A Novel LSSVM Model Integrated with GBO Algorithm to Assessment of Water Quality Parameters

Mojtaba Kadkhodazadeh · Saeed Farzin

Received: 27 April 2021 / Accepted: 20 July 2021 / Published online: 25 September 2021
© The Author(s), under exclusive licence to Springer Nature B.V. 2021

Abstract
In this study, a novel least square support vector machine (LSSVM) model integrated with gradient-based optimizer (GBO) algorithm is introduced for the assessment of water quality (WQ) parameters. For this purpose, three stations, including Ahvaz, Armand, and Gotvand in the Karun river basin, have been selected to model electrical conductivity (EC) and total dissolved solids (TDS). First, to prove the superiority of the LSSVM-GBO algorithm, the performance is evaluated with three benchmark datasets (Housing, LVST, Servo). Then, the results of the new hybrid algorithm were compared with those of artificial neural network (ANN), adaptive neuro-fuzzy interface system (ANFIS), and LSSVM algorithms. Input combination for assessment of WQ parameters EC and TDS consists of \( \text{Ca}^{+2}, \text{Cl}^{-1}, \text{Mg}^{+2}, \text{Na}^{+1}, \text{SO}_4, \text{HCO}_3, \) sodium absorption ratio (SAR), sum cation (Sum.C), sum anion (Sum.A), pH, and Q. The modeling results based on evaluation criteria showed the significant performance of LSSVM-GBO among all benchmark datasets and algorithms. Other results showed that in Ahvaz station, Sum.C, Sum.A, and Na\(^{+1}\) parameters, and in Armand and Gotvand stations, Sum.C, Sum.A, and Cl\(^{-1}\) parameters have the greatest impact on modeling EC and TDS parameters. Then, EC and TDS modeling was performed based on the best input combination and the best algorithm in different time delays. The highest accuracy of modeling EC and TDS parameters in Gotvand station was and C1 time delay.

Keywords Novel hybrid model · LSSVM · GBO · Water quality parameters · Benchmark datasets · Karun river

1 Introduction
Pollution rate and river WQ is an important issue that is very effective in human health. Rivers carrying water and nutrients are necessary and provide important resources for drinking, industrial, aquatic, recreational and agricultural consumption (Ho et al. 2019).
Therefore, they require at least an acceptable level of WQ. In recent years due to the rapid population growth and urban extension, agriculture, economic development, and increasing industrial production, pollution has increased in rivers (Busico et al. 2020); therefore, the qualitative study of water resources is one of the most critical subjects in most regions.

One of the best ways to study the problems of water pollution is modeling and analysis of WQ using modern methods such as artificial intelligence (AI). In recent years, many studies have been conducted about EC, and TDS modeling in different regions using data mining methods as these methods have many accuracies and, like physical and mathematical models, they do not need to specify multiple parameters and reduce the cost of research work. Modeling TDS and EC concentration and predicting it is essential for pollution control and water resource management (Azad et al. 2019).

In recent years, the AI models have been widely employed for WQ issues such as TDS, water quality index (WQI), dissolved oxygen (DO), chemical oxygen demand (COD), biochemical oxygen demand (BOD), EC, SAR, total hardness (TH), ammoniacal nitrogen (AN), suspended solid (SS) and pH (Emamgholizadeh et al. 2014; Tiyasha et al. 2020; Baghapour et al. 2020; Vijay and Kamaraj 2021). The common examples of these AI models include: ANN, ANFIS, support vector machine (SVM), group method of data handling (GMDH), and genetic programming (GP) (Barzegar et al. 2016; Salami et al. 2016; Haghiabi et al. 2018; Aryafar et al. 2019). The hybrid models higher accuracy than individual models, and recently, it has been widely used in assessment WQ parameters (Khosravi et al. 2018). Azad et al. (2018) and Kisi et al. (2019) reported that hybrid models ANFIS with differential evolution (DE), genetic algorithm (GA), ant colony optimization for continuous domains (ACO_R), compact genetic algorithm (CGA), and particle swarm optimization (PSO) performed well in modeling EC, SAR, and TH. In another study, Najafzadeh et al. (2019) used gene express programming (GEP), model tree (MT), and evolutionary polynomial regression (EPR) to estimate BOD, DO, and COD. The results of the models showed the relative superiority of the EPR. Deterministic and numerical models have been applied extensively to model WQ; the counting convolutional neural network (CCNN) model has high accuracy for estimating the DO concentration (Zounemat-Kermani et al. 2019). Li et al. (2019) proposed a new model combining recurrent neural network (RNN) with improved Dempster/Shafer (D-S) evidence theory (RNNs-DS) for the prediction of WQ. Results indicated that the new model has better performance. Lu and Ma (2020) offered two new hybrid models which combine complete ensemble empirical mode decomposition with adaptive noise (CEEMDAN) with an extreme gradient boosting (XGBoost) and random forest (RF) for WQ parameters prediction. Results showed the superiority of the CEEMDAN-XGBoost. Banadkooki et al. (2020) used ANN, ANFIS, SVM to predict the TDS. Also, optimization algorithms were used to train these models. In this study, the ANFIS-moth flam optimization (MFO) and ANFIS-cat swarm optimization (CSO) models showed good performance. Melesse et al. (2020) applied two individual M5 prime (M5P) and random forest (RF) and eight novel hybrid algorithms to predict EC. The results showed that hybrid algorithms have minor errors.

This study, for the first time, proposed a novel hybrid model by combining the LSSVM model and GBO algorithm to achieve accurate estimation of EC and TDS at three hydrometric stations with different kinds of climate and WQ in the Karun river area (as a case study). For prove the superiority of the LSSVM-GBO algorithm, the performance is evaluated with benchmark datasets. Then, the results of LSSVM-GBO are
compared with the ANN, ANFIS, and LSSVM algorithms to demonstrate the ability and accuracy of the proposed algorithm. After modeling EC and TDS, based on different inputs in three stations, the best input combination and the best algorithm are selected. Finally, EC and TDS modeling is performed based on the best input combination and the best algorithm in different time delays. The study concludes that the LSSVM-GBO algorithm that integrates a novel meta-heuristic optimization algorithm and machine learning can raise the accuracy of the assessment of WQ parameters, while the proposed model has the potential to analyze other engineering problems. Other sections of the paper are as follows:

In Sect. 2 first introduces the study area, benchmark datasets, ANN model, ANFIS model, LSSVM model, optimization algorithm, and hybrid of LSSVM and GBO. Afterward, data collected, evaluation criteria, and finally steps of modeling WQ parameters are presented. In Sect. 3, first, the correlation coefficients between the WQ parameters are calculated. Then presents the results of benchmark datasets. After selecting the best algorithm and the best input combination based on modeling WQ parameters, EC and TDS modeling is performed at different time delays. Finally, the main points of the present study are summed up in Sect. 4.

2 Materials and Methods

2.1 Study Area

The Karun river is situated in the southwest of Iran, and with a basin area of 67,257 km$^2$ and with a length of 95 miles, it is the most important river of Iran, which leads to the Persian Gulf. Karun river originates from Zagros mountain ranges which are stretched from northwest to southeast and is the only navigable river in Iran. The average annual precipitation in Karun is 620 mm. In the present study, three stations, including Ahvaz, Armand, and Gotvand in the Karun river, have been selected to model WQ. Ahvaz and Gotvand stations are located in Khuzestan province, and Armand station is located in Chaharmahal Bakhtiari province. The choice of stations has been such that all climates are examined. Gotvand has an arid and semi-arid climate, Ahvaz has a dry and extremely dry climate, and Armand has a humid and Mediterranean climate. Fig. 1 shows the location of the study area.

2.2 Benchmark Datasets

In this article, the performance of the proposed LSSVM-GBO algorithm is compared with that of other algorithms on three real-world regression problems. Benchmark datasets that are real-world issues are a good criterion to determine algorithms working. In recent years, many studies have been conducted to compare algorithms with real-world regression problems (Breiman 2001; Zhang and Yang 2015; Henríquez and Ruz 2017). Specifications of benchmark datasets are shown in Table S1 in Supplementary Information. Also, Fig. S1 (Supplementary Information) shows the process of changing target data in benchmark datasets.
2.3 Artificial Neural Network (ANN)

ANN is a data-processing method with specific performance characteristics resembling the human brain (Lee et al. 2008). ANN has three layers: input layer, hidden layer, and output layer. ANN consists of a set of artificial neurons (Anaraki et al. 2021). Neurons and links between the neurons have a weight. Input data is processed in hidden layers and creates an output (Acharya et al. 2019). In this study, multi-layer perceptron (MLP) is used as the structure of ANN. So far, different methods have been put forward to training the network that the most famous of which is the Levenberg-Marquart method (Hadi and Tombul 2018). The relation between inputs (x) and outputs (Z) in the ANN as follows:

\[ Z = f \left( \sum_{i=1}^{n} w_i x_i + b \right) \]  \hspace{1cm} (1)

where, \( f \) is activation function, \( b \) is bias, \( w_i \) is the weight of connection, \( Z \) is the output, and \( x_i \) is the input. Figure S2 shows the structure of the ANN model.

2.4 Adaptive Neuro-Fuzzy Interface System (ANFIS)

ANFIS model method is a well-known AI method that has been used currently in WQ parameters, predicting rainfall and hydrological variables (Alizamir et al. 2020). ANFIS modeling is a reach where the combination of neural networks and fuzzy argument find their strengths (Bisht and Jangid 2011).
This model combines the advantage of both neural networks and fuzzy logic and can benefit from that at the same time (Kumar et al. 2019). ANFIS techniques can learn a system performance from enough large data sets and automatically procreate fuzzy sets to a pre-specified correctness level. The ANFIS model consists of five layers; input layers, rule layers, average layers, consequent layers, and total output layer. The main duty of the ANFIS is to optimize values of the equivalent fuzzy such that the error between the target and the actual output is minimized. Two fuzzy “if–then” rules are used as follows (Yaseen et al. 2018):

\[
\text{IF } x \text{ is } A_1 \text{ and } y \text{ is } B_1 \text{ THEN } f_1 = p_1 x + q_1 y + r_1 \tag{2}
\]

\[
\text{IF } x \text{ is } A_2 \text{ and } y \text{ is } B_2 \text{ THEN } f_2 = p_2 x + q_2 y + r_2 \tag{3}
\]

where, \(x\) and \(y\) are input variables, \(A_i\) and \(B_i\) are the linguistic labels characterized by convenient membership functions \((i = 1 \text{ or } 2)\), and \(p_i, q_i, r_i\) are the output function parameters \((i = 1 \text{ or } 2)\).

