University of Baghdad
College of Engineering
JOURNAL OF ENGINEERING

Journal of Engineering

journal homepage: www.joe.uobaghdad.edu.iq

Volume 30 Number 2 February 2024

Predicting Biochemical Oxygen Demand at the Inlet of Al-Rustumiya Wastewater Treatment Plant Using Different Mathematical Techniques

Saja Ali Abd[1], Ali Omran Al-Sulttani[2,*]

Department of Water Resources Engineering, College of Engineering, University of Baghdad, Baghdad, Iraq
saja.ali2110m@coeng.uobaghdad.edu.iq[1], ali2003ena@uobaghdad.edu.iq[2]

ABSTRACT

Water quality planning relies on Biochemical Oxygen Demand BOD. BOD testing takes five days. The Particle Swarm Optimization (PSO) is increasingly used for water resource forecasting. This work designed a PSO technique for estimating everyday BOD at Al-Rustumiya wastewater treatment facility inlet. Al-Rustumiya wastewater treatment plant provided 702 plant-scale data sets during 2012-2022. The PSO model uses the daily data of the water quality parameters, including chemical oxygen demand (COD), chloride (Cl^-), suspended solid (SS), total dissolved solids (TDS), and pH, to determine how each variable affects the daily incoming BOD. PSO and multiple linear regression (MLR) findings are compared, and their performance is evaluated using mean square error, relative absolute mistake, and coefficient of determination. PSO utilised COD, TDS, SS, pH, and Cl^- as inputs, generating a mean square error of 1029.10, an average absolute relative error of 9.41%, and a coefficient of determination of 0.89. Comparisons demonstrated that the PSO model could accurately calculate the daily BOD at Al-Rustumiya wastewater treatment plant's inlet.

Keywords: Particle Swarm Optimization, Multiple Linear Regression Model, MATLAB, Sensitivity Analysis.

*Corresponding author
Peer review under the responsibility of University of Baghdad.
https://doi.org/10.31026/j.eng.2024.02.02

Article received: 21/07/2023
Article accepted: 04/10/2023
Article published:01/02/2024

التنبؤ بطلب الأكسجين الكيميائي الحيوي عند مدخل محطة الرستمية لمعالجة مياه الصرف الصحي باستخدام تقنيات رياضية مختلفة

سجى علي عبد ، علي عمران السلطاني*

قسم هندسة الموارد المائية، كلية الهندسة، جامعة بغداد، بغداد، العراق

الخلاصة

يعتمد تخطيط جودة المياه على الطلب على الأوكسجين الكيميائي الحيوي BOD. يستغرق اختبار BOD 5 أيام. يستخدم تحسين سرب الجسيمات (PSO) بشكل متزايد للتنبؤ بالموارد المائية. صمم هذا العمل تقنية PSO لتقدير الطلب على الأوكسجين الكيميائي الحيوي اليومي في مدخل محطة معالجة مياه الصرف الصحي في الرستمية. وفرت محطة معالجة مياه الصرف الصحي في الرستمية 702 مجموعة بيانات خلال الفترة 2012–2022. يتم تغذية PSO بالعديد من البيانات اليومية متغيرات نوعية المياه ، بما في ذلك الطلب الكيميائي على الأوكسجين (COD) ، والكلوريد (Cl^-) ، والمواد الصلبة العالقة (SS) ، وإجمالي المواد الصلبة الذائبة (TDS) ، ودرجة الحموضة (pH)، لتحديد كيفية تأثير كل متغير على BOD الوارد يوميا الى المحطة. تتم مقارنة نتائج PSO والانحدار الخطي المتعدد MLR ويتم تقييم أدائها باستخدام متوسط الخطأ التربيعي والخطأ النسبي المطلق ومعامل التحديد. تم استخدام COD و TDS و SS و pH و Cl^- كمدخلات بواسطة PSO ، مما أدى إلى توليد خطأ مربع متوسط قدره 1029.10 ، ومتوسط خطأ نسبي مطلق قدره 9.41٪ ، ومعامل تحديد 0.89. أظهرت المقارنات أن نموذج PSO يمكنه حساب الطلب الأوكسجيني البيولوجي اليومي بدقة في مدخل محطة الرستمية لمعالجة مياه الصرف الصحي.

الكلمات المفتاحية: امثلية سرب الجسيمات، نموذج الانحدار الخطي المتعدد، ماتلاب، تحليل الحساسية.

1. INTRODUCTION

Variations in the integrity of water from different sources continue to be a source of concern. Consequently, effective methods for modelling water quality parameters in surface waters are required for pollution control and the implementation of necessary management **(Al-Musawi, 2016; Corominas et al., 2018; Abdallah et al., 2020; Bhagat et al., 2020)**. Several thousand distinct chemical compounds are discharged into the environment by industrial and municipal wastewaters, which are significant sources of contamination for aquatic biota **(Abbas and Hassan, 2018; Mohammed et al., 2022)**. Therefore, the significance of utilizing efficient control and monitoring methods for effluent treatment systems is widely acknowledged **(Khudair, 2019; Mohsin et al., 2021)**. Any wastewater purification facility requires a trustworthy **(Ye et al., 2021)**. The model offers an instrument to estimate the result and a foundation for regulating the process's operation **(Robles et al., 2019)**. Due to the abundance of bio-organic components that are challenging to model using a mechanical approach, this method is intricate and highly nonlinear. Using conventional experimental methods to predict the operational parameters of a plant is a time-intensive endeavour that hinders the control of such processes **(Mohammed and Al-Obaidi, 2021)**.

