ORIGINAL ARTICLE

A novel experimental and machine learning model to remove COD in a batch reactor equipped with microalgae

Atef El Jery[1] · Ayesha Noreen[2] · Mubeen Isam[3] · José Luis Arias-Gonzáles[4] · Tasaddaq Younas[5] · Nadhir Al-Ansari[6] · Saad Sh. Sammen[7]

Abstract

By using microorganisms and the microalgae Chlorella vulgaris in conjunction with sequencing batch reactors (SBRs), the performance of a wastewater treatment facility was studied. For this purpose, the effect of pH, temperature, COD_{inlet}, and air flowrate on COD removal rate and residual was investigated. A single-factorial optimization method is utilized to optimize the amount of COD removal, and the best result is obtained with a pH of 8, $COD_{inlet} = 600$ mg/l, and an airflow rate of 55 l/min. Under optimal conditions, the amount of residual COD in the effluent reached 36 mg/l, showing an augmentation in the efficiency of the desired system. Moreover, empirical correlations are proposed for double-factorial optimization of residual COD and COD removal. Also, a multilayer perceptron artificial neural network is proposed to model the process and predict the residual COD concentration. The useful technique of hyperparameter tuning is utilized to obtain the best result for the predictions. All the effective parameters, including the number of hidden layers, neurons, epochs, and batch size, are adjusted. Data from the experiments agreed well with the artificial neural network modeling results. For this modeling, the values of the correlation coefficient (R^2) and mean absolute error (MAE) were obtained as 0.98 and 2%, respectively.

Keywords Sequential batch reactor (SBR) · Wastewater treatment · Microalgae · Chlorella vulgaris · Artificial neural network

Introduction

Water is one of the most important resources used by humans. Due to the increase in industrial activities and urbanization, high-quality water availability decreases daily. Along with the decrease in water resources, the release of industrial, agricultural, and urban wastewater in water-receiving environments is another factor that threatens the world's limited water resources. Therefore, wastewater treatment and reusing treated wastewater are essential (Rangasamy et al. 2007). There are different physical, chemical, and biological methods for wastewater treatment. Biological

✉ Nadhir Al-Ansari
  nadhir.alansari@ltu.se

  Atef El Jery
  ajery@kku.edu.sa

  Ayesha Noreen
  anoreen@ankara.edu.tr

  Mubeen Isam
  mubeenisam1990@gmail.com

  José Luis Arias-Gonzáles
  Joseariasgon6@gmail.com

  Tasaddaq Younas
  tygeologist@gmail.com

  Saad Sh. Sammen
  Saad123engineer@yahoo.com

1 Department of Chemical Engineering, College of Engineering, King Khalid University, Abha 61421, Saudi Arabia

2 Department of Social Environmental Sciences, Faculty of Language History and Geography, Ankara University, Ankara, Turkey

3 Building and Construction Techniques Engineering, Al-Mustaqbal University College, Hillah, Babil 51001, Iraq

4 Department of Social Sciences, Faculty of Social Studies, Pontifical University of Peru, San Miguel, Peru

5 Hassan Al Amir Soil Analysis, Dubai, United Arab Emirates

6 Civil, Environmental and Natural Resources Engineering, Lulea University of Technology, 971 87 Lulea, Sweden

7 Department of Civil Engineering, College of Engineering, University of Diyala, Baqubah, Diyala Governorate, Iraq

processes for removing organic materials are less costly and energy-consuming than physical and chemical methods (Roodbari et al. 2022). This system was evaluated using microorganisms and the microalgae Chlorella vulgaris with sequencing batch reactors (SBRs) (Bordbar et al. 2021). This method can greatly remove pollutants due to the possibility of influencing the microbial population and its simplicity and easy operation (Alimoradi et al. 2020). In addition, flexibility, efficiency, high capacity, cost-effectiveness, and no need for complex electrical systems have made it important to use this biological treatment method in urban and industrial wastewater treatment (Mazandarani et al. 2018). In all SBR systems, the treatment process is carried out in 5 consecutive stages: filling, aeration and reaction, sedimentation, discharge of wastewater output, and stagnation (Alimoradi et al. 2023). In the filling stage, raw sewage enters the reactor and mixes with the microorganisms. In the aeration stage, various biological reactions such as nitrification, denitrification, and phosphorus separation are performed; after that, the microorganisms are separated from the treated wastewater (sedimentation stage), and the treated effluent is discharged from the reactor (Hoffmann 1998). In addition to the existing purification methods, microalgae cultivation systems can also play a valuable role in wastewater treatment because microalgae can increase the removal of nutrients, heavy metals, organic substances, and pathogens from wastewater. The noteworthy point in wastewater treatment using microalgae is their coexistence and cooperation with microorganisms in the treatment process (Alimoradi et al. 2022a). In this connection, microalgae absorb the carbon dioxide gas produced from the metabolic activities of microorganisms and use it in photosynthesis to produce various compounds.

Hamta et al. (2022) explained and analyzed the functional mechanism underlying copolymer membrane formation from both a theoretical and experimental perspective. They investigated the styrene–acrylonitrile copolymer (SAN) membrane morphology. They also evaluated the effect of different membrane formation conditions on the membrane morphology analysis. In another study Hamta et al. (2021), they did an experimental study on the microfiltration membrane separation technology to remediate oil wastewater. They used oil rejection to assess the performance of the membranes in the O/W emulsion. The rejection was calculated, and the oil content in the permeate was examined using the chemical oxygen demand test. Bozorgnezhad et al. (2015a), Hasheminasab et al. (2014) studied the proton exchange membrane fuel cell (PEMFC) as a renewable energy source and replacement for internal combustion engines. They believe characterizing and researching the two-phase flow phenomena of PEMFC as floods are essential to effective water management. In another study (Bozorgnezhad et al. 2015b), they investigated water droplets as essential for effective water management in fuel cells. In their work, optical imaging and digital image processing were used to investigate the dispersion of water droplets in the cathode channel of a transparent PEMFC with a single-serpentine flow field under various input flow parameters and operational circumstances.

On the other hand, the oxygen produced by microalgae is used in the metabolic processes of microorganisms, so if microalgae are used along with other microorganisms to perform the purification process, it will lead to the improvement of the effluent quality and thus the efficiency of the purification system (Chamgordani 2022). Below are some examples of research conducted in wastewater treatment using microalgae. Lim et al. (2010) investigated the role of microalgae in wastewater treatment. This research resulted from consuming nitrogen and phosphorus compounds in wastewater by microalgae to grow and use their capacity to remove heavy metals. Delgadillo-Mirquez et al. (2016) used Chlorella vulgaris microalgae for the biological treatment of textile wastewater and investigated the issues related to color removal and reducing chemical oxygen demand (COD) in different wastewater concentrations. This research resulted in the color removal rate from 41.8 to 50%, and COD reduction from 38.2 to 62.3%. Feng et al. (2022) showed that microalgae could be used in urban wastewater treatment before entering the natural environment and lead to improved water quality. Govarthanan et al. (2020) showed that the use of Chlorella vulgaris microalgae in industrial wastewater treatment has high efficiency. It is worth noting that in the studies conducted, only the use of microalgae was limited in the matter of purification. However, in the present research, the combined use of microorganisms for the treatment of sanitary wastewater and the development of machine learning-based models to predict residual COD in the wastewater. However, not many studies have been done in this field.

