

Removal efficiency prediction model based on the artificial neural network for pollution prevention in wastewater treatment plants

Hussein Y.H. Alnajjar and Osman Üçüncü
Karadeniz Technical University, Trabzon, Turkey

Abstract

Purpose – Artificial intelligence (AI) models are demonstrating day by day that they can find long-term solutions to improve wastewater treatment efficiency. Artificial neural networks (ANNs) are one of the most important of these models, and they are increasingly being used to forecast water resource variables. The goal of this study was to create an ANN model to estimate the removal efficiency of biological oxygen demand (BOD), total nitrogen (TN), total phosphorus (TP) and total suspended solids (TSS) at the effluent of various primary and secondary treatment methods in a wastewater treatment plant (WWTP).

Design/methodology/approach – The MATLAB App Designer model was used to generate the data set. Various combinations of wastewater quality data, such as temperature (T), TN, TP and hydraulic retention time (HRT) are used as inputs into the ANN to assess the degree of effect of each of these variables on BOD, TN, TP and TSS removal efficiency. Two of the models reflect two different types of primary treatment, while the other nine models represent different types of subsequent treatment. The ANN model's findings are compared to the MATLAB App Designer model. For evaluating model performance, mean square error (MSE) and coefficient of determination statistics (R^2) are utilized as comparative metrics.

Findings – For both training and testing, the R values for the ANN models were greater than 0.99. Based on the comparisons, it was discovered that the ANN model can be used to estimate the removal efficiency of BOD, TN, TP and TSS in WWTP and that the ANN model produces very similar and satisfying results to the APPDESIGNER model. The R -value (Correlation coefficient) of 0.9909 and the MSE of 5.962 indicate that the model is accurate. Because of the many benefits of the ANN models used in this study, it has a lot of potential as a general modeling tool for a range of other complicated process systems that are difficult to solve using conventional modeling techniques.

Originality/value – The objective of this study was to develop an ANN model that could be used to estimate the removal efficiency of pollutants such as BOD, TN, TP and TSS at the effluent of various primary and secondary treatment methods in a WWTP. In the future, the ANN could be used to design a new WWTP and forecast the removal efficiency of pollutants.

Keywords Artificial intelligence, Artificial neural network, Biological treatment, Biological oxygen demand, Total nitrogen, Total phosphorus, Total suspended solids

Paper type Research paper

Introduction

Domestic and industrial untreated wastewater is one of the most serious producing environmental pollutants. To realize the treatment of wastewater, it is very important in increasing the removal efficiency of contaminants and the stability of operation in the wastewater treatment systems. Therefore, a computing operating system can be developed to enhance the stability of wastewater treatment systems (Niu *et al.*, 2020). However, the



wastewater treatment process is extremely complicated, its mechanism is difficult to understand, and a traditional operating system based on a mathematical model could not effectively simulate the wastewater treatment process. With artificial intelligence (AI) development, AI has obvious advantages over traditional process control methods, which could model the complex nonlinear wastewater treatment processes by modeling human thinking processes (Qiao, Wang, Li, & Li, 2018).

So, neural networks (NNs) are one of the most rapidly developing AI technologies. On the other hand, they can not only describe complex phenomena by mapping nonlinear functions, but they also have the advantage of being self-learning and self-adaptive, allowing them to compensate for the shortcomings of traditional control systems (Shin, Kim, Yu, Kim, & Hwang, 2019). In recent years, NN has been capable of successfully modeling wastewater treatment.

In the literature, it was an intermittent cycle extended aeration-sequential batch reactor, and artificial neural network (ANN) models were developed to predict faecal coliform and total coliform elimination (ICEAS-SBR). This network was developed using wastewater influent pH, biochemical oxygen demand (BOD), chemical oxygen demand (COD), total suspended solids (TSS), oil and grease (O&G), total Kjeldahl nitrogen (TKN), ammoniacal nitrogen (AN), total phosphorus (TP), faecal coliform and total coliform. ANN models allow for the regulation of faecal coliform and total coliform levels in treated wastewater effluent, lowering public health concerns, particularly for oyster consumers (Khatri, Khatri, & Sharma, 2020).

In another previous study, ANNs were used to mimic an anaerobic fermentation process for biogas production coupled with wastewater purification in a modern wastewater treatment plant (WWTP). Based on real-scale industrial data, neural models were trained, validated and tested, considering both technological aspects of the process and the quality of treated effluent. A parameter sensitivity study revealed that the operation process factors had a greater impact on biogas yield than the wastewater quality (COD, BOD₅, TSS, Pg and Ng). The proposed ANN model can be employed as a forecasting tool, as well as in the testing of other prospective process intensification and optimization scenarios (Sakiewicz, Piotrowski, Ober, & Karwot, 2020).