BOD is one of the most essential effluent management and planning parameters. The approximate biodegradable organic substance concentration in the water specimen. It is determined by the quantity of oxygen needed by aerobic microorganisms in the specimen to convert the organic matter into a stable form. The relationship between biochemical oxygen demand (BOD) and chemical oxygen demand (COD) is significant for measuring the oxygen consumption induced by the decomposition of organic matter. The biochemical oxygen demand (BOD) and chemical oxygen demand (COD) are two important parameters used in estimating the degree of organic pollution in wastewater **(Amarasinghe et al., 2017)**. Preparing for and analysing the BOD test requires considerable time and effort. The fifth day of this process is devoted to data acquisition and analysis. **(Khazraji and Nasser, 2012; Garcia et al., 2013; Malviya and Jaspal, 2021)** Several water quality models, including traditional mechanistic approaches, have been developed to handle the most efficient water conservation practices. Most of these models require inaccessible data to be entered, making the process extremely expensive and time-consuming. **(Kennedy and Eberhart, 1995)** described a population-specific technique for uncertain optimization influenced by the social behaviour patterns of flocking birds or schools of fish.

Particle swarm optimization is a population-specific technique **(Jian-jun et al., 2013)**. A collection of particles representing potential initial solutions is used to initialize the system. These particles navigate the search space in search of the optimal fitness value. Particle swarm optimization is among the finest optimizing strategies. It results from its worldwide convergence capabilities, straightforward adoption, implementation, and durability **(Xie et al., 2012)**. PSO is a 1995 algorithm for evolutionary computation inspired by flocking birds' social conduct. Algorithm PSO is widely acknowledged and utilized to solve various optimization issues. A particle, a potential solution, is an s-dimensional vector in the PSO method **(dos Santos Coelho, 2010)**. A swarm is a random cluster of particles. The optimal position of a particle within the search space is determined by comparing its present fitness value to its maximum fitness value **(Yan et al., 2019)**. Since particle velocities and positions are modified after each iteration, the global best position is the best of all individual best positions, and the particle flight velocity is the particle's velocity in the physical analogy **(Al-Obaidi, 2020)**. PSO is suitable for estimating BOD at the inlet because it accounts for nonlinear interactions between predictors and predicted. Effectiveness of multiple linear regression (MLR) and PSO models compared.

This work aims to use the PSO model to obtain the BOD values at the inlet of the wastewater treatment plants directly without waiting for five days.

2. DESCRIPTION OF DATA

In the current work, 702 BOD recording data from the Al-Rustumiya wastewater treatment plant inlet that have been tested and analyzed were used to generate the dataset for the proposed numerical model. The input vectors (COD, SS, pH, Cl^{-1}, TDS) and the output variable (BOD) are selected. The minimum value of three between the number of dataset items and the number of variable inputs required to model acceptableness and ratios above five are recommended **(Frank and Todeschini, 1994)**. This ratio for the training sets in the current case study was 562/5 or 112.4, which can be accepted since it exceeds the recommended value. Five hundred sixty-two records (80%) from the 702 datasets were considered for the construction process, while the remaining 140 recordings (20%) were used to validate the PSO models. **Table 1** provides descriptive information for the samples.

Table 1. Statistics describing the variables utilized in the development of the model.

Description		Mean	Min.	Max.	Standard Error	Standard Deviation	Sample Variance
Input	COD, ppm	359.78	140	942	4.427	117.28	13754
	SS, ppm	350.09	102	982	6.428	170.32	29010
	pH	7.3260	5.15	7.98	0.012	0.3242	0.1051
	Cl^-, ppm	321.26	192	551	1.8981	50.297	2529.7
	TDS, ppm	1222.76	829	1792	5.333	141.30	19967.1
Output	Total BOD, ppm	Total BOD of wastewater treatment plant inlet					

3. PARTICLE SWARM OPTIMISATION (PSO) ALGORITHM

Each particle's position and velocity can be updated following the Eqs. (1) and (2) throughout the entire search. Procedure **(Al-Sulttani et al., 2022)**.

$$V_i(t+1) = wV_i(t) + c_1 Rand(\cdot)_1 \mid pbest_i t - X_i(t) \mid + c_2 Rand(\cdot)_2 \mid gbest_i t - X_i(t) \mid \quad (1)$$

$$X_i(t+1) = X_i(t) + V_i(t+1) \quad (2)$$

where:

X_i and V_i indicate the location and velocity of individual particles accordingly.