In recent years, the use of data-driven methods has been significantly developed due to the existence of a large amount of data collected on the quantity and quality of wastewater and the need to reuse unconventional waters due to the lack of water resources (Singh et al. 2023; Zaboli et al. 2022; Agarwal et al. 2023; Alimoradi et al. 2022b, 2021; Naderi et al. 2021; Chamgordani et al. 2019). Among the data-oriented methods, artificial neural networks are very important and useful for water and wastewater quality problems with complex and nonlinear behavior. By using artificial neural networks, there is no need to describe the phenomena involved in the process mathematically. It can be used to simulate and predict nonlinear and complex behavior between inputs and outputs, which is useful in wastewater treatment processes. The application of these models can provide the basis for improving the efficiency of wastewater treatment plants and enable the rational and economic

efficiency of unconventional water resources and planning for the optimal combined use of these waters. The importance of using these models is very necessary due to the limitation of water resources and environmental problems. Nasr et al. (2012) predicted the performance of the system using an ANN. This study showed that the artificial neural network could predict the treatment plant's performance. Pai et al. (2007) designed a three-layer artificial neural network SBR to predict COD concentration and the number of suspended solids in the reactor. Kundu et al. (2013) used a multilayer artificial neural network to evaluate the reactor performance of wastewater treatment. For this purpose, they predicted the removal efficiency of COD and ammonia nitrogen with input variables. Based on the experimental results, the output error of the model was 33.3%, and the coefficient of determination (R^2) was higher than 0.94, the slope of the regression line was close to 1, and the width from the origin was close to zero, which shows the high potential of the artificial neural network in prediction. Luccarini et al. (2010) used the artificial neural network method to model the SBR reactor in their study. The presented neural network model with a multilayer perceptron structure was able to predict COD and volatile suspended solids concentrations during one cycle of this reactor and showed the high accuracy and ability of the neural network. Khatri et al. (2020) used a multilayer perceptron ANN to predict the COD concentration of the SBR reactor. Bagheri et al. (2015) investigated the performance of rock shore SBR in wastewater treatment using a multilayer perceptron neural network and radial basis function neural network. The network input parameters included mixed liquid suspended solids, dissolved oxygen, total suspended solids, total nitrogen, and temperature. Both training processes were very successful in modeling SBR. The training data set and modeling accuracy evaluation data showed a relatively good match between the experimental and modeling outputs. Rodríguez-Rángel et al. (2022) proposed a potential method for producing biofuels is the one-stage generation of carbohydrate-enriched microalgal biomass in wastewater. As a result, a model to predict biomass production in wastewater treatment systems was developed in their work using an analysis of artificial intelligence methods. They used five artificial intelligence techniques—artificial neural networks (ANNs), convolutional neural networks (CNN), long short-term memory networks (LSTMs), K-nearest neighbors (KNN), and random forest (RF)—to model carbohydrates. Singh and Mishra (2021) used machine learning algorithms to quickly find patterns and offer the ideal combinations in large datasets. They employed a decision tree algorithm to identify the effects and the best combination of predictor variables, including microalgae class, precultivation stage deciding factors, and operating variables, resulting in high biomass productivity and wastewater treatment capability. Alimoradi and Shams (2017) also carried out an optimization procedure on the subcooled flow boiling. The used an ANN model and genetic algorithm to predict the average vapor volume fraction in the outlet.

As a result of the inadequate performance of power plants' treatment plants, this research focuses on the use of microorganisms and Chlorella vulgaris microalgae together to improve the treatment of sanitary wastewater. As a first step, a pilot unit of the SBR reactor with microorganisms and Chlorella vulgaris microalgae will be constructed. To determine the optimal value of the COD index, variable parameters are checked, including pH, temperature, COD concentration in the incoming wastewater, and air flow rate. The development of an artificial neural network model based on a multilayer structure is another aspect of this research. In wastewater treatment processes, effluent quality prediction can assist in the improvement of the operation of wastewater treatment plants by minimizing microbial risks (Wang et al. 2021). Also, empirical correlations are proposed to find the best ranges of operating conditions for both residual COD and COD removal. By using artificial neural networks, there is no need to describe the phenomena involved in the process mathematically, and it can be useful in simulating and increasing the scale of complex biological systems (Iliopoulou et al. 2022).

Materials and methods

This reactor consisted of a wastewater tank, an aerobic and a sedimentation tank. This research deals with sanitary wastewater from a power plant, which is primarily human waste, washing water, and kitchen garbage. In the course of checking the pollution load of the incoming wastewater weekly, it was found that COD values at the beginning of the week, which coincided with an increase in the company's employee base, reached a maximum value of 1000 mg/l, and the minimum is 350 mg/l. The plexiglass tanks of aerobic and sedimentation zones had a diameter of 25 cm and 23 cm, respectively. The height of these tanks is 70 cm. As the sludge and suspended matter are settled in the clarifier area, the effluent is pumped into a pond. Also, the sludge on the tank's bottom was removed after several working cycles. As part of this reactor's flow cycle, there were 6 min of filling, 558 min of aeration, 30 min of sedimentation, and 6 min of discharge. Here is a view of the pilot reactor for an SBR used in Fig. 1.

In this research, distilled water was used to dilute the primary wastewater in different concentrations of COD, and COD concentration was measured using special vials, using the spectrophotometric method and spectrophotometer made by Hach. In order to aerate, an air compressor of 35 kPa, 100 W, and a maximum output flow rate of 100 ± 2 l/min, and a rotameter made by Fischer with the ability to measure

Fig. 1 The schematic view of an SRB pilot
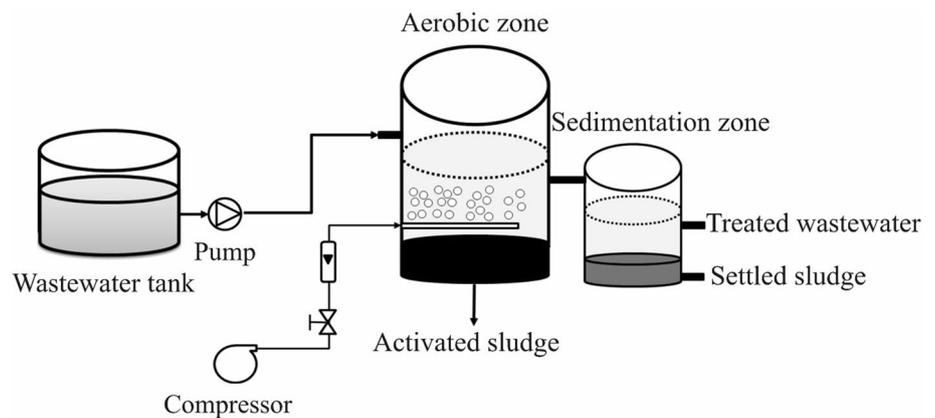
Table 1 The features of SBR

Parameter	Value
Time (hr)	2–10
Air flowrate (l/min)	15–75
pH (−)	4–10
Temperature (°C)	25–50
Inlet COD (mg/l)	350–1000

the flow intensity of 0–100 l/min are used to measure the airflow rate. Aqualytic made a digital thermometer for measuring the wastewater temperature and controlling its temperature using hot water. A Hach pH meter measured the pH of the system, and if necessary, sodium bicarbonate or sulfuric acid, both of which have a purity of 98%, was injected into the wastewater with an Alldos pump dosing system. For pumping wastewater to the reactor, a floor pump with a nominal flow rate of 5 m³/hr was considered.

Startup and operation of the SBR reactor

Chlorella vulgaris microalgae were prepared according to their quality and availability, and both were transferred to the reactor. The first and most important section of biological wastewater treatment is the adaptability of microorganisms to the wastewater. During the initial startup of the reactor, a small load is exerted to the reactor to improve the sludge activity. Initially, 350 mg/l of COD entered the system at startup. Microorganisms and microalgae adapted to the existing conditions for 20 days after the reactor was started. After that, the residual COD values in the effluent were obtained in 10-h cycles and at different pH, temperature, inlet COD, and air flowrate. To perform the desired tests, sampling was done from the system at one-hour intervals. There are a number of variables affecting the operation of the reactor studied, as shown in Table 1.

Results and discussion

Since wastewater treatment is affected by various factors, in the present study, a single-factorial method is used for the investigation of pH, inlet COD, temperature, and aeration. To calculate the COD removal, the following equation is employed.

$$\text{COD Removal} = \frac{\text{COD}_{\text{inlet}} - \text{Residual COD}}{\text{COD}_{\text{inlet}}} \tag{1}$$

