



Suspended sediment load prediction using artificial neural network and ant lion optimization algorithm

Fatemeh Barzegari Banadkooki¹ · Mohammad Ehteram² · Ali Najah Ahmed³ · Fang Yenn Teo⁴ · Mahboube Ebrahimi¹ · Chow Ming Fai⁵ · Yuk Feng Huang⁶ · Ahmed El-Shafie^{7,8}

Received: 2 January 2020 / Accepted: 23 June 2020 / Published online: 3 July 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Suspended sediment load (SSL) estimation is a required exercise in water resource management. This article proposes the use of hybrid artificial neural network (ANN) models, for the prediction of SSL, based on previous SSL values. Different input scenarios of daily SSL were used to evaluate the capacity of the ANN-ant lion optimization (ALO), ANN-bat algorithm (BA) and ANN-particle swarm optimization (PSO). The Gorganrood basin in Iran was selected for this study. First, the lagged SSL data were used as the inputs to the models. Next, the rainfall and temperature data were used. Optimization algorithms were used to fine-tune the parameters of the ANN model. Three statistical indexes were used to evaluate the accuracy of the models: the root-mean-square error (RMSE), mean absolute error (MAE) and Nash-Sutcliffe efficiency (NSE). An uncertainty analysis of the predicting models was performed to evaluate the capability of the hybrid ANN models. A comparison of models indicated that the ANN-ALO improved the RMSE accuracy of the ANN-BA and ANN-PSO models by 18% and 26%, respectively. Based on the uncertainty analysis, it can be surmised that the ANN-ALO has an acceptable degree of uncertainty in predicting daily SSL. Generally, the results indicate that the ANN-ALO is applicable for a variety of water resource management operations.

Keywords River suspended sediment load · Artificial neural network · Ant lion optimization · Bat algorithm · Particle swarm optimization · Sensitivity analysis

Responsible editor: Marcus Schulz

✉ Ali Najah Ahmed
mahfoodh@uniten.edu.my

Fatemeh Barzegari Banadkooki
barzegari@pnu.ac.ir

Mohammad Ehteram
mohammdehteram@semnan.ac.ir

Fang Yenn Teo
Teo@nottingham.edu.my

Mahboube Ebrahimi
ebrahimi_mahboube@yahoo.com

Chow Ming Fai
Chowmf@uniten.edu.my

Yuk Feng Huang
huangyf@utar.edu.my

Ahmed El-Shafie
elshafie@um.edu.my

¹ Agricultural Department, Payam Noor University, Tehran, Iran

² Department of Water Engineering and Hydraulic Structures, Faculty of Civil Engineering, Semnan University, Semnan, Iran

³ Institute of Energy Infrastructure (IEI), Universiti Tenaga Nasional (UNITEN), 43000 Kajang, Selangor, Malaysia

⁴ Faculty of Science and Engineering, University of Nottingham Malaysia, 43500 Semenyih, Selangor, Malaysia

⁵ Institute of Sustainable Energy (ISE), Universiti Tenaga Nasional (UNITEN), 43000 Kajang, Selangor, Malaysia

⁶ Department of Civil Engineering, Lee Kong Chian Faculty of Engineering and Science, Universiti Tunku Abdul Rahman, 43200, Kajang, Selangor, Malaysia

⁷ Department of Civil Engineering, Faculty of Engineering, University of Malaya (UM), 50603, Kuala Lumpur, Malaysia

⁸ National Water Center, United Arab Emirates University, 15551, Al Ain, United Arab Emirates

Introduction

Sediment transport, which is among the most important aspects of river flow, is significantly influenced by the flow condition and the size of the particles (Ehteram et al. 2019c). The characteristics of sedimentation in fluvial systems can be better understood, through a process known as suspended sediment load (SSL) estimation. Sedimentation stems from erosion and weathering (Misset et al. 2019). Suspended sediment load and sediment transportation are affected by geological and geomorphological issues (Chang et al. 2019). The SSL prediction represents a major issue in the fields of hydraulic engineering and river basin development. The influence of different meteorological parameters (rainfall, temperature and discharge), on sediment transport, renders the modelling of SSL a complicated undertaking (Samantaray and Ghose 2018). In order to overcome the issues of complexity and nonlinearity that come with SSL prediction, data-driven models are often brought into the picture. The selection of an appropriate model for SSL prediction has proven to be an exceedingly difficult task (Talebi et al. 2016). According to the literature, empirical models fall short when it comes to comparison with accurate SSL predictions (Khan et al. 2018; Khosravi et al. 2018; Samantaray and Ghose 2018; Chang et al. 2019; Misset et al. 2019). As such, the prediction of the SSL based on local data is deemed the way forward for an accurate SSL prediction (Jothiprakash and Garg 2009). Jothiprakash and Garg (2009) used artificial neural network (ANN) to predict the volume of sediment in a reservoir. They opined that the accuracy level of ANN is above that of a regression model. The major benefit that comes with ANN usage is its capacity to model mathematical relationships, between input and output data, through a learning process, without prior knowledge of the problems (Melesse et al. 2011; Nourani and Andalib 2015; Ehteram et al. 2018; Ehteram et al. 2019c). To prepare an adequate architecture of ANN, several trial and error procedures need to be carried out. Also required is extensive input data for the training process (Ehteram et al. 2019c). Melesse et al. (2011) used ANN, multiple linear regression (MLR) and multiple nonlinear regression (MNL) to estimate SSL. According to the results attained, the ANN estimations for most simulations were better than those realized through the MLR and MNL. Kisi et al. (2012) employed genetic programming (GP), ANN and adaptive neuro-fuzzy inference system (ANFIS) to estimate SSL. They concluded that GP is superior to the ANFIS and ANN models. Kakaei Lafdani et al. (2013) investigated the estimation of daily SSL with a support vector machine (SVM) and ANN. Streamflows and rainfall were employed as the input data. They reported that the ANN model is significantly more accurate than other SSL estimation models. Vafakhah (2012) used the ANFIS and ANN models for the modelling of SSL. They forwarded that the ANFIS

model provides a better estimation than the ANN model. Tayfur et al. (2013) investigated the potential of ANN-genetic algorithm (GA) for predicting SSL. They found the performance of ANN-GA to be superior to that of ANN. Afan et al. (2014) used radial basis function neural network (RBFNN), and multilayer perceptron (MLP) neural network, for SSL modelling. They claimed that the MLP model delivers better accuracy than the RBFNN model. Nourani and Andalib (2015) examined the capabilities of least square support vector machine (LSSVM) and ANN models in estimating SSL. The results indicate that the LLSVM model provides better accuracy than the ANN model. Talebi et al. (2016) utilized regression trees and model trees for estimating SSL. They reported that the performance of regression trees and model trees is superior to that of the ANN model. Emamgholizadeh and Demneh (2018) examined the proficiency of GP, ANN and ANFIS for SSL modelling. The results give credence to the fact that the accuracy level of the GEP model is above that of the ANN and ANFIS models.

While data-driven models are highly regarded for predicting hydrological variables, these models require tuning through training algorithms (Banadkooki et al. 2020). In order to complete the final network training, it is essential that the entire data-driven model be fine-tuned. Of late, optimization algorithms such as bat algorithm (BA), GA, particle swarm optimization (PSO), shark algorithm (SA) and firefly algorithm (FFA) have been used to train data-driven models, in order to determine their optimal parameter values (Allawi et al. 2018; Ehteram et al. 2018a, b, 2019a, b; Farzin et al. 2018; Abobakr Yahya et al. 2019; Yousif et al. 2019; Valikhan-Anaraki et al. 2019; Najah Ahmed et al. 2019; Afan et al. 2020). According to the relevant literature, optimization algorithms increase the convergence speed of traditional training algorithms, such as the gradient descent method and the backpropagation algorithm (Allawi et al. 2018; Ehteram et al. 2018b; Farzin et al. 2018). To the best of our knowledge, no previous attempt has been made to employ an ANN model-based ant lion optimization (ALO) for predicting SSL, thus the aim of this study.

ALO was initially introduced by Mirjalili (2015). Scholars frequently used the ALO for a variety of optimization issues. These optimization issues include those related to energy consumption forecasting (Wang et al. 2018), parameter optimization of SVM (Tian et al. 2018), optimal power flow (Ali et al. 2017) and hydraulic turbines (Tharwat and Hassanien 2017). The forte of the ANN model lies in its capacity to optimize ANN parameters such as weight connections and bias values. In this paper, a novel ALO algorithm was used to train the ANN model. An innovative hybrid ANN model was introduced for the forecasting of daily SSL. The two main goals of this undertaking are (a) to evaluate the capacity of hybrid ANN models for SSL estimation and (b) to evaluate the uncertainty of the ALO, BA and PSO algorithms for the training

of ANN models. The following sections provide information on the applied models, the study location and the data involved. We close with an output analysis and an explanation on the conclusions derived.

Materials and methods

Artificial neural network

The inspiration of ANN model came from the desire to provide intelligent calculations that mimics the routine structure of brain. The ANN models are composed of computational units, namely neurons (Allawi et al. 2018; Ehteram et al. 2018b; Farzin et al. 2018). Each neuron is linked to other neurons by means of weight connections. The ANN model is mostly made up of three layers: an input layer, the hidden layers and an output layer. An illustration of the most frequently employed ANN model is presented in Fig. 1.

For this study, only one hidden layer network was used, as it was deemed sufficiently dynamic for the prediction of any hydrological variable. Many investigations have been conducted to determine the optimal number of hidden neurons in the ANN model. These investigations led to the introduction of several empirical equations. The following equation (Karkalos et al. 2019) was employed for this study:

$$n = \sqrt{N_{inp} + N_{out}} + a \tag{1}$$

where N_{inp} is the number of input neurons, N_{out} is the number of output neurons, a is the number between zero and 10 and n is the number of hidden neurons.

A variety of transfer functions exist which can be used to construct an artificial neural network neuron. Transfer functions are needed to describe the input-output relationship at each neuron layer. Nonlinear transfer functions are usually employed to introduce nonlinearity into the network. In order to minimize the network error, it is essential that the

components of weights and biases be fine-tuned. The ANN models use training algorithms to find the optimal value of weight connection and biases (Ehteram et al. 2018; Tikhamarine et al. 2020). The most commonly used algorithm for ANN training is the backpropagation algorithm (BPA). The BPA is a traditional method for tuning the weight and bias values. Currently, optimization algorithms are frequently applied for the training of ANN hyperparameters. It is noteworthy that the BPA may experience slow convergence and may also be trapped in local minima. The BPA executes an iterative gradient descent procedure to decrease the network error.

Ant lion optimization

The ALO is capable of mimicking the hunting behaviour of ant lions in nature. An ant lion (ANL) moves along a circular path to dig a hole in sand (Mirjalili 2015). The ant lion larvae hide at the bottom of the pit (trap) and wait for ants (prey). With the arrival of the prey (ant), the ant lions pull it into the pit by throwing sand towards the edge of the hole. The performance of the algorithm is based on the following assumptions (Wang et al. 2018):

- 1- Ants (the prey) move randomly in the search space, and the traps of the ANLs affect the movement of the ants.
- 2- The best fitness ANL makes a large hole.
- 3- The ability to hunt an ant, by an elite ANL, is proportional to the fitness of that ANL.
- 4- Each ant (AN) may be hunted by each ANL in each iteration (Fig. 2).

