ISEIS

Journal of Environmental Informatics 40(2) 84-94 (2022)

Journal of Environmental Informatics

www.iseis.org/jei

Dye Pollutant Removal from Synthetic Wastewater: A New Modeling and Predicting Approach Based on Experimental Data Analysis, Kriging Interpolation Method, and Computational Intelligence Techniques

M. Valikhan Anaraki[1], F. Mahmoudian[2], F. Nabizadeh Chianeh[2], and S. Farzin[1 *]

[1] *Department of Water Engineering and Hydraulic Structures, Faculty of Civil Engineering, Semnan University, Semnan, Iran*
[2] *Department of Applied Chemistry, Faculty of Chemistry, Semnan University, Semnan, Iran*

Received 11 March 2021; revised 30 May 2021; accepted 26 September 2021; published online 05 January 2022

ABSTRACT. In the present study, a new approach by coupling the interpolation method with computation-based technique (data-mining algorithms and an optimization algorithm) is introduced for modeling and optimization removal of Reactive Orange 7 (RO7) dye removal from synthetic wastewater. To this end, four significant factors like pH, electrolyte concentration, current density, and electrolysis time are considered as input variables. Thus, modeling of RO7 removal is implemented using eight data mining algorithms including multivariate linear regression (MLR), ridge regression (RR), multivariate nonlinear regression (MNLR), artificial neural network (ANN), classification and regression tree (CART), k nearest neighbor (KNN), random forest (RF), and support vector machine (SVM). These algorithms require a large data set for creating reliable results. However, creating a large number of experimental data request consuming high cost and time. Hence, the interpolation methods of kriging (KRG) and inverse distance weight (IDW) are applied for generating more data, whereas KRG has more accuracy than IDW by increasing the 47.080, 36.914, and 1.77% in *MAE*, *RMSE*, and *R* values, respectively. Then, the data mining algorithms are used for modeling the decolorization efficiency (DE) based on the original data and new data from KRG. It is found that using new data leads to significantly increasing accuracy (94.47, 96.43, 1.52, and 2.77% for *MAE*, *RMSE*, *R* and R^2, respectively) of DE modeling. Also, SVM has demonstrated the highest accuracy out of all data mining algorithms (by increasing the 97.13, 98.30, and 14.42% in *MAE*, *RMSE*, and R^2 values, respectively). Another challenge in the removal of RO7 from synthetic wastewater is predicting the maximum removal amount and optimal input variables. For this purpose, the hybrid of SVM and whale optimization algorithm (WOA) is employed. Finally, SVM-WOA has predicted the maximum of DE (91%) by optimal values of 4.2, 1.5 g/L, 4.2 mA/cm^2, and 18 min for pH, C, I, and Time, respectively. In light of the high performance of the introduced approach for modeling removal process and predicting optimal conditions of removal process, this approach can be suggested for the removal of other pollutants from wastewater when the number of experimental data set is limited.

Keywords: electrochemical decolorization, computational intelligence techniques, Kriging interpolation, Reactive Orange 7

1. Introduction

Water pollution by diverse organic compounds from various industries is an important and alarming issue worldwide. The development of the synthetic dye industry over the past years has led to increased production of colored wastewater containing toxic, refractory, and non-biodegradable organic pollutants. Furthermore, dyes' presence in effluents prevents the passage of light into the water body; Thus, it has an adverse effect on the equilibrium of the quality status of the ecological environment and human health (Moreira et al., 2017; Nidheesh et al., 2018). Therefore, the treatment of these colored wastewaters before discharge to the environment is essential.

There are several ways to treat colored wastewaters, whereas a promising next-generation technology for the conventional wastewater treatment methods can be the electrochemical advanced oxidation processes (EAOPs). According to generated various oxidizing species during the EAOPs along with its advantages such as cost-effectiveness, versatility, simplicity, amenability of automation, and environmental compatibility as well as high removal efficiency without producing any secondary pollutants, this method has become an appropriate alternative for efficient degradation of a wide range of organic pollutants. (Chianeh and Parsa, 2016; Santos et al., 2016; Siedlecka et al., 2018). Besides, various parameters such as the type and concentration of the supporting electrolyte along with the anode material significantly determine the types of reaction involved in the removal process (Yang et al., 2012; Chianeh and Parsa, 2014).

Considering the complications associated with the relationship of the above-mentioned parameters, it is difficult to explain the accurate relationship between these parameters. On the other hand, it is costly and time-consuming to study different conditions and obtain sufficient experimental data that encompasses most conditions (Wan et al., 2019). Thus, modeling and optimization of wastewater treatment processes utilizing artificial

* Corresponding author. Tel.: (+98) 9174112891; fax: (+98) 2333321005.
E-mail address: Saeed.farzin@semnan.ac.ir (S. Farzin).

ISSN: 1726-2135 print/1684-8799 online
 doi:10.3808/jei.202200473.

intelligence (AI) due to their unique advantages (like a low number of parameters, low computation time and no need for boundary and initial condition), with a focus on achieving the maximum removal efficiency of various pollutants, especially dye pollutants, has received much attention (Fan et al., 2018; Mossavi et al., 2019). Some papers investigated the performance of this method in modeling of different wastewater treatment processes such as using artificial neural network (ANN) with 54 samples for adsorption of Methyl Orange (MO) (Tanhaei et al., 2016), Acid Red 33 electrochemical decolorization modeling by ANN with 78 samples (Chianeh et al., 2017), modeling of Rhodamine B remediation from aqueous solutions via the phytoextraction method by ANN (with 154 samples) and random forest (RF) (Kooh et al., 2019), utilizing of support vector machine (SVM) with 249 samples for modeling the photocatalytic degradation of methyl tert-butyl ether (Oyehan et al., 2019), reduction of phosphorus modeling by linear regression, ANN, and M5P with 106 number of data (Kumar and Deswal, 2020), modeling of Pb^{+2} adsorption by ANN and multivariate linear regression (MLR) using 20 samples (Ashrafi et al., 2020), and adsorption Pb(II), Ni(II), and Cu(II) modeling through ANN with 476 samples (Hanandeh et al., 2021).

The strengths and weaknesses associated with using state-of-the-art optimization algorithms, employing different data mining, presenting interpretable relations, sensitivity analyzing, and using sufficient data in some studies of wastewater treatment have been outlined in Table B.1 in Appendix B. The limitation in the amount of experimental data and using classical data mining techniques for the prediction of optimal parameters are the most important drawbacks of all mentioned studies, but the modeling of proposed approaches requires a large number of data to provide a trustworthy model. In addition, as mentioned previously, increasing experimental data for the modeling process require high cost and time. In this regard, interpolating data using different methods like kriging (KRG) and inverse distance weighting (IDW) can be applied to overcome this problem. The literature review about the interpolation method is presented in Appendix C. On the other hand, in the case of the optimization of parameter values in the EAOPs modeling, the whale optimization algorithm (WOA) as a new and powerful optimization algorithm proposed by Mirjalili and Lewis (2016), can be a suitable candidate. It is worth noting that, the WOA has been inspired by the hunting behavior of the whales and also has successfully been used in different fields (Mohammadi et al., 2019; Anaraki et al., 2021).

According to our best knowledge, there is a lack of a comprehensive study of pollution modeling based on all strength points such as utilizing different data mining, presenting interpretable relations, using sensitivity analyzing for determining the best parameters of data mining, employing sufficient data for training data mining, and predicting optimal treatment conditions. It is very important to consider strength points like these during modeling processes based on experimental data such as electrochemical removal data. While only one or two of the aforementioned strengths were taken into account in most of the investigated studies.