The effect of inlet COD

It is important to investigate how changes in the sewage pollution load affect the performance of the treatment system, given that it may decrease or increase during the year. In order to test whether the change in pollution load impacts the system, wastewater with different CODs was prepared. The wastewater preparation with different CODs was obtained by diluting the wastewater in different concentrations. Figure 2a shows the variation of residual COD concentration in terms of time in various incoming wastewater COD, pH equal to 8, T of 35 °C, and Q of 85 l/min. As illustrated in Fig. 2a, the incoming wastewater has become more polluted, increasing residual COD. It is necessary to control the incoming COD load in an almost constant range. A continuous and adequate substrate would be provided for microorganisms and microalgal growth. As mentioned earlier, the maximum COD value of the power plant inlet wastewater was 1000 mg/l, and the minimum was 350 mg/l. The optimal COD for wastewater was considered to be 600 mg/l. In a study by Kapdan and Oztekin (Kapdan and Oztekin 2006), the effect of changing the initial concentration of COD in the range of 400–1800 mg/l on the rate of COD removal from textile industry wastewater was investigated. According to the obtained results, the highest COD removal occurred in the initial COD of 500 mg/l. At concentrations higher than 500 mg/l, the COD removal efficiency decreased.
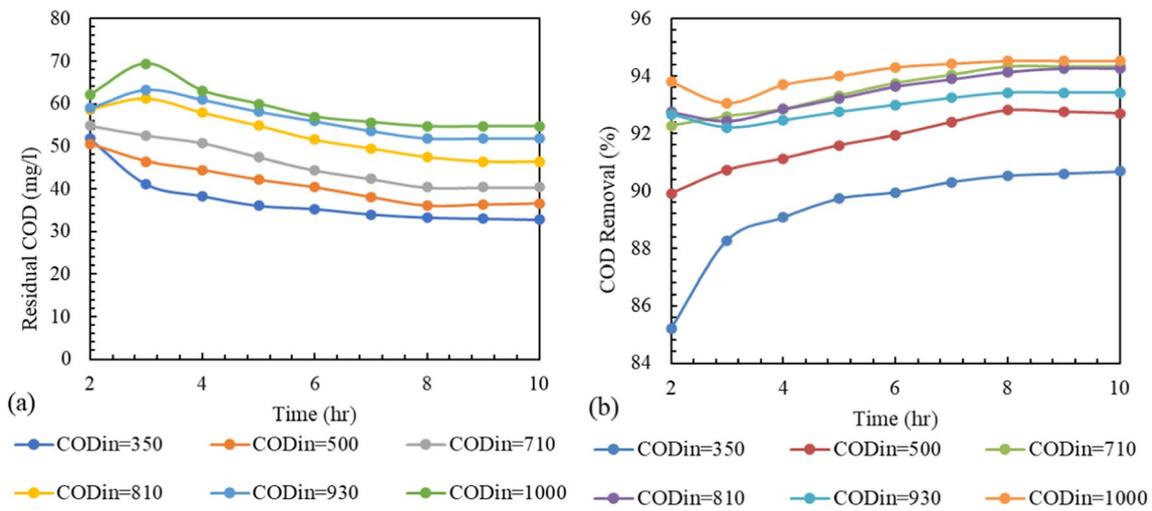
Fig. 2 The effect of COD inlet on **a** residual COD and **b** COD removal

In Fig. 2b, the amount of COD removal is presented. As can be seen, the highest efficiency of COD removal is achieved with inlet COD equal to 1000 mg/l. As the inlet COD increases, the COD removal increases. This is also observed when time is passed. In other words, the removal continues as the wastewater is kept in the reactor, but the slope of the increase is not the same, and it levels off after 6 h.

The effect of pH on the performance of the purification system

According to the report of Guo et al. (2015), the importance of pH characteristics is due to its relationship with biological activities. For this purpose, wastewaters with different pH were prepared to investigate the effect of pH change on the treatment system's performance. The amount of different pH was controlled by injecting a bicarbonate solution. Figure 3a shows the variation of residual COD concentration in terms of time and different pHs. The inlet wastewater temperature was 35 °C, the inlet COD was 600 mg/l, and the air flow rate was 85 l/min. After 2 h from the test, the amount of remaining COD is almost equal to 54 mg/l, as shown in Fig. 3a. The time at low pH began to increase after this period, reaching 123 l/min even at pH 4. Due to the fact that a process of digestion of organic substances within wastewater occurs at the beginning of the purification process with microorganisms and microalgae, it results in the formation of organic acids that further decrease the pH level of the system, resulting in an increase in residual COD and a decrease in the activity of microorganisms and microalgae. With the
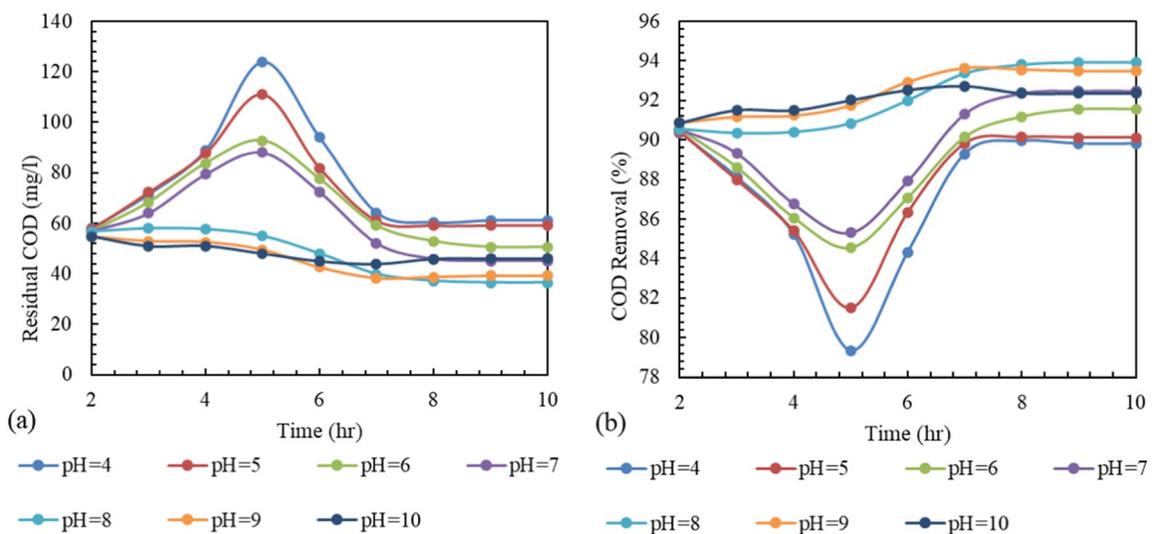
Fig. 3 The effect of pH on **a** residual COD and **b** COD removal

decomposition of acids generated and the decomposition of proteins and fats, pH and microorganism, and microalgal activity increase, and the residual COD decreases and is fixed at about 60 mg/l. The process of reducing residual COD at high pHs occurs with a suitable slope and eventually remains constant. A pH of 8 is optimal, and the residual COD, in this case, reaches the lowest value of 36 mg/l (See Fig. 3a). In a study by Khataee et al. (2010), the best filtration system performance using microalgae was obtained in the range of pH between 5.7 and 8.5.

The results of Fig. 3b show that the best result for COD removal is also achieved at the pH of 8. The amount of COD removal equals almost 94% after 10 h. Because of the mentioned reasons, at lower pHs, the removal drops significantly and then increases and levels off.

The effect of temperature on the performance of the purification system

Temperature is a parameter affecting the microbial activity of the system (Khataee et al. 2010). Considering that it is possible to witness inevitable changes due to various reasons, including seasonal changes and sewage temperature, necessary predictions should be made for these occasions. In the power plant, the desired temperature was set by using the boiling water of the boilers and adding it to the incoming sewage. The system's performance was tested at several temperatures to investigate the same issue. Figure 4a and b shows the variation of residual COD and COD removal in terms of time at various T and a pH of 8. The input COD is 600 mg/l, and the airflow rate is 85 l/min. According to Fig. 4a, 35–40 degrees is ideal for the system. Microorganisms and microalgae do not have access to oxygen and CO_2

when temperatures are high, which causes a decrease in their activity, and as a result, an increase in the residual COD is observed compared to other temperatures. This is also reflected in Fig. 4b, where the lowest removal is for an inlet temperature of 50 °C.

During low temperatures, however, residual COD increases because the advancement of microorganisms would drop. In most seasons of the year, the incoming sewage temperature is approximately 35 degrees, according to data from the power plant. On the other hand, according to Fig. 4a and b, there is no noticeable difference between the residual COD at temperatures of 35 and 40 °C, so this process's optimal temperature is considered 35 °C. The results of Azeez's research (Azeez 2010) show that using Chlorella vulgaris microalgae, the optimal temperature at which the highest COD removal occurred was 35 °C, which is consistent with the results obtained in this research.

