

Adnan, Rana Muhammad et al.

Article — Published Version

Comparison of improved relevance vector machines for streamflow predictions

Journal of Forecasting

Provided in Cooperation with:

John Wiley & Sons

Suggested Citation: Adnan, Rana Muhammad et al. (2023) : Comparison of improved relevance vector machines for streamflow predictions, Journal of Forecasting, ISSN 1099-131X, Wiley, Hoboken, NJ, Vol. 43, Iss. 1, pp. 159-181, <https://doi.org/10.1002/for.3028>

This Version is available at:

<https://hdl.handle.net/10419/288211>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc/4.0/>

RESEARCH ARTICLE

Comparison of improved relevance vector machines for streamflow predictions

Rana Muhammad Adnan¹  | Reham R. Mostafa² | Hong-Liang Dai¹ | Ehsan Mansouri³ | Ozgur Kisi^{4,5}  | Mohammad Zounemat-Kermani⁶

¹School of Economics and Statistics, Guangzhou University, Guangzhou, China

²Department of Information Systems, Faculty of Computers and Information Sciences, Mansoura University, Mansoura, Egypt

³Department of Computer and Technology, Birjand University of Medical Sciences, Birjand, Iran

⁴Department of Civil Engineering, Lübeck University of Applied Science, Lübeck, Germany

⁵Department of Civil Engineering, School of Technology, Ilia State University, Tbilisi, Georgia

⁶Department of Water Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

Correspondence

Hong-Liang Dai, School of Economics and Statistics, Guangzhou University, Guangzhou, 510006, China.
Email: hldai618@gzhu.edu.cn

Ozgur Kisi, Department of Civil Engineering, School of Technology, Ilia State University, 0162, Tbilisi, Georgia.
Email: ozgur.kisi@th-luebeck.de

Funding information

National Social Science Foundation of China, Grant/Award Number: 18BTJ029; Key Projects Of National Statistical Science Research Projects, Grant/Award Number: 2020LZ10; Guangdong Natural Science Research Projects, Grant/Award Number: 2023A1515011520; Guangzhou Municipal Education Bureau, Grant/Award Number: 202235324

Abstract

This study investigates the feasibility of relevance vector machine tuned with dwarf mongoose optimization algorithm in modeling monthly streamflow. The proposed method is compared with relevance vector machines tuned by particle swarm optimization, whale optimization, marine predators algorithms, and single relevance vector machine methods. Various lagged values of hydroclimatic data (e.g., precipitation, temperature, and streamflow) are used as inputs to the models. The relevance vector machine tuned with dwarf mongoose optimization algorithm improved the efficiency of single method in monthly streamflow prediction. It is found that the integrating metaheuristic algorithms into single relevance vector machine improves the prediction efficiency, and among the input combinations, the lagged streamflow data are found to be the most effective variable on current streamflow whereas precipitation has the least effect.

KEYWORDS

dwarf mongoose optimization algorithm, hydroclimatic data, relevance vector machine, streamflow prediction

1 | INTRODUCTION

Streamflow prediction has been one of the most complex and challenging problems in hydrological systems, such as water resources management, assessment of risk, controlling natural disasters (viz., flash floods), and

allocation of water due to the nonlinear distribution of river flow patterns (Bennett et al., 2017; Goshime et al., 2020; Kisi & Cigizoglu, 2007; Tyralis et al., 2021). Even though there are various methodologies for dealing with the streamflow prediction problem, it is worth mentioning that over the last two decades, machine learning

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2023 The Authors. *Journal of Forecasting* published by John Wiley & Sons Ltd.

(ML) models (like artificial neural networks) have gained more attention than physically based, numerical, or traditional statistical techniques to enhance streamflow prediction accuracy and precision as a result of their adaptability (Hasanpour Kashani et al., 2015). Ideally, supervised ML methods can learn the connection between input and output datasets and then save the knowledge for predicting the unseen outputs (Hussain et al., 2020). In this regard, ML modeling has been expanded and applied successfully to predict streamflow in the past few years (Malik et al., 2020; Sayari et al., 2022; Zhou et al., 2018; Zounemat-Kermani et al., 2021).

Among the well-known ML models, support vector machines (SVMs) are considered powerful alternative kernel-based models developed by Cortes and Vapnik (1995) for classification problems and later extended to prediction model learning as support vector regression (SVR) models (Smola & Schölkopf, 2004). A few years later, Tipping (2001) introduced the relevance vector machines (RVMs) to alter the primary idea of the SVM in a Bayesian context to overcome the potential weaknesses of the SVM. The fundamental function of RVM models is data center, which produces equivalent and often superior results than kernel-based learning machines. Several successful modeling cases have been reported in the literature for the application of RVM (Safari et al., 2022; Tao et al., 2021). Examples include Bui et al. (2018) who proposed an integrated approach according to the RVM to assess landslide susceptibility. They also used the experimental data, and the SVM and logistic regression results to evaluate their model's effectiveness. The performance of the proposed model was promising in both the training and testing phases.

Constructing and developing robust and efficient ML models highly depend on their optimization algorithm during the training and validation process. In light of this important issue, nature-inspired optimization algorithms have become popular to focus on solving complex optimization issues like multidimensional and multimodal problems. These models are highly efficient with less computational effort and time consumption. In accordance with the context of this study, several nature-inspired optimization algorithms (e.g., particle swarm optimization [PSO], whale optimization algorithm [WOA], marine predators algorithm (MPA), and dwarf mongoose optimization algorithm [DMOA]) are considered for literature review.

Kennedy and Eberhart (1995) developed the PSO algorithm based on swarm behavior like fish, bird flocking, and schooling. Su et al. (2022) applied the PSO algorithm in the sewage treatment system to improve energy efficiency. The optimal control solution based on the PSO algorithm was proposed, which can be applied to save

electrical energy. Wang et al. (2022) used the PSO algorithm to optimize regional water resource allocation in Yinchuan city of Ningxia, China. The allocation schemes of water resources are simulated in three different precipitation scenarios in 2025. Their results indicate the practical significance of their model.

Mirjalili and Lewis (2016) proposed a swarm intelligence established on humpback whales' hunting behavior, called whale optimization algorithm (WOA). They tested the algorithm with 29 mathematical and 6 structural optimization problems and proved that WOA is more efficient than other nature-inspired algorithms and conventional methods. Mohammadi and Mehdizadeh (2016) modeled reference evapotranspiration on a daily scale with a novel hybrid SVR model with a WOA algorithm. They combined approaches such as random forests (RF), relief, Pearson's correlation, and principal component analysis with the SVR model. The results demonstrated that the hybrid RF, SVR, and WOA models performed better than other models. Liu et al. (2022) presented WOA-based point cloud data (WOAPCD) for inspection of sewage systems. This model was validated with actual sewer datasets and reconstructed the 3D model accurately. The proposed model proved more efficient than PSO in modeling speed and fitting error.

Faramarzi et al. (2020) presented the MPA inspired by the behavior of predators in attacking their prey. This algorithm was tested by 58 mathematical benchmark functions and 3 different types of optimization algorithms like genetic algorithm and PSO, gravitational search algorithm, cuckoo search, as well as other recently developed algorithms (e.g., salp swarm algorithm [SSA]). The results of the MPA were more accurate than most of the applied algorithms. In another study, Abdel-Basset et al. (2021) applied the MPA method for solving multi-objective optimization by introducing four variants of the MPA. The models introduced were mentioned as multiobjective marine predators' algorithm (MMPA), a combination of dominance strategy based on exploration-exploitation and MPA (M-MMPA), a hybrid model consisting of M-MMPA and Gaussian-based mutation, and an integrated model of Nelder–Mead simplex method with M-MMPA (M-MMPA-NMM). All the proposed models performed well compared with the studied multi-objective optimizer. DMOA (DMO) is another new nature-inspired algorithm developed by Agushaka et al. (2022) to find a solution to the classical and CEC 2020 benchmark functions and 12 semi-real constrained engineering optimization criteria. The social behavior of dwarf mongooses was imitated in the DMO method. This model was compared with seven optimization algorithms and performed better in five statistical indices.

The principal aim of this paper is threefold: (1) the enhancement of streamflow prediction using a novel integrated strategy by embedding RVM with nature-inspired metaheuristic methods such as PSO, WOA, MPA, and DMOA; (2) evaluating the suggested techniques as hybrid models to its individual counterpart, RVM; (3) comparing the efficiency of the models based on previous streamflow data and a wide range of statistical procedures. To the best of our knowledge, none of the previous studies have investigated the above-mentioned integration of RVM with the proposed optimization techniques, particularly, the recent MPA and DMOA methods.

2 | CASE STUDY

In this study, Neelum Basin is selected as case study (Figure 1). Neelum catchment is located in northern areas of Pakistan with coordinates of 73.24° – 75.23° E and latitudes 34.14° – 35.8° N. The basin has total drainage area of 7438 km^2 with elevation variation of 753 – 5724 m and 144 km long basin span length. Basin is selected due to its geographically important location. Neelum River is an transboundary river which originates from Kahmir Valley in the Indian region and called as Kishanganga River in India. Basin 69% area is in Pakistan whereas 31% area lies in Indian side. Precipitation occurred in the basin in form of snow in the snow form and basin have key glaciers such as Saranwali, Shonthar, Parbat,

Dewarian, Rati Gali, and Mianwitch. Snowmelt from these glaciers contributes a big portion in the summer runoff. Therefore, streamflow forecasting in this basin is challenging due to less information and data availability about snowmelt. The basin also has key importance due to its contribution to the hydropower generation capacity of Pakistan. It can be seen from the figure that the Neelum basin is the largest upstream tributary of Mangla Catchment that have Pakistan's second largest dam, that is, Mangla dam with a generation capacity of 1150 MW . In addition, recently, Pakistan built another dam in the Neelum basin with a generation capacity of 969 MW , that is, called as Neelum–Jhelum Hydropower Plant. Therefore, accurate estimation of river flows of this basin is very necessary for water management in this key basin. In addition, precise forecasting of streamflow of this basin can aid the hydropower engineer in proper scheduling and operation of both reservoirs. For the application of selected ML models to forecast streamflows of the Neelum basin, monthly streamflow, precipitation, and air temperature data of Muzaffarabad gauging and the climatic station is collected from the Water And Power Development Authority (WAPDA) in Pakistan for the duration of 1971–2010. For the selected models' application, the obtained 480 months are divided into the ratio of 75% data (360 months) for training and of 25% data (120 months) for testing purposes. The basic statistical summary of the climatic and streamflow data is reported in Table 1.

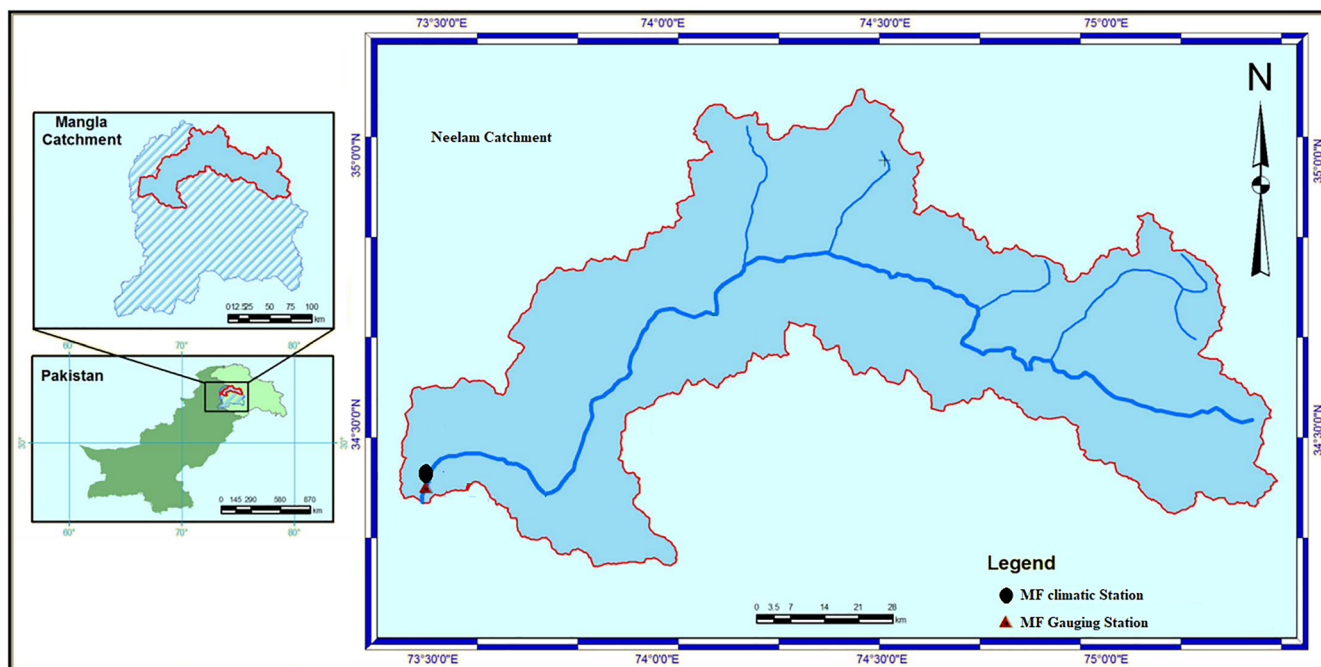


FIGURE 1 Study area.

