Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2021.26481

209 (2021) 242–253
January

Determining the best and simple intelligent models for evaluating BOD_5 of Ahvaz wastewater treatment plant

Mojtaba Ghaed Rahmati, Parvaneh Tishehzan*, Hadi Moazed

Faculty of Water and Environmental Engineering, Shahid Chamran University of Ahvaz, Ahvaz, Postal Code: 6135743311, Iran, emails: ptishehzan@yahoo.com/partishehzan@scu.ac.ir (P. Tishehzan), Mojtaba.ghaedrahmati@gmail.com (M. Ghaed Rahmati), Hmoazed955@yahoo.com (H. Moazed)

Received 14 January 2020; Accepted 22 August 2020

ABSTRACT

The biochemical oxygen demand (BOD_5) could be used as an indication of wastewater treatment quality, but measuring BOD_5 is very time-consuming and costly. Ahvaz wastewater treatment plant (A-WWTP) plays a pivotal role in reducing the input load to the Karun River and it is very important to check its efficiency. Thus, the most critical parameters affecting the BOD_5 were determined using the linear regression and stepwise method. The capability of the multivariate linear regression model (MLR), feed-forward artificial neural network (FF-ANN), and adaptive neuro-fuzzy inference system (ANFIS) were investigated with different architectures and inputs to predict the effluent BOD_5 of A-WWTP (for daily and monthly modes). These architectures had two, three, four, or five inputs. The results of the MLR revealed that the maximum correlation coefficients (R) for training and testing were 0.916 and 0.864 (daily), and 0.809 and 0.793 on a monthly basis, respectively. The maximum R in FF-ANN for training and testing was 0.960 and 0.906 (daily basis), and 0.921 and 0.849 (monthly basis), respectively. Meanwhile, the maximum R in ANFIS for training and testing was 0.980 and 0.933 daily, and 0.968 and 0.927 monthly, respectively. The results indicated that the three models are appropriate, but the ANFIS is a more accurate model. In addition, based on conditions and available wastewater qualitative parameters, all of the architectures can be used to estimate the output BOD_5.

Keywords: Adaptive neuro-fuzzy inference system; Biochemical oxygen demand; Feed-forward artificial neural network; Regression; Wastewater

1. Introduction

The raw urban wastewater contains various pollutants and has an enormous environmental impact. Thus, to meet the regulation standard for discharge in surface water resources or for reusing the treated wastewater in agriculture and industry, the wastewater treatment plants have been established.

One of the major indications of treated plant performance is biochemical oxygen demand (BOD_5). Thus, measuring and monitoring BOD_5 continuously in wastewater treatment plants are essential. However, laboratory methods for BOD_5 measuring are time-consuming and costly. Thus, advanced forecasting techniques are employed for the assessment.

One of the most straightforward modeling techniques is the use of multivariate regression. Belhaj et al. [1] used the multivariate linear regression (MLR) to model the behavior of the Sfax wastewater treatment plant in Southeastern Tunisia. They developed models with a high level of approximation and accuracy, with the coefficients of determination (R^2) of 0.973, 0.946, and 0.925 for BOD_5, chemical oxygen demand (COD), and total suspended solids (TSS), respectively.

On the other side, the artificial neural network (ANN) is flexible and straightforward to analyze the non-linear

* Corresponding author.

problems of environmental science [2]. The study of literature from 1995 to 2019 indicated that the prediction ability of artificial intelligence technology should be strengthened by modifying important parameters of the wastewater treatment process, to provide operators with the opportunity to efficiently manage parameter shocks and ensure wastewater discharge water quality standards [3]. In recent years, many researchers have been using ANN to predict some parameters such as BOD_5 of industrial and domestic wastewater plants effluent [4–12]. Zare Abyaneh [13] investigated the efficiency of MLR and ANN models in the prediction of two major wastewater quality parameters, BOD_5 and COD, in the Ekbatan wastewater treatment plant (Tehran). The performance of these models was evaluated using the coefficient of correlation (*R*), and root means square error (RMSE). The results showed that the ANN model performed far better than the MLR. Also, Gaza wastewater treatment efficiency was determined considering influent input values of pH, temperature (*T*), BOD_5, COD, and TSS with influent output values of BOD_5, COD, and TSS. The results revealed that the performance of the ANN model was better than that of the MLR model [14].

There are several methods in the training of ANNs that affect ANN performance. In this regard, there are three methods called multi-layered (ML-ANN), teaching-learning based algorithm, and artificial bee colony algorithm (ABC-ANN) which had been applied to estimate BOD_5 of the wastewater treatment plant in Turkey [15]. The input flow (*Q*), *T*, pH, COD, suspended sediment, total phosphorus (TP), total nitrogen (TN), and electrical conductivity (EC) was used as the input parameters to estimate the BOD_5. The result indicated that the ML-ANN method provided the best estimation of both training and test series with $R^2 = 0.8924$ and $R^2 = 0.8442$, respectively.

The results of these studies revealed that ANN could predict wastewater plant efficiency, but data uncertainty caused researchers to develop other models such as a genetic algorithm to simulate the behavior of wastewater treatment plants (WWTP) [2,16]. Shokri et al. [17] applied two models of Mamdani fuzzy and Sugeno for evaluating the Tabriz wastewater treatment plant. They obtained the *R* of 0.91 and 0.94 for BOD_5 and TSS, respectively. These researchers concluded that both models could verify the performance of the treatment plant as well as the problems related to the treatment plants, including uncertainty. Thus, they suggested fuzzy methods when studying the treatment plants. Nadiri et al. [18] introduced a supervised committee fuzzy logic (SCFL) model as a predictive ensemble model for effluent water quality. The SCFL model used an ANN to combine forecasted water quality resulting from individual fuzzy logic (Takagi-Sugeno, Mamdani, and Larsen). Civelekoglu et al. [19] compared ANN and adaptive neuro-fuzzy inference system (ANFIS) for estimating the effluent COD. The results overall indicated that the ANFIS modeling approach may be suitable to describe the relationship between wastewater quality parameters and may have potential applications for performance prediction and control of aerobic biological processes in wastewater treatment plants. A comparison of the results of ANFIS and ANN models with an identical structure indicated that both models were suitable for activated sludge system simulation, but ANFIS is more efficient and offers better results than the ANN model [20]. Elsewhere, three different artificial intelligence-based non-linear models, that is, feed-forward neural network, ANFIS, support vector machine approaches, and MLR methods were applied for predicting the performance of the Nicosia wastewater treatment plant. The obtained results of single models proved that the ANFIS model provides effective outcomes in comparison with individual models [21].

Akilandeswari and Kavitha [22] used the ANFIS and compared it to the MLR model to estimate the output wastewater BOD_5 in the textile industry. They found that the ANFIS had better performance than the MLR model.

Research examined the operation of aerobic granular sludge (AGS) reactors via ANFIS and support vector regression. The results indicated the potential of artificial intelligence for developing predictive models for the AGS process and provided insight into the selection of the appropriate algorithms for these models [23].