Refer to other research for more information on the ANFIS algorithm (Jang 1993; Milan et al. 2021). Figure S3 shows the structure of the ANFIS model.

### 2.5 Least Square Support Vector Machine (LSSVM)

LSSVM is an implementation of a support vector machine for classification and pattern identification, regression analysis, and the problem of learning a ranking function. The advantages of LSSVM include high precision, mathematical tractability, and direct geometric commentary. The algorithm converts the nonlinear relationship between inputs and outputs to a linear relationship (Keshtegar et al. 2019). Equation (4) shows the prevailing relationship between input and output in LSSVM:

\[
M = \sum_{i=1}^{n} k(x, x_i) \alpha_i + b \tag{4}
\]

where, \(M\) is the output, \(\alpha_i\) is the weighting coefficient of input data, \(b\) is the bias, \(k(x, x_i)\) is the value of the kernel function for different inputs. The LSSVM tries to minimize the difference between measured data and estimated data. The parameters \(\alpha_i\) and \(b\) are computed as follows (Farzin and Valikhan Anaraki 2021):

\[
\begin{bmatrix}
\text{kernel} & \bar{I}^T \\
\bar{I} & \text{kernel} + C^{-1}I
\end{bmatrix}
\begin{bmatrix}
b \\ \alpha
\end{bmatrix}
= \begin{bmatrix}
0 \\
M
\end{bmatrix} \tag{5}
\]

C is regulation parameter; the parameters \(\alpha, M, I, \bar{I}\) are computed as follows:

\[
\alpha = \begin{bmatrix}
\alpha_1 \\
\vdots \\
\alpha_n
\end{bmatrix}, \quad \bar{I} = \begin{bmatrix}
1 \\
\vdots \\
1
\end{bmatrix}, \quad M = \begin{bmatrix}
M_1 \\
\vdots \\
M_n
\end{bmatrix}, \quad I = \text{diag}(1,1,\ldots,1) \tag{6}
\]

The radial basis function (RBF) is used as a kernel function, as follows (Farzin et al. 2020):

\[
k(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right) \tag{7}
\]

Figure 2 shows the structure of the LSSVM model.
2.6 Optimization Algorithm

2.6.1 GBO Algorithm

The GBO was proposed as a meta-heuristic optimization algorithm by Ahmadianfar et al. (2020) and showed that GBO has a more promising operation capability than other optimization algorithms. The GBO uses two main operators: gradient search rule (GSR) and local escaping operator (LEO) (Hassan et al. 2021). GSR uses the slope-based method to reach better positions in the search space. Simultaneous operation of two stages of exploration and exploitation and creating a suitable balance between these two processes causes the optimal performance of this algorithm. (Olorunda and Engelbrecht 2008; Patel and Savsani 2015; Draa et al. 2015).

2.6.2 Gradient Search Rule (GSR)

GSR is the core of the GBO algorithm. The duty of GSR is to find better opportunities and increase convergence rate acceleration (DM) in the search space. Therefore, the equation to update the current vector position is \( x_n^m \)

\[
X_{1n}^m = x_n^m - \text{randn} \times \rho_1 \times \frac{2\Delta x \times x_n^m}{(y_{p_n}^m - y_{q_n}^m + \epsilon)} + \text{rand} \times \rho_2 \times (x_{\text{best}} - x_n^m) \tag{8}
\]

where, randn is a random number, \( \epsilon \) is a small number \([0, 0.10]\), rand is a random number in \([0, 1]\), and \( x_{\text{best}} \) is the best solution. \( \rho_1 \) is a factor to balance between two stages of...
exploration and exploitation, $\rho_2$ is a random parameter, $\Delta x$ equal to the difference between the best solution ($x_{\text{best}}$) and a randomly selected position ($x^m_{r1}$). By replacing the position of the best vector ($x_{\text{best}}$) with the current vector ($x^m_n$) in Eq. (8), the new vector ($X^m_{2n}$) is obtained as follows:

$$X^m_{2n} = x_{\text{best}} - \text{rand} \times \rho_1 \times \frac{2\Delta x \times x^m_n}{(yF^m_n - yq^m_n + \varepsilon)} + \text{rand} \times \rho_2 \times (x^m_{r1} - x^m_{r2})$$  \hspace{1cm} (9)

Due to the positions $X^m_{1n}$, $X^m_{2n}$, $X^m_n$, the new solution at the next iteration ($x^{m+1}_n$) can be defined as:

$$x^{m+1}_n = r_a \times (r_b \times X^m_{1n} + (1 - r_b) \times X^m_{2n}) + (1 - r_a) \times X^m_{3n}$$  \hspace{1cm} (10)

$$X^m_{3n} = X^m_n - \rho_1 \times (X^m_{2n} - X^m_{1n})$$  \hspace{1cm} (11)

$r_a$ and $r_b$ are two random numbers in [0, 1]. Figure 3 shows the structure of the GBO model.
2.6.3 Local Escaping Operator (LEO)

The LEO operator helps the algorithm to exit local optima points and accelerate the convergence of the algorithm. The LEO is capable of solving complex problems in the GBO algorithm. By using several solutions, the LEO generates a solution with superior performance ($X_{m_{LEO}}^m$). The solution $X_{m_{LEO}}^m$ is produced as follows:

1: if rand $<$ pr
   if rand $<$ 0.5
2: $X_{m_{LEO}}^m = X_{n}^{m+1} + f_1 \times (u_1 \times x_{best} - u_2 \times x_{k}^m) + f_2 \times \rho_1 \times (u_3 \times (X_{2}^{m} - X_{1}^{m}) + u_2 \times (x_{n}^{m} - x_{k}^m))/2$
3: $X_{n}^{m+1} = X_{m_{LEO}}^m$
4: Else
5: $X_{m_{LEO}}^m = x_{best} + f_1 \times (u_1 \times x_{best} - u_2 \times x_{k}^m) + f_2 \times \rho_1 \times (u_3 \times (X_{2}^{m} - X_{1}^{m}) + u_2 \times (x_{n}^{m} - x_{k}^m))/2$
6: $X_{n}^{m+1} = X_{m_{LEO}}^m$
7: End
8: End

where, $f_1$ is uniform random number $\in [-1,1]$, $f_2$ is a random number from a normal distribution with a mean of 0 and standard deviation of 1, pr is the probability, while $u_1$, $u_2$, and $u_3$ are random numbers.

For more details, see Ahmadianfar et al. (2020).

2.7 Hybrid of LSSVM and GBO

The parameters $C$ and $\sigma$ have a notable effect on the performance of the LSSVM model. In this study, the optimization algorithm GBO was used to find the optimal value of the LSSVM parameters. In the hybrid algorithm of LSSVM and GBO, the values of the LSSVM parameters are considered as decision variables. Also, the scheme of LSSVM-GBO is shown in Fig. 4. The steps of the LSSVM-GBO algorithm are described as follows:

1. Test and training data are randomly selected from the available data.
2. The initial parameters of the optimization algorithms GBO (the number of iterations and the population size) are randomly determined.
3. The LSSVM parameters (initial population) are initialized, and the GBO algorithm finds the optimal solution (values of parameters $C$ and $\sigma$) in the search space.
4. After obtaining the optimal answer of the LSSVM parameters, training data and test data are used to obtain the LSSVM optimization model and to evaluate the predictive ability of the LSSVM optimization model.

2.8 Data Collected

According to Table S2, in this study, the WQ data from three Karun river stations were used for EC and TDS modeling. Since each combination of inputs can have a unique effect on the accuracy of the results, this reason, 13 combinations of inputs with 0–12 months period time delay have been created for the inputs of the algorithms. Table S3 shows the details of time delays on the best inputs.

2.9 Evaluation Criteria

In the present study, four evaluation criteria, including mean absolute error (MAE), relative root mean square error (RRMSE), correlation coefficient (R), and coefficient of determination $R^2$, were calculated in testing phases. Expressions for these measures are given as follows (Zhu et al. 2020):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |O_i - P_i|$$

(13)
where, N: the number of data, O: observed values, P: predicted values.

### 2.10 Assessment of WQ Parameters

Figure 5 depicts the general framework for the assessment of WQ parameters. The steps of WQ modeling are described as follows:

1. Comparison of the performance of algorithms with three benchmarks.

\[
\text{RRMSE} = \sqrt{\frac{\sum_{i=1}^{N} (O_i - P_i)^2}{N}} / \text{SD}(P) \tag{14}
\]

\[
R = \frac{\sum_{i=1}^{N} (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N} (P_i - \bar{P})^2 \sum_{i=1}^{N} (O_i - \bar{O})}} \tag{15}
\]

\[
R^2 = \left[ \frac{\sum_{i=1}^{N} (P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{N} (P_i - \bar{P})^2 \sum_{i=1}^{N} (O_i - \bar{O})}} \right]^2 \tag{16}
\]

Fig. 5 Flowchart for modeling WQ parameter
2. 70% of the data is considered for the training data and 30% for the test data.
3. WQ is modeled based on the training data and the test data with the mentioned algorithms.
4. According to the evaluation criteria, the best algorithm and the best input combination are determined.
5. Creating time delays in the best algorithm and best input combination.
6. Calculating evaluation criteria in time delays and select the best time delay.
7. Finally, the EC and TDS time series are calculated.

3 Results and Discussion

3.1 Calculate the Correlation Coefficient

Tables 1, 2, and 3 show the correlation between WQ parameters in three stations of Ahvaz, Armand, and Gotvand. Based on the results of the correlation matrix, the highest amount of correlation of EC and TDS parameters with input parameters related to Sum.A and Sum.C at three stations. The correlation between EC and inputs is greater. The higher the correlation between inputs and outputs, the higher the modeling accuracy.