Hence, a new approach has been developed for the first time in the current study for electrochemical removal modeling of the colored wastewater containing Reactive Orange 7 (RO7) as a dye sample pollutant. Unlike other investigated studies, this approach incorporates all of the state-of-the-art optimization algorithms, employing different data mining techniques, presenting interpretable relations, and using sufficient data. Indeed, the presented approach in this work, regarding to the insufficient experimental data, the interpolating methods were used for generating more data without consuming time and additional cost to conduct more experiments for achieving reliable models. For this reason, the performance of KRG and IDW interpolation methods is compared. Furthermore, based on the complexity of the relationships of the parameters in electrochemical removal processes, the MLR, RR, MNLR CART, RF, ANN, KNN, and SVM techniques are used to model the electrochemical removal of RO7. Subsequently, since the prediction of optimal conditions leading to the highest levels of pollutant removal is still a matter of debate, the hybrid of WOA and best-obtained modeling method is used to solve this issue. Therefore, the presented approach in this work is the combination of the KRG, a computation intelligent method based on data mining techniques (MLR, RR, MNLR, CART, RF, ANN, KNN, and SVM) and optimization algorithm (WOA).

2. Materials and Methods

2.1. Data Description

For this study, the experimental data were collected from the 31 proposed conditions by central composite design (CCD) for the RO7 removal process using Ti/MWCNT (multi-walled carbon nanotubes) anode that the experiments were carried out by Chianeh and Avestan (2020). It's worth noting that MWCNT with beneficial features such as superb mechanical strength and chemical stability and also large specific surface area have attracted great attention as the anode material for electrochemical degradation of various pollutants (Chianeh and Parsa, 2015; Duan et al., 2019; Esmaelian et al., 2019). The Ti/MWCNT anode was prepared by electrophoretic deposition (EPD) method. Besides, response surface methodology (RSM) using CCD was employed to investigate the relationship and interaction between the effective parameters (pH, electrolyte concentration (C), current density (I), and electrolysis time) as well as to achieve the optimum condition for the electrochemical decolorization of RO7. The preparation steps of the Ti/MWCNT electrode and the electrochemical removal process are depicted in Figure D.1 in Appendix D. Therefore, in this study, obtained data are used as inputs for decolorization efficiency (DE) modeling. In Table D.1 in Appendix D, the original experimental analysis data and their statistical characteritics are presented. As can be seen from Table D.1, the maximum values of average and standard deviation are related to DE with values of 35.099 and 22.354%, respectively. While the minimum of mentioned values is related to C and with values of 1.50 and 0.49 g/L. Furthermore, due to the uncertainty of some experiments (Rows 1, 6, 11, 12, 15, 19, and 23 in Table D.1), there are different decolorization efficiency values, therefore the average of the above seven data has been used. Thus, the final number of original data used in this work is 25.

2.2. Methodology

In the present study, a new approach is presented for dye pollutant removal from synthetic wastewater. This approach is based on the combination of experimental data analysis (EDA), interpolation method (IM), and computational intelligence techniques include different data mining methods and optimization algorithm. The presented approach consists of four main steps:

Processing Data: Experimental data including inputs and output are randomly divided into training (70%) period and testing (30%) period. It is worth noting that, using randomly dividing the data can avoid over fitting. Furthermore, it is evident from Table D.1 that the range of data is different. Therefore, the inputs data is scaled to [0, 1] as follows:

$$X_{scale} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{1}$$

In which, X_{scale}, X, X_{min} and X_{max} are scaled inputs data, original inputs data, minimum and maximum of inputs data, respectively. Scaling data leads to lower computation costs and the efficiency of modeling.

Generating New Data: The IDW and KRG methods can be used for generating new data from original data by interpolating. For this purpose, the following approach is used: (a) First, IDW and KRG methods are trained based on the experimental data. Then, the cross-validation method is employed for training the IDW and KRG methods (Stone, 1974; Rodriguez et al., 2010). This method involves executing IDW and KRG 25 times (number of new original experimental data), where 24 rows of data are considered for training the interoplation methods in each execution. Afterward, interpolation methods estimate the one remaining data. In the next step, the mean of evaluation criteria between DE and estimated DE during all 25 executions is considered as final evaluation criteria's in the testing period. Subsequently, the appropriate interpolation method with accurate evaluation criteria's in the testing period is selected for generating new data. (b) In this part of the mentioned approach, there is a black-box model (best-trained interpolation method) generates new DE data for each input set including pH, C, I, and time. Input sets data can be defined by the user, however, input data imposed to IDW or KRG must be in an acceptable range. Indeed, it is assumed in the present study that stated ranges are similar to original input data, and the lower and upper bounds of newly generated input data are equal to the minimum and maximum of the original input data. Then N new input data are randomly generated and are imposed to the best interpolation method. After that, the best interpolation model generates new data based on the random input data. It is worth mentioning that random inputs are generated by Equation 2. This equation generates a random number between the lower and upper bound of the original inputs data:

$$X_{new} = lb + (ub - lb) \times \vec{R} \tag{2}$$

where X_{new} is new data, lb is lower bound of experimental inputs, ub is upper bound of experimental inputs, and $\vec{R}$ is random vector. Numerous reliable studies, such as (Ehteram et al., 2018; Farzin and Anaraki, 2021; Ferdowsi et al., 2021; Karami et al., 2021), employed this equation to generate random numbers between specific lower and upper bounds. The pseudo-code of generating new data is shown in Figure 1. (c) New DE data are generated by applying the generated inputs to the best interoplation method. Thus, the new data are merged with original data to achieve more reliability.

Set the number of new data(700 data in the present study)
for i=1: number of new data
$X^{i}_{new} = lb + (ub - lb) * \vec{R}$
The interpolation model estimate new DE based on X^{i}_{new}
end
Save new DE Data

Figure 1. The pseudo-code of generating new data.

DE Modeling: The original and interpolated data are applied to data mining techniques include multivariate linear regression (MLR), ridge regression (RR), multivariate nonlinear regression (MNLR), artificial neural network (ANN), classifycation and regression tree (CART), k nearest neighbor (KNN), support vector machine (SVM), and random forest (RF). Then, each data mining technique estimates the DE. (a) Sensitivity analysis: it is important to note that each data mining technique has some parameters that can affect the final results. Accordingly, sensitivity analysis is used to determine the parameters of each algorithm. (b) Evaluation of techniques: the best data mining technique is selected based on the evaluation criteria including mean absolute error (*MAE*), root mean square error (*RMSE*), and correlation coefficient (*R*), determination coefficient (R^2).

Prediction of optimal DE: In finally, the hybrid of SVM and WOA (SVM-WOA) algorithm is employed to predict maximum DE and finding the optimum input variables (pH, C, I, and time). It is worth pointing out that the hybrid of one data-driven method and WOA uses applications of both mentioned methods without changing the accuracy of the data-driven. Indeed, to estimate the objective function (or DE) for each whale position, WOA requires a data-driven method. This data-driven method must have more accuracy than other investigated algorithms to generate reliable results for optimization DE. The scheme of the presented approach is illustrated in Figure 2. Additionally, the explanations of MLR, MNLR, RR, ANN, CART, KNN, and RF are presented in Appendix F to Appendix L.

