the first input scenario exhibited the lowest prediction performance, with RMSE, MAE, NSE, and  $R^2$  values of 775.7 kW, 178 kW, 0.48, and 0.71, respectively, for the test data. In summary,

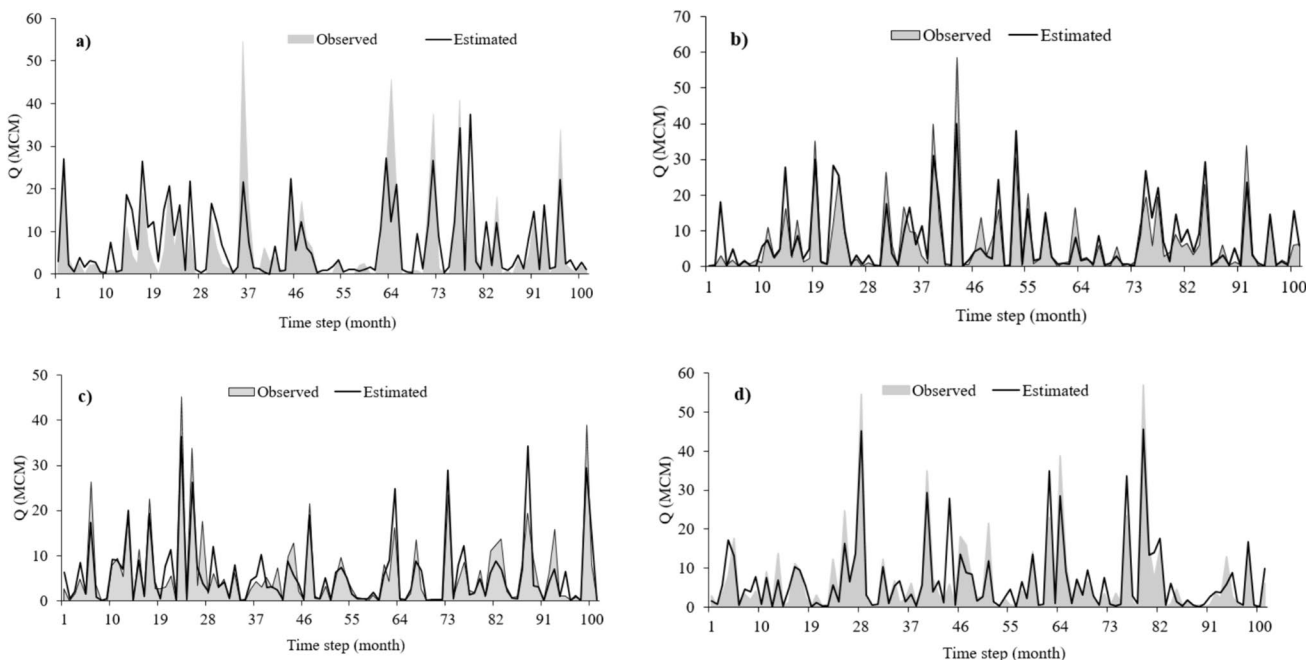


Fig. 8 Time-series plot of observational and predicted data by machine learning models for the prediction of dam inflow, a ANFIS, b SVR, c HHO-ANFIS, d HHO-SVR