3.2 Testing Algorithms with Benchmark Datasets

According to Table 4, the performance of the LSSVM-GBO algorithm is compared with that of other algorithms on three real-world regression problems. Results showed that the hybrid model provides much better accuracy than the ANN, ANFIS, and LSSVM model. Values of MAE, RRMSE, R in the Housing dataset were 5.30, 0.91, 0.44, respectively. Also, in the LVST dataset, they were 99.94, 0.20, 0.98, respectively, and in the Servo
### Table 2: Correlation coefficient results between input and output parameters in Armand station

| Parameters | Correlation matrix |
|------------|--------------------|
|            | Q  | Ca$^{+2}$ | Cl$^{-1}$ | SAR | HCO$_3$ | Sum.A | Sum.C | Mg$^{+2}$ | Na$^{+1}$ | SO$_4$ | pH | EC | TDS |
| Q          | 1.00 |          |          |     |          |       |       |          |          |       |    |    |     |
| Ca$^{+2}$  | -0.01 | 1.00 |          |     |          |       |       |          |          |       |    |    |     |
| Cl$^{-1}$  | -0.44 | 0.26 | 1.00 |     |          |       |       |          |          |       |    |    |     |
| SAR        | -0.4 | 0.06 | 0.79 | 1.00 |          |       |       |          |          |       |    |    |     |
| HCO$_3$    | -0.04 | 0.23 | -0.03 | -0.03 | 1.00 |       |       |          |          |       |    |    |     |
| Sum.A      | -0.35 | 0.52 | -0.84 | 0.63 | 0.17 | 1.00 |       |          |          |       |    |    |     |
| Sum.C      | -0.34 | 0.56 | 0.84 | 0.61 | 0.17 | 0.98 | 1.00 |          |          |       |    |    |     |
| Mg$^{+2}$  | -0.18 | 0.10 | 0.45 | 0.15 | 0.17 | 0.65 | 0.66 | 1.00 |          |       |    |    |     |
| Na$^{+1}$  | -0.02 | 0.04 | 0.01 | 0.90 | 0.02 | 0.02 | 0.98 | 1.00 |          |       |    |    |     |
| SO$_4$     | -0.02 | 0.04 | 0.08 | 0.13 | 0.02 | -0.02 | 0.01 | 0.07 | 0.04 | 1.00 |    |    |     |
| pH         | 0.01 | 0.16 | 0.08 | 0.13 | 0.02 | -0.02 | 0    | 0.01 | 0.07 | 0.04 | 1.00 |    |    |     |
| EC         | 0.35 | 0.52 | 0.85 | 0.63 | 0.18 | 0.94 | 0.95 | 0.61 | -0.01 | 0.03 | 0.05 | 1.00 |    |     |
| TDS        | -0.21 | 0.23 | 0.43 | 0.29 | 0.59 | 0.84 | 0.84 | 0.92 | 0.57 | 0.58 | 0.07 | 0.88 | 1.00 |     |
| Parameters | Q   | Ca$^{+2}$ | Cl$^{-1}$ | SAR | HCO$_3$ | Sum.A | Sum.C | Mg$^{+2}$ | Na$^{+1}$ | SO$_4$ | pH | EC | TDS |
|------------|-----|-----------|-----------|-----|---------|-------|-------|-----------|-----------|-------|-----|----|-----|
| Q          | 1.00|           |           |     |         |       |       |            |            |       |     |    |     |
| Ca$^{+2}$  | 0.13| 1.00      |           |     |         |       |       |            |            |       |     |    |     |
| Cl$^{-1}$  | -0.30| 0.49     | 1.00      |     |         |       |       |            |            |       |     |    |     |
| SAR        | -0.35| 0.30     | 0.95      | 1.00|         |       |       |            |            |       |     |    |     |
| HCO$_3$    | -0.01| 0.29     | 0.12      | 0.05| 1.00    |       |       |            |            |       |     |    |     |
| Sum.A      | -0.20| 0.69     | 0.95      | 0.85| 0.21    | 1.00  |       |            |            |       |     |    |     |
| Sum.C      | -0.20| 0.70     | 0.95      | 0.85| 0.21    | 0.99  | 1.00  |            |            |       |     |    |     |
| Mg$^{+2}$  | -0.17| 0.20     | 0.14      | 0.33| 0.17    | 0.53  | 0.53  | 1.00      |            |       |     |    |     |
| Na$^{+1}$  | -0.29| 0.49     | 0.99      | 0.96| 0.12    | 0.95  | 0.95  | 0.42      | 1.00      |       |     |    |     |
| SO$_4$     | 0.07 | 0.81     | 0.55      | 0.40| 0       | 0.75  | 0.75  | 0.51      | 0.57      | 1.00  |     |    |     |
| pH         | 0.05 | 0.18     | 0.06      | 0   | -0.27   | -0.11 | -0.11 | -0.08     | -0.05     | -0.10 | 1.00|    |     |
| EC         | -0.22| 0.67     | 0.95      | 0.87| 0.20    | 0.98  | 0.98  | 0.49      | 0.95      | 0.71  | -0.09| 1.00|     |
| TDS        | -0.20| 0.67     | 0.94      | 0.85| 0.18    | 0.97  | 0.97  | 0.48      | 0.93      | 0.71  | -0.10| 0.98| 1.00|
dataset, were 0.46, 0.41, 0.91, respectively. In Fig. 6, the modeling accuracy for the LSSVM-GBO algorithm and benchmark datasets has been shown.

### 3.3 Select the Best Input Combination and the Best Algorithm

In this section, by estimating EC and TDS using different inputs, the best input combination and the best algorithm are identified. According to the results obtained, at all stations, the LSSVM-GBO algorithm has the highest accuracy.

In Table 5, the results of the algorithms are listed to the EC modeling of the Ahvaz station. Based on the results of Table 5, The modeling results showed that the Sum.C parameter has the most impact in modeling EC. In the optimal hybrid model, values of MAE, RRMSE, and R were 74.30, 0.14, 0.99, respectively. In Table 6, the evaluation criteria of the listed to the EC modeling of the Armand station. Sum.C parameter has the most impact in modeling EC in this station. Also, the value of MAE is equal to 19.67, RRMSE equal to 0.22, R equal to 0.98. In Table 7, the results of the EC
calculated by different algorithms and the combination of different inputs in the Gotvand station are compared. Also, the Sum.C parameter has the greatest impact in modeling EC in the Gotvand station. Values of MAE, RRMSE, R were 57.12, 0.16, 0.99, respectively.

The results of TDS modeling in Ahvaz station by ANN, ANFIS, LSSVM, LSSVM-GBO algorithms are shown in Table 8. Sum.C parameter has the most significant impact in modeling TDS. Values of MAE, RRMSE, R were 74.89, 0.43, 0.90, respectively. The results of TDS modeling in the Armand station are shown in Table 9. The modeling results showed that the Sum.C parameter has the most significant impact on modeling TDS. Values of MAE, RRMSE, R were 16.99, 0.26, 0.97, respectively. In Table 10, the results of the TDS calculated by different algorithms and the combination of different inputs in the Gotvand station are compared. Sum.C parameter has the most significant impact on modeling TDS in the Gotvand station. Values of MAE, RRMSE, R were 37.18, 0.17, 0.99, respectively.
### Table 5 Results of the algorithms to the EC modeling of the Ahvaz station

| Algorithms       | Evaluation criteria | Inputs (with Q) | Ca\(^{+2}\) | Cl\(^{-1}\) | SAR | HCO\(_3\) | Sum.C | Mg\(^{+2}\) | Na\(^{+1}\) | SO\(_4\) | Sum.A | pH |
|------------------|---------------------|-----------------|-------------|------------|-----|-----------|------|-----------|-----------|--------|-------|-----|
| ANN-train        | MAE                 |                 | 160.94      | 78.59      | 148.70 | 259.64    | 51.56 | 208.07    | 317.28    | 137.53 | 58.71 | 293.11 |
|                  | RRMSE               |                 | 0.52        | 0.25       | 0.48 | 0.86      | 0.17  | 0.65      | 0.99      | 0.45   | 0.22  | 0.91 |
|                  | R                   |                 | 0.85        | 0.96       | 0.87 | 0.55      | 0.98  | 0.75      | 0.07      | 0.88   | 0.97  | 0.40 |
| ANFIS-train      | MAE                 |                 | 173.53      | 88.88      | 150.45 | 277.59    | 52.29 | 202.75    | 79.50     | 139.38 | 61.9  | 265.01 |
|                  | RRMSE               |                 | 0.59        | 0.26       | 0.49 | 0.85      | 0.18  | 0.64      | 0.26      | 0.48   | 0.25  | 0.82 |
|                  | R                   |                 | 0.80        | 0.96       | 0.87 | 0.51      | 0.98  | 0.76      | 0.96      | 0.87   | 0.96  | 0.55 |
| LSSVM-train      | MAE                 |                 | 146.42      | 84.86      | 146.29 | 242.13    | 56.08 | 205.92    | 81.45     | 134.25 | 63.64 | 258.61 |
|                  | RRMSE               |                 | 0.48        | 0.27       | 0.47 | 0.76      | 0.18  | 0.65      | 0.28      | 0.44   | 0.25  | 0.80 |
|                  | R                   |                 | 0.87        | 0.96       | 0.88 | 0.64      | 0.98  | 0.75      | 0.95      | 0.89   | 0.96  | 0.59 |
| LSSVM-GBO train  | MAE                 |                 | 156.98      | 84.78      | 139.43 | 201.06    | 56.11 | 203.48    | 81.59     | 133.13 | 63.68 | 251.47 |
|                  | RRMSE               |                 | 0.50        | 0.27       | 0.45 | 0.65      | 0.18  | 0.64      | 0.28      | 0.44   | 0.25  | 0.78 |
|                  | R                   |                 | 0.86        | 0.96       | 0.89 | 0.76      | 0.98  | 0.76      | 0.95      | 0.89   | 0.96  | 0.61 |
| ANN-test         | MAE                 |                 | 207.11      | 189.41     | 288.93 | 412.77    | 78.552 | 345.42   | 649.09    | 259.48 | 71.97 | 615.56 |
|                  | RRMSE               |                 | 0.52        | 0.56       | 0.57 | 0.85      | 0.15  | 0.69      | 1.25      | 0.65   | 0.15  | 1.21 |
|                  | R                   |                 | 0.86        | 0.88       | 0.90 | 0.69      | 0.99  | 0.75      | -0.02     | 0.78   | 0.98  | 0.45 |
| ANFIS-test       | MAE                 |                 | 248.17      | 143.18     | 285.56 | 521.88    | 78.38 | 330.95    | 125.59    | 201.68 | 73.36 | 589.68 |
|                  | RRMSE               |                 | 0.50        | 0.27       | 0.55 | 1.03      | 0.15  | 0.66      | 0.24      | 0.43   | 0.14  | 1.16 |
|                  | R                   |                 | 0.87        | 0.97       | 0.92 | 0.74      | 0.99  | 0.77      | 0.97      | 0.90   | 0.98  | 0.57 |
| LSSVM-test       | MAE                 |                 | 200.93      | 141.84     | 276.8  | 439.53    | 78.34 | 291.2     | 130.73    | 196.89 | 72.68 | 527.96 |
|                  | RRMSE               |                 | 0.43        | 0.27       | 0.54 | 0.89      | 0.15  | 0.60      | 0.25      | 0.42   | 0.14  | 1.02 |
|                  | R                   |                 | 0.90        | 0.98       | 0.93 | 0.69      | 0.98  | 0.81      | 0.97      | 0.91   | 0.97  | 0.66 |
| LSSVM-GBO test   | MAE                 |                 | 201.49      | 141.43     | 269.39 | 405.04    | 74.3  | 288.34    | 129.23    | 194.43 | 72.47 | 497.10 |
|                  | RRMSE               |                 | 0.42        | 0.26       | 0.52 | 0.85      | 0.14  | 0.59      | 0.24      | 0.42   | 0.15  | 0.97 |
|                  | R                   |                 | 0.91        | 0.97       | 0.92 | 0.69      | 0.99  | 0.82      | 0.98      | 0.91   | 0.98  | 0.71 |
### Table 6: Results of the algorithms to the EC modeling of the Armand station

