Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2021.26950

218 (2021) 164–176
April

The influence of aeration scheme and aeration rate on the permeate flux for wastewater treatment using membrane bioreactors: experimental and artificial neural network modeling

Jasir Jawad[a], Alaa H. Hawari[b,*], Syed Javaid Zaidi[a], Abdulkarim Almukdad[b]

[a]*Centre for Advanced Materials, Qatar University, P.O. Box: 2713, Doha, Qatar, Tel. +(974) 5044-1796; email: jasir.jawad@qu.edu.qa (J. Jawad), Tel. +(974) 5013-9969; email: szaidi@qu.edu.qa (S.J. Zaidi)*
[b]*Department of Civil and Architectural Engineering, Qatar University, P.O. Box: 2713, Doha, Qatar, Tel. +(974) 3393-1555; emails: a.hawari@qu.edu.qa (A.H. Hawari), aa1304575@student.qu.edu.qa (A. Almukdad)*

Received 24 August 2020; Accepted 11 December 2020

ABSTRACT

In this paper, the effect of time, aeration scheme, aeration rate, and mixed liquor suspended solid (MLSS) concentration on the permeate flux in membrane bioreactors have been studied. Aeration rates of 0.5, 1.0, and 1.5 LPM were tested with MLSS concentrations of 5, 10, and 15 g/L. Furthermore, a continuous and pulsed aeration scheme (5 min on and 5 min off) was tested. The experimental data were used to develop an artificial neural network (ANN) model, which showed excellent agreement (R^2 = 0.9963) with the data. The average normalized flux decreased as the MLSS concentration increased from 5 to 15 g/L and increased as the aeration rate increased from 0.5 to 1.5 LPM. No clear correlation was found between the aeration schemes and the average normalized flux. ANN weights analysis revealed the order of importance was time > aeration rate > MLSS concentration > aeration scheme.

Keywords: Neural networks; Membrane bioreactors; Aeration scheme; Membrane separation; Wastewater treatment

1. Introduction

Due to the increased water scarcity, it is expected that by 2025 half of the world's population will be affected by the lack of clean water [1]. Therefore, in addition to conventional desalination techniques for seawater and brackish water, recovery, and reuse of wastewater have also become important. Membrane bioreactors (MBRs) are considered an economical and a compact technique which have attracted great attention from industrial and municipal wastewater treatment plants in the past few decades [2]. The MBR system combines a biological treatment process and a membrane filtration unit. The MBR technology offers many advantages over the conventional activated sludge process, such as providing low amounts of waste sludge and high effluent quality [3]. Two of the major problems encountered in this process are energy consumption and membrane fouling [4,5]. Over 90% of the energy requirements in a submerged MBR process is due to aeration [6]. The concentration of dissolved oxygen affects the efficiency of MBR. At low dissolved oxygen concentration, the activity of the microorganisms is low, which leads to the low efficiency of the MBR process. More than a required aeration rate must be provided to maintain good effluent quality, which also results in high energy consumption [7]. Further research showed that intermittent aeration prevented fouling and preserved bioflocs while improving the filtration process [8,9]. The pulsed aeration approach showed enhanced permeate flux and also prevents floc damage that lowers the membrane fouling [10,11]. The increase in the MLSS concentration results in lower

* Corresponding author.

permeate flux due to the increased fouling in the membranes [12,13]. High MLSS concentration results in higher cake resistance and increased the dynamic viscosity of the solution, which leads to higher internal friction. Although most studies showed that MLSS concentration had an adverse effect on the permeate flux, however, Le-Clech et al. [14] showed that MLSS had no effect on fouling up to a threshold concentration value. Several studies have been performed to simulate and model the MBR process in an attempt to optimize the performance of the process. In the last few decades, artificial intelligence and machine learning have shown a high capability for problem-solving in environmental engineering problems. These techniques have been used to model and optimize pollutant removal in water treatment processes [15]. One of the highly used machine learning techniques is the artificial neural network (ANN). ANN is a data-driven technique and a black-box model that is not governed by physical laws [16]. The aim of ANN is to develop a relationship between a set of input variables and one or more output variables. ANN holds few advantages over the conventional theoretical models such as the capability of modeling highly non-linear systems, working with noisy or incomplete data, less computational time, and the ability to train and update the existing model with new data when available [17].

Several studies have been conducted in the literature to model the MBR process using different types of ANN. Chen et al. [18] presented the applicability of two-phase anaerobic digestion (TPAD) MBR for the treatment of pharmaceutical wastewater. ANN was used to investigate the removal of chemical oxygen demand (COD) by evaluating the performance of three multi-layer perceptron (MLP) type neural networks. Liu et al. [19] used principal component analysis (PCA) to identify dissolved oxygen (DO), MLSS, pH, temperature, and suction power as the most impactful variables on the permeate flux in MBRs. The permeate flux was predicted using the ANN model with particle swarm optimization (PSO) to study membrane fouling in MBR. In another research, a conventional MLP type model was compared to the PSO-based ANN model that showed the later to be more superior in the prediction of fouling resistance in MBRs [2]. Similarly, Li et al. [20] used a genetic algorithm (GA) process for the optimization of the weights in an ANN model to study membrane fouling. Pendashteh et al. [21] modeled the MBR treatment of hypersaline oily wastewater using ANN to predict several effluent characteristics. The relative importance showed that the organic loading rate (OLR) had the most effect, whereas the total dissolved solids (TDS) had the least effect on the model outputs. Mirbagheri et al. [22] developed a radial basis function (RBF) neural network to predict effluent concentration for the treatment of a combined municipal and industrial wastewater. In another research, the authors also concluded that MLP had better prediction and generalization ability than RBF neural networks when trained using GA to predict transmembrane pressure (TMP) and permeate flux in MBRs [3]. The Levenberg–Marquardt method showed the least relative error when used to model the treatment of petrochemical wastewater to predict TMP and COD using ANN in MBRs [23]. The sensitivity analysis revealed that the hydraulic retention time (HRT) and MLSS had the highest impact on the outputs. Schmitt et al. [5] studied membrane fouling in an anoxic–aerobic MBR by predicting TMP based on several input variables. The study showed that typical operating conditions such as MLSS, DO, pH, and COD could not be linked with the TMP, but total nitrogen (TN) and total phosphorus (TP) concentrations had higher impact on the TMP.

Geissler et al. [24] presented an Elman neural network (ENN) model to predict the permeate flux in an MBR process with less than 2.7% average deviation. A semi-empirical model described the filtration resistance based on the operating parameters to estimate the permeability decline. Furthermore, the sensitivity analysis of the model showed the dominant effect of backwash pressure on the stability of the flux. A similar neural network was developed by Li and Wang [25] to study membrane fouling for the treatment of sewage wastewater. The degree of membrane fouling was represented by the permeate flux. Moreover, the ENN model was compared to an MLP type model that showed, due to its recursive nature, the ENN was more accurate in the predictions. Chen et al. [26] and Zhao et al. [27] predicted interfacial energies to study membrane fouling in MBR using RBF neural networks. The computational time for the Extended Derjaguin–Landau–Verwey–Overbeek (XDLVO) approach for the calculation of interfacial energy was 3 times higher than the RBF model. Cai et al. [28–31] developed multiple wavelet neural networks to study the effect of hydraulic retention, salinity, and pH on the MBR treatment of marine sewage. The different studies showed that the wavelet network was more accurate than the corresponding MLP network. Table 1 summarizes the details of the neural network developed in the literature for the MBR process.