Random walk of ants

The ALO has two groups of search agents: ant (prey) and ant lion. The best search agents (ant lions) never change their

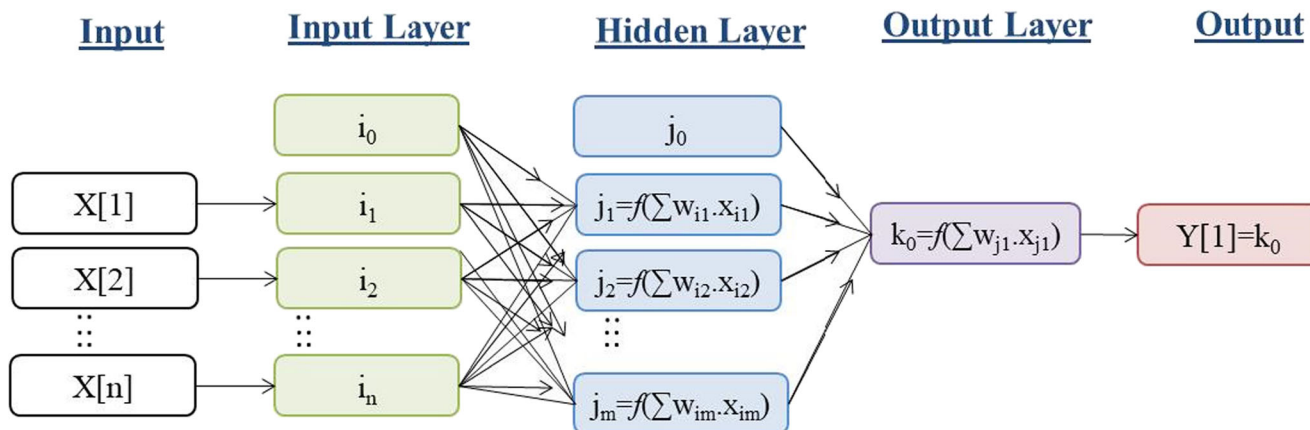
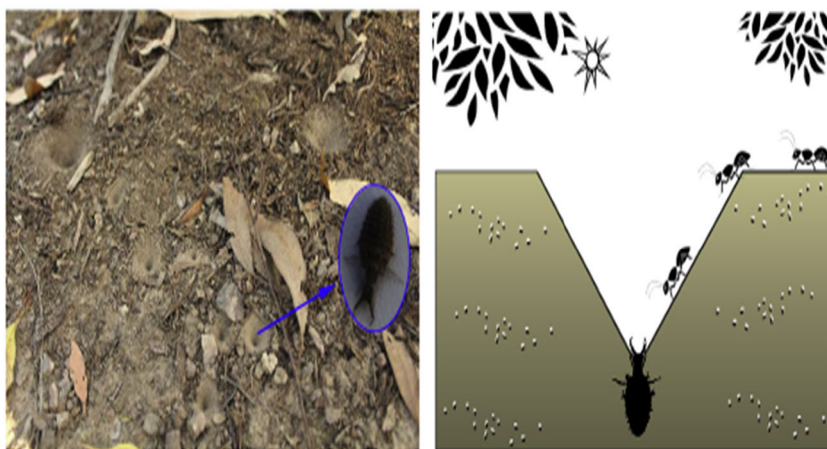


Fig. 1 Schematic structure of ANN (three-layer MLP)

Fig. 2 Hunting behaviour of ANLs



locations, except during the substitution of a special ant. Ant agents can perform random walks in the search space (Mirjalili 2015). The ants trapped in a hole may be hunted by the ant lions. Through random walks, ants update their position around the search space (Mirjalili 2015).

$$X_t = [0, \text{cumsum}(2r_t(t_1)-1), \text{cumsum}(2r_t(t_2)-1), \dots, \text{cumsum}(2r_t(t_n)-1)] \quad (2)$$

where t is the step of random walk, cumsum is the cumulative sum, n is the maximum number of iterations, n is the population size and r is the stochastic function which is computed through the following equation (Mirjalili 2015):

$$r(t) = \begin{cases} 1 \leftarrow \text{if}(\text{rand} > 0.5) \\ 0 \leftarrow \text{if}(\text{rand} \leq 0.5) \end{cases} \quad (3)$$

Equation (3) ensures that the ants move within the boundaries of the search space

$$R_i^j = \frac{(X_i - a_i^t) \times (d_i^t - c_i^t)}{(b_i^t - a_i)} + c_i^t \quad (4)$$

where a_i^t is the minimum step of random walk (RW), X_i^t is the position of i th ant at t th iteration, b_i is the maximum step of random walk, R_i^j is the position of ant i after performing the random walk to nearby ant lion j , c_i^t is the minimum i th variable at t th iteration and d_i^t is the maximum i th variable at t th iteration (Fig. 2).

Trapping hole of the ANL

As mentioned above, the traps of the ANLs affect the random walks of the ANs. The following equation is used to mathematically model this hypothesis (Wang et al. 2018):

$$d_i^t = \text{Ant lion}_j^t + d^t \quad (5)$$

$$c_i^t = \text{Ant lion}_j^t + c^t \quad (6)$$

where c^t is the minimum of all variables at t th iteration, c_i^t is

the minimum of all variables for i th ant, d^t is the maximum of all variables at t th iteration and d_i^t is the maximum of all variables for i th ant. Figure 3 shows the random walk of an ant.

Building the trap

A roulette wheel mechanism is used to simulate the ANL’s hunting ability. As shown in Fig. 3, the ANs are hypothesized to be trapped in only one chosen ANL pit. The ALO algorithm uses the roulette wheel mechanism to select ANLs based on their fitness during optimization. The fitter ANLs are more likely to be successful in hunting an ant (Mirjalili 2015).

Sliding ants towards the ant lion

When the ANL realizes that there is an AN in the trap, it throws sand towards the edge of the hole to pull the prey (ant) in. This behaviour is simulated through the following equation:

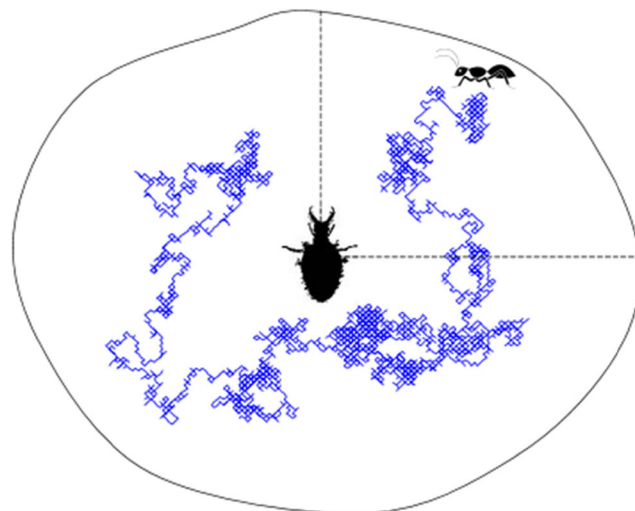


Fig. 3 Random walk of an ant

$$\begin{aligned} c^t &= \frac{c^t}{I} \\ d^t &= \frac{d^t}{I} \end{aligned} \quad (7)$$

where c^t is the minimum of all variables at t th iteration, d^t is the maximum of all variables at t th iteration and I is the ratio. The computation of the I parameter is portrayed in Fig. 4.

Catching prey and rebuilding the trap

In order to increase its chances of catching new AN, the ANL has to alter its position to the position of the hunted ant. This behaviour is simulated in Eq. (10)

$$\text{Ant lion}_j^t = \begin{cases} \text{Ant}_i^t \leftarrow \text{if } (f(\text{Ant}_i^t)) < (f(\text{Ant lion}_j^t)) \\ \text{Ant lion}_j^t \leftarrow \text{otherwise} \end{cases} \quad (8)$$

where Ant_i^t is the location of j th ANL, Ant_i^t is the position of i th ANL and t is the current iteration.

Elitism

The ALO has two groups of search agents: ant (prey) and ant lions. The best search agents (ant lions) only alter their locations during the substitution of a special ant. Ant agents can perform random walks in the search space (Mirjalili 2015). The ant lions hunt ants trapped in a hole. The ant lion with a superior fitness level in each iteration is regarded an elite ant lion. The following equation is used to attain the position of the ants (Mirjalili 2015):

$$\text{Ant}_i^t = \frac{R_A^t + R_e^t}{2} \quad (9)$$

where R_A^t is the random walk around the ant lion (ANL) chosen by way of the roulette mechanism (RM), R_e^t is the random walk around the elite ant lion at t th iteration and Ant_i^t is the location of i th ant lion at t th iteration (Fig. 4).

Bat algorithm

BA is an optimization algorithm for global optimization. This algorithm mimics the behaviour of bats. Each bat flies randomly with a velocity (v_i) at location (solution) x_i with a changing frequency or loudness (A_i). When the bat seeks and finds its prey, its frequency, pulsation rate (r) and loudness are altered. The position, velocity and frequency of bats are updated as follows (Cui et al. 2019):

$$f_l = f_{\min} + (f_{\max} - f_{\min}) \times \beta \quad (10)$$

$$y_l(t) = y_l(t-1) + v_l(t), t = 1, \dots, T \quad (11)$$

$$v_l(t) = [y_l(t-1) - Y_*] \times f_l, t = 1, \dots, T \quad (12)$$

where f_{\min} is the minimum frequency, f_{\max} is the maximum frequency, β is the random value, $y_l(t)$ is the position at time step (t), Y_* is the best solution, f_l is the frequency and T is the total period of assessment.

Subsequent to the selection of a solution from among the best current solutions, a new solution for each bat is derived locally through Eq. (13) (random walk)

$$y(t) = y(t-1) + \varepsilon A(t), t = 1, \dots, T \quad (13)$$

where $A(t)$ is the average loudness and ε is the random number. Figure 5 shows the flowchart of the BA. The pulsation rate increases when a bat finds its prey, while the loudness decreases. The loudness and pulsation rate values are updated as follows:

$$r_l^{t+1} = r_l^0 [1 - \exp(-\gamma t)] A_l^{t+1} = \alpha A_l^t \quad (14)$$

where γ and α are the constant values.

Particle swarm optimization

The PSO is applied in a wide variety of areas. This includes rock strength estimation, energy demand forecasting, coefficient discharge estimation, water resource management and classifier selection (Wang et al. 2018). PSO's group of particles seeks out the best global (g_{best}) and personal (p_{best}) positions. The updating of the position, and velocity of particles in this algorithm, is based on the following equation:

$$V_{\text{new}} = \omega \times V + C_1 \times r_1 (p_{\text{best}} - X) + C_2 \times r_2 (g_{\text{best}} - X) \quad (15)$$

$$X_{\text{new}} = X + V_{\text{new}} \quad (16)$$

where X is the current position of the particles, C_1 and C_2 are the acceleration coefficients, r_1 and r_2 are the random values and ω is the inertia weight.

Hybrid ANN and optimization models

As mentioned previously, the output of ANN is affected by the bias and weight values. The aim of any optimizer is to determine the optimal values of the ANN parameters. To achieve this goal in the hybrid ANN and optimization models, optimization algorithms are employed to optimize the weight and bias values. These weight and bias values are then used as inputs to the optimization algorithms. Put plainly, each agent in each optimization algorithm is a representative of a bias or weight value. The optimization-simulation process usually begins with the initialization of a collection of random agents. The hybrid ANN optimization models are trained with the initial position of the agents (i.e. with their original weights and bias parameters). At the next level, the fitness of the agents is computed. An objective function (i.e. the root-mean-square error) is defined to decrease the hybrid model

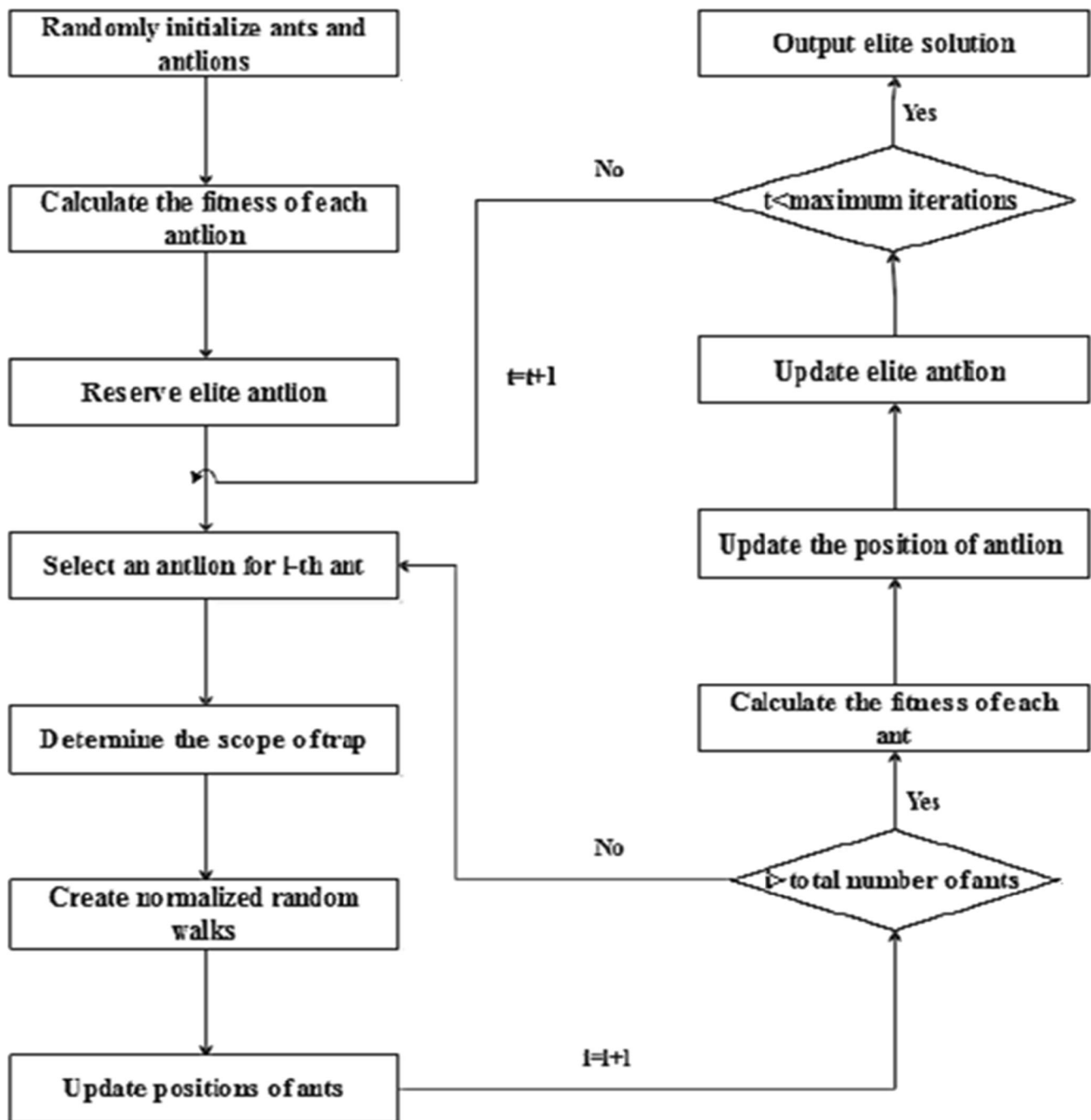


Fig. 4 Flowchart of ant lion optimization

error. The calculated error is decreased through alterations in the position of the agents. This process continues until a stopping criterion is met.

