Modeling and predicting suspended sediment load under climate change conditions: a new hybridization strategy
Saeed Farzin and Mahdi Valikhan Anaraki

ABSTRACT
In the present study, for the first time, a new strategy based on a combination of the hybrid least-squares support-vector machine (LS-SVM) and flower pollination algorithm (FPA), average 24 general circulation model (GCM) output, and delta change factor method has been developed to achieve the impacts of climate change on runoff and suspended sediment load (SSL) in the Lighvan Basin in the period (2020–2099). Also, the results of modeling were compared to those of LS-SVM and adaptive neuro-fuzzy inference system (ANFIS) methods. The comparison of runoff and SSL modeling results showed that the LS-SVM-FPA algorithm had the best results and the ANFIS algorithm had the worst results. After the acceptable performance of the LS-SVM-FPA algorithm was proved, the algorithm was used to predict runoff and SSL under climate change conditions based on ensemble GCM outputs for periods (2020–2034, 2035–2049, 2070–2084, and 2085–2099) under three scenarios of RCP2.6, RCP4.5, and RCP8.5. The results showed a decrease in the runoff in all periods and scenarios, except for the two near periods under the RCP2.6 scenario for runoff. The predicted runoff and SSL time series also showed that the SSL values were lower than the average observation period, except for 2036–2039 (up to an 8% increase in 2038).

Key words | climate change, flower pollination algorithm, hybridization strategy, least-squares support-vector machine, suspended sediment loads

HIGHLIGHTS
- A hybrid strategy has been developed for modeling and predicting suspended sediment load (SSL) under climate change.
- A hybrid strategy was compared to standalone LS-SVM and adaptive neuro-fuzzy inference system.
- This approach has the potential to model and predict various hydrological variables.
- The ensemble general circulation model of fifth report and RCP scenarios were employed for predicting SSL.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (http://creativecommons.org/licenses/by/4.0/).
INTRODUCTION

The suspended sediment movements are important in different fields, such as water resource management, water structure designs, and river and dam engineering. Modeling the amount of suspended sediment in the river is an essential issue to design water storage and flow control facilities, such as dams and canals. Also, the suspended sediments affect the quality of drinking water requirements of residential areas and the quality of water requirements of agriculture and industry. On the other hand, the suspended sediments are the result of complex and nonlinear flow processes in the river (Nourani et al. 2019). Therefore, modeling the nonlinear relationship between suspended sediments and river flow using different nonlinear methods has become one of the important challenges for different scientific societies, such as engineering and water resources management.

Meanwhile, machine learning algorithms are more popular than physical and mathematical methods, due to their high accuracy, lower cost, and few number parameters. In recent years, machine learning algorithms have been successfully applied in modeling various water resources and hydrology problems. Dariane & Azimi (2016) used the adaptive neuro-fuzzy inference system (ANFIS), and Seifi & Riahi (2020) used the least-square support-vector machine (LS-SVM) for modeling different parameters in water resource management and hydrology. Kumar et al. (2019) employed multiple linear regression, artificial neural network (ANN), radial basis function neural network, classification and regression tree, M5 model tree, and LS-SVM for daily suspended sediment modeling. Results of this study showed that LS-SVM and ANN produced better results.

Machine learning has a good performance in modeling various parameters of water resources and hydrology such as SSL. Some effective parameters of machine learning can only be determined by trial and error. However, the trial and error method generally requires a great deal of time and cost. In this case, metaheuristic optimization methods with more accuracy and lower computational cost than the trial and error method are useful approaches. Recently, hybrid machine learning algorithms and metaheuristic optimization algorithms have gained popularity in recent years. To model various hydrological parameters, Meshram et al. (2019) used the hybrid feed-forward neural networks with particle swarm and particle swarm-gravitational search algorithms, and Yaseen et al. (2018) used the hybrid of LS-SVM and BAT algorithm. The flower pollination algorithm (FPA) (Yang 2012) was also effective in increasing the accuracy of machine learning methods, such as ANN (Wang et al. 2017) and ANFIS (Farrokhi et al. 2020). Also, in the study conducted by Wu et al. (2019), the Extreme Learning Machine (ELM) optimized by the FPA was used to estimate the reference evapotranspiration. In this study, the hybrid ELM and flower pollination had more accuracy than the hybrid of ELM with genetic and ant algorithms. Also, mentioned methods had successful applications in other fields such as the optimal operation of the reservoir (Ehteram et al. 2018; Mohammadi...
et al. 2019) and the optimal design of the open channel (Farzin & Valikhan Anaraki 2020).

Naturally, there is a specified pattern in sediment transportation and flow in the river. However, the phenomenon of climate change has increased the speed of the hydrological cycle and changed the magnitude and temporal pattern of runoff with the increasing temperature and changing precipitation patterns. This can increase the concentration of sediments in the rivers and, consequently, the riverbed instability, and it can damage the structures around the river, as well as causing problems for living organisms. Therefore, the effect of climate change is necessary to study on the suspended sediment load (SSL).

The effect of climate change on the amount of SSL has been investigated by Azari et al. (2016) using the Soil and Water Assessment Tool (SWAT) model and different outputs of general circulation models (GCMs), such as HadCM3 and CGCM2. Zhou et al. (2017) used the SWAT model, ensemble GCMs, and the RCP4.5 and RCP8.5 (Representative Concentration Pathway) scenarios. Thus, the projection of flow and sediment yield under coordinated climate change and urbanization scenarios was investigated in Korea (Kim et al. 2017). The impact of climate change on the hydrological parameters such as sediment load was reviewed in the Blue Nile basin, Ethiopia (Gelete et al. 2019). The changes in flow and sediment load were estimated in a poorly gauged Brahmaputra river basin under an extreme climate scenario (Haque et al. 2020).

