

## Research Article

# Machine Learning-Based Sine-Cosine Algorithm for Wastewater Quality Assessment Using Activated Carbon

Hadeel Alsolai,<sup>1</sup> Mashaal M. Asiri,<sup>2</sup> Rana Alabdan,<sup>3</sup> Mohammed Abdullah Al-Hagery,<sup>4</sup> Anwer Mustafa Hilal ,<sup>5</sup> Mohammed Rizwanullah,<sup>5</sup> Abdelwahed Motwakel,<sup>5</sup> and Ishfaq Yaseen<sup>5</sup>

<sup>1</sup>Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia

<sup>2</sup>Department of Computer Science, College of Science & Art at Mahayil, King Khalid University, Saudi Arabia

<sup>3</sup>Department of Information Technology, College of Computer and Information Science, Majmaah University, Al-Majmaah 11952, Saudi Arabia

<sup>4</sup>Department of Computer Science, College of Computer, Qassim University, Saudi Arabia

<sup>5</sup>Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, AlKharj, Saudi Arabia

Correspondence should be addressed to Anwer Mustafa Hilal; [a.hilal@psau.edu.sa](mailto:a.hilal@psau.edu.sa)

Received 13 January 2022; Revised 17 February 2022; Accepted 6 April 2022; Published 9 May 2022

Academic Editor: S.K. Khadheer Pasha

Copyright © 2022 Hadeel Alsolai et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Activated carbon is one of the most highly proven adsorbents for organic chemicals from wastewater. It acts as a filter and adsorbs various chemicals from the wastewater. It has large pore size and strong adsorptive capacity. The quality of wastewater is generally determined by chemical oxygen demand (COD), biochemical oxygen demand (BOD<sub>5</sub>), total suspended solids (TSSs), total phosphorus (TP), and total nitrogen (TN). Wastewater contaminant measurement is significant for saving aquatic life and reusing treated water. Adsorption of contaminants that contribute for wastewater quality indicators uses machine learning algorithm for prediction. Many research works have been done, and the issues are inefficiency and time consuming in the adsorption of contaminants by activated carbon in wastewater management. To overcome these issues, this paper introduces hybrid technique of Voting-Based Extreme Learning Machine with sine cosine algorithm (VELM-SCA). The accuracy of VELM-SCA algorithm in classification of water quality status produced improved accuracy is 0.97.

## 1. Introduction

Removal of dissolved chemical components and pollutants from the wastewater plant by surface phenomenon is called as adsorption. The wastewater treatment is a complicated process that ensures effective purification of wastewater. The adsorption of the organic pollutant to maintain the reduction of pollutant level in the wastewater treatment plant is an essential one. In the wastewater, treatment of plants depends to the indicator values of parameters which describe the quality of wastewater. For maintaining adequate

level of indicators of wastewater quality in the aspect of reducing the pollutant in wastewater, many techniques are available for predicting both quality and quantity of influent wastewater. The wastewater quality indicator parameters are total nitrogen (TN), total phosphorus (TP), biochemical oxygen demand (BOD<sub>5</sub>), total suspended solids (TSS), and chemical oxygen demand (COD).

United nation sustainable goal 6 that states “ensure access to water and sanitation for all” motivates us to research on wastewater treatment. Survey points that one person out of every three does not have safe drinking infra-

structure. Water is basic necessities for all living creature. It is essential to monitor wastewater and its contaminants present in it.

Adsorption utilizes carbon-based material like activated carbon, and biochar was one of the prominent methods in the wastewater treatment [1]. To implement this, so many existing algorithms like random forest (RF), decision tree (DT), Gradient Boosting algorithm (GBA) are available. The issues in the existing algorithms are high efficiency but inaccurate in the reduction of pollutant in the wastewater plants. To overcome this issue, this paper promotes hybrid technique of Voting-Based Extreme Learning Machine with sine cosine algorithm (VELM-SCA). It produces high variation in the prediction of carbon in the wastewater plants and gives high variation in quality and quantity of reduction of the pollutant in the wastewater. The concept of voting ELM is a majority voting system which classified the adsorption of carbon in the wastewater. To improve the accuracy rate, this voting approach is used [2]. For improving the exploitation sine-cosine, the algorithm (SCA) allows to select the prediction of carbon by using search operations in a mathematical function of the sine and cosine [3].

Machine learning algorithm plays a powerful tool in elaborating the complicated multivariate and interpretation for optimizing the adsorption of the dissolved component in the wastewater system. The algorithms like response to surface methodology (RSM), and boosted regression tree (BRT) and general regression neural network (GRNN) were applied in the prediction of carbon in the wastewater model [4, 5]. This paper [6] proposed the polluted natural environment for a long period because of the usage of persistent contaminants which can easily enter into the surface and groundwater. The most powerful and reusable adsorbent is the activated carbon for wastewater treatment [7]. The main contributions of this paper are as follows:

- (1) Implementing the hybrid technique of the proposed method by using VELM based on the concepts of prediction of carbon absorption in the wastewater by taking decision making of the majority of the voting system
- (2) To improve the accuracy of sine-cosine, a mathematical function-based optimization algorithm is implemented

The paper has been organized as follows: Section 2 describes the review of the literature, Section 3 introduces the prediction of wastewater treatment using VELM-SCA, Section 4 discusses the experimented results, and Section 5 concludes the paper with future directions.