**Table 5** The performance of scenarios and machine learning models in the prediction of HEPP

Models	Scenarios	RMSE		MAE		NSE		R <sup>2</sup>	
		training	test	training	test	training	test	training	test
SVR	S <sub>1</sub>	724.2	580.6	1.9	160.6	0.53	0.73	0.54	0.76
	S <sub>2</sub>	580.1	762.3	23.7	42.4	0.72	0.49	0.72	0.49
	<b>S<sub>3</sub></b>	<b>371.4</b>	<b>423.7</b>	<b>15.1</b>	<b>65.4</b>	<b>0.88</b>	<b>0.85</b>	<b>0.88</b>	<b>0.86</b>
	S <sub>4</sub>	360.3	468.2	87.5	227.5	0.90	0.69	0.91	0.78
	S <sub>5</sub>	389.1	551.6	21.5	88.9	0.88	0.70	0.88	0.74
HHO-SVR	S <sub>1</sub>	775.1	604.6	0.00	4.80	0.48	0.70	0.48	0.71
	S <sub>2</sub>	593.8	570.6	9.95	53.0	0.70	0.68	0.71	0.69
	S <sub>3</sub>	393.8	487.7	12.6	77.4	0.87	0.77	0.87	0.78
	S <sub>4</sub>	341.1	409.7	11.3	8.40	0.90	0.87	0.90	0.88
	<b>S<sub>5</sub></b>	<b>330.8</b>	<b>392.1</b>	<b>17.7</b>	<b>74.0</b>	<b>0.91</b>	<b>0.86</b>	<b>0.91</b>	<b>0.88</b>

The selected model and scenario are highlighted in bold

the results indicate that HHO-SVR outperformed SVR, suggesting that HHO effectively enhances the prediction results of SVR.

In the following, the time-series analysis of the selected scenario of each model was done for the test data to better understand their performance. The results of the time series of observed and predicted HEPP values used for the test data showed that the HHO-SVR model estimated the test data more accurately than the SVR in the selected scenario (Fig. 9). As it is clear from the results of the models, both models had a good estimate of the base values of HEPP, in other words, the models have well estimated the minimum values of HEPP. However, in predicting the maximum values, the HHO-SVR model has performed better than the SVR. For example, HHO-SVR provided somewhat appropriate performance in predicting steps 8, 26, and 58. However, in SVR, some maximum values of HEPP, in some steps such as steps 15 and 30, are not predicted with an acceptable performance. The error between observed and predicted values in both models follows a normal distribution with standard deviations of 388 and 422 for HHO-SVR and SVR, respectively. The average error in SVR also shows that this model has underestimated the values, especially in the maximum values (mean error = -65), while HHO-SVR has overestimated the values (mean error = 74). The results also show that the errors in HHO-SVR are always between -1500 and -750, while this error ranges between -1000 and -1750 in SVR.