| Algorithms       | Evaluation criteria | Inputs (with Q) | Ca$^{2+}$ | Cl$^{-1}$ | SAR | HCO$_3$ | Sum.C | Mg$^{2+}$ | Na$^{+1}$ | SO$_4$ | Sum.A | pH  |
|------------------|---------------------|-----------------|-----------|-----------|-----|---------|-------|-----------|-----------|--------|-------|-----|
| ANN-train        | MAE                 |                 | 58.00     | 39.78     | 56.88| 58.00   | 27.71 | 62.67     | 123.33    | 67.36  | 28.99 | 66.09|
| ANN-train        | RRMSE               |                 | 0.69      | 0.47      | 0.67| 0.68    | 0.33  | 0.73      | 1.36      | 0.78   | 0.36  | 0.77|
| ANN-train        | R                   |                 | 0.72      | 0.87      | 0.73| 0.72    | 0.94  | 0.68      | 0.46      | 0.61   | 0.93  | 0.63|
| ANN-train        | MAE                 |                 | 58.09     | 38.73     | 56.88| 64.37   | 27.70 | 60.04     | 49.21     | 64.66  | 29.61 | 65.97|
| ANFIS-train      | MAE                 |                 | 58.09     | 38.73     | 56.88| 64.37   | 27.70 | 60.04     | 49.21     | 64.66  | 29.61 | 65.97|
| ANFIS-train      | RRMSE               |                 | 0.69      | 0.46      | 0.66| 0.76    | 0.33  | 0.71      | 0.58      | 0.75   | 0.37  | 0.97|
| ANFIS-train      | R                   |                 | 0.71      | 0.88      | 0.74| 0.64    | 0.94  | 0.70      | 0.81      | 0.65   | 0.92  | 0.62|
| LSSVM-train      | MAE                 |                 | 61.42     | 37.99     | 61.47| 58.63   | 28.10 | 58.69     | 61.02     | 69.44  | 30.32 | 62.22|
| LSSVM-train      | RRMSE               |                 | 0.72      | 0.46      | 0.72| 0.69    | 0.33  | 0.70      | 0.71      | 0.80   | 0.37  | 0.73|
| LSSVM-train      | R                   |                 | 0.68      | 0.88      | 0.68| 0.71    | 0.94  | 0.71      | 0.69      | 0.59   | 0.92  | 0.67|
| LSSVM-GBO train  | MAE                 |                 | 61.33     | 37.48     | 61.36| 57.78   | 28.01 | 57.76     | 41.79     | 68.97  | 30.30 | 47.76|
| LSSVM-GBO train  | RRMSE               |                 | 0.72      | 0.45      | 0.72| 0.69    | 0.33  | 0.68      | 0.51      | 0.79   | 0.38  | 0.60|
| LSSVM-GBO train  | R                   |                 | 0.68      | 0.88      | 0.69| 0.72    | 0.94  | 0.72      | 0.86      | 0.60   | 0.92  | 0.79|
| ANN-test         | MAE                 |                 | 73.22     | 47.02     | 83.24| 87.02   | 19.58 | 77.94     | 91.68     | 95.36  | 25.74 | 96.09|
| ANN-test         | RRMSE               |                 | 0.81      | 0.65      | 0.93| 0.93    | 0.22  | 0.88      | 0.94      | 1.01   | 0.28  | 1.03|
| ANN-test         | R                   |                 | 0.74      | 0.80      | 0.75| 0.70    | 0.98  | 0.64      | 0.55      | 0.68   | 0.97  | 0.67|
| ANFIS-test       | MAE                 |                 | 81.99     | 50.6      | 85.03| 103.07  | 21.18 | 83.71     | 74.54     | 104.37 | 24.12 | 102.44|
| ANFIS-test       | RRMSE               |                 | 0.88      | 0.65      | 0.99| 1.44    | 0.25  | 0.93      | 0.96      | 1.10   | 0.29  | 1.09|
| ANFIS-test       | R                   |                 | 0.72      | 0.81      | 0.71| 0.21    | 0.97  | 0.59      | 0.70      | 0.66   | 0.97  | 0.62|
| LSSVM-test       | MAE                 |                 | 80.61     | 48.13     | 86.01| 85.13   | 19.87 | 75.5      | 97.04     | 95.88  | 22.99 | 97.32|
| LSSVM-test       | RRMSE               |                 | 0.87      | 0.63      | 0.95| 0.91    | 0.23  | 0.80      | 1.02      | 0.99   | 0.27  | 1.03|
| LSSVM-test       | R                   |                 | 0.74      | 0.82      | 0.75| 0.73    | 0.97  | 0.76      | 0.66      | 0.77   | 0.97  | 0.66|
| LSSVM-GBO test   | MAE                 |                 | 80.33     | 47.05     | 85.72| 82.66   | 19.67 | 72.39     | 73.43     | 95.31  | 22.57 | 93.33|
| LSSVM-GBO test   | RRMSE               |                 | 0.78      | 0.61      | 0.95| 0.85    | 0.22  | 0.78      | 0.84      | 0.98   | 0.26  | 1.02|
| LSSVM-GBO test   | R                   |                 | 0.74      | 0.83      | 0.75| 0.80    | 0.98  | 0.78      | 0.76      | 0.77   | 0.98  | 0.63|
### Table 7: Results of the algorithms to the EC modeling of the Gotvand station

| Algorithms          | Evaluation criteria | Inputs (with Q) | Ca\(^{2+}\) | Cl\(^{-1}\) | SAR | HCO\(_3\) | Sum.C | Mg\(^{2+}\) | Na\(^{+1}\) | SO\(_4\) | Sum.A | pH  |
|---------------------|----------------------|-----------------|----------|-----------|-----|---------|-------|----------|---------|--------|-------|-----|
| ANN-train           | MAE                  |                 | 125.02   | 208       | 81.88 | 186.25  | 37.77 | 164.36   | 64.53   | 114.73 | 36.53 | 215.98 |
| ANN-train           | RRMSE                |                 | 0.61     | 0.99      | 0.40  | 0.93    | 0.18  | 0.84     | 0.34    | 0.54   | 0.18  | 1.00 |
| ANN-train           | R                    |                 | 0.79     | 0.05      | 0.91  | 0.36    | 0.98  | 0.51     | 0.94    | 0.84   | 0.98  | 0.11 |
| ANFIS-train         | MAE                  |                 | 129.77   | 53.08     | 85.62 | 181.42  | 38.13 | 162.08   | 55.64   | 120.67 | 37.53 | 182.7 |
| ANFIS-train         | RRMSE                |                 | 0.64     | 0.25      | 0.45  | 0.92    | 0.18  | 0.85     | 0.27    | 0.58   | 0.18  | 0.91 |
| ANFIS-train         | R                    |                 | 0.76     | 0.96      | 0.89  | 0.37    | 0.98  | 0.51     | 0.96    | 0.81   | 0.98  | 0.40 |
| LSSVM-train         | MAE                  |                 | 119.12   | 53.06     | 83.56 | 183.98  | 39.76 | 168.76   | 55.04   | 126.52 | 38.68 | 208.22 |
| LSSVM-train         | RRMSE                |                 | 0.58     | 0.26      | 0.43  | 0.94    | 0.19  | 0.86     | 0.26    | 0.59   | 0.18  | 0.99 |
| LSSVM-train         | R                    |                 | 0.86     | 0.96      | 0.90  | 0.33    | 0.98  | 0.49     | 0.96    | 0.80   | 0.98  | 0.26 |
| LSSVM-GBO train     | MAE                  |                 | 118.53   | 52.91     | 83.74 | 173.44  | 39.94 | 167.7    | 53.11   | 126.5  | 38.8  | 183.76 |
| LSSVM-GBO train     | RRMSE                |                 | 0.58     | 0.25      | 0.42  | 0.89    | 0.19  | 0.86     | 0.26    | 0.59   | 0.18  | 0.91 |
| LSSVM-GBO train     | R                    |                 | 0.81     | 0.96      | 0.90  | 0.44    | 0.98  | 0.49     | 0.96    | 0.80   | 0.98  | 0.42 |
| ANN-test            | MAE                  |                 | 210.01   | 358.73    | 151.98 | 322.94  | 58.92 | 294.14   | 102.11  | 217.83 | 56.25 | 385.13 |
| ANN-test            | RRMSE                |                 | 0.64     | 1.08      | 0.51  | 1.00    | 0.17  | 0.93     | 0.36    | 0.67   | 0.16  | 1.06 |
| ANN-test            | R                    |                 | 0.77     | 0.42      | 0.87  | 0.38    | 0.98  | 0.48     | 0.93    | 0.73   | 0.98  | –0.16|
| ANFIS-test          | MAE                  |                 | 224.85   | 89.71     | 153.88 | 310.32  | 60.25 | 278.79   | 98.53   | 216.57 | 53.09 | 339.99 |
| ANFIS-test          | RRMSE                |                 | 0.67     | 0.30      | 0.52  | 1.00    | 0.16  | 0.89     | 0.32    | 0.68   | 0.15  | 1.08 |
| ANFIS-test          | R                    |                 | 0.74     | 0.95      | 0.86  | 0.39    | 0.99  | 0.51     | 0.94    | 0.73   | 0.99  | 0.16 |
| LSSVM-test          | MAE                  |                 | 212.53   | 81.86     | 141.81 | 315.53  | 57.95 | 283.33   | 88.94   | 228.46 | 52.72 | 358.75 |
| LSSVM-test          | RRMSE                |                 | 0.65     | 0.27      | 0.49  | 1.00    | 0.17  | 0.88     | 0.29    | 0.69   | 0.15  | 1.08 |
| LSSVM-test          | R                    |                 | 0.76     | 0.96      | 0.88  | 0.44    | 0.97  | 0.52     | 0.95    | 0.73   | 0.96  | –0.01|
| LSSVM-GBO test      | MAE                  |                 | 211.02   | 81.57     | 141.08 | 312.16  | 57.12 | 282.43   | 88.69   | 228.49 | 52.77 | 336.5 |
| LSSVM-GBO test      | RRMSE                |                 | 0.64     | 0.26      | 0.49  | 0.99    | 0.16  | 0.88     | 0.29    | 0.69   | 0.14  | 1.07 |
| LSSVM-GBO test      | R                    |                 | 0.76     | 0.97      | 0.88  | 0.41    | 0.99  | 0.52     | 0.95    | 0.72   | 0.99  | 0.17 |
## Table 8  Results of the algorithms to the TDS modeling of the Ahvaz station