Previous studies simulating wastewater treatment efficiency have provided the best architecture for each study. Nevertheless, sometimes, operators of WWTP could not measure all wastewater quality parameters. The work problems of WWTP, especially in the city of Ahvaz, require that we have several forecasting models with different numbers of input data for different conditions in order to control the efficiency of the treatment plant for better management. Ahvaz wastewater treatment plant (A-WWTP) plays a substantial role in reducing the input load to the Karun River, but not all of the quality parameters are measured daily. Knowing several intelligent models with different architectures and inputs can help operators of WWTP to evaluate the wastewater treatment performance. Simplicity and user-friendliness of the model are two important indicators for developing and applying models. Therefore, in this study, we tried to examine the simplest and most widely used models. Accordingly, this paper focused on four architectures for three models (MLR, feed-forward artificial neural network (FF-ANN), and ANFIS). Further, these models were applied to simulate BOD_5 for both daily and monthly modes.

2. Materials and methods

2.1. Study area

The study area covers A-WWTP, Iran. The data used were for 8 y (96 months) leading to March 2018. The wastewater sampling is not done every day. On the other hand, some data were outlier or the measurement was not complete. Therefore, after the initial data analysis, 676 data were obtained for daily simulation. The total number of monthly data is 96 months. The primary processes of this treatment plant include screening, equalization tank, primary settling tank, aeration lagoon, and activated sludge reactor. The flow rate is 50,000 cubic meters per day. Most of the wastewater quality parameters are measured manually. All chemical analytical methods are according to the standard methods.

2.2. Multivariate linear regression

Statistical methods, such as MLR models, are simple and useful tools for investigating any relationship between dependent and independent parameters [24]. MLR is based on the least-squares. In the best model, the sum of square error between observed and predicted parameters should be minimum. In this study, the MLR was used to model the linear relationship between a dependent parameter and two or more independent parameters.

2.3. Artificial neural network

ANN is a simplified model of the human brain. This model applies as a mathematical structure to display the nonlinear relations between the inputs and outputs. A neural network is composed of neural cells called neuron and communication units called an axon. The neurons of ANN are a simple form of biological neurons. Each ANN is composed of three layers, including input, output, and hidden layer. There are some neurons as processing units on each of these layers, which are connected through some weighted connections. The operations of each neuron are as follows: (1) the neuron collects all the inputs arrived to the cell, (2) the neuron reduces the neuron threshold value, (3) the neuron passes them across a stimulus function or activity function, and (4) the neuron output is created. Activity functions are used for transferring the outputs of each layer to the next layer. The feed-forward neural network modeling technique is the most widely used ANN type in water resources applications [25].

In this paper, the FF-ANN model was applied with 12 train functions and three activity functions (tansig, logsig, and purelin functions). For avoiding overtraining, the maximum number of RMSE increase and the number of repetitions was considered 6 and 1,000, respectively. The best structure was selected based on the best statistical indices.

2.4. Adaptive neuro-fuzzy inference system

Fuzzy logic does not have a systematic process for fuzzy controllers. Thus, Jang et al. [26] presented the ANFIS. ANFIS acts based on the changes in the value and range of membership functions at different iterations to achieve an appropriate network based on minimum error. ANFIS uses Takagi–Sugeno inference method. The number and type of inputs and form of membership functions are the factors affecting the neural fuzzy model [26].

In this study, for simulation with ANFIS, eight input membership functions and two output membership functions were used (linear and constant). Also, the error tolerance equaled zero, and the number of replications was 1–50. The best structure was selected based on the best statistical indices.

2.5. Training and testing dataset

Experimental data sets were either divided into three parts (training, validation, and testing) or two parts (training and testing) [3]. According to Hamed et al. [5], Moral et al. [27], and Solgi et al. [28], data are classified into training and test data in each type of neural network with a favorite method and type of architecture. For this purpose, 75% of data were used for training and 25% for the test in this study.

2.6. Input data preparation for FF-ANN and ANFIS

Since entering raw data reduces the network speed and accuracy, the data normalization method was used to prevent the minimization of weights and early saturation of neurons. Based on the normalization method, each number was converted into a number between zero and one to be used in the neural network function [29]. Researchers use different equations for normalization, and here the following five equations were used to study the accuracy of equations. Eqs. (1) and (5) were presented by Fathi et al. [30], Eq. (2) by Nourani and Komasi [31], Eq. (3) by Asadi et al. [32], and Eq. (4) by Haghdadi et al. [33]. After normalizing the current data by Eqs. (1)–(5) and obtaining the linear regression, it was found that Eq. (1) has the best performance for data normalization.

$$y = 0.5 + \left(0.5 \times \left(\frac{X - \bar{X}}{X_{\max} - X_{\min}}\right)\right) \tag{1}$$

$$y = \left(\frac{X - X_{\min}}{X_{\max} - X_{\min}}\right) \tag{2}$$

$$y = 0.05 + \left(0.95 \times \left(\frac{X - X_{\min}}{X_{\max} - X_{\min}}\right)\right) \tag{3}$$

$$y = 0.1 + \left(0.8 \times \left(\frac{X - X_{\min}}{X_{\max} - X_{\min}}\right)\right) \tag{4}$$

$$y = 0.5 + \left(0.5 \times \left(\frac{X - X_{\min}}{X_{\max} - X_{\min}}\right)\right) \tag{5}$$

where X represents the desired data, $\bar{X}$ denotes the data mean, $X_{\max}$ shows the maximum data, $X_{\min}$ shows the minimum data, and y is the normalized data. The most significant parameters affecting the output wastewater BOD_5 were determined using the linear regression and stepwise method via SPSS 21 software. The MLR model was investigated with SPSS 21 software and the MATLAB 2013b software was used to develop FF-ANN and ANFIS models.

2.7. Statistical evaluation indices

The RMSE (Eq. (6)) and R index (Eq. 7) were used for comparing the created models and selecting the best model to estimate the BOD_5 value of the A-WWTP in daily and monthly periods.

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(T_i - \text{O}_i\right)^2} \tag{6}$$

$$R = \frac{\sum_{i=1}^{n}\left(T_i - \bar{T}\right)\left(\text{O}_i - \bar{\text{O}}\right)}{\sqrt{\sum_{i=1}^{N}\left(T_i - \bar{T}\right)^2\left(\text{O}_i - \bar{\text{O}}\right)^2}} \tag{7}$$

In these equations, T_i represents the observed values, O_i indicates the model output values, $\bar{T}$ and $\bar{O}$ denote the average of T_i and O_i, and, N reflects the number of data.

3. Results and discussion

3.1. Estimation of daily output BOD_5

A large number of independent parameters affect effluent BOD_5 such as T, volume (V), turbidity, pH, EC, COD, TSS, BOD_5, sludge volume index (SVI) and dissolved oxygen (DO) of input and output wastewater, primal sedimentation basin, as well as the first and second aerated lagoons. Table 1 describes the statistics of the treatment plant data used for developing MLR, FF-ANN, and ANFIS models.

The removal efficiency of BOD_5 in this plant was 89.6%. This result showed the excellent performance of A-WWTP when it was compared with WWTP in Yazd or Gaze. The removal percentage of BOD_5 in Yazd and Gaze was 74.6% and 79%, respectively [14,34].