2.2.1. Inverse Distance Weight (IDW)

IDW is an interpolation method with the following main equation: (Johnston et al., 2001):

$$Y(x) = \sum_{i=1}^{n} W_i Z_i \tag{3}$$

where W is weight of each observed value, and n is number of points around the unmeasured value. W based on the inverse of the distance between unmeasured and observed values, the W is calculated as follows:

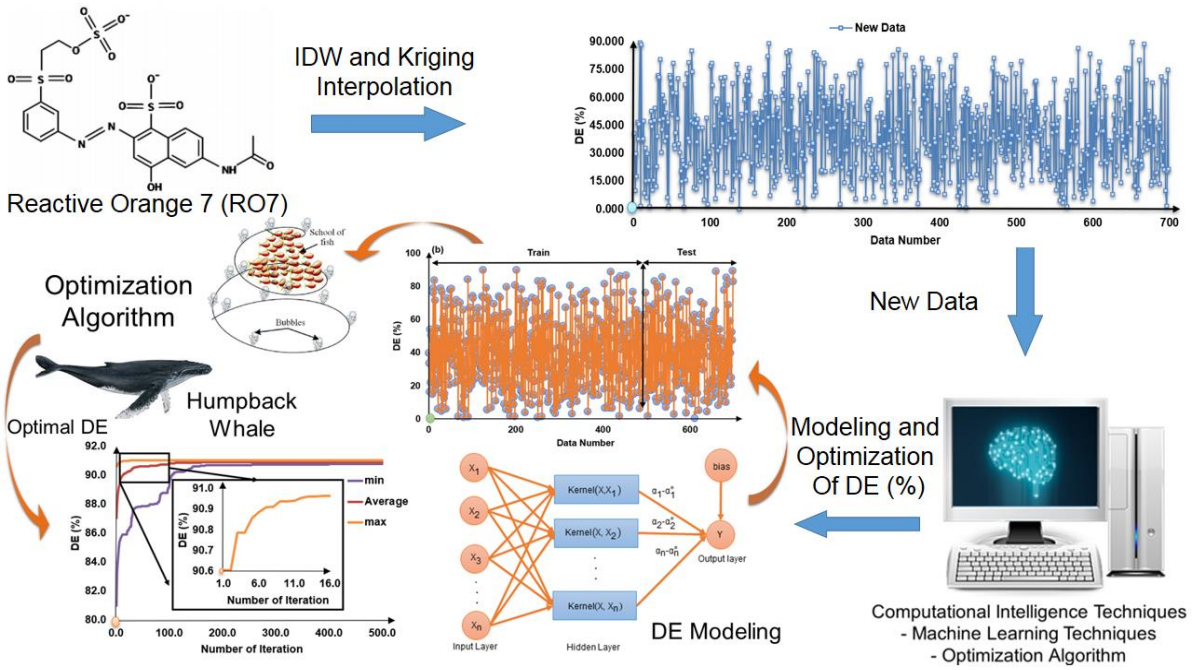

Figure 2. Presented approach.

$$W_i = \left(1/d_i\right)^p \Big/ \sum_{j=1}^{n}\left(1/d_j\right)^p \tag{4}$$

In which, d is Euclidian distance between unmeasured value and observed value. p is power that impacts weight.

2.2.2. Kriging (KRG)

One of the most known interpolation methods is KRG that estimate the unmeasured values by the following equation (Simpson et al., 2001):

$$Y(x) = \sum_{i=1}^{n} W_i f_i(x) + Z(x) \tag{5}$$

where Z is a realization of a stochastic process term with zero mean and. The spatial correlation function can be calculated as follows (Simpson et al., 2001):

$$Cov[Z(x_i), Z(x_j)] = \sigma^2 R\left(x_i, x_j\right) \tag{6}$$

In which, σ^2 is process variance and R denotes the correlation. There are different correlation functions, and Sacks et al. (1989) has presented the Gaussian correlation function, which is often used. The Gaussian correlation function is given below (Simpson et al., 2001):

$$R = \prod_{l=1}^{nx} \exp\left(-\theta_l \left(x_l^i - x_l^j\right)^2\right) \forall\ \theta_l \in R^+ \tag{7}$$

where θ and nx are hyper parameter and number of inputs, respectively. For the term of $W_i\ f_i(x)$, there are different types includes constant, linear, and quadratic that we used constant type such as Simpson et al. (2001). More details on the KRG method can be found here (Sacks et al., 1989).

2.2.3. Support Vector Machine (SVM)

SVM is a data mining technique that is developed by (Vapnik and Mukherjee, 2000). SVM is composed of only three layers (input layer, one hidden layer, and output layer). The main equation in SVM is as follows:

$$Y_{\text{mod}} = \sum_{i=1}^{n}\left(\alpha_i - \alpha_i^*\right) Kernel\left(x, x_i\right) + bias \tag{8}$$

where α_i and α_i^* are lagrangian multiplier, x is input variable, x_i is i^{th} input variable, *Kernel* is kernel function, n is number of training data. In SVM, lagrangian multiplier and bias are determined based on the quadric optimization method by minimizeing the empirical error and complexity of the algorithm. SVM use kernel function for mapping features from a higher dimension space to a lower dimension space. This matter helps SVM to consider the nonlinear relationship between inputs and outputs as linear. While there are different kernel functions, the radial basis kernel function (RBF) frequently has been used in many studies such as (Karamouz et al., 2009; Shiri et al., 2020). Therefore, in the current study, the RBF kernel function is used. The RBF formula is as follows:

$$Kernel = \exp\left(-\left\|x - x_i\right\|^2 \Big/ 2\sigma^2\right) \tag{9}$$

In which, σ is kernel width. As mentioned above, SVM used a global optimization algorithm in its structure. In this sense, SVM is not trapped in local optimum like ANN, which used a gradient-based optimization algorithm. Besides, since the SVM is composed of three layers, there is no need to determine the structure of this algorithm. Another beneficial feature of the SVM is it has only two parameters to execute, which include cost coefficient or C (is refered to minimizing the empirical error), and kernel width or gamma. It should also be noted that C and gamma must be determined carefully because both parameters have high effects on the final performance of SVM. For more details see (Chen et al., 2019; Cui et al., 2021). The structure of SVM is shown in Figure 3.

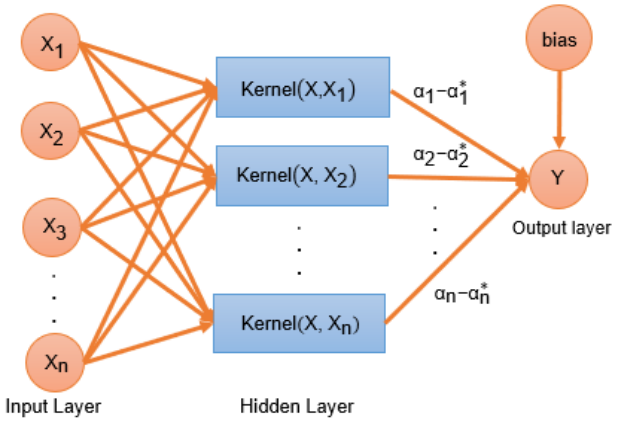

Figure 3. The SVM structure.

2.2.4. Whale Optimization Algorithm (WOA)

WOA is one of the state-of-the-art optimization algorithms that is introduced by (Mirjalili et al., 2016). This algorithm is inspired by the hunting behavior of humpback whales. WOA is comprised of three operators including Encircling Prey, Bubble-Net Attacking Method, Search for prey, these operators are explained as follows:

Encircling Prey: In this operator, the humpback whale moves to the prey position. In WOA, the best position of whales obtained so far has been considered the prey position. This behavior is down as follows (Mirjalili and Lewis, 2016):

$$\vec{D} = \left|\vec{C}\vec{X}^{*} - \vec{X}\left(t\right)\right| \tag{10}$$

$$\vec{X}\left(t+1\right) = \vec{X}^{*} - \vec{A}.\vec{D} \tag{11}$$

where $\vec{D}$ and $\vec{A}$ are coefficient vectors and $\vec{X}^{*}$, $\vec{X}(t)$ and $\vec{X}(t+1)$ are best obtained position so far, current position and new position of whales, respectively. The $\vec{A}$ and $\vec{C}$ vectors are computed as follows (Mirjalili and Lewis, 2016):

$$\vec{A} = 2\vec{a}\vec{r} - \vec{a} \tag{12}$$

$$\vec{C} = 2\vec{r} \tag{13}$$

where $\vec{a}$ linearly decreases from 2 to 0 by increasing the number of iteration and $\vec{r}$ is a random vector between 0 to 1.