| Algorithms        | Evaluation criteria | Inputs ( with Q) | Ca²⁺ | Cl⁻¹ | SAR  | HCO₃  | Sum.C | Mg²⁺ | Na⁺¹ | SO₄  | Sum.A | pH    |
|-------------------|---------------------|------------------|------|------|------|-------|-------|------|------|------|-------|-------|
| ANN-train         | MAE                 |                  | 109.18 | 68.14 | 108.29 | 297.2 | 63.82 | 162.43 | 65.48 | 189.27 | 53.11 | 174.13 |
| ANN-train         | RRMSE               |                  | 0.52  | 0.33  | 0.52  | 1.20  | 0.33  | 0.71  | 0.33  | 0.86  | 0.27  | 0.81  |
| ANN-train         | R                   |                  | 0.84  | 0.94  | 0.85  | 0.39  | 0.94  | 0.71  | 0.94  | 0.50  | 0.96  | 0.57  |
| ANFIS-train       | MAE                 |                  | 122.52 | 68.44 | 110.05 | 190.37 | 49.53 | 142.62 | 67.54 | 96.59  | 54.98 | 184.75 |
| ANFIS-train       | RRMSE               |                  | 0.61  | 0.32  | 0.54  | 0.86  | 0.25  | 0.65  | 0.33  | 0.49  | 0.31  | 0.54  |
| ANFIS-train       | R                   |                  | 0.79  | 0.94  | 0.84  | 0.49  | 0.96  | 0.75  | 0.94  | 0.86  | 0.94  | 0.51  |
| LSSVM-train       | MAE                 |                  | 113.69 | 68.28 | 108.19 | 166.08 | 52.06 | 139.55 | 67.72 | 92.57  | 57.56 | 166.71 |
| LSSVM-train       | RRMSE               |                  | 0.54  | 0.33  | 0.52  | 0.78  | 0.26  | 0.65  | 0.34  | 0.46  | 0.31  | 0.76  |
| LSSVM-train       | R                   |                  | 0.53  | 0.94  | 0.85  | 0.61  | 0.96  | 0.75  | 0.93  | 0.88  | 0.95  | 0.64  |
| LSSVM-GBO train   | MAE                 |                  | 111.87 | 64.51 | 103.84 | 137.39 | 50.75 | 141.98 | 64.45 | 91.08  | 54.90 | 172.41 |
| LSSVM-GBO train   | RRMSE               |                  | 0.53  | 0.31  | 0.50  | 0.66  | 0.26  | 0.66  | 0.32  | 0.45  | 0.29  | 0.79  |
| LSSVM-GBO train   | R                   |                  | 0.84  | 0.94  | 0.86  | 0.75  | 0.96  | 0.74  | 0.94  | 0.88  | 0.95  | 0.60  |
| ANN-test          | MAE                 |                  | 143.64 | 124.12 | 193.7  | 331.70 | 120.64 | 206.04 | 124.79 | 243.75 | 73.51 | 376.84 |
| ANN-test          | RRMSE               |                  | 0.57  | 0.49  | 0.64  | 0.94  | 0.51  | 0.69  | 0.51  | 0.74  | 0.45  | 1.12  |
| ANN-test          | R                   |                  | 0.82  | 0.88  | 0.82  | 0.60  | 0.85  | 0.73  | 0.86  | 0.71  | 0.89  | 0.54  |
| ANFIS-test        | MAE                 |                  | 174.68 | 124.26 | 204.25 | 359.21 | 74.58  | 231.96 | 113.38 | 149.41 | 72.68 | 403.95 |
| ANFIS-test        | RRMSE               |                  | 0.60  | 0.53  | 0.66  | 1.06  | 0.48  | 0.78  | 0.51  | 0.57  | 0.47  | 1.19  |
| ANFIS-test        | R                   |                  | 0.80  | 0.85  | 0.82  | 0.65  | 0.88  | 0.66  | 0.86  | 0.82  | 0.88  | 0.43  |
| LSSVM-test        | MAE                 |                  | 146.18 | 126.34 | 199.45 | 309.07 | 75.00  | 216.71 | 113.66 | 144.07 | 74.9  | 363.57 |
| LSSVM-test        | RRMSE               |                  | 0.55  | 0.50  | 0.66  | 0.95  | 0.44  | 0.72  | 0.48  | 0.55  | 0.44  | 1.06  |
| LSSVM-test        | R                   |                  | 0.83  | 0.87  | 0.81  | 0.60  | 0.89  | 0.72  | 0.87  | 0.83  | 0.89  | 0.60  |
| LSSVM-GBO test    | MAE                 |                  | 143.68 | 125.58 | 195.84 | 308.62 | 74.89  | 204.91 | 113.00 | 143.46 | 77.87 | 347.75 |
| LSSVM-GBO test    | RRMSE               |                  | 0.54  | 0.50  | 0.64  | 0.95  | 0.43  | 0.70  | 0.48  | 0.55  | 0.43  | 1.03  |
| LSSVM-GBO test    | R                   |                  | 0.83  | 0.88  | 0.82  | 0.64  | 0.90  | 0.73  | 0.87  | 0.83  | 0.90  | 0.57  |
Table 9 Results of the algorithms to the TDS modeling of the Armand station

| Algorithms   | Evaluation criteria | Inputs ( with Q) | Ca\(^{+2}\) | Cl\(^{-1}\) | SAR | HCO\(_3\) | Sum.C | Mg\(^{+2}\) | Na\(^{+1}\) | SO\(_4\) | Sum.A | pH |
|--------------|---------------------|------------------|------------|------------|-----|---------|-------|---------|---------|-------|------|----|
| ANN-train    | MAE                 |                  | 58.38      | 52.27      | 75.64| 60.44   | 40.40 | 65.80   | 43.44   | 44.44 | 56.29| 60.96|
| ANN-train    | RRMSE               |                  | 0.92       | 0.92       | 1.00| 0.43    | 0.76  | 0.96    | 0.29    | 0.30  | 0.88 | 0.91|
| ANN-train    | R                   |                  | 0.38       | 0.36       | 0.15| 0.94    | 0.64  | 0.26    | 0.95    | 0.95  | 0.51 | 0.41|
| ANFIS-train  | MAE                 |                  | 59.58      | 52.88      | 59.69| 64.87   | 47.35 | 59.29   | 31.33   | 66.31 | 48.77| 63.43|
| ANFIS-train  | RRMSE               |                  | 0.92       | 0.91       | 0.92| 0.66    | 0.91  | 0.92    | 0.21    | 0.61  | 0.91 | 0.92|
| ANFIS-train  | R                   |                  | 0.37       | 0.40       | 0.37| 0.74    | 0.40  | 0.37    | 0.97    | 0.78  | 0.40 | 0.38|
| LSSVM-train  | MAE                 |                  | 59.11      | 47.18      | 60.34| 41.68   | 41.77 | 60.67   | 39.87   | 57.08 | 43.68| 65.33|
| LSSVM-train  | RRMSE               |                  | 0.94       | 0.92       | 0.94| 0.33    | 0.92  | 0.94    | 0.27    | 0.54  | 0.92 | 0.94|
| LSSVM-train  | R                   |                  | 0.34       | 0.37       | 0.32| 0.95    | 0.38  | 0.34    | 0.96    | 0.84  | 0.38 | 0.32|
| LSSVM-GBO train | MAE           |                  | 58.88      | 47.82      | 60.56| 43.03   | 19.92 | 60.78   | 29.65   | 55.96 | 43.58| 65.28|
| LSSVM-GBO train | RRMSE        |                  | 0.93       | 0.92       | 0.94| 0.35    | 0.40  | 0.94    | 0.20    | 0.50  | 0.92 | 0.94|
| LSSVM-GBO train | R            |                  | 0.35       | 0.38       | 0.32| 0.94    | 0.91  | 0.33    | 0.97    | 0.87  | 0.38 | 0.33|
| ANN-test     | MAE                 |                  | 46.50      | 35.35      | 53.67| 51.34   | 18.27 | 61.13   | 64.58   | 63.64 | 28.18| 51.39|
| ANN-test     | RRMSE               |                  | 0.77       | 0.70       | 0.88| 0.83    | 0.31  | 0.97    | 1.05    | 1.05  | 0.45 | 0.83|
| ANN-test     | R                   |                  | 0.71       | 0.78       | 0.51| 0.74    | 0.97  | 0.44    | 0.64    | 0.66  | 0.89 | 0.71|
| ANFIS-test   | MAE                 |                  | 45.50      | 42.88      | 46.26| 42.95   | 26.58 | 41.03   | 48.38   | 60.26 | 32.96| 50.67|
| ANFIS-test   | RRMSE               |                  | 0.79       | 0.72       | 0.84| 0.74    | 0.54  | 0.71    | 0.94    | 1.01  | 1.20 | 0.88|
| ANFIS-test   | R                   |                  | 0.68       | 0.80       | 0.72| 0.77    | 0.85  | 0.76    | 0.71    | 0.69  | 0.52 | 0.66|
| LSSVM-test   | MAE                 |                  | 43.38      | 31.27      | 43.12| 53.85   | 17.33 | 42.56   | 61.74   | 58.04 | 17.09| 51.31|
| LSSVM-test   | RRMSE               |                  | 0.70       | 0.61       | 0.71| 0.85    | 0.27  | 0.70    | 0.98    | 0.93  | 0.26 | 0.81|
| LSSVM-test   | R                   |                  | 0.76       | 0.81       | 0.80| 0.71    | 0.96  | 0.73    | 0.71    | 0.75  | 0.96 | 0.71|
| LSSVM-GBO test | MAE            |                  | 43.34      | 29.97      | 43.19| 52.17   | 16.99 | 42.67   | 51.91   | 57.13 | 16.93| 50.67|
| LSSVM-GBO test | RRMSE       |                  | 0.70       | 0.56       | 0.70| 0.83    | 0.26  | 0.70    | 0.89    | 0.92  | 0.25 | 0.80|
| LSSVM-GBO test | R            |                  | 0.76       | 0.82       | 0.81| 0.77    | 0.97  | 0.74    | 0.68    | 0.75  | 0.97 | 0.71|
Table 10 Results of the algorithms to the TDS modeling of the Gotvand station