Case study

An agricultural basin located in the north of Iran was selected for this case study. The Gorganrood basin is positioned at longitude 10' 54° to 26' 56° N and latitude 35' 36°

to 38' 15° E, as shown in Fig. 6. Except for the period between May and July, a wet climate with moderate temperatures prevails in this area. The average annual rainfall is 515 mm, while the average annual temperature is 17.8 °C. The lowest and highest rainfall levels occur in August and March, respectively. August is the hottest month and January the coldest.

The construction of accurate SSL prediction models involves the generation of input data for all the soft computing models, by means of statistically significant

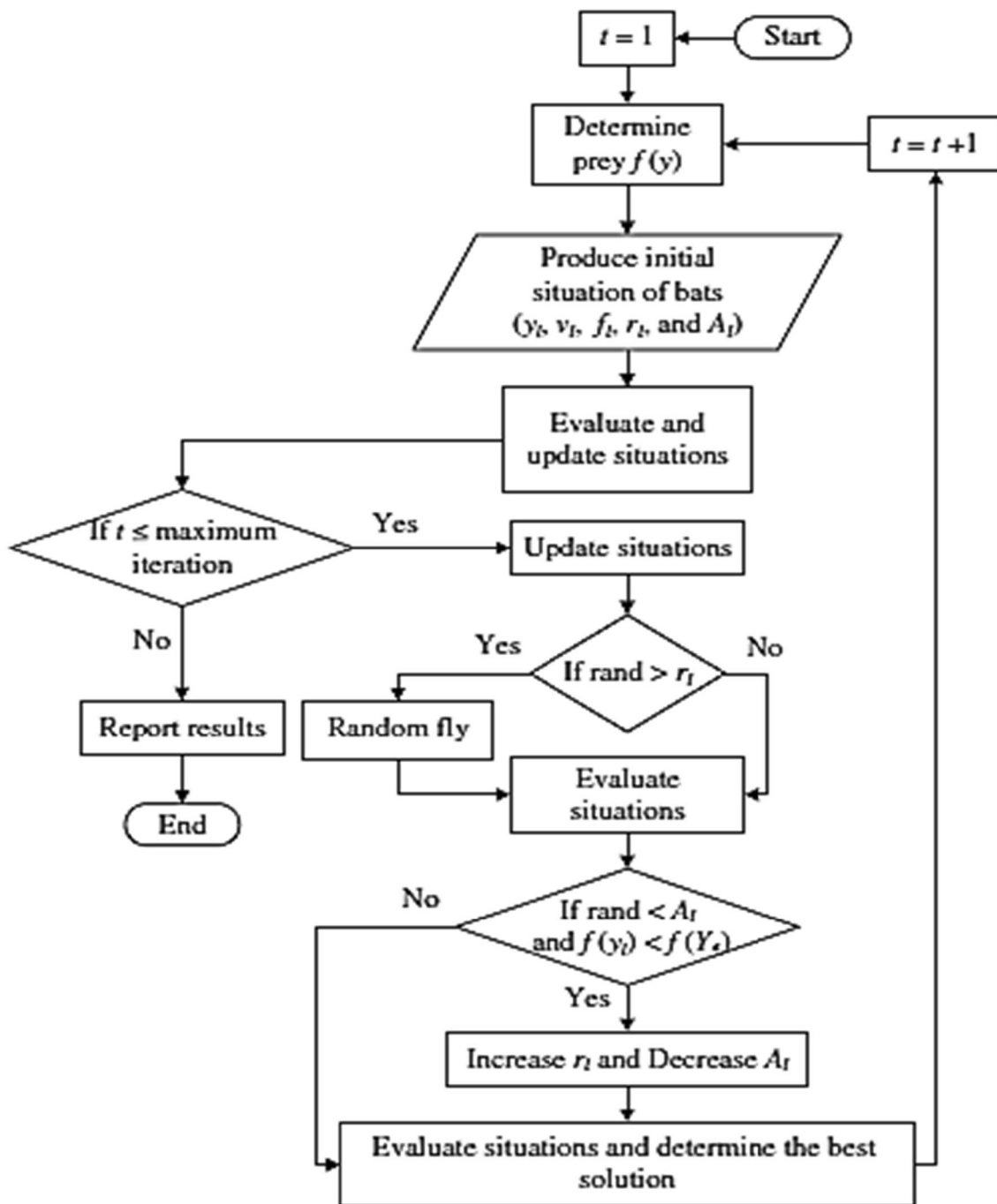


Fig. 5 Flowchart of bat algorithm

lagged combinations of historical SSL. These inputs were used for a 1-day-ahead forecast of SSL. In this study, first, the soft computing models were evaluated using the data at Narmab station in the Goorganrood basin. Three hundred eighty daily data were used from the Narmab station. Then, climate data of the Salian station, Narmab station, Tamar station and Basir Abad station were used to validate the models. Three thousand six hundred daily data of these stations were used as the inputs to the models; 70% and 30% of data were used

for the training and testing levels, respectively, as shown in Figs. 7 and 8. Different data splitting ratios have been examined to achieve the best data splitting ratio. However, the best results were obtained for 70% of data and 30% of data for the training and testing levels, respectively. Figure 9 shows the best objective function value obtained for different sizes of data through the ANN model. The results indicated that the best results are obtained by the 70% and 30% of data for the training and testing levels, respectively.

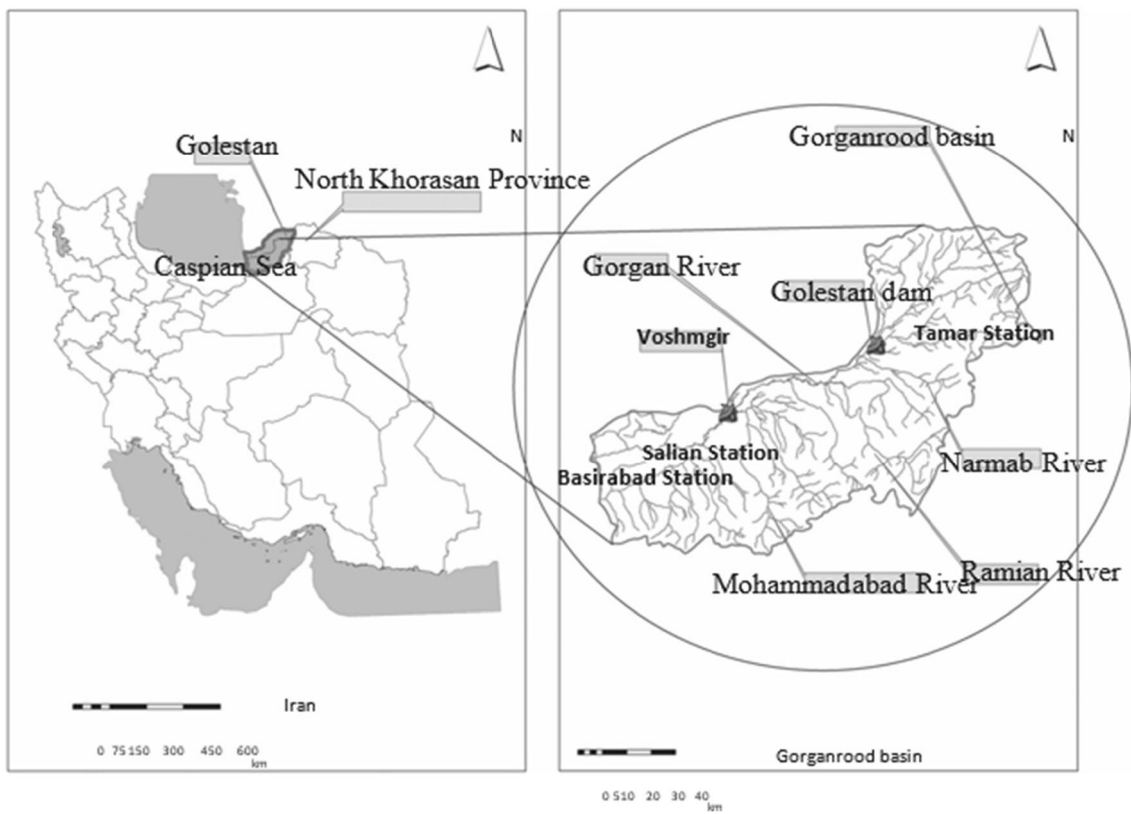


Fig. 6 Location of the case study

Table 1 shows the statistical characteristics of data. The mean, maximum and minimum values at the Narmab station is 41,000, 65,000 and 2800, respectively. In this study, the SSL was estimated with and without meteorological data. First, Eq. (18) was

used to estimate SSL without meteorological data. Then, the SSL was estimated using meteorological data.

Statistical analysis generated a unique set of inputs, with a total of five different input combinations derived by way of

Fig. 7 Time series of SSL for the training stage

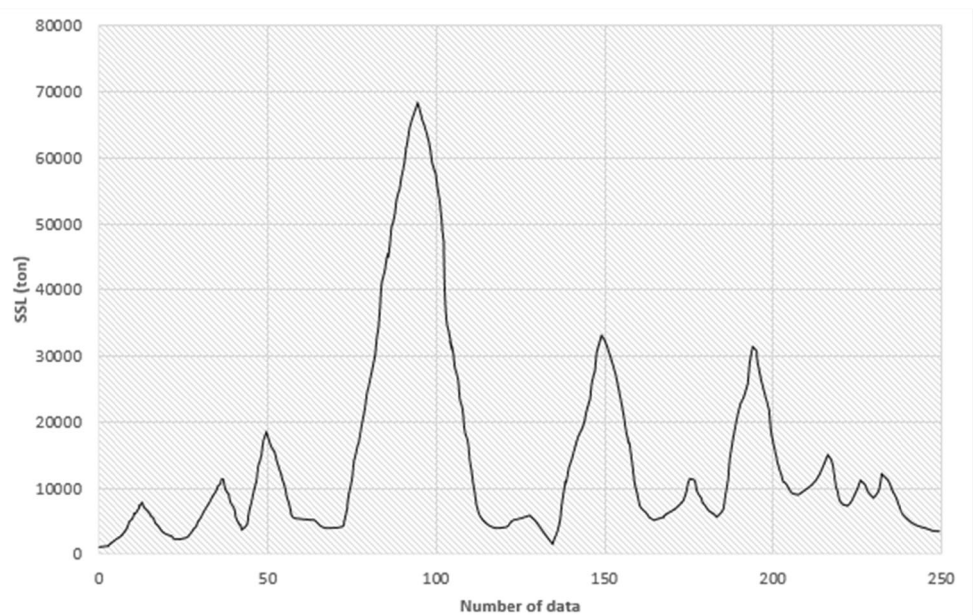
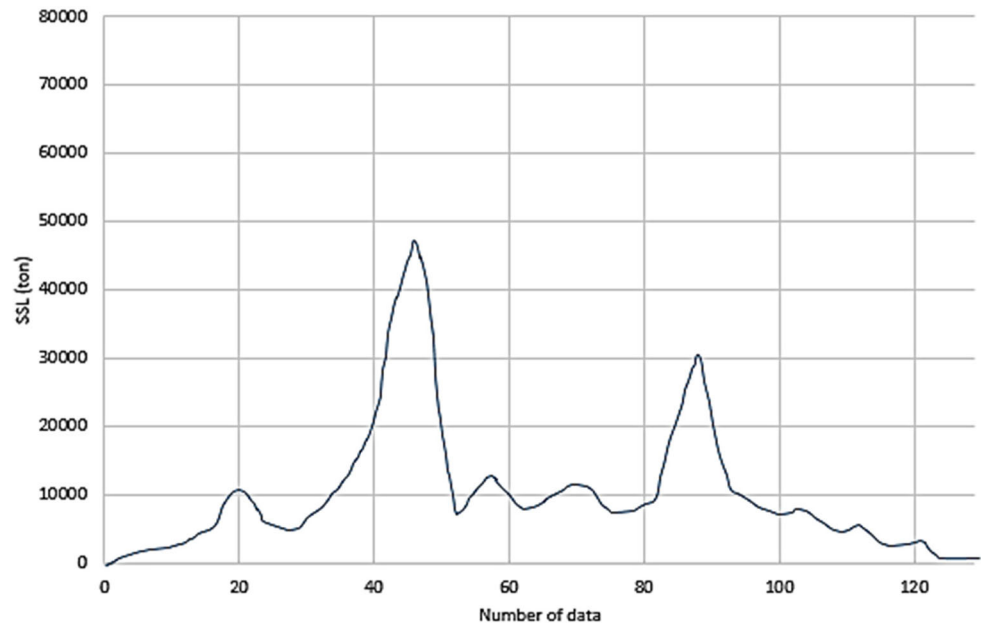


Fig. 8 Time series of SSL for the testing stage



the periodicity (day) and historical SSL. In this study, first, the soft computing models are evaluated using

$$\begin{aligned}
 \text{SSL}(t) &= f[\text{SSL}(t-1)], (1 : \text{Number of scenarios}) \\
 \text{SSL}(t) &= f[\text{SSL}(t-1), \text{SSL}(t-2)], (2 : \text{Number of scenarios}) \\
 \text{SSL}(t) &= f[\text{SSL}(t-1), \text{SSL}(t-2), \text{SSL}(t-3)], (3 : \text{Number of scenarios}) \\
 \text{SSL}(t) &= f[\text{SSL}(t-1), \text{SSL}(t-2), \text{SSL}(t-3), \text{SSL}(t-4)], (4 : \text{Number of scenarios}) \\
 \text{SSL}(t) &= f[\text{SSL}(t-1), \text{SSL}(t-2), \text{SSL}(t-3), \text{SSL}(t-4), \text{SSL}(t-5)], (5 : \text{Number of scenarios})
 \end{aligned} \tag{17}$$

where $t-1$ is the pervious-day SSL, $t-2$ is the previous 2-day SSL, $t-3$ is the previous 3-day SLL, $t-4$ is the previous 4-day SLL and $t-5$ is the previous 5-day SSL.