Bubble-Net Attacking Method: In nature, whales swim to prey and create the shrinking spiral patch shape around prey as simultaneously. Since these two movements are simultaneous, WOA considers the same probability for each one of the movements. This behavior is down as follows (Mirjalili and Lewis, 2016):

$$\vec{X}\left(t+1\right) = \begin{cases} \vec{X}^{*} - \vec{A}.\vec{D} & if\ P < 0.5 \\ \vec{D}'e^{bl}.\cos(2\pi l) + \vec{X}^{*} & if\ P \geq 0.5 \end{cases} \tag{14}$$

In which, $\vec{D}' = \left|\vec{X}^{*} - \vec{X}\left(t\right)\right|$ and b are constant values for defining the logarithmic shape of spiral, l is a random number in the range of -1 to 1, and P is a random number.

Search for Prey: In this behavior, the position of whales is updated based on the random position as follows (Mirjalili and Lewis, 2016):

$$\vec{D} = \left|\vec{C}\vec{X}_{rand} - \vec{X}\left(t\right)\right| \tag{15}$$

$$\vec{X}\left(t+1\right) = \vec{X}_{rand} - \vec{A}.\vec{D} \tag{16}$$

In which, $\vec{X}_{rand}$ is a random position from search space. In the present study, $\vec{X}(t)$ is input data including pH, C, I, and time and the best position is the position of the whale with maximum DE. Indeed, each whale in WOA has one position (pH, C, I, and time) and also has one objective function (value of DE). However, WOA has not able to estimate DE for each position of the whale. For this reason, there is a need to estimate the DE using one modeling method. Hence, in this study, the objective function, or in other words, the value of DE, is estimated using the SVM method as a more accurate model.

```
Fit the SVM model based on the pH, C, I, and Time data as inputs
and DE data as a target
Set initial parameters of WOA and generate an initial population
Run obtained SVM model for each Whale and calculate objective
function (DE) for that
Select the best population with maximized DE
for It=1:Maximum Iteration
    for n=1:Search agent number
        update position of whale (pH, C, I, and Time) by equations
        14-20
        Run SVM and predict DE (objective function)
        Update the best position of whales obtained so far
    end
end
Return maximize DE and correspond solution (pH, C, I, and Time)
```

Figure 4. The pseudo-code of SVM-WOA.

2.2.5. Hybrid of SVM and WOA (SVM-WOA)

In the present study, the SVM-WOA is employed for predicting optimal DE. In this method, mentioned parameters namely pH, C, I, and time are considered as decision variables, as well as the DE value is considered as the objective function. In SVM-

WOA, the SVM is first trained based on the inputs including pH, C, I, and Time along with DE as the target data. Following, the trained model is used to predict the DE for each input set. In the next step, WOA generates the initial population, which each search agent in this population has one position (pH, C, I, and time) and one objective function (estimated DE by SVM). In this regard, the position of each whale is the input set of SVM, and trained SVM predicts the DE for this input set. After generating the initial population, the position of whales is updated based on the equations 14 ~ 20 to achieve maximum DE and correspond input data. The pseudo-code of SVM-WOA for predicting optimal DE is demonstrated in Figure 4.

2.2.6. Evaluation Criteria

In this study, the *MAE*, *RMSE*, *R*, and R^2 (Kvålseth, 1985; Kasuya, 2019) are used to evaluate the interpolation and data mining techniques. These evaluation criteria are frequently employed by many studies such as (Azad et al., 2019a, b; Valikhan-Anaraki et al., 2019). The formula of *MAE*, *RMSE*, and R^2 criteria are presented in Appendix E.

3. Results and Discussion

3.1. Sensitivity Analysis of Data

The Pearson correlation coefficient is used as a way to measure the influence of input variables on output variables (DE) as well as input variables on each other as shown in Figure 5. According to the results of this Figure, input variables have no correlation with each other which leads to efficient modeling because of input variables are independent of each other. As seen in Figure 5, the magnitude of the correlation between input and output variables is varied from 0.07 to 0.63. It is worth noting that, the maximum correlation magnitude is related to Time, whereas the minimum correlation magnitude is related to C (g/L). In addition, the negative and positive correlations show the inverse and direct relationships between inputs and DE efficiency, respectively. With this in mind, by increasing all input variables except pH, enhancement of DE efficiency can be achieved. In contrast, reducing the pH value can lead to an increase in the DE efficiency as a target. Besides, sensitivity analysis of input variables in Appendix M (Table M.1) shows that pH, C, I, and Time are the best input combination for modeling DE.

3.2. Sensitivity Analysis of Proposed Techniques

Interpolation and data mining techniques have some parameters for executing that affect their final result significantly. Hence, determining these parameters as carefully is an inevitable issue. In the present study, the best parameters of interpolation and data mining techniques include IDW, KRG, RR, ANN, KNN, SVM, and RF are obtained by sensitivity analysis. In this process, different values are assigned to considered parameters, then mentioned techniques are executed based on the assigned values. Afterward, RMSE is measured for each assigned value, and the best parameter value is selected in each method. The sensitivity analysis of IDW, KRG, RR, ANN, KNN, SVM, and RF are presented in Appendix N (Figures N.1 ~ N.5). As seen, for IDW, the best value of the power parameter is equal to 28, and for KRG, the best value of $tetha_0$ is 0.037. Also, the best value of Gamma in RR for original data is equal to 0.02 and for new data equal to 0.01. Additionally, the best value of RMSE of original data in ANN is related to lbfgs learning algorithm with 6 and 18 neurons for first and second layers, whereas the minimum RMSE for new data is obtained from lbfgs learning algorithm with 20 and 16 neurons in first and second layers, respectively. Furthermore, in KNN, the minimum value of RMSE is related to neighbor number 6 for both original and new data. In SVM, the best value of C for both original and new data is 1,000, and the best value of Gamma for original and new data is 0.2 and 1.2, respectively.

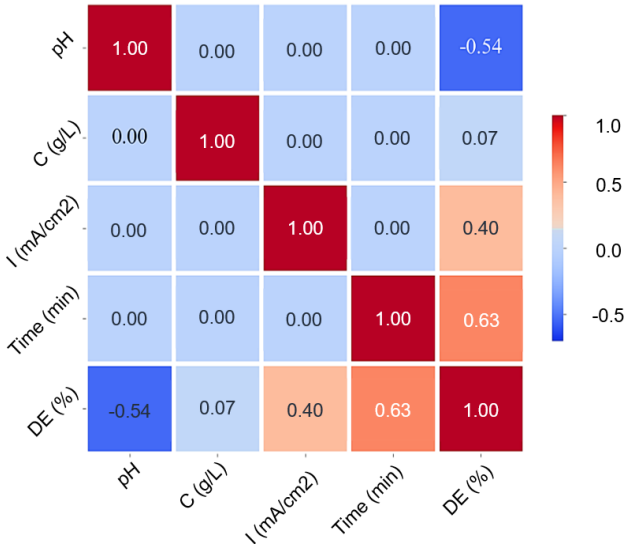

Figure 5. Sensitivity analysis of data.