## 2. Review of Literature

Adsorption is the removal of the organic and inorganic pollutants on the surface of the adsorbents as well as in wastewater treatment. In recent years, searching for low-cost adsorbents in the pollutant component which are available in the water and wastewater treatment and activated carbon

plays a vital role in the adsorption of pollutant material [8]. Adsorbents are categorized into two types, namely, natural and synthetic adsorbents. Natural adsorbents are clays, charcoal, zeolites, clay minerals, and ores. Synthetic adsorbents are formed from household wastage, agricultural product wastage, industrial wastage, polymeric adsorbents, and sewage sludge.

Activated carbon acts as an adsorbent for both natural and synthetic adsorbents because of it is an easily available product and cheap, enhancing the properties of adsorption capabilities. Similarly, granular activated carbon (GAC) and powdered activated carbons (PAC) are commonly available adsorbents in the removal of organic and inorganic and removal of industrial wastage water in the land surface area [9]. In the carbon-based adsorbent, predicting the capacity of adsorption is by (Q, mg/g), and it is compared with various influencers of properties of carbon-based properties which are total carbon content (C, wt.%), the molar ratio of oxygen to carbon (O/C), the molar ratio of hydrogen to carbon (H/C), molar ratio of oxygen and nitrogen to carbon [(O+N)/C], ash content (ash, wt.%), point of zero charges (pHpzc), and Brunauer-Emmett-Teller surface area (SBET, m<sup>2</sup>/g) [1]. This paper [10] proposed the sine-cosine algorithm (SCA) for improving the metaheuristic search. In improving the concept of ELM, this voting ELM can be implemented. This VELM has difference between samples of class forming various vector set values. The multiple independent ELMs are impended for taking majority voting method for making decisions [11]. Table 1 shows the survey on activated carbon in various techniques.

## 3. Proposed Methodology

Nowadays, domestic wastage and industrial wastage pollute the eco-friendly environment, which has been adsorbed by activated carbon. This wastage contains soluble and insoluble content and chemicals. In this proposed work, we implement the hybrid of Voting-Based Extreme Learning Machine with sine-cosine algorithm (VELM-SCA) to improve the efficiency in the prediction of wastewater treatment plants [20–29]. The architecture of the proposed work (VELM-SCA) is given in Figure 1.

Figure 1 describes three phases, namely, data collection, preprocessing, and analysis of VELM-SCA.

**3.1. Data Collection.** In this work, the data were collected from the wastewater treatment plant (WWTP) in Busan, Korea, for three years from January 2008 to December 2010 [19], which are total nitrogen (TN) and total phosphorous (TP) elements, chemical oxygen demand (COD), and biochemical demand (BCD).

**3.2. Preprocessing.** To enhance the quality in the prediction of carbon in the wastewater, plant needed normalization as a preprocessing technique. Data normalization is based on linear normal function by using:

$$\text{nor}_{i,j} = \frac{\text{ver}_{ij} - \text{Min}_k}{\text{Max}_k - \text{Min}_k}, \quad (1)$$

TABLE 1: Survey on techniques used in activated carbon.

Author	Applications
Aravind et.al (2021) [12]	Using morphological operator for activated carbon for dye removal
Alshabib et.al (2021) [13]	Removal of methylene blue from wastewater using DFT in activated carbon
Wang et.al (2021) [14]	Adsorption of dyes on an activated carbon
Zhu et.al (2021) [1]	Carbon-based materials for tetracycline and sulfamethoxazole adsorption
F. Yang et.al (2020) [15]	Biochar derived dissolved organic matter on adsorption of sulfamethoxazole
Li et.al (2020) [16]	Magnetic biochar in water treatment
Xiang et.al (2019) [17]	Carbon-based materials as adsorbents for antibiotic removal
Hasan, et.al (2019) [18]	Activated carbon composite and its testing for methyl orange removal
Kaetzl et.al (2019) [19]	Biochar using wastewater treatment

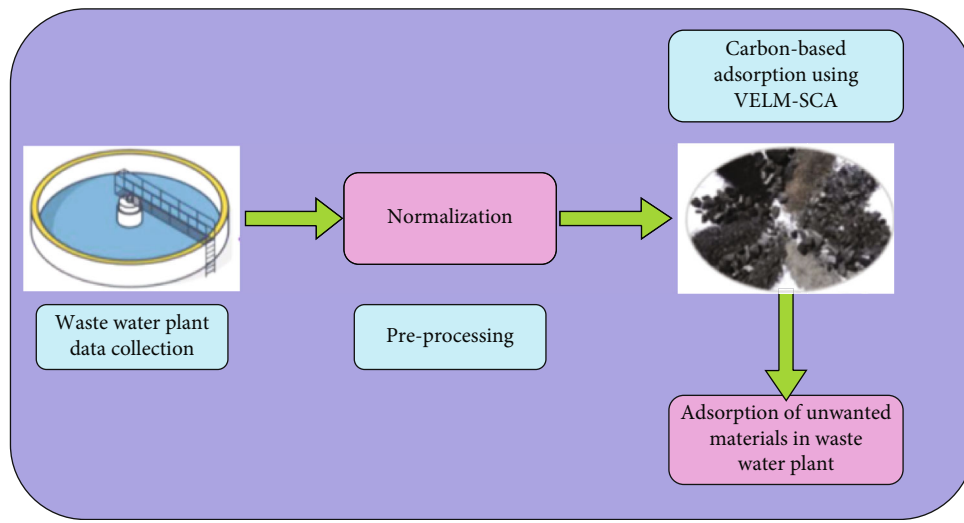


FIGURE 1: Architecture of VELM-SCA.