The principal component analysis

In this study, the principal component analysis (PCA) is used to reduce the dimensionality of input data. The PCA is used to select the appropriate climate data for predicting SSL at the Salian station, Narmab station, Tamar station and Basir Abad station. The PCA converts the given input data into a smaller number of independent variables (Noori et al. 2011). The new set of variables is named principal components (PCs). A set of principals has a number of features, including the following: (1) all PCs are uncorrelated, (2) each PC is a linear integration of input variables and (3) the first PC explains most of the variability in the data (Noori et al. 2011).

$$Z = a_{i1}X_1 + a_{i2}x_2 + \dots + a_{ip}X_p \tag{18}$$

where Z_i represents PCs, X is the input variable and a is the eigenvector. The dimension of new future space is

determined by the eigenvalues. The variance of data along the new feature space is described by the eigenvalues (Noori et al. 2011).

$$|R - \lambda I| = 0 \tag{19}$$

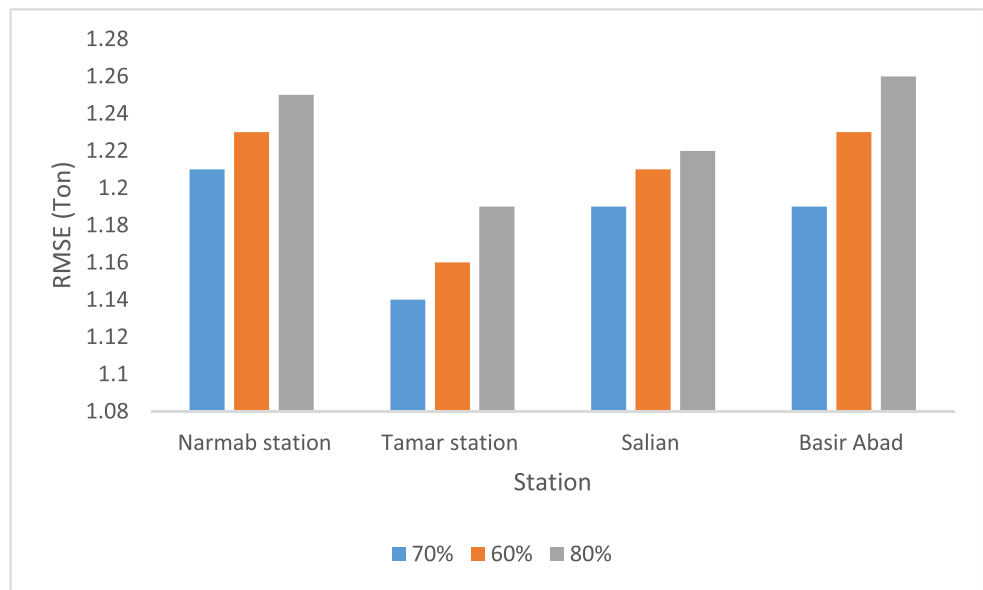
where I is the unit matrix, R is the variance-covariance matrix and λ is the eigenvalues.

The sample size of data should be investigated before modellers use the PCA. Kaiser-Meyer-Olkin (KMO) coefficient is computed to investigate the adequacy of sample sizes. The minimum value for KMO should be 0.5. In this study, the value of KMO was 0.85, 0.84, 0.79 and 0.65 for the Salian station, Tamar station, Basir Abad station and Narmab station, respectively (Tayfur et al. 2013).

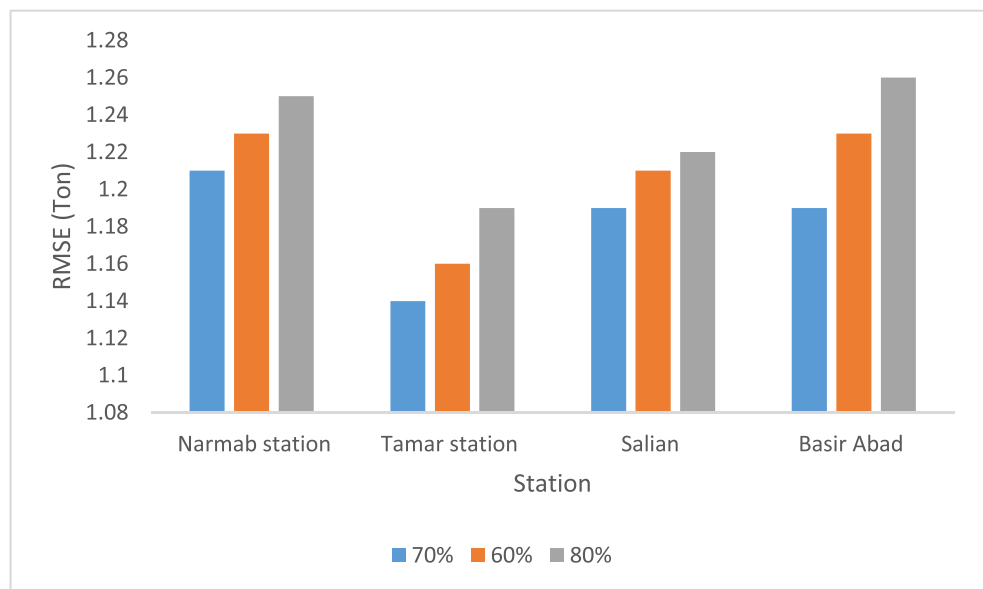
$$\text{KMO} = \frac{\sum(\text{correlation})^2}{\sum(\text{correlation})^2 + \sum(\text{partial correlation})^2} \tag{20}$$

The lagged meteorological data of temperature, discharge and rainfall were used for a 1-day-ahead forecast of SSL.

Fig. 9 a, b Sensitivity analysis for the data splitting ratios



(a)



(b)

1. $Q(t - 1)$, $Q(t - 2)$ and $Q(t - 3)$ are the previous-day discharge, two-previous-day discharge and three-previous-day discharge, respectively.
2. $R(t - 1)$, $R(t - 2)$ and $R(t - 3)$ are the previous-day rainfall, two-previous-day rainfall and three-previous-day rainfall, respectively.
3. $T(t - 1)$, $T(t - 2)$ and $T(t - 3)$ are the previous-day temperature, two-previous-day temperature and three-previous-day temperature, respectively.

Tables 2, 3, 4 and 5 show the descriptive statistics of the PCs and eigenvectors obtained through the PCA application.

Narmab station

The cumulative variance proportion (CVP) shows the contribution rate of PCs. The CVP of 0.90 or more than 0.90 is significant (Zhao and Huang 2017). As observed in Table 2,

Table 1 Statistical indices of the collected data for SSL and meteorological inputs

Parameter	Station			
	Narmab River	Tamar	Salian station	Basir Abad
SSL (tons)				
Min	28,000	25,000	24,550	25,200
Max	65,000	63,000	64,500	67,000
Mean	41,000	43,000	40,500	42,000
Standard deviation	32.23	31.12	30.98	34.23
Rainfall (mm)				
Min	12	10	23	18
Max	45	67	45	66
Mean	18	34	32	34
Temperature (°C)				
Min	8	12	10	14
Max	38	38	33	39
Mean	23	22	25	27
Discharge (m ³ /s)				
Min	10	14	12	12
Mean	14	18	15	18
Max	25	29	21	28

the analysis of first four principal components summed up a contribution rate of 93%. The eigenvectors evaluate the coefficients for the formation of PCs. The coefficients of 0.75 or more than 0.75 have the most effects on the PCs (Zhao and Huang 2017). The results indicated that $Q(t-1)$, $R(t-1)$ and $T(t-1)$ had the most effect on the first four PCs in the Narmab station.

Basir Abad station

As observed in Table 3, the first four principal components summed up a contribution rate of 91%. The results indicated that the $Q(t-1)$, $R(t-1)$ and $T(t-1)$ had the most effect on the first four PCs. Thus, the first four principal components were used as the inputs to the models.

Table 2 Outputs of the PCA model for the selection of the appropriate metrological inputs for the Narmab River station

Inputs	PC (1)	PC (2)	PC (3)	PC (4)	PC (5)	PC (6)	PC (7)	PC (8)	PC (9)	PC (10)	PC (11)	PC (12)
Eigenvectors obtained through the application of PCA												
$Q(t-1)$	0.91	0.86	0.85	0.83	0.78	0.54	0.45	0.22	0.20	0.19	0.17	0.15
$R(t-1)$	0.89	0.80	0.79	0.78	0.76	0.51	0.26	0.24	0.20	0.19	0.17	0.14
$T(t-1)$	0.78	0.77	0.76	0.75	0.74	0.50	0.30	0.29	0.18	0.18	0.16	0.12
$Q(t-2)$	0.70	0.69	0.63	0.55	0.52	0.49	0.29	0.28	0.17	0.15	0.12	0.11
$R(t-2)$	0.65	0.65	0.60	0.44	0.50	0.48	0.18	0.17	0.16	0.14	0.12	0.10
$T(t-2)$	0.54	0.53	0.45	0.43	0.42	0.40	0.21	0.15	0.12	0.10	0.09	0.09
$Q(t-3)$	0.43	0.40	0.39	0.38	0.37	0.29	0.16	0.14	0.09	0.08	0.07	0.07
$R(t-3)$	0.40	0.39	0.38	0.35	0.34	0.28	0.14	0.12	0.07	0.05	0.05	0.05
$T(t-3)$	0.39	0.35	0.35	0.32	0.32	0.26	0.12	0.11	0.05	0.02	0.02	0.02
$Q(t-3)$	0.36	0.34	0.32	0.30	0.31	0.25	0.10	0.10	0.01	0.01	0.01	0.01
$R(t-3)$	0.32	0.30	0.30	0.29	0.30	0.23	0.09	0.09	0.01	0.01	0.01	0.01
$T(t-3)$	0.30	0.28	0.29	0.27	0.25	0.22	0.08	0.07	0.01	0.01	0.01	0.01
Variance proportion (%)	42	24	14	10	5	3	0.5	0.5	0.25	0.25	0.25	0.25

Table 3 Outputs of the PCA model for the selection of the appropriate metrological inputs for the Basir Abad station

Inputs	PC (1)	PC (2)	PC (3)	PC (4)	PC (5)	PC (6)	PC (7)	PC (8)	PC (9)	PC (10)	PC (11)	PC (12)
Eigenvectors obtained through the application of PCA												
$Q(t - 1)$	0.90	0.84	0.85	0.84	0.69	0.67	0.32	0.30	0.29	0.21	0.19	0.18
$R(t - 1)$	0.86	0.83	0.79	0.77	0.55	0.52	0.30	0.28	0.26	0.19	0.17	0.16
$T(t - 1)$	0.75	0.76	0.76	0.75	0.54	0.51	0.28	0.27	0.20	0.17	0.16	0.14
$Q(t - 2)$	0.69	0.65	0.63	0.62	0.53	0.50	0.22	0.21	0.19	0.16	0.15	0.13
$R(t - 2)$	0.64	0.62	0.60	0.58	0.50	0.49	0.20	0.18	0.17	0.14	0.12	0.10
$T(t - 2)$	0.53	0.55	0.54	0.53	0.49	0.44	0.19	0.18	0.16	0.10	0.09	0.09
$Q(t - 3)$	0.42	0.41	0.40	0.39	0.38	0.36	0.17	0.16	0.15	0.08	0.07	0.07
$R(t - 3)$	0.39	0.38	0.37	0.36	0.37	0.35	0.16	0.15	0.14	0.05	0.05	0.05
$T(t - 3)$	0.38	0.37	0.36	0.34	0.35	0.34	0.15	0.14	0.13	0.02	0.02	0.02
$Q(t - 3)$	0.35	0.34	0.35	0.33	0.34	0.32	0.14	0.12	0.11	0.01	0.01	0.01
$R(t - 3)$	0.30	0.29	0.34	0.32	0.32	0.31	0.12	0.11	0.10	0.01	0.01	0.01
$T(t - 3)$	0.29	0.28	0.33	0.30	0.30	0.29	0.10	0.09	0.07	0.01	0.01	0.01
Variance proportion (%)	40 ≥90	25	14	12	3	3	1	1	0.25	0.25	0.25	0.25

Salian station

As observed in Table 4, the first two PCs summed up a contribution rate of 90%. Thus, they were used as the inputs to the models. The results indicated that the $Q(t - 1)$, $R(t - 1)$ and $T(t - 1)$ had the most effect on the first two PCs.