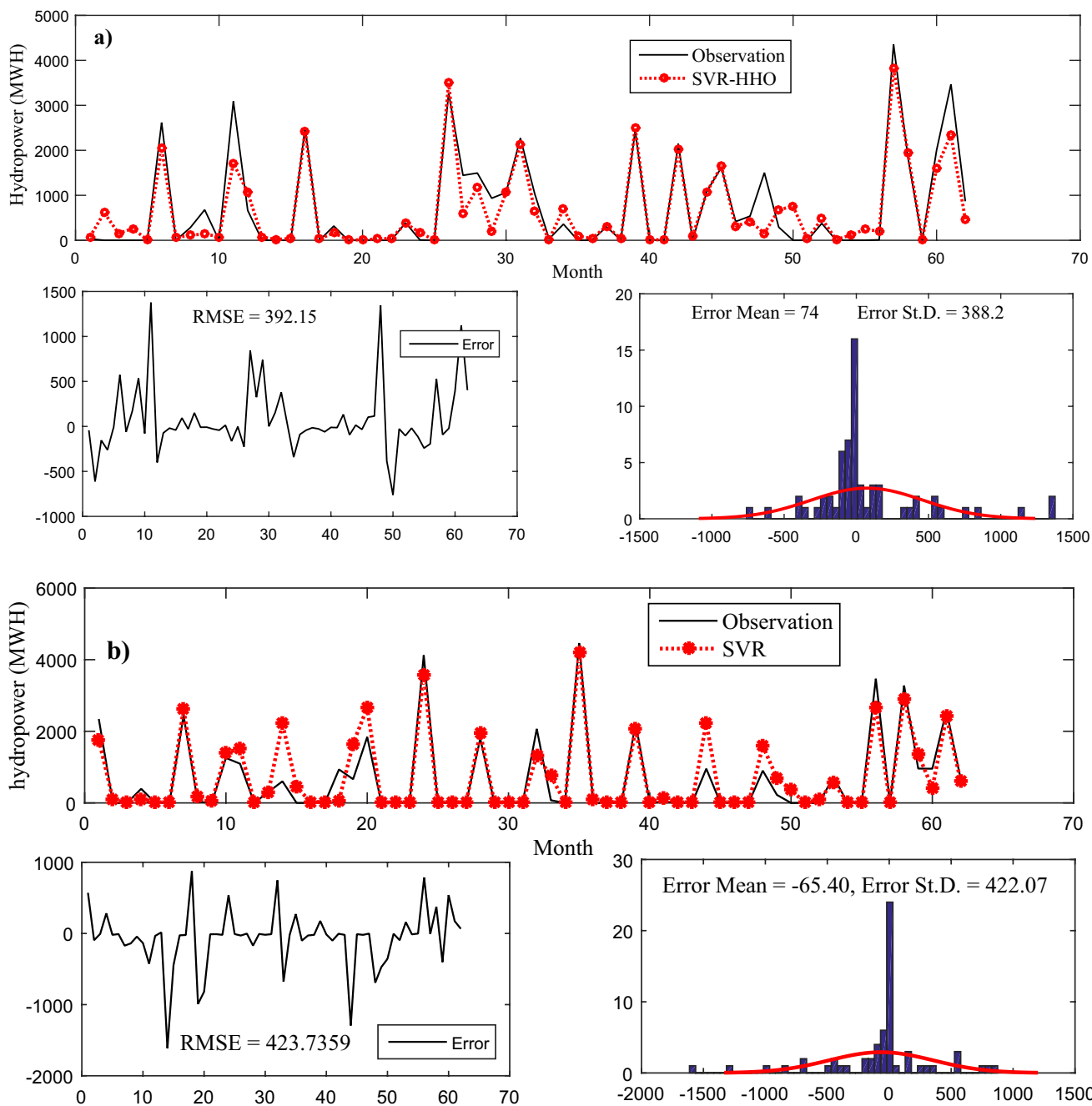
Among the other graphs examined in this research is the scatterplot, which expresses the trend and intensity of differences between predicted and observed values. According to Fig. 10, an acceptable agreement was observed between the actual and predicted values. Although R<sup>2</sup> values in both models were always greater than 0.85 and the positions of the points were close to the regression line, the results of the values and visual evaluation show that the performance of HHO-SVR was more appropriate.

Based on the position of the models relative to the observed data in Taylor's diagram (Fig. 11), the HHO-SVR model appears to be a better model than the SVR model. In this diagram, the vertical and horizontal axes represent the values of the standard deviation, while the arcs inside the quarter circle indicate the RMSD, and the correlation coefficient can be calculated from the quarter circle arc. The results indicate that both models achieved correlation coefficients between 0.90 and 0.93, demonstrating their appropriate efficiency in predicting HEPP. The RMSD values for HHO-SVR and SVR were 420 kW and 435 kW, respectively. Additionally, the standard deviation of the output of the HHO-SVR model was closer to that of the observational data. In conclusion, based on these observations, it can be inferred that the performance of the HHO-SVR model was more appropriate compared to SVR.

## Discussion

In arid and semi-arid areas such as Mahabad, due to the decrease in the river flow to the dam and the heat of the air, the estimation of HEPP from the dams always faces serious challenges. A sudden outage of municipal electricity causes irreparable damage to home electrical systems, which is expensive to compensate for. Therefore, its proper estimation according to the amount of the river in this season can prevent such damage to a large extent. Therefore, the approach proposed in this work can be an effective help to managers and decision-makers.