As mentioned earlier aeration accounts for almost 90% of the energy consumption in MBRs. To the best of our knowledge, the modeling of the aeration rate and the aeration scheme in MBRs have not been modeled or studied using neural networks. The neural network weights can be used to study the impact of the aeration scheme, such as continuous or pulsed, on the permeate flux. The aeration method has a direct impact on the amount of energy consumed. Therefore, the selection of an optimum aeration scheme and aeration rate at a specific MLSS concentration will lead to a higher permeate flux as well as lower energy consumption. In this work, ANNs were used to model MBRs with the input variables being: time, aeration scheme, aeration rate, and MLSS and the output variable is the permeate flux. The experiments were performed using a lab-scale MBR setup at aeration rates of 0.5, 1, and 1.5 LPM, whereas the MLSS concentration was 5, 10, and 15 g/L. Both continuous and pulsed (5 min on 5 min off) schemes were used for each set of experiments. An optimal MLP-ANN model was selected out of 40 trained networks built based on the experimental data. Finally, the sensitivity analysis of the model was conducted by analyzing the weights of the optimal network to show the relative importance of each variable.

2. Materials and methods

2.1. Experimental setup

The schematic diagram of the experimental setup is presented in Fig. 1. The bench-scale MBR system consisted

Table 1
Summary of the neural network models for MBR in literature

Process	Input	Output	Network topology	Activation	Training algorithm	Performance	References
Pilot plant MBR with submerged capillary modules	TMP, rate of TMP change, TMP backwash, filtration time, backwash time, solid retention time, total suspended solids (TSS), temperature, and oxygen decay rate	Permeate flux	9-55-1	–	–	Average deviation <2.7%	[24]
Pilot treatment of pharmaceutical wastewater using TPAD-MBR	Oxidation–reduction potential, hydraulic retention time, pH, COD loading rate, and DO concentration	Effluent COD of MBR	5-4-3 5-4-2 5-4-1	tan-sigmoid	Gradient descent with momentum and adaptive learning rate	MSE = 0.008546	[18]
Treatment of submarine wastewater using a combined membrane process	Hydraulic retention time, pH, COD loading rate, DO concentration, temperature, pressure, turbidity, and TSS	COD, ammonia nitrogen, turbidity, anionic surfactants	–	log-sigmoid	–	Average error: COD = 5.14% NH_4^+–N = 6.20% RMSE: Turbidity =2.76% LAS = 1.41%	[32]
Ultrafiltration of printing and dyeing wastewater	MLSS, DO, pH, temperature, and suction power	Permeate flux	5-12-1	–	Gradient descent method/ particle swarm optimization	Average error = 3.59%	[19]
Hypersaline oily wastewater treatment using MBR	TDS, OLR, reaction time, and time	COD, TOC, MLSS, oil in sludge	4-9-4	tan-sigmoid	Batch back propagation	R^2 = 0.97339	[21]
MBR sewage treatment plant	MLSS, total resistance, and operating pressure	Permeate flux	–	log-sigmoid	Levenberg–Marquardt Genetic algorithm	Average error = 0.0740 Average error = 0.0331	[20]
MBR pilot plant with combined aeration and filtration	Time, TSS, COD, sludge retention time (SRT), and MLSS	TMP, permeate flux	5-10-2	tan-sigmoid Gaussian	Levenberg–Marquardt	R^2 = 0.98, 0.93	[3]
Treatment of combined municipal and industrial wastewater	Influent concentration, HRT, MLSS, TDS, and pH	Effluent concentration	5-5-1	Gaussian function	–	R^2 = 0.99882	[22]

(Continued)

Table 1 Continued

Process	Input	Output	Network topology	Activation	Training algorithm	Performance	References
Treatment of petrochemical wastewater	MLSS, HRT, and time	TMP, COD concentration	3-17-2	tan-sigmoid	Levenberg–Marquardt	R^2 = 0.9999	[23]
Sewage treatment	MLSS, operating pressure, total resistance, pH, COD, and temperature	Permeate flux	6-10-1	log-sigmoid	Gradient descent method	Error = 0.0509	[25]
Anoxic–aerobic MBR treatment of domestic wastewater	Permeate flux, COD MBR, MLSS MBR, and OLR	TMP	4-5-1	log-sigmoid	Levenberg–Marquardt	R^2 = 0.854	[5]
MBR pilot plant	Contact angle of three probe liquids, zeta potential, and separation distance	Interfacial energy	5-8-1 5-160-1 5-110-1	Gaussian function	Gram–Schmidt and adaptive gradient descent	R^2 = 1	[26]
Different MBR studies in literature	Temperature, permeate flux, MLSS, and TMP	Fouling resistance	–	–	Levenberg–Marquardt	R^2 = 0.9682	[2]
Air-lift multilevel circulation MBR treatment of marine domestic sewage	COD, HRT, MLSS, and pH	Effluent COD	4-10-1	tan-sigmoid	Gradient descent with momentum and adaptive learning rate	Average error = 3.67%	[31]
Ship sewage treatment using aerobic–anaerobic micro-sludge MBR system	COD or TN loading rate, pH, F/M, and salinity	Effluent COD, TN	4-3-1 4-3-1	sigmoid wavelet function	–	Average error (MLP) = 7%, 9.72% Average error (WNN) = 2.1%, 3.8%	[30]
Treatment of ship sewage using an air-lift multilevel circulation membrane reactor	Influent COD, influent NH_4^+–N, salinity	Effluent COD Effluent NH_4^+–N	3-2-1	wavelet	–	Average error = 4.04% Average error = 7.90%	[29]
Treatment of ship sewage using an air-lift multilevel circulation membrane reactor	TN and COD loading rate, HRT, pH, and F/M	Effluent COD TN	5-4-1 5-4-1	wavelet	–	Average error = 4.1% Average error = 7.1%	[28]