$Rand(\cdot)_1$ and $Rand(\cdot)_2$ are evenly dispersed at random digits among 0 and 1 that are essentially equal.

pbest represents the optimal location of every particle in the distance.

gbest symbolises the optimal placement of each particle globally **(Poli and Blackwell, 2007).** Acceleration coefficients c_1 and c_2 are a term 'reliability' variables which represent the level of assurance in the optimal solution discovered by a single particle (c_1 – the cognitive parameter) and by virtue of horde as a whole (c_2 -social the parameter). w represents the inertia mass in Eq. (1), which was added to enhance a convergence process at the iterative process. The following amount of weight is the factor of scaling used to manage the swarm's exploration abilities. It modifies a present velocity value, influencing the most recent velocity vector. The inertial weight was missing from the initial PSO algorithm but was subsequently included **(Shi and Eberhart, 1998)**.

4. CONVERGENCE CRITERIA

Due to the repetitive character of the PSO searches, convergent terms can be used to terminate the optimization process. The most commonly accepted convergent metrics are the maximum number of PSO rounds and a required minimal estimation error for an objective function's optimal value. The degree of difficulty of the optimization issues defines the most iterations, whereas the second criterion implies that the optimal global value has already been determined. Evaluating or adjusting the algorithm in equations for which the optimal value is already known is necessary. However, this cannot be applied to actual structure optimization problems in which the optimal solution is unknown. **Table 2** lists the primary PSO parameters, whereas **Table 3** enumerates and discusses the PSO convergent variables used in this investigation.

Table 2. Principal particle swarm optimization variables. **(Lavanya and Udgata, 2011)**

Description	Details
The number of variables, N	The median limit is 10 to 40. The amount of specific difficult or special issues can be raised to between 50 and 100.
The dimension of variables, D	It depends on the optimization problem.
Weight of inertia, w	Generally, (w) has been adjusted to a number less than 1, and w = 0.70 is considered ideal for promoting quicker convergence. It can also be changed while running trials.
x^U are vectors comprising the minimum and maximum amounts of the n-design particles	The optimization problem dictates their existence. In general, a variety of particle dimension ranges can be employed.
cognition and societal considerations	Usually, c1 = c2 = 1.494. Other values may also be utilized, providing that 0 < c1 + c2 < 4.

Table 3. Particle swarm optimization convergence variables **(Lavanya and Udgata, 2011)**

Description	Details
A minimum relative enhancement (fm) of the objective function's value	Convergence has occurred (including the present run) when it is equal or falls below fm.
T max is the highest number of trials for the termination criterion.	In combination with other PSO variables (D, N) determined by the complication of the problem to be optimized.
Number of trials (kf) for which a convergence check is satisfied by the relative improvement in the function of the objective.	If the relative progress of the targeted function during the previous kf iterations exceeds a predetermined threshold, the kf iterations continue.

5. PROPOSED PSO MODEL FOR BOD OF INLET OF WASTEWATER TREATMENT PLANT

Typically, the optimization process employs a gradient-based local search algorithm. A starting point derived from a global search is required for the optimization process to be successful. Robust training procedures require initialization and optimization processes. **(Xie et al., 2012; Ye, 2017)** explained how the PSO algorithm can determine the optimal BOD at the inlet of Al-Rustumiya wastewater treatment plant.

1. Each particle in the problem's hyperspace is assigned a random place to begin the swarm's initialization.
2. The proposed model's objective function is evaluated for each particle.
3. Each particle's objective function value is compared to its pbest (pbest represents the optimal location of every particle in the distance). If the current value is greater than the pbest value, the current particle position, X_i, is set as pbest, and the current value is pbest.
4. It is determined which particle has the highest objective function value. It has been determined that its objective function returns the value gbest (gbest is the optimal placement to each particle globally) and that its location is gbest.
5. All particle positions and velocities are revised depending on the Eqs. (1) and (2).

6. Steps 2 to 5 will be reiterated till one of the standards for convergence (an utmost number of repetitions or an adequately acceptable objective function value) is met.

MATLAB code was used to simulate the proposed model to optimise the cost amount model for construction projects. Equation (3) is the proposed model to be optimized:

$$\mathrm{BOD} = \mathrm{F}_1 \times \mathrm{COD} + \mathrm{F}_2 \times \mathrm{SS} + \mathrm{F}_3 \times \mathrm{pH} + \mathrm{F}_4 \times \mathrm{Cl}^{-1} + \mathrm{F}_5 \times \mathrm{TDS} \quad (3)$$

where F_1, F_2, F_3, F_4, and F_5 are the unknown coefficients.

The primary objective of particle swarm optimization to optimize a BOD paradigm is to identify the optimal set of coefficients in a solution space. Consequently, there was little difference between the real BOD of building initiatives and the estimated final form of the optimized formulations. The PSO algorithm modifies its procedure until either a suitable guest or a predetermined maximum iteration has been completed. The parameters of the PSO model are shown in **Table 4.** The size of the swarm was modified to ascertain the most effective particle count in terms of convergence and processing time. In this study, 20, 40, 60, 80, and 100 swarm sizes were used to evaluate the accuracy of the suggested design. The number of repetitions is set at 500 because the differences in the objective function become constant after 87 rounds.

Table 4. Parameters of the PSO.