Input: sample of training data set  $S = \{(a_n, b_n)\}, n = 1, 2 \dots, N$  with label of  $b_n \in \{1, 2, \dots, C\}$  ;  
 $w(t)$  is the water quality indicator of wastewater plant  $t$ ;  $w$  is the total number of organic materials in the wastewater at  $t$ ;  $p$  is the crossover probability and its mutation.  
 Output: prediction of carbon in the wastewater plant using max voting decision.  
 Step 1: randomly select the water quality indicator parameter  $w(t)$   
 Step 2: evaluate fitness value of  $w(t)$  using  

$$\text{fit} = \frac{\sum_{k=1}^K \text{weight}_k \text{acc}_k}{\sum_{k=1}^K \text{weight}_k} \cdot (2)$$
  
 Step 3: choose the next parameter of water quality indicator  $w(t)$ . Crossover and mutation on  $w(t)$  are used. Repeat the following steps until the criteria of genetic algorithm (GA) is met.  
 (1) Probability of crossover  $\text{prob}_c$  is determined by  $(1 - \text{prob}_c)w$  members of  $w(t)$ , and its fitness value is selected probabilistically to  $w(t + 1)$ .  
 (2) Probability of chromosomes is  $\text{prob}_m$  in  $w(t + 1)$  for the mutation, and store the weights of  $\{wt_1^{\text{opt}}, wt_2^{\text{opt}}, \dots, wt_k^{\text{opt}}\}$  which are the outputs, and predict its label of chemical component in the wastewater plant.  
 Step 4: voting-based ELM is the produced output taken as decision making of multiple ELMs by using  

$$\varphi^{\text{wt}}(a, S) = \text{argmax}_{j=1}^C \sum_{k=1}^K wt_k^{\text{opt}} D_{k,j}. (3)$$

ALGORITHM 1: VELMSA.

Step 1: initialize the random population solutions,  $X_1, X_2, \dots, X_n$ .  
 Step 2: by using Equation (4), the value of  $\text{rnd}_1$  is used to switch from explore to exploit. It gradually decreases from 2 to 0.  
 Step 3: while ( $f < f_{\text{maxit}}$ ) do.  
 Step 4: compute the fitness value fit for each solution.  
 Step 5: express  $X_d(f)$  that has the best fitness value  $\text{fit}_d$ .  
 Step 6: update  $\text{rnd}_1, \text{rnd}_2, \text{rnd}_3$ , and  $\text{rnd}_4$  values.  
 Step 7: use Equation (4) to update  $X_i(f)$   
 Step 8: end while  
 Step 9: output  $X_d(f)$ .

ALGORITHM 2: SCA.

Input: sample of training dataset  $S = \{(a_n, b_n)\}$ ,  $n = 1, 2 \dots, N$  with label of  $b_n \in \{1, 2, \dots, C\}$  ;  
 $w(t)$  is the water quality indicator of wastewater plant  $t$ ;  $w$  is the total number of organic materials in the wastewater at  $t$ ;  $p$  is the crossover probability and its mutation.  
 Output: prediction of carbon in the wastewater plant.  
 Step 1: randomly select the water quality indicator parameter  $w(t)$   
 Step 2: evaluate fitness value of  $w(t)$  using Equation (2).  
 Step 3: choose the next parameter of water quality indicator  $w(t)$ . Crossover and mutation on  $w(t)$  are calculated.  
 Step 4: using Algorithm 2, determine the best fitness  $\text{fit}_d$ .  
 Step 5: update the best solution out of all water quality indicator parameter value.  
 Step 6: update the position of adsorption of carbon in the wastewater plant by using  
 If  $\text{rnd}_3 > 1$ , then  $\text{rnd}_3 + X_d(f)$ . It merges all related carbon particle.  
 Step 7: Else  $\text{rnd}_3 < 1$  discard that particle.  
 Step 8: End.

ALGORITHM 3: VELM-SCA.

TABLE 2: VELM-SCA metric measures.

Algorithm	SRCC	R <sup>2</sup>	MAE	RMSE
VELM	0.89	0.82	2.43	4.56
SCA	0.91	0.93	1.67	3.87
VELM-SCA	0.96	0.97	1.22	2.45

TABLE 3: Metric measures' report.

Parameter	VELM	SCA	VELM-SCA
TP	156	180	140
TN	123	148	70
FP	85	60	30
FN	75	45	25
Accuracy	0.69	0.88	0.97
Precision	0.78	0.89	0.95
Recall	0.74	0.93	0.96

where  $\text{nor}_{ij}$  is the normalized value for each data point;  
 $\text{ver}_{ij}$  is the sample data and  $\text{Min}_k, \text{Max}_k$ .