	Mean	Min.	Max	Skewness	Std. dev.
Precipitation					
Whole dataset	123.9	0	732.3	1.918	114.2
Training dataset	126.1	0	732.3	1.903	117.4
Testing dataset	117.5	0	613.8	1.938	103.9
Temperature					
Whole dataset	20.60	7.132	34.66	-0.219	7.193
Training dataset	20.46	8.118	34.66	-0.184	7.216
Testing dataset	21.04	7.132	30.52	-0.329	7.107
Streamflow					
Whole dataset	324.6	33.64	1319.9	1.063	300.6
Training dataset	336.6	41.45	1319.9	1.020	311.6
Testing dataset	288.9	33.64	1064.1	1.129	261.9

TABLE 1 The statistical parameters of the applied data.

3 | METHODS

3.1 | ML using metaheuristics

Through the use of ML techniques, metaheuristics can extract valuable information from created data. Such knowledge is incorporated into the search process to help metaheuristics make more intelligent decisions, which further improves the quality and robustness of their solutions. In ML, algorithms are developed to infer facts from data and learn new tasks. Recently, metaheuristics and ML techniques have been integrated extensively (Karimi-Mamaghan et al., 2022).

A metaheuristic optimization algorithm combines heuristics with a priori knowledge about the problem origin in order to solve a wide class of optimization problems. An evolutionary algorithm and a swarm intelligence algorithm are two types of metaheuristic algorithms. During the extensive research in evolutionary and swarm intelligence computation, numerous hybrid optimization algorithms have been proposed that combine major features of existing optimization algorithms. The development of unified models was driven in part by a great variety of metaheuristic algorithms and their common characteristics (stochastic behavior and population structure).

A majority of these search and optimization methods are inspired by a particular event or process in nature as the source of inspiration. In recent decades, researchers have developed robust metaheuristic techniques such as simulated annealing, evolutionary algorithms, PSO, and ant colony optimization as a result of identifying and formulating the similarities between algorithms and processes they are modeled on. It is important to deal with the uncertainty in supply and demand of water resources,

as well as how to optimally allocate them between refineries, recycling plants, and water production facilities when managing water supply and wastewater collection (Fathollahi-Fard et al., 2020; Sakib et al., 2021).

Water was allocated to agricultural, urban, and hydroelectric sectors in the Middle East using a linear programming model in 2003 (Goldman & Saykally, 2004). They identified sustainable allocations through a cooperative game theory that all stakeholders were willing to accept in this regard. Based on different amounts of energy prices and economic efficiency, their model determines the allocation of revenues generated by cooperation between players. It was the first time that a study was developed with multiobjective linear programming method for allocating water resources considering economic costs, wastewater reductions, and demand level optimizations (Fattahi & Fayyaz, 2010). In 2011, to cover all economic aspects of water distribution network allocation, a nonlinear mixed integer programming model was developed and used in another definitive optimization model (Verleye & Aghezzaf, 2011). Using an integrated reservoir management system to change reservoir operations according to climate change conditions, researchers developed a new reservoir operation system (Eum et al., 2012). Three methods were used to manage reservoirs, namely, the nearest neighbor model, the hydrological model, and the differential evolutionary optimization model. They used six probabilistic scenarios in their research. Based on their results, it has been demonstrated that the integrated management system can provide reservoir operation curves that are adaptive to future climate scenarios, reflecting the hydrological characteristics for future climate scenarios. Based on probabilistic scenarios, Kang and Lansey created a model that uses a multiobjective optimization platform. Water consumption was

expected to grow, resulting in conflicting goals (Kang & Lansey, 2013). A model based on game theory was used to study the effects of water scarcity and conflict on international conflict (Steinbrueckh, 2014). Researchers presented a scenario-based model in 2015 to predict water demand and the relationship between water pressure in pipes and its uncertainty (Pérez et al., 2015). An additional study developed a multiobjective scenario-based model that pursued three goals using nonlinear programming. Mortazavi-Naeini et al. (2015) utilized multiobjective scenario-based models to develop the model. There was a focus on minimizing all operating and structural costs of the water distribution system as well as the expected value of the system in light of climate change scenarios. Using probabilistic programming, Mo et al. (2015) analyzed river and groundwater scenarios. In a subsequent study, a model was developed for minimizing the cost of the water supply chain because of hydraulic uncertainty of water fluctuations (Schwartz et al., 2016). A Bayesian network model was proposed in 2021 to predict and analyze disasters that may occur in any water network, taking into account legal, environmental, safety, political, social, and economic factors (Sakib et al., 2021).

According to Rezaei and Safavi (2020), the GuASPSO algorithm is a novel variant of the PSO algorithm. To achieve global particle guidance, the local guide particles in this algorithm are all grouped into several clusters. GuASPSO's exploration and exploitation phases are balanced through gradual decreases of this chance as the iterations go on. To boost the chance of particles not being trapped in local optima, Liu et al. (2020) randomly perturbed the acceleration coefficients of the PSO, using Gaussian white noise with adjustable intensity. The RVM ML model will be used in this study in order to optimize parameters of this ML RVM model by implementing PSO, WOA, MPA, and DMOA algorithms. Below is a brief description of RVM's ML and meta heuristics algorithms.

3.2 | RVM

RVMs are ML algorithms that use Bayesian inference to achieve parsimonious regression and classification solutions. RVMs provide probabilistic classification and have the same functional form as the SVMs. With a certain covariance function, it is equivalent to a Gaussian process. RVM's Bayesian formulation avoids the free parameters of SVMs, which are usually optimized through cross-validation. However, RVMs are subject to local minima because they use an expectation maximization-like learning method. SVMs do not use sequential minimal optimization algorithms, which are guaranteed to

find a global optimum in standard optimization algorithms. It is possible for RVM to achieve a level of generalization accuracy comparable with that of the well-established support vector approach (and even better) while utilizing dramatically fewer kernel functions. In a practical implementation, this means a significant reduction in memory and computation. Besides the need to choose the type of kernel and any associated parameters, we also benefit from the absence of any additional nuisance parameters (Tipping, 2000). Basic architecture of RVM model is shown in Figure 2.

3.3 | PSO

PSO is a population-based stochastic optimization technique. PSOs begin by creating populations, or swarms, of particles. By placing the particles in corresponding positions, we can identify possible optimization solutions. This method is based on the concept that each particle has a main characteristic, which is its velocity (Adnan et al., 2021; Kennedy & Eberhart, 2022). Based on the equations below, we can calculate the particle i 's position and velocity at the epoch t (Adnan et al., 2021).

$$x_{i,t} = \{x_{i1}, x_{i2}, \dots, x_{in}\} \quad (1)$$

$$v_{i,t} = \{v_{i1}, v_{i2}, \dots, v_{in}\} \quad (2)$$

In this example, n represents the number of parameters. It is necessary to update particle position and velocity iteratively as follows:

$$v_{i,t+1} = X * ((v_{i,t} + c_1 r_1 (pbest_{i,t} - x_{i,t}) + c_2 r_2 (gbest_t - x_{i,t}))) \quad (3)$$

$$x_{i,t+1} = x_{i,t} + v_{i,t+1} \quad (4)$$

The next and current iterations, the particle's velocity, and position at iteration $t + 1$, t , $v_{i,t}$, and $x_{i,t}$, respectively. Thus, $pbest_{i,t}$ and $gbest_{i,t}$ represent an individual's previous and global best positions, respectively. Additionally, r_1 and r_2 are arbitrarily selected numbers that vary from 0 to 1, and c_1 and c_2 signal the influence of social and cognitive components. Finally, X is usually considered to be 0.729 (Menad et al., 2019) as a convergence factor. PSO's final step is to update the best position and determine the best individual from the entire swarm for each individual. Using the equations below will minimize the problem (Adnan et al., 2021). Figure 3 explains the basic flow chart of PSO algorithm.

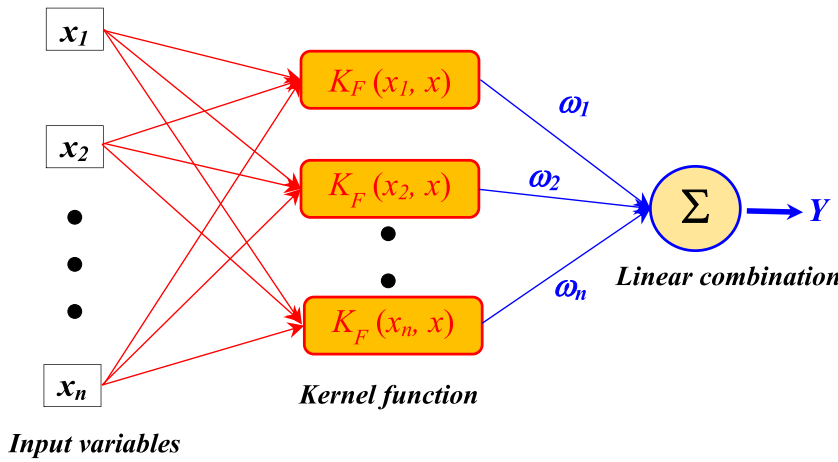


FIGURE 2 Relevance vector machine architecture.

$$pbest_{i,t+1} = \begin{cases} pbest_{i,t}, & \text{if } f(pbest_{i,t}) \leq f(x_{i,t+1}) \\ x_{i,t+1}, & \end{cases} \quad (5)$$

$$gbest_{t+1} = \min\{f(pbest_{i,t+1})\} \quad (6)$$

Other variable is defined by two basic functions with a range of inputs. Here is the variable Y , which is mapped from variable X using threshold c .

$$Y = \max(0, X - c) \quad (7)$$

$$Y = \max(0, c - X) \quad (8)$$

3.4 | WOA

Based on whale bubble net feeding behavior, WOA (Mirjalili & Lewis, 2016) is a metaheuristic swarm intelligence optimization algorithm introduced in 2016. Figure 4a illustrates how the whale approaches the prey by gradually contracting and encircling the prey over the spiral path. In order to prevent the prey from swimming through the bubble net, the whale continuously spits bubbles upward as it approaches (Fan et al., 2021). According to Figure 4b, WOA parameters can be optimized by three mechanisms: encircling, spiral updating, and global searching. WOA mainly consists of two steps: (1) surround shrinkage and (2) spiral update. Here is a brief description of each process.

3.4.1 | Surround shrinkage

This stage also sets the optimal unit in the population as the target as individual members update their positions

and approach the prey. Modeling in mathematics looks like this (Fan et al., 2021):

$$X(t+1) = X_p(t) - A \cdot D \quad (9)$$

There are two coefficient vectors, \mathbf{A} and \mathbf{B} , and there whales are identified by their position vectors, $X(t)$; t is the number of iterations now happening; $X_p(t)$ is the position vector of the prey; $X(t)$ is the position vector of the whale; and X is the position of the whale.

$$\begin{cases} \mathbf{A} = 2a \cdot r_1 - a \\ \mathbf{B} = 2r_2 \end{cases} \quad (10)$$

The random numbers r_1 and r_2 are in the range $[0, 1]$, and a represents the control parameter. Each iteration decreases a from 2 to 0, namely,

$$a = \left(1 - \frac{t}{t_{\max}}\right) \quad (11)$$

where maximum iterations are expressed as t_{\max} .

3.4.2 | Spiral update

Spiral update mechanism is mathematically described as follows (Fan et al., 2021):

$$X(t+1) = X_p(t) + D' \cdot e^{bl} \cdot \cos(2\pi l) \quad (12)$$

$$D' = |X_p(t) - X(t)| \quad (13)$$

where D' represents the distance between the search target and the search individual, b represents the shape parameter, and l represents the random number between

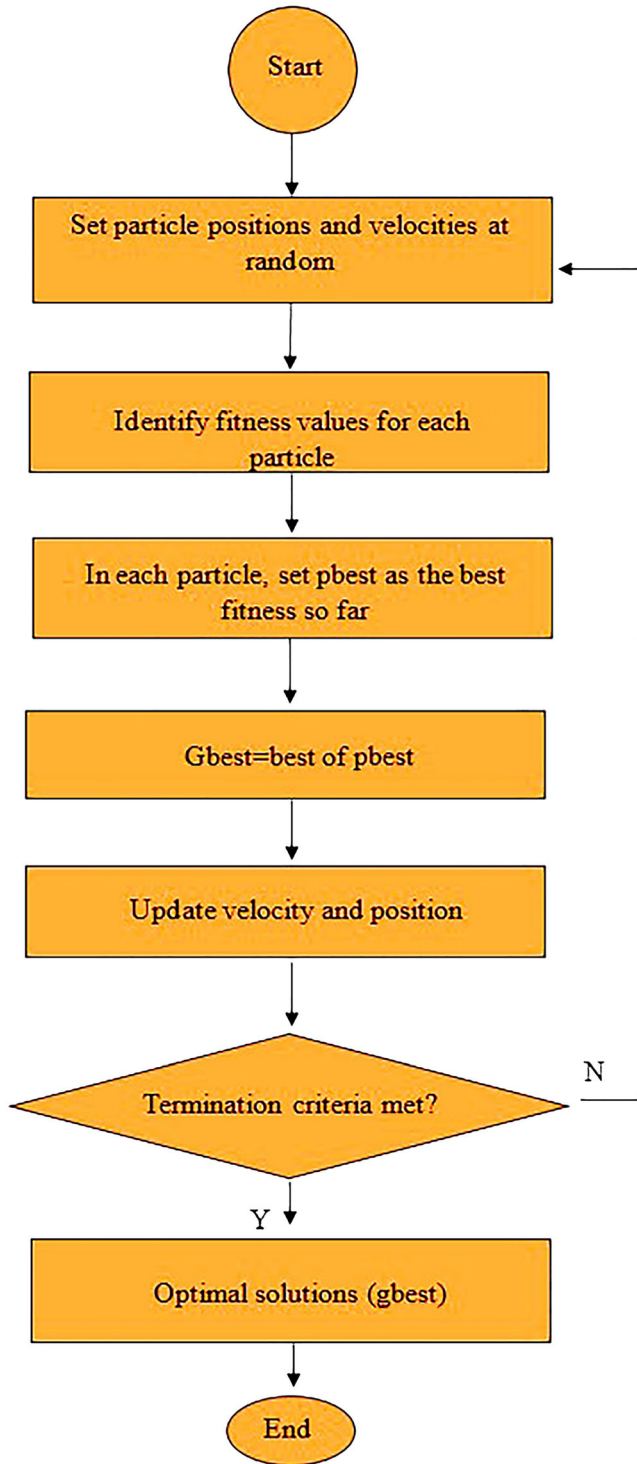


FIGURE 3 Flow chart of the particle swarm optimization algorithm (Roy et al., 2022).