The effect of aeration on the performance of the filtration system

In aerobic purification systems, aeration must be carried out optimally to create suitable conditions for the growth of microorganisms. When SBRs are used for purification, aeration becomes very important in COD removal efficiency. However, there has been limited research on the influence of this parameter on the COD removal rate (Chang et al. 2014). The desired wastewater with an input COD of 600 mg/l, pH equal to 8, and temperature of 35 °C was placed in different aeration conditions. A diffuser was used for aeration. An air compressor was used according to the characteristics of the diffuser used in the pilot and the pressure required for aeration. Figure 5 shows the variation of residual COD
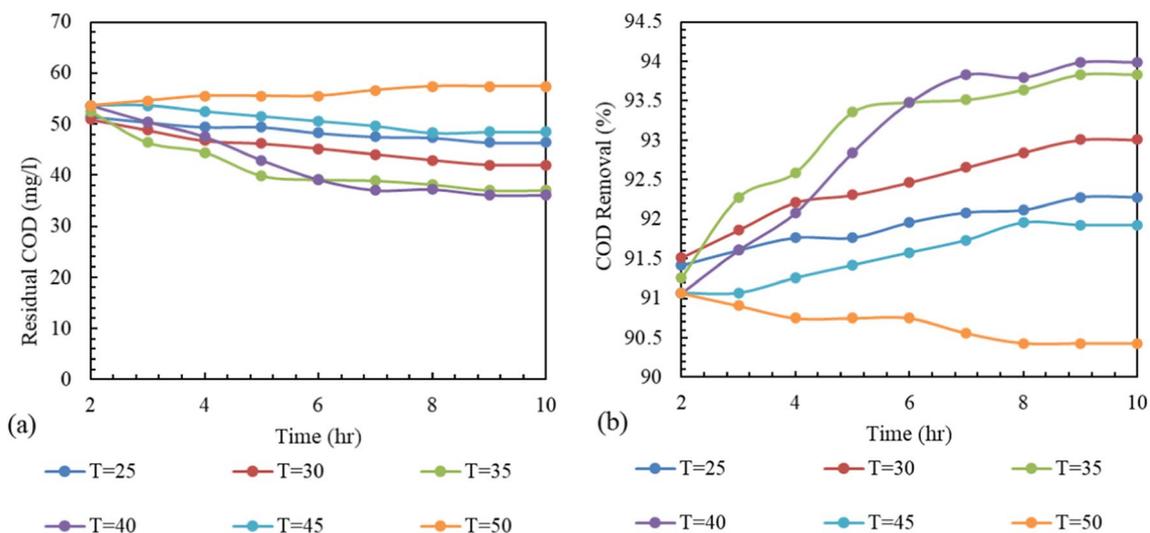

Fig. 4 The effect of temperature on **a** residual COD and **b** COD removal

concentration according to time and amount of different aerations.

As can be seen in Fig. 5a and b, the increase in Q has developed the system, and with the increase in the amount of aeration, the residual COD has decreased, and removal increased. Nevertheless, this is not very noticeable for Q higher than 55 l/min because higher flow rate, in addition to increasing costs, sometimes causes cell failure and, as a result, decreases the advancement of microorganisms. The optimum flow rate was selected as 55 l/min. In this discharge, the remaining COD has reached the lowest value. The results of studies by Primasari et al. (2011) also showed that COD removal increases with an increasing aeration rate. In this research, by increasing the amount of aeration from 1.5 to 2 l/min, the COD removal rate increased from 85.1 to 97.1%.

Optimization methods

Single-factorial optimization

As mentioned, the optimal case was obtained by examining the impact of different parameters on COD removal residual. These values included pH = 8, $T = 35$ °C, $COD_{inlet} = 600$ mg/l, and $Q = 55$ l/min. Under these conditions, the residual COD in the effluent reached 36 mg/l.

The correlations and optimization

In this section, high-accuracy empirical correlations for residual COD and COD removal are introduced. These models are based on time and other 4 effective parameters. The general polynomial equation of the models is shown as Eq. (2) (Azeez 2010).

$$ResidualCOD = A_1 + A_2x + A_3y + A_4x^2 + A_5xy + A_6y^2 + A_7x^3 + A_8x^2y + A_9xy^2 + A_{10}y^3$$

$$(2)$$

In Eq. (8), x is time, and y is the other parameters studied in the present study. The values for the constants for each y are presented in Table 2.

The optimized values of residual COD are presented in Fig. 6 based on the correlation of Eq. (2). As can be seen in Fig. 6a, the lowest amount of residual COD is observed when the inlet COD is minimum and the time is equal to 10 h. Therefore, based on the results of Fig. 6a, the best case is at inlet COD of 300 mg/l and after 10 h. Generally, the residual COD decreased as time passed, and using the minimum amount of inlet COD seems to be the best option.

In Fig. 6b, the effect of pH and time is investigated. To achieve the best result for residual COD, the best choice for pH is from 8 to 9, and the time should be around 9 h.

The effect of temperature and time is studied in Fig. 6c. The result shows that the residual COD minimized at early stages of the process. By increasing time the residual increased in general, so it seems logical to use inlet temperature of 50 °C at around 1 to 2 h.

Figure 6d indicates the effects of air flowrate and time, where the best case is around the flow rate of 40 to 60 l/min and the time higher than 6 h. It is understood that by increasing the air flow rate, residual COD is minimized. Moreover, as the time passed, a similar effect was observed.
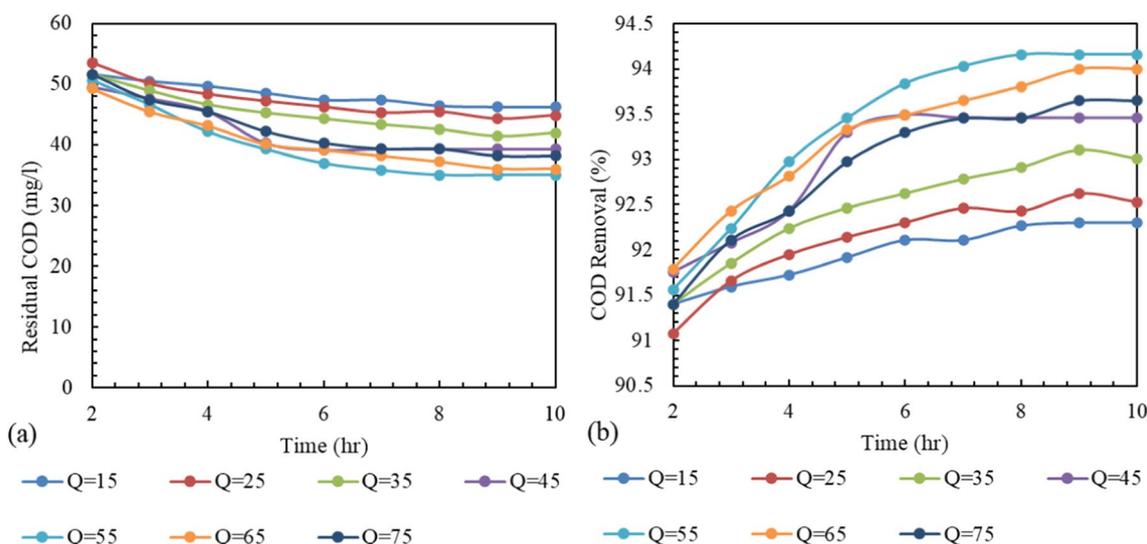

Fig. 5 The effect of temperature on **a** residual COD and **b** COD removal

Table 2 The constant values for residual COD for different y and the evaluation of the correlations using RSME and R^2