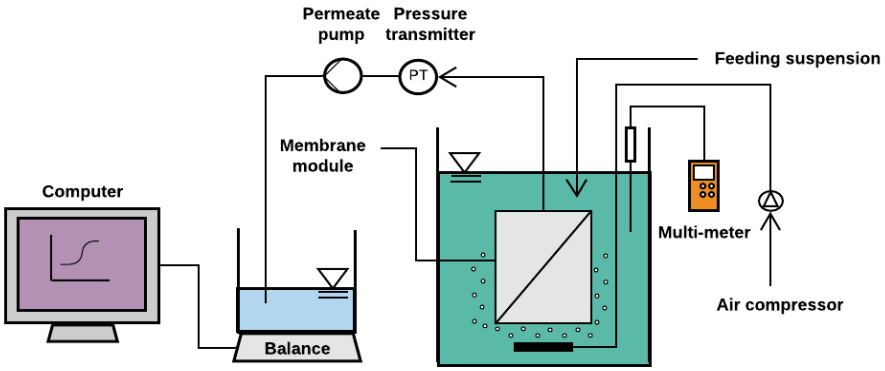

Fig. 1. Schematic diagram of the MBR system.

of a 5 L bioreactor filled with bioparticle suspension, in which a chlorinated polyethylene (CPE) membrane module was submerged. The membrane module has an effective area of 162 cm^2 (81 cm^2 on both sides) with a nominal pore size of 0.4 μm. The membrane module consists of two membranes welded onto two acrylonitrile butadiene styrene (ABS) frames supported by two 6 mm plastic spacers. The membranes were welded on the ABS frames using an ultrasonic welding machine (HiQ LOGIC 1200, Herrmann, Karlsbad, Germany). During the experiments, the water level in the reactor was monitored and maintained fixed to ensure that the concentration of the activated sludge was not changed and the membrane was kept submerged in the solution. The oxygen in the bioreactor was supplied using an air diffuser placed at the bottom of the reactor, which was connected to an air compressor. In this study, two aeration schemes, that is, continuous and pulsed at different flow rates, have been used. The pulsed aeration was supplied at an intermittent time (5 min on, 5 min off). The permeate was collected using a suction pump at 0.1 bar (Gear Pump Drive, Model 75211-15, Cole Parmar Instrument) controlled by a pressure transducer (CPT 2500, Wika) to maintain a constant pressure during all experiments.

The biomass suspension was obtained from a local wastewater treatment plant (Doha South Wastewater Treatment Plant in Qatar). The pH of the activated sludge was 6.9 ± 0.1 at 21.9°C, and the initial temperature of the activated sludge was 25°C ± 2°C. The pH, temperature, and dissolved oxygen were continuously measured during the experiments using a multimeter (WTW Multi3430, Weilheim, Germany). The experimental procedure is given below:

- The initial permeate flux (J_{PW}) was determined using pure distilled water by recording the permeate weight until a constant flux was obtained. The weight of the permeate was measured using a digital balance (EW-11017-04 Ohaus Ranger™ Scale) connected to a computer.
- The membrane was then submerged in the bioreactor and recorded the permeate weight over 2 min interval for 90 mins. The permeate flux was determined using Eq. (1).

$$J_{AS} = \frac{\left(W_t - W_0\right)}{A_m \Delta t} \quad (1)$$

where J_{AS} is the membrane flux, W_0 and W_t are the initial weight and weight at time t of the permeate, A_m is the membrane area, and Δt is the time interval.

- The normalized permeate flux for the MBR process was calculated using the ratio of permeate flux (J_{AS}) and the initial permeate flux using pure water (J_{PW}), that is, J_{AS}/J_{PW}.
- After each experiment, the membrane was washed with distilled water and the new clean permeate flux was measure again.
- The tests were conducted at different MLSS concentrations (MLSS = 5, 10, and 15 g/L) and aeration rate (AER = 0.5, 1.0, and 1.5 LPM) by repeating steps 1–4.
- Each experiment was repeated several times.

2.2. Neural network model

In this work, a feed-forward neural network, also called MLP, was used to model the MBR process. The network consists of an input layer, a hidden layer, and an output layer. The number of neurons in the input and the output layer is equal to the number of input and output variables of the model. The hidden layer is composed of hidden neurons that develop a relationship between the input and output neurons with the help of connection weights and biases. The four input variables were time, aeration scheme, MLSS concentration, and aeration rate to predict the output variable, which was the normalized permeate flux. The development of the model is carried out in three stages, that is, training, validation, and test stage. As ANN is a data-driven model, a total of 792 data points were randomly divided into three datasets, which consisted of a training data set (80%), validation data set (10%), and test data set (10%) for each stage. The description of the input and the output variables is given in Table 2. Figs. 2a and b show the ANN network architecture used in this study and a close-up of the hidden neuron that shows the calculation of its output, respectively.

The experimental data were initially normalized to avoid numerical overfitting caused by the very small or large weights associated with the neurons. The MATLAB's "mapminmax" function (Eq. (2)) was employed for the normalization of the input and output variables in the range of –1 to 1.

$$y = \frac{2\left(x - x_{\min}\right)}{\left(x_{\max} - x_{\min}\right)} - 1 \quad (2)$$

where y is the normalized value, $x_{\max}$ and $x_{\min}$ is the maximum and minimum value of the variable, respectively.

During the training stage, the training dataset was used to develop the neural network model. Initially, the weights and biases are assigned to the neurons using the Nguyen–Widrow initialization method for the calculation of the neuron's output [33]. For consistency and comparison, the initialization of the weights was fixed for all the trained networks, as randomization could lead to inconsistent performance [5]. The next step was the optimization of the ANN weights to obtain a high-performance model. Due to its high performance and convergence speed, Levenberg–Marquardt (LM) algorithm was selected to train the network and update the weights [23,34,35]. The neuron's output for the hidden and output layer is evaluated using Eq. (3):

$$a_{ij} = f_j\left(\sum_{k=1}^{n_{(j-1)}}\left(a_{k(j-1)}w_{ki(j-1)}\right) + b_{ij}\right) \quad (3)$$

where a_{ij} is the output of the neuron i in the layer j, f_j is the activation or transfer function of the layer j, $a_{k(j-1)}$ and $w_{ki(j-1)}$ are the output and the weight of neuron from the previous layer, respectively, $n_{(j-1)}$ is the neurons in the hidden layer (j–1) and b_{ij} is the bias associated with the neuron.

The presence of an activation function enables the network to model non-linear systems [5]. In this work, logistic sigmoid (log-sigmoid) function (Eq. (4)) or hyperbolic tangent sigmoid (tan-sigmoid) function (Eq. (5)) was used in the hidden layer, whereas pure linear (purelin) function (Eq. (6)) was used for the output layer. The graphical representation of the functions is illustrated in Fig. 3.

$$f(u) = \frac{1}{1 + e^{-u}} \quad (4)$$

$$f(u) = \frac{e^{2u} - 1}{e^{2u} + 1} \quad (5)$$

$$f(u) = u \quad (6)$$

The ANN weights are updated by minimizing the error between the targeted and predicted output. The learning process stops when one of the stopping criteria is reached, that is, minimum error, validation checks, etc. The performance of the trained network can be evaluated using the indicators listed in Table 3.

