Applied Soft Computing 131 (2022) 109739

Contents lists available at ScienceDirect

Applied Soft Computing

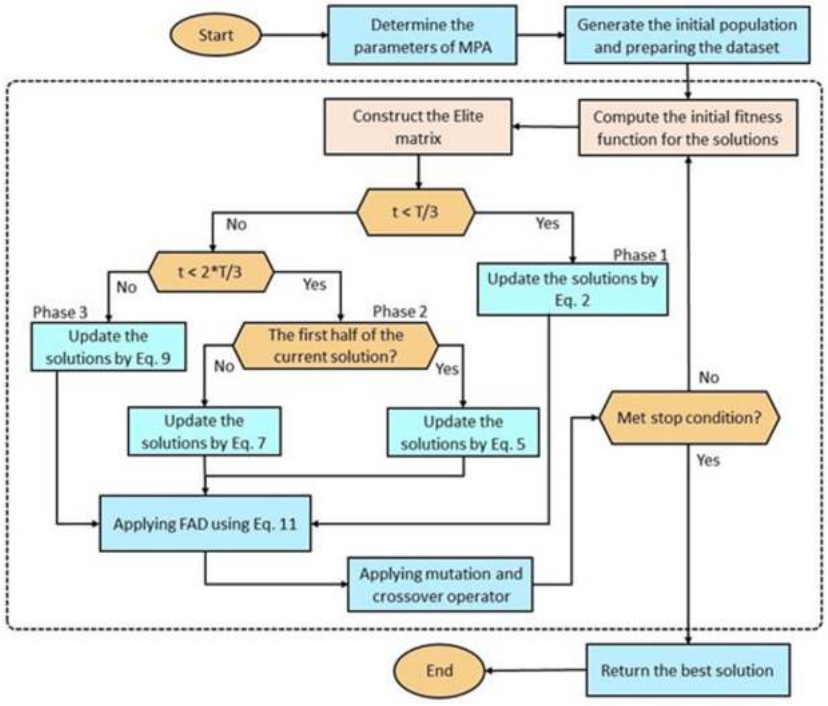
journal homepage: www.elsevier.com/locate/asoc

The viability of extended marine predators algorithm-based artificial neural networks for streamflow prediction

Rana Muhammad Adnan Ikram [a], Ahmed A. Ewees [b,*], Kulwinder Singh Parmar [c], Zaher Mundher Yaseen [f], Shamsuddin Shahid [d], Ozgur Kisi [d,e,**]

[a] School of Economics and Statistics, Guangzhou University, Guangzhou 510006, China
[b] Department of Computer, Damietta University, Damietta 34517, Egypt
[c] Department of Mathematics, IKG Punjab Technical University, Jalandhar, Kapurthala, India
[d] School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia (UTM), 81310 Johor Bahru, Malaysia
[e] School of Technology, Ilia State University, 0162, Tbilisi, Georgia
[f] Civil and Environmental Engineering Department, King Fahd University of Petroleum & Minerals, Dhahran 31261, Saudi Arabia

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 13 April 2021
Received in revised form 13 October 2022

ABSTRACT

Precise streamflow prediction is necessary for better planning and managing available water and future water resources, especially for high altitude mountainous glacier melting affected basins in the climate change context. In the current study, a novel hybridized machine learning method, extended marine predators algorithm (EMPA)-based ANN (ANN-EMPA), is developed for streamflow

* Corresponding author.
** Corresponding author at: School of Technology, Ilia State University, 0162, Tbilisi, Georgia.
E-mail addresses: rana@gzhu.edu.cn (R.M.A. Ikram), ewees@du.edu.eg (A.A. Ewees), kulmaths@gmail.com (K.S. Parmar), z.yaseen@kfupm.edu.sa (Z.M. Yaseen), sshahid@utm.my (S. Shahid), ozgur.kisi@iliauni.edu.ge (O. Kisi).

https://doi.org/10.1016/j.asoc.2022.109739

Accepted 18 October 2022
Available online 28 October 2022

Keywords:
Streamflow forecasting
Extended marine predators algorithm
Artificial neural networks
Optimization methods

estimation in the Upper Indus Basin, a key mountainous glacier melt affected basin of Pakistan. The prediction accuracy of the novel metaheuristic algorithm (EMPA) was also compared with several benchmark metaheuristic algorithms, including the marine predators algorithm (MPA), particle swarm optimization (PSO), genetic algorithm (GA), and grey wolf optimization (GWO). The results revealed that the newly developed hybridized ANN-EMPA outperformed the other hybrid ANN methods in streamflow prediction. ANN-EMPA improved the root mean square error, mean absolute error and Nash–Sutcliffe efficiency of ANN-PSO by 4.8, 4.1 and 0.5%, ANN-GA by 6.2, 5.6 and 0.6%, ANN-GWO by 3.7, 4.4 and 0.5%, and ANN-MPA by 3.2, 7.5 and 0.3%, respectively. Month number (MN) was also examined as input to the best models to assess its impact on the prediction precision. Obtained results showed that MN generally slightly improved the models' accuracy. Results also showed that temperature-based inputs provided better prediction accuracy than only streamflow as inputs. Therefore, the ANN-EMPA model can be used for streamflow estimation from temperature data only when long-term streamflow data is unavailable.

Nomenclature section

Abbreviations

ANFIS	Adaptive neuro-fuzzy inference system
ANN	Artificial neural networks
ARMA	Auto Regressive Moving Average
BRTs	Bagged regression trees
BNN	Bayesian neural network
CARTs	Classification and regression trees
CF	Control factor
ERNN	Elman Recurrent Neural Network
EGEP	Ensemble gene expression programming
EMPA	Extended Marine Predators Algorithm
FAD	Fish Aggregating Devices Effect
GA	Genetic Algorithms
GP	Genetic programming
GWO	Grey wolf optimization
W	Inertia weight
LSSVM	Least squares support vector machine
LSTM	Long short-term memory
ML	Machine learning
UIB	Upper Indus Basin
MPA	Marine Predators Algorithm
Max	Maximum
Min.	Minimum
MLR	Multiple linear regression
MARS	Multivariate adaptive regression splines
pm	Mutation Percentage
mu	Mutation probability
PSO	Particle swarm optimization
R	Random number
Std. dev.	Standard deviation
SS	Stepsize
GBRTs	Stochastic gradient boosted regression trees
S_{flow}	Streamflow
SVR	Support vector regression
R_B	Vector with random number
v	Velocity
WAPDA	Water and Power Development Authority
SWAT	Water Assessment Tool

Parameters

C1	Cognitive coefficient
C2	Cognitive coefficient
P	Constant
N	Data quantity
$\overrightarrow{E_i}$	Elite location
α	Grey wolf leaders Alpha
Ω	Grey wolf position Omega
I	Iterations
$\overrightarrow{R}_L$	Levy's motion
$\overrightarrow{X_{max}}$	Maximum bounds of the dimensions.
$\overrightarrow{X_{min}}$	Minimum bounds of the dimensions.
f	Object function
$\overrightarrow{U}$	Random vector
Q_o	Real values
ω	Rest of wolves Omega
Q	Streamflow
β	Subordinate wolves Beta
Q_c	Target values
$\overrightarrow{R}$	Uniform number in [0,1]
δ	Variable to control the differential variation
T	Various temperature

Symbols

$\otimes$	Entry wise multiplication

1. Introduction

Streamflow (S_{flow}) forecasting is crucial for most water resources management and river sustainability practices [1]. It is also important to water resource planners, hydrologists, emergency response providers, water system managers, and policymakers. S_{flow} is significantly influenced by multifaceted stochastic processes, such as temperature, rainfall and seasonal changes, and nonlinear watershed responses [2,3]. Hence, S_{flow} forecasting is tedious, especially with sufficient lead times. Pagano et al. [4] outlined the challenges to accurate S_{flow} forecasting as the automation of real-time data acquisition process, S_{flow} models being mere simplifications of the actual processes, and uncertainty in S_{flow} prediction processes. Two rainfall-runoff modelling approaches are currently available for S_{flow} forecasting; these are the physical-based and data-based models. The physical-based models include the relevant physical laws that govern S_{flow} generation and watershed response. Therefore, they require enough effort for

implementation, parameterization, and calibration. Sometimes, they also may need many experimental data. The S_{flow} forecasting performance of process-based models is more reliable than the other prediction techniques [5]. Still, the major issues are their simplification and a large volume of data requirements [6].

The data-driven computerized models, including machine learning (ML) and statistical models, revealed a sophisticated progression in modelling S_{flow} [7–10]. These empirical models are easy to implement since they are based on previous observations and are not dependent on the physical processes [11]. The statistical models strive to establish the input–output relationship with any internal process assumption [12]. The statistical models are easier to develop and offer more reliable predictions when trained on robust and representative datasets. S_{flow} forecasting using different statistical models has been widely reported. Some statistical models include the Auto Regressive Moving Average (ARMA), wavelet support vectors, and functional regressions [13–16]. The auto-regressive or time series models rely on the observed S_{flow} and precipitation to approximate the complex nonlinear hydrologic processes [17,18]. Time series models assume the present event's dependence on the previous events, often accustomed to seasonality and error functions [19]. Training and testing time series models also demand a large observational dataset [19,20]. Hence, they show less accuracy if the input data is not within the calibration data range [21]. However, auto-regressive models have been successfully applied for forecasting many hydrologic processes. For instance, the ARMA model has been used by Toth et al. [22] for real-time flood prediction. The study reported better performance of ARMA model than the simple physically-based models.

Several ML algorithms have also been applied in the modelling of nonlinear hydrologic processes; for instance, artificial neural networks (ANN) [23,24], support vector regression (SVR) [25,26], genetic programming (GP) [27], and adaptive neuro-fuzzy inference system (ANFIS) [28,29]. They have been investigated for different hydrological applications, such as sediment transport modelling [30–32], rainfall-runoff modelling [9,33], and drought prediction [34–37].

Sudheer et al. [38] reported the ANN-based models for daily rainfall-runoff prediction in India. The study first identified the exogenous input parameters relating to different kinds of autocorrelation in S_{flow} and rainfall data series. Their results showed ANN-based models' better performance than the multiple linear regression (MLR) and ARMA models. Demirel et al. [39] reported ANN-based daily S_{flow} forecasting in Portugal; they reported better suitability of the ANN models than the semi-distributed Soil and Water Assessment Tool (SWAT) model in predicting peak discharge. The use of three ML models (Bayesian neural network (BNN), SVR and Gaussian process) for daily S_{flow} forecasting of a catchment of British Columbia has been reported by Rasouli et al. [40]. The study considered MLR the benchmark model and showed nonlinear models' superiority to MLR, with BNN as the best among the nonlinear models.

Erdal and Karakurt [41] investigated two ensemble learning models' performance, bagging and stochastic gradient boosting, in S_{flow} forecasting and compared their performance with the classification and regression trees (CARTs) and SVR models. The monthly S_{flow} forecasting performance of the bagged regression trees (BRTs) and stochastic gradient boosted regression trees (GBRTs) models were found better than those of CART and SVR models. In another study, the monthly S_{flow} forecasting performance of three ANN models was evaluated by Mehr et al. [42]. Their results revealed the generalized regression ANN, radial basis function, and feed-forward back-propagation models are suitable for modelling successive station S_{flow}. Yaseen et al. [43] developed a hybrid ML model by merging the Grey model with a rolling S_{flow} forecasting method over the tropical region. The study also integrated two ANN algorithms, i.e., Elman Recurrent Neural Network (ERNN) and back-propagation and found the hybrid ML model with ERNN to achieve better prediction accuracy than the model with back-propagation algorithm. S_{flow} pattern forecasting using three ML models (multivariate adaptive regression splines (MARS), the M5 Model Tree, and the least squares support vector machine (LSSVM)) was examined by Kisi et al. [44] in Turkey. The study showed the LSSVM model to provide better S_{flow} forecasting relying on climate signal information.