### 3.3. Analysis of VELMSCA

**3.3.1. Evolutionary-Based Voting ELM.** Extreme learning machine (ELM) is based on the structure of input layer-single hidden layer with feed forward neural network. This voting-based ELM employs the concept of multiple indepen-

dent ELM training models, and the final output decision is based on majority voting. The main advantage of using this VELM is the improvement in high quality of detecting available of carbon in the wastewater plant. In addition to this, it lowered the variance. The Algorithm 1 steps of VELM are given below:

Algorithm 1 describes that as per Equation 3,  $S$  contains samples of  $S = \{(a_n, b_n)\}$  and  $n = 1, 2, 3 \dots, N$ , where  $b_n$  is class label. Here, assume that  $a$  is an input and  $b$  is predicted by  $\varphi^{\text{wt}}(a, S)$ . As per voting based ELM, it predicts output of multiple ELM training models. If  $\varphi^{\text{wt}}(a, S)$  predicts the  $k^{\text{th}}$  classifier, then  $D_{k,j} \in \{0, 1\}$  and  $k = 1, 2, 3 \dots$ , and  $K$  and its ensembled decision (max voting) is defined by  $\varphi^{\text{wt}}(a, S) = \arg \max_{j=1}^c \sum_{k=1}^K \text{wt}_k^{\text{opt}} D_{k,j}$ .

**3.3.2. Sine-Cosine Algorithm (SCA).** For getting the best prediction of carbon adsorption in the wastewater plant, this SCA function is implemented. It is based on the mathematical function of sine and cosine. SCA algorithm employs searching process by using sine or cosine equation as

$$X_i(f) = X_i(f) + \text{rnd}_1 \times \text{sinsin}(\text{rnd}_2) \times |\text{rnd}_3 X_d(f) - X_i(f)|, \quad (2)$$

$$X_i(f) = X_i(f) + \text{rnd}_1 \times \text{coscos}(\text{rnd}_2) \times |\text{rnd}_3 X_d(f) - X_i(f)|. \quad (3)$$

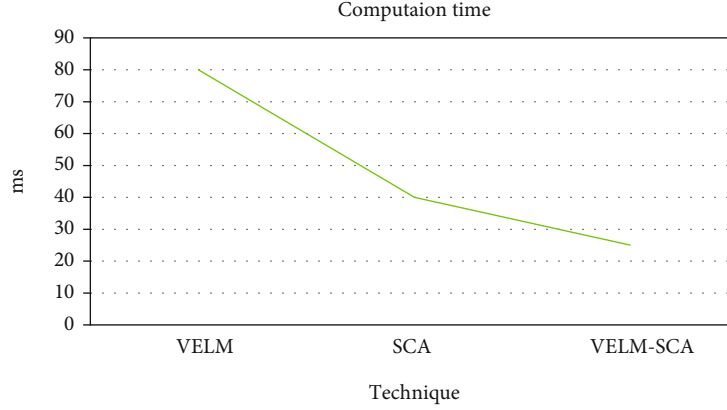


FIGURE 2: Computation time.

TABLE 4: Selection of parametric values.

Influent parameters	Effluent parameters
Influent of COD	Effluent of COD
Influent of pH	Effluent of EC
Influent of N	Effluent of pH
Influent of SO4	Effluent of N
Influent of PO4	Effluent of SO4
Influent of EC	Effluent of PO4

TABLE 5: Range of values of parameters describing influent wastewater.

Parameters	Minimum	Maximum	Average
TN (mg/dm <sup>3</sup> )	19.8	99	70.4
TP (mg/dm <sup>3</sup> )	3.8	38.6	13.55
BOD (mg/dm <sup>3</sup> )	40.5	792	378
COD (mg/dm <sup>3</sup> )	169	2520	930.1
TSS (mg/dm <sup>3</sup> )	82	1145	440
Q (mg/dm <sup>3</sup> )	26983	66883	38758

Combining Equations (2) and (3) as

$$\begin{aligned}
 X_i(f) = & \{X_i(f) + \text{rnd}_1 \times \text{sinsin}(\text{rnd}_2) \times |\text{rnd}_3 X_d(f) \\
 & - X_i(f)| \text{if } \text{rnd}_4 < 0.5 X_i(f) + \text{rnd}_1 \times \text{coscos}(\text{rnd}_2) \\
 & \times |\text{rnd}_3 X_d(f) - X_i(f)| \text{if } \text{rnd}_4 \geq 0.5,
 \end{aligned}
 \tag{4}$$

where  $X_d$  is the destination solution,  $X_i(f)$  is current solution  $\|$  that indicates random variables of  $\text{rnd}_1, \text{rnd}_2, \text{rnd}_3$ , and,  $\text{rnd}_4$  which contains absolute values. The parameter  $\text{rnd}_1$  is a random variable to determine the path of next solution, and it may be towards or outside

between  $X_i(f)$  and  $X_d(f)$ . That is switch between exploration and exploitation. The update of  $\text{rnd}_1$  is given by

$$\text{rnd}_1 = c - \text{it} \frac{c}{T}, \tag{5}$$

where  $c$  is a constant value;  $T$  is maximum number of iterations; and it is the current iteration. The random variable  $\text{rnd}_2$  is used to identify the next movement towards destination solution or best solution  $X_d(f)$ . Also, random variable  $\text{rnd}_3$  adds the random weight to  $X_d(f)$ , and it is stochastically emphasized  $\text{rnd}_3 > 1$  or deemphasize  $\text{rnd}_3 < 1$ . The random variable  $\text{rnd}_4$  is used to switch between the function sine and cosine using Equation (4). The algorithm 2 describes the SCA algorithm.

Algorithm 2 describes that the concept of SCA is randomly choose the solution from the set of solution with various dimensions. For every solution it computes fitness values. The output of best solution of SCA is best fitness.

3.3.3. *Proposed Hybrid of Voting ELM with Sine-Cosine Algorithm (VELM-SCA)*. VELM-SCA enhances the accuracy and efficiency in terms of classification of the wastewater quality and reduces the computation time for both training and testing dataset. Procedure for VELM-SCA is given below.