3.3. Obtained Interpretable Relation for MLR, RR, and MNLR

In this section, the linear relations obtained by MLR, and RR as well as nonlinear relation obtained by MNLR are presented. Relations 17 and 18 are related to MLR for original and new data, respectively. Relations 19 and 20 are obtained by RR and are related to original data and new data, respectively. The nonlinear relations 21 and 22 are computed by MNLR for new data and original data, respectively:

$$DE = 43.449 - 41.151 \times pH - 1.098 \times C + 21.996 \times I + 25.825 \times Time \quad (17)$$

$$DE = 39.457 - 50.702 \times pH - 4.244 \times C + 25.545 \times I + 28.480 \times Time \quad (18)$$

$$DE = 42.429 - 38.965 \times pH - 0.590 \times C + 21.245 \times I + 25.053 \times Time \quad (19)$$

$$DE = 39.457 - 50.690 \times pH - 4.243 \times C + 25.539 \times I + 28.473 \times Time \quad (20)$$

$$DE = -143.275 + 12.849 \times pH + 52.299 \times C + 29.473 \times I + 10.637 \times Time - 1.411 \times pH \times C - 1.663 \times pH \times I - 0.379 \times pH \times Time + 4.760 \times C \times I - 1.305 \times C \times Time - 0.355 \times I \times Time - 0.546 \times pH^2 - 13.952 \times C^2 - 2.201 \times I^2 - 0.115 \times Time^2 \quad (21)$$

$$DE = -154.819 + 14.649 \times pH + 30.985 \times C + 35.302 \times I + 13.168 \times Time + 1.179 \times pH \times C - 2.699 \times pH \times I - 0.853 \times pH \times Time + 5.053 \times C \times I - 0.744 \times C \times Time - 0.104 \times I \times Time - 0.438 \times pH^2 - 14.124 \times C^2 - 2.428 \times I^2 - 0.169 \times Time^2 \quad (22)$$

3.4. Selecting the Best Interpolation Method

In this section, to determine the best interpolation method between the IDW and KRG, different evaluation criteria include *MAE*, *RMSE*, and R^2 have been used which the obtained results are compared in Table 1. A comparison of the results reveals that the KRG leads to better results than IDW with 47.080, 36.914, and 3.415% more accurate values of *MAE*, *RMSE*, and R^2, respectively. Although both IDW and KRG methods use surrounding points to obtain unmeasured points, KRG generates more accurate results using Gaussian correlation function for computing unmeasured values instead of inverse distance.

Table 1. The Interpolation Method Accuracy

Method	*MAE*	*RMSE*	R^2
IDW	12.485	14.550	0.905
KRG	6.607	9.179	0.921
KRG vs. IDW (%)	-47.080	-36.914	3.415

Table 2. Generated Data by KRG

Row	pH	C (g/L)	I (mA/cm^2)	Time (min)	DE (%)
1	8.346	0.848	3.117	13.969	40.833
2	10.600	2.171	4.066	12.699	9.693
3	9.342	1.587	4.806	18.174	29.625
4	9.535	2.293	3.067	5.475	17.553
5	4.391	0.946	2.567	9.863	46.817
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
695	6.561	1.390	2.781	11.677	51.117
696	10.631	2.224	4.091	18.385	1.544
697	9.532	0.922	3.918	9.434	20.608
698	3.621	0.921	3.337	14.865	74.709
699	9.178	2.001	3.136	5.159	21.364
700	3.356	0.673	2.393	7.652	31.889
Min	3.000	0.500	1.000	0.000	0.000
Max	11.000	2.500	5.000	20.000	89.756
Average	6.972	1.500	3.046	12.456	39.200
Standard deviation	2.300	0.567	1.170	4.294	21.365

3.5. Generating New Data by KRG

Generated new data by the KRG method (best-interpolated method) is tabulated in Table 2. According to the results of this table, the range of new data is equal to the original data which is due to the combination of the original and interpolated data. Besides, the maximum values of average and standard deviation of new data are related to DE, and the minimum value of the mentioned parameters is related to C, which is similar to the original data in Table D.1. Moreover, the average and standard deviation of the new data are close to those of the original data. Therefore, using KRG for generating new data maintains the distribution of original data. This matter shows the acceptable results of KRG that are good for the training of KRG.

In the following, Figure 6 illustrates the new DE data which is scattered over a range of 0 to 81.120. Based on this Figure, the dispersion of new data of DE is high, therefore it is expected to linear data mining techniques have results with less accurate than nonlinear data mining techniques.

3.6. Computational Intelligence Techniques Results

The results of comparing data mining techniques for original and new data are set out in Table 3. It is apparent from this Table that the best results for original data are related to the ANN technique (by considering *MAE*, *RMSE*, *R* and R^2 criteria). In terms of accuracy in DE modeling, other techniques included SVM, RR, and MLR are also ranked second, third, and fourth, respectively. While, SVM has the first rank (based on *MAE*, *RMSE*, *R* and R^2 criteria) for new data, followed by ANN and MNLR. Moreover, the results of RR and MLR with a small difference are the weakest. The *MAE*, *RMSE*, *R* and R^2 for SVM than RR and MLR are more accurate up to 97.13, 98.30, 14.42 and 50.83%, respectively. In Table 3, by comparing the original and new data results, it can be concluded that generating new data in all data mining techniques except linear technique (MLR and RR) leads to increased accuracy of DE modeling. For instance, using the best data mining technique obtained for new data (SVM) compared to using the ANN as the best data mining technique obtained for original data has improved *MAE*, *RMSE*, *R* and R^2 by 95.764, 77.395, 1.52 and 2.77%, respectively. This improvement of accuracy can be attributed to the increasing number of data, as well as enhancement of oscillation and nonlinearity of data (Figure 6). The increasing number of data leads to better training of data mining techniques, but on the other hand, the oscillation of data and complexity of modeling increase. Therefore, nonlinear data mining techniques such as SVM and MNLR by increasing the number of data estimate the nonlinearity relationship between inputs and outputs with good accuracy. However, MLR and RR are failed to estimate the nonlinearity relationship between input and output for new data. Furthermore, the ANN algorithm has high performance for original and new data that it is for processing data in multilayers and the ability for solving linear and nonlinear problems. The superiority of SVM over ANN, MNLR, KNN, RF and CART is using the global optimization method (quadratic programming) for finding their weights of inputs, unlike ANN that used gradient-based optimization methods. In addition, using the Kernel function helps SVM consider nonlinear relationship between

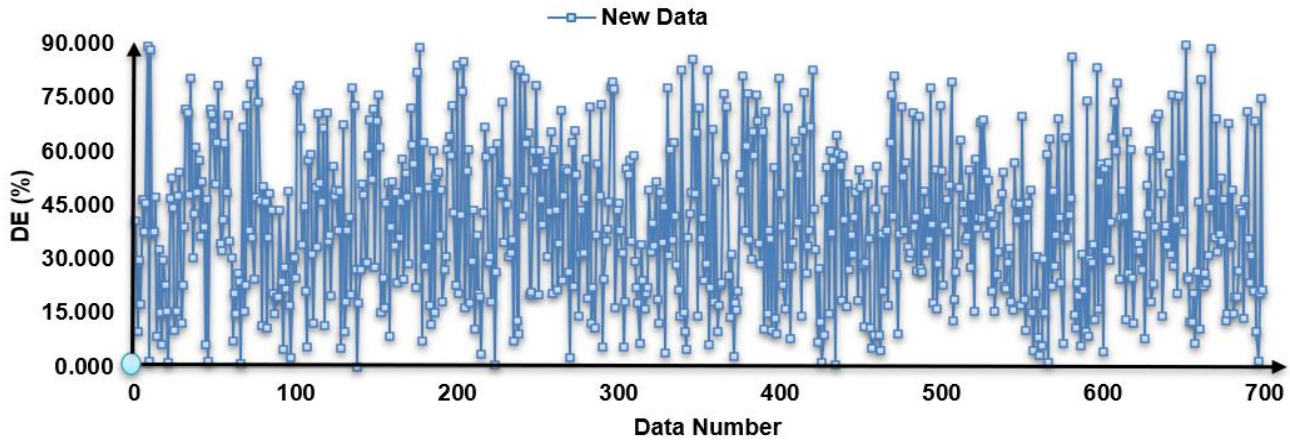

Figure 6. The variation of DE data.