Machine learning models have demonstrated their efficacy as efficient and reliable tools for predicting time series, as evidenced in the literature (Ghorbani et al., 2020; Tikhamarine et al. 2019; Arya Azar et al. 2021c). In this study, both ANFIS and SVR were employed to forecast dam inflow and HEPP. While both models exhibited relatively good performance in estimating dam inflow, ANFIS

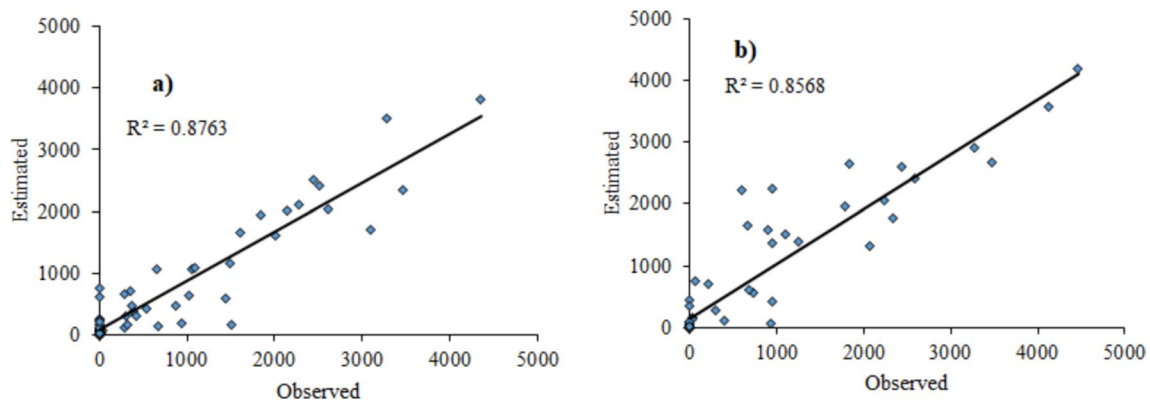


**Fig. 9** Time-series graph of observed and predicted data and error graph of each model using the selected scenarios

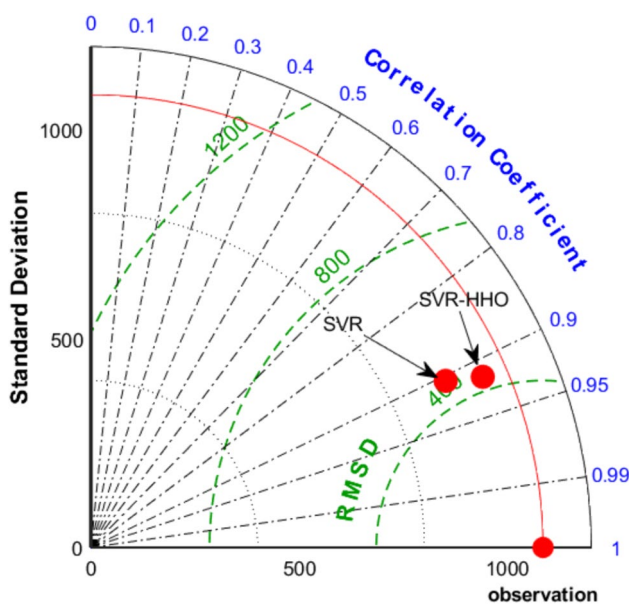
was found to be less accurate, consistent with findings from previous research by Arya Azar et al. (2021b, c, d). One advantage of the SVR model is its reduced susceptibility to getting trapped in local optimal points during parameter optimization, leading to improved computational efficiency. Although ANFIS partially mitigates uncertainty by employing fuzzy logic for prediction, it typically yields weaker results, as supported by previous research and the outcomes of this study. However, it is worth noting that by using evolutionary algorithms to find

appropriate adjustment parameters, the prediction performance of ANFIS can be readily enhanced.

The efficacy of the HHO algorithm was assessed in enhancing the predictive capabilities of ANFIS and SVR models for dam inflow forecasting. Both HHO-ANFIS and HHO-SVR hybrid models outperformed their respective single models in prediction accuracy. These findings align with previous research by Sammen et al. (2020), Arya Azar et al. (2021a, b, c), and Milan et al. (2021), indicating that metaheuristic algorithms can indeed enhance the



**Fig. 10** Scatterplot of the observed and predicted data of selected scenarios in each model



**Fig. 11** Taylor's diagram of the models in predicting the HEPP

performance of base models such as ANN and ANFIS. Specifically, the results demonstrated that the HHO algorithm exhibited superior performance in improving ANFIS training, underscoring its effectiveness in identifying optimal values for ANFIS parameters. This finding corroborates similar research outcomes in the literature, such as those by Arya Azar et al. (2021a, b, c) and Paryani et al. (2021), which also validated the performance of HHO-ANFIS in estimating hydrological and environmental parameters.