Nevertheless, the application of all of them causes model complexity. Thus, the most critical parameters affecting the output wastewater BOD_5 were determined. The output wastewater BOD_5 was used as the dependent variable while the other parameters mentioned above were used as the independent variables to generate the model. Thus, four different architectures were considered based on the number of different inputs (given the highest R and R^2 and lowest Std. error) for producing MLR, FF-ANN, and ANFIS (Table 2). The results indicated that the most important parameters affecting the output wastewater BOD_5 were the COD of input and primal sedimentation basin, input BOD_5, input EC, and DO of the primal aerated lagoon. Most researchers have applied one architecture with several data of the input parameter. Türkmenler and Pala [10] used five influent parameters (Q, BOD_5, COD, TP, and TN) for effluent BOD_5 simulation.

In this research, for daily BOD_5 simulation, 500 data were used for training and 176 data for testing.

3.1.1. Modeling by MLR model for BOD_5

The MLR model for different architectures was investigated with SPSS 21 software. Tables 3 and 4 indicate the results and the obtained regression equations, respectively.

According to Table 3, architecture 4 has the highest R and lowest RMSE in comparison with the other architectures at two steps of training and testing. Figs. 1 and 2 show the performance of MLR in predicting BOD_5 compare the values of output BOD_5 estimated from MLR to its observed values at training and testing steps over time. The line crossing the points and axis with equation $y = ax$ was presented in Figs. 1 and 2. As the coefficient in equation $y = ax$ was closer to one, the model could have a better estimate of the observed data. The created model will estimate lower values than their corresponding observed values if the coefficient is less than one. On the contrary, the generated model will determine larger values than their corresponding observed values if the coefficient is greater than one. Therefore, the scattering of points around the bisector line of the coordinate axes is less, and equation $y = ax$ will have a better match if the coefficient is closer to one. Thus, in addition to R and RMSE for recognizing the best model, the scattering of points around the bisector line of the coordinate axes and the slope of the line crossing these points can be used. The observed and estimated values will have a better match if the slope value in the line equation (Fig. 1) is close to one.

The MLR model had an acceptable performance in predicting daily BOD_5 of A–WWTP based on the values of R,

Table 1
Descriptive statistics of the treatment plant data

	Parameter	Minimum	Maximum	Average	Std. deviation
Input	T, °C	14.6	34.9	24.5	3.9
	Turbidity, NTU	24	894	123.9	85.1
	pH	6.8	8.2	7.3	0.17
	EC, µmhos/cm	1,076	8,530	4,643.9	1,106.3
	TSS, mg/L	33	776	137.9	84.1
	COD, mg/L	67	1,958	327.4	126.1
	BOD, mg/L	50	625	187.1	64.3
Output	TSS, mg/L	4	119	29.8	11.7
	BOD, mg/L	3	93	19.3	9.4
Aeration	DO1, mg/L	0.03	27.5	2.4	1.6
	DO2, mg/L	0.03	27.6	2.1	1.6
Primal sedimentation basin	SVI	15.6	1,777	168.8	175.4
	V, m^3	10	990	433.05	288.2
	COD, mg/L	65	411	249.7	65.2
	BOD, mg/L	50	300	147.67	37.45
	TSS, mg/L	7.1	304	65.38	23.79
	Turbidity, NTU	1.7	708	65.24	39.63
	EC, µmhos/cm	27	8,650	4,187.6	813.27

Table 2
Different architectures of input parameters to models (daily)

Architecture	Input	Output	R	Adjust R^2	Standard error of the estimate
1	BOD-I[a], COD-P[b]	BOD-O[c]	0.883	0.779	0.01651
2	BOD-I, COD-P, COD-I	BOD-O	0.898	0.805	0.01922
3	BOD-I, COD-P, COD-I, EC-I	BOD-O	0.904	0.816	0.01916
4	BOD-I, COD-P, COD-I, EC-I, DO-1[d]	BOD-O	0.907	0.822	0.01910

[a]I – input; [b]P – primal sedimentation basin; [c]O – output; [d]l – first aerated lagoon.

Table 3
Results of the MLR model for different architectures of daily BOD_5

Architecture	R at the training level	R at the test level	RMSE at the training level	RMSE at the test level
1	0.902	0.856	0.0216	0.0116
2	0.911	0.856	0.0212	0.0117
3	0.914	0.858	0.0211	0.0117
4	0.916	0.864	0.0210	0.0115

Table 4
MLR equations for daily and monthly BOD_5

Index	Architecture	Equation
Daily BOD_5	One	1.236COD-P + 0.089BOD-I + 0.097
	Two	1.243COD-P + 0.196BOD-I – 0.141COD-I + 0.088
	Three	1.247COD-P + 0.182BOD-I– 0.124COD_i + 0.03EC + 0.1
	Four	1.248COD-P + 0.178BOD-I – 0.121COD-I + 0.034EC – 0.022DO + 0.094
Monthly BOD_5	One	1.141COD-P + 0.374BOD-I + 0.152
	Two	0.979COD-P + 0.36BOD-I + 0.205TUR-P + 0.161
	Three	0.955COD-P + 0.481BOD-I + 0.256TUR-P + 0.15COD-I + 0.16
	Four	0.995COD-P + 0.454BOD-I + 0.257TUR-P – 0.168COD-I + 0.137SVI + 0.201

P – Primal sedimentation basin; I – Input.

RMSE, and functional correlation of observed data and estimated data (Figs. 1 and 2). Belhaj et al. [1] could estimate BOD_5 in Sfax WWTP (Tunisia) using MLR with R^2 = 0.973. They concluded that the MLR model verified BOD_5 of this treatment plant with a rather good approximation.

3.1.2. Modeling results by FF-ANN for daily data

Neural Network Toolbox in MATLAB (2013b) software was used to implement the neural network model. Modeling was conducted based on different defined architectures by selecting different neurons. The number of neurons for the input layer was assumed to be equal to the number of network inputs; the number of neurons for the output layer was considered as one; and the number of neurons for the middle layer was examined from 6 to 20 using a trial-and-error approach. The number of neurons was selected based on preliminary information, as reported in Table 5 for different architectures. The results of the FF-ANN model developed for training and verifying the output BOD_5 for different architectures are presented in Figs. 3 and 4. The comparison of Tables 3 and 5 showed that the FF-ANN model was better than the MLR model. This result concurs with Hamada et al. [14] findings.

3.1.3. Modeling results by ANFIS for daily data

ANFIS was modeled by the ANFIS toolbox in MATLAB software. Different structures in each different architecture and iterations were studied.