Values	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	RMSE	R^2
COD$_{inlet}$	45.24	− 3.896	0.0131	0.2237	− 0.0009453	2.14e-05	0	0	0	0	1.82	0.96
pH	− 234.5	79.99	91.45	− 9.736	− 7.172	− 11.84	0.3703	0.3435	0.2177	0.5157	7.53	0.85
T	90.7	− 1.895	0.04748	− 0.0382	0.1428	− 0.003023	− 4.363e-05	0.000475	− 0.002055	4.047e-05	0.3576	0.90
Q	48.53	− 1.903	0.8803	0.2463	− 0.08309	− 0.02259	− 0.009028	0.002978	0.0003417	0.0001784	0.87	0.97
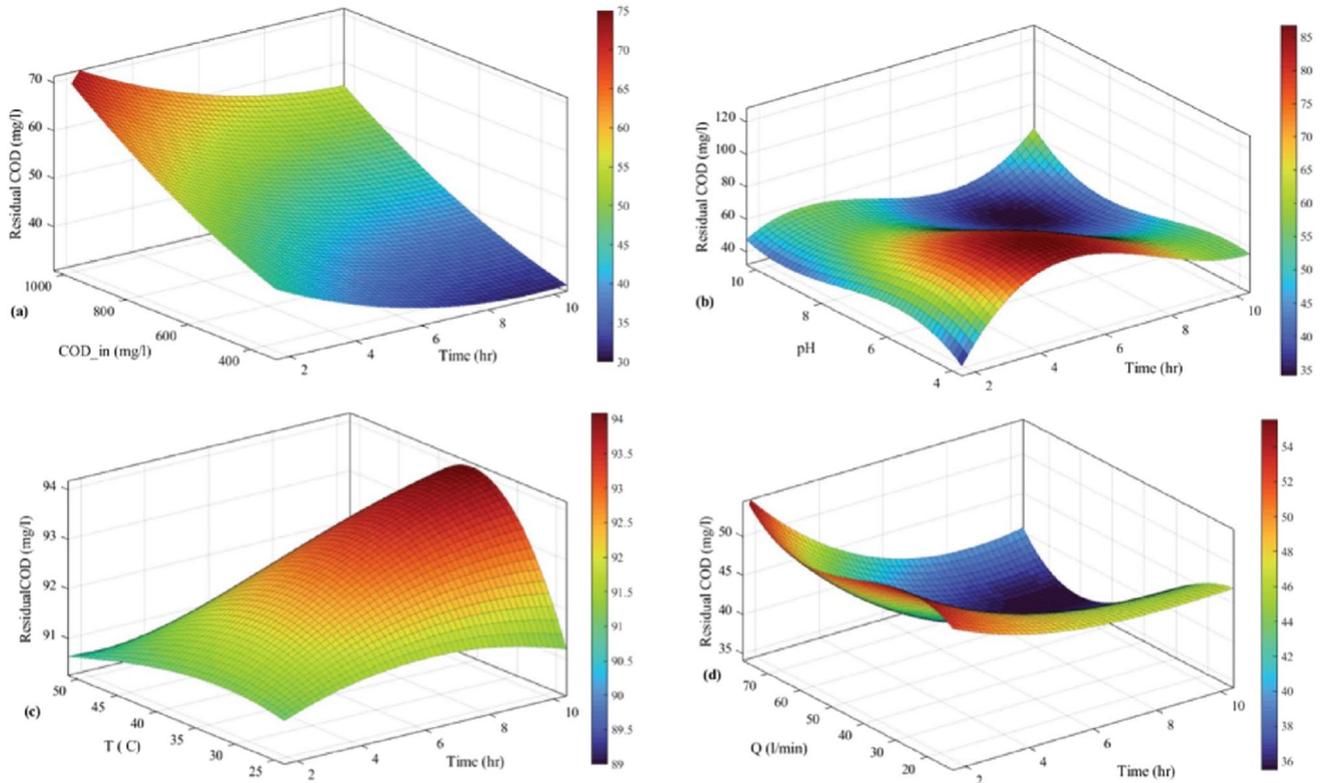
Fig. 6 Residual COD based on **a** inlet COD and time, **b** pH and time, **c** temperature and time, and **d** air flowrate and time

A similar analysis is presented for COD removal in Fig. 7. The coefficients of the correlation for COD removal are presented in Table 3, but the equation is the same as Eq. 1.

In the case of COD removal, we are looking for the maximum in contrast to residual COD. As can be seen in Fig. 7a, the highest COD removal is observed when the initial COD concentration is in the range of 800 to 1000 mg/l with a time higher than 8 h. It is concluded that increasing inlet COD would result in higher COD removal, but the required time also increases.

In Fig. 7b, the effect of pH and time is investigated. It is clear that the best range for COD removal is around pH of 7–9. However, to achieve the best case, at least 6 h is required.

The effect of inlet temperature and time is studied in Fig. 7c. The results show that the optimum case is when the inlet temperature is equal to 35 °C and the time is 10 h. As we deviate from this point, the COD removal decreases significantly.

Figure 7d indicates the effects of air flow rate and time. It seems like time plays a more dominant role in this comparison because as time increases, the COD removal augments. In addition, the highest removal is achieved at a flowrate of 50 l/min and higher.
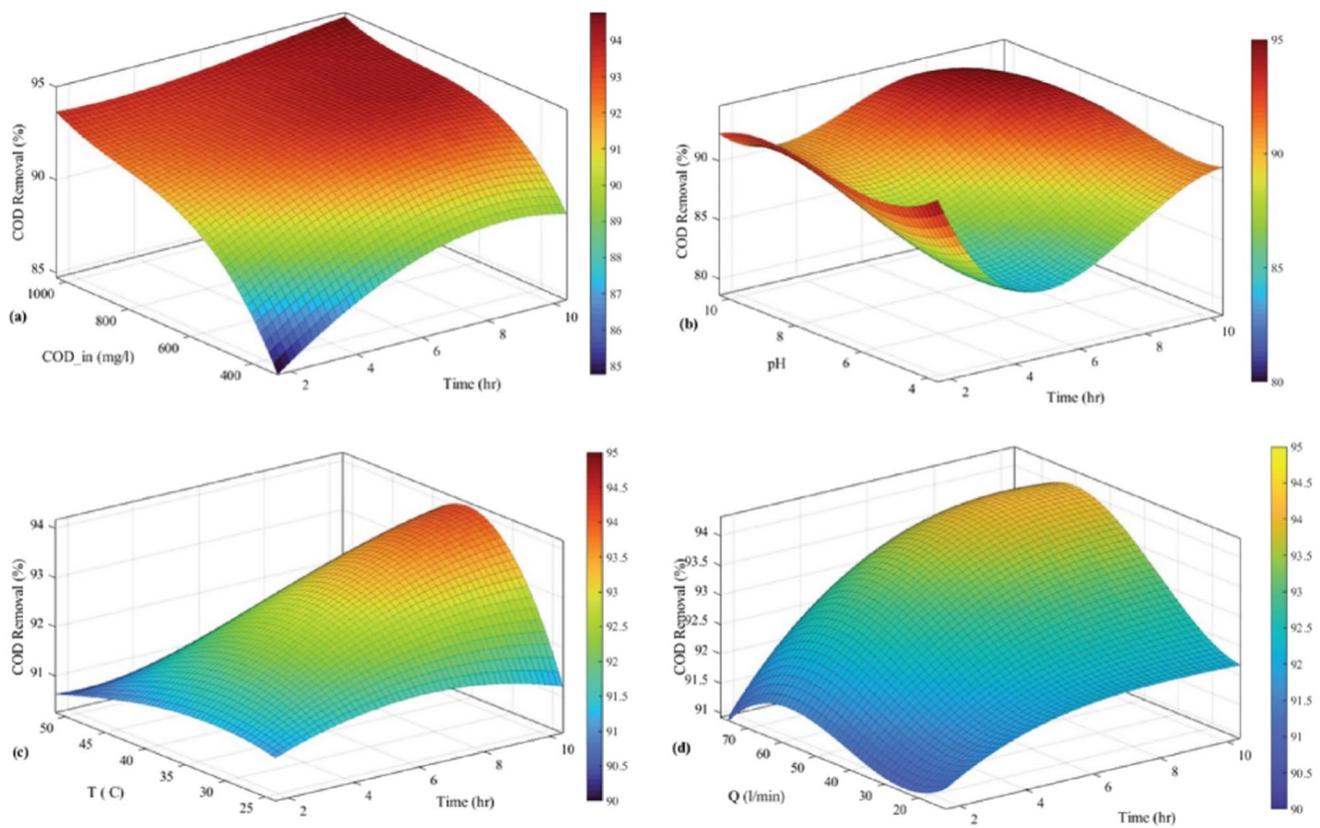
Fig. 7 COD removal based on **a** COD inlet and time, **b** pH and time, **c** temperature and time, and **d** air flowrate and time

Table 3 The constant values for residual COD for different y and the evaluation of the correlations using RSME and R^2

Values	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	RMSE	R^2
COD$_{inlet}$	62.41	2.501	0.09082	−0.1043	−0.00327	−9.967e-05	−0.0030	0.00017	4.89e-07	4.011e-08	0.4625	0.95
pH	140.3	−16.43	−14.05	2.07	1.325	1.802	−0.0776	−0.075	−0.0315	−0.08019	1.366	0.83
T	90.95	−1.904	0.02308	−0.0359	0.1429	−0.0022	−0.00015	0.000452	−0.00205	3.297e-05	0.3398	0.91
Q	91.88	0.3398	−0.1476	−0.045	0.0141	0.00377	0.00168	−0.000504	−5.9e-05	−2.98e-05	0.1373	0.98

Modeling results with artificial neural network

Reactor modeling with artificial neural network

A neural network is a mathematical structure that replicates biological neural systems. Each network consists of several interconnected units acting simultaneously. A neuron is a unit such as this. In neural networks, neurons constitute the smallest unit of processing information (Eskandari et al. 2022). A neural network model is generally composed of three layers: input, hidden, and output. Networks are trained using neurons with different transfer functions, which find connections between inputs and outputs (Ongen et al. 2013; Alimoradi et al. 2022c). The net input to the transfer function is the result of multiplying the input vector by its weights and adding the error value (Eskandari et al. 2022).

$$n_i = b_i + \sum_{j=1}^{m} \left(W_{in}(j,i) \times In_j \right) i = 1, 2, \dots, N' \quad (3)$$

where n_i is the input of each neuron, b_i is the error vector to the network, m is the number of network inputs, $W_{in}(j,i)$ is the network weights matrix (Eskandari et al. 2022).

$$a_i = f(n_i) \quad (4)$$

The output of the model is presented in Eq. (5) (Alimoradi et al. 2022c).

$$out = f' \left(b_{out} + \sum \left(W_{out}(i) \times a_i \right) \right) i = 1, 2, \ldots, N' \qquad (5)$$

In this regard, f' is the activation function of the output layer (Coulibaly et al. 2000).