| Algorithms          | Evaluation criteria | Inputs (with Q) | Ca\(^{+2}\) | Cl\(^{-1}\) | SAR | HCO\(_3\) | Sum.C | Mg\(^{+2}\) | Na\(^{+1}\) | SO\(_4\) | Sum.A | pH |
|---------------------|---------------------|-----------------|-------------|-------------|-----|-----------|-------|------------|-----------|--------|-------|----|
| ANN-train          | MAE                 |                 | 75.87       | 43.29       | 66.17 | 116.10    | 37.02 | 103.77     | 44.12     | 75.57  | 37.14 | 121.53 |
| ANN-train          | RRMSE               |                 | 0.56        | 0.32        | 0.49 | 0.93      | 0.27  | 0.85       | 0.33      | 0.54   | 0.28  | 0.93 |
| ANN-train          | R                   |                 | 0.82        | 0.94        | 0.87 | 0.35      | 0.96  | 0.51       | 0.94      | 0.83   | 0.95  | 0.35 |
| ANFIS-train        | MAE                 |                 | 85.45       | 45.49       | 62.11 | 118.18    | 35.98 | 105.11     | 45.68     | 80.19  | 37.89 | 119.74 |
| ANFIS-train        | RRMSE               |                 | 0.65        | 0.34        | 0.50 | 0.92      | 0.27  | 0.85       | 0.35      | 0.59   | 0.28  | 0.92 |
| ANFIS-train        | R                   |                 | 0.75        | 0.93        | 0.86 | 0.37      | 0.96  | 0.51       | 0.93      | 0.8    | 0.95  | 0.36 |
| LSSVM-train        | MAE                 |                 | 80.87       | 44.61       | 61.21 | 119.32    | 36.43 | 106.35     | 44.17     | 84.23  | 38.25 | 133.33 |
| LSSVM-train        | RRMSE               |                 | 0.60        | 0.34        | 0.47 | 0.94      | 0.27  | 0.86       | 0.33      | 0.60   | 0.29  | 0.99 |
| LSSVM-train        | R                   |                 | 0.79        | 0.93        | 0.87 | 0.32      | 0.96  | 0.50       | 0.94      | 0.79   | 0.95  | 0.24 |
| LSSVM-GBO train    | MAE                 |                 | 81.67       | 44.64       | 61.16 | 113.12    | 36.42 | 106.86     | 44.29     | 84.19  | 38.18 | 121.91 |
| LSSVM-GBO train    | RRMSE               |                 | 0.60        | 0.33        | 0.47 | 0.90      | 0.27  | 0.86       | 0.34      | 0.60   | 0.28  | 0.92 |
| LSSVM-GBO train    | R                   |                 | 0.79        | 0.94        | 0.88 | 0.43      | 0.96  | 0.49       | 0.94      | 0.80   | 0.96  | 0.40 |
| ANN-test           | MAE                 |                 | 131.02      | 57.30       | 91.96 | 205.02    | 39.25 | 192.06     | 94.33     | 133.08 | 35.76 | 218.16 |
| ANN-test           | RRMSE               |                 | 0.65        | 0.31        | 0.50 | 1.05      | 0.17  | 0.99       | 0.92      | 0.68   | 0.17  | 1.12 |
| ANN-test           | R                   |                 | 0.76        | 0.95        | 0.87 | 0.40      | 0.98  | 0.46       | 0.79      | 0.73   | 0.98  | 0.11 |
| ANFIS-test         | MAE                 |                 | 140.35      | 61.50       | 103.84 | 201.23   | 37.76 | 176.10     | 67.56     | 136.37 | 36.69 | 217.05 |
| ANFIS-test         | RRMSE               |                 | 0.68        | 0.32        | 0.56 | 1.04      | 0.17  | 0.92       | 0.35      | 0.70   | 0.17  | 1.12 |
| ANFIS-test         | R                   |                 | 0.74        | 0.95        | 0.86 | 0.37      | 0.98  | 0.50       | 0.94      | 0.73   | 0.98  | 0.16 |
| LSSVM-test         | MAE                 |                 | 131.60      | 54.69       | 95.79 | 202.32   | 38.19 | 178.72     | 61.36     | 142.94 | 36.26 | 227.64 |
| LSSVM-test         | RRMSE               |                 | 0.66        | 0.29        | 0.53 | 1.03      | 0.19  | 0.92       | 0.31      | 0.71   | 0.18  | 1.11 |
| LSSVM-test         | R                   |                 | 0.76        | 0.96        | 0.88 | 0.43      | 0.97  | 0.50       | 0.95      | 0.72   | 0.98  | -0.04|
| LSSVM-GBO test     | MAE                 |                 | 132.44      | 54.95       | 94.93 | 199.90   | 37.18 | 179.06     | 61.23     | 142.88 | 36.31 | 216.91 |
| LSSVM-GBO test     | RRMSE               |                 | 0.65        | 0.28        | 0.52 | 1.02      | 0.17  | 0.91       | 0.31      | 0.71   | 0.17  | 1.10 |
| LSSVM-GBO test     | R                   |                 | 0.77        | 0.97        | 0.89 | 0.41      | 0.99  | 0.51       | 0.96      | 0.73   | 0.99  | 0.16 |
| Model       | Evaluation criteria |   |   |   |   |   |   |   |   |   |   |   |   |
|------------|---------------------|---|---|---|---|---|---|---|---|---|---|---|---|
|            | MAE train           | R | MAE test                     | R | MAE train           | R | MAE train           | R | MAE train           | R | MAE train           | R |
|            | RRMSE train         |   | RRMSE test                   |   | RRMSE train         |   | RRMSE train         |   | RRMSE train         |   | RRMSE train         |   |
| Ahvaz-EC   | Ahvaz-TDS           |   | Ahvaz-TDS                     |   | Ahvaz-TDS           |   | Ahvaz-TDS           |   | Ahvaz-TDS           |   | Ahvaz-TDS           |   |
| M1         | 50.11 0.17          | 0.97 | 70.02 0.13 | 0.99 | 50.42 0.26 | 0.96 | 74.81 0.43 | 0.90 |
| M2         | 245.11 0.79         | 0.60 | 416.39 0.83 | 0.67 | 151.74 0.73 | 0.67 | 271.10 0.83 | 0.64 |
| M3         | 230.44 0.73         | 0.68 | 354.87 0.70 | 0.74 | 151.51 0.71 | 0.69 | 247.52 0.75 | 0.68 |
| M4         | 228.54 0.71         | 0.69 | 351.80 0.71 | 0.72 | 148.27 0.70 | 0.71 | 253.33 0.78 | 0.64 |
| M5         | 217.53 0.68         | 0.73 | 359.54 0.69 | 0.73 | 140.30 0.66 | 0.75 | 252.64 0.75 | 0.67 |
| M6         | 215.33 0.68         | 0.73 | 370.72 0.70 | 0.72 | 136.82 0.64 | 0.76 | 255.82 0.77 | 0.65 |
| M7         | 222.18 0.68         | 0.72 | 371.67 0.71 | 0.71 | 139.17 0.65 | 0.75 | 258.39 0.78 | 0.63 |
| M8         | 226.48 0.69         | 0.71 | 370.88 0.71 | 0.70 | 139.15 0.64 | 0.76 | 258.34 0.79 | 0.61 |
| M9         | 169.27 0.53         | 0.85 | 375.03 0.69 | 0.72 | 128.48 0.60 | 0.80 | 256.72 0.78 | 0.63 |
| M10        | 174.70 0.54         | 0.84 | 380.85 0.70 | 0.71 | 124.31 0.58 | 0.81 | 254.08 0.77 | 0.63 |
| M11        | 180.67 0.55         | 0.83 | 380.35 0.70 | 0.71 | 120.24 0.56 | 0.83 | 253.09 0.77 | 0.63 |
| M12        | 219.50 0.67         | 0.74 | 387.89 0.73 | 0.68 | 144.87 0.66 | 0.74 | 261.84 0.80 | 0.60 |
| M13        | 230.07 0.70         | 0.71 | 384.31 0.72 | 0.69 | 147.39 0.67 | 0.74 | 262.28 0.80 | 0.60 |