Tamar station

As observed in Table 5, the first two PCs summed up a contribution rate of 90%. Thus, they were used as inputs to the

models. The results indicated that the $Q(t - 1)$, $R(t - 1)$ and $T(t - 1)$ had the most effect on the first two PCs.

Simultaneous parameter and input uncertainty estimation

In this study, an integrated framework was used to quantify the parameter and input uncertainty of the used soft computing models. The parameters of ANN models (bias values and weight connections) can be considered as the source of uncertainty. The input data (temperature, rainfall and discharge) can be considered as another source of uncertainty. Thus, it is

Table 4 Outputs of the PCA model for the selection of the appropriate metrological inputs for the Salian station

Inputs	PC (1)	PC (2)	PC (3)	PC (4)	PC (5)	PC (6)	PC (7)	PC (8)	PC (9)	PC (10)	PC (11)	PC (12)
Eigenvectors obtained through the application of PCA												
$Q(t - 1)$	0.89	0.87	0.84	0.80	0.79	0.45	0.26	0.25	0.24	0.23	0.20	0.15
$R(t - 1)$	0.88	0.82	0.80	0.76	0.74	0.42	0.26	0.25	0.23	0.22	0.19	0.14
$T(t - 1)$	0.76	0.75	0.70	0.65	0.63	0.31	0.22	0.20	0.20	0.19	0.17	0.12
$Q(t - 2)$	0.55	0.54	0.52	0.50	0.48	0.29	0.21	0.20	0.20	0.19	0.14	0.11
$R(t - 2)$	0.50	0.49	0.47	0.46	0.45	0.26	0.20	0.19	0.19	0.18	0.14	0.10
$T(t - 2)$	0.45	0.47	0.46	0.45	0.44	0.27	0.19	0.18	0.18	0.17	0.12	0.09
$Q(t - 3)$	0.44	0.43	0.40	0.39	0.37	0.29	0.18	0.17	0.17	0.15	0.10	0.07
$R(t - 3)$	0.42	0.42	0.39	0.37	0.36	0.30	0.12	0.09	0.09	0.08	0.05	0.05
$T(t - 3)$	0.39	0.40	0.38	0.36	0.35	0.32	0.14	0.06	0.06	0.05	0.02	0.02
$Q(t - 3)$	0.38	0.39	0.37	0.35	0.34	0.16	0.08	0.01	0.01	0.01	0.01	0.01
$R(t - 3)$	0.36	0.35	0.33	0.32	0.31	0.15	0.07	0.01	0.01	0.01	0.01	0.01
$T(t - 3)$	0.35	0.34	0.32	0.30	0.29	0.17	0.05	0.01	0.01	0.01	0.01	0.01
Variance proportion (%)	67	23	4	3	2	0.45	0.22	0.10	0.10	0.07	0.03	0.03

Table 5 Outputs of the PCA model for the selection of the appropriate metrological inputs for the Tamar station

Inputs	PC (1)	PC (2)	PC (3)	PC (4)	PC (5)	PC (6)	PC (7)	PC (8)	PC (9)	PC (10)	PC (11)	PC (12)
Eigenvectors obtained through the application of PCA												
$Q(t - 1)$	0.86	0.82	0.67	0.54	0.49	0.39	0.38	0.30	0.25	0.18	0.15	0.14
$R(t - 1)$	0.85	0.84	0.65	0.49	0.47	0.36	0.35	0.29	0.28	0.17	0.14	0.13
$T(t - 1)$	0.75	0.75	0.60	0.48	0.44	0.35	0.32	0.26	0.25	0.14	0.14	0.12
$Q(t - 2)$	0.73	0.72	0.54	0.46	0.45	0.32	0.30	0.25	0.23	0.12	0.11	0.11
$R(t - 2)$	0.67	0.56	0.50	0.45	0.43	0.29	0.27	0.24	0.22	0.19	0.10	0.10
$T(t - 2)$	0.65	0.45	0.45	0.42	0.39	0.27	0.25	0.23	0.21	0.18	0.09	0.09
$Q(t - 3)$	0.54	0.33	0.32	0.30	0.29	0.26	0.24	0.20	0.19	0.09	0.08	0.08
$R(t - 3)$	0.43	0.32	0.28	0.26	0.25	0.25	0.23	0.19	0.18	0.08	0.07	0.07
$T(t - 3)$	0.36	0.30	0.27	0.25	0.22	0.22	0.20	0.18	0.17	0.07	0.06	0.06
$Q(t - 3)$	0.35	0.29	0.25	0.24	0.21	0.21	0.19	0.17	0.15	0.06	0.05	0.05
$R(t - 3)$	0.32	0.27	0.24	0.23	0.20	0.20	0.17	0.16	0.14	0.05	0.04	0.04
$T(t - 3)$	0.29	0.26	0.23	0.22	0.19	0.19	0.15	0.15	0.13		0.02	0.02
Variance proportion (%)	68	22	5	2	2	0.45	0.22	0.10	0.10	0.07	0.03	0.03

necessary to quantify the input and parameter uncertainty simultaneously. A soft computing model can be defined as follows (Mustafa et al. 2018):

$$SSL = M(\overline{IN}, \theta) + e \tag{21}$$

where SSL is the SSL (output) matrix of the model (M), \overline{IN} is the input matrix of the model, e is the residual error and θ is the model parameter. The error input model is used to quantify the uncertainty of input data (Mustafa et al. 2018)

$$IN_i = \overline{IN}_i \times m_i \tag{22}$$

where \overline{IN}_{ij} is the observed input data for the i th day and m_i is the respective input multiplier. The multipliers are assumed as additional latent variables (ALVs). The ALVs are computed along with model parameters. In this study, a uniform prior probability distribution was used for each input parameter and model parameter.

The inability of the standalone and hybrid ANN structures is considered as another source of uncertainty. The Bayesian model is widely used to account for the model structural uncertainty. Traditionally, the residual errors in SSL modelling are normally distributed with mean zero and constant standard deviation. Thus, the likelihood equation can be defined as follows (Mustafa et al. 2018):

$$L(\theta|IN, \widehat{SSL}) = \prod_{t=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2} (\widehat{ssl}_t - ssl_t)^2\right) = \left(\frac{1}{\sqrt{2\pi}\sigma^2}\right)^N \exp\left(-\frac{1}{2\sigma^2} \sum_{t=1}^N (\widehat{ssl}_t - ssl_t)^2\right) \tag{23}$$

where \widehat{SSL} is the measured suspended sediment load, ssl_t is the estimated suspended sediment load at time t , \widehat{ssl}_t is the measured SSL at time t , σ is the standard deviation and N is the total number of time steps. In this article, the differential evolution adaptive metropolis (DREAM) algorithm was used as a sampler to sample the posterior distribution using likelihood function. The DREAM algorithm is known as a multi-chain Markov Chain Monte Carlo method (Vrugt 2016). The DREAM algorithm is widely used to estimate the uncertainty arising from the input parameter and model parameter (Vrugt 2016; Vrugt et al. 2009). A Markov chain is generated by the DREAM algorithm. This chain is ergodic with the unique stationary distribution. More explanations about the methods are observed in Vrugt (2016).

Bayesian model averaging

To combine predictions from multiple predictive models, the Bayesian model averaging (BMA) should be used (Mustafa et al. 2020).

The uncertainty of each model’s prediction can be considered by the BMA. An average prediction along with an associated prediction distribution is provided by the BMA (Mustafa et al. 2018). The posterior distribution of the BMA predictions can be computed as follows:

$$SSL_j = \sum_{k=1}^K \mu_k F_{jk} + e_j \tag{24}$$

$$p(SSL_j|F_{jk}) = \sum_{k=1}^k p(SSL_j|F_{jk}, N_k)p(N_k|F_{jk})$$

where SSL_j is the weighted average prediction of SSL, F_{jk} is the point prediction of each model for j (1, 2, ..., n observations) and k (1, 2, ..., K is the number of models, μ_k is the weight vector of each model, $p(SSL_j|F_{jk}, N_k)$ is the posterior predictive distortion of SSL_j on F_{jk} under the regarded model N_k and $p(N_k|F_{jk})$ is the posterior probability of the respective model N_k . The BMA predictive mean and variance of SSL are conditional to the discrete ensemble of the suggested models.

$$E[SSL_j|F_{jk}] = E_N[E(SSL_j)|F_{jk}, N] = \sum_{k=1}^K E[SSL_j|F_{jk}, N_k] p(N_k|F_{jk}) \tag{25}$$

$$\text{Var}[SSL_j|F_{jk}] = E_N[\text{Var}(y_j|F_{j,k}, N)] + \text{Var}[E(y_j)|F_{j,k}, N] = \sum_{k=1}^k \text{Var}[SSL_j|F_{jk}, N_k] p(N_k|F_{jk}) + \sum_{k=1}^k (E[SSL_j|F_{j,k}, N_k] - E(SSL_j|F_{jk}))^2 p(N_k|F_{jk}) \tag{26}$$

where $E(SSL_j) | F_{jk}, N$ is the expected value of SSL and $\text{Var}(y_j|F_{j,k}, N)$ is the variance of SSL_j on F_{jk} under the regarded model N_k .

The BMA can provide the accurate estimations if the widths and standard deviation are accurately estimated. The log-likelihood function is used to estimate the weights and standard deviations.

$$L(\mu_{\text{BMA}}, \sigma_{\text{BMA}}|F, \text{SSL})$$

$$= \sum_{j=1}^n \log \left[\sum_{k=1}^k \mu_k \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left[\frac{-1}{2}\right] \sigma_k^{-2} (SSL_j - F_{kj})^2 \right] \tag{27}$$

In this study, the fully Bayesian approach using input uncertainty multipliers is integrated with the Bayesian model averaging as follows:

- 1- A number of predictive models (ANN-ALO, ANN-BA and ANN-PSO) are suggested to estimate the SSL.
- 2- The uncertainty is computed using model parameters and input data (Mustafa et al. 2020).
- 3- An error input model is defined using Eq. (6).
- 4- The hydrologically rational pristine domains are chosen for the input data (Temperature, rainfall, and discharge) and model parameters (weight values and bias connections) (Mustafa et al. 2018).
- 5- A likelihood function is considered such as section ‘Simultaneous parameter and input uncertainty estimation’ (Eq. (23)).
- 6- The DREAM algorithm is used to compute input multipliers and model parameters.
- 7- A pre-determined number of outputs are provided for each model, using the parameter values found from level 2 to 6 (Mustafa et al. 2020).
- 8- The DREAM algorithm is used to compute the model weights and variances of each ensemble member of as described in section ‘Bayesian model averaging’.

- 9- The weights for all chosen ensemble members of each predictive model are summed to calculate the model weights.
- 10- Finally, multimodal forecasts are found by evaluating predictive mean and variance using Eqs. (25) and (26), respectively.

The following indices are used to evaluate the performance of the soft computing models:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |S_{i0} - S_{ic}| \tag{28}$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N ((S_{i0}) - (S_{ic}))^2} \tag{29}$$

$$\text{NSE} = 1 - \frac{\sum_{i=1}^N (S_{i0} - S_{ic})^2}{\sum_{i=1}^N (S_{ic} - \bar{S}_{ic})^2} \tag{30}$$

$$p = \frac{1}{N} \text{count}(SSL|SSL_L \leq SSL \leq SSL_U) \times 100 \tag{31}$$

$$d = \frac{\bar{d}_x}{\sigma_x} \tag{32}$$

$$\bar{d}_x = \frac{1}{N} \sum_{i=1}^N (SSL_U - SSL_L)$$

where RMSE is the root-mean-square error, NSE is the Nash-Sutcliffe efficiency, MAE is the mean absolute error, N is the data count, SSL_L is the lower value of SSL, SSL_U is the upper value of SLL, p is the percentage of observed data bracketed by the 95% prediction uncertainty (95PPU), σ_x is the standard deviation and d is the average width of confidence interval.