The proposed approach can also be used for similar regions and areas. This approach is highly recommended, especially for areas that are in arid and semi-arid conditions or in areas where extensive data collection is not possible. Of course, such researches also have uncertainties, which can be investigated in future research. Climate change and drought

are among the extreme events that have a significant impact on water resources, which can be evaluated in the future. Finally, the practical use of this approach is recommended for proper management of water resources by managers.

## Conclusions

This study aimed to investigate the performance of various machine learning models in estimating dam inflow and HEPP, focusing on Mahabad Dam as the study area. For forecasting purposes, input variables such as temperature and inflow from the previous month were utilized in different combinations, forming input scenarios. ANFIS and SVR models, as well as their optimized versions using the HHO algorithm, were employed for prediction. The findings revealed that these models yielded acceptable performance, which can be valuable for hydrological studies and dam allocation management. The utilization of such tools can prove effective in surface water and groundwater management, as well as in the prediction of meteorological parameters, due to their ability to provide reliable results in a short timeframe. The two hybrid models, HHO-ANFIS and HHO-SVR, exhibited improvements over the base models, particularly enhancing the performance of the ANFIS model. Both Taylor's and ridgeline plots corroborated these findings, indicating the superior performance of the hybrid model. Combining variables such as dam inflow and outflow, precipitation, and evaporation contributed to enhanced accuracy in HEPP forecasting results. Furthermore, the results highlighted the efficacy of the HHO algorithm in determining optimal parameter values for both ANFIS and SVR models. Future research could explore uncertainties in machine learning model results and the impact of climate change on daily-scale variations in dam inflow and evaporation. This would contribute to a deeper understanding of

the dynamics involved and aid in developing more robust forecasting methodologies.

**Authors' contribution** All authors contributed to the study's conception, design, and writing. All authors read and approved the final manuscript.