| Armand-EC  | Armand-TDS          |   |   |   |   |   |   |   |   |   |   |   |   |
| M1         | 24.44 0.29          | 0.95 | 26.26 0.31 | 0.95 | 17.91 0.38 | 0.92 | 17.46 0.32 | 0.95 |
| M2         | 59.94 0.71          | 0.69 | 94.99 0.99 | 0.50 | 60.46 0.93 | 0.38 | 57.26 0.92 | 0.47 |
| M3         | 52.71 0.63          | 0.77 | 93.31 1.03 | 0.48 | 62.29 0.93 | 0.37 | 57.19 0.92 | 0.46 |
| M4         | 52.09 0.64          | 0.77 | 93.67 1.04 | 0.51 | 61.74 0.93 | 0.37 | 56.22 0.91 | 0.49 |
| M5         | 46.69 0.56          | 0.83 | 94.73 1.06 | 0.46 | 61.40 0.93 | 0.37 | 69.38 0.92 | 0.51 |
| M6         | 25.96 0.30          | 0.94 | 93.80 1.04 | 0.43 | 61.35 0.93 | 0.38 | 55.20 0.90 | 0.52 |
| M7         | 11.61 0.15          | 0.99 | 89.57 1.02 | 0.48 | 60.49 0.92 | 0.39 | 54.06 0.89 | 0.54 |
| M8         | 61.59 0.71          | 0.71 | 96.34 1.06 | 0.48 | 61.40 0.92 | 0.39 | 53.73 0.88 | 0.51 |
| M9         | 61.58 0.70          | 0.71 | 92.14 1.01 | 0.47 | 62.29 0.92 | 0.40 | 53.57 0.88 | 0.49 |
| Model | Evaluation criteria |  |  |  |  |  |  |  |  |  |  |  |
|-------|------------------|---|---|---|---|---|---|---|---|---|---|---|
|       | MAE train | RRMSE train | R train | MAE test | RRMSE test | R test | MAE train | RRMSE train | R train | MAE test | RRMSE test | R test |
| M10   | 61.96      | 0.71       | 0.71       | 88.14      | 0.96       | 0.51       | 62.11      | 0.92       | 0.42       | 51.75      | 0.85       | 0.52       |
| M11   | 63.61      | 0.72       | 0.70       | 86.63      | 0.94       | 0.52       | 62.82      | 0.92       | 0.42       | 52.21      | 0.85       | 0.52       |
| M12   | 64.20      | 0.73       | 0.69       | 86.55      | 0.94       | 0.52       | 64.17      | 0.93       | 0.40       | 52.98      | 0.86       | 0.51       |
| M13   | 63.09      | 0.72       | 0.70       | 87.29      | 0.93       | 0.50       | 64.27      | 0.93       | 0.40       | 53.17      | 0.86       | 0.51       |
|       | Gotvand-EC |            |            |            |            |            | Gotvand-TDS |            |            |            |            |            |
| M1    | 35.97      | 0.16       | 0.98       | 49.86      | 0.14       | 0.99       | 35.41      | 0.26       | 0.96       | 33.86      | 0.16       | 0.99       |
| M2    | 157.38     | 0.84       | 0.53       | 300.32     | 0.94       | 0.44       | 105.07     | 0.85       | 0.52       | 190.61     | 0.97       | 0.42       |
| M3    | 156.11     | 0.84       | 0.54       | 301.95     | 0.94       | 0.43       | 104.51     | 0.84       | 0.52       | 193.74     | 0.98       | 0.41       |
| M4    | 155.98     | 0.83       | 0.54       | 307.75     | 0.95       | 0.45       | 108.56     | 0.87       | 0.48       | 197.36     | 0.98       | 0.39       |
| M5    | 150.32     | 0.82       | 0.56       | 302.90     | 0.94       | 0.46       | 100.75     | 0.82       | 0.56       | 196.47     | 0.98       | 0.42       |
| M6    | 149.46     | 0.81       | 0.58       | 298.85     | 0.93       | 0.46       | 100.62     | 0.82       | 0.57       | 193.99     | 0.95       | 0.44       |
| M7    | 145.44     | 0.79       | 0.61       | 304.51     | 0.94       | 0.46       | 97.67      | 0.80       | 0.60       | 194.14     | 0.97       | 0.44       |
| M8    | 151.48     | 0.79       | 0.61       | 306.34     | 0.95       | 0.43       | 101.54     | 0.80       | 0.59       | 192.85     | 0.97       | 0.41       |
| M9    | 157.30     | 0.81       | 0.58       | 308.54     | 0.94       | 0.41       | 104.93     | 0.82       | 0.57       | 193.34     | 0.96       | 0.40       |
| M10   | 171.59     | 0.86       | 0.50       | 314.31     | 0.96       | 0.38       | 112.90     | 0.87       | 0.48       | 196.02     | 0.98       | 0.36       |
| M11   | 147.79     | 0.77       | 0.64       | 314.99     | 0.97       | 0.37       | 112.23     | 0.87       | 0.49       | 197.47     | 0.99       | 0.35       |
| M12   | 169.71     | 0.86       | 0.51       | 311.01     | 0.96       | 0.40       | 111.05     | 0.87       | 0.49       | 194.50     | 0.98       | 0.38       |
| M13   | 157.53     | 0.61       | 0.79       | 305.38     | 0.94       | 0.40       | 105.46     | 0.81       | 0.59       | 190.87     | 0.96       | 0.40       |
3.4 EC and TDS Modeling

After detecting the best algorithm (LSSVM-GBO) and best input combination (in Ahvaz station, Sum.C, Sum.A, Na⁺¹, and Q parameters, and in Gotvand and Armand stations, Sum.C, Sum.A, Cl⁻¹, and Q parameters), in Table 11, the effect of time delay on EC and TDS modeling results is investigated. For this purpose, 13 different input combinations with time delays of 0–12 months are defined. According to the evaluation criteria demonstrated in Table 11, the LSSVM-GBO algorithm at all station in combination C1 has better accuracy than other combinations.

In Fig. 7, the modeling accuracy for the LSSVM-GBO algorithm and the EC parameter in Ahvaz, Armand, and Gotvand stations has been shown. According to Fig. 7, in the testing period, the highest and lowest accuracy is related to Gotvand (MAE = 49.86, RRMSE = 0.14, R = 0.99) and Armand stations (MAE = 26.26, RRMSE = 0.31, R = 0.95),

![Comparison of scatter plots by LSSVM-GBO algorithm.](image)

**Fig. 7** Comparison of scatter plots by LSSVM-GBO algorithm. **a** Ahvaz-train, **b** Ahvaz-test, **c** Armand-train, **d** Armand-test, **e** Gotvand-train, **f** Gotvand-test
respectively. The modeling results showed that in most cases, the difference between the estimated values and the observed values is very small. The high value of the correlation coefficient means the positive effect of GBO in improving LSSVM performance.

In Fig. 8, the modeling accuracy for the LSSVM-GBO algorithm and the TDS parameter in three stations has been shown. According to Fig. 8, in the training period, the highest and lowest accuracy is related to Gotvand (MAE = 35.41, RRMSE = 0.26, R = 0.96) and Armand stations (MAE = 17.91, RRMSE = 0.38, R = 0.92) respectively, and in the testing period, the highest and lowest accuracy is related to Gotvand (MAE = 33.86, RRMSE = 0.16, R = 0.99) and Ahvaz stations (MAE = 74.81, RRMSE = 0.43, R = 0.90), respectively. According to Fig. 8, since most of the values have accumulated around the semiconductor line, the new hybrid algorithm accuracy is high. The results indicate the significant performance of the proposed new algorithm in modeling quality parameters.

Fig. 8 Comparison of scatter plots by LSSVM-GBO algorithm. a Ahvaz-train, b Ahvaz-test, c Armand-train, d Armand-test, e Gotvand-train, f Gotvand-test
Fig. 9 The results of the EC and TDS time series model. a Ahvaz-EC, b Armand-EC, c Gotvand-EC, d Ahvaz-TDS, e Armand-TDS, f Gotvand-TDS
In Fig. 9, the results of the EC and TDS time series model based on the best input combination, best time delay, and best algorithm (LSSVM-GBO) are compared in three stations. According to these figures, the amount of EC and TDS fluctuations is well modeled by the LSSVM-GBO algorithm, which indicates the high accuracy of this algorithm.

4 Conclusion

Accurate assessment of WQ parameters is an important issue due to the impact on various environmental factors. The introduction of a new hybrid model can potentially give an effective solution in this regard. In this research, a novel LSSVM model integrated with the GBO algorithm was used to estimate EC and TDS values in the Karun river in three hydrometric stations of Gotvand, Ahvaz, and Armand. In the first step, the performance of the proposed LSSVM-GBO algorithm is compared with that of other algorithms on three real-world regression problems (Housing, LVST, Servo). The modeling results showed that LSSVM-GBO indicates the highest accuracy on three regression problems. Values of MAE, RRMSE, R in the Housing dataset were 5.30, 0.91, 0.44, respectively. Also, in the LVST dataset were 99.94, 0.20, 0.98, respectively, and in the Servo dataset were 0.46, 0.41, 0.91, respectively. Eleven input combinations including, \( \text{Ca}^{+2}, \text{Cl}^{-1}, \text{Mg}^{+2}, \text{Na}^{+1}, \text{SO}_4, \text{HCO}_3, \text{SAR}, \text{Sum.C}, \text{Sum.A}, \text{pH}, \text{Q} \) was used to model the quality parameters of EC and TDS. The modeling results based on evaluation criteria showed the most significant performance of LSSVM-GBO among all algorithms. The modeling results showed that Sum.C, Sum.A, Na\(^{+1}\), and Cl\(^{-1}\) parameters have the most noticeable impact on modeling EC and TDS parameters. Based on the results, the highest accuracy of modeling EC and TDS parameter in the Gotvand station was in input combination C1. Values of MAE, RRMSE, R in modeling EC were 49.86, 0.14, 0.99, respectively, and in modeling TDS were 33.86, 0.16, 0.99, respectively. Examination of EC and TDS time series modeled by LSSVM-GBO and observational time series showed a high correlation between modeling and observational results. The LSSVM-GBO algorithm has various advantages, such as high estimation accuracy, a balance between exploration and exploitation, and fast convergence, the ability to find a global solution, and easy implementation. The GBO does not fall into the trap of local optima due to the use of a local escaping operator. Also, the GBO uses the direction of movement term to move towards the solution. Therefore, GBO, by tuning the LSSVM parameters, can more accurately evaluate WQ parameters and other engineering problems.

Supplementary Information  The online version contains supplementary material available at https://doi.org/10.1007/s11269-021-02913-4.

Funding  The research has not been supported through any funds.

Data Availability  All data generated or used during the study are applicable if requested. This article contains supplementary information file.

Declarations

Ethics Approval  Not applicable.
Consent to Participate  The authors have made a significant contribution to this manuscript, have seen and approved the final manuscript.