Parameters	Value
Swarm size	20, 40, 60, 80, and 100
Target error	1e-05
Iteration	500
C_1	1.494
C_2	1.494
W	0.7

The optimal coefficient factor values recommended by the proposed PSO model for the various swarm sizes are displayed in **Table 5**. There is a reasonable degree of concordance between the numerous testing methodologies, according to **Table 4**. The findings prove that the actual BOD model was more precise for the 100 colonies. Based on experimental results. **Table 5** demonstrates that the proposed PSO technique accurately predicts the BOD. The proposed model's estimations yielded a mean value of the estimated BOD 1.136, a standard deviation of 0.161%, and a coefficient of variation of 14.16%, confirming its precision and consistency. Comparisons of experimental data and model predictions for BOD are depicted in **Fig. 1**, indicating the recommended model's dependability in general.

According to **(Pimentel-Gomes, 2000)**, the coefficient of variation (CoV) value indicates the precision of the connection among the data outputs and inputs, with CoV amounts lower than 10%, 20–30%, and more than 30% corresponding to large precision, small precision, and small accuracy, accordingly. CoV for the suggested design was 14.16%, indicating high accuracy. The proposed model produced a value close to the mean of BOD estimation (1.136), which is near 1.0. In addition, the coefficient of determination (R^2) value of 0.8816 (shown in **Fig. 1**) and the CoV indicate that the observed and predicted BOD values correspond well. Based on these results, it is possible to conclude accurately estimates the BOD values while considering various parameters **(Al-Sulttani et al., 2017; Baki and Egemen, 2018).**

Following is the final recommended PSO model that has been optimized.

$$BOD_{predicted} = 0.41626 \times COD + 0.021169 \times SS + 0.998515 \times pH + 0.133357 \times Cl^{-1} + 0.029989 \times TDS \quad (4)$$

Table 5. Parameters used in the PSO -BOD model setting.

Factor	Swarm size				
	20	40	60	80	100
F1	0.309028	0.461140	0.410611	0.416260	0.476689
F2	0.056366	0.068520	0.011028	0.021169	0.048658
F3	0.849137	0.859876	0.950416	0.998515	0.361388
F4	0.163533	0.116927	0.178992	0.133357	0.007468
F5	0.032334	0.019243	0.030591	0.029989	0.059327
M	0.993	1.098	1.089	1.046	1.136
SD	0.167	0.159	0.156	0.149	0.161
CoV%	16.84%	14.50%	14.31%	14.20%	14.16%

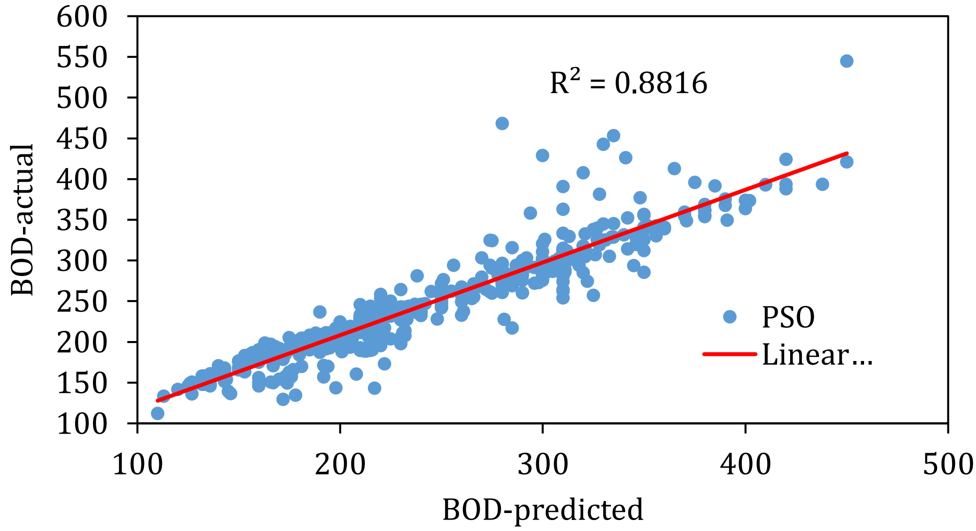

Figure 1. Comparisons between the predicted and experimental BOD values for the PSO model.

6. MULTIPLE LINEAR REGRESSION (MLR)

If it is believed that Y, the dependent variable, is influenced by m independent variables X_1, X_2 ...., X_m and a linear equation is chosen to represent their relationship, the regression equation for Y can be expressed as follows:

$$y = a + b_1 x_1 + b_2 x_2 + \cdots b_m x_m \quad (5)$$

y in this equation represents the variable's expected value. Y when the values of the independent variables are $X_1 = x_1,\ X_2 = x_2 \dots X_m = x_m$
Similar to straightforward regression, the regression coefficients a, b1, b2,..., bm are determined by minimizing the sum of the e_{yi} distances of the observation points to the plane as defined by the equation of regression **(Ross, 2020)**.

$$\sum_{i=1}^{N} e_{yi}^2 = \sum_{i=1}^{N} (y_i - a - b_1 x_{1i} - b_2 x_{2i} - b_m x_{mi})^2 \quad (6)$$

In this study, the coefficients of the regressions are determined by the least square method.