The results of the ANFIS developed for verifying the output BOD_5 for different architectures are outlined in Table 6. Figs. 5 and 6 indicate the results of ANFIS for training and testing the output BOD_5 for different architectures. Based on the obtained results, the ANFIS had more relevant results than the ANN. In research, the comparison of ANN, ANFIS, and MLR models for estimating the adsorption efficiency of biochar for the removal of Cu(II) ions showed that the ANFIS model with Gaussian membership function and fuzzy set combination of [4 5 2 3] was the best method,

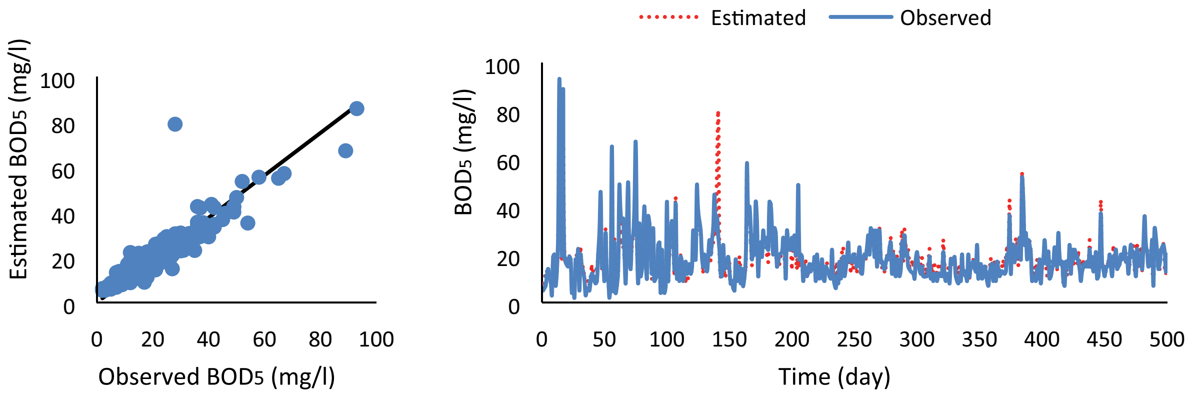

Fig. 1. Comparison of the values of the best MLR structure and the observed values of daily output BOD_5 at the training level.

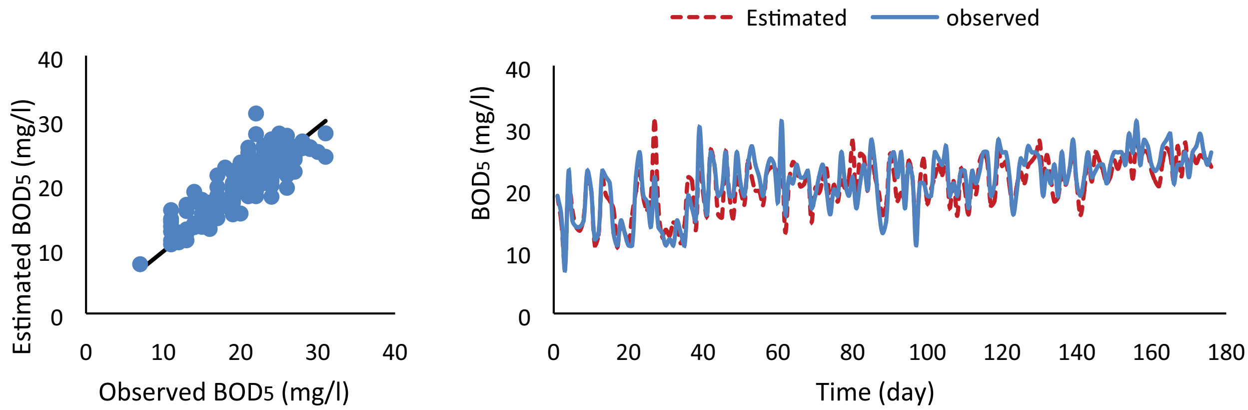

Fig. 2. Comparison of the values of the best MLR structure and the observed values of daily output BOD_5 at the test level.

Table 5
FF-ANN model results for different architectures of the daily period

Architecture	*R* at the training level	*R* at the test level	RMSE at the training level	RMSE at the test level
1	0.941	0.854	0.0149	0.0123
2	0.945	0.860	0.0143	0.0118
3	0.950	0.878	0.0131	0.0112
4	0.956	0.897	0.0126	0.0110

with an accuracy of 90.24% and 87.06% for the training and test datasets, respectively [35].

3.2. Estimation of monthly output BOD_5

Initially, useful parameters were determined by the backward method, and four different architectures were evaluated based on different inputs to make the nonlinear regression, FF-ANN, and ANFIS to estimate the wastewater monthly output BOD_5 (Table 7). The most critical parameters affecting the wastewater monthly output BOD_5 included the input COD, primal sedimentation basin, input BOD_5, turbidity of a primal sedimentation basin, and SVI. In this research, for monthly BOD_5 simulation, 72 data were used for training and 24 data for testing.

3.2.1. Modeling via MLR for monthly BOD_5

The MLR model was implemented by SPSS 21 software for the four different architectures mentioned above (Table 7). Table 4 reports the obtained equations.

As Table 8 indicates, architecture 4 has a higher *R* and a lower RMSE in comparison with the other architectures at the training and test levels. The results of architecture 4 as the best-obtained model for the training and test are presented in Figs. 7 and 8. Although the *R* values for both

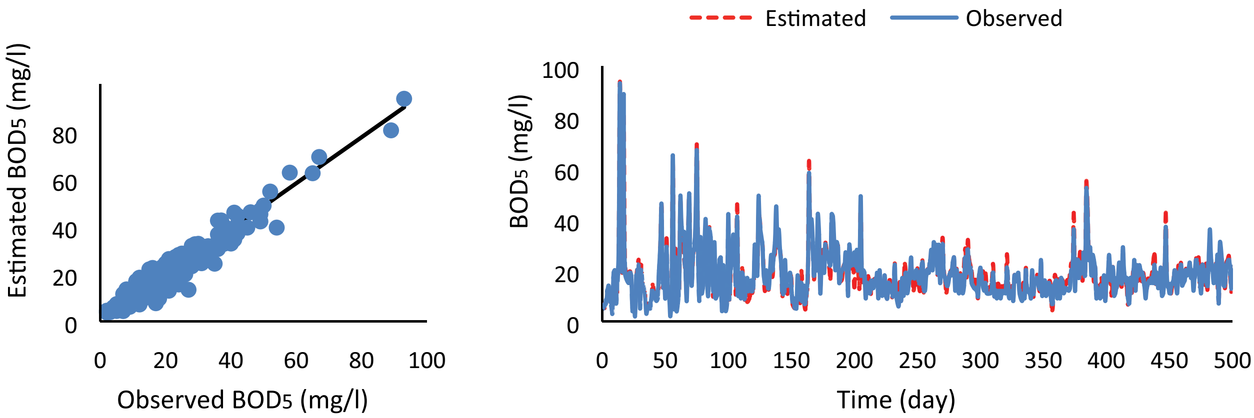

Fig. 3. Comparison of the values of the FF-ANN model and the observed values of daily output BOD_5 at the training level.

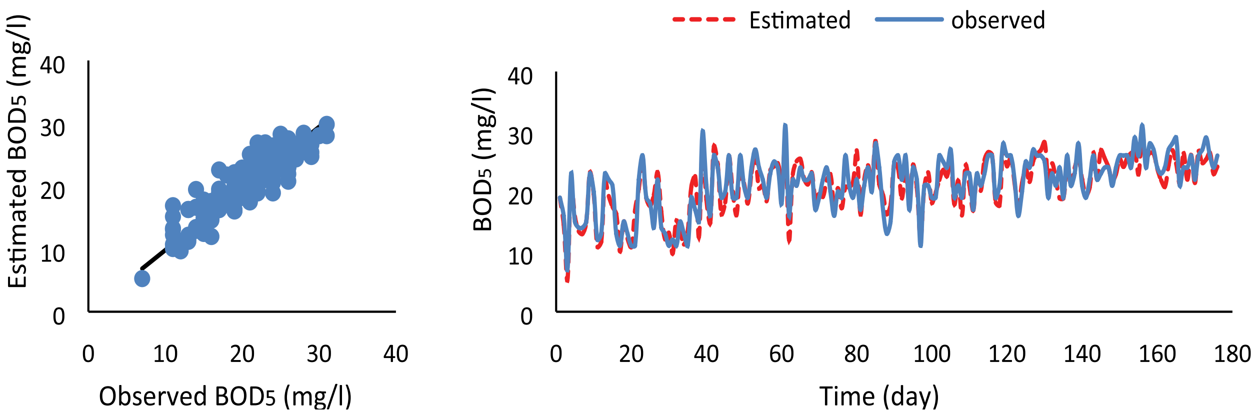

Fig. 4. Comparison of the values of the FF-ANN model and the observed values of daily output BOD_5 at the test level.