During the validation stage, the trained network is subjected to an unseen validation data set. This stage helps prevent overfitting of the model and serves as a stopping criterion. If the validation is unacceptable, the training stage continues to update the weights until it is satisfied. Lastly, the final stage of the ANN model is the test or prediction stage. The test stage does not participate

Table 2
Input and output variable description for the training of the ANN model

	Variables	Value
Input	Time, min	86
	Aeration scheme	Continuous (1)
		Pulsed (0)
	MLSS concentration, g/L	5
		10
		15
	Aeration rate, LPM	0.5
		1
		1.5
Output	Normalized permeate flux	–

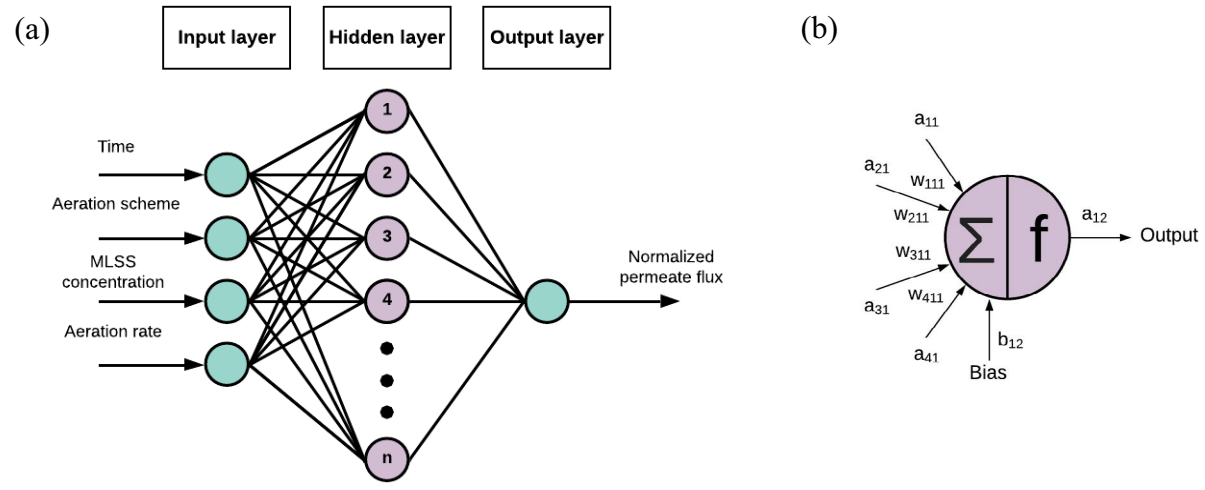

Fig. 2. (a) Network architecture of the ANN model and (b) close-up of the hidden neuron.

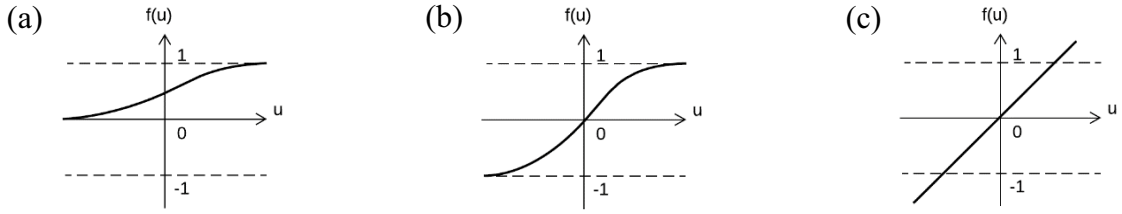

Fig. 3. Graphical representation of the activation functions (a) log-sigmoid, (b) tan-sigmoid, and (c) purelin.

Table 3
Performance indicators to compare the efficiency of the trained models

Performance indicator	Equation
Mean square error (MSE)	$\mathrm{MSE}=\frac{1}{N}\sum_{i=1}^{N}\left(y_{p,i}-y_{t,i}\right)^2$
Root mean square error (RMSE)	$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(y_{p,i}-y_{t,i}\right)^2}$
Determination coefficient	$R^2=1-\frac{\sum_{i=1}^{N}\left(y_{p,i}-y_{t,i}\right)}{\sum_{i=1}^{N}\left(y_{p,i}-y_{m}\right)}$
Adjusted determination coefficient	$R^2\ \mathrm{adjusted}=1-\frac{\left(1-R^2\right)\left(N-1\right)}{N-k-1}$

$y_{p,i}$ and $y_{t,i}$ are the predicted and target values of membrane flux, respectively; N is the total number of data points; y_m is the mean of the actual value of water; k is the number of input variables.

in the training of the network; however, it shows the generalization capability of the model when subjected to unseen data.

The value of the connection weights of the network may not hold any physical significance. Therefore, the weight partitioning methodology was used to calculate the relative importance (I_m) of each input variable [36,37]. The relative importance is given using Eq. (7) to demonstrate the sensitivity of the ANN model.

$$I_m\%=\frac{\sum_{j=1}^{n_h}\left\{\left(\frac{i_{mj}}{\sum_{k=1}^{n_m}i_{kj}}\right)O_j\right\}}{\sum_{l=1}^{n_m}\left[\sum_{j=1}^{n_h}\left\{\left(\frac{i_{mj}}{\sum_{k=1}^{n_m}i_{kj}}\right)O_j\right\}\right]}\times 100 \tag{7}$$

where m is the number of input neurons, n_h the number of hidden neurons, i_{mj} is the absolute value of connection weights between the input m and hidden neuron j, and O_j is the absolute value of connection weights between the hidden neuron j, and the output.

3. Results and discussion

3.1. Neural network model development

A total of 40 neural network models were trained, and their MSE was recorded at each stage. The hidden neurons in the network were changed from 1 to 20 neurons, with two different activation functions. The output neuron was always a pure linear function. Fig. 4a shows the MSE in each stage during the ANN modeling with a log-sigmoid function as the activation function for a different number of neurons. Except for the MSE peaks observed at 1, 8, and 14 neurons, increasing the number of neurons resulted in a decrease in the error. Similarly, Fig. 4b shows the rest of the 20 models with a tan-sigmoid function as the activation function. The error was high for neurons 1–3 in the hidden layer. An increase in the number of neurons showed a much smoother decrease in error as compared to the log-sigmoid function. The selection criteria for the optimum model was based on choosing the network with the least error in the training stage, without compromising on the efficiency in the validation and test stage. Therefore, a 15 hidden neuron network with a tan-sigmoid function (topology: 4-7-1) was selected as the optimum network, which is also encircled in Fig. 4b.

For the optimum network, the MSE during the training and updating of the weights in each stage is reflected in Fig. 5. The error trend for each stage was very similar, which showed that the training data set accurately represented both validation and test data sets, even though the data was divided randomly. The training was stopped at 47 epochs when a best validation stage performance of 1.188 was reached. Further training resulted in a decrease in the validation performance which indicates the over-training of the model.