7. SENSITIVITY ANALYSIS

Sensitivity analysis is of the utmost importance for identifying the essential input variables. Sensitivity analysis is an effective method for evaluating the contribution of each predictor variable in the proposed model. **(Gandomi et al., 2013; Alavi and Gandomi, 2011)** proposed a sensitivity analysis procedure utilized in the current study to achieve this objective. Based on the following expressions, the output's sensitivity percentage to each input parameter is determined (**Gandomi et al., 2013)**:

$$N_i = f_{max}(x_i) - f_{min}(x_i) \quad (7)$$

$$S_i = \frac{N_i}{\sum_{j=1}^{n} N_j} \times 100 \quad (8)$$

where $f_{min}(x_i)$ and $f_{max}(x_i)$, accordingly, denote the lowest and highest value of the predicted output over the i th input domain, with other variables held constant at their respective mean values of the estimated BOD. **Table 6** presents the sensitivity analysis results conducted on the proposed models.
The chemical oxygen demand (COD) has the greatest impact on the biochemical oxygen demand (BOD) values in the proposed model, while chloride (Cl^-) is the second-most influential term. In addition, the pH has the least impact on BOD values.

Table 6. Sensitivity analysis results.

Sensitivity %	COD	Cl^-	TDS	SS	pH
PSO	81.31	7.19	5.97	4.79	0.73
MLR	85.48	6.33	5.26	2.60	0.33

8. COMPARISON OF THE RESULTS BETWEEN PSO AND MLR

To evaluate the proposed PSO and MLR models, 20% (140 records) of the total datasets are utilized. These data were not used in the model development procedure. This section presents the results derived from verification records and the suggested models. Moreover, the results demonstrated that the proposed composite model (PSO) outperforms the MLR model.
The standard deviation (SD) is measured as the data variance is assessed. The lower the data variance, the smaller the SD, and vice versa. Consequently, the coefficient of variation (CoV) measures the actual quantity of relative variation and reflects the accuracy of output and input data. According to **(Pimentel-Gomes, 2000)**, a CoV value of less than 10% denotes

high precision, whereas values of 20–30% and greater than 30% denote low precision. It was found that the CoV values for the two models (PSO and MLR) were 11.83% and 12.59%, correspondingly, with a high degree of precision in establishing objective principles. In addition, a value close to 1.0 was attained for both models' mean values of the estimated BOD (1.03 and 1.1). The PSO is marginally more accurate than the MLR technique.
Evaluating a model's suitability to estimate BOD appears crucial to consider both the mean and distribution of prediction errors of the estimated BOD. **(Chang et al., 2012)** used global statistics (R^2 and mean square error MSE) that lack information regarding error distribution as this study's statistical performance evaluation criterion. The robustness of the model is evaluated using additional performance evaluation criteria, such as Average Absolute Relative Error (AARE), which shows the BOD performance index. It is evaluated such as **(Dogan et al., 2008)**:

$$\mathrm{AARE} = \frac{1}{\mathrm{N}} \sum_{\mathrm{p}=1}^{\mathrm{n}} |\mathrm{RE}| \tag{9}$$

In which,

$$\mathrm{RE} = \frac{t_p - o_p}{t_p} \times 100 \tag{10}$$

t_p stands for the pattern's observable BOD of p^{th} testing, o_p for the BOD predicted by PSO for the pattern of p^{th} testing, and N for all testing patterns combined.
The AARE value decreases with increasing efficacy. The performance management of PSO outputs was evaluated by evaluating the coefficient of determination (R^2).

$$R^2 = \frac{BOD_o - BOD_s}{BOD_s} \tag{11}$$

where:

$$BOD_o = \sum_{p=1}^{n} (t_p - t_{mean})^2 \tag{12}$$

$$BOD_s = \sum_{i=1}^{n} (t_p - O_p)^2 \tag{13}$$

where t_{mean} represents the average BOD, MSE is defined as the average square error:

$$\mathrm{MSE} = \frac{1}{\mathrm{N}} \sum_{i=1}^{n} (t_p - O_p)^2 \tag{14}$$

The performance criteria for the MLR model and PSO model's test results are presented in **Table 7**. **Fig. 2** illustrates, based on the test data set, how effectively the MLR and a specified PSO model predicted BOD.

Table 7. The PSO model and MLR model's performance.

Performance	PSO model	MLR model
AARE (%)	9.41	12.99
MSE	1029.10	1236.91
R^2	0.86	0.78
Mean	1.03	1.1
SD	0.12	0.13
CoV (%)	11.83	12.59