Although many ML-based S_{flow} forecasting models have been developed for different climate zones and time scales, the motivation to explore new models in this research domain is still in progress. Recently, an ensemble gene expression programming (EGEP) model has been introduced by Rahmani-Rezaeieh et al. [45] for 1- and 2-day ahead S_{flow} prediction. The new model's performance was benchmarked with classic GP and ANN models. The new EGEP model performed better than the comparative models. Cheng et al. [46] reported daily and monthly S_{flow} forecasting using ANN and long short-term memory (LSTM) models with a long lead-time. The study showed that the LSTM model achieved better lead-time S_{flow} prediction but was inefficient in multi-month prediction due to the unavailability of a large monthly dataset for model training. Hence, LSTM was presented as a suitable tool for daily S_{flow} forecasting to make strategic water resource management decisions. S_{flow} forecasting using a SWAT-Variable Source Area model (SWAT-VSA), ANN, ARMA, and a Bayesian ensemble model has been reported by Wagena et al. [47]. The study found that the total S_{flow} and peak flow is better predicted by the SWAT-VSA and ANN models while low flows are underpredicted. The hybridization of two or more models demonstrated a remarkable progression in S_{flow} forecasting by several researchers in the literature [48–52]. Recent development conducted on monthly S_{flow} forecasting using a hybrid Gaussian mixture model-extreme gradient boosting (GMM-XGBoost) model was performed by Ni et al. [53]. The GMM was used for S_{flow} clustering into several groups based on a tree-based model and XGBoost to fit each cluster's data. The novel algorithm's accuracy was benchmarked with standalone SVM and XGBoost models. The study's outcome suggested the suitability of the XGBoost for S_{flow} forecasting as it generally performed better than the SVM. The developed GMM-XGBoost model was also a better alternative for reliable and accurate predictions to aid optimal water resources management. Based on the reported literature review, the current research attempted to develop a new hybrid ML model by integrating ANN with the newly explored nature-inspired optimization algorithm, Marine Predators Algorithm (MPA) for S_{flow} forecasting. First, the MPA was enhanced based on the mutation and crossover operators (EMPA), and then hybridized with ANN to produce a robust hybrid predictive model. Several optimization algorithms were used as the benchmark for validation, such as particle swarm optimization (PSO), GA: genetic algorithm, GWO: grey wolf optimization.

The paper is organized as follows. A brief description of the case study area is provided in Section 2. The data used in the study and their sources are also provided in this section. This follows the method section, where all the ML and optimization algorithms used in this study are described. Besides, the new method proposed for improving S_{flow} prediction is also described with a flowchart. The obtained results are presented in Section 4. Finally, conclusions made from the study are given in Section 5.

2. Case study

The Upper Indus Basin (UIB) of northern Pakistan was selected as a case study for the current study. The UIB is considered the breadbasket of Pakistan due to its crucial role in

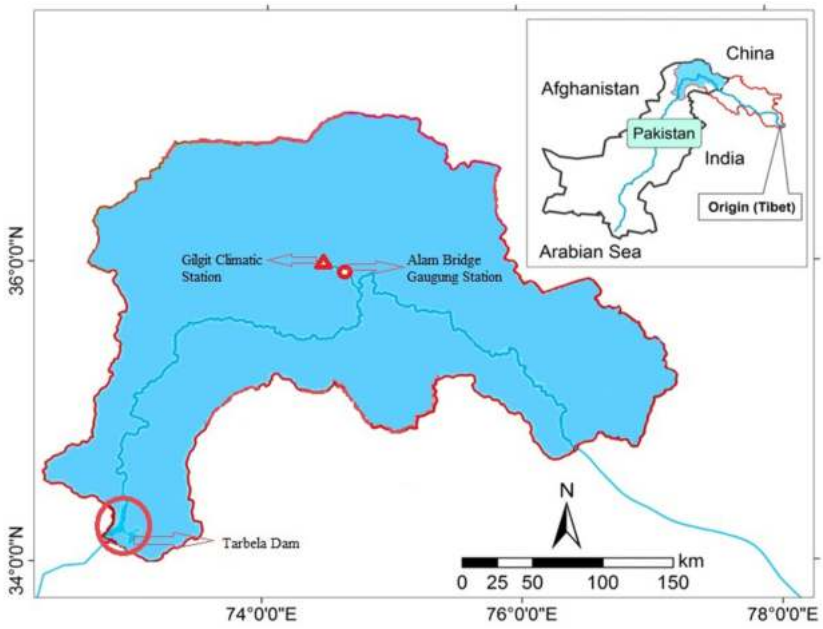


Fig. 1. The location of the Upper Indus Basin in northern Pakistan.

Pakistan's agriculture and socio-economic development. The UIB is located between 32.48°–37.07°N and 72.11°–81.83°E, covering a basin area of 65,400 km ^{2} with a mean altitude of 3750 m. The UIB is surrounded by the world's three highest mountainous, Hindukush, Karakorum and Himalayas. One of Pakistan's main hydropower installed projects, the Tarbela dam with 4888 MW generation capacity, is situated at the basin's upstream. Therefore, this basin is also called the Tarbela basin, with a river length of 1150 km above the Tarbela reservoir.

The UIB contribute 70% of the flow to the Tarbela reservoir, mainly through glacier melt. Hence, the temperature has a significant impact on the basin's streamflow. Due to a large amount of glacier melt, accurate streamflow estimation is challenging in this mountainous snow-fed basin. Therefore, this study selected this basin to apply the proposed robust models. For streamflow data, hydraulic gauging station Alam bridge (with 1280 elevation) in the UIB basin was chosen (Fig. 1). This station is located at the confluence of two main streams of UIB (Hunza and Gilgit). Therefore, it is a key location in the basin.

The temperature data recorded at Gilgit meteorological station (at 1460 Elevation) and streamflow data at Alam Bridge Streamflow Station for 35 years (1974–2008) were collected from the Water and Power Development Authority (WAPDA) of Pakistan. A brief statistical summary of the data is provided in Table 1. For better application of selected methods, 35 years dataset was divided into three proportions with a ratio of training 70% (1974–1998), validation 15% (1999–2003) and testing 15% (2004–2008), as shown in Fig. 2.

Table 1
The statistical summary of the data used in the study for the period 1974–2008.

Statistics	Whole data	Training	Validation+Testing
	Streamflow, m^3/s		
Mean	611.1	606.4	622.7
Min	76.1	79.6	76.1
Max	2738	2738	2541
Median	233.0	220.2	274.0
Skewness	1.253	1.289	1.160
Std. dev	674.6	687.2	642.1
	Temperature, °C		
Mean	15.8	15.7	16.1
Min	1.10	1.10	2.41
Max	30.7	30.7	27.9
Median	16.3	16.1	16.6
Skewness	−0.090	−0.062	−0.164
Std. dev	8.17	8.27	7.90

3. Methods

A brief description of the ML method (ANN) and optimization algorithms (PSO, GA, DE, GWO, MPA and EMPA) is given in this section.

3.1. Artificial Neural Networks (ANN)

This algorithm was formulated by Pitts [54], based on a computational model for neural networks called threshold logic. Rosenblatt [55] proposed a single layer perceptron to model human brain functioning, recognize objects, and process visual data. The learning capabilities and pattern-matching properties of ANN endorsed them to solve many problems that otherwise would have been very difficult or impossible to solve by standard statistical methods and computational techniques [56]. In addition to getting perceptions of the human brain's functionality, ANNs can be an independent tool. Thus till the 1980s, ANNs were used for various purposes. ANN works by establishing a connection between different nodes called neurons arranged in different layers interconnected and assigned with randomized weights capable of solving simple mathematical calculations [57]. These nodes identify the relation between the targeted values and the input data sets [58]. Input signals received by these neurons are organized in different layers to convert them into a single output. This output is further transferred to the other layer to conduct the same process and provide the output [59].

Generally, ANN comprises input, hidden, and output layers (Fig. 3). Each layer consists of several nodes and neurons with assigned weights to perform simple operations to compute the output [60]. ANNs are adaptive systems and thus suitable for modelling the constantly changing population, segmentation, classification [61]. ANN has been widely used in solving real-life

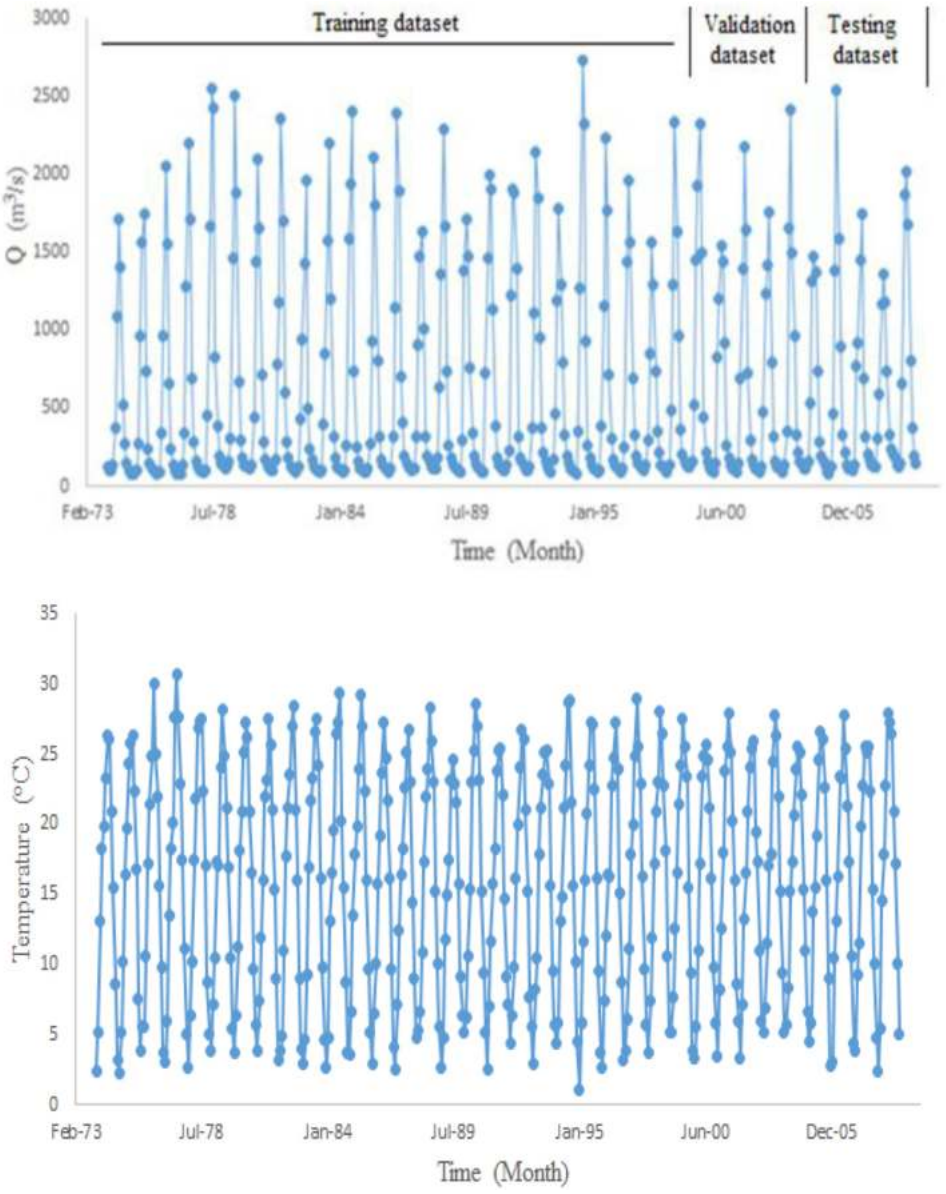


Fig. 2. Time series of streamflow and temperature data used in the study.

problems such as forecasting precipitation, droughts, streamflow, groundwater level, pollutant concentration etc. [62–67].

3.2. Genetic Algorithms (GA)

GA is an optimization and search method developed based upon the characteristics of living beings' biological evolution. Holland [68] introduced the theoretical concept of this probabilistic model in 1970. It forms the underlying people into excellent people and makes every individual liable. The fitness function measures the degree of every individual's brilliance adapt to the environment. GA has been applied in many research areas like environmental, statistics, applied mathematics, and expert systems for solving intrigue problems. The decision search hyperplanes of GA is a k-dimensional issue. The hereditary calculations ideally apportion preliminaries to the hyperplanes to find the ideal solution [69,70]. The main formation of GA is as below.