Table 6
Results of ANFIS for different architectures of the daily period

Architecture	*R* at the training level	*R* at the test level	RMSE at the training level	RMSE at the test level
1	0.948	0.868	0.0134	0.0112
2	0.962	0.881	0.0113	0.0098
3	0.971	0.899	0.0105	0.0092
4	0.979	0.930	0.0099	0.0083

levels of training and test are lower than the linear regression values for daily data (training = 0.916 and test = 0.864), the closeness of these coefficients (training = 0.810 and test = 0.794) indicates that the linear regression model for monthly data could estimate the data of test level with very good accuracy.

3.2.2. Modeling results for an FF-ANN for monthly data

Modeling was conducted based on different defined architectures (Table 7) and by selecting different numbers of neurons. Meanwhile, the number of input layer neurons was assumed to equal to the number of network inputs, the number of output layer networks was evaluated as one, and the number of middle layers was estimated as trial and error from 6 to 20 (Table 9). Figs. 9 and 10 illustrate the results of FF-ANN for training and testing the output BOD_5 for different architectures.

3.2.3. Modeling results via ANFIS for monthly data

ANFIS toolbox in MATLAB software was used for modeling via the ANFIS. Different structures in each different architecture and iterations were studied. Table 10 and Figs. 11 and 12 present the results of ANFIS for training and testing the output BOD_5 for different architectures.

4. Conclusion

The study investigated a simple MLR model and two artificial intelligence models (FF-ANN and ANFIS) with 4 architectures (requiring 2–5 inputs). The results indicated

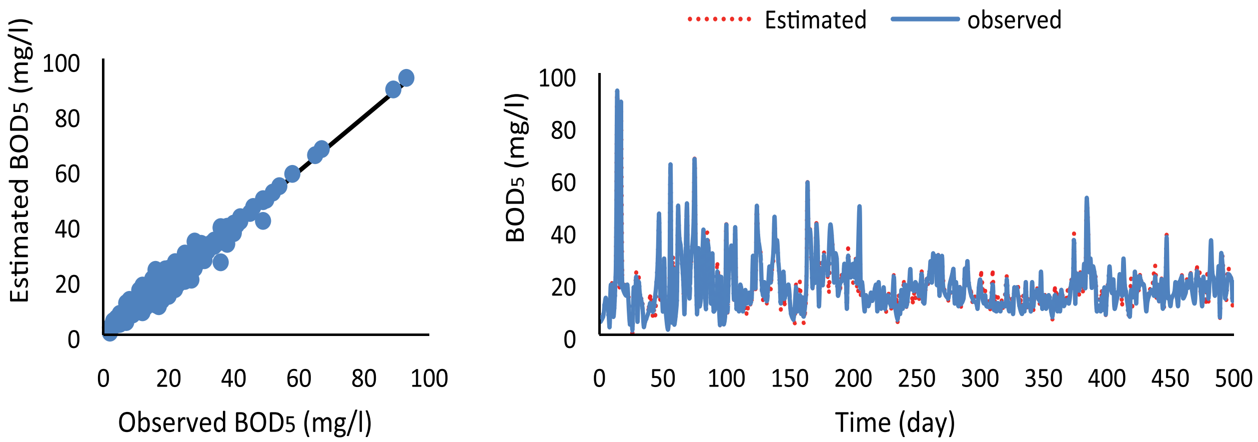

Fig. 5. Comparison of the values of ANFIS and the observed values of daily output BOD_5 at the training level.

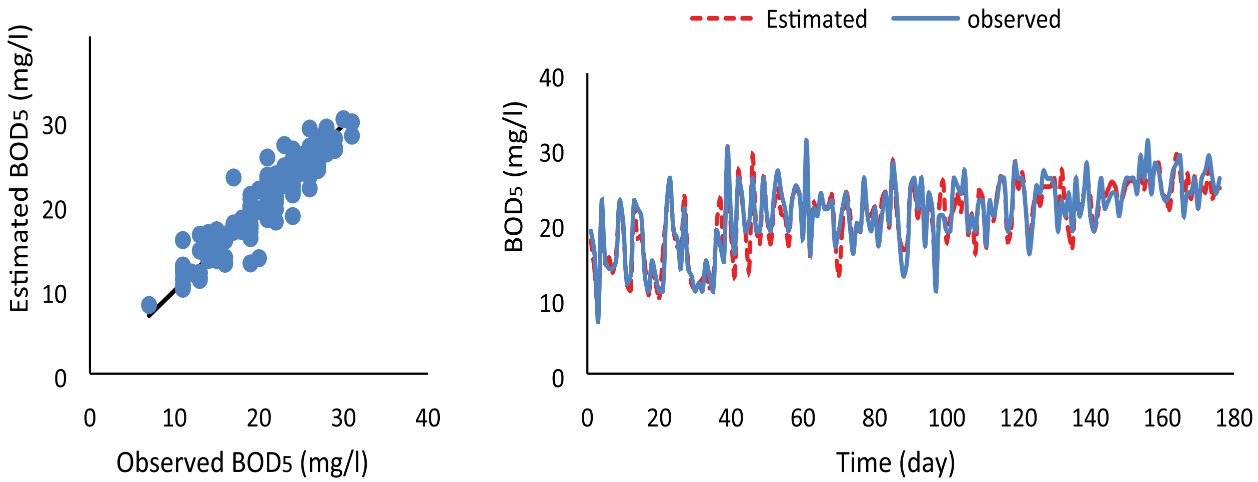

Fig. 6. Comparison of ANFIS values and the observed values of daily output BOD_5 at the test level.

Table 7
Different architectures of input parameters to monthly BOD_5 models

Architecture	Input	Output
1	BOD_5-I, COD-P	BOD_5-O
2	BOD_5-I, COD-P, TUR-P	BOD_5-O
3	BOD_5-I, COD-P, TUR-P, COD-I	BOD_5-O
4	BOD_5-I, COD-P, TUR-P, COD-I, SVI	BOD_5-O

O – Output; I – Input; P – Primal sedimentation basin.