Algorithm 3 describes that the training of the dataset  $S = \{(a_n, b_n)\}$ , with input parameter values. Calculate the fitness for each parameter and choose the best fitness that denotes the carbon content in the wastewater plant. Update the value and choose the next set of data. Update the position of carbon particle using new random variable  $\text{rnd}_3 > 1$ , and it merges all related carbon particle.

## 4. Result and Discussion

4.1. *Performance Metric Measures*. The metric measures for implementing VELM-SCA of Spearman's rank correlation coefficient (SPCC), the coefficient of determination ( $R^2$ ), root mean square error (RMSE), and mean absolute error (MAE) are discuss in the following sections.

TABLE 6: Parametric values of MAE, MAPE, and  $R^2$  using hybrid of VELM-SCA.

Influent indicators	VELM			SCA			VELM-SCA		
	MAE	MAPE	$R^2$	MAE	MAPE	$R^2$	MAE	MAPE	$R^2$
TN	5.04	9.67	0.46	8.17	11.54	0.34	4.23	8.45	0.46
TP	1.31	12.77	0.48	2.61	17.77	0.46	1.45	12.76	0.46
BOD	44.22	13.69	0.52	49.12	15.91	0.76	40.32	12.41	0.52
COD	121.34	15.09	0.45	149.34	15.12	0.69	106.24	10.09	0.42
TSS	56.05	15.03	0.59	77.05	25.03	0.47	44.05	11.03	0.42

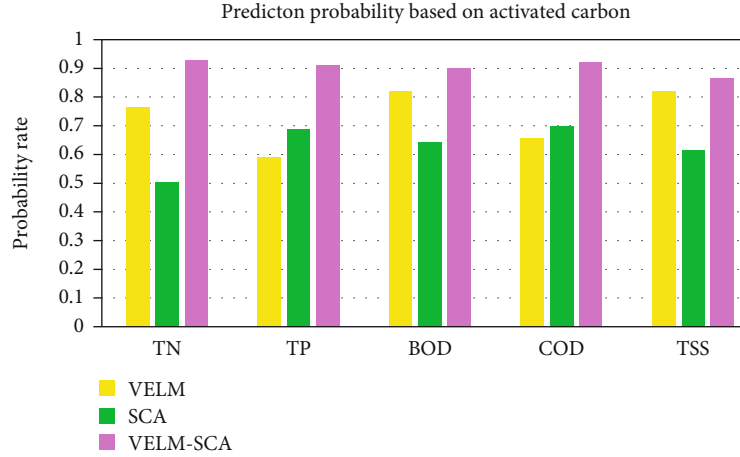


FIGURE 3: Probability prediction of carbon.

TABLE 7: Effectiveness.

Algorithms	Effectiveness				
	TN	TP	BOD	COD	TSS
VELM	75	45	77	45	68
SCA	54	43	45	48	82
VELM -SCA	91	78	92	93	90

4.1.1. *Spearman's Rank Correlation Coefficient (SPCC)*. It is the relationship the between actual and predicted variables that is in relationship between positive or negative. The formula is

$$SRCC = 1 - \left( \frac{6\sum d^2}{n^3 - n} \right), \quad (6)$$

where  $d$  is the difference in ranks between two variables, and  $n$  is the number of samples. The range of values is between -1 and +1.

4.1.2. *Coefficient of Determination ( $R^2$ )*. It explains the fraction of variations in its output, and it is defined by

$$R^2 = \frac{\text{Explained variations}}{\text{Total number of variations}}. \quad (7)$$

The maximum and minimum values of  $R^2$  are 1 and 0. The higher the value it indicates, the best the model.

4.1.3. *Mean Absolute Error (MAE)*. The MAE is defined as an average of absolute differences between the model outputs and the actual outputs. The MAE can be calculated using the following formula:

$$MAE = |\hat{y} - y|, \quad (8)$$

where  $\hat{y}$  is the predicted value, and  $y$  is the actual value.

4.1.4. *Root Mean Square Error (RMSE)*. It is an error prediction measure. It calculates the overall error of this system, and it can be calculated as

$$RMSE = \sqrt{\frac{\sum_{n=1}^N (\hat{y} - y)^2}{N}}. \quad (9)$$

In this work, VELM-SCA, the metric measures its performance based on accuracy, specificity, and sensitivity,