To the best of our knowledge, there are few studies for modeling runoff and SSL. Added to this, some of these studies considered climate change in predicting runoff and SSL. Hence, in the present study, for the first time, a new hybridization strategy has been developed for modeling and predicting runoff and SSL under climate change conditions. For this purpose, the hybrid of LS-SVM and FPA (LS-SVM-FPA) is employed for modeling and predicting runoff and SSL under climate change conditions. In this regard, the average of 24 GCM outputs, namely Ensemble GCM, is used for considering climate change conditions. Moreover, the delta change factor method is used for downscaling outputs of ensemble GCM. Thus, it is worth mentioning that after modeling, the results are compared with standalone LS-SVM and ANFIS to demonstrate the ability of LS-SVM-FPA in modeling runoff and SSL.

### MATERIALS AND METHODS

#### Case study and the used data

Lighvanghai Basin is located in the northwest of Iran and the southern city of Tabriz, and on the northern slopes of Sahand Mountain. This basin is the sub-basin of the Urmia Lake, and it is extended from eastern lengths of 46°20′0″–46°27′30″ to the north latitude of 37°42′55″–37°49′30″. The basin drains an area of 76 km², with an average discharge of 24.6 MCM. The basin has a rainy and humid climate. The maximum amount of runoff in the Lighvan Basin is related to spring. The average slope of the Lighvan Basin is 11%, and therefore, the soil of this basin is under severe erosion. In the present study, the Lighvan Basin is investigated because the Lighvan River is one of the largest subdivisions of the Lighvan Chai River, which discharges to Urmia Lake. Figure 1 illustrates the Lighvan basin location in the Urmia Lake Basin and in Iran.

In the present study, the runoff and SSL are modeled using meteorological data of the Lighvan Tabriz and Sahand synoptic stations and runoff and SSL data of the Lighvan hydrometric station. The statistical properties of the survey data are shown in Table 1. Also, to investigate the effect of climate change on runoff and SSL, the large-scale precipitation and temperature data of ensemble GCMs from the fifth report under three scenarios, including RCP2.6, RCP4.5, and RCP8.5, are used, downloaded from supported and distributed data in Canadian climate change scenarios network (http://climate-scenarios.canada.ca/?page=gridded-data). The ensemble model is obtained by averaging from 24 GCMs from the fifth report, considering equal weight (equal to 1) for each GCM. The names of these 24 GCMs and their weights for contracting the ensemble GCM are found in http://www.cccsn.ec.gc.ca. Table 2 indicates the characteristics of the three considered scenarios. Based on this table, RCP8.5 is the worst scenario. In the present study, based on the available statistical data, 15 years of 1990–2004 is considered as the observation period. This length of the observation period is common and sufficient in the different studies. For example, the length of the observed period in the study of Ashofteh et al. (2017) was equal to 14 years and in the study of
Valikhan-Anaraki et al. (2019) was equal to 12 years. The statistical periods of 2020–2034, 2034–2049, 2070–2084, and 2085–2099 are considered to evaluate the effect of climate change in the near and far future.

**The proposed strategy for the prediction of SSL under climate change**

In the present study, a new strategy has been developed for modeling and predicting SSL under climate change

| Location of Lighavan Basin in Lake Urmia Basin and Iran country. |

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Statistical characteristics of data set for investigated stations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Station name</td>
<td>Station type</td>
</tr>
<tr>
<td>Lighavan</td>
<td>Synoptic</td>
</tr>
<tr>
<td>Synoptic</td>
<td>$T_{\text{mean}}$ (°C)</td>
</tr>
<tr>
<td>Tabriz</td>
<td>Synoptic</td>
</tr>
<tr>
<td>Synoptic</td>
<td>$T_{\text{min}}$ (°C)</td>
</tr>
<tr>
<td>Synoptic</td>
<td>$T_{\text{mean}}$ (°C)</td>
</tr>
<tr>
<td>Synoptic</td>
<td>$T_{\max}$ (°C)</td>
</tr>
<tr>
<td>Sahand</td>
<td>Synoptic</td>
</tr>
<tr>
<td>Synoptic</td>
<td>$T_{\text{min}}$ (°C)</td>
</tr>
<tr>
<td>Synoptic</td>
<td>$T_{\text{mean}}$ (°C)</td>
</tr>
<tr>
<td>Synoptic</td>
<td>$T_{\max}$ (°C)</td>
</tr>
<tr>
<td>Ligavan</td>
<td>Hydrometric</td>
</tr>
<tr>
<td>Hydrometric</td>
<td>SSL (ton·day$^{-1}$)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Characteristics of the RCP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenario</td>
<td>Descriptions</td>
</tr>
<tr>
<td>RCP2.6</td>
<td>Before 2100, the maximum radiation will be 3 W·m$^{-2}$, which will reach 2.6 W·m$^{-2}$ by 2100. The annual temperature rises to 1.5 °C.</td>
</tr>
<tr>
<td>RCP4.5</td>
<td>It has more stability with less oscillations. The amount of radiation in 2100 will be 4.5 W·m$^{-2}$. The annual temperature rises to 2.4 °C.</td>
</tr>
<tr>
<td>RCP8.5</td>
<td>In 2100, radiation is about 8.5 W·m$^{-2}$. Also, the annual temperature will rise by 4.9 °C.</td>
</tr>
</tbody>
</table>
conditions. In the proposed strategy, the hybrid of LS-SVM and FPA, LS-SVM, and ANFIS is used for modeling and predicting runoff and SSL, and the delta change factor method is used for the downscaling large-scale precipitation and temperature of the ensemble GCM. In this strategy, there are two sections, including modeling and predicting (Figure 2(a)). In the modeling section, the observed precipitation, temperature, runoff, and SSL data are standardized, lagged with different lagging times, and divided into training and testing periods. The standardized relation is as follows:

\[
X_{\text{new}} = \frac{X - X_{\text{mean}}}{\text{std}(X)}
\]

where \(X_{\text{new}}\) is standardized data, \(X\) is original data, \(X_{\text{mean}}\) is the mean of data in the training period, and std(\(X\)) is the standard deviation of data in the training period. Afterward, observed precipitation and observed temperature data in different stations are used as input data, and observed runoff data are used as target data for rainfall–runoff modeling by LS-SVM-FPA, LS-SVM, and ANFIS in the observed period. After rainfall–runoff modeling, the best algorithm is selected (the best algorithm is selected based on the evaluation criteria), and modeled runoff data from this algorithm are used as input data and observed SSL data used as target data for runoff–SSL modeling by LS-SVM-FPA, LS-SVM, and ANFIS in the observed period. Then, the best runoff–SSL model is chosen according to evaluation criteria. In predicting section, the large-scale precipitation and temperature data of the ensemble GCM are standardized based on training period and downscaled to a local scale (precipitation and temperature data in the modeling section) by the delta change factor method in future periods. Then, the downscaled precipitation and temperature data for future periods are applied to the best rainfall–runoff model, and runoff data are predicted for future periods. Afterward, the predicted runoff data are used as inputs of the best runoff–SSL model, and SSL data are predicted for future periods. The scheme of the proposed method for predicting SSL is shown in Figure 2(b).

### Least-square support-vector machine

LS-SVM is presented by (Suykens et al. 2002) for solving complex classification and regression problems. The LS-SVM is an improved method of standard SVM. LS-SVM considers the relationship between inputs and outputs as follows:

\[
Y(x) = W_t \phi(x) + b
\]

where \(Y(x)\) is the model output, \(x\) is the input vector, \(\phi(x)\) is a nonlinear function, \(W_t\) is the weight of the input vector, and \(b\) is the bias of the model. In this relation, the \(W_t\) and \(b\) are unknown. In LS-SVM, to find these two parameters the following objective function is solved (Adnan et al. 2020):

\[
\text{Minimize: } (1/2)W^TW + (1/2)C \sum_{t=1}^{n} e_t
\]

Subject to: \(Y(x) = W_t \phi(x) + b + e_t\)

where \(e_t\) is the error parameter, and \(C\) is the penalty coefficient. This objective function can be solved by the optimization algorithms. To handle constraints in a defined objective function, the Lagrangian method is used. By using the Lagrangian method, the final objective function is given as follows:

\[
\text{Minimize: } (1/2)W^TW + (1/2)C \sum_{t=1}^{n} e_t - \sum_{t=1}^{n} \alpha_1 [W_t \phi(x) + b + e_t - O_t]
\]

where \(\alpha_1\) is known as Lagrangian coefficients, and \(O_t\) is the observation output. By taking partial derivation from relation (4) to \(W, b, e,\) and \(\alpha\) (Deo et al. 2017):

\[
\begin{align*}
\sum_{t=1}^{n} \alpha_t Y(x_t) & = 0 \\
\sum_{t=1}^{n} \alpha_t & = 0 \\
\alpha_t & = C e_t \\
W^T Y(x_t) + b + e_t - O_t & = 0
\end{align*}
\]
Figure 2 | The proposed strategy for predicting SSL: (a) flowchart and (b) scheme.
In SVM, to solve such objective functions, the quadratic optimization method is used. However, unlike the SVM method, LS-SVM uses the least-square optimization method for solving this objective function. It is worth mentioning that the least-square optimization method has lower computational cost and more efficiency than the quadratic optimization method. To solve relations (4) and (5) by the least-square optimization method, these relations must be rewritten as follows:

\[
\begin{bmatrix}
0 \\
1 
\end{bmatrix}^T \begin{bmatrix}
1 & 1 \\
\text{Kernel} + C^{-1}I 
\end{bmatrix} \begin{bmatrix}
\alpha \\
b
\end{bmatrix} = \begin{bmatrix}
y_1 \\
y 
\end{bmatrix}
\]

where

\[
y = \begin{bmatrix}
y_1 \\
\vdots \\
y_n
\end{bmatrix}, \quad \alpha = \begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix}, \quad 1 = \begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix}, \quad \text{and } I = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 1
\end{bmatrix}
\]

Thus, Kernel is known as the kernel function. After solving relation (6), the final relation between inputs and outputs is rewritten as follows (Farzin et al. 2020):

\[
Y(x) = \alpha_t \text{Kernel} (x(t), x) + b
\]

The kernel function is a nonlinear function with different types, including linear, polynomial, and radial base kernel functions. According to the study conducted by Ghosh (2010), the radial kernel function is more accurate than other mentioned kernel functions. Therefore, the radial base kernel function is used as follows:

\[
\text{Kernel}(x_i, x_j) = \exp \left( -\frac{\|x_i - x_j\|^2}{2\sigma^2} \right)
\]

Here, \(\sigma\) represents the width of the kernel function. In the LS-SVM algorithm, a model can be developed only by determining two parameters of \(C\) and \(\sigma\). These two parameters have a significant impact on the final results of LS-SVM. However, there is no specific method for determining these parameters. Figure 3 shows the schematic of LS-SVM.