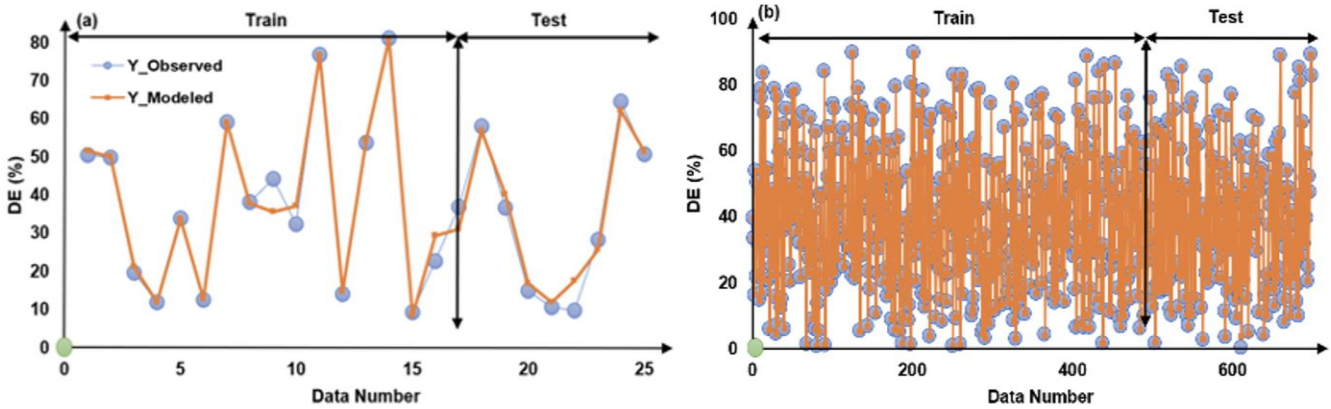

Figure 7. Comparing best data mining technique and observed data for (a) Original data and ANN output and (b) New data and SVM output.

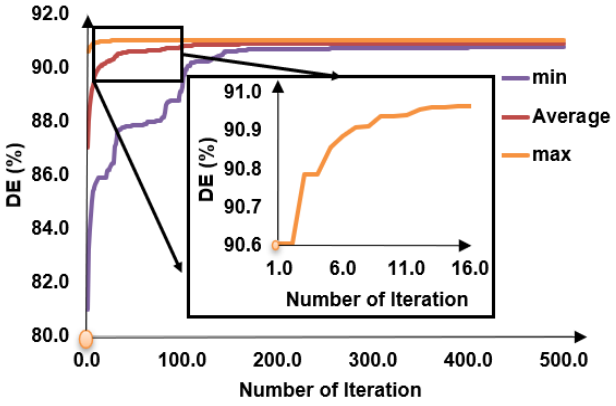

Figure 8. Convergence curve of minimum, average, and maximum for optimized DE.

inputs and output as a linear relationship. Thus, the superiority of MNLR over KNN and RF is to consider the nonlinear relationships between inputs and outputs using the quadratic of inputs and interaction of inputs. According to other results of Table 3, KNN and RF have close results and are better than CART, which mainly due to the nonlinear structure of KNN using similarity of neighbors in this technique as well as the use of multi-tree and multi times random resampling the dataset in RF. Also, the better accuracy of CART than MLR and RR is related to partitioning data into subsets that help to solve nonlinear problems.

In the following, the observed and modeled DE for original (by ANN) and new data (by SVM) are illustrated in Figure 7. From this Figure, the minimum and maximum values of DE are modeled with good accuracy. Besides, the accuracy of DE modeling for new data is more than for the original data, as evidenced in Table 3.

3.7. Determination of Optimal Conditions and Validation of the Model

In finally, the SVM-WOA is applied for predicting the optimal value of input variables (pH, C, I, and Time) and the maximum value of DE. For this purpose, the population size and the maximum number of iteration are set to 50 and 500 based on the study of (Mirjalili and Lewis, 2016). Thus, to reduce the uncertainty of the SVM-WOA results, it is run 15 times. Figure 8 demonstrates the convergence curve of minimum, average, and maximum results for 15 random runs. It can be seen that

Table 3. The Data Mining Results

Data	Algorithm	*MAE*		*RMSE*		*R*		R^2	
		Train	Test	Train	Test	Train	Test	Train	Test
Original	MLR	8.416	9.659	7.336	7.910	0.919	0.909	0.845	0.785
	RR	8.459	9.597	7.339	7.864	0.919	0.915	0.843	0.794
	MNLR	1.810	8.313	2.310	10.244	0.994	0.912	0.988	0.492
	ANN	3.281	5.370	1.924	4.056	0.988	0.985	0.976	0.973
	CART	0.000	19.793	0.000	17.077	1.000	0.569	1.000	0.015
	KNN	13.388	14.362	11.667	12.450	0.852	0.938	0.607	0.673
	SVM	2.512	5.207	1.061	4.073	0.993	0.975	0.986	0.969
	RF	5.442	17.724	4.154	16.023	0.980	0.726	0.935	0.343
New Data	MLR	10.566	10.346	8.536	8.530	0.870	0.874	0.730	0.663
	RR	10.566	10.347	8.536	8.530	0.870	0.874	0.757	0.761
	MNLR	2.974	3.064	3.909	4.228	0.984	0.978	0.967	0.956
	ANN	0.628	1.399	0.492	1.040	1.000	0.998	0.999	0.996
	CART	0.000	8.314	0.000	6.332	1.000	0.922	1.000	0.909
	KNN	4.227	5.049	3.385	3.726	0.982	0.976	0.961	0.943
	SVM	0.103	0.297	0.085	0.145	1.000	1.000	1.000	1.000
	RF	1.750	5.084	1.343	3.732	0.997	0.973	0.993	0.942
Best for new data vs. Best for original data (%)		-96.86	-94.47	-95.58	-96.43	1.21	1.52	2.45	2.77

Table 4. The SVM-WOA Optimization Results

Iteration number	Run Number											
	1	2	3	4	5	…	13	14	15	min	average	max
1	89.8	82.7	86.7	86.4	89.4	…	84.0	90.6	88.7	81.0	87.0	90.6
2	89.8	83.8	89.0	88.8	89.4	…	87.3	90.6	90.2	83.2	88.0	90.6
3	90.8	84.4	89.5	89.9	90.0	…	88.3	90.8	90.2	83.9	88.6	90.8
4	90.8	84.5	89.9	90.4	90.2	…	88.6	90.8	90.3	84.5	88.9	90.8
5	90.9	85.3	89.9	90.6	90.2	…	89.1	90.8	90.4	85.0	89.2	90.9
6	90.9	85.4	89.9	90.6	90.3	…	89.8	90.8	90.5	85.4	89.3	90.9
7	90.9	85.5	90.0	90.6	90.4	…	89.9	90.9	90.7	85.5	89.5	90.9
8	90.9	85.5	90.0	90.7	90.4	…	90.5	90.9	90.7	85.5	89.6	90.9
9	90.9	85.6	90.0	90.7	90.4	…	90.5	90.9	90.7	85.6	89.7	90.9
10	90.9	85.7	90.0	90.8	90.4	…	90.5	90.9	90.7	85.7	89.8	90.9
11	90.9	85.8	90.0	90.8	90.4	…	90.5	90.9	90.8	85.8	89.9	90.9
12	91.0	85.9	90.0	90.8	90.4	…	90.5	90.9	90.8	85.9	90.0	91.0
13	91.0	85.9	90.1	90.8	90.4	…	90.5	90.9	90.8	85.9	90.0	91.0
14	91.0	85.9	90.1	90.9	90.4	…	90.5	90.9	90.8	85.9	90.1	91.0
15	91.0	85.9	90.1	90.9	90.4	…	90.5	90.9	90.8	85.9	90.1	91.0
.	.	.	.	.	.	…	.	.	.	.	.	.
.	.	.	.	.	.		.	.	.	.	.	.
.	.	.	.	.	.		.	.	.	.	.	.
500	91.0	90.8	91.0	91.0	90.8	…	91.0	91.0	91.0	90.8	90.9	91.0

the results converged to a DE value of 91% with a maximum of 15 iterations.

In Table 4, the DE for 15 random runs of SVM-WOA are presented. According to the results, the difference between minimum (DE = 90.8%), average (DE = 90.9%), and maximum DE (DE = 91%) of 15 random runs is very Low (up to 0.2%), which shows the high performance and reliable results of SVM-WOA algorithm. Afterward, in order to validate the statistical optimization results, a confirmation experiment was done by setting the operational parameters at their optimal values as tabulated in Table 5. The removal efficiency of RO7 (DE%) was 88.7%, which had good agreement with predicted values by the SVM-WOA model. Therefore, this closeness proved that the proposed model could describe accurately the electrochemical removal of RO7 on Ti/MWCNT anode as well as the used prediction model was appropriate to optimize the operational parameters.