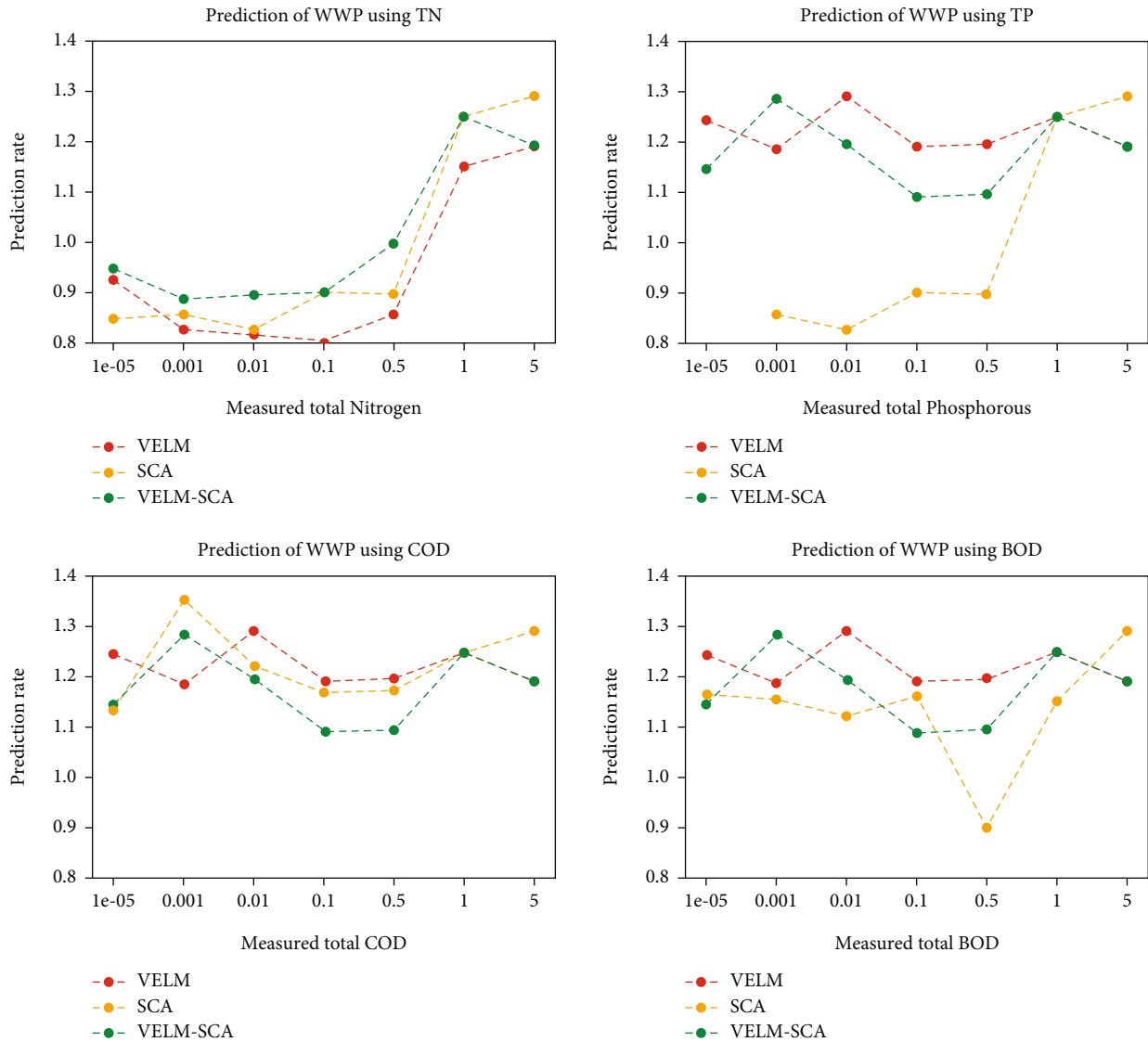


FIGURE 4: Statistical analysis of prediction of various indicators.

and precision and F-score evaluation matrices were employed.

$$\begin{aligned}
 \text{accuracy} &= \frac{TP + TN}{TP + TN + FP + FN}, \\
 \text{precision} &= \frac{TP}{TP + FP}, \\
 \text{recall} &= \frac{TP}{TP + FN}, \\
 \text{F1 - Score} &= \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}.
 \end{aligned}
 \tag{10}$$

Table 2 shows that performance metric measures of VELM-SCA.

From Table 2, it seems that the performance of evaluation is based on SRCC, RMSE, MAE, and  $R^2$  implemented in VELM, SCA, and our proposed work (VELM-SCA) algorithms. For the SRCC shows strongly correlation of 0.96. In RMSE, VELM got 4.56, SCA got 3.87, and VELM-SCA got 2.45. For the  $R^2$  value, proposed work VELM-SCA got 0.97. For the mean absolute error, VELM got 2.43, SCA got 1.67, and VELM-SCA got 1.22.

From the above Table 3, the accuracy of VELM-SCA (proposed work) is higher as compared to other classifier algorithms of VELM and SCA. In Table 3, the next higher accuracy is SCA which is also closer to VELM-SCA. Figure 2 shows the computation time for  $\text{CO}_2$  adsorption in various algorithms like VELM, SCA, and VELM-SCA.

Figure 2 shows that the proposed work needs less computation time for the adsorption of carbon. Table 4 shows that parameter selection. In the wastewater plants, these

parameters are used to predict the parametric values in the dataset.

Table 4 shows that in the influent parameter values of wastewater quantity and quality, the ranges of variation were established [20]. The parametric values in Table 5 indicate substantial variation in TN, TP, BOD, COD, TSS, and Q values. These indicators of wastewater quality constitute input data for the models of VELM-SCA describing the changes in carbon, nitrogen, phosphorus, and biochemical changes in the compounds of bioreactors.

By using the Table 5 parametric values, our proposed work VELM-SCA predicted activated carbon-based TN, TP, BOD, COD, and TSS values using mean absolute error (MAE), mean relative error (MAPE), and coefficient of determination ( $R^2$ ). The analysis of these parameters are shown in Table 6.