**Flower pollination algorithm**

The FPA is proposed by Yang (2012) to solve various engineering problems. This algorithm is inspired by the process of pollination and the reproduction of flowering plants. In the FPA algorithm, it is assumed that each plant has only one flower and each flower produces only one pollen. Also, each flower or flower pollen is equivalent to a search agent. In the natural flower pollination process, there are two forms of pollination, including cross-pollination (by biological agents) and self-pollination (by wind). In the FPA, the cross-pollination process is considered as the global optimization process. Since the global optimization process is performed throughout the problem search space, the Levy flight method (which results in larger displacements) is used to simulate this motion. Global search on FPA algorithm is performed by the following relation:

\[
X_{i+1}^t = X_i^t + L(X_i^t - X^*)
\]

where \(X_{i+1}^t, X_i^t, X^*\), and \(L\) represent the new position, the current position, the best position, and the parameter corresponding to Levy flight, respectively.

Self-pollination in the FPA algorithm is also considered as a local search. The following relation is used to accomplish this action:

\[
X_{i+1}^t = X_i^t + \epsilon(X_i^t - X^*)
\]

where \(\epsilon\) is a random number with a uniform distribution between 0 and 1. Therefore, relation (6) leads to smaller motions, and it is considered as a local search. There is also a specific probability of \(p\) to down one of each global
or local search. Hence, the positions of search agents in FPA are updated as follows:

\[
\begin{align*}
X_{t+1}^i &= X_t^i + L(X_t^i - X^*) \quad \text{if } \text{rand} > p \\
X_{t+1}^i &= X_t^i + \varepsilon(X_t^i - X^*) \quad \text{if } \text{rand} < p
\end{align*}
\]

where \( p \) is the probability of change (between global and local search). Figure 4 demonstrates the pseudo code of FPA. In the present study, the parameters of the maximum number of iterations, population size, and the probability of change were considered equal to 500, 20, and 0.8 by trial and error, respectively. See Yang (2012) for more information.

### The hybrid of LS-SVM and FPA

As mentioned, the LS-SVM has two parameters of \( C \) and \( \sigma \), while there is no specific method for determining these parameters. These two parameters have a high impact on the final accuracy of LS-SVM. Therefore, in the present study, FPA has been used to optimally determine these parameters. (1) For this purpose, first, the parameters of \( C \) and \( \sigma \) are randomly assigned to each of the FPA search agents. (2) Then,

---

**Determine FPA parameters: the maximum number of iteration, population size, and probability of change (P)**

**Generate initial population**

**Calculate fitness function for each search agent**

**Chose best solution according to fitness function of search agents**

**For** it=1:maxium number of iteration

**For** i=1:population size

**Generate random number (rand)**

**If** rand>P

**Update position of i\(^{th}\) search agent based on the global search (relation 8)**

**Else**

**Update position of i\(^{th}\) search agent based on the local search (relation 9)**

**End**

**Calculate the fitness function for i\(^{th}\) search agent**

**Update the current best solution**

**End**

**End**

**Return best solution**

---

**Figure 4 |** The pseudo code of FPA.
LS-SVM for each search agent \((C, \sigma)\) is executed. (3) Afterward, the fitness function \((R^2)\) is calculated for each search agent, and the best position or search agent is determined with the more fitness function. The fitness function is calculated as follows:

\[
\text{Max: fitness}(C, \sigma) = R^2_{\text{Test}}(Y, O) \tag{12}
\]

The position of the search agents will evolve based on relationships (10) and (11), and steps 2 to 4 will be repeated until the termination condition is reached. Finally, the optimal solution problem \((C, \sigma)\) is returned. Figure 5 demonstrates the schematic of the LS-SVM and FPA hybrids.

**Adaptive neuro-fuzzy inference system**

The ANFIS algorithm was first developed (Jang 1993) by combining the fuzzy method and ANN. ANN uses the learning ability of ANN and the inference of the fuzzy method to detect nonlinear relationships between inputs and outputs (Azad et al. 2019). There are three methods, including grid partition, subtractive clustering, and fuzzy c-mean clustering. However, according to studies conducted by Azad et al. (2019), the fuzzy c-mean clustering method is more accurate. Therefore, in the present study, the aforementioned method is used to create a fuzzy model in ANFIS. For more information about ANFIS, see Azad et al. (2018) and Salimi et al. (2020).

**Downscaling of precipitation and temperature**

One of the most common, but simple methods to downscale the outputs of GCMs, is the delta change factor method, which has been used in many studies such as Ehteram et al. (2018b). In this method, the monthly average data are used instead of using direct outputs of GCMs. Also, the changes in temperature and precipitation under climate change are calculated in this method as follows:

\[
T_f = T_{\text{obs},t} + (T_{\text{GCM,fut},t} - T_{\text{GCM,base},t}) \tag{13}
\]

\[
P_f = P_{\text{obs},t} \times \left( \frac{P_{\text{GCM,fut},t}}{P_{\text{GCM,base},t}} \right) \tag{14}
\]

In which, \(T_f\) is the future temperature, \(P_f\) is the future precipitation, \(T_{\text{obs},t}\) is the observed temperature, \(P_{\text{obs},t}\) is the
observation precipitation, $T_{GCM,\ fut}$ is the monthly large-scale average of predicted temperature, $P_{GCM,\ fut}$ is the monthly average of large-scale precipitation projections, $T_{GCM,\ bas}$ is the monthly average of large-scale temperature in the base period, $P_{GCM,\ bas}$ is the monthly average of large-scale precipitation in the base period, and $t$ indicates the $t$th month. For more information, refer to Singh & Saravanan (2020).