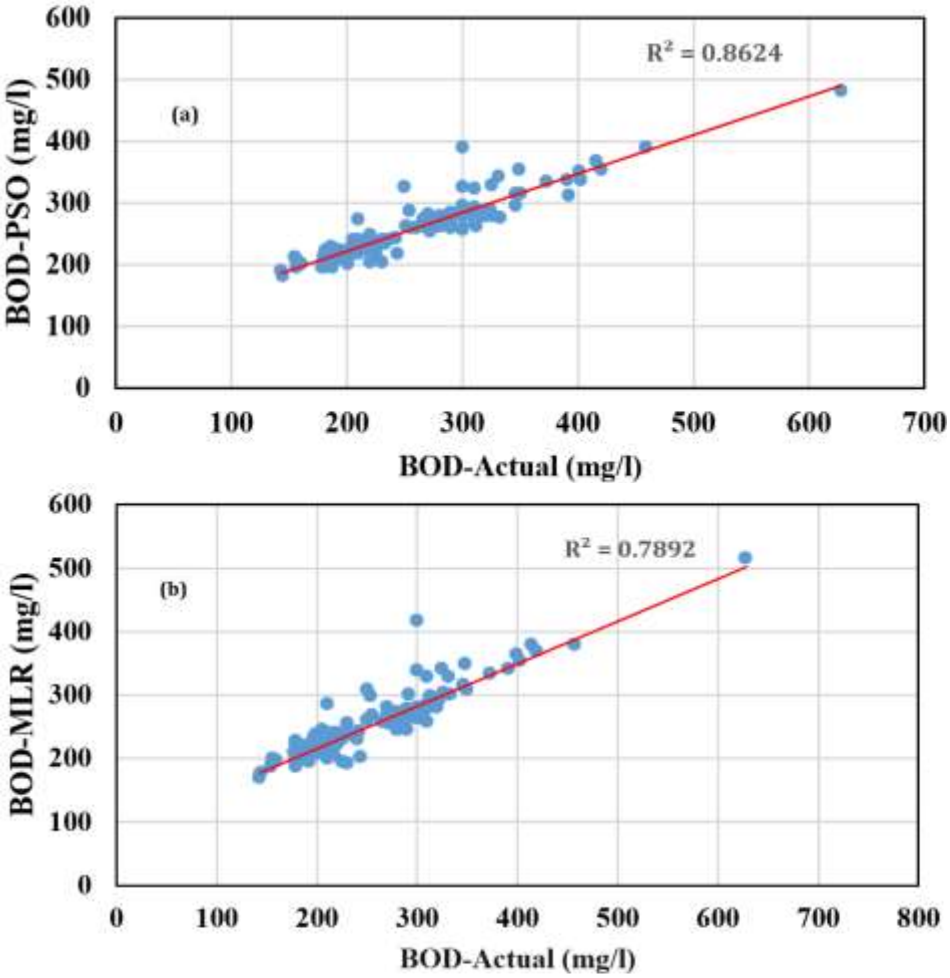

Figure 2. Comparison between actual and predicted values for the proposed models, a) PSO, b) MLR.

As depicted in **Fig. 2**, **a and b** have an acceptable R^2 value of 0.86 and 0.78, respectively, which indicates a close correlation with the empirically determined values. The comparison of the two numbers demonstrates this. Furthermore, PSO performs statistically better than MLR when estimating BOD.

Furthermore, the Absolute Relative Error ARE may be the most common method for evaluating the predictive capacities of the corresponding error variations **(Bagheri et al., 2012)**. ARE could be anticipated as follows:

$$ARE\% = \left|\frac{BOD_{act.} - BOD_{pred.}}{BOD_{act.}}\right| \times 100 \quad (15)$$

According to Eq. (15), the frequency should decrease proportionally as ARE% increases. As shown in **Fig. 3**, the proposed model has the lowest ARE for the largest frequency (less than 15%) and the highest ARE for the smallest frequency (greater than 20-25%). Therefore, the error distribution of the two predicted models is extremely satisfactory.

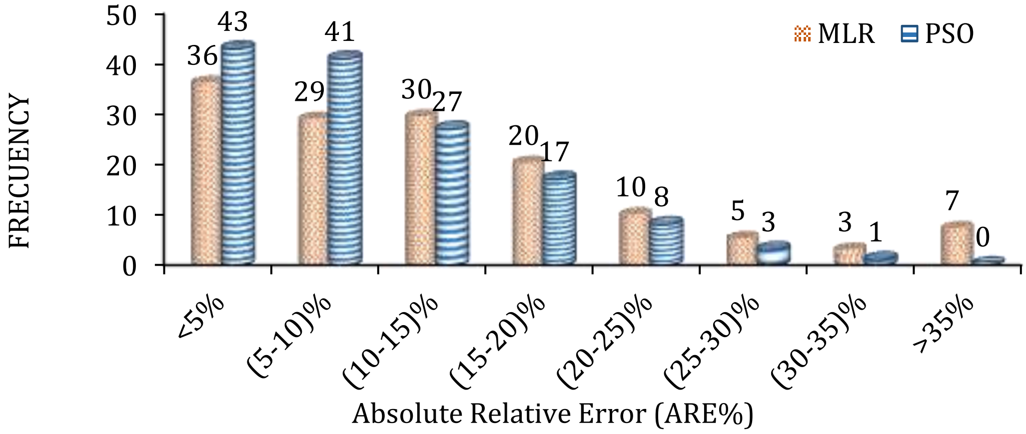

Figure 3. Absolute relative error (ARE) distribution for proposed models (BOD_{PSO} and BOD_{MLR}).