In one more study, the Jamnagar Municipal Corporation Sewage Treatment Plant (JMC-STP) was investigated in order to create a feedforward artificial neural network (FF-ANN) model. It was an alternative to the flexible physical, chemical and biological treatment process simulations for JMC-STP modeling and prediction. pH, BOD, COD, TSS, total Kjeldahl nitrogen (TKN), AN and TP were the expected effluent parameters. FF-ANN models were assessed using the MAD (mean absolute deviation), MSE (mean square error), RMSE (root mean square error) and MAPE (mean absolute percentage error). This serves as a helpful tool for the management of the plant to maximize the treatment quality while improving the plant's efficiency and dependability (Khatri, Khatri, & Sharma, 2019).

AI models have also been used to anticipate BOD elimination; a hyperbolic design equation was created using the ANN predictions. This equation combines zero and first-order kinetics. The results of the ANNs and the model design equation were compared to data from the literature and found to be reasonably accurate. The elimination of COD was shown to be highly linked with the removal of BOD. A formula for predicting COD elimination was also developed (Akratos, Papaspyros, & Tsihrintzis, 2008).

ANN models can be used to solve a variety of modeling problems in rivers, lakes, WWTPs, groundwater, ponds and streams (Chen, Song, Liu, Yang, & Li, 2020).

Another study describes the creation of a synthetic neural network model for predicting annual BOD values using widely accessible sustainability and economic/industrial parameters as inputs, after which the initial general regression neural network (GRNN) model was trained, validated and tested using 20 inputs. The proposed GRNN model can be

beneficial as a tool to support the decision-making process on sustainable development at a regional, national and worldwide level, it was determined in the end (Šiljić, Antanasijević, Perić-Grujić, Ristić, & Pocajt, 2014).

Various treatment processes, including air flotation, chemical coagulation, sedimentation and biological treatment through a fully mixed activated sludge process in a water purification process, were used to treat the entire effluent and waste disposal in the detergent industry before the soft computing techniques. Then a feed-forward with five layers The backpropagation ANN model was successfully used to optimize the proposed models, yielding the lowest root mean square error (0.066), mean square error (0.0043) and greatest R^2 value (0.996); these values demonstrate that the predicted and experimental responses were similar, and ANN may be used to describe the process (Jana, Bhunia, Das Adhikary, & Bej, 2022).

In this study, the purpose is developing the ANN-based models to predict the removal efficiency of BOD, TN, TP and TSS in a variety of primary and biological treatment systems in a WWTP. To achieve this purpose the neuron numbers in the hidden layer were varied to create models with different ANN topologies. For training and independent validation, the correlation coefficient and MSE were used to assess the performance of the created ANN models. So the removal effectiveness of BOD, TN, TP and TSS in the effluent for disposal is monitored using these models. Also the created ANN models can be used to manage the BOD, TN, TP and TSS removal efficiency to limit the danger of public health and discharge high-quality treated wastewater to receiving water bodies. Thus, it can be said that AI models, particularly NNs, make it possible to manage the operation of treatment plants and control pollutants without having to conduct traditional tests. This saves time and money, and in the future, it will aid designers in creating treatment plants more quickly, with higher treatment efficiency.

Materials and methods

The idea of duplicating the working principles of the brain on digital computers gave rise to the concept of ANNs, and the initial studies concentrated on mathematical modeling of the biological cells that make up the brain, referred to as neurons in the literature (akdağ & Karahan, 2014).

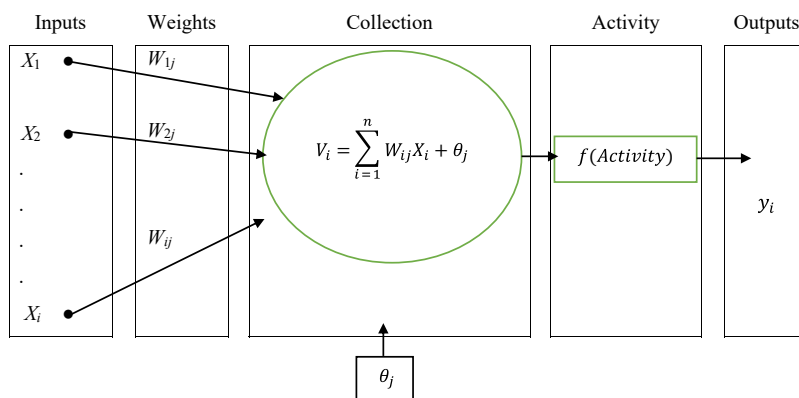
The operation of biological neurons served as the inspiration for the numerical method known as the ANN. An input signal vector x_i with the values 1, 2, ..., L is received by neuron m from a total of L input channels. The neuron then calculates the weighted sum of components x_i by multiplying each component x_i by the coefficient w_{mi} that reflects the significance of the input channel i as shown in Figure 1 (Cardoso, de Almeida, Dias, & Coelho, 2008).