### Evaluation criteria

In the present study, the accuracy of the runoff and SSL modeling results are evaluated using evaluation criteria, including determination coefficient ($R^2$), relative root-mean-square error (RRMSE), and Nash-Sutcliffe efficiency (NSE) coefficient, which are presented as follows:

\[
R^2 = \left[ \frac{\sum_{i=1}^{N} (O_i - \bar{O})(Y_i - \bar{Y})}{\sum_{i=1}^{N} (O_i - \bar{O})^2 \sum_{i=1}^{N} (Y_i - \bar{Y})^2} \right]^2 \quad (15)
\]

\[
RRMSE = \sqrt{\frac{\sum_{i=1}^{N} (Y_i - O_i)^2}{N \cdot Std(O)}} \quad (16)
\]

\[
NSE = 1 - \frac{\sum_{i=1}^{N} (O_i - Y_i)^2}{\sum_{i=1}^{N} (O_i - \bar{O})^2} \quad (17)
\]

where $O$, $O$, $Y$, $\bar{Y}$, $Std(O)$, and $N$ represent observational data, mean observational data, output, mean outputs, the standard deviation of observational data, and the number of data, respectively. The Taylor diagram is one of the graphical methods to determine the skill of algorithms (Taylor 2001). This method summarizes the criteria for correlation coefficient $R$, root-mean-square deviation (RMSD), and the standard deviation (SD) in a diagram. The RMSD and $R$ values are also obtained from the following relations:

\[
\text{RMSD} = \sqrt{\frac{\sum_{i=1}^{N} (Y_i - O_i)^2}{N}} \quad (18)
\]

**RESULTS AND DISCUSSION**

### Runoff and SSL modeling

In this study, the runoff is modeled using the data of precipitation and mean temperature at the Lighvan station, precipitation, mean temperature, minimum temperature, maximum temperature, and solar radiation at Tabriz stations, as well as the mean temperature, minimum temperature, and maximum temperature at the Sahand station. Since each combination of inputs can have a different effect on the accuracy of the results, 11 combinations with 0–10-month period lag time are defined for the inputs of the investigated algorithms. Table 3 indicates the results of runoff modeling based on the different evaluation criteria and using three algorithms, including LS-SVM-FPA, LS-SVM, and ANFIS. According to the results of this table, the best results for LS-SVM-FPA algorithms are related to the combination 10 in the training and testing period (bold and underline values). Also, the best results in the training and testing period for the LS-SVM and ANFIS algorithms are related to combination 3 (bold values). Besides, the LS-SVM-FPA algorithm and combination 10 have the most accurate results in both training and testing periods compared to the other investigated algorithms and combinations. Also, the LS-SVM algorithm is ranked second place in terms of accurate evaluation criteria. The values of $R^2$, RRMSE, and NSE for the LS-SVM-FPA algorithm and combination 10 (the best combination for LS-SVM-FPA) are, respectively, 6.10, 10.64, and 6.49%, which are more accurate than LS-SVM and combination 3 (the best combination for LS-SVM). Furthermore, based on the mentioned parameters, the superiority of LS-SVM-FPA over ANFIS is equal to 38.10, 33.33, and 38.98%, respectively.
Generally, the accuracy of the algorithms is decreased by increasing the number of inputs and the complexity of the model. The accuracy of LS-SVM and ANFIS algorithms is also decreased by increasing the number of inputs from combination 3 onwards. However, the accuracy of the LS-SVM-FPA algorithm is increased by increasing the number of inputs from combination 1 to combination 10. Also, in all the investigated combinations, the accuracy of the LS-SVM-FPA algorithm is more than that of LS-SVM and ANFIS algorithms. Therefore, combining the FPA and LS-SVM leads to the increased accuracy of the runoff modeling, even with more input variables (more complexity). In this study, the runoff and SSL data of the Lighvan hydrometric station were used for runoff–SSL modeling. For this purpose, 11 runoff input combinations with time delays of 0–10 months are defined. Also, the results of the three algorithms of LS-SVM-FPA, LS-SVM, and ANFIS in runoff–SSL modeling with 11 different input combinations are compared based on the evaluation criteria. According to the evaluation criteria demonstrated in Table 4, the LS-SVM-FPA algorithm in combination 2 has better accuracy than other algorithms and combinations (bold and underline values). After that LS-SVM-FPA in combination 2, LS-SVM in combination 1 and ANFIS in combination 9 have the second and third rank in terms of accuracy (bold values). There is a significant difference between the results of LS-SVM-FPA as the best algorithm and ANFIS as the worst algorithm. The coefficients of $R^2$, RRMSE, and NSE for LS-SVM-FPA (in the best input combination) are, respectively, 73, 22, and 47%, which are more accurate than that of ANFIS (in the best input combination).

In the following, the Taylor diagram is used to graphically demonstrate the ability of algorithms to model the runoff and SSL in training and testing periods (Figure 6). The superiority of the LS-SVM-FPA over the other two methods is well indicated in this diagram. However, ANFIS is better than the LS-SVM in the training period, and LS-SVM is more similar to the observational data in the testing period. Nevertheless, it can be concluded that the LS-SVM algorithm is more accurate than ANFIS because in the testing period, the evaluation based on the new data is the decision criterion for selecting the best algorithm. In the case of SSL, Taylor diagrams in training and testing periods show the superiority of LS-SVM-FPA over
LS-SVM and ANFIS algorithms. Also, it is indicated that the LS-SVM algorithm is more accurate than the ANFIS.