9. CONCLUSIONS

Current work demonstrates the BOD modelling capacities of the PSO and MLR models. Selecting PSO structure and input variables is essential to achieving highly determined precision. Consequently, a study on sensitivity has been performed using various performance statistics to ascertain the parameters' performance levels. Based on the results, a PSO model appears be a useful instrument to forecast the inlet BOD at Al-Rustumiya wastewater treatment plant. These findings indicate that COD predicts BOD more precisely than the other four variables (TDS, pH, SS, Cl^-). In order, the remaining variables were used to estimate BOD. After applying sensitivity analysis, the following effective parameters were identified: chloride (Cl^-), total dissolved solids (TDS), suspended solids (SS), and pH. Among the evaluated input combinations, the models with COD, TDS, pH, Cl^-, and SS as inputs have the greatest performance standards. These variables are required for more precise BOD modelling, as shown. In addition, the MLR method was used to predict BOD. Based on the comparison's outcomes for PSO and MLR, respectively, AARE% (9.41, 12.99), MSE (1029.10, 1236.91), R^2 (0.86, 0.78). Mean (1.03, 1.1), SD (0.12, 0.13), CoV% (11.83, 12.59), it was determined that the PSO technique is superior to the MLR method.

REFERENCES

Abbas, A.A.A., and Hassan, F.M., 2018. Water quality assessment of Euphrates river in Qadisiyah province (Diwaniyah river), Iraq. *The Iraqi Journal of Agricultural Science*, *49*(2), pp. 251-261. Doi:10.36103/ijas.v49i2.229

Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., and Mahfood, B., 2020. Artificial intelligence applications in solid waste management: A systematic research review. Waste Manag., 109, pp. 231–246. Doi:10.1016/j.wasman.2020.04.057

Al-Musawi, N.O.A., 2016. Application of Artificial Neural Network for predicting Iron concentration in the location of Al-Wahda water treatment plant in Baghdad city. *Journal of Engineering*, *22*(9), pp. 72-82. Doi: 10.31026/j.eng.2016.09.05.

Al-Obaidi, B., 2020. Predicting municipal sewage effluent quality index using mathematical models in the Al-Rustamiya sewage treatment plant. *Journal of Engineering Science & Technology*, *15*(6), pp. 3571-3587.

Al-Sulttani, A.O., Ahsan, A., Al-Bakri, B.A., Hason, M.M., Daud, N.N.N., Idrus, S., Alawi, O.A., Macioszek, E. and Yaseen, Z.M., 2022. Double-slope solar still productivity based on the number of rubber scraper motions. *Energies*, *15*(21), P. 7881. Doi:10.3390/en15217881

Al-Sulttani, A.O., Ahsan, A., Hanoon, A. N., Rahman, A., Daud, N., and Idrus, S.J.A.E., 2017. Hourly yield prediction of a double-slope solar still hybrid with rubber scrapers in low-latitude areas based on the particle swarm optimization technique. *Applied Energy*, 203, pp. 280-303. Doi:10.1016/j.apenergy.2017.06.011

Alavi, A.H., and Gandomi, A.H., 2011. A robust data mining approach for formulation of geotechnical engineering systems. *Engineering Computations*, *28*(3), pp. 242-274. Doi:10.1108/02644401111118132

Amarasinghe, H.A.U., Gunawardena, H., and Jayatunga, Y., 2017. Correlation between biochemical oxygen demand (BOD) and chemical oxygen demand (COD) for different industrial waste waters. *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, 21(2), pp. 42-48.

Bagheri, M., Bagheri, M., Gandomi, A.H., and Golbraikh, A., 2012. Simple yet accurate prediction method for sublimation enthalpies of organic contaminants using their molecular structure. *Thermochimica acta*, *543*, pp. 96-106. Doi:10.1016/j.tca.2012.05.008

Baki, O.T., and Egemen, A., 2018. Estimation of BOD in wastewater treatment plant by using different ANN algorithms. *Membrane and Water Treatment*, *9*(6), pp. 455-462. Doi:10.12989/mwt.2018.9.6.455

Bhagat, S.K., Tung, T.M., and Yaseen, Z.M., 2020. Development of artificial intelligence for modeling wastewater heavy metal removal: State of the art, application assessment and possible future research. *Journal of Cleaner Production*, 250, P. 119473. Doi:10.1016/j.jclepro.2019.119473

Chang, C.K., Azamathulla, H.M., Zakaria, N.A., and Ghani, A.A., 2012. Appraisal of soft computing techniques in prediction of total bed material load in tropical rivers. *Journal of earth system science*, *121*, pp. 125-133.

Corominas, L., Garrido-Baserba, M., Villez, K., Olsson, G., Cortés, U., and Poch, M., 2018. Transforming data into knowledge for improved wastewater treatment operation: A critical review of techniques. *Environ. Model. Softw.*, 106, pp. 89–103. Doi:10.1016/j.envsoft.2017.11.023

Dogan, E., Ates, A.,Yilmaz, E.C., and Eren, B.J.E.P., 2008. Application of artificial neural networks to estimate wastewater treatment plant inlet biochemical oxygen demand. Environmental Progress, 27(4), pp. 439-446. Doi:10.1002/ep.10295

dos Santos Coelho, L., 2010. Gaussian quantum-behaved particle swarm optimization approaches for constrained engineering design problems. *Expert Systems with Applications*, *37*(2), pp. 1676-1683. Doi:10.1016/j.eswa.2009.06.044

Frank, I.E. and Todeschini, R., 1994. *The data analysis handbook*. Elsevier.