Table 6 Sensitivity analysis of the different algorithms

	Value of parameter	Objective function value
PSO		
Population size	100	1.15
	200	1.10
	300	1.16
	400	1.22
Inertia coefficient	0.3	1.17
	0.5	1.09
	0.7	1.20
	0.90	1.24
c_1	1.7	1.16
	1.8	1.12
	1.9	1.17
	2.0	1.19
c_2	1.7	1.16
	1.8	1.12
	1.9	1.17
	2.0	1.19
BA		
Population size	50	1.02
	100	0.999
	150	1.12
	200	1.15
Maximum frequency	3	1.19
	5	1.17
	7	0.999
	9	1.09
Minimum frequency	1	1.29
	2	1.05
	3	0.994
	4	1.09
Maximum loudness	0.30	1.19
	0.50	1.12
	0.70	0.993
	0.90	1.06
Minimum loudness	0.10	1.19
	0.20	1.17
	0.30	0.997
	0.40	1.02
ALO		
Population size	20	0.994
	40	0.810
	60	0.829
	80	0.915
Maximum number of iterations	50	0.926
	100	0.819
	150	0.821
	200	0.822

Table 7 Evaluation parameters of studied models in the training and testing levels (Narmab River)

Model	Training			Testing		
	RMSE	MAE	NSE	RMSE	MAE	NSE
ANN-ALO (5)	0.812	0.732	0.97	1.21	1.19	0.96
ANN-BA (5)	0.998	0.824	0.96	1.24	1.21	0.95
ANN-PSO (5)	1.09	0.895	0.94	1.29	1.27	0.93
ANN-ALO (4)	0.89	0.739	0.96	1.31	1.23	0.94
ANN-BA (4)	1.12	0.832	0.95	1.35	1.3	0.94
ANN-PSO (4)	1.23	0.911	0.93	1.42	1.4	0.92
ANN-ALO (3)	1.01	0.812	0.92	1.32	1.28	0.91
ANN-BA (3)	1.44	0.878	0.91	1.49	1.37	0.89
ANN-PSO (3)	1.67	0.924	0.89	1.51	1.45	0.87
ANN-ALO (2)	1.59	1.12	0.87	1.59	1.57	0.85
ANN-BA (2)	1.69	1.32	0.85	1.67	1.63	0.83
ANN-PSO (2)	1.73	1.43	0.83	1.73	1.69	0.82
ANN-ALO (1)	1.62	1.24	0.86	1.63	1.62	0.84
ANN-BA (1)	1.69	1.45	0.82	1.69	1.67	0.81
ANN-PSO (1)	1.73	1.62	0.8	1.76	1.73	0.78

Results and discussion

Sensitivity analysis

The application of the optimization algorithms, into the ANN model, requires the prior adjustment of a suitable

set of parameters. As in any optimization algorithm, the BA, ALO and PSO are relatively sensitive to setup parameters. Hence, to get the most out of optimization algorithms, it is necessary to fine-tune the parameters, before applying them to the ANN model. While the objective function value is varied for a parameter of

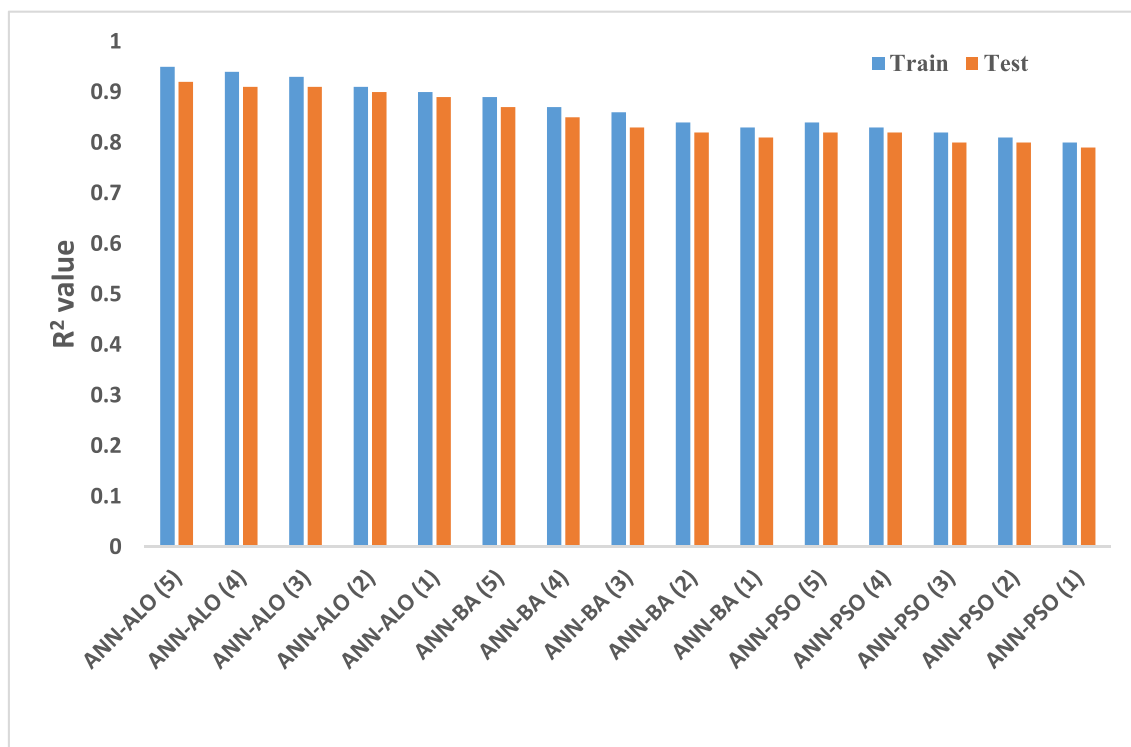
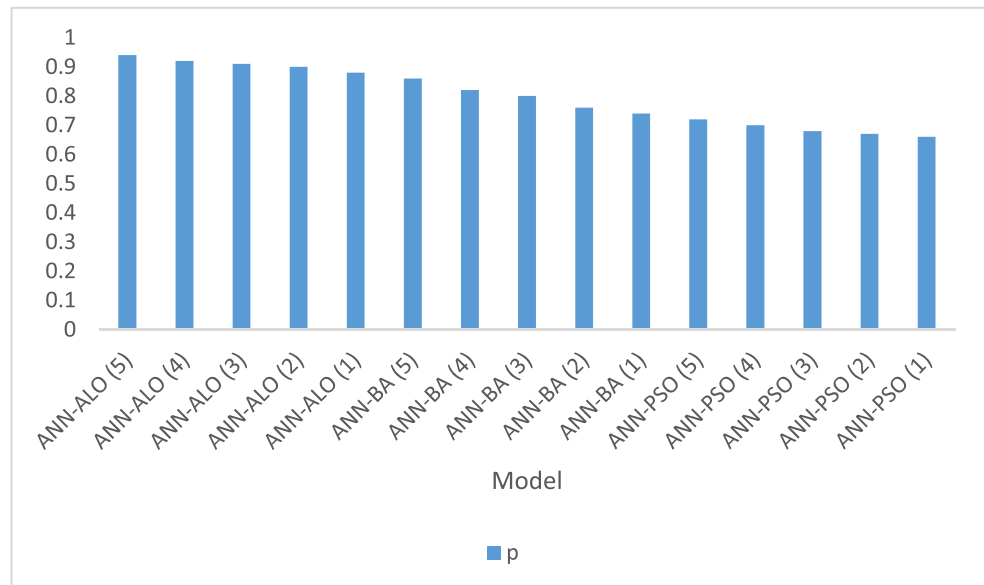
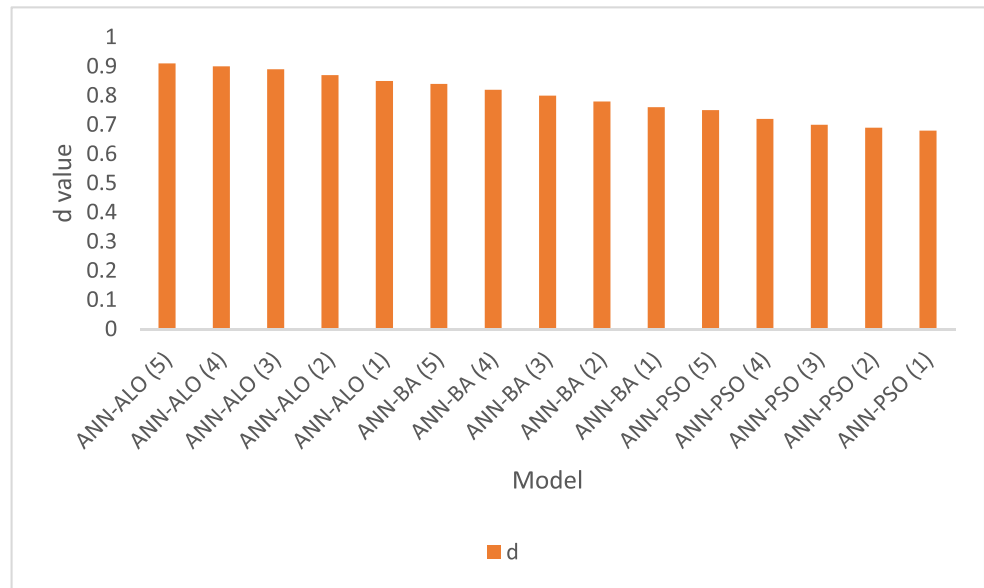


Fig. 10 Coefficient of determination (R^2) for all used models for the training and testing stages

Fig. 11 Uncertainty analysis for the proposed SSL models in the Narmab River



interest, by changing the value of that parameter, the values of other parameters in each algorithm are fixed. For this undertaking, we used RMSE as the objective function. However, depending on problem complexity, the value of the parameter of interest is altered for the running of the algorithm. Table 6 shows the results of sensitivity analysis for the optimization algorithms. It should be noted that the best result of the BA algorithm, i.e. minimum objective function value, occurs at a population size of 100, minimum frequency of 3, maximum frequency of 7, maximum loudness of 0.7 and minimum loudness of 0.3. The outputs of other algorithms are illustrated in Table 6.

Statistical results for soft computing models for the Narmab station

The outputs of the statistical parameters, for studied models in the training level, are provided in Table 7. As portrayed in this table, the periodicity significantly increased the accuracy of each model. For the ANN-ALO model, the NSE increased from 0.86 (for the first input combination) to 0.97 (for the fifth input combination). The RMSE and MAE decreased from 1.62 to 0.812 and from 1.24 to 0.732, respectively. A comparison of the ANN-ALO, ANN-BA and ANN-PSO models revealed the superiority of the ANN-ALO (5) model over the ANN-BA and ANN-PSO models. The ANN-PSO model was

Table 8 Results of models with meteorological data

Model	Station											
	Narmab River			Tamar			Basir Abad			Salian		
	RMSE	MAE	NSE	RMSE	MAE	NSE	RMSE	MAE	NSE	RMSE	MAE	NSE
Training												
ANN-ALO (3)	1.22	1.14	0.95	1.12	1.10	0.96	1.14	1.10	0.94	1.10	1.06	0.90
ANN-BA (3)	1.25	1.16	0.93	1.16	1.14	0.94	1.17	1.15	0.92	1.12	1.10	0.89
ANN-PSO (3)	1.29	1.19	0.90	1.18	1.15	0.91	1.19	1.17	0.90	1.15	1.21	0.87
Testing												
ANN-ALO (3)	1.25	1.16	0.93	1.14	1.24	0.94	1.19	1.14	0.90	1.19	1.12	0.86
ANN-BA (3)	1.23	1.19	0.89	1.18	1.29	0.92	1.23	1.24	0.89	1.23	1.19	0.85
ANN-PSO (3)	1.37	1.25	0.86	1.23	1.33	0.86	1.29	1.26	0.86	1.25	1.29	0.84

also outperformed by the ANN-BA model. According to the results attained, the ANN-ALO (5) model improved the accuracy of the ANN-BA (5) and ANN-PSO (5) models by approximately 18% and 26% for the RMSE, respectively. The results also indicate that the ANN-BA (5) model increased the accuracy of the ANN-PSO (5) model by 9.4% and 8.01% for RMSE and MAE, respectively.

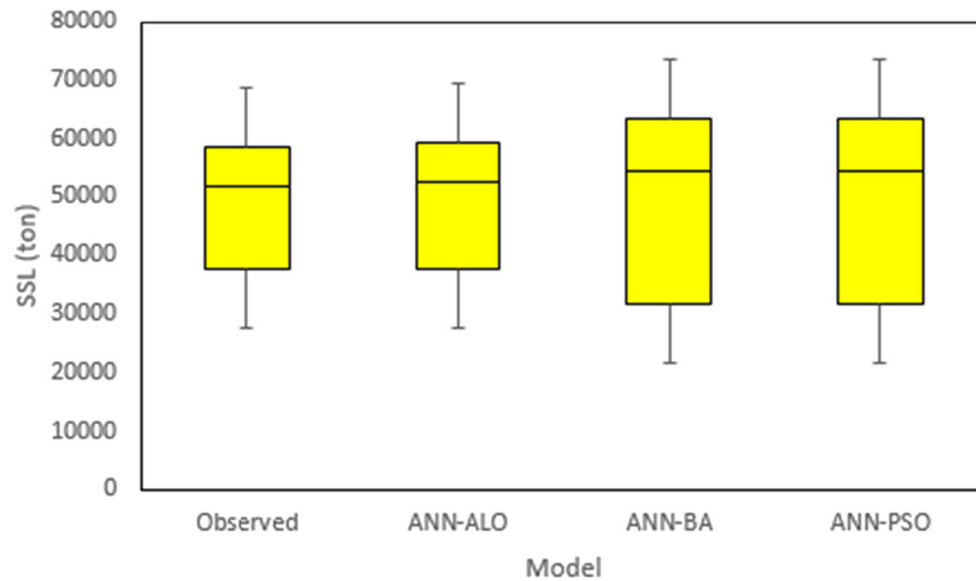
The results from investigations on the statistical parameters for models studied in the testing level are provided in Table 7. These results revealed that (a) the periodicity enhanced the accuracy of the hybrid ANN models and (b) the highest RMSE and MAE values were achieved by the ANN-PSO (1) model. From Table 6, it can be surmised that generally, the ANN-ALO (5) models are the most favourable and the ANN-PSO (1) models the least favourable. Figure 10 shows the bar chart for the soft computing models. In terms of the R^2 index, the ANN-ALO (5) model, with the values of 0.95 (training level) and 0.92 (testing level), performed better than the other models. It was also observed that the accuracy of the ANN-BA model is superior to that of the ANN-PSO model. It is noteworthy that the highest and lowest values of R^2 were attained by the ANN-ALO (5) and ANN-PSO (1) models, respectively.