The superiority of the LS-SVM-FPA results over the LS-SVM indicates the successful performance of the FPA algorithm in optimizing the LS-SVM parameters. Because FPA finds the optimal parameter of LS-SVM in such a way that modeling is done with less error and more accuracy. While in standalone LS-SVM, the parameters of LSSVM are determined by trial and error that leads to less accuracy than the hybrid of LS-SVM and FPA. This issue is observed in Tables 3 and 4 and Figure 6. The superiority of hybrid LS-SVM and optimization algorithms over standalone LS-SVM is confirmed by the results of different studies such as Anaraki et al. (2021). Thus, as mentioned earlier, LS-SVM-FPA and LS-SVM have more accurate results than ANFIS, due to the weakness of ANFIS in establishing a relationship between inputs and outputs in a complex system (Najafzadeh & Ghaemi 2019). While LS-SVM-FPA and LS-SVM by using kernel function are considered the complex and nonlinear relationship of inputs and inputs as one linear relationship, the better results of LS-SVM-based algorithms than ANFIS are observed in many studies such as Wang et al. (2017) and Farzin et al. (2020).

Figure 7 indicates the time series and distribution of observed runoff and the runoff modeled by the LS-SVM-FPA algorithm and the best input combination (combination 10). According to this figure, in the training and testing period, the modeled runoff is in good agreement with observational runoff. Therefore, the RRMSE (0.42), NSE (0.82), and $R^2$ (0.87) criteria for this algorithm are in a very good range (based on the study conducted by Moriasi et al. (2007)). Besides, according to the time series of observed and modeled runoff in Figure 7, the runoff peak data are well estimated by the LS-SVM-FPA algorithm (Moriasi et al. 2007). Figure 8 shows the time series and distribution of observed SSL and the SSL modeled by the best algorithm (LS-SVM-FPA in combination 2). As can be seen in this figure, there is only a slight difference between the time series of observed and modeled SSL. During the testing period, the values of RRMSE (0.58), NSE (0.65), and $R^2$ (0.85) are in the good range. Based on this figure, there is a linear relationship with a 45$^\circ$ slope between the observed and modeled runoff.
distribution in the two training and testing periods. According to the study implemented by Nourani et al. (2019), there is a complex and nonlinear relationship between the sediment production in a basin and the basin runoff. On the other hand, there is a significant correlation between the sediment of the basin and the amount of SSL in the previous months. However, the present study is conducted for long-term prediction under climate change conditions. Therefore, SSL data from the previous months have not been used as the inputs for SSL modeling. Also, the dispersion parameters of SSL data, such as standard deviation and range are significantly more than runoff data. These issues lead to less accuracy of SSL modeling than runoff modeling in Figures 7 and 8.

**Downscaling precipitation and temperature**

Precipitation and temperature are the most important components in the hydrology cycle. The changes in these components are the most important factors in river regime changes and, consequently, sediment production changes in the basin. Precipitation and temperature are predicted by large-scale GCMs. Therefore, in the present study, the downscaling of precipitation and temperature has been investigated using the delta change factor method to investigate the effect of climate change on river SSL more precisely. The delta change factor method converts large-scale precipitation and temperature into local-scale precipitation and temperature.

---

Figure 6 | Taylor diagram for (a) runoff in the training period, (b) runoff in the testing period, (c) SSL in the training period, and (d) SSL in the testing period.
In Table 5, effect of climate change on precipitation and mean, maximum, and minimum temperature at three stations of Lighvan, Tabriz, and Sahand is investigated, using the comparison of values predicted in four future periods (which include 2020–2034, 2035–2049, 2070–2084, and 2085–2099) with those in the observation period (1990–2004). The bold values in Table 5 show the maximum or minimum change in station and bold and underline values show maximum or minimum change in all stations. The comparison of the results shows that the maximum amount of precipitation changes (31.24%) is related to the Lighvan station in the RCP8.5 scenario and 2020–2034 period, and the minimum amount of precipitation changes (−14.18%) is related to Sahand station in the scenario RCP8.5 and 2085–2099 period. In terms of the average temperature, the Lighvan station under the RCP8.5 scenario and in the 2085–2099 period has the highest change (18.92%), and the Tabriz station under two scenarios of RCP2.6 and RCP4.5 has the lowest change (1.93%) in 2020–2034. The comparison of the predicted maximum temperature values, regarding the observation period, shows that most changes (8.96%) are for the Sahand station under the RCP8.5 scenario in 2085–2099 period, and the least changes (1.53%) are for the Tabriz station under the scenarios of RCP2.6, RCP4.5, and RCP8.5 in 2020–2034 period and under RCP2.6 scenario in periods of 2035–2049, 2070–2084, and 2085–2099. According to the minimum temperature, most changes (2.19%) are for the Tabriz station under RCP2.6, RCP4.5, and RCP8.5 scenarios in periods of 2020–2034, RCP4.5 scenario in the 2035–2049 period, and RCP2.6 in periods of 2070–2084 and 2085–2099 (Table 5). As seen, the temperature parameters in all stations for all scenarios and future periods increase. Also, the precipitation is decreased in the Lighvan station. These changes in metrological variables can lead to decrease runoff and production SSL. Thus, the downscaled temperatures are close to each other under RCP scenarios of ensemble GCM that have lower uncertainty than other GCMs due to averaging from outputs of 24 GCMs.