(a) Individual fitness evaluation
(b) Gene pool formation
(c) Recombination and mutation

The working procedure of GA is illustrated in Fig. 4. For more detail information, the readers are referred to original source [68].

3.3. Particle swarm optimization (PSO)

Particle swarm optimization (PSO) is conceptualized from natural living creatures' social movement to find a food source. This food searching conduct is related to an advanced search for non-direct conditions in an open esteemed pursuit space. Kennedy and Eberhart in 1995 [71] developed PSO as the new model for global optimization. This model has become very popular in many areas of research. Researchers worldwide use PSO by adjusting its parameters. In the PSO, particles go through the hunt space and try to find the neighbourhood's position (Fig. 5). This neighbourhood is defined for every separate particle as the subset of particles. The Euclidean neighbourhood is used in the

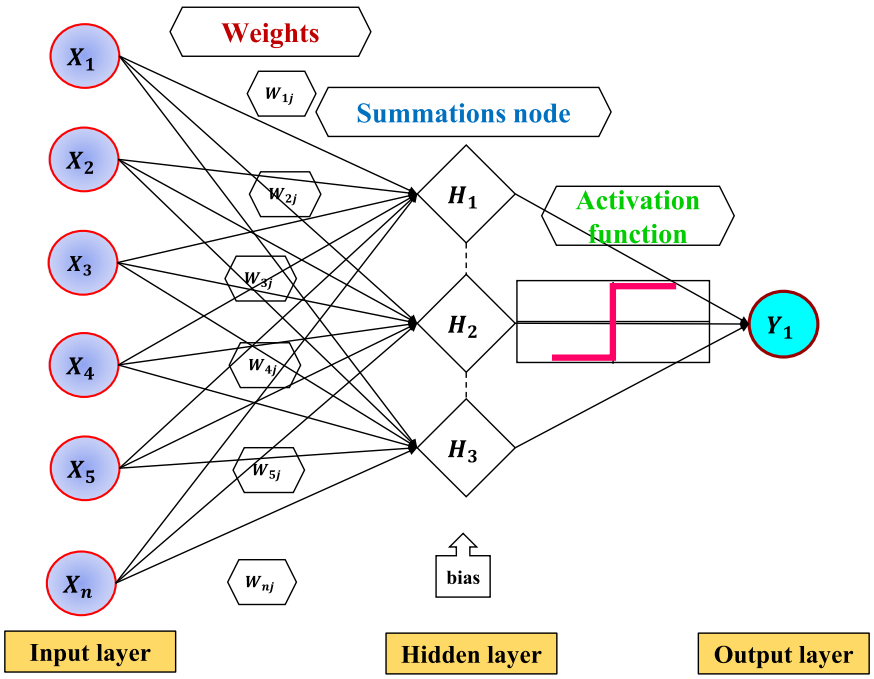


Fig. 3. A simple perceptron neural network.

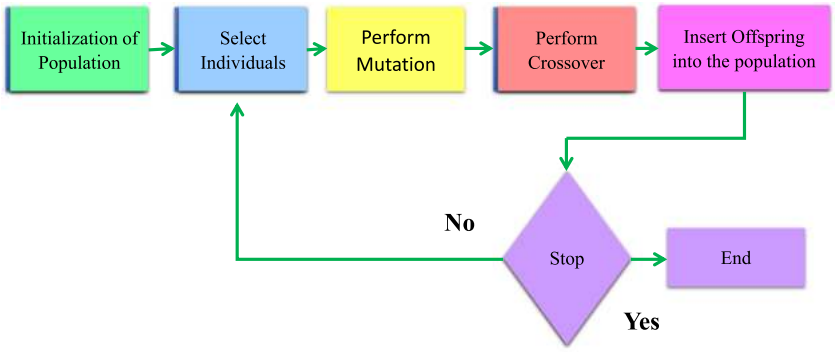


Fig. 4. The working procedure of GA.

first developed model to communicate particles to calculate the particles' distance. The Euclidean model deserted for less computationally models when the exploration centre is moved [72, 73]. The readers can reach more information from the original source [71].

3.4. Grey wolf optimization (GWO)

Grey wolf optimization (GWO) is a recently developed evolution model based on the reproduction of grey wolves. There is a higher position and management of the grey wolves in the pack. Mirjalili et al. [74] developed this GWO metaheuristic model. It imitates the hunting habits and social ladder of grey wolves by introducing α, β, Ω, and ω. For the constrained wildlife populace, the problem is finding a feasible solution for the cost-efficient level of habitat safety. These constraints need a high likelihood of getting the goal populace size. The prey position is evaluated by α, β, Ω and it updates its values around the prey. At the last position, α, β, Ω define the circle of the search space. The major aim of this GWO model is to find the prey by the shortest possible path [75–77]. The following four processes are used to find the possible shortest path.

A. Exploration B. Enclosing prey
C. Hunting D. Exploitation prey

The process to find the best shortest path is discussed using a flow chart in Fig. 6.

3.5. Marine Predators Algorithm (MPA)

The nature-inspired marine predators algorithm (MPA) optimization uses three steps to model the whole life of prey and predator. These three steps are based upon the prey and predator's velocity proportion, vital for transmitting the optimization development. A high-velocity proportion is an important characteristic in the first step, while unity and low-velocity proportions are in the second and third steps, respectively. These stages are identified with the guidelines administered on hunter and prey development while addressing hunter and prey development in nature. Besides, MPA employs another stage known as Fish Aggregating Devices Effect (FAD) to avoid local minima. The details of the stages are as follows [78].

(a) **High-Velocity proportion:** In this step of high velocity ($v \geq 10$), prey travels faster than predator as prey moves quickly to find its food until the predator moves. Here initial iterations for investigation matters are vital for the

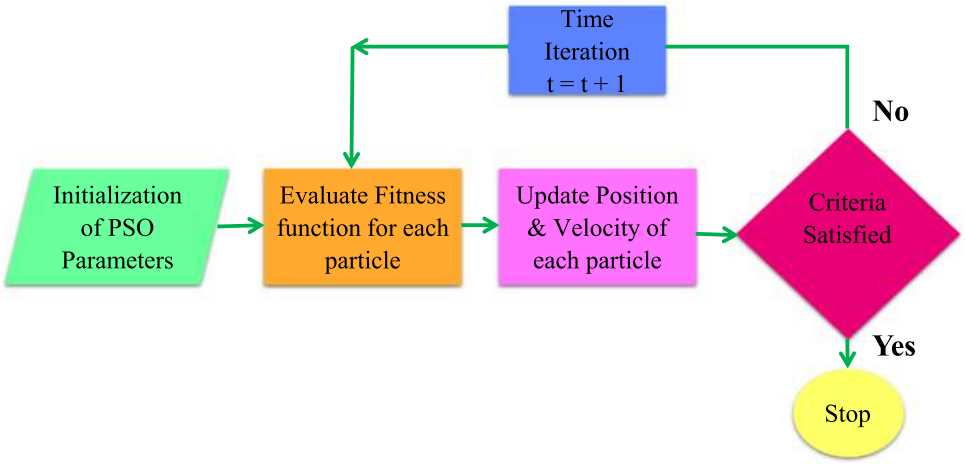


Fig. 5. The working procedure of PSO.

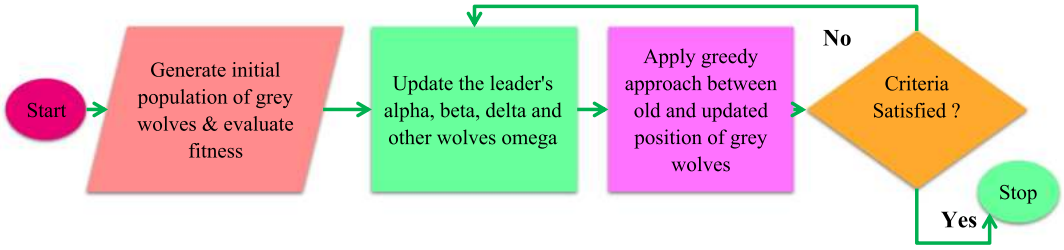


Fig. 6. The working procedure of GWO.

optimization process. Mathematically it can be described as below.

$$\text{While } I < \frac{1}{3}\text{Max}\, I \qquad I = \text{iterations} \tag{1}$$

$$\overrightarrow{SS}_i = R_B \otimes \left(\overrightarrow{E_i} - R_B \otimes \overrightarrow{P_i}\right); \quad i = 1, 2, 3, \ldots, n. \tag{2}$$

SS = stepsize, R_B = Vector having random number, $\otimes$ = entry wise multiplication

$$\overrightarrow{P_i} = \overrightarrow{P_i} + P.\overrightarrow{R} \otimes \overrightarrow{SS_i}$$

$$P = 0.5 \text{ (Constant)}, \quad \overrightarrow{R} = \text{uniform number in } [0, 1] \tag{3}$$

here I is the present iteration. In this stage, the progression size or development speed is more.

(b) **Unity-Velocity Proportion:** In the second stage, prey and predator travel speeds are the same when both search for their food. This stage is mid of optimization, with prey being accountable for exploitation and predator for exploration. This step can be developed mathematically for $v \approx 1$, and prey travels with Levy's motion and predator move in the Brownian motion as below:

$$\text{For } \frac{1}{3}\text{Max}\, I < I < \frac{2}{3}\text{Max}\, I \tag{4}$$

For the first half of the populace

$$\overrightarrow{SS}_i = \overrightarrow{R}_L \otimes \left(\overrightarrow{E_i} - \overrightarrow{R}_L \otimes \overrightarrow{P_i}\right); \quad i = 1, 2, 3, \ldots \frac{n}{2}. \tag{5}$$

$$\overrightarrow{P_i} = \overrightarrow{P_i} + P.\overrightarrow{R} \otimes \overrightarrow{SS_i} \tag{6}$$

For the second half of the populace

$$\overrightarrow{SS}_i = \overrightarrow{R}_B \otimes \left(\overrightarrow{R}_B \otimes \overrightarrow{P_i} - \overrightarrow{P_i}\right); \quad i = \frac{n}{2}, \frac{n+1}{2}, \ldots, n. \tag{7}$$

$\overrightarrow{P_i} = \overrightarrow{E_i} + P.CF \otimes \overrightarrow{SS_i}$, here $CF = \left(1 - \frac{I}{\text{Max}I}\right)^{\left(2\frac{I}{\text{Max}I}\right)}$ is controlling unit step size for predator motion. $\overrightarrow{R}_B \otimes \overrightarrow{P_i}$ emulates the hunter movement by the Brownian approach, where the prey can refresh its area depending on the Brownian hunter movement.

(c) **Low-Velocity Proportion :** The predator travels faster than the prey with the high exploitation process for optimization in the third stage. The velocity in this stage is 0.1, and the predator follows Levy's motion. This can be expressed in the mathematical form as given below:

$$\text{For } I > \frac{2}{3}\text{Max}\, I \tag{8}$$

$$\overrightarrow{SS}_i = \overrightarrow{R}_L \otimes \left(\overrightarrow{R}_L \otimes \overrightarrow{E_i} - \overrightarrow{P_i}\right); \quad i = 1, 2, 3, \ldots, n. \tag{9}$$

$$\overrightarrow{P_i} = \overrightarrow{E_i} + P.CF \otimes \overrightarrow{SS_i} \tag{10}$$

$\overrightarrow{R}_L \otimes \overrightarrow{E_i}$ depicts Levy's motion and uses Elite location to recreate the hunter movement to refresh the prey's position.

(d) **Fish Aggregating Devices Effect (FAD):** The FAD is the local optima point in this stage, whereas the MPA tries to overcome this issue. The following equation mathematically presents the FAD:

$$\overrightarrow{P_i} = \begin{cases} \overrightarrow{P_i} + CF[\overrightarrow{X_{min}} + \overrightarrow{R}] \otimes (\overrightarrow{X_{max}} - \overrightarrow{X_{min}}) \otimes \overrightarrow{U} & \text{if } r \le FAD \\ \overrightarrow{P_i} + [FAD(1-r) + r](\overrightarrow{\text{Pr}}_1 - \overrightarrow{\text{Pr}}_2) & \text{if } r > FAD \end{cases} \tag{11}$$

where FAD denotes the probability of effect of FAD. $\overrightarrow{U}$ denotes a random vector contains [0, 1]. r is a random value in [0, 1]. r_1 *and* r_2 random indexes of P. $\overrightarrow{X_{max}}$ *and* $\overrightarrow{X_{min}}$ are maximum and minimum bounds of the dimensions.