$[-1, 1]$. However, the process of whales constricting the circular path and the process of whales advancing along the circular path are synchronized during hunting. Based on this, the probability p of group encirclement and spiral renewal are equal, each of which is 0.5, as shown below (Fan et al., 2021):

$$X(t+1) = \begin{cases} X_p(t) - A \cdot |B \cdot X_p(t) - X(t)|, & p \leq 0.5 \\ X_p(t) + D' \cdot e^{bl} \cos(2\pi l), & p > 0.5 \end{cases} \quad (14)$$

A random number between 0 and 1 is used as p . The whale's hunting behavior is modeled at this stage. A randomly selected search individual is used to update the search location, not the optimal search individual. The random selection of search individuals is then removed from the global search so that other individuals can be found. Here is the mathematical model:

$$X(t+1) = X_{rand}(t) - A \cdot |B \cdot X_{rand}(t) - X(t)| \quad (15)$$

where $X_{rand}(t)$ is the optimal individual position randomly selected by the population.

3.5 | MPA

Foraging strategies such as Lévy's and Brown's movements in ocean predators as well as encounter rate policies in biological interactions between predators and prey are the major inspirations for MPA (Faramarzi et al., 2020). Foraging strategy and encounter rate policy in marine ecosystems are determined by the rules that naturally govern optimal foraging strategies.

3.5.1 | MPA formulation

In compliance with most metaheuristics, MPA utilizes a population-based approach by which the initial solution is disposed across the search space uniformly as a first step in order to reach a better solution:

$$X_0 = X_{\min} + rand(X_{\max} - X_{\min}) \quad (16)$$

Assume that X_{\min} and X_{\max} represent lower and upper bounds for variables and $rand$ represents a uniform random vector between 0 and 1. Top predators are said to be better at foraging in nature based on the survival of the fittest theory. A matrix called Elite is constructed by selecting the fittest solution as the top predator. Using the information about the prey's position, these arrays search and find the prey (Faramarzi et al., 2020).

$$Elite = \begin{pmatrix} X_{1,1}^I & X_{1,2}^I & \dots & X_{1,d}^I \\ X_{2,1}^I & X_{2,2}^I & \dots & X_{2,d}^I \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ X_{n,1}^I & X_{n,2}^I & \dots & X_{n,d}^I \end{pmatrix}_{n \times d} \quad (17)$$

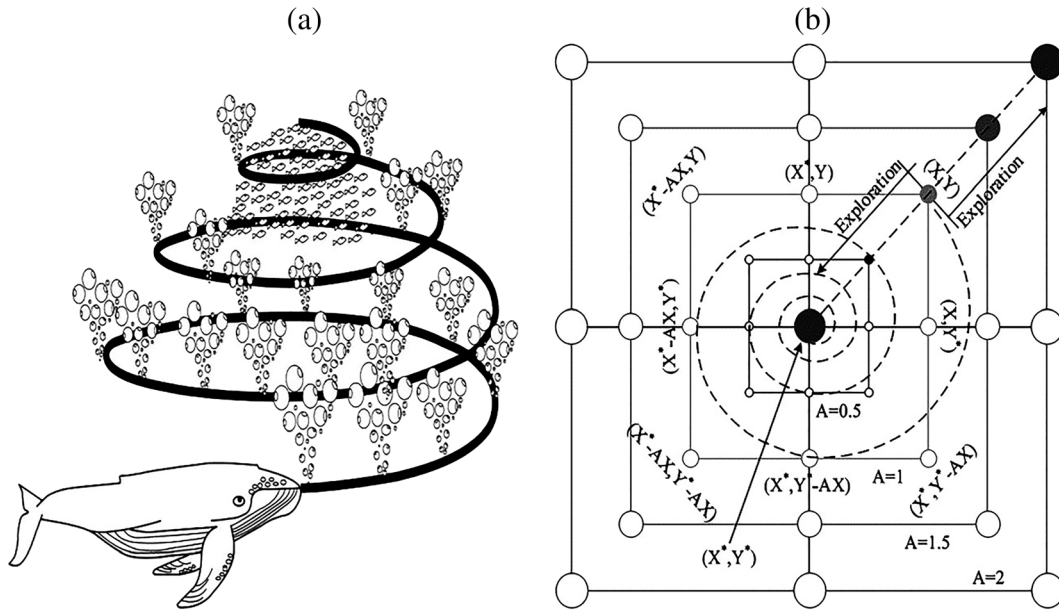


FIGURE 4 Whale optimization algorithm (a) feeding behavior; (b) parameter optimization mechanism (Fan et al., 2021).

As a result of replicating X_i n times, the Elite matrix is constructed representing the top predator vector. Search agents are counted by n while dimensions are counted by d . Search agents include both predators and prey. Predators look for their prey at the same time as their prey looks for its own food. The Elite will be updated at the end of each iteration if a better predator replaces the top predator. The predators update their positions based on another matrix called Prey, which has the same dimension as Elite. A simple explanation could be that after initialization, the fittest of the Prey (predator) becomes the Elite. Here is the Prey:

$$\text{Prey} = \begin{pmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,d} \\ X_{2,1} & X_{2,2} & \dots & X_{2,d} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ X_{n,1} & X_{n,2} & \dots & X_{n,d} \end{pmatrix}_{n \times d} \quad (18)$$

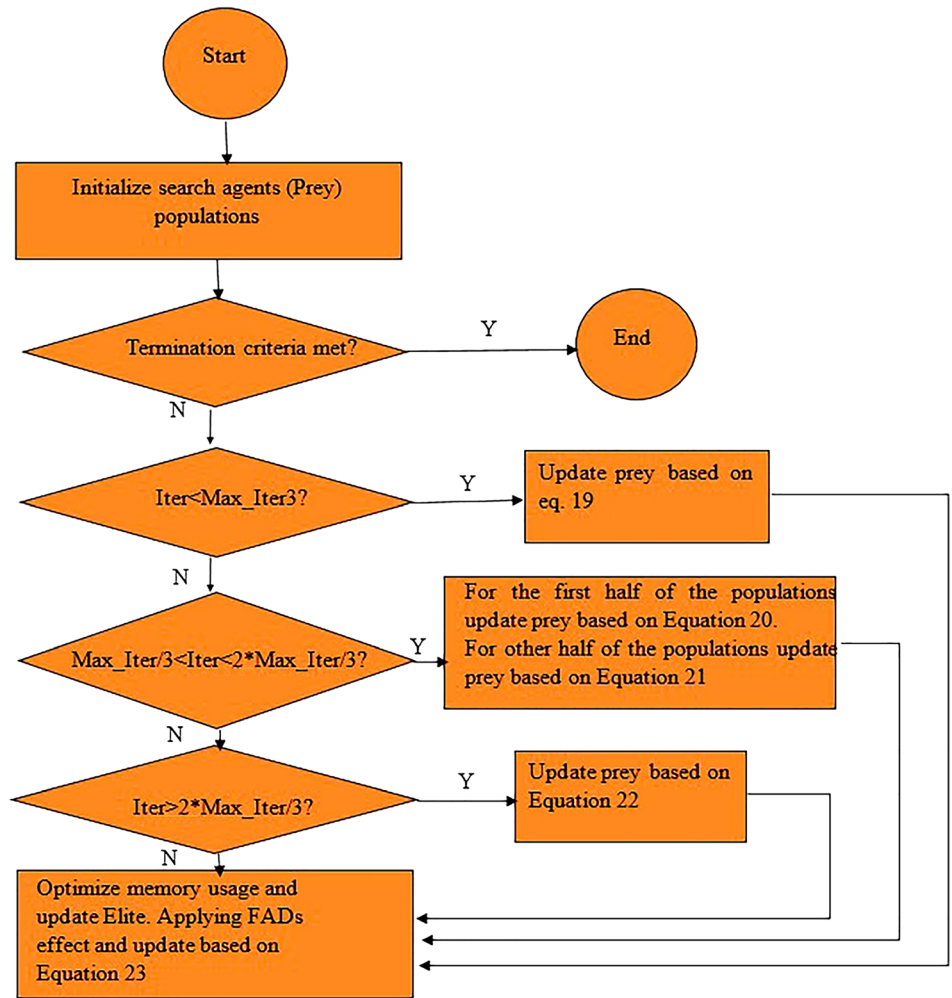
The dimensions of the j th prey are given by X_{ij} in Equation (18). This two matrices are mainly and directly involved in the entire optimization process. Figure 5 illustrates the flowchart of MPA.

3.5.2 | MPA exploration phase

Prey moves in Brownian motion during the first phase of optimization. Due to the uniform distribution of preys

within the domain in its initial iterations and the relatively large distance between predators and preys, Brownian motion can allow preys to explore their neighborhoods separately, resulting in a better exploration. Preys with new positions are then evaluated for fitness, and if the new position is more fitting than the old one, it is replaced. Prey can interpret their fitted positions as abundance food areas, and the saving procedure can be interpreted as their memory recalling abundance food areas. A predator is considered to be more successful at foraging for its food than its prey. By this, it means that if the fitness value of the prey is better than that of the top predator, the prey will replace it. As prey search for food, predators are now ready to start foraging. Optimization begins here in the second phase. Exploration into exploitation is the purpose of this phase. Both predators and prey search for food in this phase to take advantage of exploration and exploitation. Exploration is managed by half the population, while exploitation is handled by the other half. In Brownian motion, the predator searches for its prey and takes a long jump if it cannot find any food in the immediate neighborhood. In Lévy motion, the prey searches its immediate neighborhood, and if it cannot find food, it switches to the Lévy strategy. FADs combined with Lévy strategy long steps greatly help MPA avoid local optima stagnation and improve its performance since the predator and prey locations are relatively close and the step length is shorter than the previous phase. The algorithm needs high exploitation capabilities as it reaches its final stage of optimization. To more efficiently search a certain neighborhood, the predator switches from Brownian to Lévy strategy in this phase.

FIGURE 5 Flow chart of marine predators algorithm. FAD, fish aggregating device.



During this phase, predators are able to use adaptive defined convergence factors (CFs) to limit their searches within a particular neighborhood in order to avoid wasting time and energy searching in nonpromising regions of the domain due to the long step sizes of the Lévy strategy.

3.5.3 | MPA exploitation phase

Optimization in MPA is conducted in three main phases that consider different velocity ratios while simulating the whole life cycle of a predator and prey. When the predator is moving faster than the prey (1), when the predator and prey are moving at approximately the same speed (2), and when the predator is moving faster than the prey (3). Phases are defined and assigned a specific iteration period. A set of steps that mimic the movement of predators and prey in nature is based on rules that govern the nature of predators and prey movement.

Step 1: If a predator is moving faster than a prey or when there is a high velocity ratio. A scenario such as this occurs in the early stages of optimization, when exploration plays a particularly important role. The best strategy for predators in high-velocity ratios ($v \geq 10$) is not to move at all. This rule can be mathematically represented as follows:

$$\begin{aligned} \overrightarrow{stepsize}_i &= \overrightarrow{R}_B \otimes \left(\overrightarrow{Elite}_i - \overrightarrow{R}_B \otimes \overrightarrow{Prey}_i \right) \\ \overrightarrow{Prey}_i &= \overrightarrow{Prey}_i + P \cdot \overrightarrow{R} \otimes \overrightarrow{stepsize}_i \end{aligned} \quad (19)$$

Based on the Brownian motion model, R_B represents random numbers based on normal distributions. Enterwise multiplications are represented by the notation \otimes . Prey's movement is simulated by multiplying rb by prey. R_B is a vector of uniform random numbers in $[0,1]$ with $P=0.5$ as a constant number. Typically, this scenario occurs in the first third of iterations for exploration

abilities that are high. A current iteration is called *Iter*, while a maximum iteration is called *Max_iter*.