Fig. 6 shows the scatter plot for the predicted and targeted output values in all three stages of the ANN modeling. The correlation coefficient (R-value) for the training, validation, and test stage was 0.99814, 0.99824, and 0.99792, respectively. For the complete experimental data, the R-value was 0.99816. A positive R-value closer to 1 shows a high correlation between the predicted and targeted output. This can be observed in Fig. 6 as the majority of the data points fall on the best fit line for each plot at different stages of modeling. The performance of the network can further be evaluated using MSE, RMSE, R^2, and

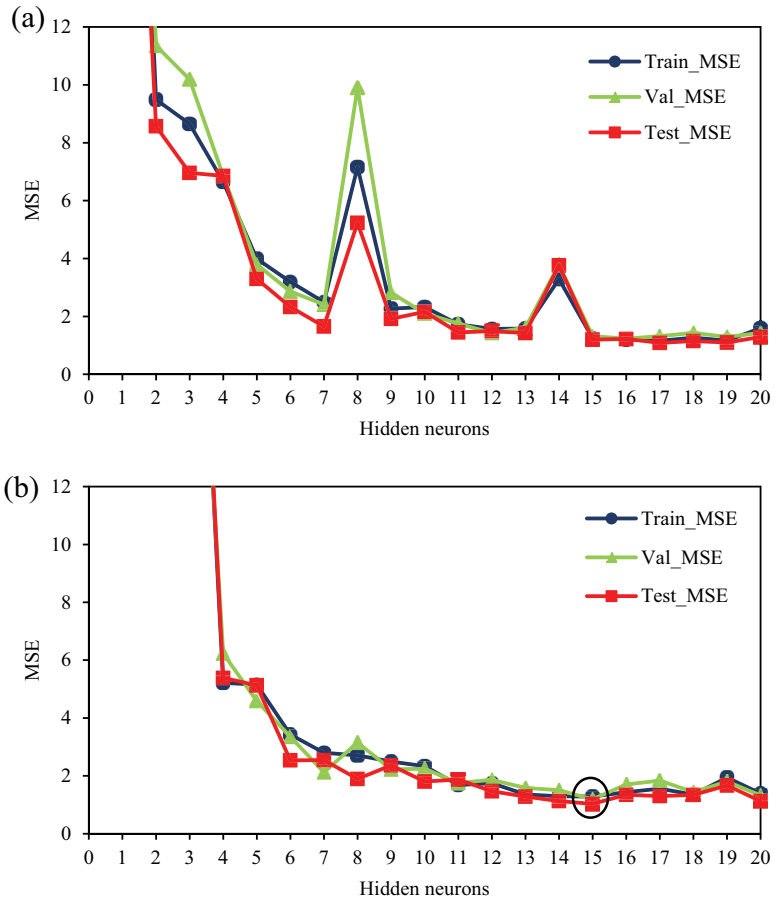

Fig. 4. Simulation summary for the selection of optimum model using (a) log-sigmoid function and (b) tan-sigmoid function.

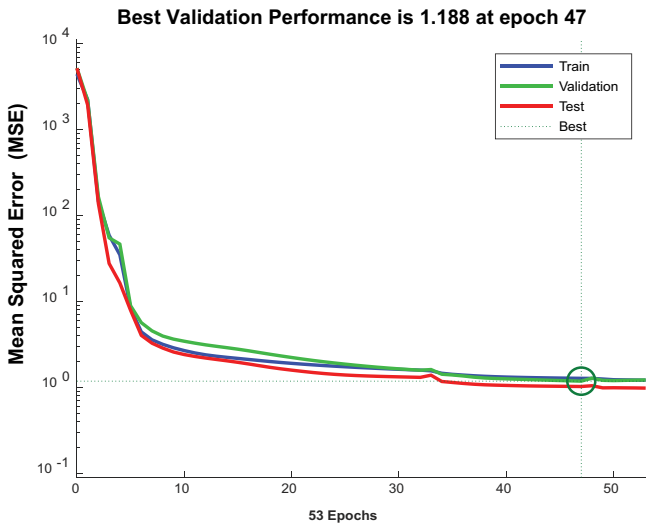

Fig. 5. Evolution of MSE in all three stages during the training of the optimum network.

adjusted R^2 as summarized in Table 4. The targeted MSE was set to zero for the training of the network. The lowest possible MSE for training, validation, and test stages was 1.281, 1.188, and 1.029, respectively. The low-value of

Table 4
Performance of the optimum network in different stages of ANN modeling

Data set	MSE	RMSE	R	R^2	Adj. R^2
Training	1.281	1.131	0.99814	0.996283	0.996265
Validation	1.188	1.090	0.99824	0.996483	0.996465
Test	1.029	1.015	0.99792	0.995844	0.995823
All	1.247	1.117	0.99816	0.996323	0.996305

the RMSE of each stage was comparable and, therefore, did not show any evidence of over-fitting. The determination coefficient (R^2) was greater than 0.99 for all the cases. This means that more than 99% of the variance in the data was explained by the model. Lastly, the adjusted R^2 was also greater than 99%, which showed that the input variables have a significant impact on the model output.

3.2. Impact of input variables on the normalized permeate flux

The aeration process accounts for most of the energy consumption in the MBR process. Therefore, it is essential to optimize the aeration process to minimize the energy requirements while maximizing the permeate flux. Figs. 7a–c represent the impact of aeration rate and

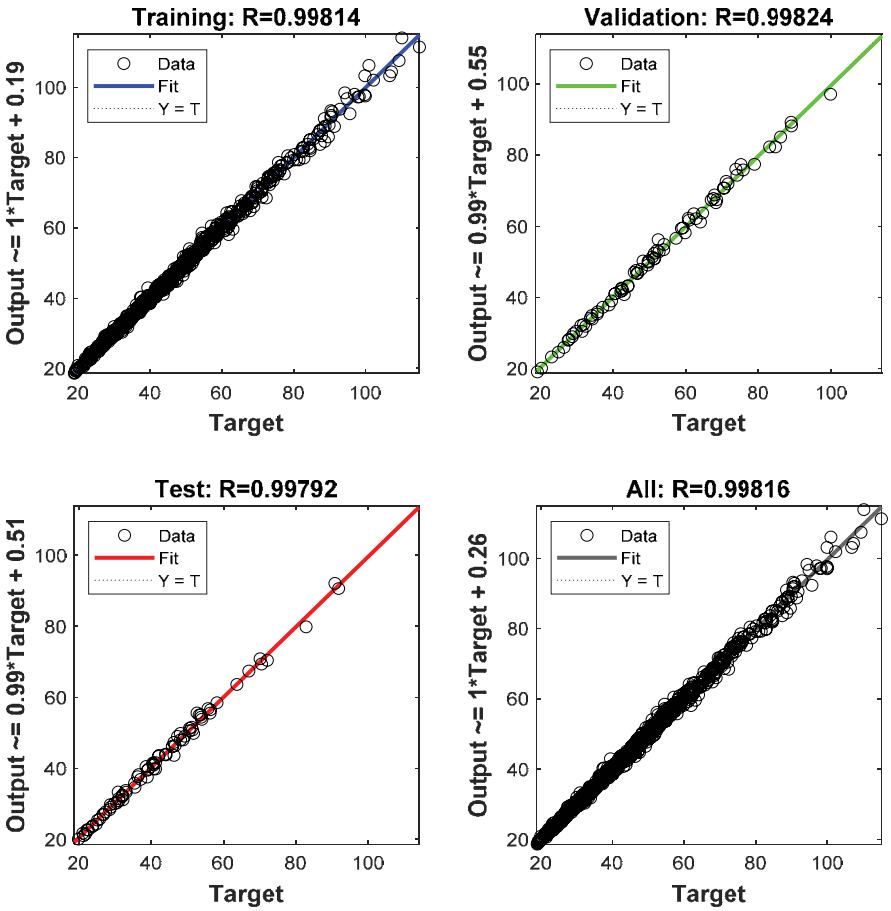

Fig. 6. Comparison of the targeted and predicted outputs in different stages of the ANN modeling.