3.6. Extended Marine Predators Algorithm (EMPA)

The extended MPA is also a nature-inspired algorithm of optimization. The performance of the new proposed EMPA model is better than the baseline MPA model. Minimax improvement issues can depict numerous plan problems. Traditional strategies for tackling these problems are restricted to the problem. As of late, developmental calculations, especially coevolutionary advancement methods, have been applied to the minimax problem. Another strategy for improving the minimax method utilizing transformative calculations is recently proposed. This calculation is better and has shown a vigorous post position improvement.

This paper proposes a extended version of MPA by extending the searching using the mutation and crossover operators to facilitate more flexibility in discovering new points in the search space. Thus, it overcomes the limitations of MPA like local optimum trapping and premature convergence.

In EMPA, the original MPA is modified using a mutation operator in its structure which generates a mutation vector v_{mu} using the following formula:

$$v_{mu,i} = v_{p1} + \delta \times (v_{p2} - v_{p3}) \tag{12}$$

where $i = 1, 2, \ldots, N$. v_1, v_2, and v_3 are selected randomly from the population. $\delta \in [0, 2]$ to control the differential variation.

The mutation vector is subsequently used in the crossover stage by randomly updating a point in the current solution depending on a crossover probability, which is set to 0.2 in this study. The quality of the new vector is checked using the fitness function. The new solution is used if the produced value is better than the old fitness value; otherwise, the old one is retained. Furthermore, the EMPA is applied to optimize the weights and biases parameters of the ANN. The ANN-EMPA begins by defining all parameters and receiving the experimental data divided into the train, validate, and test sets. The EMPA searches for the ANN model's best weights and biases parameters to optimize the ANN. The results are evaluated using Eq. (13) to test the candidate parameters' quality.

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(Q_{oi} - Q_{ci})^2 \tag{13}$$

where Q_o defines the real values, Q_c defines the target values, and N defines the sample size. The candidate parameters are chosen based on the smallest error. The EMPA evaluates the parameters until reaching the maximum number of fitness function evaluations. The selected parameters are then passed to optimize the ANN model to start the testing phase. In this study, the EMPA is tested in predicting daily streamflow. Three performance measures were used to evaluate the performance of EMPA. Fig. 7 illustrates the entire process of the EMPA.

The computational complexity of the EMPA can be expressed as in Eq. (14).

$$O(t(n * d + mo + f * n)) \tag{14}$$

where mo $= O(t*d)$, t denotes the iterations number, n represents population size, f denotes the objective function, and d is the dimension of the problem.

4. Application and results

A novel ML method, ANN-EMPA, was applied for streamflow prediction using temperature and streamflow data as inputs. The outcomes of the new method were compared with the ANN hybridized with marine predators algorithm (MPA), grey wolf optimization (GWO), genetic algorithm (GA) and particle swarm optimization (PSO) using the following statistical indices:

Table 2
Training and VT results of the ANN-PSO model.

Algorithm	Hyperparameters	Value
EMPA	FADs	0.2
	P	0.5
	beta	1.5
	mutation	0.02
MPA	FADs	0.2
	P	0.5
	beta	1.5
PSO	w	1
	wDamp	0.99
	C1	1
	C2	2
GA	pc	0.8
	gamma	0.2
	pm	0.3
	mu	0.02
	beta	8
GWO	a	in [0 2]

$$\text{RMSE: Root Mean Square Error} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[(Q_o)_i - (Q_C)_i]^2} \tag{15}$$

$$MAE: MeanAbsoluteError = \frac{1}{N}\sum_{i=1}^{N}|(Q_o)_i - (Q_C)_i| \tag{16}$$

$$NSE: Nash - Sutcliffe = 1 - \frac{\sum_{i=1}^{N}[(Q_o)_i - (Q_C)_i]^2}{\sum_{i=1}^{N}[(Q_o)_i - \overline{Q}_o]^2}, -\infty < NSE \leq 1 \tag{17}$$

where $Q_o, Q_c, \overline{Q}_o$ are observed, computed and mean observed streamflows, respectively, and N refers to data quantity. Initial parameter values of each metaheuristic algorithm are reported in Table 2. The general parameters of the experiments are as follows: the population size is equal to 25, the objective function is the RMSE, and the stopping criterion is 1250 fitness function evaluation. The experiments were conducted using Matlab 2014, running on "PC - Core i3" with "MS Windows 10" and 8 MB of ram. We optimized both weights and biases parameters for the ANN model. The parameter were set to: learning rate = 0.2, hidden layer = 10, max epochs = 1000, and the number of weights and biases = 101.

Temperature and streamflow data were split into three parts, 70% for training, 15% for validation and 15% for testing. Various temperature (T) and streamflow (Q) combinations based on correlation analysis were first considered as inputs (Table 3). For inputs combinations of streamflow and temperature, autocorrelation function (ACF) and cross autocorrelation function (CCF) is utilized on the time series of Q and T to evaluate the prominent inputs combinations. From the ACF plots of Q and CCF plots between Q and T, it is clearly visible that Q at one time lag, eleventh and twelfth time lags have prominent influence on current streamflow value, whereas the CCF plots clearly demonstrates the T prominent influence on time t, at 1, 11 and 12 previous time lags. Therefore on the basis of both plots, five inputs combinations selected i.e. (i) Qt-1, Qt-11; (ii) Qt-1, Qt-11, Qt-12; (iii) Tt, Tt-1; (iv) Tt, Tt-1, Tt-11 and (v) Tt, Tt-1, Tt-11, Tt-12.

Then, input combinations were obtained by combining the optimal inputs of T and Q. Training and validation+test (VT) efficiencies of the ANN-PSO are summarized in Table 3. The method

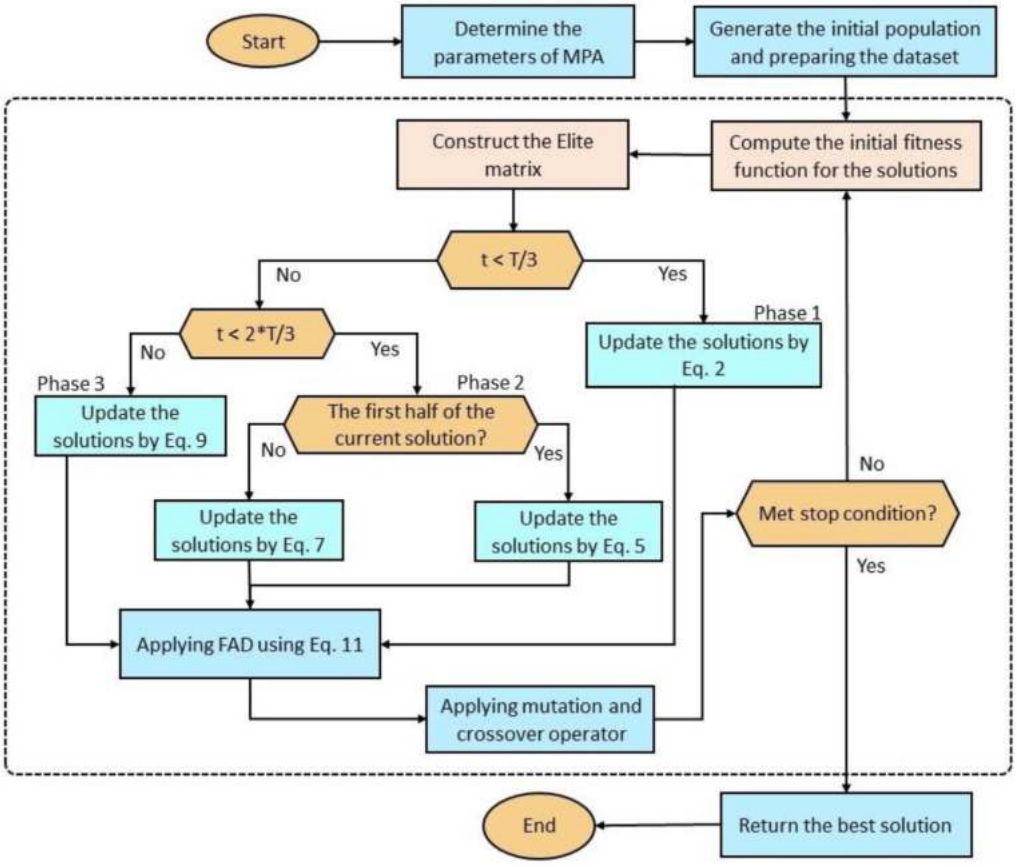


Fig. 7. The flowchart of the proposed EMPA.

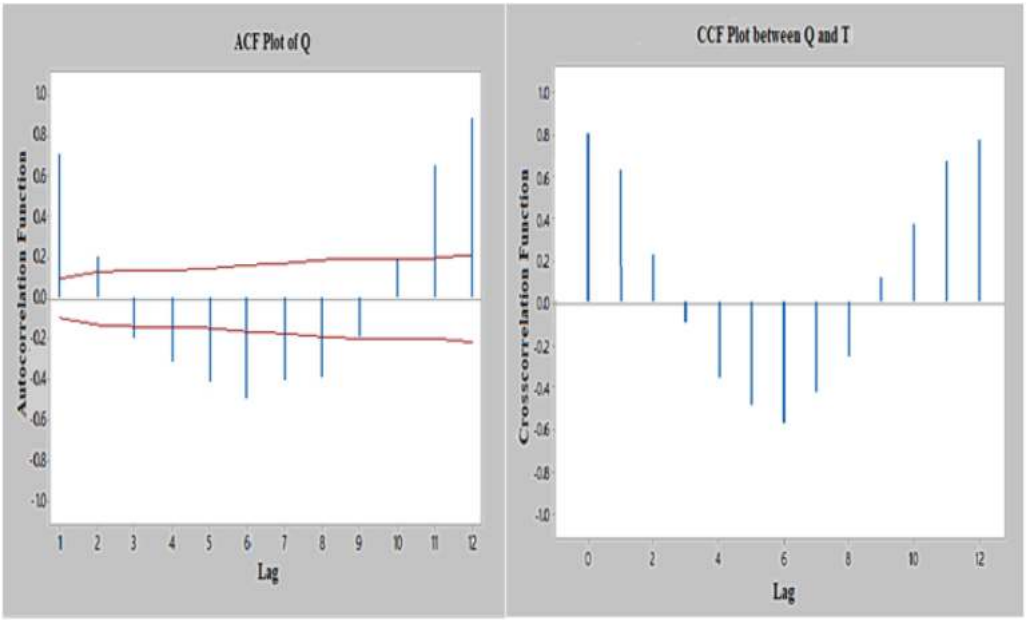


Fig. 8. ACF and CCF Plots.

showed the lowest RMSE (136.39 m^3/s) and MAE (88.06 m^3/s) and the highest NSE (0.947) for the last input combination (Qt-1, Qt-11, Qt-12, Tt, Tt-1, Tt-11, Tt-12) in the VT stage. The best T-based ANN-PSO model showed better performance than the optimal Q-based models. Combining T with Q as inputs considerably improved the model's efficiency in both training and VT stages. For example, in the VT stage, RMSE, MAE and NSE improvement of the optimal T-based model was 22.7, 13.7 and 3.3%, respectively.