Step 2: When predators and prey move at the same pace. Both of them appear to be searching for their prey. In this section, the exploration is transiently converted into exploitation during the intermediate stage of optimization. In this phase, both exploration and exploitation matters. In this phase, prey is responsible for exploitation and predator for exploration.

$$\begin{aligned}\overrightarrow{stepsize}_i &= \overrightarrow{R}_L \otimes (\overrightarrow{Elite}_i - \overrightarrow{R}_L \otimes \overrightarrow{Prey}_i) \\ \overrightarrow{Prey}_i &= \overrightarrow{Prey}_i + P \cdot \overrightarrow{R} \otimes \overrightarrow{stepsize}_i\end{aligned}\quad (20)$$

\overrightarrow{R}_L represents the Lévy movement through a vector of random numbers based on the Lévy distribution. In light of the fact that most Levy distribution step sizes are associated with small steps, this section offers a useful tool for exploiting the distribution. This study assumes the following for the second half of the populations:

$$\begin{aligned}\overrightarrow{stepsize}_i &= \overrightarrow{R}_B \otimes (\overrightarrow{Elite}_i - \overrightarrow{Prey}_i) \\ \overrightarrow{Prey}_i &= \overrightarrow{Elite}_i + P \cdot CF \otimes \overrightarrow{stepsize}_i\end{aligned}\quad (21)$$

For predator movements, CF controls step size as an adaptive parameter.

Step 3: In cases where predators are moving faster than prey, or when the predators have a low velocity ratio. Usually, this scenario occurs in the last phase of optimization, when the exploitation capability is high.

$$\begin{aligned}\overrightarrow{stepsize}_i &= \overrightarrow{R}_L \otimes (\overrightarrow{R}_L \otimes \overrightarrow{Elite}_i - \overrightarrow{Prey}_i) \\ \overrightarrow{Prey}_i &= \overrightarrow{Elite}_i + P \cdot CF \otimes \overrightarrow{stepsize}_i\end{aligned}\quad (22)$$

3.5.4 | Eddy formation and fish aggregating devices' effect

Eddy formation or fish aggregating devices (FADs) are two other factors that cause behavioral changes in marine predators. In search space, FADs are considered local optima, and their effect is to trap users at these points. In order to avoid stagnation in local optima,

these longer jumps should be considered during simulation. As a result, the FADs effect can be mathematically expressed as follows:

$$\overrightarrow{Prey}_i = \begin{cases} \overrightarrow{Prey}_i + CF [\overrightarrow{X}_{min} + \overrightarrow{R} \otimes (\overrightarrow{X}_{max} - \overrightarrow{X}_{min})] \otimes \overrightarrow{U} & \text{if } r \leq \text{FADs} \\ \overrightarrow{Prey}_i + [FADs(1-r) + r] (\overrightarrow{Prey}_{r1} - \overrightarrow{Prey}_{r2}) & \text{if } r > \text{FADs} \end{cases}\quad (23)$$

In which FADs = 0.2 is the probability that FAD will affect the optimization process. The binary vector \mathbf{U} contains arrays of zeros and ones. A random vector is generated from [0,1] and its array is transformed into a zero-valued array if it is less than 0.2 and a one-valued array if it is greater than 0.2. The uniform random number r is generated from [0,1]. The lower and upper bounds of the dimensions are represented by the vectors X_{min} and X_{max} . Prey matrix random indexes are indicated by the subscripts $r1$ and $r2$.

3.6 | DMOA

By simulating dwarf mongoose behavioral compensation, the proposed algorithm can be used to simulate this process. The compensatory behavior adaptations include reduction of prey size, use of babysitters, and seminomadic lifestyle. For the purpose of implementing our model, the dwarf mongoose social structure is stratified into three groups: the alpha group, scouts, and babysitters. A seminomadic way of life is a result of compensation behavioral adaptations among the groups, which leads to a territory large enough for the entire group to live in harmony. Figure 6 illustrates the flowchart of DMOA.

A general framework for DMOs (Agushaka et al., 2022) to formulate the optimization processes is described below.

3.6.1 | Alpha group

After the population has been initialized, each solution's fitness is calculated. According to Equation (24), the alpha female (α) is selected based on the probability value for each population fitness (Agushaka et al., 2022).

$$\alpha = \frac{fit_i}{\sum_{i=1}^n fit_i}\quad (24)$$

As a result, the n -bs of the alpha group is the number of mongooses. Babysitters are referred to as *bs*. *Peep* is a

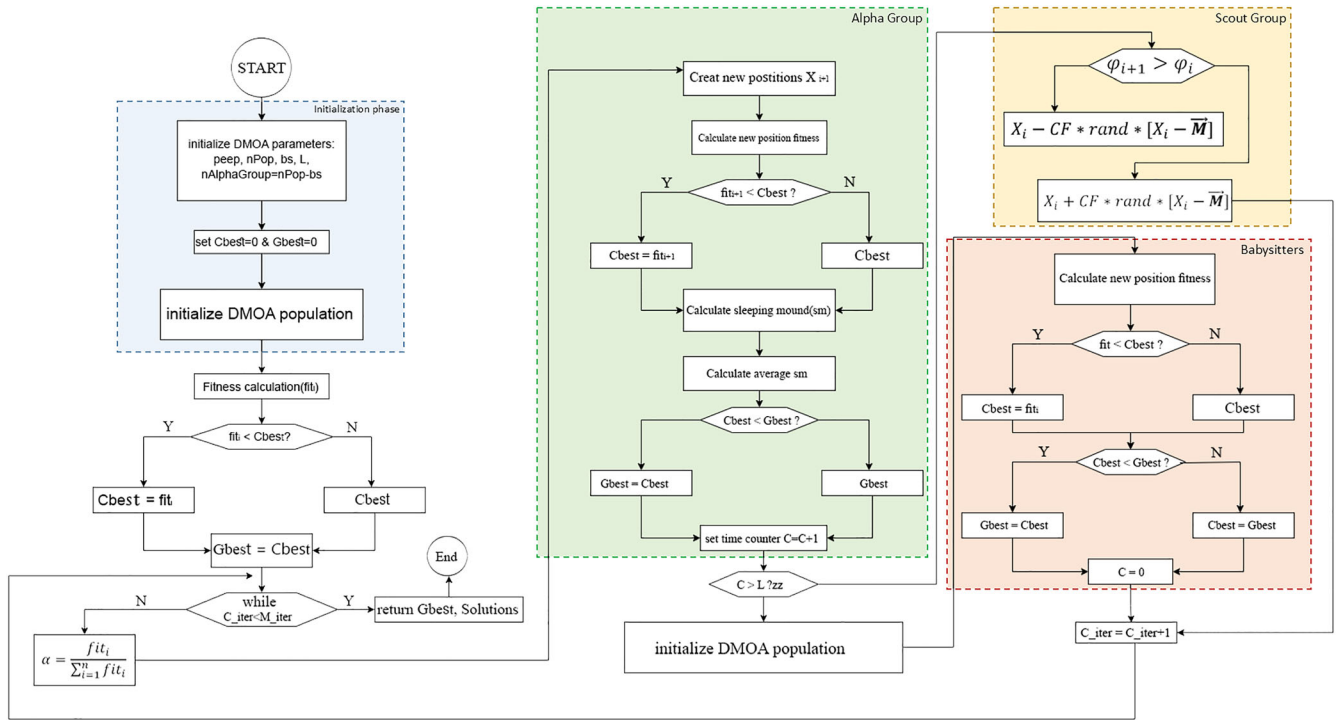


FIGURE 6 Flow chart of dwarf mongoose optimization algorithm.

vocalization made by the alpha female in order to keep the family on a path. Initially, every mongoose sleeps in the sleeping mound, \emptyset . DMOs use Equation (25) to produce candidate food positions.

$$X_{i+1} = X_i + \Phi * peep \quad (25)$$

where Φ is a uniformly distributed random number $[-1, 1]$, after every iteration, the sleeping mound is as given in Equation (26).

$$sm_i = \frac{fit_{i+1} - fit_i}{\max\{|fit_{i+1}|, |fit_i|\}} \quad (26)$$

Based on Equation (27), we can determine the average value of the sleeping mounds found.

$$\varphi = \frac{\sum_{i=1}^n sm_i}{n} \quad (27)$$

Once the babysitter exchange criterion is met, the algorithm moves into the scouting phase.

3.6.2 | Scout group

Because mongooses do not return to previous sleeping mounds, the scouts look for the next sleeping mound.

For our model, scouting and foraging occur simultaneously (Chou & Truong, 2021). It is modeled as an attempt to find a new sleeping mound that is evaluated by its success or failure. Thus, the movement of the mongooses is determined by their overall performance. A new sleeping mound will be discovered if the family forages far enough. In Equation (28), a mongoose scout is simulated (Agushaka et al., 2022).

$$X_{i+1} = \begin{cases} x_i - CF * \Phi * rand * [X_i - \bar{M}] & \text{if } \varphi_{i+1} > \varphi_i \\ x_i + CF * \Phi * rand * [X_i - \bar{M}] & \text{else} \end{cases} \quad (28)$$

where, rand is a random number between $[0, 1]$, With each iteration.

$CF = \left(1 - \frac{iter}{Max_{iter}}\right) \left(2 - \frac{iter}{Max_{iter}}\right)$ decreases linearly as the group's collective-volitive movement is controlled. $\bar{M} = \sum_{i=1}^n \frac{X_i * sm_i}{X_i}$ determines the direction in which the mongoose moves to the new sleeping mound.

3.6.3 | The babysitters

A babysitter is a subordinate group member who stays with the young every day and rotates with the alpha female (mother) to lead the group on foraging excursions. Midday and evening are usually times when she returns to suckle the young. As the population size increases, the number of babysitters also increases; their effect on

the algorithm is that they reduce the overall population size. The babysitter exchange parameter resets the family's food source and scouting information previously held by the replacing members. With zero fitness weight, the alpha group's average weight is reduced in the next iteration, thereby emphasizing exploitation by hindering group movement.

The following unique attributes of DMO make it theoretically superior to some algorithms in literature in finding optimum solutions to different optimization problems:

- DMO randomly generates and improves candidate solutions based on the exploration and exploitation ability of dwarf mongooses and their compensatory adaptations.
- In search of food or sleeping mounds, the dwarf mongoose explores different regions of the problem search space.
- Due to its resemblance to dwarf mongooses, which are incapable of capturing large prey for family consumption but can acquire enough food to sustain themselves individually, the DMO exploits promising regions of the search space.
- DMOs can only be tuned for one parameter.

4 | APPLICATION AND RESULTS

The accuracy of new hybrid method, RVM tuned with DMOA (RVM-DMOA), is investigated in modeling monthly streamflow using hydroclimatic data as input. The model efficiency is tested by comparing with hybrid RVM tuned by PSO (RVM-PSO), RVM tuned by WOA (RVM-WOA), RVM tuned by MPA (RVM-MPA), and single RVM methods. Various lagged precipitation, temperature, and streamflow data were employed as inputs to the models, and their outcomes were assessed based on the following criteria:

$$\begin{aligned} RMSE : \text{Root Mean Square Error} \\ = \sqrt{\frac{1}{N} \sum_{i=1}^N [(Y_0)_i - (Y_C)_i]^2} \end{aligned} \quad (29)$$

$$MAE : \text{Mean Absolute Error} = \frac{1}{N} \sum_{i=1}^N |(Y_0)_i - (Y_C)_i| \quad (30)$$

$$\begin{aligned} NSE : \text{Nash - Sutcliffe Efficiency} \\ = 1 - \frac{\sum_{i=1}^N [(Y_0)_i - (Y_C)_i]^2}{\sum_{i=1}^N [(Y_0)_i - \bar{Y}_0]^2}, \quad -\infty < NSE \leq 1 \end{aligned} \quad (31)$$

$$\begin{aligned} R^2 : \text{Determination Coefficient} \\ = \left[\frac{\sum_{i=1}^N (Y_o - \bar{Y}_o)(Y_c - \bar{Y}_c)}{\sqrt{\sum_{i=1}^N (Y_o - \bar{Y}_o)^2 (Y_c - \bar{Y}_c)^2}} \right]^2 \end{aligned} \quad (32)$$

where Y_c, Y_o, \bar{Y}_o, N are the calculated, measured, and mean of the observed streamflow and data quantity, respectively. Table 2 provides some information about the parameters used in metaheuristic algorithms. As clearly seen from the table, 30 populations and 100 iterations were used in all algorithms, and they were run 20 times to obtain more robust outcomes. Data were divided in two subsets, 75% for training and 25% for testing.

4.1 | Results

Training and testing statistics are tabulated in Table 3 for single RVM models in predicting monthly streamflow. The combinations of inputs were decided by observing auto-correlation, partial auto-correlation, and cross-correlation functions. In the table, $P_{t-1}, T_{t-1},$ and Q_{t-1} indicate the precipitation, temperature and streamflow values of previous month and vice versa. A seen from

TABLE 2 Parameter settings for all algorithms.

Algorithm	Parameter	
PSO	Cognitive component (c_1)	2
	Social component (c_2)	2
	Inertia weight	0.2–0.9
WOA	a_1	Variable decreases linearly from 2 to 0
	a_2	Linearly decreases from –1 to –2
MPA	FADs	0.2
	P	0.5
	β	1.5
DMOA	—	
Common settings	Population	30
	Number of iterations	100
	Number of runs for each algorithm	20

Abbreviations: DMO, dwarf mongoose optimization; FAD, fish aggregating device; MPA, marine predators algorithm; PSO, particle swarm optimization; WOA, whale optimization algorithm.