In artificial neural network modeling, network prediction accuracy depends on the number of hidden layers, neurons, transfer functions, and network training, so these variables are chosen to optimize the network structure. However, in the present study, the hyperparameter tuning procedure is employed to bring more accuracy for the model. The number of neurons in the input and output layers is equal to the number of input and output variables of the process, respectively, and the optimal number of neurons in the hidden layer was obtained by a single-factorial optimization method. The values of pH, temperature, inlet COD, and air flow rate were considered as the input parameters, and COD removal and residual were the output parameters. The activation functions are mainly linear, sigmoid, and ReLU (Sebti et al. 2019; Pisa et al. 2019). Figure 8 shows an outline of the ANN used in the present research.

The Adam optimizer algorithm was used in this study to train the desired network. The choice of this algorithm was due to its convergence speed and high efficiency in network optimization and training (Sahoo and Ray 2006).

Using this algorithm, network output data are compared with experimental data. A correction is made to the model's weights based on the calculated error in previous layers. This process of modifying the weights in the network continues until the best weights that create the correct output for the system are identified and selected. In fact, this method obtains proper weights by regularly correcting the error. Once the most appropriate weights are determined, the network training stops, and the corresponding weights are fixed. The weights are applied to new data input related to model accuracy evaluation. In this case, the efficiency of the network is judged by comparing the model's results with the experimental values. In this study, we have used Python to implement the neural network. The obtained data from the experiments are randomly split, so seventy percent of the total data is used for training the model. The remaining data are used for evaluation.

Since in any experimental studies, machine and human error causes some of the obtained data to be illogical, in this research, all the experimental data, which included 240 data, were checked, and inappropriate and illogical data from the dataset were removed. At this stage, 210 data were obtained, which were normalized for modeling. In order to increase the speed of convergence.

Furthermore, the accuracy of the neural network requires that the input and output data be normalized. Using this relationship, the data were normalized between 0 and 1.
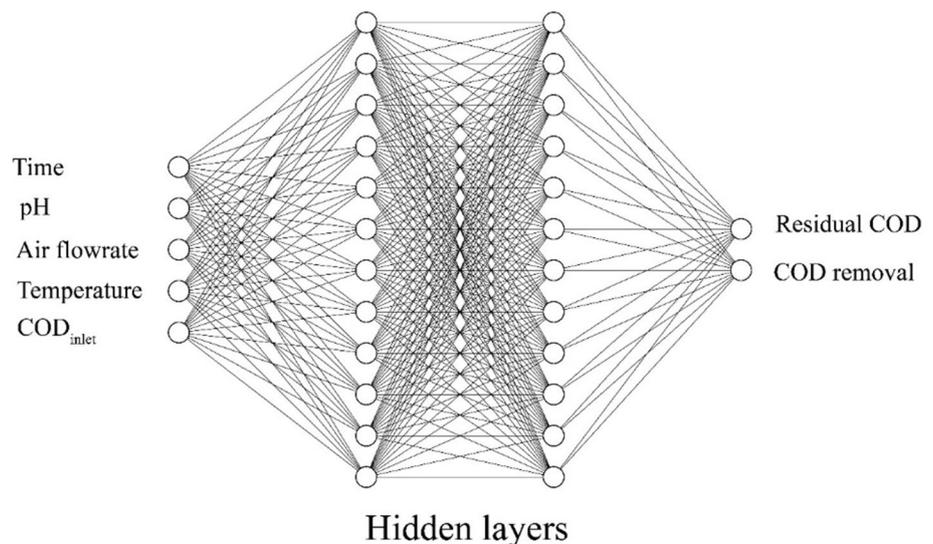
$$\overline{X} = \frac{X_i - X_{min}}{X_{max} - X_{min}} \qquad (6)$$

where $\overline{X}$ is the normalized data, X_i is the desired initial data, and X_{min} and X_{max} are the min and max of data in desired series (Xi et al. 2011).

To quantitatively examine the performance of the artificial neural network, the measures of mean absolute error (MAE) and correlation coefficient (R^2) were used as follows.

$$RSME = \frac{1}{N} \sum_{i=1}^{N} \left(Y_{pred} - Y_{exp} \right) \qquad (7)$$

Fig. 8 The ANN model of the present study

Time
pH
Air flowrate
Temperature
COD_{inlet}

Residual COD
COD removal

Hidden layers

Table 4 Hyperparameter tuning for layer structure

Number	Hidden layers	MAE (%)	R^2
1	(32)	3.12	0.96
2	(32,64)	2.99	0.97
3	(32,64,32)	2.84	0.97
4	(32,64,64,32)	2.64	0.97
5	(32,64,128,64,32)	2.48	0.98
6*	(32,64,128,128,64,32)	2.16	0.98
7	(32,64,128,256,128,64,32)	2.18	0.98
8	(32,64,128,256,256,128,64,32)	2.19	0.98
9	(32,64,128,256,512,256,128,64,32)	2.22	0.98

Table 5 Hyperparameter tuning for activation function

Model	Activation function	Mean absolute error (%)	R^2
1	Linear	2.91	0.97
2*	ReLU	2.23	0.98
3	Sigmoid	2.44	0.98

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(Y_{\text{pred}} - Y_{\text{exp}}\right)^2}{\sum_{i=1}^{N}\left(Y_{\text{exp}} - \overline{Y}\right)^2} \tag{8}$$

In the above relationships, Y_{pred} is the modeling output. If the MAE tends to zero and the correlation coefficient tends to 1, it will indicate the high accuracy of ANN (Ly et al. 2022; Vyas et al. 2015).

A total of 210 data were utilized in order to model a wastewater treatment system. In the first modeling stage, the number of hidden layers is investigated; generally, by adding more layers, the accuracy of the model increases. However, this could be a trap because by having a deeper network, the chances of falling into overfitting increase. Overfitting is a significant problem for most deep-learning models. Therefore, in the present study, we have used possible methods to avoid this problem. We have used dropouts and L2 regularizers to prevent this error. Table 4 presents the results for the number of hidden layers, and the optimum model is selected with the lowest MAE.

When the hidden layers are determined, the activation function of the output layer is investigated. The activation function for all the hidden layers is ReLU due to its high accuracy, but for the output layer, the sigmoid function seems to be the best case, according to Table 5.

Another hyperparameter that needs adjusting is the batch size. The batch size is the number of datapoints entering each feedforward process. Table 6 shows the results of models by the variation of the batch size.

Table 6 Hyperparameter tuning for batch size

Model	Batch size	Mean absolute error (%)	R^2
1	2	2.97	0.97
2	4	2.45	0.98
3	8	2.34	0.98
4	16	2.29	0.98
5*	32	2.23	0.98
6	64	2.86	0.97

Table 7 Hyperparameter tuning for epochs

Model	Epochs	Mean absolute error (%)	R^2
1	1500	6.25	0.88
2	2500	4.68	0.91
3	6000	3.45	0.95
4	15,000	2.58	0.98
5*	25,000	2.06	0.98
6	35,000	2.09	0.98
7	45,000	2.08	0.98
8	55,000	2.09	0.98

Table 8 The ANN settings for final models

| Hyperparameter	Residual COD	COD removal
Layers' structure	(32,64,128,64,32)	(32,64,128,128,64,32)
Batch size	16	32
Epochs	45,000	25,000
Activation function	Sigmoid	ReLU

According to the results, 32 is the optimal batch size, so this is chosen as the final model. In Table 7, the number of epochs is investigated to find the best result for this parameter.

Obviously, the selected parameter for the number of epochs is 25,000 because it has the least error and the highest R^2. Table 8 presents the final model after careful examination and setting of the hyperparameters.