scheme on the normalized permeate flux for the different MLSS concentrations. The Fig. 7 also show the predictions from the ANN model validating against the experimental data with excellent agreement. The change of normalized permeate flux with time for different aeration rates and schemes at 5 g/L MLSS concentration is shown in Fig. 7a. The normalized flux increased with an increase in the aeration rate from 0.5 to 1.5 LPM for both the pulsed and the continuous aeration modes. This is possibly due to the turbulence created on the surface of the membrane that prevents fouling [14,38]. In the beginning, the normalized flux for the pulsed aeration scheme was higher than the continuous scheme. However, the rate of decrease in the normalized flux for pulsed aeration was higher than continuous aeration. Therefore, by the end of the experiment, the normalized flux for continuous aeration became greater than the pulsed aeration scheme. Fig. 7b presents the change of normalized flux for different aeration rates at 10 g/L MLSS concentration. Similar to the previous case, the aeration rate showed a proportional impact on the normalized flux. The increase in the MLSS concentration resulted in a slight decrease in the initial normalized flux. As the MLSS concentration increased from 5 to 10 g/L, the rate of decrease of the normalized flux over the experimental time also increased. As the MLSS concentration increased more aeration (more oxygen) was required to sustain the biomass. The lack of oxygen causes the death of the biomass and promotes fouling. Fig. 7c shows the change in the normalized flux for different aeration rates and schemes at 15 g/L. A high MLSS concentration of 15 g/L led to a further increase in the rate of decrease of the normalized flux throughout the experiment. At this concentration, the only significant impact on the normalized flux was that of the aeration rate. The pulsed and continuous aeration scheme showed comparable results.

For each experiment, the average normalized flux is listed in Table 5. It can be seen from Table 5 that the average normalized flux decreases as the MLSS concentration increases from 5 to 15 g/L. It can be also seen from Table 5, that the average normalized flux increases as the aeration rate increases from 0.5 to 1.0 LPM. No clear correlation was observed between the impact of the aeration scheme (i.e., pulsed or continuous) on the average normalized flux. It was noticed, that the difference of the average normalized flux when using the pulsed aeration scheme compared to when using the continuous aeration scheme ranged between a minimum value of 0.5% at MLSS concentration 5 g/L and aeration rate 1.0 LPM to a maximum value of 21.1% at MLSS concentration 15 g/L and aeration rate 1.0 LPM. The highest average normalized flux was 63.5% obtained with a continuous aeration rate at 1.5 LPM and 5 g/L MLSS. The corresponding average normalized

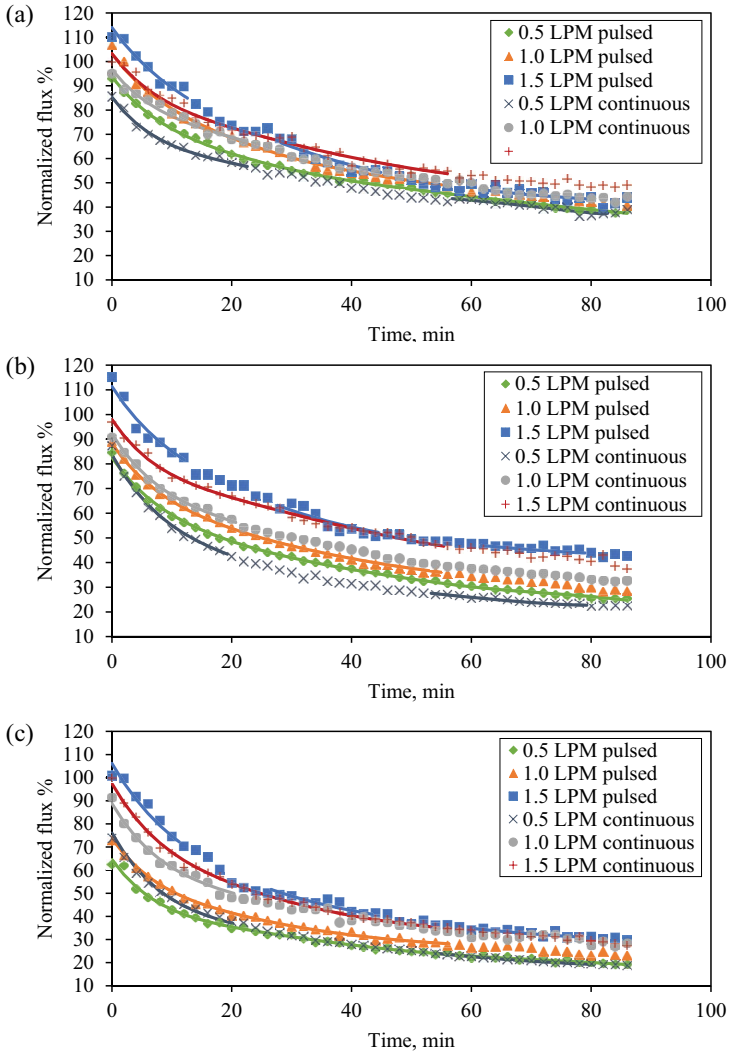

Fig. 7. Experimental and predicted normalized permeate flux for different aeration rates and schemes at MLSS concentrations of (a) 5 g/L, (b) 10 g/L, and (c) 15 g/L. The dots represent the experimental data, whereas the continuous and dashed lines represent neural network simulation.

flux with the pulsed aeration was 62.1% only 2.25% less than the continuous aeration scheme. It should be noted that the aeration scheme will have a direct impact on the energy requirements. Where using a pulsed aeration scheme will reduce the energy requirements by half. Therefore, a sensitivity analysis was performed to study the relative importance of each studied variable.