Table 8
Results of the MLR model for different architectures of monthly BOD_5

Architecture	*R* at the training level	*R* at the test level	RMSE at the raining level	RMSE at the test level
1	0.790	0.782	0.0658	0.0427
2	0.795	0.790	0.0643	0.0421
3	0.802	0.791	0.0634	0.0411
4	0.810	0.794	0.0618	0.0396

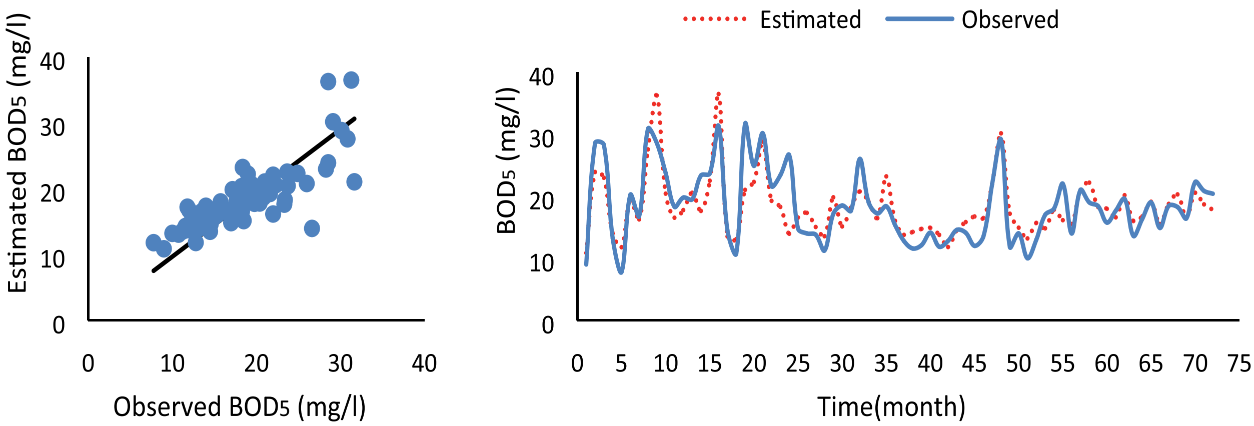

Fig. 7. Comparison of the values of the best linear regression structure and the observed values of monthly output BOD_5 at the training level.

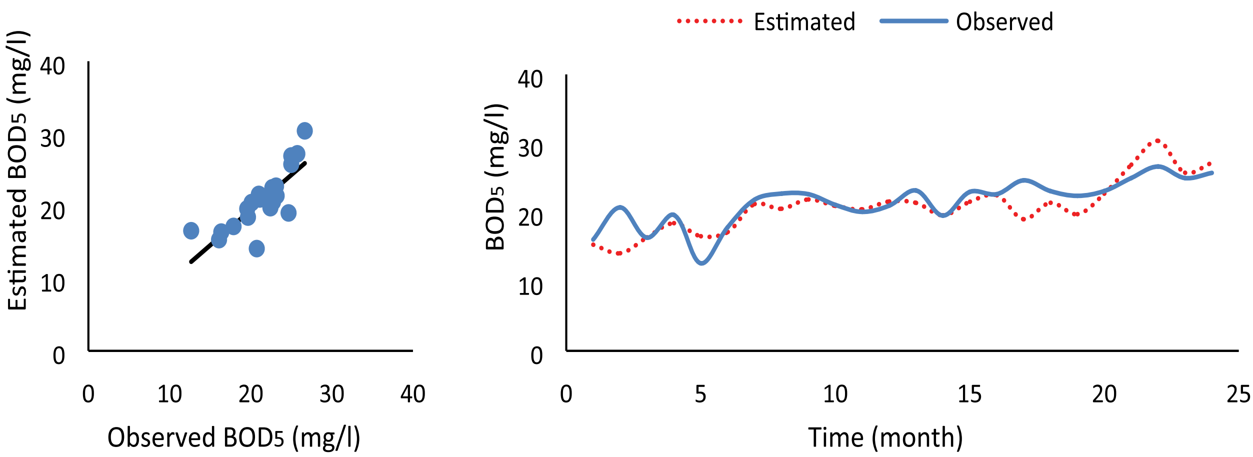

Fig. 8. Comparison of the values of the best linear regression structure and the observed values of monthly output BOD_5 at the test level.

Table 9
Results of FF-ANN for different architectures of monthly BOD_5

Architecture	*R* at the training level	*R* at the test level	RMSE at the training level	RMSE at the test level
1	0.858	0.826	0.0591	0.0365
2	0.865	0.834	0.0540	0.0314
3	0.876	0.843	0.0506	0.0308
4	0.905	0.854	0.0463	0.0290

Table 10
Monthly results of ANFIS for different architectures

Architecture	*R* at the training level	*R* at the test level	RMSE at the training level	RMSE at the test level
1	0.872	0.835	0.0499	0.0377
2	0.898	0.840	0.0440	0.0324
3	0.930	0.862	0.0317	0.0269
4	0.956	0.927	0.0210	0.0174

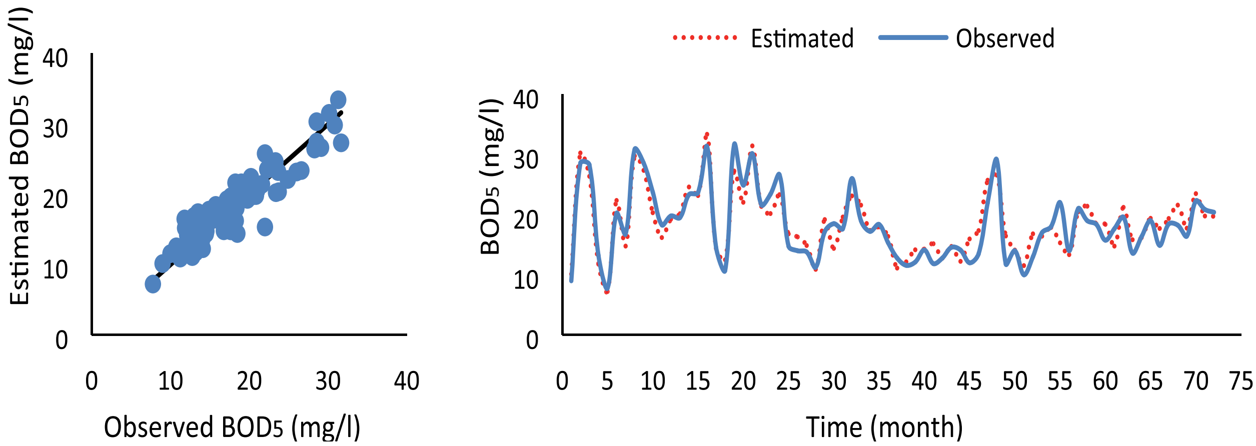

Fig. 9. Comparison of the values of architecture 4 in FF-ANN and the observed values of monthly BOD_5 (training level).