Model visualization is done using the predicted results with respect to experimental ones. The best predictions should fit on the 45° line ($y = x$). The variation from this line shows the error of the models in predicting. Figure 9 shows the consistency of the actual COD residual values with the results obtained from the artificial neural network.
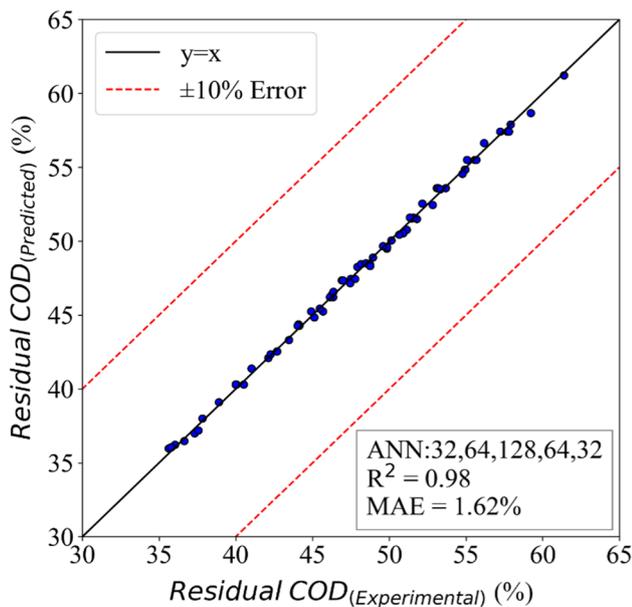
Fig. 9 The ANN prediction of residual COD
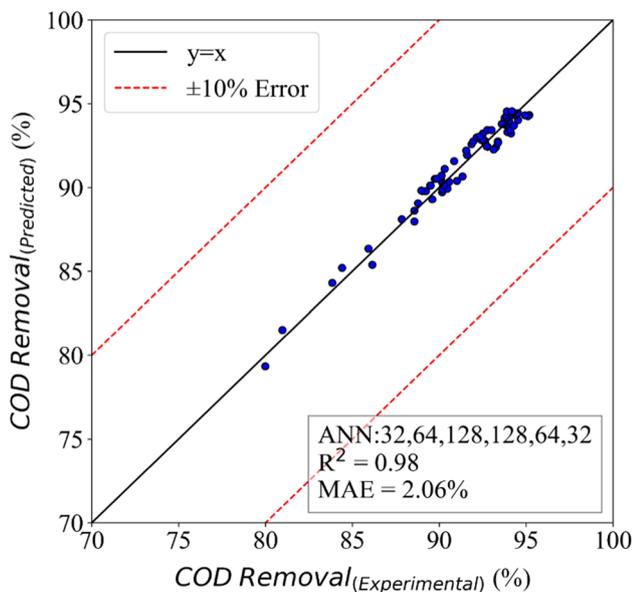
Fig. 10 ANN prediction of COD removal

Also, the COD removal is predicted using the ANN model. Figure 10 presents the results of COD removal with both experimental and predictive results.

In Table 9, there is a comparison between the experimental residual COD values with the results obtained from modeling in optimized conditions. It should be noted that this point was excluded from the training process of ANN models. The data point of the investigated case in Table 9 is $T = 35\,°C$, $COD_{inlet} = 600\,mg/l$, $pH = 8$, and $Q = 35\,l/min$.

Table 9 The comparison of experimental and predicted results for a local optimum condition

Time (hour)	Residual COD (mg/l)	
	Experimental results	ANN prediction
2	50.63	51.25
3	46.59	46.71
4	42.16	42.52
5	39.27	39.13
6	38.19	38.39
7	37.43	37.41
8	36.31	36.32
9	36.30	36.31
10	36.29	36.30

Comparing the computational cost of the ANN model with the experimental procedure, we concluded that the machine learning algorithm is very more effective. The time required for the ANN models is 10% of the time to carry out the experiment.

Conclusion

This research was conducted with the aim of improving the quality of wastewater by using wastewater with microalgae in the wastewater treatment of a combined cycle power plant. Using variables like pH, COD_{inlet}, temperature, and air flow rate, the SBR system was designed. In order to achieve optimal residual COD and COD removal values, the optimum of them was selected by analyzing the pilot system values. The results of the single-factorial optimization method showed that the system has the best performance at a pH of 8, a temperature of 35 °C, an inlet COD of 600 mg/l, and an air flowrate of 55 l/min. Under optimal conditions, the amount of residual COD in the effluent reached 36 mg/l, which indicates an increase in the efficiency of the desired system. These results indicate the optimal conditions for biological transformations by a collection of microorganisms and microalgae. Using time and residual COD, we proposed a double-factorial optimization for this parameter as well as the COD removal. The correlations were presented to model the changes of COD removal and residual in a range of mentioned parameters. In order to further investigate the SBR reactor, the study was also modeled with an artificial neural network. The input data of the neural network includes inlet COD, pH, temperature, and Q. Machine learning-based predictive models are also proposed for residual COD and COD removal, and the results had MAE and R^2 lower than 3% and higher than 0.97, respectively. The results of the predictive models showed satisfying agreement with the experimental results. The computational costs of the ANN models were

about 10% of the experimental procedure, so this method is way more efficient, and it is also very accurate. In the future research, other algorithms could be used to predict the objective function using the data from the present study. The comparative study of the machine learning model could be used to find the best model.

Declarations

Conflict of interest All authors declare no conflict of interest.

Ethical approval The authors did this research in compliance with ethical standards.

References

Agarwal S, Singh AP, Mathur S (2023) Removal of COD and color from textile industrial wastewater using wheat straw activated carbon: an application of response surface and artificial neural network modeling. Environ Sci Pollut Res. https://doi.org/10.1007/s11356-022-25066-2

Alimoradi H, Shams M (2017) Optimization of subcooled flow boiling in a vertical pipe by using artificial neural network and multi objective genetic algorithm. Appl Therm Eng 111:1039–1051

Alimoradi H, Soltani M, Shahali P, Moradi Kashkooli F, Larizadeh R, Raahemifar K, Adibi M, Ghasemi B (2020) Experimental investigation on improvement of wet cooling tower efficiency with diverse packing compaction using ANN-PSO algorithm. Energies 14(1):167

Alimoradi H, Shams M, Ashgriz N, Bozorgnezhad A (2021) A novel scheme for simulating the effect of microstructure surface roughness on the heat transfer characteristics of subcooled flow boiling. Case Stud Thermal Eng 24:100829

Alimoradi H, Shams M, Ashgriz N (2022a) Bubble behavior and nucleation site density in subcooled flow boiling using a novel method for simulating the microstructure of surface roughness. Korean J Chem Eng 39(11):2945–2958

Alimoradi H, Zaboli S, Shams M (2022b) Numerical simulation of surface vibration effects on improvement of pool boiling heat transfer characteristics of nanofluid. Korean J Chem Eng 39(1):69–85

Alimoradi H, Eskandari E, Pourbagian M, Shams M (2022c) A parametric study of subcooled flow boiling of Al2O3/water nanofluid using numerical simulation and artificial neural networks. Nanoscale Microscale Thermophys Eng 26(2–3):129–159

Alimoradi H, Shams M, Ashgriz N (2023) Enhancement in the pool boiling heat transfer of copper surface by applying electrophoretic deposited graphene oxide coatings. Int J Multiph Flow 159:104350

Azeez RA (2010) A study on the effect of temperature on the treatment of industrial wastewater using chlorella vulgaris alga. Algae 8:9

Bagheri M, Mirbagheri SA, Ehteshami M, Bagheri Z (2015) Modeling of a sequencing batch reactor treating municipal wastewater using multilayer perceptron and radial basis function artificial neural networks. Process Saf Environ Prot 93:111–123

Bordbar M, Naderi N, Alimoradi Chamgordani M (2021) The relation of entanglement to the number of qubits and interactions between them for different graph states. Indian J Phys 95:901–909

Bozorgnezhad A, Shams M, Ahmadi G, Kanani H, Hasheminasab M (2015a) The experimental study of water accumulation in PEMFC cathode channel. In: Fluids engineering division summer meeting (vol 57212, p V001T22A004). American Society of Mechanical Engineers

Bozorgnezhad A, Shams M, Kanani H, Hasheminasab M (2015b) Experimental investigation on dispersion of water droplets in the single-serpentine channel of a PEM Fuel cell. J Dispersion Sci Technol 36(8):1190–1197

Chamgordani MA (2022) The entanglement properties of superposition of fermionic coherent states. Int J Theor Phys 61(2):33