3.3. Sensitivity analysis

The optimal weights and biases associated with the neurons in the ANN model (4-7-1) are reported in Table 6. These weights were used to study the sensitivity of the model with respect to each input variable. The relative importance was calculated using Eq. (7) and illustrated as a pie chart in Fig. 8. The order of importance was evaluated to be time > aeration rate > MLSS concentration > aeration scheme. It can be seen from Fig. 8 that MLSS concentration, aeration scheme, and aeration rate had almost

Table 5
Comparison of average normalized permeate flux for various experiments at different aeration rate and scheme

MLSS concentration (g/L)	Aeration rate (LPM)	Average normalized flux (%)	
		Pulsed	Continuous
5	0.5	53.9	50.7
	1.0	58.7	59.0
	1.5	62.1	63.5
10	0.5	40.6	36.0
	1.0	45.1	48.3
	1.5	60.4	56.6
15	0.5	30.2	31.5
	1.0	35.5	43.0
	1.5	48.5	45.5

Table 6
Optimal values of weights and biases obtained in the training stage of the ANN model

Input weight matrix, IW	IW{1,1}				Bias vector, *b*	*b*{1}
{Destination: hidden layer	–0.1437	–0.1368	0.2770	–0.1264	{Destination: hid-	–1.7448
Source: inputs}	1.2812	0.6457	–1.8222	0.6545	den layer}	0.2067
	–4.4576	0.6730	2.8959	–0.9634		–0.0176
	1.4653	0.2406	–1.3937	0.4676		–0.3087
	0.8841	1.8539	–1.5731	0.5219		–0.4707
	–2.3992	1.4726	1.9476	–0.8672		0.0657
	–2.6486	–4.4092	–1.1459	–2.4493		–0.0258
	1.1772	3.6587	–4.5651	2.0152		0.1434
	2.1834	0.1113	3.5622	–0.4646		–0.0697
	3.3130	–0.0798	–0.0944	0.0570		–1.9276
	–1.2773	0.4904	–0.2035	0.2683		0.2786
	1.1701	1.3307	–0.5737	0.4330		–0.4539
	0.8585	–0.1999	–0.2020	–1.1597		–1.7091
	–0.9446	4.1958	5.2075	1.0702		0.0852
	0.7060	–1.6574	–0.2342	–2.2866		1.1536
Layer weight matrix, LW	$LW\{2,1\}^T$				Bias scalar, *b*	*b*{1}
{Destination: output layer	–0.3671				{Destination:	1.9504
Source: hidden layer}	–1.2167				output layer}	
	2.9207					
	–0.7894					
	–0.0418					
	1.1741					
	–5.3413					
	1.4996					
	1.2389					
	4.3080					
	–1.0224					
	2.4474					
	2.8002					
	–3.4759					
	5.0468					

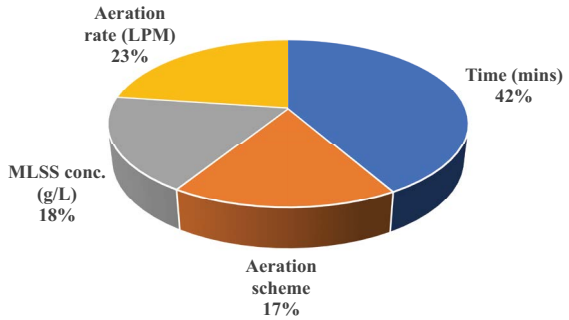

Fig. 8. Relative importance of each input variable is based on the weights in the optimum network.

the same importance at 18%, 17%, and 23%, respectively. Time was relatively the most important variable (42%) which had a negative impact on the normalized flux, as time increases, fouling occurs in the membrane which leads to the decline of the permeate flux. Taking energy into account, the aeration scheme (i.e., pulsed aeration) would decrease the energy consumption by half. As the weights of MLSS concentration, aeration rate, and aeration scheme were close to each other, it should be noted the aeration scheme has a higher impact on energy consumption.

4. Conclusion

The developed ANN model showed excellent agreement with the experimental data with an overall RMSE and R^2 of 1.117 and 0.9963, respectively. The average normalized flux decreased as the MLSS concentration increased from 5 to 15 g/L and it increased as the aeration rate increased from 0.5 to 1.0 LPM. No clear correlation was observed between the impact of the aeration scheme (i.e., pulsed or continuous) on the average normalized flux. Analysis of the ANN revealed the order of importance was time > aeration rate > MLSS concentration > aeration scheme.

Acknowledgments

This research is made possible by UREP award (UREP22-007-2-003) from Qatar National Research Fund (QNRF). The statements made herein are solely the responsibility of the authors.