Consent to Publish  The authors have agreed to publish the study in water resource management.

Conflict of Interests  The authors declare that they have no conflict of interest.

References

Acharya R, Pal J, Das D, Chaudhuri S (2019) Long-range forecast of Indian summer monsoon rainfall using an artificial neural network model. Meteorol Appl 26:347–361. https://doi.org/10.1002/met.1766
Ahmadianfar I, Bozorg-Haddad O, Chu X (2020) Gradient-based optimizer: a new metaheuristic optimization algorithm. Inf Sci 540:131–159. https://doi.org/10.1016/j.ins.2020.06.037
Alizamir M, Kisi O, Adnan RM, Kuriqi A (2020) Modelling reference evapotranspiration by combining neuro-fuzzy and evolutionary strategies. Acta Geophys 68:1113–1126. https://doi.org/10.1007/s11600-020-00446-9
Anaraki MV, Farzin S, Mousavi SF, Karami H (2021) Uncertainty analysis of climate change impacts on flood frequency by using hybrid machine learning methods. Water Resour Manag 35:199–223. https://doi.org/10.1007/s11269-020-02719-w
Aryafar A, Khosravi V, Zarepourfard H, Rooki R (2019) Evolving genetic programming and other AI-based models for estimating groundwater quality parameters of the Khezri plain. Eastern Iran Environ Earth Sci 78:69. https://doi.org/10.1007/s12665-019-8092-8
Azad A, Karami H, Farzin S, Saeedian A, Kashif H, Sayyahi F (2018) Prediction of water quality parameters using ANFIS optimized by intelligence algorithms (case study: Gorganrood river). KSCE J Civ Eng 22:2206–2213. https://doi.org/10.1007/s12205-017-1703-6
Azad A, Karami H, Farzin S, Mousavi SF, Kisi O (2019) Modeling river water quality parameters using modified adaptive neuro fuzzy inference system. Water Science and Engineering 12:45–54. https://doi.org/10.1016/j.wse.2018.11.001
Baghapour MA, Shooshtarian MR, Zarghami M (2020) Process mining approach of a new water quality index for long-term assessment under uncertainty using consensus-based fuzzy decision support system. Water Resour Manag 34:1155–1172. https://doi.org/10.1007/s11269-020-02489-5
Barzegar R, Adamowski J, Asghari Moghaddam A (2016) Application of wavelet-artificial intelligence hybrid models for water quality prediction: a case study in Aji-Chay river. Iran Stoch Environ Risk Assess 30:1797–1819. https://doi.org/10.1007/s00477-016-1213-y
Banadkooki FB, Ehteram M, Panahi F, Sammen SS, Othman FB, Ahmed E-S (2020) Estimation of total dissolved solids (TDS) using new hybrid machine learning models. J Hydrol 587:124989. https://doi.org/10.1016/j.jhydrol.2020.124989
Bisht DCS, Jangid A (2011) Discharge modelling using adaptive neuro-fuzzy inference system. IJAST Journal 31:99–114
Breiman L (2001) Random forests. Mach Learn 45:5–32. https://doi.org/10.1023/A:1010933404324
Busico G, Kazakis N, Cuoco E, Colombani N, Tedesco D, Voudouris K, Mastrocicco M (2020) A novel hybrid method of specific vulnerability to anthropogenic pollution using multivariate statistical and regression analyses. Water Res 171:115386. https://doi.org/10.1016/j.watres.2019.115386
Draa A, Bouzoubia S, Boukhalfa I (2015) A sinusoidal differential evolution algorithm for numerical optimization. Appl Soft Comput 27:99–126. https://doi.org/10.1016/j.asoc.2014.11.003
Emamgholizadeh S, Kashi H, Maroofpoor I, Zalaghi E (2014) Prediction of water quality parameters of Karoon river (Iran) by artificial intelligence-based models. Int J Environ Sci Technol 11:645–656. https://doi.org/10.1007/s13762-013-0378-x
Farzin S, Valikhan Anaraki M (2021) Modeling and predicting suspended sediment load under climate change conditions: a new hybridization strategy. J Water Climate Change. https://doi.org/10.2166/wcc.2021.317
Farzin S, Chianeh FN, Anaraki MV, Mahmoudian F (2020) Introducing a framework for modeling of drug electrochemical removal from wastewater based on data mining algorithms, scatter interpolation method, and multi criteria decision analysis (DID). J Clean Prod 266:122075. https://doi.org/10.1016/j.jclepro.2020.122075
Hadi SJ, Tombul M (2018) Streamflow forecasting using four wavelet transformation combinations approaches with data-driven models: a comparative study. Water Resour Manag 32:4661–4679. https://doi.org/10.1007/s11269-018-2077-3

Haghiabi AH, Nasrolahi AH, Parsaie A (2018) Water quality prediction using machine learning methods. Water Qual Res J 53:3–13. https://doi.org/10.2166/wqrj.2018.025

Hassan MH, Kamel S, El-Dabah MA, Rezk H (2021) A novel solution methodology based on a modified gradient-based optimizer for parameter estimation of photovoltaic models. J Electron 10:472. https://doi.org/10.3390/electronics10040472

Henriquez PA, Ruiz GA (2017) Extreme learning machine with a deterministic assignment of hidden weights in two parallel layers. Neurocomputing 226:109–116. https://doi.org/10.1016/j.neucom.2016.11.040

Ho JY et al (2019) Towards a time and cost-effective approach to water quality index class prediction. J Environ Manage 575:148–165. https://doi.org/10.1016/j.jhydrol.2019.05.016

Jang JSR (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23:665–685

Keshtegar B, Heddam S, Kisi O, Zhu SP (2019) Modeling total dissolved gas (TDG) concentration at Columbia river basin dams: high-order response surface method (H-RSM) vs M5Tree, LSSVM, and MARS, Arab J Geosci. https://doi.org/10.1007/s12517-019-4687-3

Khosravi Kh, Mao L, Kisi O, Yaseen ZM, Shahid Sh (2018) Quantifying hourly suspended sediment load using data mining models: case study of a glaciated Andean catchment in Chile. J Hydrol 567:165–179. https://doi.org/10.1016/j.jhydrol.2018.10.015

Kisi O, Azad A, Kashfi H, Saeedian A, Hashemi SAA, Ghorbani S (2019) Modeling groundwater quality parameters using hybrid neuro-fuzzy methods. Water Resour Manag 33:847–861. https://doi.org/10.1007/s11269-018-2147-6

Kumar A, Kumar P, Singh VK (2019) Evaluating different machine learning models for runoff and suspended sediment simulation. Water Resour Manag 33:1217–1231. https://doi.org/10.1007/s11269-018-2178-z

Lee KT, Hung WC, Meng CC (2008) Deterministic insight into ANN model performance for storm runoff simulation. Water Resour Manag 22:67–82. https://doi.org/10.1007/s11269-006-9144-x

Li L, Jiang P, Xu H, Lin G, Guo D, Wu H (2019) Water quality prediction based on recurrent neural network and improved evidence theory: a case study of Qiantang river, China. Environ Sci Pollut Res Int 26:19879–19896. https://doi.org/10.1007/s11356-019-05116-y

Lu H, Ma X (2020) Hybrid decision tree-based machine learning models for short-term water quality prediction. Chemosphere 249:126169. https://doi.org/10.1016/j.chemosphere.2020.126169

Melesse AM, Khosravi K, Tiefenbacher JP, Heddam S, Kim S, Mosavi A, Pham BT (2020) River water salinity prediction using hybrid machine learning models. Water 12:2951. https://doi.org/10.3390/w12102951

Milan SG, Roozbahani A, Arya Azar N, Javadi S (2021) Development of adaptive neuro fuzzy system-evolutionary algorithms hybrid models (ANFIS-EA) for prediction of optimal groundwater exploitation. J Hydrol 598:126258. https://doi.org/10.1016/j.jhydrol.2021.126258

Najafzadeh M, Ghaemi A, Emamgholizadeh S (2019) Prediction of water quality parameters using evolutionary computing-based formulations. Int J Environ Sci Technol 16:6377–6396. https://doi.org/10.1007/s13762-018-2049-4

Olorunda O, Engelbrecht AP (2008) Measuring exploration/exploitation in particle swarms using swarm diversity. IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence) 10221910:1128–1134. https://doi.org/10.1109/CEC.2008.4630938

Patel VK, Savsani VJ (2015) Heat transfer search (HTS): a novel optimization algorithm. Inf Sci 324:217–246. https://doi.org/10.1016/j.ins.2015.06.044

Salami ES, Salari M, Elhteshami M, Bidokhti NT, Ghadimi H (2016) Application of artificial neural networks and mathematical modeling for the prediction of water quality variables (case study: southwest of Iran). Desalin Water Treat 57:27073–27084. https://doi.org/10.1080/19443994.2016.1167624

Tiyasha TTM, Yaseen ZM (2020) A survey on river water quality modelling using artificial intelligence models: 2000–2020. J Hydrol 585:124670. https://doi.org/10.1016/j.jhydrol.2020.124670

Vijay S, Kamaraj K (2021) Prediction of water quality index in drinking water distribution system using activation functions based ANN. Water Resour Manag 35:535–553. https://doi.org/10.1007/s11269-020-02729-8

Yaseen ZM, Ghareb MI, Ebtelaj I, Bonakdari H, Siddique R, Heddam S, Yusif AA, Deo R (2018) Rainfall pattern forecasting using novel hybrid intelligent model based ANFIS-FFA. Water Resour Manag 32:105–122. https://doi.org/10.1007/s11269-017-1797-0

Zhang P, Yang Zh (2015) A robust AdaBoost RT based ensemble extreme learning machine. Math Problems Eng. https://doi.org/10.1155/2015/260970
Zhu B, Feng Y, Gong D, Jiang S, Zhao L, Cui N (2020) Hybrid particle swarm optimization with extreme learning machine for daily reference evapotranspiration prediction from limited climatic data. Comput Electron Agric 173:105430. https://doi.org/10.1016/j.compag.2020.105430

Zounemat-Kermani M, Seo Y, Kim S, Ghorbani MA, Samadianfard S, Naghshara S, Kim NW, Singh VP (2019) Can decomposition approaches always enhance soft computing models? Predicting the dissolved oxygen concentration in the St Johns river, Florida. Appl Sci. https://doi.org/10.3390/app9122534

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