Table 3, the models with 1, 11, and 12 lagged P , T , and Q values have the best accuracy in both training and testing periods, and only 1 lag produces insufficient predictions. It is clear that the lagged Q inputs have the highest influence on the current Q value and it is followed by the temperature data. The best P -based RVM model respectively provided the root mean square error (RMSE), mean absolute error (MAE), R^2 , and Nash–Sutcliffe efficiency (NSE) as 243.6 m^3/s , 191.6 m^3/s , 0.1606, and 0.1343 while the best T -based model correspondingly had the 184.2 m^3/s , 144.8 m^3/s , 0.6047, and 0.5052; and Q -based

model produced the lowest RMSE (121 m^3/s) and MAE (74.3 m^3/s) and the highest R^2 (0.7896) and NSE (0.7865) in modeling monthly streamflow.

Table 4 sums up the comparison statistics of the hybrid RVM–PSO models in predicting monthly streamflow. Here also, the lagged Q inputs produced better models' efficiency whereas the worst models belong to P inputs. The Q -based RVM–PSO acted as the best model with the lowest RMSE (117.3 m^3/s), MAE (73.8 m^3/s), and the highest R^2 (0.8124) and NSE (0.7993) in streamflow prediction, and it is followed by the T -based model

TABLE 3 Training and test statistics of the RVM models for streamflow prediction.

Model inputs	Training period				Test period			
	RMSE, m^3/s	MAE, m^3/s	R^2	NSE	RMSE, m^3/s	MAE, m^3/s	R^2	NSE
<i>P</i> inputs								
P_{t-1}	307.8	247.1	0.0500	0.0240	255.6	219.4	0.0783	0.0474
P_{t-1}, P_{t-11}	277.5	220.9	0.2085	0.2066	247.8	196.9	0.1550	0.1047
$P_{t-1}, P_{t-11}, P_{t-12}$	273.6	214.4	0.2291	0.2291	243.6	191.6	0.1606	0.1343
<i>T</i> inputs								
T_{t-1}	279.1	223.2	0.1974	0.1974	256.5	214.5	0.1312	0.0554
T_{t-1}, T_{t-11}	185.6	143.9	0.6461	0.6453	187.1	148.0	0.6140	0.4896
$T_{t-1}, T_{t-11}, T_{t-12}$	187.7	143.5	0.6370	0.6370	184.2	144.8	0.6047	0.5052
<i>Q</i> inputs								
Q_{t-1}	189.1	146.0	0.6317	0.6316	171.2	126.3	0.5829	0.5725
Q_{t-1}, Q_{t-11}	114.1	73.6	0.8668	0.8658	121.2	75.7	0.7542	0.7538
$Q_{t-1}, Q_{t-11}, Q_{t-12}$	111.0	69.6	0.8731	0.8731	121.0	74.3	0.7896	0.7865

Abbreviations: MAE, mean absolute error; NSE, Nash–Sutcliffe efficiency; RMSE, root mean square error; RVM, relevance vector machine.

TABLE 4 Training and test statistics of the RVM–PSO models for streamflow prediction.

Model inputs	Training period				Test period			
	RMSE, m^3/s	MAE, m^3/s	R^2	NSE	RMSE, m^3/s	MAE, m^3/s	R^2	NSE
<i>P</i> inputs								
P_{t-1}	301.8	243.9	0.0744	0.0615	254.7	217.1	0.0791	0.0543
P_{t-1}, P_{t-11}	275.1	216.5	0.2215	0.2206	245.7	193.4	0.1569	0.1196
$P_{t-1}, P_{t-11}, P_{t-12}$	270.0	211.3	0.2504	0.2492	241.0	191.1	0.1742	0.1531
<i>T</i> inputs								
T_{t-1}	275.0	220.4	0.2163	0.2110	254.4	211.3	0.1517	0.0561
T_{t-1}, T_{t-11}	179.2	134.3	0.6692	0.6692	182.2	139.8	0.6234	0.5158
$T_{t-1}, T_{t-11}, T_{t-12}$	184.4	144.5	0.6499	0.6499	182.7	144.7	0.6149	0.5134
<i>Q</i> inputs								
Q_{t-1}	187.7	142.3	0.6348	0.6347	165.4	124.5	0.6067	0.6005
Q_{t-1}, Q_{t-11}	108.6	66.7	0.8809	0.8785	117.3	73.8	0.8124	0.7993
$Q_{t-1}, Q_{t-11}, Q_{t-12}$	109.3	66.2	0.8775	0.8770	119.5	73.6	0.7998	0.7917

Abbreviations: MAE, mean absolute error; NSE, Nash–Sutcliffe efficiency; RMSE, root mean square error; RVM, relevance vector machine.

having RMSE, MAE, R^2 , and NSE of 182.2 m³/s, 139.8 m³/s, 0.6234, and 0.5158, respectively. Comparison with single RVM reveals that integrating PSO slightly improves its efficiency in both training and testing periods.

The results of hybrid RVM–WOA models are enlisted in Table 5 for predicting monthly streamflow. The models with only Q inputs act better than the P - and T -based models in both training and testing periods with respect to various assessment statistics. The based Q -based model had the lowest RMSE (116.5 m³/s) and MAE (72.5 m³/s) and the highest R^2 (0.8167) and NSE (0.8021) in the test period, and it was followed by the best T -based model with RMSE, MAE, R^2 , and NSE of 174 m³/s, 117.6 m³/s, 0.7149, and 0.5584, respectively. Applying WOA algorithm in tuning RVM parameters improved its accuracy in monthly streamflow prediction (compare Tables 3 and 5); improvements in RMSE, MAE, R^2 , and NSE of the best Q -based model are 3.7%, 2.4%, 3.4%, and 2% in the test period, respectively.

Table 6 reports the training and testing outcomes of the hybrid RVM–MPA models in monthly streamflow prediction. Similar to the RVM, RVM–PSO, and RVM–WOA models, the Q -based inputs produced the best predictions whereas the worst outcomes belong to the P -based inputs. It is clear from Tables 3 and 6 that the MPA algorithm improved the efficiency of single RVM in both training and testing periods; improvements in RMSE, MAE, R^2 , and NSE of the best Q -based model are 6.6%, 7.3%, 4.5%, and 4.7% in the test period, respectively. The assessment criteria of the proposed RVM–DMOA models

are enlisted in Table 7 for training and testing periods. As found for the previous RVM-based models, here also lagged Q seem to be the most effective inputs on current streamflow, T and P , respectively, follow it. DMOA algorithm performed better than the other algorithms in improving single RVM for monthly streamflow prediction; improvements in RMSE, MAE, R^2 , and NSE of the best Q -based model are 7.5%, 10.8%, 6.7%, and 5.2% in the test period, respectively.

Training and testing statistics of the all models with the optimal input combinations are compared in Table 8. In the table, OPT. P shows the optimal precipitation input obtained for each method and vice versa. It is clear from the table that combining all P , T , and Q inputs provides the best prediction accuracy and the RVM–DMOA is superior to the other models. According to the last input combinations, the proposed model improved the RMSE of RVM, RVM–PSO, RVM–WOA, and RVM–MPA by 7.5%, 3.9%, 3.9%, and 1% in the test period, respectively. The ML models are also compared in Table 9 in peak streamflow prediction for the test period. The streamflow values higher than 715 m³/s were selected as threshold for comparison. As evident from Table 9, the proposed RVM–DMOA model provided superior accuracy with the least relative error in peak streamflow prediction, and it should also be noted that the use of metaheuristic algorithms improves the ability of single RVM model.

The RVM-based methods are visually compared in Figures 7–11 for the test period. Time variation graph provided in Figure 7 illustrates that the predictions of the

TABLE 5 Training and test statistics of the RVM–PSO models for streamflow prediction.

Model inputs	Training period				Test period			
	RMSE, m ³ /s	MAE, m ³ /s	R^2	NSE	RMSE, m ³ /s	MAE, m ³ /s	R^2	NSE
<i>P</i> inputs								
P_{t-1}	297.1	242.2	0.0908	0.0908	252.5	214.3	0.0971	0.0704
P_{t-1}, P_{t-11}	268.1	206.0	0.2599	0.2597	242.1	190.2	0.1696	0.1453
$P_{t-1}, P_{t-11}, P_{t-12}$	268.8	210.5	0.2556	0.2556	240.7	189.0	0.1744	0.1555
<i>T</i> inputs								
T_{t-1}	276.8	216.2	0.2256	0.2205	251.4	210.6	0.1561	0.0787
T_{t-1}, T_{t-11}	174.5	131.1	0.6879	0.6864	175.8	137.5	0.6661	0.5493
$T_{t-1}, T_{t-11}, T_{t-12}$	167.3	127.9	0.7193	0.7117	174.0	117.6	0.7149	0.5584
<i>Q</i> inputs								
Q_{t-1}	185.2	141.1	0.6391	0.6391	163.7	123.6	0.6119	0.6091
Q_{t-1}, Q_{t-11}	106.6	64.0	0.8831	0.8830	117.1	70.2	0.8011	0.8001
$Q_{t-1}, Q_{t-11}, Q_{t-12}$	104.5	64.7	0.8825	0.8817	116.5	72.5	0.8167	0.8021

Abbreviations: MAE, mean absolute error; NSE, Nash–Sutcliffe efficiency; PSO, particle swarm optimization; RMSE, root mean square error; RVM, relevance vector machine.

TABLE 6 Training and test statistics of the RVM–MPA models for streamflow prediction.

Model inputs	Training period				Test period			
	RMSE, m ³ /s	MAE, m ³ /s	R ²	NSE	RMSE, m ³ /s	MAE, m ³ /s	R ²	NSE
<i>P</i> inputs								
P_{t-1}	296.6	241.9	0.0939	0.0937	251.3	213.2	0.1142	0.0790
P_{t-1}, P_{t-11}	261.7	203.1	0.2990	0.2947	239.3	185.1	0.1929	0.1647
$P_{t-1}, P_{t-11}, P_{t-12}$	259.4	201.9	0.3077	0.3069	235.8	186.2	0.2149	0.1891
<i>T</i> inputs								
T_{t-1}	270.0	211.0	0.2490	0.2488	247.9	207.6	0.1594	0.1037
T_{t-1}, T_{t-11}	159.7	110.9	0.7394	0.7371	171.5	121.8	0.7090	0.5711
$T_{t-1}, T_{t-11}, T_{t-12}$	160.6	108.0	0.7384	0.7343	164.6	106.3	0.7490	0.6048
<i>Q</i> inputs								
Q_{t-1}	182.4	138.8	0.6488	0.6483	162.1	120.9	0.6246	0.6168
Q_{t-1}, Q_{t-11}	103.5	62.5	0.8907	0.8895	114.6	69.7	0.8114	0.8085
$Q_{t-1}, Q_{t-11}, Q_{t-12}$	102.8	63.5	0.8977	0.8969	113.0	68.9	0.8252	0.8237

Abbreviations: MAE, mean absolute error; MPA, marine predators algorithm; NSE, Nash–Sutcliffe efficiency; RMSE, root mean square error; RVM, relevance vector machine.

TABLE 7 Training and test statistics of the RVM–DMOA models for streamflow prediction.

Model inputs	Training period				Test period			
	RMSE, m ³ /s	MAE, m ³ /s	R ²	NSE	RMSE, m ³ /s	MAE, m ³ /s	R ²	NSE
<i>P</i> inputs								
P_{t-1}	293.6	238.3	0.1118	0.1116	250.1	202.6	0.1188	0.0875
P_{t-1}, P_{t-11}	252.0	189.5	0.3458	0.3458	230.3	180.4	0.2410	0.2265
$P_{t-1}, P_{t-11}, P_{t-12}$	245.8	187.2	0.3778	0.3778	232.1	177.2	0.2583	0.2145
<i>T</i> inputs								
T_{t-1}	261.7	202.1	0.2944	0.2944	245.3	204.6	0.1686	0.1227
T_{t-1}, T_{t-11}	122.5	78.5	0.8454	0.8454	164.7	112.6	0.7476	0.6042
$T_{t-1}, T_{t-11}, T_{t-12}$	119.7	74.5	0.8524	0.8524	159.6	124.7	0.6958	0.6285
<i>Q</i> inputs								
Q_{t-1}	179.8	134.0	0.6671	0.6671	161.5	121.5	0.6251	0.6196
Q_{t-1}, Q_{t-11}	93.6	52.1	0.9097	0.9097	111.9	66.3	0.8426	0.8274
$Q_{t-1}, Q_{t-11}, Q_{t-12}$	94.2	53.8	0.9085	0.9085	112.5	67.2	0.8322	0.8254

Abbreviations: DMOA, dwarf mongoose optimization algorithm; MAE, mean absolute error; NSE, Nash–Sutcliffe efficiency; RMSE, root mean square error; RVM, relevance vector machine.