**Funding** Authors state no funding is involved.

**Data availability** The research data associated with this study are available in the paper.

## Declarations

**Conflict of interest** The authors of this paper declare that they have no conflict of interest.

**Ethical approval** There are no ethical issues.

**Consent to publish** All authors agree to publish.

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

## References

- Abbasi A, Firouzi B, Sendur P (2021) On the application of Harris hawks optimization (HHO) algorithm to the design of microchannel heat sinks. *Eng Comput* 37:1409–1428
- Ahmed F, Siwar C, Begum RA (2014) Water resources in Malaysia: Issues and challenges. *J Food Agric Environ* 12(2):1100–1104
- Arya Azar N, Ghordoyee Milan S, Kardan N (2021a) Development of a hybrid ANN-evolutionary algorithms models to predict the Froude number in open channel flows in modeling of sediment transport. *Environ Water Eng* 7(1):73–87
- Arya Azar N, Ghordoyee Milan S, Kayhomayoon Z (2021b) Predicting monthly evaporation from dam reservoirs using LS-SVR and ANFIS optimized by Harris hawks optimization algorithm. *Environ Monit Assess* 193(11):1–14
- Arya Azar N, Kardan N, Ghordoyee Milan S (2021c) Developing the artificial neural network–evolutionary algorithms hybrid models (ANN–EA) to predict the daily evaporation from dam reservoirs. *Eng Comput* 39:1–19
- Arya Azar N, Milan SG, Kayhomayoon Z (2021d) The prediction of longitudinal dispersion coefficient in natural streams using LS-SVM and ANFIS optimized by Harris hawk optimization algorithm. *J Contam Hydrol* 240:103781
- Babaei M, Moeini R, Ehsanzadeh E (2019) Artificial neural network and support vector machine models for inflow prediction of dam reservoir (case study: Zayandehroud dam reservoir). *Water Resour Manage* 33:2203–2218
- Banadkooki FB, Ehteram M, Panahi F, Sammen SS, Othman FB, Ahmed ES (2020) Estimation of total dissolved solids (TDS) using new hybrid machine learning models. *J Hydrol* 587:124989
- Barzola-Monteses J, Gomez-Romero J, Espinoza-Andaluz M, Fajardo W (2022) Hydropower production prediction using artificial neural networks: an Ecuadorian application case. *Neural Comput Appl* 34(16):13253–13266
- Bilgili M, Ozbek A, Sahin B, Kahraman A (2015) An overview of renewable electric power capacity and progress in new technologies in the world. *Renew Sustain Energy Rev* 49:323–334
- Choubin B, Moradi E, Golshan M, Adamowski J, Sajedi-Hosseini F, Mosavi A (2019) An ensemble prediction of flood susceptibility using multivariate discriminant analysis, classification and regression trees, and support vector machines. *Sci Total Environ* 651:2087–2096
- Clerici A, Alimonti G (2015) World energy resources. *EPJ WEB Conf* 98:01001
- Dehghani M, Riahi-Madvar H, Hooshyaripour F, Mosavi A, Shamshirband S, Zavadskas EK, Chau KW (2019) Prediction of hydropower generation using grey wolf optimization adaptive neuro-fuzzy inference system. *Energies* 12(2):289
- Dmitrieva K (2015) Forecasting of a hydropower plant energy production (Master's thesis)
- Enayati SM, Najarchi M, Mohammadpour O, Mirhosseini SM (2022) Development of a hybrid adaptive neuro fuzzy inference system—Harris hawks optimizer (ANFIS-HHO) for inlet flow to the dam reservoir prediction. *Environ Water Eng* 8(4):796–809
- Esmaili M, Aliniaiefard S, Mashal M, Asefpour Vakilian K, Ghorbanzadeh P, Azadegan B, Seif M, Didaran F (2021) Assessment of adaptive neuro-fuzzy inference system (ANFIS) to predict production and water productivity of lettuce in response to different light intensities and CO<sub>2</sub> concentrations. *Agric Water Manag* 258:107201
- Fathian F, Fard AF, Ouarda TB, Dinpashoh Y, Nadoushani SM (2019) Modeling streamflow time series using nonlinear SETAR-GARCH models. *J Hydrol* 573:82–97
- Firat M, Güngör M (2007) River flow estimation using adaptive neuro fuzzy inference system. *Math Comput Simul* 75(3–4):87–96
- Ghorbani MA, Khatibi R, Singh VP, Kahya E, Ruskeepää H, Saggi MK, Jani R (2020) Continuous monitoring of suspended sediment concentrations using image analytics and deriving inherent correlations by machine learning. *Sci Rep* 10(1):8589
- Guo J, Zhou J, Qin H, Zou Q, Li Q (2011) Monthly streamflow forecasting based on improved support vector machine model. *Expert Syst Appl* 38(10):13073–13081
- Hanoon MS, Ahmed AN, Razzaq A, Oudah AY, Alkhayyat A, Huang YF, El-Shafie A (2023) Prediction of hydropower generation via machine learning algorithms at three Gorges Dam. *China Ain Shams Eng J* 14(4):101919
- Hashemi A, Asefpour Vakilian K, Khazaei J, Massah J (2014) An artificial neural network modeling for force control system of a robotic pruning machine. *J Inf Org Sci* 38(1):35–41
- Heidari AA, Mirjalili S, Faris H, Aljarah I, Mafarja M, Chen H (2019) Harris hawks optimization: algorithm and applications. *Futur Gener Comput Syst* 97:849–872
- Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. *IEEE Trans Syst Man Cybern* 23(3):665–685
- Jimeno-Sáez P, Senent-Aparicio J, Pérez-Sánchez J, Pulido-Velazquez D (2018) A comparison of SWAT and ANN models for daily runoff simulation in different climatic zones of peninsular Spain. *Water* 10(2):192
- Kayhomayoon Z, Azar NA, Milan SG, Moghaddam HK, Berndtsson R (2021a) Novel approach for predicting groundwater storage loss using machine learning. *J Environ Manage* 296:113237