Table 5. Predicted Optimal Conditions and Observed DE%

pH	NaCl (g/L)	I (mA/cm^2)	Time (min)	Predicted DE%	Observed DE%
4.2	1.5	4.2	18	91%	88.7 %

4. Conclusions

In the present study, the interpolation methods with computation intelligent techniques include data mining algorithms and an optimization algorithm were combined for modeling and optimization of the electrochemical removal process. In this regard, IDW and KRG were used for generating new data. Then, data mining techniques include MLR, RR, MNLR, ANN, CART, KNN, SVM, and RF were employed for modeling the electrochemical removal of RO7 as a case study. Subsequently, in order to predict optimal DE, the hybrid of SVM and WOA was utilized. The main findings of the present study are as follows:

(1) KRG had better results than IDW based on the *MAE*, *RMSE*, and R^2 (by 36.914% more accurate *RMSE*). Hence, this method was used for generating new data.

(2) Generating new data leads to enhance DE modeling accuracy compared with using original data with increases of 95.764, 77.395, 1.52 and 2.77% in *MAE*, *RMSE*, *R* and R^2 values, respectively.

(3) The SVM provided more accurate results than other data mining techniques with *MAE*, *RMSE*, and R^2 up to 97.13, 98.30, and 14.42%, respectively.

(4) The interpretable relationship between DE and pH, C, I, and time was established.

(5) The DE optimum value of 91% was obtained through SVM-WOA. The standard deviation of 15 random runs and the difference between the minimum and maximum results of SVM-WOA was low. According to SVM-WOA, the best solution was 4.2, 1.5 g/L, 4.2 mA /cm^2, and 18 min for pH, C, I, and time, respectively.

The proposed approach had more reliable results and higher accuracy, without the requiring of conducting additional experiments to obtain sufficient data. Therefore, this proposed approach has a high potential to achieve the maximum electrochemical removal of RO7 as a model pollutant.

References

Anaraki, M.V., Farzin, S., Mousavi, S.F. and Karami, H. (2021). Uncertainty Analysis of Climate Change Impacts on Flood Frequency by Using Hybrid Machine Learning Methods. *Water Resour. Manag.*, 35(1), 199-223. https://doi.org/10.1007/s11269-020-02719-w

Ashrafi, M., Borzuie, H., Bagherian, G., Chamjangali, M.A. and Nikoofard, H. (2020). Artificial neural network and multiple linear regression for modeling sorption of Pb^{2+} ions from aqueous solutions onto modified walnut shell. *Sep. Sci. Technol.*, 55(2), 222-233. https://doi.org/10.1080/01496395.2019.1577437

Azad, A., Karami, H., Farzin, S., Mousavi, S.F. and Kisi, O. (2019). Modeling river water quality parameters using modified adaptive neuro fuzzy inference system. *Water Sci. Eng.*, 12(1), 45-54. https://doi.org/10.1016/j.wse.2018.11.001

Azad, A., Manoochehri, M., Kashi, H., Farzin, S., Karami, H., Nourani, V. and Shiri, J. (2019). Comparative evaluation of intelligent algorithms to improve adaptive neuro-fuzzy inference system performance in precipitation modelling. *J. Hydrol.*, 571, 214-224. https://doi.org/10.1016/j.jhydrol.2019.01.062

Chen, J.L., Yang, H., Lv, M.Q., Xiao, Z.L. and Wu, S.J. (2019). Estimation of monthly pan evaporation using support vector machine in Three Gorges Reservoir Area, China. *Theor. Appl. Climatol.*, 138(1), 1095-1107. https://doi.org/10.1007/s00704-019-02871-3

Chianeh, F.N. and Avestan, M.S. (2020). Application of central composite design for electrochemical oxidation of reactive dye on Ti/MWCNT electrode. *J. Iran. Chem. Soc.*, 17(5), 1073-1085. https://doi.org/10.1007/s13738-019-01834-9

Chianeh, F.N. and Parsa, J.B. (2014). Degradation of azo dye from aqueous solutions using nano-SnO_2/Ti electrode prepared by electrophoretic deposition method: Experimental Design. *Chem. Eng. Res. Des.*, 92(11), 2740-2748. https://doi.org/10.1016/j.cherd.2014.03.001

Chianeh, F.N. and Parsa, J.B. (2015). Degradation of azo dye in aqueous solution using Ti anode coated with MWCNTs-TiO_2. *J. Iran. Chem. Soc.*, 12(1), 175-182. https://doi.org/10.1007/s13738-014-0471-1

Chianeh, F.N. and Parsa, J.B. (2016). Decolorization of azo dye C.I. Acid Red 33 from aqueous solutions by anodic oxidation on MWCNTs/Ti electrodes. *Desalin. Water Treat.*, 57(43), 20574-20581. https://doi.org/10.1080/19443994.2015.1110716

Chianeh, F.N., Parsa, J.B. and Vahidian, H.R. (2017). Artificial neural network modeling for removal of azo dye from aqueous solutions by Ti anode coated with multiwall carbon nanotubes. *Environ. Prog. Sustain. Energy*, 36(6), 1778-1784. https://doi.org/10.1002/ep.126

Cui, X., Goff, T., Cui, S., Menefee, D., Wu, Q., Rajan, N., Nair, S., Phillips, N. and Walker, F. (2021). Predicting carbon and water vapor fluxes using machine learning and novel feature ranking algorithms. *Sci. Total Environ.*, 775, 145130. https://doi.org/10.1016/j.scitotenv.2021.145130

Duan, P.Z., Gao, S.H., Li, X., Sun, Z.R. and Hu, X. (2019). Preparation of CeO_2-ZrO_2 and titanium dioxide coated carbon nanotube electrode for electrochemical degradation of ceftazidime from aqueous solution. *J. Electroanal. Chem.*, 841, 10-20. https://doi.or-g/10.1016/j.jelechem.2019.03.061

Esmaelian, M., Chianeh, F.N. and Asghari, A. (2019). Degradation of ciprofloxacin using electrochemical oxidation by Ti/nanoSnO_2-MWCNT electrode: Optimization and modelling through central composite design. *J. Ind. Eng. Chem.*, 78, 97-105. https://doi.org-/10.1016/j.jiec.2019.06.032

Ehteram, M., Karami, H., Mousavi, S.F., Farzin, S., Celeste, A.B. and Shafie, A.E. (2018). Reservoir operation by a new evolutionary algorithm: Kidney Algorithm. *Water Resour. Manag.*, 32(14), 4681-4706. https://doi.org/10.1007/s11269-018-2078-2

Fan, M.Y., Hu, J.W., Cao, R.S., Ruan, W.Q. and Wei, X.H. (2018). A review on experimental design for pollutants removal in water treatment with the aid of artificial intelligence. *Chemosphere*, 200, 330-343. https://doi.org/10.1016/j.chemosphere.2018.02.111

Farzin, S. and Anaraki, M.V. (2021). Optimal construction of an open channel by considering different conditions and uncertainty: application of evolutionary methods. *Eng. Optim.*, 53(7), 1173-1191. https://doi.org/10.1080/0305215X.2020.1775825

Ferdowsi, A., Valikhan-Anaraki, M., Mousavi, S.F., Farzin, S. and Mir- jalili, S. (2021). Developing a model for multi-objective optimiza- tion of open channels and labyrinth weirs: Theory and application in Isfahan Irrigation Networks. *Flow Meas. Instrum.*, 80, 101971. https://doi.org/10.1016/j.flowmeasinst.2021.101971

Hanandeh, A.E., Mahdi, Z. and Imtiaz, M.S. (2021). Modelling of the adsorption of Pb, Cu and Ni ions from single and multi-component aqueous solutions by date seed derived biochar: Comparison of six machine learning approaches. *Environ. Res.*, 192, 110338. https://-doi.org/10.1016/j.envres.2020.110338

Johnston, K., Ver Hoef, J.M., Krivoruchko, K. and Lucas, N. (2001). *Using ArcGIS geostatistical analyst.* Esri Redlands, pp 1-48.