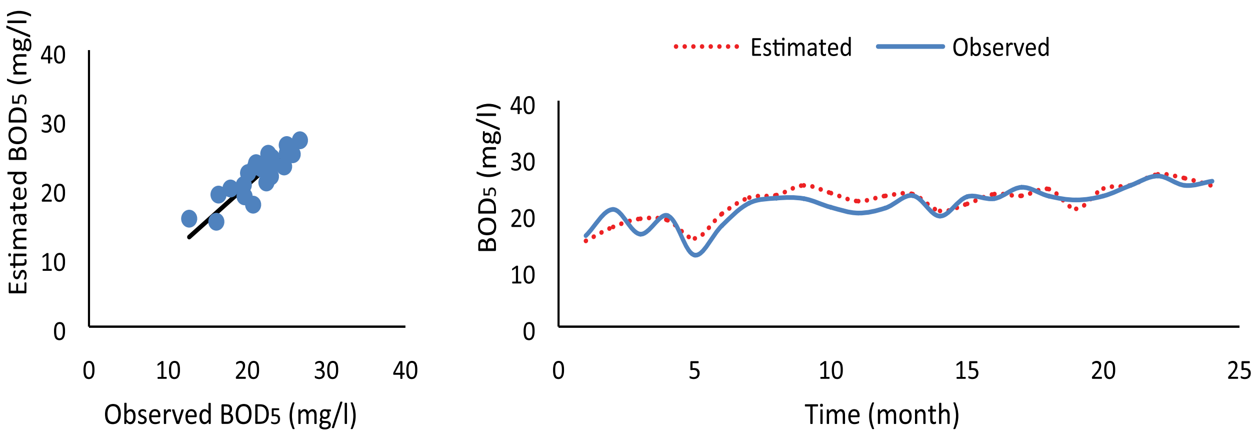

Fig. 10. Comparison of the values of architecture 4 in ANN and the observed values of monthly BOD_5 (test level).

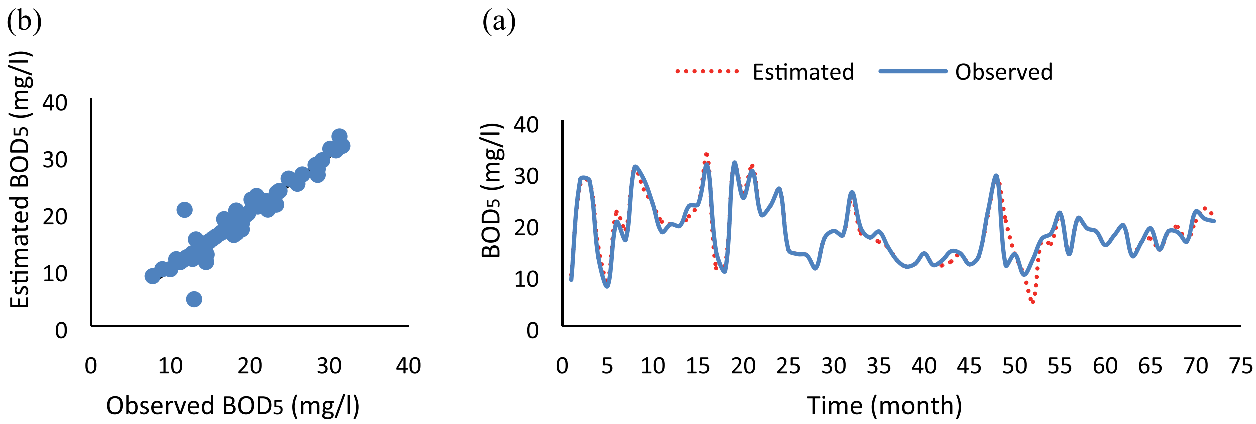

Fig. 11. Comparison of the values of architecture 4 in ANFIS and the observed values of monthly BOD_5 (training level).

that although architecture 4 had better results in both periods and three models, other architectures had acceptable solutions close to the observed values. Different architectures can be used based on conditions and available wastewater qualitative parameters to estimate the output BOD_5 in treatment plants. Furthermore, based on the obtained results, the ANFIS had a better performance for daily and monthly periods at training and test levels as compared to the regression and FF-ANN. Note that the performance of the three models was more appropriate in estimating the daily output BOD_5 of this treatment plant as compared to the monthly type. According to the results

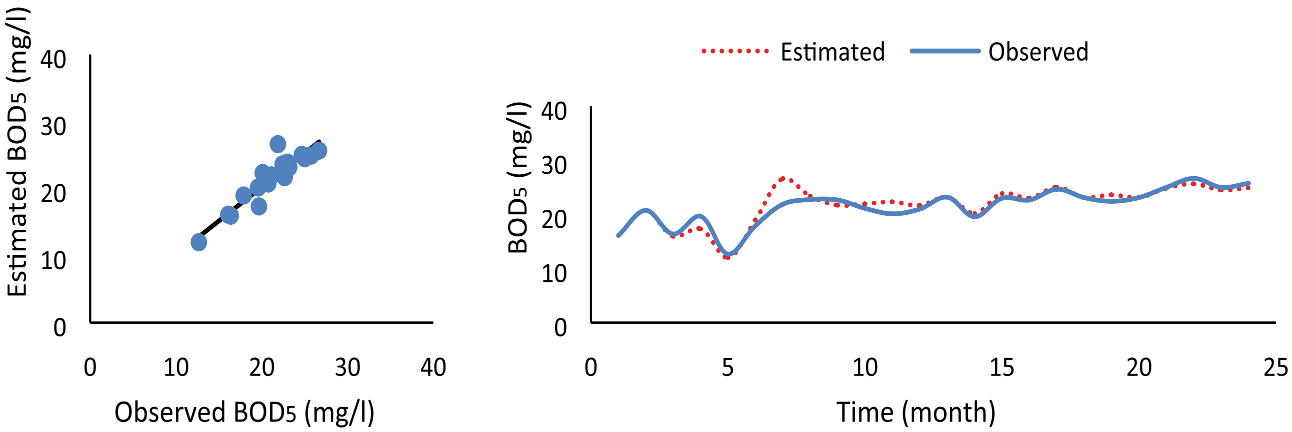

Fig. 12. Comparison of the values of architecture 4 in ANFIS and the observed values of monthly BOD_5 (test level).

of this study, other models of artificial intelligence such as genetic algorithms as well as hybrid models are proposed to evaluate the quality of effluent and technical and economic management of the treatment plant.

Acknowledgments

The authors would like to appreciate the Ahwaz Water and Waste Water Company, which cooperated with this project. Also, we are grateful to the Research Council of the Shahid Chamran University of Ahvaz for financial support (GN.95.3.02.31400).