- Kayhomayoon Z, Ghordoyee Milan S, Arya Azar N, Kardan Moghaddam H (2021) A new approach for regional groundwater level simulation: clustering, simulation, and optimization. *Nat Resour Res* 30:1–21
- Kayhomayoon Z, Jamnani MR, Rashidi S, Milan SG, Azar NA, Berndtsson R (2023) Soft computing assessment of current and future groundwater resources under CMIP6 scenarios in north-western Iran. *Agric Water Manag* 285:108369
- Khosravi K, Golkarian A, Booi MJ, Barzegar R, Sun W, Yaseen ZM, Mosavi A (2021) Improving daily stochastic streamflow prediction: Comparison of novel hybrid data mining algorithms. *Hydrol Sci J* 66:1457
- Kim BJ, Lee YT, Kim BH (2022) A study on the optimal deep learning model for dam inflow prediction. *Water* 14(17):2766
- Latif SD, Ahmed AN (2024) Ensuring a generalizable machine learning model for forecasting reservoir inflow in Kurdistan region of Iraq and Australia. *Environ Dev Sustain* 26(5):12513–12544
- Meshram SG, Ghorbani MA, Shamshirband S, Karimi V, Meshram C (2019) River flow prediction using hybrid PSO-GSA algorithm based on feed-forward neural network. *Soft Comput* 23(20):10429–10438
- Milan SG, Roozbahani A, Banihabib ME (2018) Fuzzy optimization model and fuzzy inference system for conjunctive use of surface and groundwater resources. *J Hydrol* 566:421–434
- Milan SG, Roozbahani A, Azar NA, Javadi S (2021) Development of adaptive neuro fuzzy inference system–Evolutionary algorithms hybrid models (ANFIS-EA) for prediction of optimal groundwater exploitation. *J Hydrol* 598:126258
- Moayedi H, Tien Bui D, Anastasios D, Kalantar B (2019) Spotted hyena optimizer and ant lion optimization in predicting the shear strength of soil. *Appl Sci* 9(22):4738
- Mohammadi B (2021) A review on the applications of machine learning for runoff modeling. *Sustain Water Resour Manag* 7(6):1–11
- Musarat MA, Alaloul WS, Rabbani MBA, Ali M, Altaf M, Fediuk R, Farooq W (2021) Kabul river flow prediction using automated ARIMA forecasting: a machine learning approach. *Sustainability* 13(19):10720
- Najafzadeh M, Homaei F, Mohamadi S (2021) Reliability evaluation of groundwater quality index using data-driven models. *Environ Sci Pollut Res* 29:1–17
- Nematollahi Z, Sanayei HRZ (2023) Developing an optimized groundwater exploitation prediction model based on the Harris hawk optimization algorithm for conjunctive use of surface water and groundwater resources. *Environ Sci Pollut Res* 30(6):16120–16139
- Nguyen MT, Sebesvari Z, Souvignet M, Bachofer F, Braun A, Garschagen M, Hagenlocher M (2021) Understanding and assessing flood risk in Vietnam: Current status, persisting gaps, and future directions. *J Flood Risk Manag* 14(2):e12689
- Noorbeh P, Roozbahani A, Kardan Moghaddam H (2020) Annual and monthly dam inflow prediction using Bayesian networks. *Water Resour Manage* 34:2933–2951
- Obahoundje S, Diedhiou A, Akpoti K, Kouassi KL, Ofosu EA, Kouame DGM (2024) Predicting climate-driven changes in reservoir inflows and hydropower in Côte d'Ivoire using machine learning modeling. *Energy* 302:131849
- Paryani S, Neshat A, Pradhan B (2021) Improvement of landslide spatial modeling using machine learning methods and two Harris hawks and bat algorithms. *Egypt J Remote Sens Space Sci* 24(3):845–855
- Peng K, Feng K, Chen B, Shan Y, Zhang N, Wang P, Li J (2023) The global power sector's low-carbon transition may enhance sustainable development goal achievement. *Nat Commun* 14(1):3144