Garcia, S.N., Clubbs, R.L., Stanley, J.K., Scheffe, B., Yelderman Jr, J.C., and Brooks, B.W., 2013. Comparative analysis of effluent water quality from a municipal treatment plant and two on-site wastewater treatment systems. *Chemosphere*, *92*(1), pp. 38-44. Doi:10.1016/j.chemosphere.2013.03.007

Jian-jun, X., Yan-chao, X., Li-mei, Y., Hai-long, Z.H.A.O., Zhi-gang, S.U.N., and Li-li, B.A.I., 2013. Research on the method of optimal PMU placement. *International Journal of Online Engineering*, *9*(6), pp. 24-29. Doi:10.3991/ijoe.v9iS7.3189

Kennedy, J., and Eberhart, R., 1995. Particle swarm optimization. *Proceedings of ICNN'95-international conference on neural networks*, 4, pp. 1942-1948. IEEE. Doi:10.1109/ICNN.1995.488968

Khazraji, S.D.A., and Nasser, N.O., 2012. Mathematical model for BOD in waste water discharges from Al Dora refinery in Baghdad. *Journal of Engineering*, *18*(12), pp. 1297-1306. Doi:10.31026/j.eng.2012.12.01.

Khudair, B.H., 2019. Influent flow rate effect on sewage pump station performance based on organic and sediment loading. *Journal of Engineering*, *25*(9), pp. 1-11. Doi:10.31026/j.eng.2019.09.1

Lavanya, D., and Udgata, S.K., 2011. Swarm intelligence based localization in wireless sensor networks. In *Multi-disciplinary Trends in Artificial Intelligence: 5th International Workshop, MIWAI 2011, Hyderabad, India, December 7-9, 2011. Proceedings 5,* pp. 317-328. Springer Berlin Heidelberg. Doi:10.1007/978-3-642-25725-4_28

Malviya, A., and Jaspal, D., 2021. Artificial intelligence as an upcoming technology in wastewater treatment: a comprehensive review. *Environmental Technology Reviews*, *10*(1), pp. 177-187. Doi:10.1080/21622515.2021.1913242

Mohammed, R., and Al-Obaidi, B., 2021. Treatability influence of municipal sewage effluent on surface water quality assessment based on Nemerow pollution index using an artificial neural network. In *IOP Conference series: earth and environmental science,* 877(1), P. 012008. IOP Publishing. Doi:10.1088/1755-1315/877/1/012008

Mohammed, R., Huseen, B., and Hashim, A., 2022. Effluent quality assessment monitoring of al-rustamiya sewage treatment plant using Geographical Information System (GIS). In *IOP Conference Series: Earth and Environmental Science,* 1002(1), P. 012011. IOP Publishing. Doi:10.3390/en15217881

Mohsin, R.M., Khudair, B.H., and Mohammed, A.H., 2021. Effective quality control of a municipal wastewater treatment plant using Geographic information systems: A Review. *Journal of Engineering*, *27*(7), pp. 66-72. Doi:10.31026/j.eng.2021.07.06

Pimentel-Gomes, F., 2000. Course of experimental statistics. *Piracicaba: FEALQ*, *15*.

Poli, R., Kennedy, J. and Blackwell, T., 2007. Particle swarm optimization: An overview. *Swarm intelligence*, *1*, pp. 33-57. Doi:10.1007/s11721-007-0002-0

Robles-Rodriguez, C.E., Ben-Ayed, A., Bernier, J., Rocher, V., and Dochain, D., 2019. Management of an integrated network of wastewater treatment plants for improving water quality in a river basin. *IFAC-Papers OnLine*, *52*(1), pp. 358-363. Doi:10.1016/j.ifacol.2019.06.088

Ross, S.M., 2020. *Introduction to probability and statistics for engineers and scientists*. Academic press.

Shi, Y., and Eberhart, R., 1998, May. A modified particle swarm optimizer. *IEEE international conference on evolutionary computation proceedings. IEEE world congress on computational intelligence (Cat. No. 98TH8360), May,* pp. 69-73. IEEE.

Xie, F., Wang, Q.J., and Li, G.L., 2012. Optimization research of FOC based on PSO of induction motors. *15th international conference on electrical machines and systems (ICEMS), October,* pp. 1-4. IEEE. https://ieeexplore.ieee.org/document/6401720

Yan, J., Chen, X., Yu, Y., and Zhang, X., 2019. Application of a parallel particle swarm optimization-long short term memory model to improve water quality data. *Water*, 11(7), P.1317. Doi:10.3390/w11071317

Ye, X., 2017. *Coupling of multi-agent based simulation and particle swarm optimization for environmental planning and decision making*. Doctoral dissertation, Memorial University of Newfoundland.

Ye, X., Chen, B., Storesund, R., and Zhang, B., 2021. System control and optimization in wastewater treatment: a particle swarm optimization (PSO) approach. *Soft Computing Techniques in Solid Waste and Wastewater Management,* pp. 393-407. Doi:10.1016/B978-0-12-824463-0.00027-6.