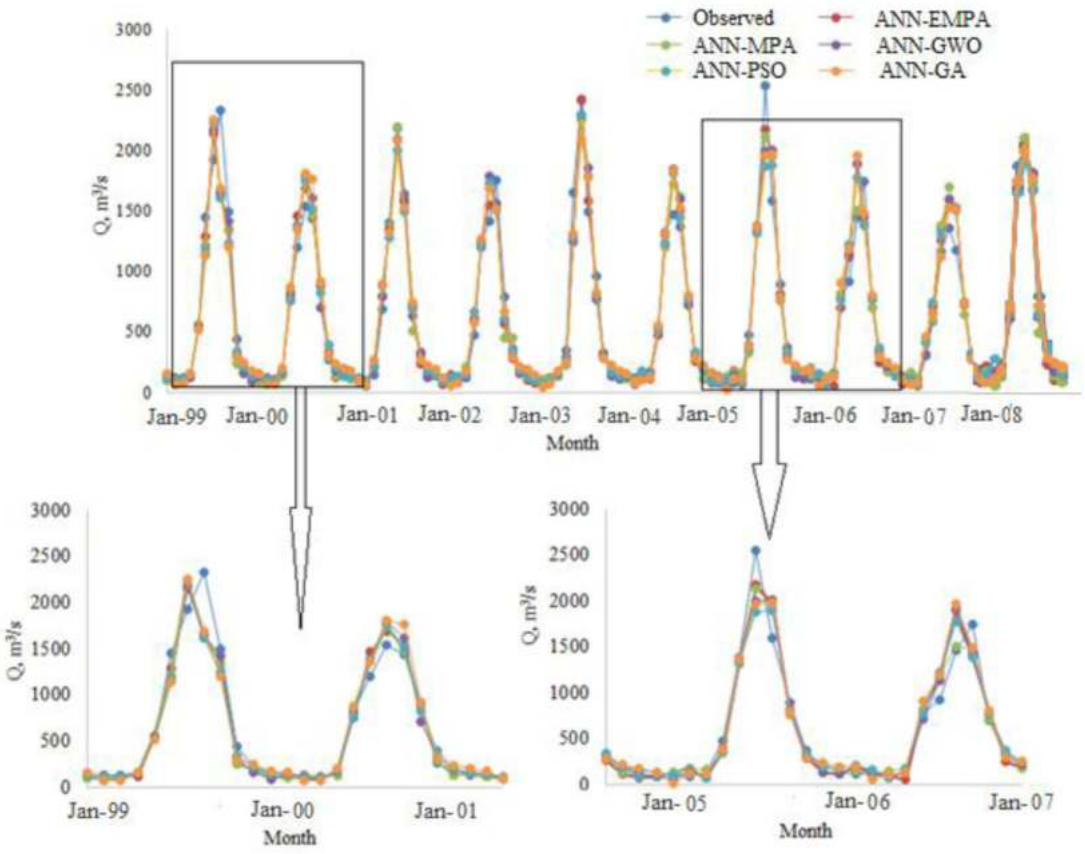


Fig. 9. Time variation graphs of the observed and predicted streamflows by different ANN-based models for the best input combination in the validation +testing period.

Table 3
Training and VT results of the ANN-PSO model.

Input combinations	Training			Validation + Testing		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Qt-1, Qt-11	179.30	98.71	0.928	196.50	113.80	0.896
Qt-1, Qt-11, Qt-12	146.40	82.52	0.954	183.20	103.10	0.909
Tt, Tt-1	180.30	102.30	0.920	196.20	119.20	0.898
Tt, Tt-1, Tt-11	152.50	87.84	0.951	178.40	108.10	0.914
Tt, Tt-1, Tt-11, Tt-12	146.90	85.74	0.940	176.50	102.10	0.917
Qt-1, Qt-11, Qt-12 , Tt, Tt-1, Tt-11, Tt-12 (best Q, best T)	128.50	69.62	0.959	**136.39**	**88.06**	**0.947**
Mean	**155.60**	**87.78**	**0.942**	**177.90**	**105.70**	**0.914**

Table 4 gives the accuracy of ANN-GA method for the training and VT stages. This method also showed the best efficiency (RMSE: 139.03 m^3/s, MAE: 89.08 m^3/s and NSE: 0.945 in testing stage) for the last input combination (Qt-1, Qt-11, Qt-12, Tt, Tt-1, Tt-11, Tt-12). Here also T-based ANN-GA model exhibited relatively better performance than the Q-based model. Merging both inputs (optimal T and optimal Q) improved the model performance in the test stage by 21.7, 12.7 and 3.2% in RMSE, MAE and NSE, respectively. It is clear from Tables 3 and 4 that the ANN-PSO prediction accuracy was higher than the ANN-PSO in S_{flow} simulation (training stage) and estimation (VT stage).

As indicated in Table 5, the difference between T-input and Q-input models was higher for the ANN-GWO. The differences were 21.6, 26.4 and 4.8% in RMSE, MAE and NSE, respectively. Merging the optimal T and Q inputs improved RMSE while the T-input ANN-GWO showed lower MAE and higher NSE than the all input model.

Table 6 reports the performance of ANN-MPA models for six input cases. As evident from the table, the model accuracy was almost the same for the optimal T-input (Tt, Tt-1, Tt-11, Tt-12) and Q-input (Qt-1, Qt-11, Qt-12). Merging the optimal T and Q inputs improved the accuracy significantly; RMSE and MAE decreased by 23.7 and 15.6%, respectively, and NSE increased by 3.5%.

Q-input model showed higher performance for ANN-EMPA (Table 7) than the T-input model. The improvement in RMSE, MAE and NSE were 4.4, 2.6 and 0.7%, respectively. The ANN-EMPA with the best T and Q inputs decreased RMSE of the Q-input model from 166.889 to 131.80 m3/s or by 21%, MAE from 97.697 to 84.44 m^3/s or by 13.6%, and decreased the NSE from 0.924 to 0.9532 or by 3.2%. For all input cases, the ANN-EMPA method performed better than the others in the training and VT stages, except for the 5th input combination (the best T-input case) of the ANN-GWO. The performance of the methods was generally

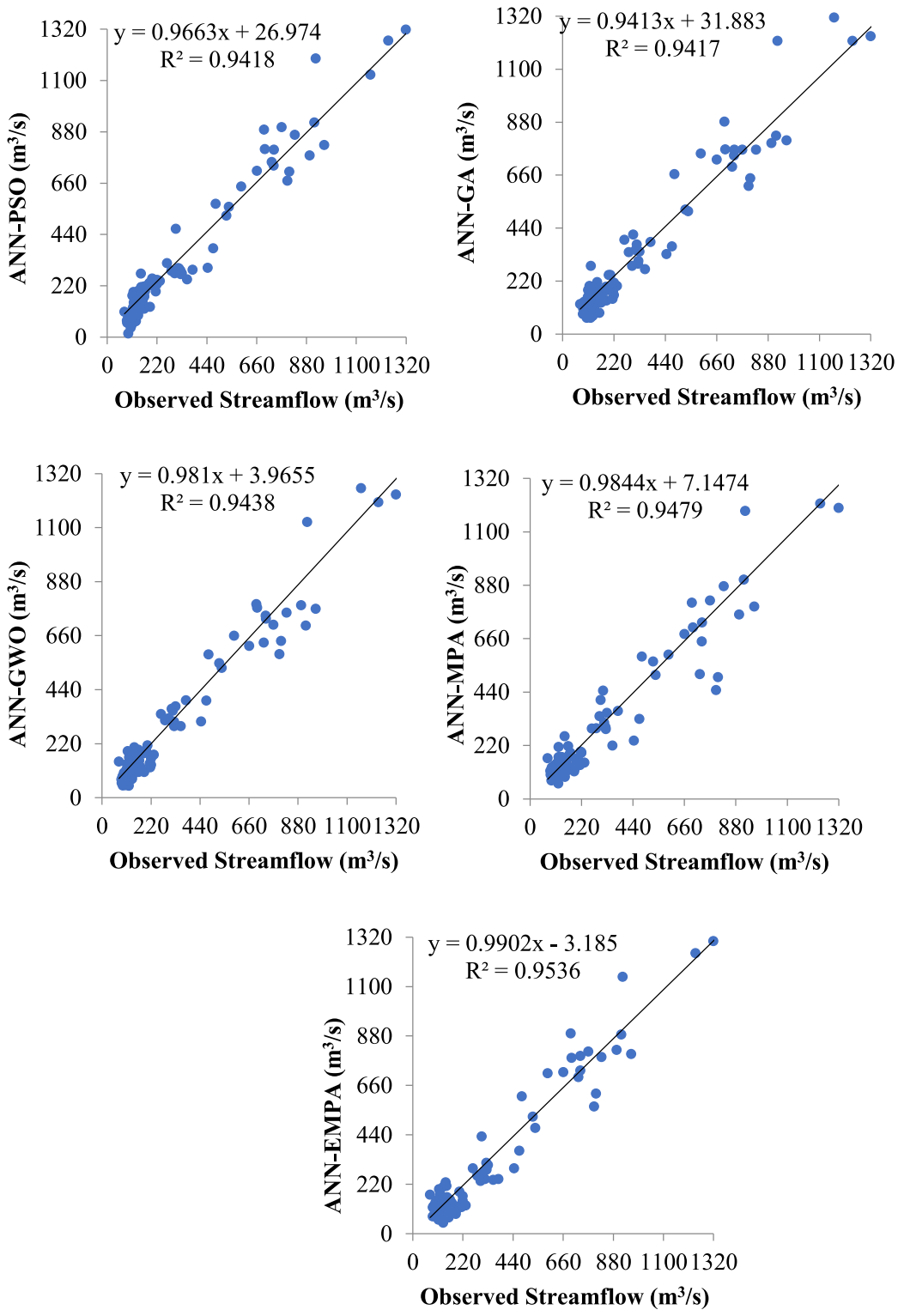


Fig. 10. Scatterplots of the observed and predicted streamflows by different ANN-based models for the best input combination in the validation +testing period.

consistent, and they provided the best accuracy for T and Q inputs together. It is observed from Tables 3 to 7 that the performance of the methods was generally good for only T input with NSE varies from 0.916 to 0.958. This is very important as S_{flow} measurement is often challenging due to technical problems, particularly in developing countries. In case of lacking Q data, only temperature data measured using a simple thermometer can be utilized to estimate streamflow.

The ANN-PSO, ANN-GA, ANN-GWO, ANN-MPA and ANN-EMPA methods are compared in Table 8 for the last input combination. In this table, periodicity input (month number, MN) was also used as input to see its influence on models' performances in both training and VT stages. Bold values in the table indicate the best models (lowest RMSE and MAE and the highest NSE).

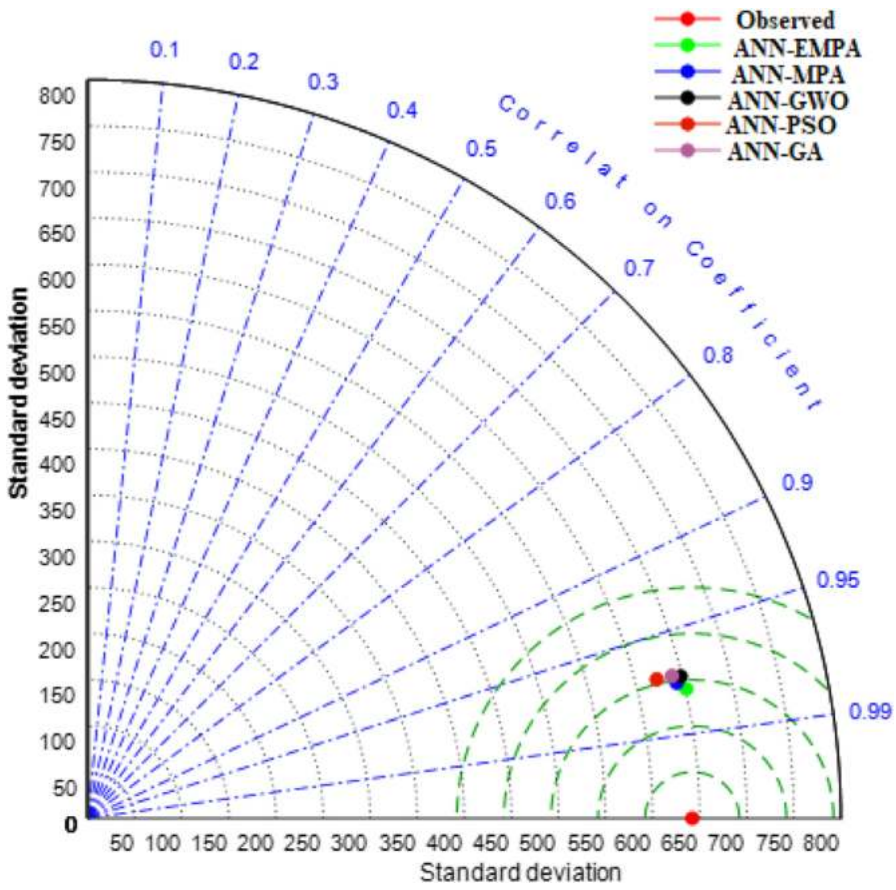


Fig. 11. Taylor diagrams of the estimated streamflows by different ANN based models in the validation +testing period.