- Poul AK, Shourian M, Ebrahimi H (2019) A comparative study of MLR, KNN, ANN and ANFIS models with wavelet transform in monthly stream flow prediction. *Water Resour Manage* 33(8):2907–2923
- Puttinaovarat S, Horkaew P (2020) Flood forecasting system based on integrated big and crowdsource data by using machine learning techniques. *IEEE Access* 8:5885–5905
- Rahimzad M, Moghaddam Nia A, Zolfanoon H, Soltani J, Danandeh Mehr A, Kwon HH (2021) Performance comparison of an LSTM-based deep learning model versus conventional machine learning algorithms for streamflow forecasting. *Water Resour Manage* 35(12):4167–4187
- Sahin ME, Ozbay Karakus M (2024) Smart hydropower management: utilizing machine learning and deep learning method to enhance dam's energy generation efficiency. *Neural Comput Appl* 36:1–17
- Samadinfard S, Jarhan S, Salwana E, Mosavi A, Shamshirband S, Akib S (2019) Support vector regression integrated with fruit fly optimization algorithm for river flow forecasting in Lake Urmia Basin. *Water* 11(9):1934
- Sammen SS, Jalut QH, Nama AH (2020) Hydrological study and analysis for proposed Al-Arkhama Dam, Iraq. In *IOP Conference Series: Materials Science and Engineering* (Vol. 737, No. 1, p. 012160). IOP Publishing
- Seifi A, Riahi-Madvar H (2019) Improving one-dimensional pollution dispersion modeling in rivers using ANFIS and ANN-based GA optimized models. *Environ Sci Pollut Res* 26:867–885
- Sharma B, Goel NK (2024) Streamflow prediction using support vector regression machine learning model for Tehri Dam. *Appl Water Sci* 14(5):1–20
- Shu X, Ding W, Peng Y, Wang Z (2024) Value of long-term inflow forecast for hydropower operation: a case study in a low forecast precision region. *Energy* 298:131218
- Stefenon SF, Ribeiro MHD, Nied A, Yow KC, Mariani VC, dos Santos Coelho L, Seman LO (2022) Time series forecasting using ensemble learning methods for emergency prevention in hydroelectric power plants with dam. *Electric Power Syst Res* 202:107584
- Tabbussum R, Dar AQ (2021) Comparison of fuzzy inference algorithms for stream flow prediction. *Neural Comput Appl* 33(5):1643–1653
- Tikhmarine Y, Souag-Gamane D, Kisi O (2019) A new intelligent method for monthly streamflow prediction: hybrid wavelet support vector regression based on grey wolf optimizer (WSVR-GWO). *Arab J Geosci* 12:1–20
- Turner SW, Doering K, Voisin N (2020) Data-driven reservoir simulation in a large-scale hydrological and water resource model. *Water Resour Res* 56(10):e2020WR027902
- Vapnik V (1998) The support vector method of function estimation. In: Suykens JAK, Vandewalle J (eds) *Nonlinear modeling*. Springer, Boston, pp 55–85
- Wang S, Yu L, Tang L, Wang S (2011) A novel seasonal decomposition based least squares support vector regression ensemble learning approach for hydropower consumption forecasting in China. *Energy* 36(11):6542–6554
- Wu JS, Han J, Annambhotla S, Bryant S (2005) Artificial neural networks for forecasting watershed runoff and stream flows. *J Hydrol Eng* 10(3):216–222
- Zhang X, Wang H, Peng A, Wang W, Li B, Huang X (2020) Quantifying the uncertainties in data-driven models for reservoir inflow prediction. *Water Resour Manage* 34:1479–1493
- Zolfaghari M, Golabi MR (2021) Modeling and predicting the electricity production in hydropower using conjunction of wavelet transform, long short-term memory and random forest models. *Renew Energy* 170:1367–1381

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