As mentioned earlier, one of the main aims of this undertaking is to investigate the uncertainty of the different hybrid ANN models, by way of the p and d factors. The lagged input data (lagged SSL values) have the uncertainties due to the error measurements. Thus, it is necessary to perform the uncertainty analysis. The reduction in d parameter and the increase in p parameter resulted in a more favourable uncertainty for all the models. In the case of the ANN-ALO model, p increased from 0.66 (the first input combination) to 0.94 (the fifth input combination). The results indicate that the periodicity significantly improved the accuracy of each model. As can be observed in Fig. 11, the minimum and maximum values of p are 0.66 and 0.94 for ANN-ALO (5) and ANN-

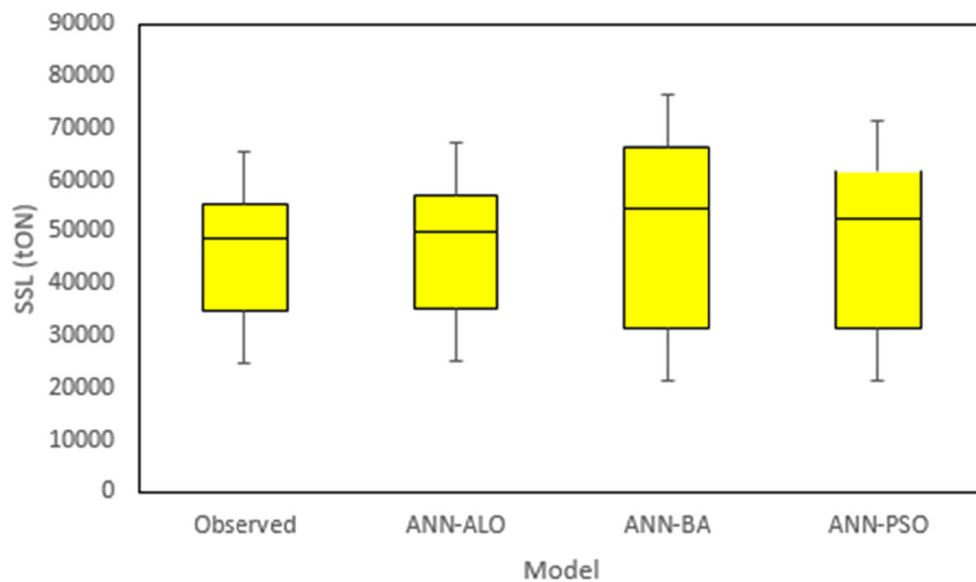
PSO (1), respectively, while the maximum and minimum values of d are 0.39 and 0.14 for ANN-ALO (5) and ANN-PSO (1), respectively. A comparison between the ANN-BA and ANN-PSO models revealed that the ANN-BA model has a higher value of p and a lower value of d .

Evaluation of ANN models using meteorological inputs

Table 8 (training) shows the training results of statistical models using meteorological input data. Comparison of ANN-ALO, ANN-BA and ANN-PSO indicated that the ANN-ALO model was better than the ANN-BA and ANN-PSO models. The ANN-ALO increased the accuracy of ANN-BA and ANN-PSO by approximately 2.4% and 5.4% for RMSE in the Narmab River, respectively. From the MAE point of view, ANN-ALO with the values of 1.14 ton, 1.10 ton, 1.10 ton and 1.06 ton had better accuracy in comparison with the ANN-BA and ANN-PSO in the Narmab River, Tamar, Basir Abad and Salian stations. Table 8 (testing) indicates the testing results using meteorological input data. The results revealed that the ANN-ALO was the optimum model, as verified by a RMSE value of 1.25, MAE of 1.16 ton and NSE of 1.93 in the Narmab River. Among the hybrid ANN models, the hybrid ANN-ALO was seen to have the highest value of NSE in the Narmab River, Tamar, Basir Abad and Salian stations. The results also indicated that the ANN-ALO and ANN-BA had the lowest MAE values in the Narmab River, Tamar, Basir Abad and Salian stations. According to Table 8, the accuracy of the ANN-PSO model with the higher values of RMSE and MAE and lower values of NSE is worse than that of the other models. Figure 12 shows the boxplot of ANN models in different stations. The results indicated that the maximum and minimum shapes of boxplots provided by the ANN-ALO followed the observed values compared to those provided by the ANN-BA and



(a)

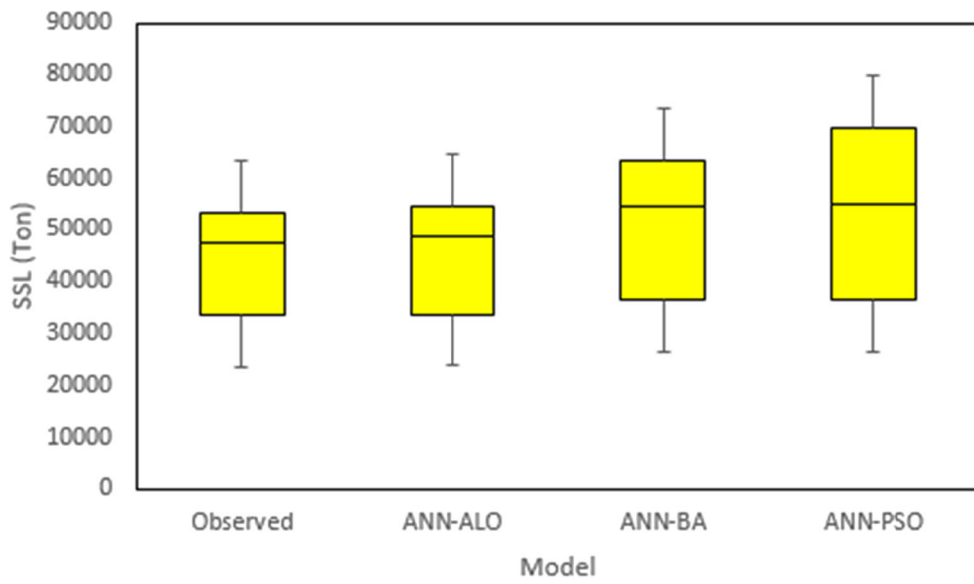


(b)

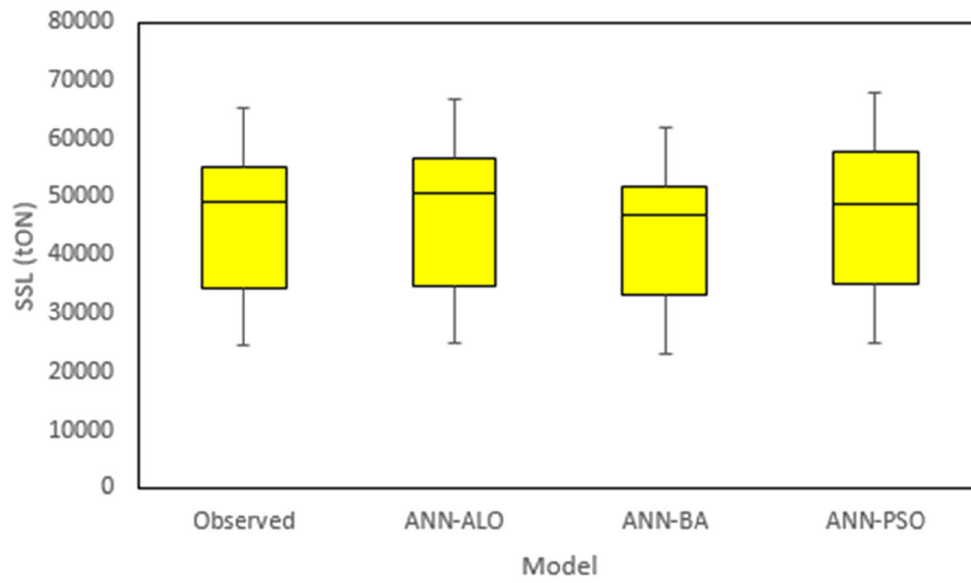
Fig. 12 The boxplots for the **a** Narmab River, **b** Basir Abad, **c** Salian and **d** Tamar stations

ANN-PSO models. It can be concluded that there is a significant difference between the ANN-PSO with the observed data. Figure 13 indicates the uncertainty results using meteorological inputs. In the case of the ANN-ALO model, the d value decreased from 0.16 (ANN-PSO) to 0.12 (ANN-ALO), from 0.19 (ANN-PSO) to 0.12 (ANN-ALO), from 0.25 (ANN-PSO) to 0.21 (ANN-ALO) and from 0.26 (ANN-

PSO) to 0.22 (ANN-ALO) in the Narmab River, Tamar, Basir Abad and Salian stations, respectively. As observed in Fig. 12, the maximum and minimum values of p are 0.91 and 0.89 in the Narmab River. A comparison between the ANN-BA and ANN-PSO models revealed that the ANN-PSO model has a lower value of p and a higher value of d in different stations.



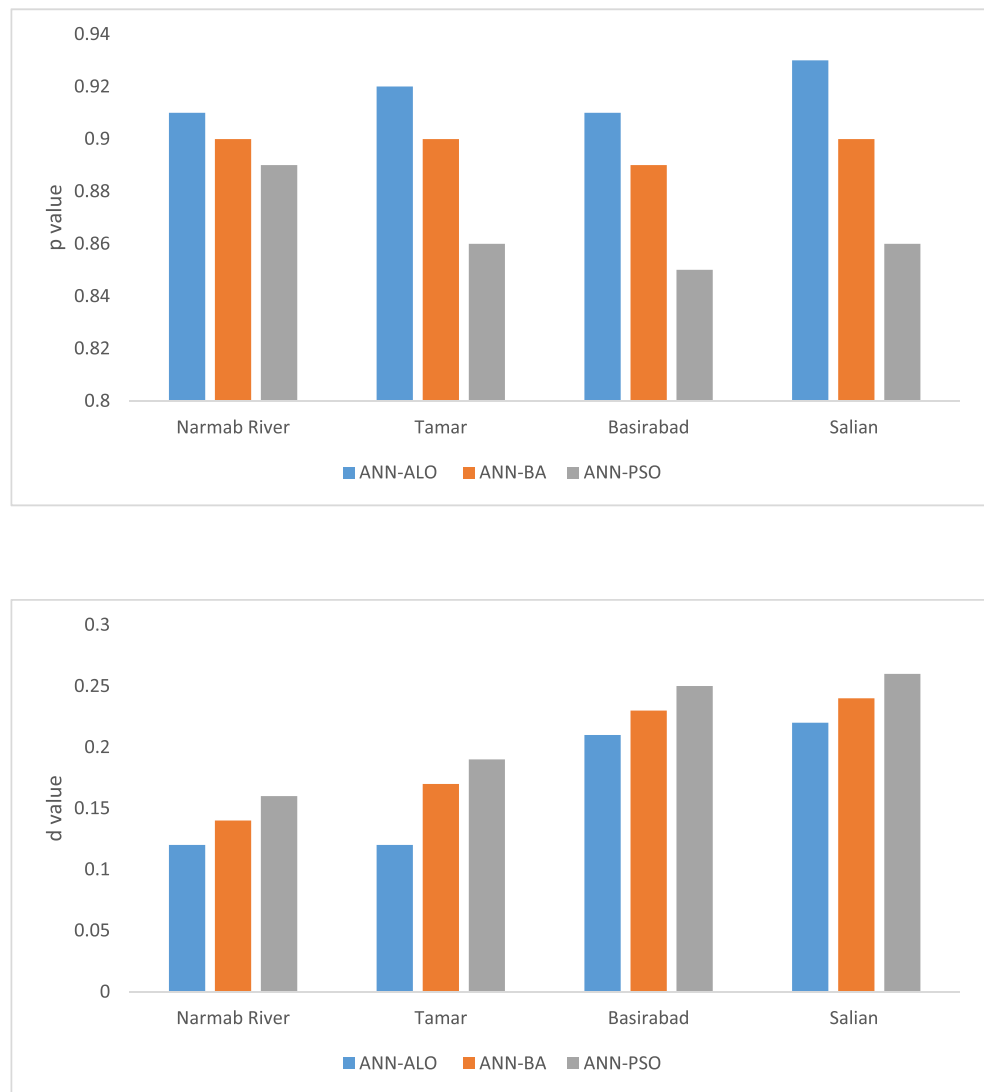
(c)



(d)

Fig. 12 continued.