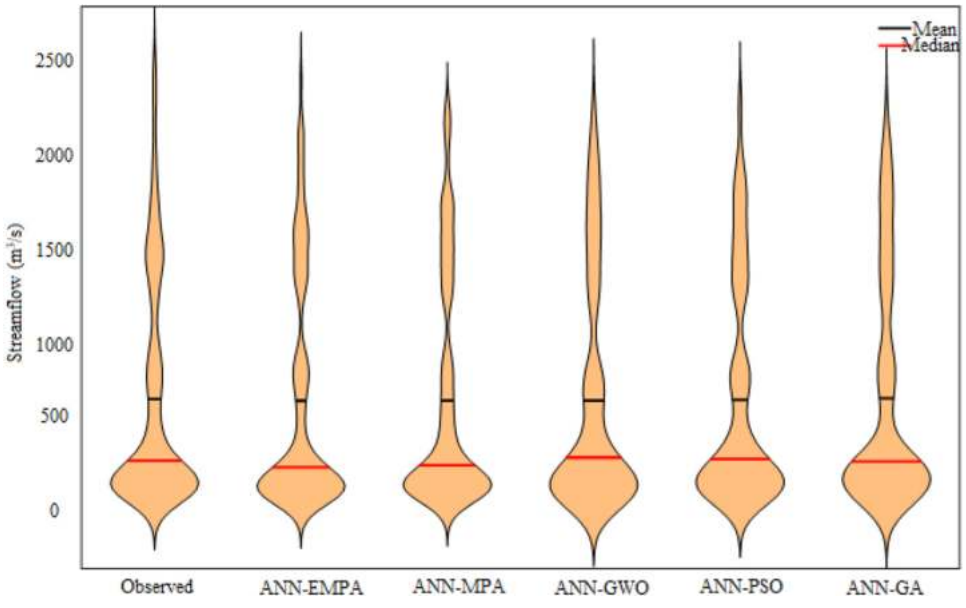


Fig. 12. Violin Charts estimated streamflows by different ANN-based models in the validation +testing period.

The MN generally improves the models' accuracy slightly (improvement in RMSE of ANN-PSO, ANN-GA, ANN-GWO, and ANN-EMPA was 0.6, 1, 0.4 and 2%, respectively), except the ANN-MPA. The ANN-EMPA improved RMSE, MAE and NSE of ANN-PSO by 4.8, 4.1 and 0.5%, ANN-GA by 6.2, 5.6 and 0.6%, ANN-GWO by 3.7, 4.4 and 0.5% and ANN-MPA by 3.2, 7.5 and 0.3%, respectively. The results clearly showed the better performance of ANN-EMPA in estimating monthly streamflow. The other models can be ranked as follows: ANN-MPA, ANN-GWO, ANN-PSO and ANN-GA, respectively. The implemented methods' performance was visually compared in Figs. 9 and 10 using hydrograph and scatterplot. Fig. 8 shows that the temporal variability of ANN-EMPA estimated S_{flow} is very close to the observed flow. Two frames were selected from the whole part as seen from the first time variation graph

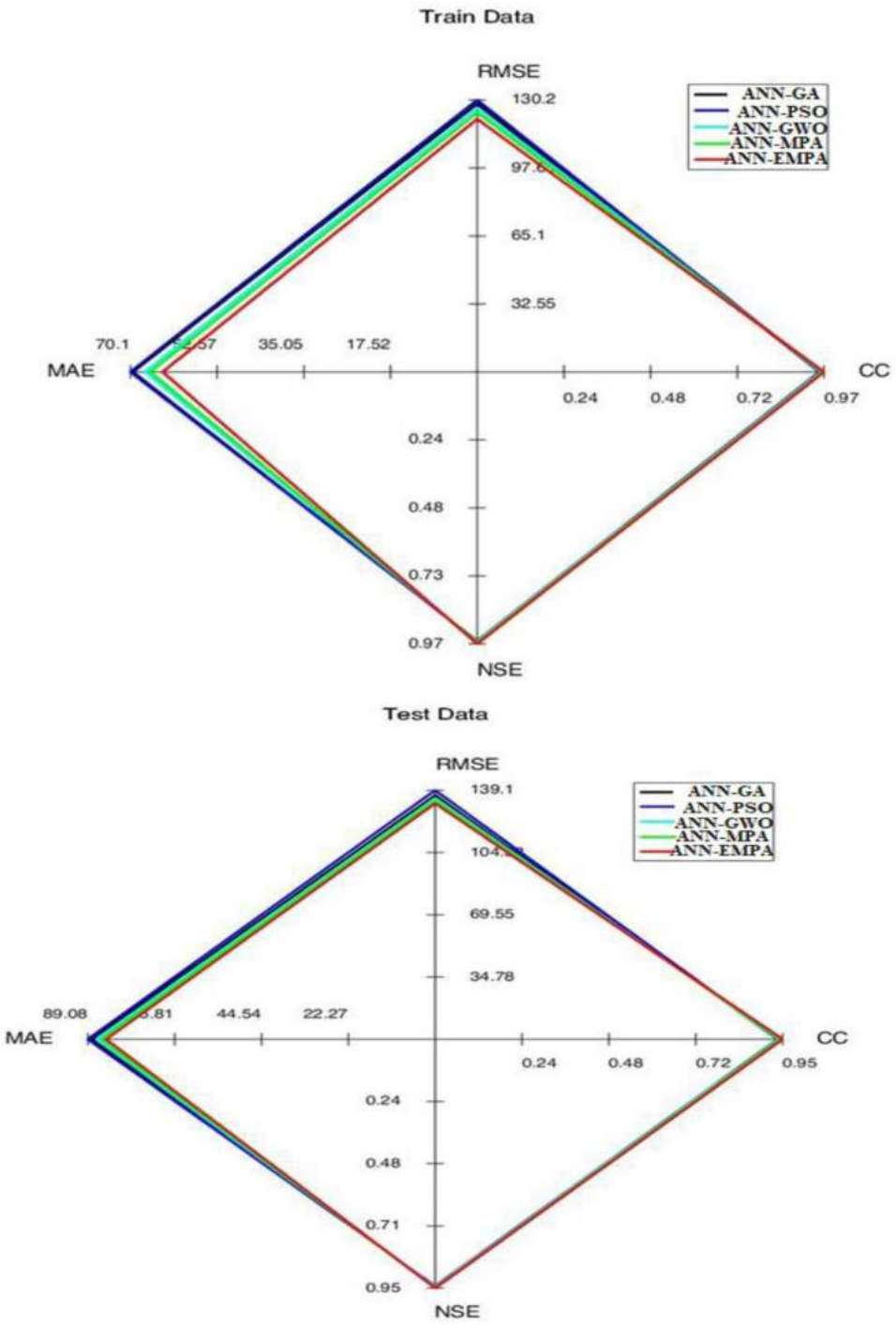


Fig. 13. Radar charts of the estimated streamflows by different ANN-based models in the training and validation +testing periods.

and these parts were zoomed in the bottom subplots. A closer look at the series revealed that the ANN-EMPA could simulate the low and most of the extremely high values well.

Fig. 9 shows that the ANN-EMPA estimates were near the observed values and less scattered than the other alternative models. The slope and bias coefficients of the ANN-EMPA were near 1 and 0, respectively and R^2 near 1 (0.9536), much better than the ANN-MPA, ANN-GWO, ANN-PSO and ANN-GA. The Taylor diagram in Fig. 11 reveals that the newly proposed method (ANN-EMPA) has lower RMSE and higher correlation, and its standard deviation is close to the observation.

Violin charts in Fig. 12 shows the similarity in the distribution of models' estimates with the observation. The similarity between the distributions of the estimated and observed streamflows was more for ANN-EMPA model. The methods were further compared using radar charts (Fig. 13). The models' differences in RMSE, MAE, NSE and correlation coefficient (CC) were visually compared using this graph. Fig. 13 also justifies the ANN-EMPA model's superiority to the ANN-MPA, ANN-GWO, ANN-PSO and ANN-GA in streamflow estimation.

Overall, the ANN-EMPA performed superior to the other methods in streamflow estimation. This indicates that EMPA can better optimize the weights/parameters of the ANN method in mapping

Table 4
Training and VT results of the ANN-GA model.

Input combinations	Training			Validation + Testing		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Qt-1, Qt-11	180.40	99.34	0.928	195.90	113.60	0.897
Qt-1, Qt-11, Qt-12	149.70	83.76	0.952	186.70	104.00	0.906
Tt, Tt-1	182.30	103.2	0.918	198.10	120.90	0.896
Tt, Tt-1, Tt-11	148.80	86.12	0.952	179.60	109.30	0.913
Tt, Tt-1, Tt-11, Tt-12	146.20	85.97	0.941	177.60	102.10	0.916
Qt-1, Qt-11, Qt-12 , Tt, Tt-1, Tt-11, Tt-12 (best Q, best T)	130.20	70.10	0.958	**139.05**	**89.08**	**0.945**
Mean	**156.30**	**88.08**	**0.941**	**179.50**	**106.40**	**0.912**

Table 5
Training and VT results of the ANN-GWO model.

Input combinations	Training			Validation + Testing		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Qt-1, Qt-11	178.60	97.43	0.930	194.60	112.90	0.898
Qt-1, Qt-11, Qt-12	144.90	82.60	0.955	177.90	101.40	0.914
Tt, Tt-1	177.30	100.4	0.929	193.20	116.20	0.901
Tt, Tt-1, Tt-11	145.60	84.49	0.954	175.80	101.50	0.919
Tt, Tt-1, Tt-11, Tt-12	145.00	84.75	0.942	175.50	100.65	0.926
Qt-1, Qt-11, Qt-12 , Tt, Tt-1, Tt-11, Tt-12 (best Q, best T)	126.30	67.13	0.961	**134.70**	**86.75**	**0.948**
Mean	**152.90**	**86.13**	**0.945**	**176.30**	**105.90**	**0.913**

Table 6
Training and VT results of the ANN-MPA model.

Input combinations	Training			Validation + Testing		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Qt-1, Qt-11	177.20	96.61	0.931	194.60	112.70	0.898
Qt-1, Qt-11, Qt-12	142.50	80.39	0.956	174.90	101.10	0.917
Tt, Tt-1	174.30	99.25	0.931	191.30	115.50	0.903
Tt, Tt-1, Tt-11	139.80	81.83	0.958	179.90	109.90	0.913
Tt, Tt-1, Tt-11, Tt-12	144.90	84.75	0.944	174.80	101.50	0.920
Qt-1, Qt-11, Qt-12, Tt, Tt-1, Tt-11, Tt-12 (best Q, best T)	124.20	66.11	0.963	**133.40**	**85.71**	**0.952**
Mean	**150.50**	**84.82**	**0.947**	**174.80**	**104.40**	**0.917**

Table 7
Training and VT results of the ANN-EMPA model.

Input combinations	Training			Validation + Testing		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Qt-1, Qt-11	176.90	95.80	0.933	193.10	112.20	0.900
Qt-1, Qt-11, Qt-12	134.10	80.01	0.961	166.90	97.69	0.924
Tt, Tt-1	172.30	98.24	0.934	189.10	114.90	0.904
Tt, Tt-1, Tt-11	142.60	83.12	0.956	177.70	107.00	0.915
Tt, Tt-1, Tt-11, Tt-12	128.50	75.73	0.964	169.50	98.30	0.928
Qt-1, Qt-11, Qt-12, Tt, Tt-1, Tt-11, Tt-12 (best Q, best T)	121.10	63.60	0.967	**131.80**	**84.44**	**0.953**
Mean	**145.90**	**82.75**	**0.952**	**172.20**	**102.70**	**0.922**

Table 8
The comparison of best models with and without periodicity for the training and VT periods.