References

[1] D. Belhaj, I. Jaabiri, N. Turki, C. Azri, M. Kallel, H. Ayadi, Descriptive and multivariable analysis of the water parameters quality of Sfax sewage treatment plant after rehabilitation, IOSR-J. Comput. Eng., 16 (2014) 81–91.
[2] F. Rafat Motevali, Sh. Danesh, H. Rajabi Mashhadi, Evaluation and Management of the Performance of Semi-Mechanical Treatment Plants by Predicting the Quality of their Output Wastewater by Genetic Algorithm or Optimized Neural Network Model, The 5th National Conference on Water, Wastewater and Waste, Tehran, 2014.
[3] L. Zhao, T.J. Dai, Z. Qiao, P.Z. Sun, J.Y. Hao, Y.K. Yang, Application of artificial intelligence to wastewater treatment: a bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse, Process Saf. Environ. Prot., 133 (2020) 169–182.
[4] M. Falah Nezhad, N. Mehrdadi, A. Torabian, S. Behboudian, Artificial neural network modeling of the effluent quality index for municipal wastewater treatment plants using quality variables: south of Tehran wastewater treatment plant, J. Water Supply Res. Technol., 65 (2016) 18–27.
[5] M.M. Hamed, M.G. Khalafallah, E.A. Hassanien, Prediction of wastewater treatment plant performance using artificial neural networks, Environ. Modell. Software, 19 (2004) 919–928.
[6] S. Heddam, H. Lamda, S. Filali, Predicting effluent biochemical oxygen demand in a wastewater treatment plant using generalized regression neural network based approach: a comparative study, Environ. Process., 3 (2016) 153–165.
[7] K.P. Oliveira-Esquerre, D.E. Seborg, R.E. Bruns, M. Mori, Application of steady-state and dynamic modeling for the prediction of the BOD of an aerated lagoon at a pulp and paper mill: Part I. Linear approaches, Chem. Eng. J., 104 (2004) 73–81.
[8] F.S. Mjalli, S. Al-Asheh, H.E. Alfadala, Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance, J. Environ. Manage., 83 (2007) 329–338.
[9] A. Mehdi Pour, M. Shokuhiyan, Effect of Input Sewage Parameters on the Estimation of the Output TSS Verification by Using Sensitivity Analysis Based on Artificial Neural Networks, 7th Civil Engineering Congress, University of Sistan and Baluchistan, Zahedan, 2012.
[10] H. Türkmenler, M. Pala, Performance assessment of advanced biological wastewater treatment plants using artificial neural networks, Int. J. Eng. Technol., 3 (2017) 151–156.
[11] M. Vyas, B. Modhera, V. Vyas, A.K. Sharma, Performance forecasting of common effluent treatment plant parameters by artificial neural network, ARPN J. Eng. Appl. Sci., 6 (2011) 38–42.
[12] H. Zare Abyaneh, M. Bayat Varkeshi, J. Bayat Varkeshi, Application of artificial neural network in evaluating Ekbatan wastewater treatment plant, Ecology, 3 (2012) 85–98.
[13] H. Zare Abyaneh, Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters, J. Environ. Health Sci. Eng., 12 (2014) 40.
[14] M. Hamada, H. Adel Zaqoot, A. Abu Jreiban, Application of artificial neural networks for the prediction of Gaza wastewater treatment plant performance-Gaza strip, J. Appl. Res. Water Wastewater, 5 (2018) 399–406.
[15] O.T. Baki, E. Aras, Estimation of BOD in wastewater treatment plant by using different ANN algorithms, Membr. Water Treat., 9 (2018) 455–462.
[16] F. Fang, B.-J. Ni, W.-M. Xie, G.-P. Sheng, S.-G. Liu, Z.-H. Tong, H.-Q. Yu, An integrated dynamic model for simulating a full-scale municipal wastewater treatment plant under fluctuating conditions, Chem. Eng. J., 160 (2010) 522–529.
[17] S. Shokri, A. Asqari Moghadam, A. Nadiri, Evaluation of the Efficiency of Tabriz Wastewater Treatment Plant by Using Different Fuzzy Systems, National Conference on Environmental Research of Iran, Shahid Mofattah Faculty, Hamedan, 2013.
[18] A.A. Nadiri, S. Shokri, F.T.C. Tsai, A.A. Moghaddam, Prediction of effluent quality parameters of a wastewater treatment plant using a supervised committee fuzzy logic model, J. Cleaner Prod., 180 (2018) 539–549.
[19] G. Civelekoglu, N.O. Yigit, E. Diamadopoulos, M. Kitis, Modelling of COD removal in a biological wastewater treatment plant using adaptive neuro-fuzzy inference system and artificial neural network, Water Sci. Technol., 60 (2009) 1475–1487.
[20] M. Vajedi, S. Shah Husseini, Active sludge process modeling by adaptive neuro-fuzzy inference system, Water Wastewater, 4 (2014) 108–111.
[21] V. Nourani, G. Elkiran, S.I. Abba, Wastewater treatment plant performance analysis using artificial intelligence–an ensemble approach, Water Sci. Technol., 78 (2018) 2064–2076.

[22] S. Akilandeswari, B. Kavitha, Determination of biochemical oxygen demand by adaptive neuro fuzzy inference system, Adv. Appl. Sci. Res., 4 (2013) 101–104.
[23] M.S. Zaghloul, R.A. Hamza, O.T. Iorhemen, J.H. Tay, Comparison of adaptive neuro-fuzzy inference systems (ANFIS) and support vector regression (SVR) for data-driven modelling of aerobic granular sludge reactors, J. Environ. Chem. Eng., 8 (2020) 103742.
[24] M.A. Razi, K. Athappilly, A comparative predictive analysis of neural networks (NNs), nonlinear regression and classification and regression tree (CART) models, Expert Syst. Appl., 29 (2005) 65–74.
[25] E. Dogan, B. Sengorur, R. Koklu, Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique, J. Environ. Manage., 90 (2009) 1229–1235.
[26] J.S.R. Jang, C.T. Sun, E. Mizutani, Neuro-fuzzy and soft computing-a computational approach to learning and machine intelligence, IEEE Trans. Autom. Control, 42 (1997) 1482–1484.
[27] H. Moral, A. Aksoy, C.F. Gokcay, Modeling of the activated sludge process by using artificial neural networks with automated architecture screening, Comput. Chem. Eng., 32 (2008) 2471–2478.
[28] A. Solgi, F. Radmanesh, H. Zarei, V. Nourani, Hybrid models performance assessment to predict flow of Gamasyab River, Int. J. Adv. Biol. Biomed. Res., 2 (2014) 1837–1846.
[29] S. Riad, J. Mania, L. Bouchaou, Y. Najjar, Rainfall-runoff model usingan artificial neural network approach, Math. Comput. Modell., 40 (2004) 839–846.
[30] P. Fathi, Y. Muhammadi, M. Homayi, Intelligent modeling of monthly flow time series into Vahdat dam in Sanandaj, J. Water Soil, 23 (2009) 209–220.
[31] V. Nourani, M.A. Komasi, A geomorphology-based ANFIS model for multi-station modeling of rainfall–runoff process, J. Hydrol., 490 (2013) 41–55.
[32] S. Asadi, J. Shahrabi, P. Abbaszadeh, S.A. Tabanmehr, A new hybrid artificial neural networks for rainfall–runoff process modeling, Neurocomputing, 121 (2013) 470–480.
[33] N. Haghdadi, A. Zarei-Hanzaki, A.R. Khalesian, H.R. Abedi, Artificial neural network modeling to predict the hot deformation behavior of an A356 aluminum alloy, Mater. Des., 49 (2013) 386–391.
[34] M. Farzadkia, M.H. Ehrampoush, S. Sadeghi, M. Kermani, M.T. Ghaneian, V. Ghelmani, E.A. Mehriziu, Performance evaluation of wastewater stabilization ponds in Yazd-Iran, Environ. Health Eng. Manage. J., 1 (2014) 7–12.
[35] Y.J. Wong, S.K. Wong, C.H. Arumugasamy, A. Chung, V. Selvarajoo Sethu, Comparative study of artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and multiple linear regression (MLR) for modeling of Cu(II) adsorption from aqueous solution using biochar derived from rambutan (*Nephelium lappaceum*) pee, Environ. Monit. Assess., 192 (2020) 439.