![Figure 7](image-url)

*Figure 7* | Results of runoff modeling with LS-SVM-FPA: (a) times series and (b) scatter.
Prediction of runoff and SSL

After the downscaling of the precipitation and temperature, which are the most important factors in determining runoff and SSL, the prediction of runoff and SSL changes in future periods has been performed under three scenarios of RCP2.6, RCP4.5, and RCP8.5. The comparison of the predicted runoff changes with the observed runoff indicates that the highest amount of runoff changes (5.19%) is related to the RCP2.6 scenario in the 2035–2049 period, and the lowest runoff changes (−36.17%) is related to the RCP8.5 scenario in the 2084–2099 period. In the case of SSL, the largest changes (5.44%) are predicted by the RCP8.5 scenario for the 2020–2034 period. However, the least amount of SSL changes (−27.08%) is related to the RCP2.6 scenario in the 2070–2084 period (Table 6). According to this table, runoff in all periods and scenarios decreases except for the first two periods under the RCP2.6 scenario. Thus, the SSL is also decreased in all periods and scenarios except for the first two periods under RCP4.5, and RCP8.5 scenarios. These increases in runoff and SSL may be due to an increase in snowmelt by increasing temperature and occurring extreme events such as heavy precipitation. However, by shifting from first future periods to last future periods, increasing temperature and decreasing the precipitation, and changing the pattern of precipitation from snow to rain, the snow accumulation in the highest of the Lighvan basin is decreased, and consequently, the runoff and SSL are decreased especially in last two periods.

The results of the monthly average of observed runoff show that the maximum monthly average runoff was observed at the Lighvan hydrometric station in spring. This pattern is also seen in the scenarios of RCP2.6, RCP4.5, and RCP8.5. However, the predicted monthly average runoff in the spring for all three scenarios in the four investigated periods is lower than or close to the similar values in the observation period. Also, the magnitude of this runoff reduction for all three scenarios in the two periods of 2020–2034 and 2035–2049 is less than in the other two periods (Figure 9).
Table 5: Downscaling of precipitation and temperature

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Lighavan</td>
<td>P (mm)</td>
<td>0.85</td>
<td>1.08</td>
<td>1.11</td>
<td>1.11</td>
<td>1.08</td>
</tr>
<tr>
<td>Change (%)</td>
<td>0</td>
<td>27.63</td>
<td>30.36</td>
<td>31.24</td>
<td>27.75</td>
<td>25.02</td>
</tr>
<tr>
<td>Tabriz</td>
<td>P (mm)</td>
<td>0.19</td>
<td>0.25</td>
<td>0.24</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Change (%)</td>
<td>0</td>
<td>29.88</td>
<td>30.64</td>
<td>31.12</td>
<td>25.20</td>
<td>24.97</td>
</tr>
<tr>
<td>Change (%)</td>
<td>0</td>
<td>18.64</td>
<td>18.64</td>
<td>18.65</td>
<td>18.66</td>
<td>18.73</td>
</tr>
<tr>
<td></td>
<td>Tmax (°C)</td>
<td>18.67</td>
<td>19.16</td>
<td>19.16</td>
<td>19.16</td>
<td>19.16</td>
</tr>
<tr>
<td>Change (%)</td>
<td>0</td>
<td>1.53</td>
<td>1.53</td>
<td>1.53</td>
<td>1.53</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td>Tmin (°C)</td>
<td>7.76</td>
<td>8.25</td>
<td>8.25</td>
<td>8.25</td>
<td>8.25</td>
</tr>
<tr>
<td>Change (%)</td>
<td>0</td>
<td>2.19</td>
<td>2.19</td>
<td>2.19</td>
<td>2.19</td>
<td>2.20</td>
</tr>
</tbody>
</table>
Therefore, it can be concluded that the runoff has a decreasing trend. It is also observed that the amount of runoff decreased by the increases in precipitation, which can be due to the increase in mean, minimum, and maximum temperature and, accordingly, evaporation. Also, in the winter and autumn seasons, when the temperature is low as a result of evaporation, the increase in precipitation has increased the predicted runoff for the two near future periods compared to the observation period. In the next two periods, the runoff in the two seasons of winter and autumn is less than the observed runoff (Figure 8).

Figure 10 illustrates the 10-year moving averages of the predicted runoff time series for the periods of 2020–2049 and 2070–2099 and under the three scenarios of RCP2.6, RCP4.5, and RCP8.5. Based on this figure, it is obvious that the predicted runoff has a descending trend, despite the fluctuation across all the periods. This is in accordance with Figure 9. Thus, this descending trend in the runoff predicted from 2020 to 2049 period and under the RCP8.5 scenario has the greater slope (−0.0046). Also, the amount

<table>
<thead>
<tr>
<th>Table 6</th>
<th>Results of prediction runoff and SSL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scenarios</td>
<td>Q (m³·s⁻¹)</td>
</tr>
<tr>
<td>RCP2.6</td>
<td>4.50</td>
</tr>
<tr>
<td>RCP4.5</td>
<td>−12.37</td>
</tr>
<tr>
<td>RCP8.5</td>
<td>−12.67</td>
</tr>
<tr>
<td>Scenarios</td>
<td>SSL (ton·day⁻¹)</td>
</tr>
<tr>
<td>RCP2.6</td>
<td>−11.95</td>
</tr>
<tr>
<td>RCP4.5</td>
<td>4.86</td>
</tr>
<tr>
<td>RCP8.5</td>
<td>5.44</td>
</tr>
</tbody>
</table>

Therefore, it can be concluded that the runoff has a decreasing trend. It is also observed that the amount of runoff decreased by the increases in precipitation, which can be due to the increase in mean, minimum, and maximum temperature and, accordingly, evaporation. Also, in the winter and autumn seasons, when the temperature is low as a result of evaporation, the increase in precipitation has increased the predicted runoff for the two near future periods compared to the observation period. In the next two periods, the runoff in the two seasons of winter and autumn is less than the observed runoff (Figure 8).