Fig. 13 The uncertainty results using meteorological inputs



Conclusions

The modelling of different SSL input scenarios (including previous sediment values) was performed to investigate the effectiveness of the ANN-ALO model. The outputs were compared with the ANN-PSO and ANN-BA models. The meteorological inputs and lagged SSLs were used as the inputs to the models. First, lagged SSLs were used to estimate the SSL in the Narmab River station. Five input combinations were used to estimate the lagged SSLs in the Narmab River station. The results indicated that the ANN-ALO (5) model improved the accuracy of the ANN-BA (5) and ANN-PSO (5) models by approximately 18% and 26% for the RMSE, respectively. Then, the meteorological inputs were used to estimate SSL in different stations. The principal component analysis was used to select the best input combinations. The ANN-ALO model was the optimum model as verified by a RMSE value of 1.22,

1.12, 1.14 and 1.10 for the training data set in the Narmab River, Tamar, Basir Abad and Salian stations. The results reflected that the ANN-PSO model was the worst model, as verified by the RMSE value of 1.29 (Narmab River), 1.18 (Tamar), 1.19 (Basir Abad) and 1.15 (Salian) for the training data set. The uncertainty results indicated that the ANN-ALO model had the highest p value and the lowest d value in different stations. The findings of the current study indicate that the hybrid ANN models have a high ability for predicting SSL. The next papers can integrate the climate models with the hybrid soft computing models to predict SSL for the future periods under climate change condition. Additionally, the ALO can be used to develop the other soft computing models such as ANFIS and SVM model for predicting SSL.

Acknowledgements The authors gratefully acknowledge the technical facility support received from the University of Malaya, Malaysia.

References

- Abobakr Yahya AS, Ahmed AN, Binti Othman F, Ibrahim RK, Afan HA, el-Shafie A, Fai CM, Hossain MS, Ehteram M, Elshafie A (2019) Water quality prediction model based support vector machine model for Ungauged River catchment under dual scenarios. *Water* 11: 1231. <https://doi.org/10.3390/w11061231>
- Afan HA, El-Shafie A, Yaseen ZM et al (2014) ANN based sediment prediction model utilizing different input scenarios. *Water Resour Manag* 29:1231–1245. <https://doi.org/10.1007/s11269-014-0870-1>
- Afan HA, Allawi MF, El-Shafie A et al (2020) Input attributes optimization using the feasibility of genetic nature inspired algorithm: application of river flow forecasting. *Sci Rep* 10. <https://doi.org/10.1038/s41598-020-61355-x>
- Ali ES, Abd Elazim SM, Abdelaziz AY (2017) Ant lion optimization algorithm for optimal location and sizing of renewable distributed generations. *Renew Energy* 101:1311–1324. <https://doi.org/10.1016/j.renene.2016.09.023>
- Allawi MF, Jaafar O, Ehteram M, Mohamad Hamzah F, el-Shafie A (2018) Synchronizing artificial intelligence models for operating the dam and reservoir system. *Water Resour Manag* 32:3373–3389. <https://doi.org/10.1007/s11269-018-1996-3>
- Banadkooki FB, Ehteram M, Ahmed AN, Teo FY, Fai CM, Afan HA, Sapitang M, el-Shafie A (2020) Enhancement of groundwater-level prediction using an integrated machine learning model optimized by whale algorithm. *Nat Resour Res*. <https://doi.org/10.1007/s11053-020-09634-2>
- Chang Q, Zhang C, Zhang S, Li B (2019) Streamflow and sediment declines in a loess hill and gully landform basin due to climate variability and anthropogenic activities. *Water* 11:2352. <https://doi.org/10.3390/w11112352>
- Cui Z, Cao Y, Cai X et al (2019) Optimal LEACH protocol with modified bat algorithm for big data sensing systems in Internet of Things. *J Parallel Distrib Comput* 132:217–229. <https://doi.org/10.1016/j.jpdc.2017.12.014>
- Ehteram M, Karami H, Mousavi SF, Farzin S, Celeste AB, Shafie AE (2018a) Reservoir operation by a new evolutionary algorithm: kidney algorithm. *Water Resour Manag* 32:4681–4706. <https://doi.org/10.1007/s11269-018-2078-2>
- Ehteram M, Singh VP, Karami H et al (2018b) Irrigation management based on reservoir operation with an improved weed algorithm. *Water* 10:1267. <https://doi.org/10.3390/w10091267>
- Ehteram M, El-Shafie AH, Hin LS et al (2019a) Toward bridging future irrigation deficits utilizing the shark algorithm integrated with a climate change model. *Appl Sci* 9:3960. <https://doi.org/10.3390/app9193960>
- Ehteram M, Ghotbi S, Kisi O et al (2019b) Investigation on the potential to integrate different artificial intelligence models with metaheuristic algorithms for improving river suspended sediment predictions. *Appl Sci* 9:4149. <https://doi.org/10.3390/app9194149>
- Ehteram M, Afan HA, Dianatikah M, Ahmed AN, Ming Fai C, Hossain MS, Allawi MF, Elshafie A (2019c) Assessing the predictability of an improved ANFIS model for monthly streamflow using lagged climate indices as predictors. *Water* 11:1130. <https://doi.org/10.3390/w11061130>
- Emamgholizadeh S, Demneh RK (2018) A comparison of artificial intelligence models for the estimation of daily suspended sediment load: a case study on the Telar and Kasilian rivers in Iran. *Water Sci Technol Water Supply* 19:165–178. <https://doi.org/10.2166/ws.2018.062>
- Ehteram M, Mousavi S-F, Karami H, Farzin S, Deo R, Othman FB, Chau KW, Sarkamaryan S, Singh VP, el-Shafie A (2018) Bat algorithm for dam–reservoir operation. *Environ Earth Sci* 77. <https://doi.org/10.1007/s12665-018-7662-5>
- Farzin S, Singh V, Karami H, Farahani N, Ehteram M, Kisi O, Allawi M, Mohd N, el-Shafie A (2018) Flood routing in river reaches using a three-parameter Muskingum model coupled with an improved bat algorithm. *Water* 10:1130. <https://doi.org/10.3390/w10091130>
- Jothiprakash V, Garg V (2009) Reservoir sedimentation estimation using artificial neural network. *J Hydrol Eng* 14:1035–1040. [https://doi.org/10.1061/\(asce\)he.1943-5584.0000075](https://doi.org/10.1061/(asce)he.1943-5584.0000075)
- Kakaei Lafdani E, Moghaddam Nia A, Ahmadi A (2013) Daily suspended sediment load prediction using artificial neural networks and support vector machines. *J Hydrol* 478:50–62. <https://doi.org/10.1016/j.jhydrol.2012.11.048>
- Karkalos NE, Efkolidis N, Kyratsis P, Markopoulos AP (2019) A comparative study between regression and neural networks for modeling Al6082-T6 alloy drilling. *Machines* 7(1):13
- Khan MYA, Hasan F, Tian F (2018) Estimation of suspended sediment load using three neural network algorithms in Ramganga River catchment of Ganga Basin, India. *Sustain Water Resour Manag* 5: 1115–1131. <https://doi.org/10.1007/s40899-018-0288-7>
- Khosravi K, Mao L, Kisi O, Yaseen ZM, Shahid S (2018) Quantifying hourly suspended sediment load using data mining models: case study of a glacierized Andean catchment in Chile. *J Hydrol* 567: 165–179. <https://doi.org/10.1016/j.jhydrol.2018.10.015>
- Kisi O, Dailr AH, Cimen M, Shiri J (2012) Suspended sediment modeling using genetic programming and soft computing techniques. *J Hydrol* 450–451:48–58. <https://doi.org/10.1016/j.jhydrol.2012.05.031>
- Melesse AM, Ahmad S, McClain ME et al (2011) Suspended sediment load prediction of river systems: an artificial neural network approach. *Agric Water Manag* 98:855–866. <https://doi.org/10.1016/j.agwat.2010.12.012>
- Mirjalili S (2015) The ant lion optimizer. *Adv Eng Softw* 83:80–98. <https://doi.org/10.1016/j.advengsoft.2015.01.010>
- Misset C, Recking A, Navratil O, Legout C, Poirel A, Cazilhac M, Briguet V, Esteves M (2019) Quantifying bed-related suspended load in gravel bed rivers through an analysis of the bedload-suspended load relationship. *Earth Surf Process Landforms*. <https://doi.org/10.1002/esp.4606>
- Mustafa SMT, Nossent J, Ghysels G, Huysmans M (2018) Estimation and impact assessment of input and parameter uncertainty in predicting groundwater flow with a fully distributed model. *Water Resour Res* 54(9):6585–6608
- Mustafa SMT, Nossent J, Ghysels G, Huysmans M (2020) Integrated Bayesian Multi-model approach to quantify input, parameter and conceptual model structure uncertainty in groundwater modeling. *Environ Model Softw* 126:104654. <https://doi.org/10.1016/j.envsoft.2020.104654>
- Najah Ahmed A, Binti Othman F, Abdulmohsin Afan H, Khaleel Ibrahim R, Ming Fai C, Shabbir Hossain M, Ehteram M, Elshafie A (2019) Machine learning methods for better water quality prediction. *J Hydrol* 578:124084. <https://doi.org/10.1016/j.jhydrol.2019.124084>
- Noori R, Karbassi AR, Moghaddamnia A, Han D, Zokaei-Ashtiani MH, Farokhnia A, Gousheh MG (2011) Assessment of input variables determination on the SVM model performance using PCA, gamma test, and forward selection techniques for monthly stream flow prediction. *J Hydrol* 401(3–4):177–189
- Nourani V, Andalib G (2015) Daily and monthly suspended sediment load predictions using wavelet based artificial intelligence approaches. *J Mt Sci* 12:85–100. <https://doi.org/10.1007/s11629-014-3121-2>
- Samantaray S, Ghose DK (2018) Evaluation of suspended sediment concentration using descent neural networks. *Procedia Comput Sci* 132: 1824–1831. <https://doi.org/10.1016/j.procs.2018.05.138>
- Talebi A, Mahjoobi J, Dastorani MT, Moosavi V (2016) Estimation of suspended sediment load using regression trees and model trees approaches (case study: Hyderabad drainage basin in Iran). *ISH J*

- Hydraul Eng 23:212–219. <https://doi.org/10.1080/09715010.2016.1264894>
- Tayfur G, Karimi Y, Singh VP (2013) Principle component analysis in conjunction with data driven methods for sediment load prediction. *Water Resour Manag* 27:2541–2554. <https://doi.org/10.1007/s11269-013-0302-7>
- Tharwat A, Hassanien AE (2017) Chaotic antlion algorithm for parameter optimization of support vector machine. *Appl Intell* 48:670–686. <https://doi.org/10.1007/s10489-017-0994-0>
- Tian T, Liu C, Guo Q, Yuan Y, Li W, Yan Q (2018) An improved ant lion optimization algorithm and its application in hydraulic turbine governing system parameter identification. *Energies* 11:95. <https://doi.org/10.3390/en11010095>
- Tikhmarine Y, Souag-Gamane D, Najah Ahmed A, Kisi O, el-Shafie A (2020) Improving artificial intelligence models accuracy for monthly streamflow forecasting using grey wolf optimization (GWO) algorithm. *J Hydrol* 582:124435. <https://doi.org/10.1016/j.jhydrol.2019.124435>
- Vafakhah M (2012) Comparison of cokriging and adaptive neuro-fuzzy inference system models for suspended sediment load forecasting. *Arab J Geosci* 6:3003–3018. <https://doi.org/10.1007/s12517-012-0550-5>
- Valikhan-Anaraki M, Mousavi S-F, Farzin S, Karami H, Ehteram M, Kisi O, Fai CM, Hossain MS, Hayder G, Ahmed AN, el-Shafie AH, Bin Hashim H, Afan HA, Lai SH, el-Shafie A (2019) Development of a novel hybrid optimization algorithm for minimizing irrigation deficiencies. *Sustainability* 11:2337. <https://doi.org/10.3390/su11082337>
- Vrugt JA (2016) Markov chain Monte Carlo simulation using the DREAM software package: theory, concepts, and MATLAB implementation. *Environ Model Softw* 75:273–316
- Vrugt JA, Ter Braak C, Diks CGH, Robinson BA, Hyman JM, Higdon D (2009) Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *Int J Nonlin Sci Num* 10(3):273–290
- Wang J, Du P, Lu H et al (2018) An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. *Appl Soft Comput* 72:321–337. <https://doi.org/10.1016/j.asoc.2018.07.022>
- Yousif A, Sulaiman S, Diop L, Ehteram M, Shahid S, al-Ansari N, Yaseen Z (2019) Open channel sluice gate scouring parameters prediction: different scenarios of dimensional and non-dimensional input parameters. *Water* 11:353. <https://doi.org/10.3390/w11020353>
- Zhao Y, Huang S (2017) Pollution characteristics of industrial construction and demolition waste. *Pollut Control Resour Recovery*:51–101

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