Models	Training			Validation + Testing		
	RMSE	MAE	NSE	RMSE	MAE	NSE
Best ANN-PSO	128.40	69.62	0.959	136.40	88.06	0.947
Best ANN-PSO, MN	140.40	75.75	0.957	**135.60**	**82.71**	**0.950**
Best ANN-GA	130.20	70.10	0.958	139.10	89.08	0.945
Best ANN-GA, MN	141.30	76.18	0.957	**137.60**	**84.01**	**0.949**
Best ANN-GWO	126.30	67.13	0.961	134.70	86.75	0.948
Best ANN-GWO, MN	139.50	74.65	0.958	**134.10**	**82.95**	**0.950**
Best ANN-MPA	124.20	66.11	0.963	**133.40**	**85.71**	**0.952**
Best ANN-MPA, MN	138.70	73.84	0.958	134.90	86.47	0.951
Best ANN-EMPA	121.10	63.6	0.968	131.80	84.44	0.953
Best ANN-EMPA, MN	132.50	69.06	0.962	**129.10**	**79.29**	**0.955**

MN: Month number.

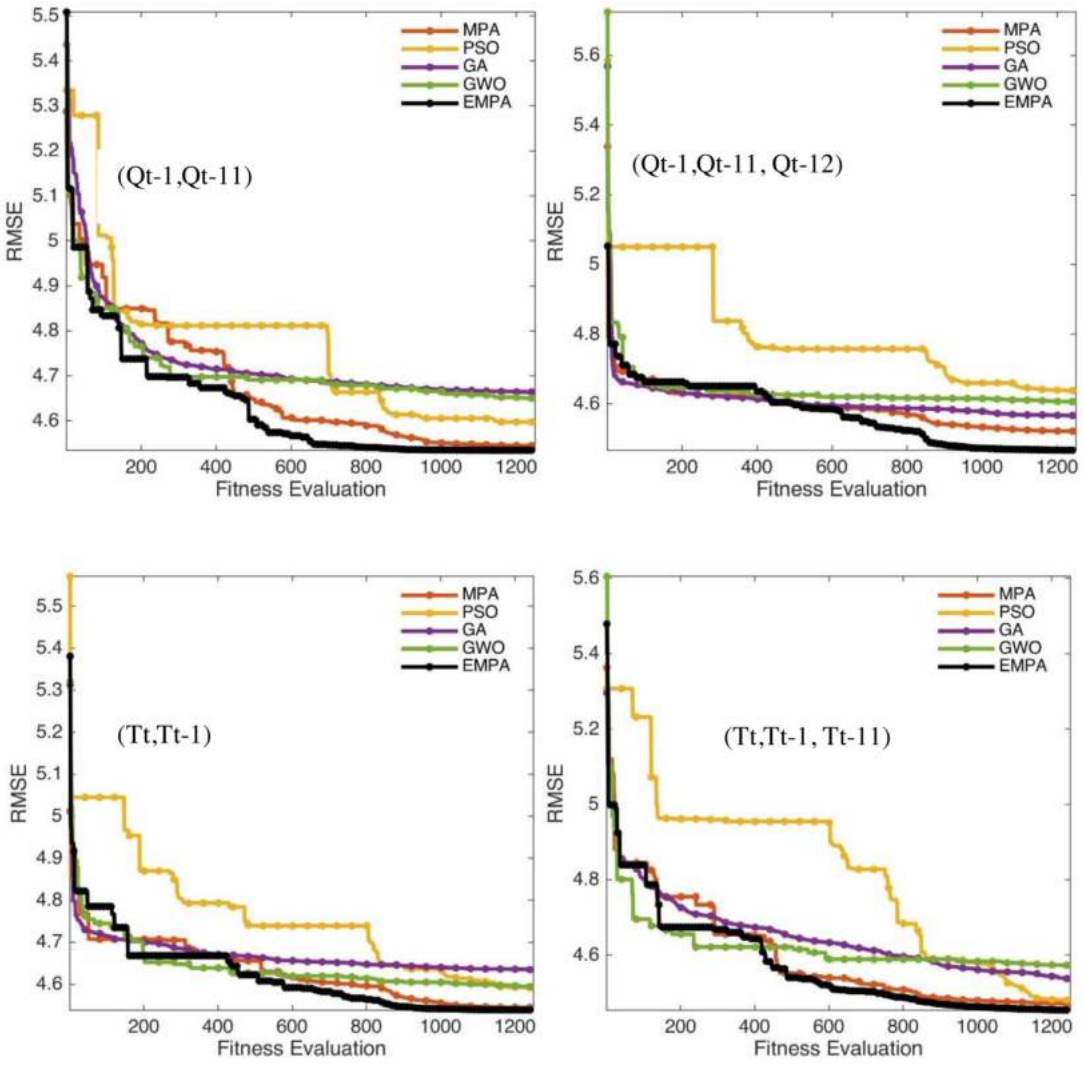


Fig. 14. Convergence accuracy of the metaheuristic algorithms utilized for streamflow prediction.

streamflow from temperature and previous streamflow information. The main advantage of the introduced algorithm is its three phases optimization. In every phase, the predator's optimum movement policy is utilized to identify the step size to catch the prey. The MPA design imitates the rules and foraging the marine predator's strategy to indicate a nature-inspired metaheuristic. This made it very close to the prototype (Faramarzi et al. 2020).

The convergence accuracy of the metaheuristic algorithms is also compared in Fig. 14 for different input combinations. It is clear from the figure that the EMPA is faster than the other algorithms and it is followed by the MPA. The PSO algorithm starts slowly and then at the end it provides better convergence compared to the GA.

5. Concluding remarks

The suitability of a new method, extended marine predators algorithm (EMPA) - based ANN (ANN-EMPA), in streamflow prediction has been investigated in this study by comparing its performance with the hybrid ANN-MPA, ANN-GWO, ANN-GA and ANN-PSO methods. Temperature and streamflow data of different lags were utilized as inputs, and optimal lags were decided based on correlation analysis. The derived conclusions of the study are as follows.

- The statistical indices and visualization methods suggested superior performance of ANN-EMPA than the other hybrid ANN methods in streamflow prediction. The accuracy rank of the methods (from the best to worst) is as follows: ANN-EMPA > ANN-MPA > ANN-GWO > ANN-PSO > ANN-GA.
- Temperature-based models showed better accuracy than streamflow-based models. It improved the NSE from 0.916 to 0.958. This is significant in practical applications, especially in regions where streamflow measurement is difficult or available data contains many missing values.
- The best models were obtained by using both temperature and streamflow (Qt-1, Qt-11, Qt-12, Tt, Tt-1, Tt-11, Tt-12) as inputs.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

[1] Z.M. Yaseen, A. El-Shafie, H.A. Afan, et al., RBFNN versus FFNN for daily river flow forecasting at Johor River, Malaysia, Neural Comput. Appl. (2015) http://dx.doi.org/10.1007/s00521-015-1952-6.

[2] S. Londhe, S. Charhate, Comparison of data-driven modelling techniques for river flow forecasting, Hydrol. Sci. J. 55 (2010) 1163–1174, http://dx.doi.org/10.1080/02626667.2010.512867.

[3] M.A. Ghorbani, R. Khatibi, V. Karimi, et al., Learning from multiple models using artificial intelligence to improve model prediction accuracies: Application to river flows, Water Resour. Manag. 32 (2018) 4201–4215, http://dx.doi.org/10.1007/s11269-018-2038-x.

[4] T.C. Pagano, A.W. Wood, M.-H. Ramos, et al., Challenges of operational river forecasting, J. Hydrometeorol. 15 (2014) 1692–1707.

[5] C.M. Zeal, D.H. Burn, S.P. Simonovic, Short term streamflow forecasting using artificial neural networks, J. Hydrol. 214 (1999) 32–48, http://dx.doi.org/10.1016/S0022-1694(98)00242-X.

[6] A. Danandeh Mehr, E. Kahya, E. Olyaie, Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique, J. Hydrol. (2013) http://dx.doi.org/10.1016/j.jhydrol.2013.10.003.

[7] A. Dahamsheh, H. Aksoy, Markov chain-incorporated artificial neural network models for forecasting monthly precipitation in arid regions, Arab. J. Sci. Eng. (2013) http://dx.doi.org/10.1007/s13369-013-0810-z.

[8] T. Yang, A.A. Asanjan, E. Welles, et al., Developing reservoir monthly inflow forecasts using artificial intelligence and climate phenomenon information, Water Resour. Res. 53 (2017) 2786–2812.

[9] F. Kratzert, D. Klotz, C. Brenner, K. Schulz, Rainfall-runoff modelling using long short-term memory (lstm) networks, 2018.

[10] J. Quilty, J. Adamowski, M. Boucher, A stochastic data-driven ensemble forecasting framework for water resources: A case study using ensemble members derived from a database of deterministic wavelet-based models, Water Resour. Res. 55 (2019) 175–202.

[11] Z. Liu, P. Zhou, X. Chen, Y. Guan, A multivariate conditional model for streamflow prediction and spatial precipitation refinement, J. Geophys. Res. Atmos. 120 (2015) 10–116.

[12] Z.M. Yaseen, S.O. Sulaiman, R.C. Deo, K.-W. Chau, An enhanced extreme learning machine model for river flow forecasting: State-of-the-art, practical applications in water resource engineering area and future research direction, J. Hydrol. 569 (2018) 387–408, http://dx.doi.org/10.1016/j.jhydrol.2018.11.069.

[13] O. Kisi, M. Cimen, A wavelet-support vector machine conjunction model for monthly streamflow forecasting, J. Hydrol. 399 (2011) 132–140, http://dx.doi.org/10.1016/j.jhydrol.2010.12.041.

[14] C.N. Babu, B.E. Reddy, A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data, Appl. Soft Comput. 23 (2014) 27–38, http://dx.doi.org/10.1016/j.asoc.2014.05.028.

[15] M. Valipour, Long-term runoff study using SARIMA and ARIMA models in the United States, Meteorol. Appl. (2015) http://dx.doi.org/10.1002/met.1491.

[16] P. Masselot, S. Dabo-Niang, F. Chebana, T.B.M.J. Ouarda, Streamflow forecasting using functional regression, J. Hydrol. (2016) http://dx.doi.org/10.1016/j.jhydrol.2016.04.048.

[17] Q. Shao, H. Wong, M. Li, W.-C. Ip, Streamflow forecasting using functional-coefficient time series model with periodic variation, J. Hydrol. 368 (2009) 88–95.

[18] L. Diop, A. Bodian, K. Djaman, et al., The influence of climatic inputs on streamflow pattern forecasting: Case study of upper Senegal river, Environ. Earth Sci. 77 (182) (2018).

[19] K. Yurekli, A. Kurunc, F. Ozturk, Application of linear stochastic models to monthly flow data of Kelkit stream, Ecol. Model. 183 (2005) 67–75.

[20] M. Khashei, M. Bijari, An artificial neural network (p, d, q) model for timeseries forecasting, Expert Syst. Appl. 37 (2010) 479–489, http://dx.doi.org/10.1016/j.eswa.2009.05.044.

[21] J. Guo, J. Zhou, H. Qin, et al., Monthly streamflow forecasting based on improved support vector machine model, Expert Syst. Appl. 38 (2011) 13073–13081, http://dx.doi.org/10.1016/j.eswa.2011.04.114.

[22] E. Toth, a. Brath, a. Montanari, Comparison of short-term rainfall prediction models for real-time flood forecasting, J. Hydrol. 239 (2000) 132–147, http://dx.doi.org/10.1016/S0022-1694(00)00344-9.

[23] D. Zhang, G. Lindholm, H. Ratnaweera, Use long short-term memory to enhance internet of things for combined sewer overflow monitoring, J. Hydrol. 556 (2018) 409–418.

[24] X. Wen, Q. Feng, R.C. Deo, et al., Two-phase extreme learning machines integrated with complete ensemble empirical mode decomposition with adaptive noise for multi-scale runoff prediction, J. Hydrol. (2019).

[25] Z. Liang, Y. Li, Y. Hu, et al., A data-driven SVR model for long-term runoff prediction and uncertainty analysis based on the Bayesian framework, Theor. Appl. Climatol. 133 (2018) 137–149.