RVM–DMOA model better follow the observed streamflow values than the other models. However, it is also clear that all models cannot predict some peaks well, there are considerable underpredictions. It is evident from the scatter graphs given in Figure 8 that the RVM–DMOA model produces the less scattered predictions compared with other models and metaheuristic algorithms improved the prediction accuracy of the single RVM model. From the Taylor graph in Figure 9, it can be said that the RVM–

DMOA model has the least RMSE and the highest correlation and its standard deviation is closest to the observed one. Figure 10 shows the violin charts in which the distribution of the RVM–DMOA predictions are closer to that of the observed streamflow compared with the other RVM-based models. The radar chart provided in Figure 11 shows the four assessment statistics at the same time, and it is clearly observed from this figure that the RVM–DMOA has the least RMSE and MAE and the highest

TABLE 8 Training and test statistics of the models for streamflow prediction using optimal input combinations.

Model inputs	Training period				Test period			
	RMSE, m ³ /s	MAE, m ³ /s	R ²	NSE	RMSE, m ³ /s	MAE, m ³ /s	R ²	NSE
RVM								
OPT. P + OPT. T	175.64	137.14	0.6822	0.6822	187.12	147.90	0.6172	0.4894
OPT. P + OPT. Q	108.41	66.68	0.8792	0.8789	119.29	70.71	0.7973	0.7925
OPT. T + OPT. Q	135.16	85.45	0.8118	0.8118	130.00	86.18	0.7536	0.7536
OPT. P + OPT. T + OPT. Q	107.67	73.10	0.8807	0.8806	112.29	73.96	0.8198	0.8161
Mean								
RVM-PSO								
OPT. P + OPT. T	167.14	130.20	0.7122	0.7122	186.89	139.81	0.6404	0.6342
OPT. P + OPT. Q	108.03	65.45	0.8798	0.8798	114.06	71.34	0.8141	0.8103
OPT. T + OPT. Q	107.65	69.42	0.8806	0.8806	112.81	73.74	0.8261	0.8143
OPT. P + OPT. T + OPT. Q	104.42	67.04	0.8877	0.8877	110.35	69.94	0.8392	0.8224
Mean								
RVM-WOA								
OPT. P + OPT. T	162.56	124.32	0.7310	0.7278	179.72	136.58	0.6526	0.6464
OPT. P + OPT. Q	107.42	64.72	0.8822	0.8811	113.72	71.43	0.8171	0.8114
OPT. T + OPT. Q	105.54	66.94	0.8854	0.8853	112.35	72.41	0.8273	0.8161
OPT. P + OPT. T + OPT. Q	102.80	65.02	0.8914	0.8911	110.10	67.66	0.8428	0.8232
Mean								
RVM-MPA								
OPT. P + OPT. T	161.19	123.04	0.7462	0.7424	178.31	134.61	0.6568	0.6490
OPT. P + OPT. Q	104.31	62.53	0.8883	0.8879	113.43	71.00	0.8180	0.8124
OPT. T + OPT. Q	103.53	64.76	0.8900	0.8896	111.70	70.02	0.8332	0.8181
OPT. P + OPT. T + OPT. Q	101.70	64.27	0.8938	0.8934	109.33	66.15	0.8494	0.8357
Mean								
RVM-DMOA								
OPT. P + OPT. T	103.47	64.67	0.8897	0.8897	162.92	105.75	0.7246	0.7130
OPT. P + OPT. Q	83.45	47.73	0.9283	0.9283	112.58	69.78	0.8200	0.8152
OPT. T + OPT. Q	73.82	43.09	0.9439	0.9439	104.39	64.00	0.8437	0.8411
OPT. P + OPT. T + OPT. Q	67.49	42.48	0.9531	0.9531	98.54	61.24	0.8647	0.8584
Mean								

Abbreviations: DMOA, dwarf mongoose optimization algorithm; MAE, mean absolute error; MPA, marine predators algorithm; NSE, Nash-Sutcliffe efficiency; RMSE, root mean square error; RVM, relevance vector machine; PSO, particle swarm optimization; WOA, whale optimization algorithm.

correlation coefficient and NSE in predicting monthly streamflow. The improvements by applying metaheuristic algorithms are also clearly seen from the chart.

4.2 | Discussion

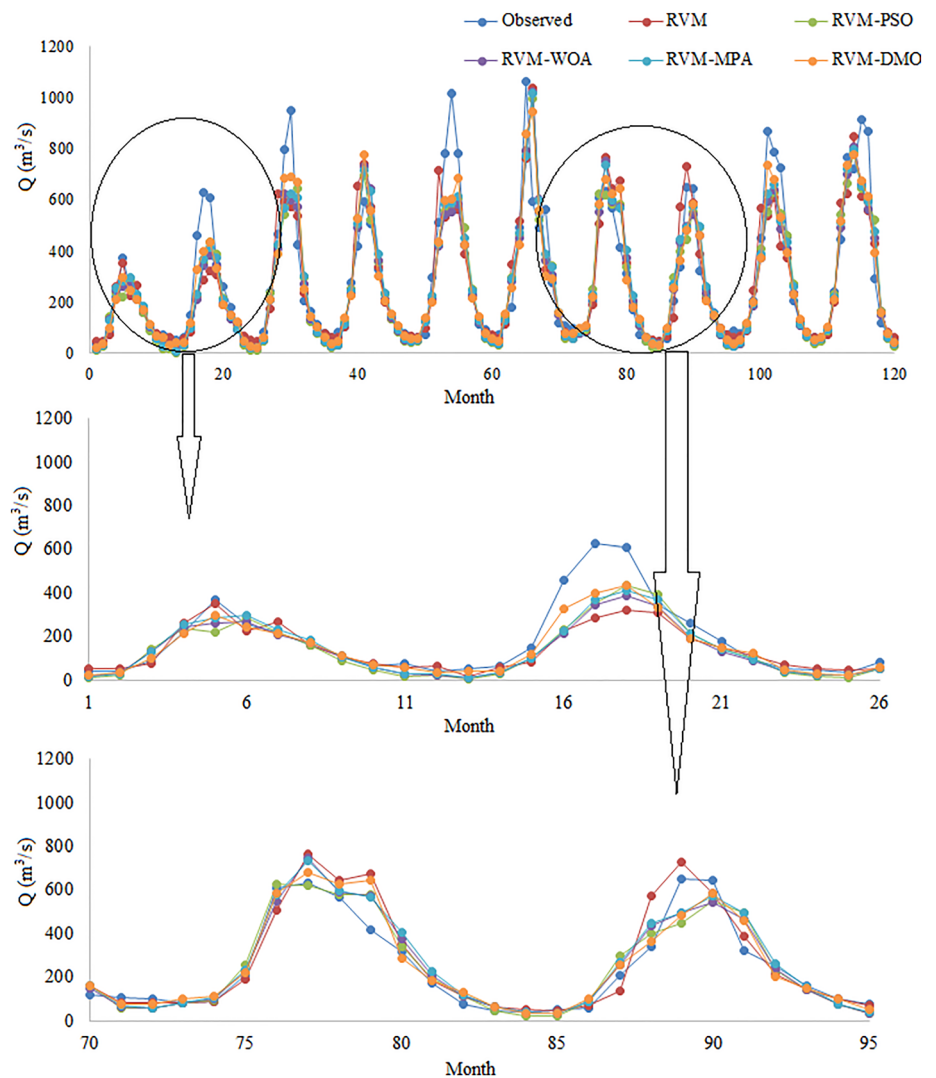
The feasibility of RVM tuned with DMOA was investigated in predicting monthly streamflow and its outcomes were compared with hybrid and single RVM models.

Various lagged values of precipitation, temperature, and streamflow data were utilized as input to the models to see their influence on the current streamflow. According to the result, previous streamflow values were found to be most effective on current streamflow, and it was followed by monthly temperature data as a second-best affective variable. As evident from Table 1, the precipitation data have more skewed distribution compared with temperature data, and this is very effective on ML models' accuracy as also mentioned by the previous

TABLE 9 The comparison of different RVM-based models in peak streamflow prediction for the test period.

Date	Observed values Peaks >715	Estimates, m ³ /s					Relative error, %				
		RVM	RVM- PSO	RVM- WOA	RVM- MPA	RVM- DMOA	RVM	RVM- PSO	RVM- WOA	RVM- MPA	RVM- DMOA
May 3	794.4	590.7	543.2	624.5	568.6	684.7	25.6	31.6	21.4	28.4	13.8
June 3	948.8	570.1	600.7	596.1	624.8	691.9	39.9	36.7	37.2	34.1	27.1
May 5	783.0	530.4	557.7	542.5	584.2	600.1	32.3	28.8	30.7	25.4	23.4
June 5	1016.7	557.8	582.9	553.3	587.9	602.7	45.1	42.7	45.6	42.2	40.7
July 5	783.8	559.8	609.8	583.1	615.3	684.3	28.6	22.2	25.6	21.5	12.7
May 6	1064.1	760.7	792.4	791.1	769.3	860.4	28.5	25.5	25.7	27.7	19.1
May 9	866.7	538.2	550.3	593.2	621.3	734.4	37.9	36.5	31.6	28.3	15.3
June 9	787.4	607.7	629.5	638.4	658.5	678.4	22.8	20.1	18.9	16.4	13.8
May 10	766.3	622.4	664.0	699.6	722.0	735.0	18.8	13.4	8.7	5.8	4.1
June 10	718.4	850.5	785.2	807.9	795.5	775.1	-18.4	-9.3	-12.5	-10.7	-7.9
July 10	916.6	611.7	641.4	657.4	655.3	673.4	33.3	30.0	28.3	28.5	26.5
August 10	869.0	556.7	596.8	559.8	593.8	615.0	35.9	31.3	35.6	31.7	29.2
Absolute error							367.2	328.1	321.6	300.7	233.7

FIGURE 7 Time variation graphs of the observed and predicted streamflow by different relevance vector machine (RVM)-based models in the test period using best input combination. DMO, dwarf mongoose optimization; MPA, marine predators algorithm; PSO, particle swarm optimization; WOA, whale optimization algorithm.



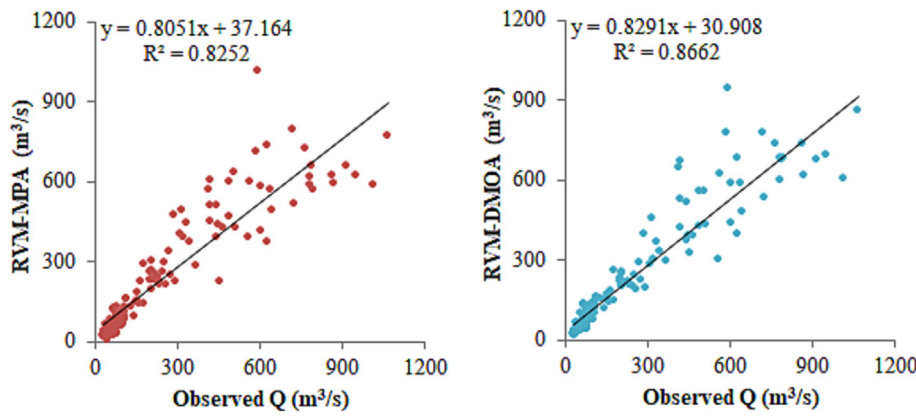


FIGURE 8 Scatterplots of the observed and predicted streamflow by different relevance vector machine (RVM)-based models in the test period using best input combination.

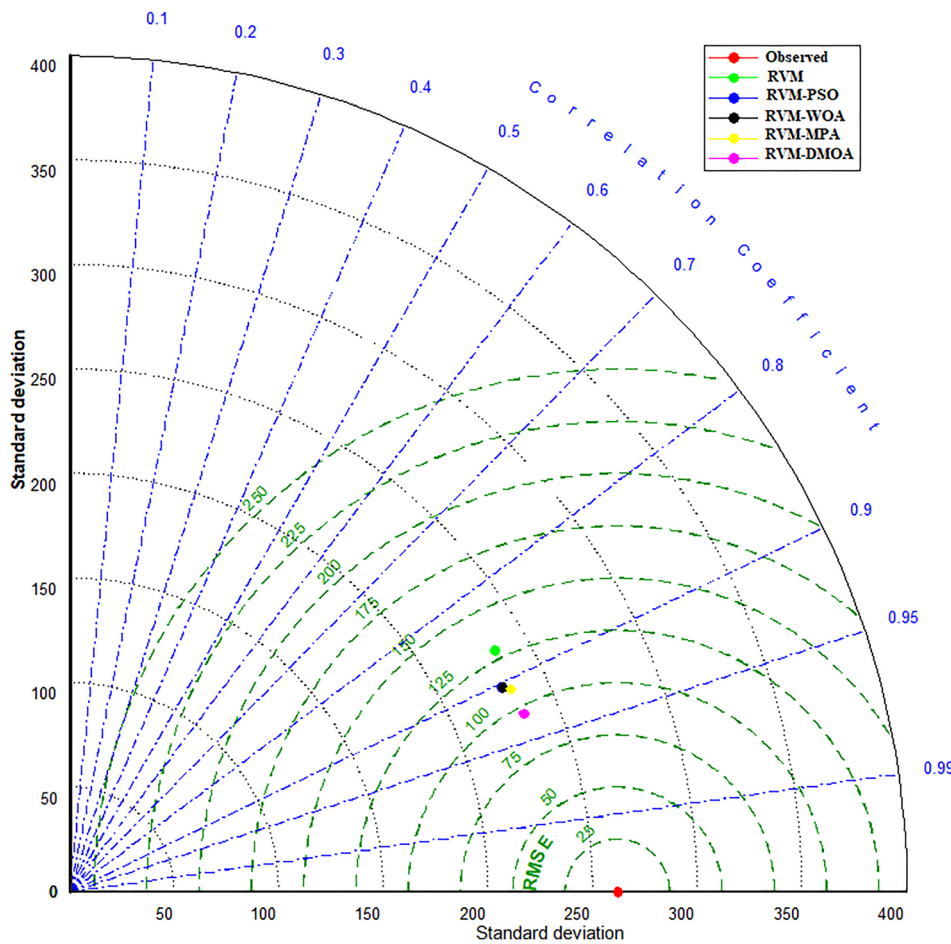


FIGURE 9 Taylor diagrams of the predicted streamflow by different relevance vector machine (RVM)-based models in the test period using best input combination. DMO, dwarf mongoose optimization; MPA, marine predators algorithm; PSO, particle swarm optimization; WOA, whale optimization algorithm.

literature (e.g., Adnan et al., 2020; Kisi & Aytok, 2013; Kisi & Parmar, 2016). Good predictions were obtained from the RVM-DMOA model by using only precipitation and temperature-based models (NSE is higher than 0.70 in the test period). Such models can be very essential in monthly streamflow prediction especially for the developing countries because in case of missing Q data, lagged values cannot be used in prediction. This is also reported by the previous research (Adnan et al., 2020).