The analysis of Table 4 shows that TN prediction activated carbon based on VELM inMAE = 5.04 mg/dm<sup>3</sup> and SCA inMAE = 8.17 mg/dm<sup>3</sup>, and fusion of VELM-SCA gives MAE = 4.23 mg/dm<sup>3</sup>, similar for MAPE and VELM inMAPE = 99.67 mg/dm<sup>3</sup>, and SCA inMAPE = 11.54 mg/dm<sup>3</sup> and hybrid of VELM-SCA gives MAPE = 8.45 mg/dm<sup>3</sup>. In addition, TP prediction activated carbon based on VELM inMAE = 1.31 mg/dm<sup>3</sup> and SCA inMAE = 2.61 mg/dm<sup>3</sup>, and hybrid of VELM-SCA gives MAE = 1.45 mg/dm<sup>3</sup>, similar for MAPE and VELM inMAPE = 12.77 mg/dm<sup>3</sup> and SCA inMAPE = 17.77 mg/dm<sup>3</sup>, and hybrid of VELM-SCA gives MAPE = 12.76 mg/dm<sup>3</sup>. Prediction of wastewater quality requires parametric indicators of TN, and TP shown that our proposed work VELM-SCA gives better performance. Similarly, for other indicators of BOD and COD, TSS also shows that the lowest prediction errors were observed for the VELM-SCA approach. Figure 3 shows the probability of prediction based on carbon in the wastewater plant for our proposed work.

From Figure 3, it seems that our proposed work VELM-SCA gives the better probability prediction of wastewater treatment for plants. Table 6 shows the effectiveness of various machine learning algorithm in terms of various influent indicators by using

$$\text{Effectiveness} = \frac{N - L}{N} \times 100, \quad (11)$$

where  $N$  is the total number of testing data, and  $L$  is the total number of losing test data.

From Table 7, it seems that the effectiveness of our proposed work produces better result compared with other existing algorithms. Figure 4 shows the VELM-SCA model for predicting the influent water based on activated carbon with various indicators.

In Figure 4, it seems that our proposed work gives effective prediction of WWTP that measured various indicators for the treatment of plants, which is the convergence curves of VELM-SCA for all datasets. Our proposed work has a fast convergence compared with other two algorithms.

## 5. Conclusion

The prediction of wastewater treatment plants used the adsorption of unwanted influent indicators using activated carbon. In this work, VELM-SCA uses the predicting influent indicators in the wastewater and identify the maximum and minimum point values with boundary set values. It also implemented various parameters like total phosphorous (TP), total nitrogen (TN), BOD, COD, and TSS. VELM-SCA accurately predicts the unwanted influent indicators and filtered it using activated carbon with minimum error prediction and also proves the efficient and robust performance. The accuracy of VELM-SCA algorithm in classification of water quality status that produced highest accuracy is 0.97. Our proposed work VELM-SCA outperforms the best result compared with existing classifier. In future to work this will be extended by using various ML algorithms and also upgrade our work in various influent and Effluent indicators for predicting the wastewater treatment for plants.

## Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work under grant numbers (RGP 2/45/43). Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2022R303), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. The authors would like to thank the Deanship of Scientific Research at Majmaah University for supporting this work under Project numbers (R-2022-50 and R-2022-119).

## References

- [1] X. Zhu, Z. Wan, D. C. W. Tsang et al., "Machine learning for the selection of carbon-based materials for tetracycline and sulfamethoxazole adsorption," *Chemical Engineering Journal*, vol. 406, article 126782, 2021.
- [2] M. Mengcan, C. Xiaofang, and X. Yongfang, "Constrained voting extreme learning machine and its application," *Journal of Systems Engineering and Electronics*, vol. 32, no. 1, pp. 209–219, 2021.
- [3] M. Z. Rehman, A. Khan, R. Ghazali, M. Aamir, and N. M. Nawari, "A new multi sine-cosine algorithm for unconstrained optimization problems," *PLoS One*, vol. 16, no. 8, article e0255269, 2021.
- [4] X. Zhu, D. C. W. Tsang, L. Wang et al., "Machine learning exploration of the critical factors for CO<sub>2</sub> adsorption capacity on porous carbon materials at different pressures," *Journal of Cleaner Production*, vol. 273, article 122915, 2020.