Figure 9 illustrates the monthly average of observed and predicted runoff: (a) RCP2.6, (b) RCP4.5, and (c) RCP8.5.

![Figure 9](http://iwaponline.com/jwcc/article-pdf/doi/10.2166/wcc.2021.317/868384/jwc2021317.pdf)
of predicted runoff in the 2070–2099 period is less than the runoff in the observation period. As mentioned before, this descending trend in runoff is due to the increase in predicted temperature. Also, by more attention to Figure 9, it can be seen that the moving average of annual runoff oscillates. This oscillation may be for the decrease in precipitation in all three scenarios. Thus, increasing temperature parameters, such as minimum, mean, and maximum temperature, can lead to more snowmelt, change in the pattern of precipitation from snow to rain, and increase in extreme events such as heavy precipitation.

Figure 11 shows the monthly mean SSL in the observation and future periods. According to this figure, the predicted SSL is less than the SSL observed in all months of the year, except for May, June, and July. This is in full accordance with Figure 9 because, in this figure, the amount of future runoff is lower than the forecasted runoff.

However, the amount of SSL in May, June, and July is increased since the SSL is modeled based on the runoff with 0- and 1-month time delays, and the SSL in the mentioned months has been increased by the increase of runoff in the previous months, and more runoff leads to the more production of SSL. As can also be seen in Figure 11, SSL is highly sensitive to runoff changes, such that the SSL is significantly changed by small changes in the runoff of a month or the runoff of the previous months. This issue can be attributed to different reasons such as the type of soil in the upstream of the Lighvan station and the bed material of the river. Thus, the results of all scenarios in Figure 11 are close to each other. The reason for these
results could be due to the close results of downscaled temperature parameters in Table 6.

Figure 12 demonstrates the 10-year moving average of the SSL predicted in future periods of 2020–2049 and 2070–2099 and under the scenarios of RCP2.6, RCP4.5, and RCP8.5. Regarding this figure, the amount of SSL has a descending trend in the future periods and under all three investigated scenarios. The slope change of the SSL is also lower in the 2070–2099 period, in comparison with the 2020–2049 period. This can be due to the reduction of runoff. The comparison of the investigated scenarios shows that the slope of runoff changes is more under the RCP8.5 scenario (0.0478 for 2020–2049 and 0.02525 for 2070–2099) and is less under the RCP2.6 scenario (0.0367 for 2020–2049 and 0.0562 for the 2070–2099 periods). By comparing the 10 years moving average of SSL and runoff, it can be concluded that both of them follow the same trend. Such that in the period of 2020–2050, both Figures 9 and 11 will have a peak in 2038 (0.94 m³·s⁻¹ and 14.5 ton·day⁻¹ for runoff and SSL, respectively). Accordingly, in all three scenarios, the SSL in the years 2036–2039 is more than the observed SSL, which could cause damage to the hydraulic structures. Thus, the slope change of SSL is more than runoff in all scenarios and periods, which is due to more sensitivity of SSL to changes of runoff.

CONCLUSION

In the present study, the hybrid of LS-SVM algorithm and FPA (LS-SVM-FPA) were used to model and predict the runoff and SSL. The results were also compared using LS-
SVM and ANFIS algorithms. The comparison of the results based on evaluation criteria and the Taylor diagram showed that LS-SVM-FPA has superiority in modeling both runoff and SSL parameters so that the NSE coefficient was increased up to 47% by the LS-SVM-FPA algorithm in SSL modeling. Also, the results of runoff and SSL modeling by the LS-SVM-FPA algorithm based on $R^2$, RRMSE, and NSE criteria were in a good range. These values for runoff modeling were equal to 0.87, 0.42, and 0.82, respectively. Also, in SSL modeling, the values of $R^2$, RRMSE, and NSE criteria were equal to 0.85, 0.58, and 0.65, respectively.

The results of runoff and SSL prediction, using ensemble GCM for future periods (2020–2034, 2035–2049, 2070–2084, and 2085–2099) and under three scenarios of RCP2.6, RCP4.5, and RCP8.5, indicate reduction runoff up to 36.17% in all periods and scenarios, except for the first two periods under the RCP2.6 scenario. In the first two periods including 2020–2034 and 2035–2049 under RCP2.6, the runoff increased 4.5 and 5.19% to the base period. Also, the runoff changes were more frequent in the periods of 2070–2084 and 2085–2099, in comparison with the other studied periods. In the case of SSL, the predicted results also showed a decrease in SSL (up to 25.36%) compared to the observation period, except for the first two periods under RCP4.5 and RCP8.5. In these periods, SSL increased up to 5.44%. Furthermore, runoff and SSL had a descending trend in all periods and scenarios. Nevertheless, between 2036 and 2039, the amounts of runoff and SSL were greater than the observed values, which may cause problems for downstream structures in the Lighvan Basin. The maximum value of runoff and SSL in this period was about 0.94 m$^3$·s$^{-1}$ and 14.5 ton·day$^{-1}$, respectively. For future studies, it is suggested that some structures in the
downstream of the Lighvan station can be designed by considering these values of runoff and SSL.

CONFLICT OF INTEREST

None declared.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES


Deo, R. C., Kisi, O. & Singh, V. P. 2017 Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and m5tree model. Atmospheric Research 184, 149–175.