[26] E. Meng, S. Huang, Q. Huang, et al., A robust method for non-stationary streamflow prediction based on improved EMD-SVM model, J. Hydrol. 568 (2019) 462–478, http://dx.doi.org/10.1016/j.jhydrol.2018.11.015.

[27] M. Ravansalar, T. Rajaee, O. Kisi, Wavelet-linear genetic programming: A new approach for modeling monthly streamflow, J. Hydrol. 549 (2017) 461–475, http://dx.doi.org/10.1016/j.jhydrol.2017.04.018.

[28] Z.M. Yaseen, I. Ebtehaj, H. Bonakdari, et al., Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model, J. Hydrol. (2017).

[29] M. Dehghani, A. Seifi, H. Riahi-Madvar, Novel forecasting models for immediate-short-term to long-term influent flow prediction by combining ANFIS and Grey Wolf optimization, J. Hydrol. 576 (2019) 698–725.

[30] I. Ebtehaj, H. Bonakdari, A.H. Zaji, et al., Design of a new hybrid artificial neural network method based on decision trees for calculating the Froude number in rigid rectangular channels, J. Hydrol. Hydromech. 64 (2016) 252–260.

[31] I. Ebtehaj, H. Bonakdari, A.H. Zaji, An expert system with radial basis function neural network based on decision trees for predicting sediment transport in sewers, Water. Sci. Technol. 74 (2016) 176–183.

[32] I. Ebtehaj, H. Bonakdari, A.H. Zaji, A nonlinear simulation method based on a combination of multilayer perceptron and decision trees for predicting non-deposition sediment transport, Water Sci. Technol. Water Supply 16 (2016) 1198–1206.

[33] Z. Xiang, J. Yan, I. Demir, A rainfall-runoff model with LSTM-based sequence-to-sequence learning, Water Resour. Res. 56 (2020) e2019WR025326.

[34] R. Zhang, Z.-Y. Chen, L.-J. Xu, C.-Q. Ou, Meteorological drought forecasting based on a statistical model with machine learning techniques in Shaanxi province, China. Sci. Total Environ. 665 (2019) 338–346.

[35] O. Rahmati, F. Falah, K.S. Dayal, et al., Machine learning approaches for spatial modeling of agricultural droughts in the south-east region of Queensland Australia, Sci. Total Environ. 699 (2020) 134230, http://dx.doi.org/10.1016/j.scitotenv.2019.134230.

[36] A. Danandeh Mehr, A. Rikhtehgar Ghiasi, Z.M. Yaseen, A.U. Sorman, L. Abualigah, A novel intelligent deep learning predictive model for meteorological drought forecasting, J. Ambient Intell. Humaniz. Comput. (2022) 1–15.

[37] A.D. Mehr, B. Vaheddoost, B. Mohammadi, ENN-SA: A novel neuro-annealing model for multi-station drought prediction, Comput. Geosci. 145 (2020) 104622.

[38] K.P. Sudheer, A.K. Gosain, K.S. Ramasastri, A data-driven algorithm for constructing artificial neural network rainfall-runoff models, Hydrol. Process 16 (2002) 1325–1330, http://dx.doi.org/10.1002/hyp.554.

[39] M.C. Demirel, A. Venancio, E. Kahya, Flow forecast by SWAT model and ANN in Pracana basin, Portugal. Adv. Eng. Softw 40 (2009) 467–473, http://dx.doi.org/10.1016/j.advengsoft.2008.08.002.

[40] K. Rasouli, W.W. Hsieh, A.J. Cannon, Daily streamflow forecasting by machine learning methods with weather and climate inputs, J. Hydrol. (2012) 414–415:284–293, http://dx.doi.org/10.1016/j.jhydrol.2011.10.039.

[41] H.I. Erdal, O. Karakurt, Advancing monthly streamflow prediction accuracy of CART models using ensemble learning paradigms, J. Hydrol. 477 (2013) 119–128, http://dx.doi.org/10.1016/j.jhydrol.2012.11.015.

[42] a. Danandeh Mehr, E. Kahya, a. Şahin, M.J. Nazemosadat, Successive-station monthly streamflow prediction using different artificial neural network algorithms, Int. J. Environ. Sci. Technol. (2014) http://dx.doi.org/10.1007/s13762-014-0613-0.

[43] Z.M. Yaseen, M. Fu, C. Wang, et al., Application of the hybrid artificial neural network coupled with rolling mechanism and grey model algorithms for streamflow forecasting over multiple time horizons, Water Resour. Manag. 32 (2018) 1883–1899, http://dx.doi.org/10.1007/s11269-018-1909-5.

[44] O. Kisi, B. Choubin, R.C. Deo, Z.M. Yaseen, Incorporating synoptic-scale climate signals for streamflow modelling over the Mediterranean region using machine learning models, Hydrol. Sci. J. (2019).

[45] A. Rahmani-Rezaeieh, M. Mohammadi, A. Danandeh Mehr, Ensemble gene expression programming: A new approach for evolution of parsimonious streamflow forecasting model, Theor. Appl. Climatol. (2020) http://dx.doi.org/10.1007/s00704-019-02982-x.

[46] M. Cheng, F. Fang, T. Kinouchi, et al., Long lead-time daily and monthly streamflow forecasting using machine learning methods, J. Hydrol. 590 (2020) 125376.

[47] M.B. Wagena, D. Goering, A.S. Collick, et al., Comparison of short-term streamflow forecasting using stochastic time series, neural networks, process-based, and Bayesian models, Environ. Model Softw. 126 (2020) 104669.

[48] Z.M. Yaseen, I. Ebtehaj, H. Bonakdari, et al., Novel approach for streamflow forecasting using a hybrid ANFIS-FFA model, J. Hydrol. (2017).

[49] Z.M. Yaseen, W.H.M.W. Mohtar, A.M.S. Ameen, et al., Implementation of univariate paradigm for streamflow simulation using hybrid data-driven model: Case study in tropical region, IEEE Access 7 (2019) 74471–74481.

[50] Feng Z. kai, Niu W. jing, Tang Z. yang, et al., Monthly runoff time series prediction by variational mode decomposition and support vector machine based on quantum-behaved particle swarm optimization, J. Hydrol. (2020) http://dx.doi.org/10.1016/j.jhydrol.2020.124627.
[51] A. Danandeh Mehr, S. Ghadimi, H. Marttila, A. Torabi Haghighi, A new evolutionary time series model for streamflow forecasting in boreal lake-river systems, Theor. Appl. Climatol. (2022) 1–14.
[52] A.D. Mehr, A.H. Gandomi, MSGP-LASSO: An improved multi-stage genetic programming model for streamflow prediction, Inform. Sci. 561 (2021) 181–195.
[53] L. Ni, D. Wang, J. Wu, et al., Streamflow forecasting using extreme gradient boosting model coupled with Gaussian mixture model, J. Hydrol. 124901 (2020).
[54] W.S. McCulloch, W. Pitts, A logical calculus of the ideas immanent in nervous activity, Bull. Math. Biophys. 5 (4) (1943) 115–133.
[55] F. Rosenblatt, The perceptron: A probabilistic model for information storage and organization in the brain, Psychol. Rev. 65 (6) (1958) 386–408.
[56] T. Kavzoglu, P.M. Mather, The use of backpropagating artificial neural networks in land cover classification, Int. J. Remote Sens. 24 (23) (2003) 4907–4938.
[57] D. Luongvinh, Y. Kwon, Behavioral modeling of power amplifiers using fully recurrent neural networks, In IEEE MTT-S International Microwave Symposium Digest. 17 (2005) 1–4.
[58] Faruk DÖ, A hybrid neural network and ARIMA model for water quality time series prediction, Eng. Appl. Artif. Intell. 23 (4) (2010) 586–594.
[59] V. Moosavi, M. Vafakhah, B. Shirmohammadi, N. Behnia, A wavelet-ANFIS hybrid model for groundwater level forecasting for different prediction periods, Water Resourc. Manag. 27 (5) (2013) 1301–1321.
[60] H. Relvas, A.I. Miranda, An urban air quality modeling system to support decisionmaking: Design and implementation, Air Qual., Atmos. Health 11 (7) (2018) 815–824.
[61] P.L. Mao, R.K. Aggarwal, A novel approach to the classification of the transient phenomena in power transformers using combined wavelet transform and neural network, IEEE Trans. Power Deliv. 16 (4) (2001) 654–660.
[62] O. Kisi, Neural networks and wavelet conjunction model for intermittent streamflow forecasting, J. Hydrol. Eng. 14 (8) (2009) 773–782.
[63] K.S. Parmar, R. Bhardwaj, River water prediction modeling using neural networks, fuzzy and wavelet coupled model, Water Resour. Manag. 29 (1) (2015) 17–33.
[64] O. Kisi, K.S. Parmar, Application of least square support vector machine and multivariate adaptive regression spline models in long term prediction of river water pollution, J. Hydrol. 534 (2016) 104–112.
[65] L.M. Ventura, F. de Oliveira Pinto, L.M. Soares, A.S. Luna, A. Gioda, Forecast of daily PM2.5 concentrations applying artificial neural networks and Holt–Winters models, Air Qual., Atmos. Health 12 (3) (2019) 317–325.
[66] Rana Muhammad Adnan, Zhongmin Liang, Kulwinder Singh Parmar, Kirti Soni, Ozgur Kisi, Modeling monthly streamflow in mountainous basin by MARS, GMDH-NN and DENFIS using hydroclimatic data, Neural Comput. Appl. (2020).
[67] N. Singh, S.B. Singh, Hybrid algorithm of particle swarm optimization and Grey Wolf optimizer for improving convergence performance, J. Appl. Math. (2017) 2030489, http://dx.doi.org/10.1155/2017/2030489, pages 15, 2017.
[68] J.H. Holland, Adaptation in Natural and Artificial Systems, University of Michigan Press, Ann Arbor, MI, 1975.
[69] G.E. Liepins, M.R. Hilliard, Genetic algorithms: Foundations and applications, Ann. Oper. Res. 21 (1989) 31–57, http://dx.doi.org/10.1007/BF02022092.
[70] D.S. Weile, E. Michielssen, Genetic algorithm optimization applied to electromagnetics: A review, IEEE Trans. Antennas Propag. 45 (3) (1997) 343–353, http://dx.doi.org/10.1109/8.558650.
[71] J. Kennedy, R.C. Eberhart, Particle swarm optimization, in: Proc. IEEE International Conference on Neural Networks, Vol. IV, Perth, Australia, IEEE Service Center, Piscataway, NJ, 1995, pp. 1942–1948.
[72] R. Poli, J. Kennedy, T. Blackwell, Particle swarm optimization, Swarm Intell. 1 (2007) 33–57, http://dx.doi.org/10.1007/s11721-007-0002-0.
[73] D. Wang, D. Tan, L. Liu, Particle swarm optimization algorithm: An overview, Soft Comput. 22 (2018) 387–408, http://dx.doi.org/10.1007/s00500-016-2474-6.
[74] S. Mirjalili, S.M. Mirjalili, A. Lewis, Grey Wolf optimizer, Adv. Eng. Softw. 69 (2014) 46–61, http://dx.doi.org/10.1016/j.advengsoft.2013.12.007.
[75] E. Emary, H.M. Zawbaa, A.L. Hassanien, Binary Grey Wolf optimization approaches for feature selection, Neurocomputing 172 (2016) 371–381, http://dx.doi.org/10.1016/j.neucom.2015.06.083.
[76] Q. Al-Tashi, S.J. Abdul Kadir, H.M. Rais, S. Mirjalili, H. Alhussian, Binary optimization using hybrid Grey Wolf optimization for feature selection, IEEE Access 7 (2019) 39496–39508, http://dx.doi.org/10.1109/ACCESS.2019.2906757.
[77] Zj Teng, Jl Lv, Lw Guo, An improved hybrid Grey Wolf optimization algorithm, Soft Comput. 23 (2019) 6617–6631, http://dx.doi.org/10.1007/s00500-018-3310-y.
[78] A. Faramarzi, M. Heidarinejad, S. Mirjalili, A.H. Gandomi, Marine predators algorithm: A nature-inspired metaheuristic, Expert Syst. Appl. 152 (2020).