Karami, H., Anaraki, M.V., Farzin, S. and Mirjalili, S. (2021). Flow direction algorithm (FDA): a novel optimization approach for solving optimization problems. *Comput. Ind. Eng.*, 156(4), 107224. https://

doi.org/10.1016/j.cie.2021.107224

Karamouz, M., Ahmadi, A. and Moridi, A. (2009). Probabilistic reservoir operation using Bayesian stochastic model and support vector machine. *Adv. Water Resour.*, 32(11), 1588-1600. https://doi.org/10.1016/j.advwatres.2009.08.003

Kasuya, E. (2019). On the use of r and r squared in correlation and regression. *Ecol.Res.*, 34(1), 235-236. https://doi.org/10.1111/144-0-1703.1011

Kooh, M.R.R., Dahri, M.K., Lim, L.B.L., Lim, L.H. and Lee, S.L. (2019). Phytoextraction capability of Azolla pinnata in the removal of rhodamine B from aqueous solution: artificial neural network and random forests approaches. *Appl. Water Sci.*, 9(4), 80. https://doi.org/10.1007/s13201-019-0960-6

Kumar, S. and Deswal, S. (2020). Phytoremediation capabilities of Salvinia molesta, water hyacinth, water lettuce, and duckweed to reduce phosphorus in rice mill wastewater. *Int. J. Phytoremediat.*, 22 (11), 1097-1109. https://doi.org/10.1080/15226514.2020.173172 9

Kvålseth, T.O. (1985). Cautionary note about R^2. *Am. Stat.*, 39(4), 279-285. https://doi.org/10.1080/00031305.1985.10479448

Mirjalili, S. and Lewis, A. (2016). The Whale Optimization Algorithm. *Adv. Eng. Softw.*, 95, 51-67. https://doi.org/10.1016/j.advengsoft.2016.01.008

Mohammadi, M., Farzin, S., Mousavi, S.F. and Karami, H. (2019). Investigation of a new hybrid optimization algorithm performance in the optimal operation of multi-reservoir bbenchmark systems. *Water Resour. Manag.*, 33(14), 4767-4782. https://doi.org/10.1007/s11269-019-02393-7

Moreira, F.C., Boaventura, R.A.R., Brillas, E. and Vilar, V.J.P. (2017). Electrochemical advanced oxidation processes: a review on their application to synthetic and real wastewaters. *Appl. Catal. B: Envi ron.*, 202, 217-261. https://doi.org/10.1016/j.apcatb.2016.08. 037

Mossavi, E., Sabzevari, M.H., Ghaedi, M., Azqhandi, M.H.A. and Hosseini, S.J. (2019). A rapid and efficient sono-chemistry process for removal of pollutant: Statistical modeling study. *Polyhedron*, 171, 65-76. https://doi.org/10.1016/j.poly.2019.07.002

Nidheesh, P.V, Zhou, M. and Oturan, M.A. (2018). An overview on the removal of synthetic dyes from water by electrochemical advanced oxidation processes. *Chemosphere*, 197(1), 210-227. https://doi.org/10.1016/j.chemosphere.2017.12.195

Oyehan, T.A., Liadi, M.A. and Alade, I.O. (2019). Modeling the efficiency of TiO_2 photocatalytic degradation of MTBE in contaminated water: A support vector regression approach. *SN Appl. Sci.*, 1(5), 386. https://doi.org/10.1007/s42452-019-0417-4

Rodriguez, J.D., Perez, A. and Lozano, J.A. (2010). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE T. Pattern Anal.*, 32(3), 569-575. https://doi.org/10.1109/TPAMI.2009.187

Sacks, J., Welch, W.J., Mitchell, T.J. and Wynn, H.P. (1989). Design and analysis of computer experiments. *Stat. Sci.*, 4(4), 409-423. https://doi.org/10.1214/ss/1177012413

Santos, M.J.R., Medeiros, M.C., Oliveira, T.M.B.F., Morais, C.C.O., Mazzetto, S.E., Martínez-Huitle, C.A. and Castro, S.S.L. (2016). Electrooxidation of cardanol on mixed metal oxide (RuO_2-TiO_2 and IrO_2-RuO_2-TiO_2) coated titanium anodes: insights into recalcitrant phenolic compounds. *Electrochim. Acta*, 212, 95-101. https://doi.org/10.1016/j.electacta.2016.06.145

Shiri, J., Zounemat-Kermani, M., Kisi, O. and Karimi, S.M. (2020). Comprehensive assessment of 12 soft computing approaches for modelling reference evapotranspiration in humid locations. *Meteorol. Appl.*, 27(1), 1-15. https://doi.org/10.1002/met.1841

Siedlecka, E.M., Ofiarska, A., Borzyszkowska, A.F., Bialk-Bielinska, A., Stepnowski, P. and Pieczynska, A. (2018). Cytostatic drug removal using electrochemical oxidation with BDD electrode: Degradation pathway and toxicity. *Water Res.*, 144(1), 235-245. https://doi.org/10.1016/j.watres.2018.07.035

Simpson, T.W., Lin, D.K.J., and Chen, W. (2001). Sampling strategies for computer experiments: design and analysis. *Int. J. Reliab. Appl.*, 2(3), 209-240.

Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *J. R. Stat. Soc., Ser. B, Methodol.*, 36(2), 111-133. https://doi.org/10.1111/j.2517-6161.1974.tb00994.x

Tanhaei, B., Ayati, A., Lahtinen, M., Vaziri, B.M. and Sillanpää, M. (2016). A magnetic mesoporous chitosan based core-shells biopolymer for anionic dye adsorption: Kinetic and isothermal study and application of ANN. *J. Appl. Polym. Sci.*, 133(22), 43466. https://doi.org/10.1002/app.43466

Valikhan-Anaraki, M., Mousavi, S.F., Farzin, S., Karami, H., Ehteram, M., Kisi, O., Fai, C.M., Hossain, M.S., Hayder, G., Ahmed, A.N., El-Shafie, A.H., Hashim, H.B., Afan, H.A., Lai, S.H. and El-Shafie, A. (2019). Development of a Novel Hybrid Optimization Algorithm for Minimizing Irrigation Deficiencies. *Sustain.*, 11(8), 2337. https://doi.org/10.3390/su11082337

Vapnik, V. and Mukherjee, S. (2000). Support vector method for multivariate density estimation. *Proc. Adv. Neural Inf. Process. Syst.*, 659-665.

Wan, J.T., Jin, C.J., Liu, B.H., She, Z.L., Gao, M.C. and Wang, Z.Y. (2019). Electrochemical oxidation of sulfamethoxazole using Ti/SnO_2-Sb/Co-PbO_2 electrode through ANN-PSO. *J. Serb. Chem. Soc.*, 84(7), 713-727. https://doi.org/10.2298/JSC1808100-25W

Wu, W.Y., Huang, Z.H., Hu, Z.T., He, C. and Lim, T.T. (2017). High performance duplex-structured SnO_2 anode modified by SnO_2-Sb-CNT composite for bisphenol A removal: electrochemical oxidation enhanced by adsorption. *Sep. Purif. Technol.*, 179, 25-35. https://doi.org/10.1016/j.seppur.2017.01.032

Yang, S.Y., Choo, Y.S., Kim, S., Lim, S.K., Lee, J. and Park, H. (2012). Boosting the electrocatalytic activities of SnO_2 electrodes for remediation of aqueous pollutants by doping with various metals. *Appl. Catal. B: Environ.*, 111-112, 317-325. https://doi.org/10.1016/j.apcatb.2011.10.014