- [5] M. Foroughi, M. H. A. Azghandi, and S. Kakhki, "Bio-inspired, high, and fast adsorption of tetracycline from aqueous media using Fe<sub>3</sub>O<sub>4</sub>-g-CN@PEI- $\beta$ -CD nanocomposite: modeling by response surface methodology (RSM), boosted regression tree (BRT), and general regression neural network (GRNN)," *Journal of Hazardous Materials*, vol. 388, article 121769, 2020.
- [6] M. Bahrami, M. J. Amiri, and B. Beigzadeh, "Adsorption of 2,4-dichlorophenoxyacetic acid using rice husk biochar, granular activated carbon, and multi-walled carbon nanotubes in a fixed bed column system," *Water Science and Technology*, vol. 78, no. 8, pp. 1812–1821, 2018.
- [7] A. Spaltro, M. Pila, S. Simonetti et al., "Adsorption and removal of phenoxy acetic herbicides from water by using commercial activated carbons: experimental and computational studies," *Journal of Contaminant Hydrology*, vol. 218, pp. 84–93, 2018.
- [8] G. Crini, "Recent developments in polysaccharide-based materials used as adsorbents in wastewater treatment," *Progress in Polymer Science*, vol. 30, no. 1, pp. 38–70, 2005.
- [9] C. Moreno-Castilla, "Adsorption of organic molecules from aqueous solutions on carbon materials," *Carbon*, vol. 42, no. 1, pp. 83–94, 2004.
- [10] S. Mirjalili, "SCA: a sine cosine algorithm for solving optimization problems," *Knowledge-Based Systems*, vol. 96, pp. 120–133, 2016.
- [11] R. Polikar, "Ensemble based systems in decision making," *IEEE Circuits and Systems Magazine*, vol. 6, no. 3, pp. 21–45, 2006.
- [12] M. Aravind and M. Amalanathan, "Structural, morphological, and optical properties of country egg shell derived activated carbon for dye removal," *Materials Today: Proceedings*, vol. 43, pp. 1491–1495, 2021.
- [13] M. Alshabib, M. A. Oluwadamilare, A. Tanimu, I. Abdulazeez, K. Alhooshani, and S. A. Ganiyu, "Experimental and DFT investigation of ceria-nanocomposite decorated AC derived from groundnut shell for efficient removal of methylene-blue from wastewater effluent," *Applied Surface Science*, vol. 536, article 147749, 2021.
- [14] H. Wang, Z. Li, S. Yahyaoui et al., "Effective adsorption of dyes on an activated carbon prepared from carboxymethyl cellulose: experiments, characterization and advanced modelling," *Chemical Engineering Journal*, vol. 417, article 128116, 2021.
- [15] F. Yang, Q. Zhang, H. Jian et al., "Effect of biochar-derived dissolved organic matter on adsorption of sulfamethoxazole and chloramphenicol," *Journal of Hazardous Materials*, vol. 396, article 122598, 2020.
- [16] X. Li, C. Wang, J. Zhang, J. Liu, B. Liu, and G. Chen, "Preparation and application of magnetic biochar in water treatment: a critical review," *Science of the Total Environment*, vol. 711, article 134847, 2020.
- [17] Y. Xiang, Z. Xu, Y. Wei et al., "Carbon-based materials as adsorbent for antibiotics removal: mechanisms and influencing factors," *Journal of Environmental Management*, vol. 237, pp. 128–138, 2019.
- [18] M. Hasan, M. M. Rashid, M. M. Hossain et al., "Fabrication of polyaniline/activated carbon composite and its testing for methyl orange removal: Optimization, equilibrium, isotherm and kinetic study," *Polymer Testing*, vol. 77, article 105909, 2019.
- [19] K. Kaetzl, M. Lübken, G. Uzun et al., "On-farm wastewater treatment using biochar from local agroresidues reduces pathogens from irrigation water for safer food production in developing countries," *Science of the Total Environment*, vol. 682, pp. 601–610, 2019.
- [20] S. Moosavi, O. Manta, Y. A. El-Badry, E. E. Hussein, Z. M. El-Bahy, and M. Fawzi, "A Study on Machine Learning Methods' Application for Dye Adsorption Prediction onto Agricultural Waste Activated Carbon," *Nanomaterials*, vol. 11, p. 2734, 2021.
- [21] B. Szeląg, K. Barbusiński, J. Studziński, and L. Bartkiewicz, "Prediction of wastewater quality indicators at the inflow to the wastewater treatment plant using data mining methods," *E3S Web of Conferences*, vol. 22, article 00174, 2017.
- [22] A. S. Mohammed, K. Venkatachalam, S. Hubálovský, P. Trojovský, and P. Prabu, "Smart Edge Computing for 5 g/6 g Satellite IOT for Reducing Inter Transmission Delay," *Mobile Networks and Applications*, vol. 178, pp. 1–10, 2022.
- [23] N. Bacanin, M. Zivkovic, T. Bezdán, K. Venkatachalam, and M. Abouhawwash, "Modified firefly algorithm for workflow scheduling in cloud-edge environment," *Neural Computing and Applications*, pp. 1–26, 2022.
- [24] K. Venkatachalam, P. Prabu, A. S. Alluhaidan, S. Hubálovský, and P. Trojovský, "Deep belief neural network for 5G diabetes monitoring in big data on edge IoT," *Mobile Networks and Applications*, pp. 1–10, 2022.
- [25] N. Bacanin, B. Timea, K. Venkatachalam, and F. Al-Turjman, "Optimized convolutional neural network by firefly algorithm for magnetic resonance image classification of glioma brain tumor grade," *Journal of Real-Time Image Processing*, vol. 18, no. 4, pp. 1085–1098, 2021.
- [26] K. Venkatachalam, P. Prabu, A. Almutairi, and M. Abouhawwash, "Secure biometric authentication with de-duplication on distributed cloud storage," *PeerJ Computer Science*, vol. 7, article e569, 2021.
- [27] K. Venkatachalam, S. Siuly, N. Bacanin, S. Hubalovsky, and P. Trojovsky, "An efficient Gabor Walsh-Hadamard transform based approach for retrieving brain tumor images from MRI," *IEEE Access*, vol. 9, pp. 119078–119089, 2021.
- [28] S. Ramamoorthy, G. Ravikumar, B. Saravana Balaji, S. Balakrishnan, and K. Venkatachalam, "MCAMO: multi constraint aware multi-objective resource scheduling optimization technique for cloud infrastructure services," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 6, pp. 5909–5916, 2021.
- [29] R. S. Ponnagal, S. Karthick, B. Dhiyanesh, S. Balakrishnan, and K. Venkatachalam, "Optimized virtual network function provisioning technique for mobile edge cloud computing," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 6, pp. 5807–5815, 2021.