From visual comparison, it was found that all the ML models considerably underpredict the monthly peak

streamflow. This can be explained by the limited number of samples for the peak values. Time interval is monthly, and we have limited number of peak streamflow data, and the ML models are data-driven and cannot be well learned from the phenomenon (Adnan et al., 2020). The other important issue is the data range used in training and testing periods. As clearly seen from Table 1, the range of training datasets (33.64–1064.1 m³/s) for streamflow do not cover the range of testing period (41.45–1319.9 m³/s). This is also justified by the previous research of Kisi and Parmar, 2016 and Adnan et al., 2020.

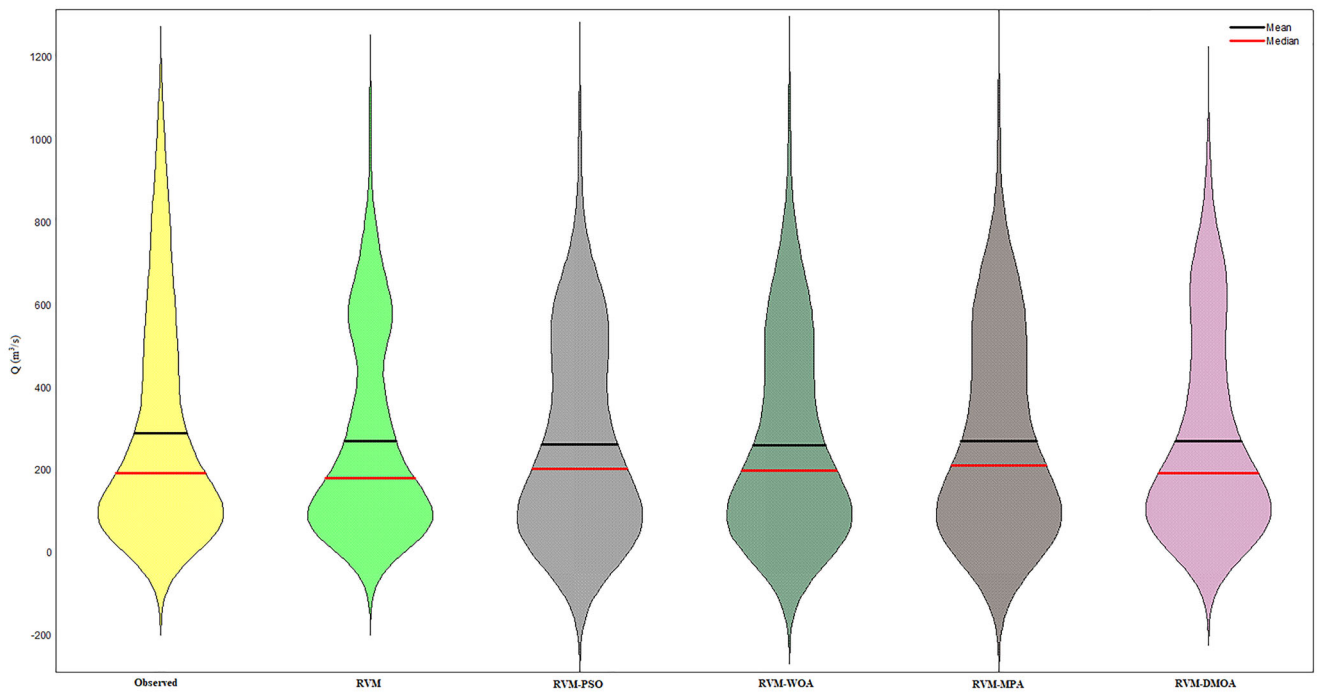
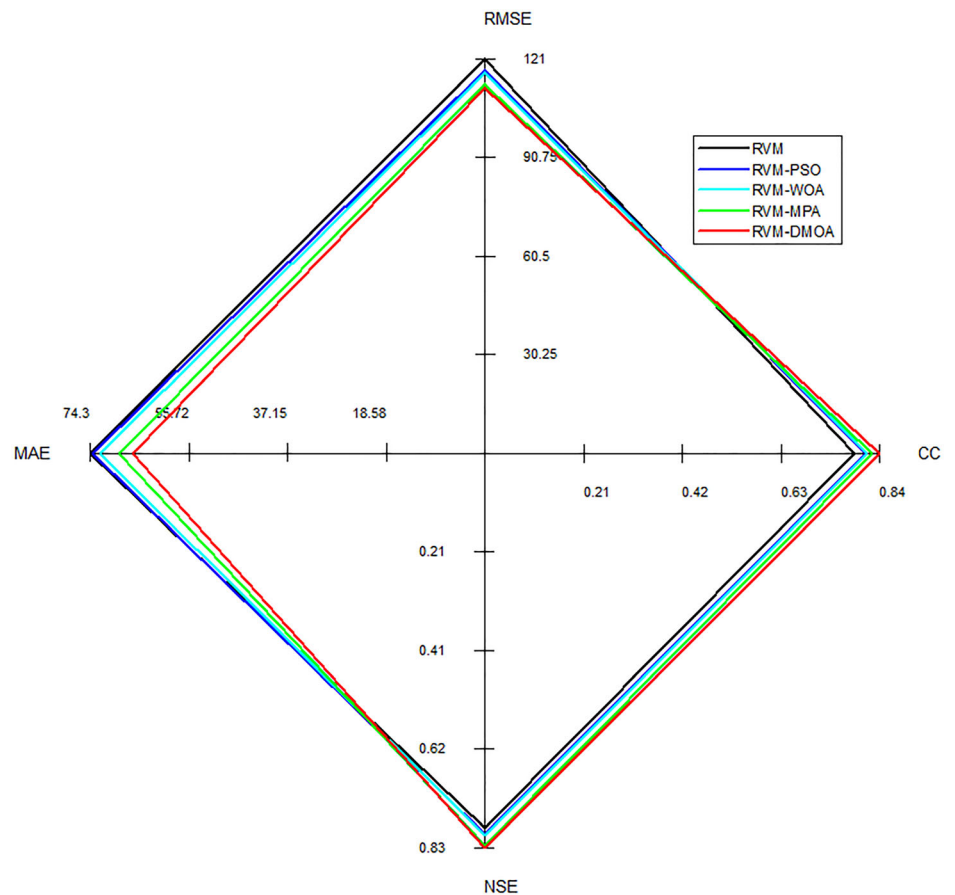


FIGURE 10 Violin charts of the predicted streamflow by different relevance vector machine (RVM)-based models in the test period using best input combination. DMO, dwarf mongoose optimization; MPA, marine predators algorithm; PSO, particle swarm optimization; WOA, whale optimization algorithm.

FIGURE 11 Radar chart of the predicted streamflow by different relevance vector machine (RVM)-based models in the test period using best input combination. CC, correlation coefficient; DMO, dwarf mongoose optimization; MAE, mean absolute error; MPA, marine predators algorithm; NSE, Nash-Sutcliffe efficiency; PSO, particle swarm optimization; WOA, whale optimization algorithm.



The results revealed that the metaheuristic algorithms are very useful in improving RVM model in predicting monthly streamflow. The DMOA algorithm was found to be superior to the other alternatives and provided the best predictions. The main advantage of this algorithm is having the ability of exploratory and exploitation and its compensatory adaptation. This algorithm stochastically generates and improves a set of candidate solutions for a given optimization problem (Agushaka et al., 2022).

5 | CONCLUDING REMARKS

In this study, the ability of hybrid RVM–DMOA was investigated for predicting monthly streamflow using hydroclimatic data (e.g., precipitation, temperature, and streamflow) as input. The prediction outcomes were compared with other RVM-based methods (e.g., RVM, RVM–GWO, RVM–WOA, and RVM–MRFO). The RVM–DMOA improved the accuracy of single RVM in monthly streamflow prediction. The improvements in RMSE, MAE, R^2 , and NSE of RVM in testing period were found to be 7.5%, 10.8%, 6.7%, and 5.2% by applying RVM–DMOA, respectively. The study revealed that the employing metaheuristic algorithms improves the efficiency of single RVM in monthly streamflow prediction. Among the input combinations, the lagged streamflow data were found to be the most effective variable on current streamflow while precipitation had the least effect. On the other hand, the best accuracy was obtained from the models comprising lagged precipitation, temperature, and streamflow. Combination of precipitation and temperature inputs also produced satisfactory predictions, which is very essential in the regions where streamflow data have gaps or in case of missing data.

AUTHOR CONTRIBUTIONS

The authors have equally contributed to the study. All authors read and approved the original manuscript.

ACKNOWLEDGMENTS

The authors would also like to express their sincere appreciation to the associate editor and the anonymous reviewers for their comments and suggestions. This work was supported by the National Social Science Foundation of China (grant number 18BTJ029), Key Projects of National Statistical Science Research Projects (grant number 2020LZ10), General Projects of Guangdong Natural Science Research Projects (grant number 2023A1515011520) and Tertiary Education Scientific Research Project of Guangzhou Municipal Education Bureau (grant number 202235324). Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Rana Muhammad Adnan  <https://orcid.org/0000-0002-2650-8123>

Ozgun Kisi  <https://orcid.org/0000-0001-7847-5872>

REFERENCES

- Abdel-basset, M., Mohamed, R., Mirjalili, S., Chakraborty, R. K., & Ryan, M. (2021). An efficient marine predators algorithm for solving multi-objective optimization problems: Analysis and validations. *IEEE Access*, 9, 42817–42844. <https://doi.org/10.1109/ACCESS.2021.3066323>
- Adnan, R. M., Liang, Z., Heddami, S., Zounemat-Kermani, M., Kisi, O., & Li, B. (2020). Least square support vector machine and multivariate adaptive regression splines for streamflow prediction in mountainous basin using hydro-meteorological data as inputs. *Journal of Hydrology*, 586, 124371. <https://doi.org/10.1016/j.jhydrol.2019.124371>
- Adnan, R. M., Mostafa, R. R., Islam, A. R. M. T., Gorgij, A. D., Kuriqi, A., & Kisi, O. (2021). Improving drought modeling using hybrid random vector functional link methods. *Water (Basel)*, 13, 3379. <https://doi.org/10.3390/W13233379>
- Agushaka, J. O., Ezugwu, A. E., & Abualigah, L. (2022). Dwarf mongoose optimization algorithm. *Computer Methods in Applied Mechanics and Engineering*, 391, 114570. <https://doi.org/10.1016/J.CMA.2022.114570>
- Bennett, J. C., Wang, Q. J., Robertson, D. E., Schepen, A., Li, M., & Michael, K. (2017). Assessment of an ensemble seasonal streamflow forecasting system for Australia. *Hydrology and Earth System Sciences*, 21(12), 6007–6030. <https://doi.org/10.5194/hess-21-6007-2017>
- Bui, D. T., Shahabi, H., Shirzadi, A., Chapi, K., Hoang, N. D., Pham, B. T., Bui, Q. T., Tran, C. T., Panahi, M., Ahmad, B. B., & Saro, L. (2018). A novel integrated approach of relevance vector machine optimized by imperialist competitive algorithm for spatial modeling of shallow landslides. *Remote Sensing*, 10, 1538.
- Chou, J. S., & Truong, D. N. (2021). A novel metaheuristic optimizer inspired by behavior of jellyfish in ocean. *Applied Mathematics and Computation*, 389, 125535. <https://doi.org/10.1016/J.AMC.2020.125535>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>
- Eum, H. I., Vasan, A., & Simonovic, S. P. (2012). Integrated reservoir management system for flood risk assessment under climate change. *Water Resources Management*, 26, 3785–3802. <https://doi.org/10.1007/S11269-012-0103-4>
- Fan, Q., Yu, F., & Xuan, M. (2021). Transformer fault diagnosis method based on improved whale optimization algorithm to optimize support vector machine. *Energy Reports*, 7, 856–866. <https://doi.org/10.1016/J.EGYR.2021.09.188>