References

[1] F. Du, A.H. Hawari, B. Larbi, A. Ltaief, G.R. Pesch, M. Baune, J. Thöming, Fouling suppression in submerged membrane bioreactors by obstacle dielectrophoresis, J. Membr. Sci. 549 (2018) 466–473.
[2] H. Hamedi, M. Ehteshami, S.A. Mirbagheri, S. Zendehboudi, New deterministic tools to systematically investigate fouling occurrence in membrane bioreactors, Chem. Eng. Res. Des., 144 (2019) 334–353.
[3] S.A. Mirbagheri, M. Bagheri, Z. Bagheri, A.M. Kamarkhani, Evaluation and prediction of membrane fouling in a submerged membrane bioreactor with simultaneous upward and downward aeration using artificial neural network-genetic algorithm, Process Saf. Environ. Prot., 96 (2015) 111–124.
[4] A. Hosseinzadeh, J.L. Zhou, A. Altaee, M. Baziar, X. Li, Modeling water flux in osmotic membrane bioreactor by adaptive network-based fuzzy inference system and artificial neural network, Bioresour. Technol., 310 (2020), doi: 10.1016/j.biortech.2020.123391.
[5] F. Schmitt, R. Banu, I.T. Yeom, K.U. Do, Development of artificial neural networks to predict membrane fouling in an anoxic-aerobic membrane bioreactor treating domestic wastewater, Biochem. Eng. J., 133 (2018) 47–58.
[6] M. Gander, B. Jefferson, S. Judd, Aerobic MBRs for domestic wastewater treatment: a review with cost considerations, Sep. Purif. Technol., 18 (2000) 119–130.
[7] M.C. Ozturk, F. Martin Serrat, F. Teymour, Optimization of aeration profiles in the activated sludge process, Chem. Eng. Sci., 139 (2016) 1–14.
[8] D. Guibert, R. Ben Aim, H. Rabie, P. Côté, Aeration performance of immersed hollow-fiber membranes in a bentonite suspension, Desalination, 148 (2002) 395–400.
[9] R. Van Kaam, D. Anne-Archard, M. Alliet, S. Lopez, C. Albasi, Aeration mode, shear stress and sludge rheology in a submerged membrane bioreactor: some keys of energy saving, Desalination, 199 (2006) 482–484.
[10] E. Braak, M. Alliet, S. Schetrite, C. Albasi, Aeration and hydrodynamics in submerged membrane bioreactors, J. Membr. Sci., 379 (2011) 1–18.
[11] R. Van Kaam, D. Anne-Archard, M.A. Gaubert, C. Albasi, Rheological characterization of mixed liquor in a submerged membrane bioreactor: interest for process management, J. Membr. Sci., 317 (2008) 26–33.
[12] S. Buetehorn, C.N. Koh, T. Wintgens, D. Volmering, K. Vossenkaul, T. Melin, Investigating the impact of production conditions on membrane properties for MBR applications, Desalination, 231 (2008) 191–199.
[13] J. Chang, W. Liang, E. Xiao, Z. Wu, Effect of intermittent aeration on the treatment performance in a submerged membrane bioreactor, Wuhan Univ. J. Nat. Sci., 15 (2010) 455–460.
[14] P. Le-Clech, B. Jefferson, S.J. Judd, Impact of aeration, solids concentration and membrane characteristics on the hydraulic performance of a membrane bioreactor, J. Membr. Sci., 218 (2003) 117–129.
[15] M. Bagheri, A. Akbari, S.A. Mirbagheri, Advanced control of membrane fouling in filtration systems using artificial intelligence and machine learning techniques: a critical review, Process Saf. Environ. Prot., 123 (2019) 229–252.
[16] M. Hamachi, M. Cabassud, A. Davin, M. Mietton Peuchot, Dynamic modelling of crossflow microfiltration of bentonite suspension using recurrent neural networks, Chem. Eng. Process. Process Intensif., 38 (1999) 203–210.
[17] M. Asghari, A. Dashti, M. Rezakazemi, E. Jokar, H. Halakoei, Application of neural networks in membrane separation, Rev. Chem. Eng., 36 (2018) 265–310.
[18] Z. Chen, N. Ren, A. Wang, Z.P. Zhang, Y. Shi, A novel application of TPAD-MBR system to the pilot treatment of chemical synthesis-based pharmaceutical wastewater, Water Res., 42 (2008) 3385–3392.
[19] Z. Liu, D. Pan, J. Wang, S. Yang, Modelling of Membrane Fouling by PCA-PSOBP Neural Network, 2010 International Conference on Computing, Control and Industrial Engineering, IEEE, Wuhan, 2010, pp. 34–37.
[20] C. Li, Z. Yang, H. Yan, T. Wang, The application and research of the GA-BP neural network algorithm in the MBr membrane fouling, Abstr. Appl. Anal., 2014 (2014), doi: 10.1155/2014/673156.
[21] A.R. Pendashteh, A. Fakhru'l-Razi, N. Chaibakhsh, L.C. Abdullah, S.S. Madaeni, Z.Z. Abidin, Modeling of membrane bioreactor treating hypersaline oily wastewater by artificial neural network, J. Hazard. Mater. 192 (2011) 568–575.
[22] S.A. Mirbagheri, M. Bagheri, S. Boudaghpour, M. Ehteshami, Z. Bagheri, Performance evaluation and modeling of a submerged membrane bioreactor treating combined municipal and industrial wastewater using radial basis function artificial neural networks, J. Environ. Health Sci. Eng., 13 (2015) 1–15.
[23] H. Hazrati, A.H. Moghaddam, M. Rostamizadeh, The influence of hydraulic retention time on cake layer specifications in the membrane bioreactor: experimental and artificial neural network modeling, J. Environ. Chem. Eng., 5 (2017) 3005–3013.
[24] S. Geissler, T. Wintgens, T. Melin, K. Vossenkaul, C. Kullmann, Modelling approaches for filtration processes with novel submerged capillary modules in membrane bioreactors for wastewater treatment, Desalination, 178 (2005) 125–134.
[25] C. Li, X. Wang, Application of MBR Membrane Flux Prediction Based on Elman Neural Network, in: DEStech Trans. Eng. Technol. Res., 2018, pp. 365–372. Available at: https://doi.org/10.12783/dtetr/iccere2017/18308.
[26] Y. Chen, G. Yu, Y. Long, J. Teng, X. You, B.Q. Liao, H. Lin, Application of radial basis function artificial neural network to quantify interfacial energies related to membrane fouling in a membrane bioreactor, Bioresour. Technol., 293 (2019), doi: 10.1016/j.biortech.2019.122103.
[27] Z. Zhao, Y. Lou, Y. Chen, H. Lin, R. Li, G. Yu, Prediction of interfacial interactions related with membrane fouling in a membrane bioreactor based on radial basis function artificial neural network (ANN), Bioresour. Technol., 282 (2019) 262–268.
[28] Y. Cai, X. Li, A.A. Zaidi, Y. Shi, K. Zhang, R. Feng, A. Lin, C. Liu, Effect of hydraulic retention time on pollutants removal from real ship sewage treatment via a pilot-scale air-lift multilevel circulation membrane bioreactor, Chemosphere, 236 (2019), doi: 10.1016/j.chemosphere.2019.07.069.
[29] Y. Cai, A.A. Zaidi, Y. Shi, K. Zhang, X. Li, S. Xiao, A. Lin, Influence of salinity on the biological treatment of domestic ship sewage using an air-lift multilevel circulation membrane reactor, Environ. Sci. Pollut. Res., 26 (2019) 37026–37036.
[30] Y. Cai, X. Li, A.A. Zaidi, Y. Shi, K. Zhang, P. Sun, Z. Lu, Processing efficiency, simulation and enzyme activities analysis of an air-lift multilevel circulation membrane bioreactor (AMCMBR) on marine domestic sewage treatment, Period. Polytech., Chem. Eng., 63 (2019) 448–458.
[31] Y. Cai, T. Ben, A.A. Zaidi, Y. Shi, K. Zhang, A. Lin, C. Liu, Effect of pH on pollutants removal of ship sewage treatment in an innovative aerobic-anaerobic micro-sludge MBR system, Water Air Soil Pollut., 230 (2019), doi: 10.1007/s11270-019-4211-0.
[32] Z. Chen, A. Zhou, N. Ren, Y. Tian, D. Hu, Pollutants removal and simulation model of combined membrane process for wastewater treatment and reuse in submarine cabin for long voyage, J. Environ. Sci. 21 (2009) 1503–1512.
[33] D. Nguyen, B. Widrow, Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights, 1990 IJCNN International Joint Conference on Neural Networks, 1990, San Diego, CA, 1990.

[34] J.J. Moré, The Levenberg–Marquardt Algorithm: Implementation and Theory, G.A. Watson, Ed., Numerical Analysis, Lecture Notes in Mathematics, Vol. 680, Springer, Berlin, Heidelberg, 1978, pp. 105–116.
[35] M. Barello, D. Manca, R. Patel, I.M. Mujtaba, Neural network based correlation for estimating water permeability constant in RO desalination process under fouling, Desalination. 345 (2014) 101–111.
[36] G.D. Garson, Interpreting neural-network connection weights, AI Expert, 6 (1991) 46–51.
[37] A.T.C. Goh, Seismic liquefaction potential assessed by neural networks, J. Geotech. Eng., 120 (1994), doi: 10.1061/(ASCE)0733-9410(1994)120:9(1467).
[38] H.C. Chua, T.C. Arnot, J.A. Howell, Controlling fouling in membrane bioreactors operated with a variable throughput, Desalination, 149 (2002), 225–229.