Chamgordani MA, Naderi N, Koppelaar H, Bordbar M (2019) The entanglement dynamics of superposition of fermionic coherent states in Heisenberg spin chains. Int J Mod Phys B 33(17):1950180

Chang GF, Hong W, Bian X, Liu B, Tan X, Su Y (2014) The effect of aeration rate on COD removal from high salinity wastewater in SBR process. Appl Mech Mater 522:605–608

Coulibaly P, Anctil F, Bobée B (2000) Daily reservoir inflow forecasting using artificial neural networks with stopped training approach. J Hydrol 230(3–4):244–257

Delgadillo-Mirquez L, Lopes F, Taidi B, Pareau D (2016) Nitrogen and phosphate removal from wastewater with a mixed microalgae and bacteria culture. Biotechnology Reports 11:18–26

Eskandari E, Alimoradi H, Pourbagian M, Shams M (2022) Numerical investigation and deep learning-based prediction of heat transfer characteristics and bubble dynamics of subcooled flow boiling in a vertical tube. Korean J Chem Eng 39(12):3227–3245

Feng J, Zhang Q, Tan B, Li M, Peng H, He J, Zhang Y, Su J (2022) Microbial community and metabolic characteristics evaluation in startup stage of electro-enhanced SBR for aniline wastewater treatment. J Water Process Eng 45:102489

Govarthanan M, Jeon CH, Jeon YH, Kwon JH, Bae H, Kim W (2020) Non-toxic nano approach for wastewater treatment using Chlorella vulgaris exopolysaccharides immobilized in iron-magnetic nanoparticles. Int J Biol Macromol 162:1241–1249

Guo H, Jeong K, Lim J, Jo J, Kim YM, Park JP, Kim JH, Cho KH (2015) Prediction of effluent concentration in a wastewater treatment plant using machine learning models. J Environ Sci 32:90–101

Hamta A, Ashtiani FZ, Karimi M, Sadeghi Y, MoayedFard S, Ghorabi S (2021) Copolymer membrane fabrication for highly efficient oil-in-water emulsion separation. Chem Eng Technol 44(7):1321–1326

Hamta A, Ashtiani FZ, Karimi M, Moayedfard S (2022) Asymmetric block copolymer membrane fabrication mechanism through self-assembly and non-solvent induced phase separation (SNIPS) process. Sci Rep 12(1):771

Hasheminasab M, Bozorgnezhad A, Shams M, Ahmadi G, Kanani H (2014) Simultaneous investigation of PEMFC performance and water content at different flow rates and relative humidities. In: International Conference on Nanochannels, Microchannels, and Minichannels (Vol. 46278, p. V001T07A002). American Society of Mechanical Engineers

Hoffmann JP (1998) Wastewater treatment with suspended and non-suspended algae. J Phycol 34(5):757–763

Iliopoulou A, Zkeri E, Panara A, Dasenaki M, Fountoulakis MS, Thomaidis NS, Stasinakis AS (2022) Treatment of different dairy wastewater with Chlorella sorokiniana: removal of pollutants and biomass characterization. J Chem Technol Biotechnol 97(11):3193–3201

Kapdan IK, Oztekin R (2006) The effect of hydraulic residence time and initial COD concentration on color and COD removal performance of the anaerobic–aerobic SBR system. J Hazard Mater 136(3):896–901

Khataee AR, Dehghan G, Ebadi A, Zarei M, Pourhassan M (2010) Biological treatment of a dye solution by Macroalgae Chara sp: effect of operational parameters, intermediates identification and artificial neural network modeling. Bioresour Technol 101(7):2252–2258

Khatri N, Khatri KK, Sharma A (2020) Artificial neural network modelling of faecal coliform removal in an intermittent cycle extended aeration system-sequential batch reactor based wastewater treatment plant. J Water Process Eng 37:101477

Kundu P, Debsarkar A, Mukherjee S (2013) Artificial neural network modeling for biological removal of organic carbon and nitrogen from slaughterhouse wastewater in a sequencing batch reactor. Adv Artif Neural Syst 2013:13–13

Lim SL, Chu WL, Phang SM (2010) Use of chlorella vulgaris for bioremediation of textile wastewater. Biores Technol 101(19):7314–7322

Luccarini L, Bragadin GL, Colombini G, Mancini M, Mello P, Montali M, Sottara D (2010) Formal verification of wastewater treatment processes using events detected from continuous signals by means of artificial neural networks. Case study: SBR plant. Environ Modell Softw 25(5):648–660

Ly QV, Truong VH, Ji B, Nguyen XC, Cho KH, Ngo HH, Zhang Z (2022) Exploring potential machine learning application based on big data for prediction of wastewater quality from different full-scale wastewater treatment plants. Sci Total Environ 832:154930

Mazandarani AH, Torabian A, Panahi HA (2018) Removal of codeine phosphate from water and artificial wastewater using sand modified with amine and carboxylic acid groups. Desalin Water Treat 131:261–271

Naderi N, Bordbar M, Hasanvand FK, Chamgordani MA (2021) Influence of inhomogeneous magnetic field on the qutrit teleportation via XXZ Heisenberg chain under intrinsic decoherence. Optik 247:167948

Nasr MS, Moustafa MA, Seif HA, El Kobrosy G (2012) Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-EGYPT. Alex Eng J 51(1):37–43

Ongen A, Ozcan HK, Arayıcı S (2013) An evaluation of tannery industry wastewater treatment sludge gasification by artificial neural network modeling. J Hazard Mater 263:361–366

Pai TY, Tsai YP, Lo HM, Tsai CH, Lin CY (2007) Grey and neural network prediction of suspended solids and chemical oxygen demand in hospital wastewater treatment plant effluent. Comput Chem Eng 31(10):1272–1281

Pisa I, Santín I, Vicario JL, Morell A, Vilanova R (2019) ANN-based soft sensor to predict effluent violations in wastewater treatment plants. Sensors 19(6):1280

Primasari B, Ibrahim S, Annuar MSM, Remmie LXI (2011) Aerobic treatment of oily wastewater: effect of aeration and sludge concentration to pollutant reduction and PHB accumulation. World Acad Sci Eng Technol 78:172–176

Rangasamy P, Iyer PVR, Ganesan S (2007) Anaerobic tapered fluidized bed reactor for starch wastewater treatment and modeling using multilayer perceptron neural network. J Environ Sci 19(12):1416–1423

Rodríguez-Rángel H, Arias DM, Morales-Rosales LA, Gonzalez-Huitron V, Valenzuela Partida M, García J (2022) Machine learning methods modeling carbohydrate-enriched cyanobacteria biomass production in wastewater treatment systems. Energies 15(7):2500

Roodbari M, Alimoradi H, Shams M, Aghanajafi C (2022) An experimental investigation of microstructure surface roughness on pool boiling characteristics of TiO2 nanofluid. J Therm Anal Calorim 147(4):3283–3298

Sahoo GB, Ray C (2006) Predicting flux decline in crossflow membranes using artificial neural networks and genetic algorithms. J Membr Sci 283(1–2):147–157

Sebti A, Boutra B, Trari M, Aoudjit L, Igoud S (2019) Application of artificial neural network for modeling wastewater treatment process. In: International conference in artificial intelligence in renewable energetic systems, pp 143–154. Springer, Cham

Singh V, Mishra V (2021) Exploring the effects of different combinations of predictor variables for the treatment of wastewater by microalgae and biomass production. Biochem Eng J 174:108129

Singh D, Tembhare M, Machhirake N, Kumar S (2023) Biogas generation potential of discarded food waste residue from ultra-processing activities at food manufacturing and packaging industry. Energy 263:126138

Vyas M, Modhera B, Sharma A (2015) BOD approximation for common effluent treatment plant using ANN. Asian J Water Environ Pollut 12(1):81–89

Wang J, Song A, Huang Y, Liao Q, Xia A, Zhu X, Zhu X (2021) Domesticating Chlorella vulgaris with gradually increased the concentration of digested piggery wastewater to bio-remove ammonia nitrogen. Algal Res 60:102526

Xi X, Cui Y, Wang Z, Qian J, Wang J, Yang L, Zhao S (2011) Study of dead-end microfiltration features in sequencing batch reactor (SBR) by optimized neural networks. Desalination 272(1–3):27–35

Zaboli S, Alimoradi H, Shams M (2022) Numerical investigation on improvement in pool boiling heat transfer characteristics using different nanofluid concentrations. J Therm Anal Calorim 147(19):10659–10676