- Faramarzi, A., Heidarinejad, M., Mirjalili, S., & Gandomi, A. H. (2020). Marine predators algorithm: A nature-inspired meta-heuristic. *Expert Systems with Applications*, *152*, 113377. <https://doi.org/10.1016/j.eswa.2020.113377>
- Fathollahi-Fard, A. M., Ahmadi, A., & al-e-Hashem, S. M. J. M. (2020). Sustainable closed-loop supply chain network for an integrated water supply and wastewater collection system under uncertainty. *Journal of Environmental Management*, *275*, 111277. <https://doi.org/10.1016/J.JENVMAN.2020.111277>
- Fattahi, P., & Fayyaz, S. (2010). A compromise programming model to integrated urban water management. *Water Resources Management*, *24*(6), 1211–1227. <https://doi.org/10.1007/S11269-009-9492-4>
- Goldman, N., & Saykally, R. J. (2004). Elucidating the role of many-body forces in liquid water. I. Simulations of water clusters on the VRT (ASP-W) potential surfaces. *The Journal of Chemical Physics*, *120*, 4777–4789. <https://doi.org/10.1063/1.1645777>
- Goshime, D. W., Absi, R., Haile, A. T., Ledésert, B., & Rientjes, T. (2020). Bias-corrected CHIRP satellite rainfall for water level simulation, Lake Ziway, Ethiopia. *Journal of Hydrologic Engineering*, *25*, 05020024. [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001965](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001965)
- Hasanpour Kashani, M., Daneshfaraz, R., Ghorbani, M. A., Najafi, M. R., & Kisi, O. (2015). Comparison of different methods for developing a stage discharge curve of the Kizilirmak River. *Journal of Flood Risk Management*, *8*, 71–86. <https://doi.org/10.1111/jfr3.12064>
- Hussain, D., Hussain, T., Khan, A. A., Naqvi, S. A. A., & Jamil, A. (2020). A deep learning approach for hydrological time-series prediction: A case study of Gilgit river basin. *Earth Science Informatics*, *13*(3), 915–927. <https://doi.org/10.1007/s12145-020-00477-2>
- Kang, D., & Lansey, K. (2013). Scenario-based robust optimization of regional water and wastewater infrastructure. *Journal of Water Resources Planning and Management*, *139*, 325–338. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000236](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000236)
- Karimi-Mamaghan, M., Mohammadi, M., Meyer, P., Karimi-Mamaghan, A. M., & Talbi, E. G. (2022). Machine learning at the service of meta-heuristics for solving combinatorial optimization problems: A state-of-the-art. *European Journal of Operational Research*, *296*, 393–422. <https://doi.org/10.1016/J.EJOR.2021.04.032>
- Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In: *Proceedings of the IEEE International Conference on Neural Networks*. IEEE, Piscataway, NJ, USA, 1942–1948.
- Kennedy, J., & Eberhart, R. (2022). Particle swarm optimization. *Proceedings of ICNN'95—International Conference on Neural Networks* 4. IEEE: 1942–1948. Accessed September 8. doi: <https://doi.org/10.1109/ICNN.1995.488968>
- Kisi, O., & Ayttek, A. (2013). Explicit neural network in suspended sediment load estimation. *Neural Network World*, *23*(6), 587–607. <https://doi.org/10.14311/NNW.2013.23.035>
- Kisi, O., & Cigizoglu, H. K. (2007). Comparison of different ANN techniques in river flow prediction. *Civil Engineering and Environmental Systems*, *24*(3), 211–231. <https://doi.org/10.1080/10286600600888565>
- Kisi, O., & Parmar, K. S. (2016). Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution. *Journal of Hydrology*, *534*, 104–112. <https://doi.org/10.1016/j.jhydrol.2015.12.014>
- Liu, W., Shao, Y., Chen, K., Li, C., & Luo, H. (2022). Whale optimization algorithm-based point cloud data processing method for sewer pipeline inspection. *Automation in Construction*, *141*, 104423. <https://doi.org/10.1016/j.autcon.2022.104423>
- Liu, W., Wang, Z., Zeng, N., Yuan, Y., Alsaadi, F. E., & Liu, X. (2020). A novel randomised particle swarm optimizer. *International Journal of Machine Learning and Cybernetics*, *12*, 529–540. <https://doi.org/10.1007/S13042-020-01186-4>
- Malik, A., Tikhamarine, Y., Souag-Gamane, D., Kisi, O., & Pham, Q. B. (2020). Support vector regression optimized by meta-heuristic algorithms for daily streamflow prediction. *Stochastic Environmental Research and Risk Assessment*, *34*, 1755–1773. <https://doi.org/10.1007/s00477-020-01874-1>
- Menad, N. A., Nouredine, Z., Hemmati-Sarapardeh, A., & Shamshirband, S. (2019). Modeling temperature-based oil-water relative permeability by integrating advanced intelligent models with grey wolf optimization: Application to thermal enhanced oil recovery processes. *Fuel*, *242*, 649–663. <https://doi.org/10.1016/J.FUEL.2019.01.047>
- Mirjalili, S., & Lewis, A. (2016). The whale optimization algorithm. *Advances in Engineering Software*, *95*, 51–67. <https://doi.org/10.1016/j.advengsoft.2016.01.008>
- Mo, S., Duan, H., Shen, B., & Wang, D. (2015). Interval two-stage stochastic integer programming for urban water resource management under uncertainty. *Journal of Coastal Research*, *73*, 160–165. <https://doi.org/10.2112/SI73-028.1>
- Mohammadi, B., & Mehdizadeh, S. (2016). Modeling daily reference evapotranspiration via a novel approach based on support vector regression coupled with whale optimization algorithm. *Agricultural Water Management*, *237*, 106145.
- Mortazavi-Naeini, M., Kuczera, G., Kiem, A. S., Cui, L., Henley, B., Berghout, B., & Turner, E. (2015). Robust optimization to secure urban bulk water supply against extreme drought and uncertain climate change. *Environmental Modelling and Software*, *69*, 437–451. <https://doi.org/10.1016/J.ENVSOF.2015.02.021>
- Pérez, R., Sanz, G., Cugueró, M. À., Blesa, J., & Cugueró, J. (2015). Parameter uncertainty modelling in water distribution network models. *Procedia Engineering*, *119*(1), 583–592. <https://doi.org/10.1016/J.PROENG.2015.08.911>
- Rezaei, F., & Safavi, H. R. (2020). GuASPSO: A new approach to hold a better exploration–exploitation balance in PSO algorithm. *Soft Computing*, *24*, 4855–4875. <https://doi.org/10.1007/S00500-019-04240-8/TABLES/17>
- Roy, S. M., Pareek, C. M., Machavaram, R., & Mukherjee, C. K. (2022). Optimizing the aeration performance of a perforated pooled circular stepped cascade aerator using hybrid ANN-PSO technique. *Information Processing in Agriculture*, *9*, 533–546. <https://doi.org/10.1016/J.INPA.2021.09.002>
- Safari, M. J. S., Arashloo, S. R., & Vaheddoost, B. (2022). Fast multi-output relevance vector regression for joint groundwater and lake water depth modeling. *Environmental Modelling & Software*, *154*, 105425. <https://doi.org/10.1016/j.envsoft.2022.105425>
- Sakib, N., Ibne Hossain, N. U., Nur, F., Talluri, S., Jaradat, R., & Lawrence, J. M. (2021). An assessment of probabilistic disaster in the oil and gas supply chain leveraging Bayesian belief

- network. *International Journal of Production Economics*, 235, 108107. <https://doi.org/10.1016/J.IJPE.2021.108107>
- Sayari, S., Meymand, A. M., Aldallal, A., & Zounemat-Kermani, M. (2022). Meta-learner methods in forecasting regulated and natural river flow. *Arabian Journal of Geosciences*, 15(11), 1051. <https://doi.org/10.1007/s12517-022-10274-4>
- Schwartz, R., Housh, M., & Ostfeld, A. (2016). Least-cost robust design optimization of water distribution systems under multiple loading. *Journal of Water Resources Planning and Management*, 142, 04016031. [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000670](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000670)
- Smola, A. J., & Schölkopf, B. (2004). A tutorial on support vector regression. *Statistics and Computing*, 14(3), 199–222. <https://doi.org/10.1023/B:STCO.0000035301.49549.88>
- Steinbrueckh, E. (2014). *Water scarcity and international conflict in Africa, a game theory approach*. Università Degli Studi Di Perugia, Scienze Politiche.
- Su, B., Lin, Y., Wanga, J., Quan, X., Chang, Z., & Rui, C. (2022). Sewage treatment system for improving energy efficiency based on particle swarm optimization algorithm. *Energy Reports*, 8, 8701–8708. <https://doi.org/10.1016/j.egy.2022.06.053>
- Tao, H., Al-Bedyry, N. K., Khedher, K. M., Shahid, S., & Yaseen, Z. M. (2021). River water level prediction in coastal catchment using hybridized relevance vector machine model with improved grasshopper optimization. *Journal of Hydrology*, 598, 126477. <https://doi.org/10.1016/j.jhydrol.2021.126477>
- Tipping, M. E. (2000). The relevance vector machine. In S. A. Solla, T. K. Leen, & K. R. Müller *Advances in neural information processing systems 12*. Cambridge, Mass: MIT Press.
- Tipping, M. E. (2001). Sparse Bayesian learning and the relevance vector machine. *Journal of Machine Learning Research*, 1, 211–244.
- Tyralis, H., Papacharalampous, G., & Langousis, A. (2021). Super learning for daily streamflow forecasting: Large-scale demonstration and comparison with multiple machine learning algorithms. *Neural Computing and Applications*, 33, 3053–3068. <https://doi.org/10.1007/s00521-020-05172-3>
- Verleye, D., & Aghezzaf, E. H. (2011). *Modeling and optimization of production and distribution of drinking water at VMW*. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) 6701 LNCS. (pp. 315–326). Springer. https://doi.org/10.1007/978-3-642-21527-8_37/COVER
- Wang, Z., Tian, J., & Feng, K. (2022). Optimal allocation of regional water resources based on simulated annealing particle swarm optimization algorithm. *Energy Reports*, 8, 9119–9126. <https://doi.org/10.1016/j.egy.2022.07.033>
- Zhou, J., Peng, T., Zhang, C., & Sun, N. (2018). Data pre-analysis and ensemble of various artificial neural networks for monthly streamflow forecasting. *Water*, 10, 628. <https://doi.org/10.3390/w10050628>
- Zounemat-Kermani, M., Mahdavi Meymand, A., & Hinkelmann, R. (2021). A comprehensive survey on conventional and modern neural networks: Application to river flow forecasting. *Earth Science Informatics*, 14(2), 893–911. <https://doi.org/10.1007/s12145-021-00599-1>

AUTHOR BIOGRAPHIES

Rana Muhammad Adnan is currently a full-time Postdoc Researcher at the School of Economics and Statistics, Guangzhou University. He received his MS degree in Water Resources Engineering from the University of Engineering and Technology, Lahore, Pakistan, in 2013 and the PhD degree in hydraulic engineering in 2017 from Huazhong University of Science and Technology, Wuhan, China. His main research interests include the use of Big Data with help of artificial intelligence and applied statistical methods in fields of environment, atmospheric processes, wind energy, solar energy, ocean engineering, hydrology, and water resources.

Reham R. Mostafa is working as an associate professor at the Information Systems Department, Faculty of Computers and Information Sciences, Mansoura University, Egypt. Her working area involves artificial intelligence, big data analytics, information security, computer vision, and optimization

Hong-Liang Dai is currently a professor at the School of Economics and Statistic, Guangzhou University, Guangzhou, China. He received the MS degree in applied mathematics from Wuhan University, Wuhan, China, in 2003 and the PhD degree in applied mathematics from Sun Yat-Sen University, Guangzhou, China, in 2013. His research interests include machine learning, data mining, statistical data analysis, image processing, and bioinformatics.

Ehsan Mansouri is a researcher working in Department of Computer and Technology, Birjand University of Medical Sciences, Iran. He is working on computational medicine, cloud computing, machine learning, and data grid.

Ozgur Kisi has been working as a guest professor at University of Applied Sciences, Lübeck, Germany. He received his PhD in the Institute of Science and Technology (Hydraulics Division) at the Istanbul Technical University, Turkey (2003). His research fields are developing novel algorithms and methods towards the innovative solution of hydrologic forecasting and modeling; suspended sediment modeling; forecasting, estimating, and spatial and temporal analysis of hydroclimatic variables such as precipitation, streamflow, suspended sediment, evaporation, evapotranspiration, groundwater, lake level, and water quality

parameters; hydro-informatics. He is serving as a reviewer for more than 100 journals indexed in the Science Citation Index (SCI) in the fields of hydrology, irrigation, water resources and hydro-informatics. He has authored more than 500 research articles, 15 chapters, and 30 discussions (Web of Science h-index = 73, Google Scholar h-index = 95; Scopus h-index 80).

Mohammad Zounemat-Kermani is a university lecturer and researcher specializing in hydrosociences. His research areas are focused on developing soft computing techniques and machine learning models in hydraulics and hydrology applications.

How to cite this article: Adnan, R. M., Mostafa, R. R., Dai, H.-L., Mansouri, E., Kisi, O., & Zounemat-Kermani, M. (2024). Comparison of improved relevance vector machines for streamflow predictions. *Journal of Forecasting*, 43(1), 159–181. <https://doi.org/10.1002/for.3028>