Creating the neural network

An ANN is made up of a collection of very simple and densely interconnected processors known as neurons, which are similar to biological neurons in the brain. Weighted linkages connect the neurons, transferring messages from one to the other. Through its connections, each neuron gets a variety of input signals, but it never creates more than one output signal. The outgoing link of the neuron transmits the output signal (corresponding to the biological axon). The outgoing link then separates into several branches, each of which transmits the same signal (the signal is not divided among these branches in any way). The incoming connections of other neurons in the network end the outward branches. Figure 1 symbolized the way an artificial neuron works (Negnevitsky, 2005).

The addition function operates as given in equation (1).

Figure 1.
Architecture of a
typical artificial neural
network



$$V_i = X_1 W_1 + X_2 W_2 + \dots + X_n W_n + \theta_j = \sum_{i=1}^n W_{ij} X_i + \theta_j \quad (1)$$

$$y = f(V)$$

As shown in [Figure 1](#), each input generates a change in the neuron output, and the magnitude of this change is determined by the connection gains that determine the input's effect degree, the adder's threshold value and the type of neuron activation function. Where V is the addition function; gains denoted by W_i are weight; X_i inputs; θ_j value as the threshold; y output; The f function is also called the neuron activation function. As can be seen from the equations above, since the threshold value is independent of the inputs, in cases where all inputs are zero, the value of $f(\theta)$ is observed at the neuron output instead of $f(0)$, which eliminates the necessity for the neuron output to be zero under the specified conditions. The use of the threshold value is considered in practice as an input with a value of $+1$ or -1 entering the adder with a link with a weight of θ .

The activation function of the neuron is one of the most critical aspects in influencing neuron behavior. This function analyzes the neuron's net input and calculates the output the neuron will produce in response to it ([Öztemel, 2008](#)). There are many approaches for calculating the output in this function, similar to the addition function, and not all process elements must utilize the same activation function. Depending on the type of problem and the network structure employed, different functions may be preferred. The linear function, step function, sigmoid function and hyperbolic tangent function are commonly employed as activation functions ([Yurtoğlu, 2005](#)). The mathematical expressions of three of these activation functions used are given below.

$$\text{Sigmoid type activation function: } f(V) = \frac{1}{1 + e^{-V}} \quad (2)$$

$$\text{Tangent sigmoid type activation function: } f(V) = \frac{2}{1 + e^{-2V}} - 1 \quad (3)$$

$$\text{Hyperbolic tangent type activation function: } f(V) = \frac{e^V - e^{-V}}{e^V + e^{-V}} \quad (4)$$

The input layer, the output layer and the hidden layer are the three primary layers. The input layer is the first layer, and it feeds the ANN with external data input. In statistics, these data correspond to independent variables. The number of neurons in the input layer is formed by the number of parameters affecting the problem, and the number of parameters affects the number of neurons in the input layer. The output layer is the final layer, and it is responsible for transmitting data to the outside world. In statistics, output variables correspond to dependent variables. The hidden layer is the layer that sits between the input and output layers in the model. The neurons in the hidden layer are not connected to the outside world. They simply receive and send signals to and from the input layer (Küçükocaoglu, Keskin Benli, & Kucuksozen, 2005).

Neural network training (learning)

From the moment they are born, people begin the process of “learning by doing.” The brain continues to develop during this process. Learning occurs as we live and experience by modifying synaptic connections and even generating new ones. This holds true for ANN as well. Learning occurs through training with examples; in other words, realization occurs through processing input/output data, i.e. the training algorithm modifying the synapses’ weights using that data until convergence is attained.

As in a biological neural network, learning in ANNs is the act of altering the weight values between neurons to fulfill a specified function. These weight values are initially assigned at random. As examples are displayed to ANNs, their weight values fluctuate. The goal is to determine the weight values that will result in the correct outputs for the network’s examples. When the network’s weight values are correct, it suggests the network can generalize about the events represented by the samples. “Network learning” is the process of ANNs obtaining the ability to generalize about unknown cases by extracting specific information from previous examples (Agatonovic-Kustrin and Beresford, 2000).

“Testing” the network refers to attempts to determine whether the network learns (performs) after the training is completed. Examples that the network has not seen during learning are utilized for testing. Using the connection weights calculated during training, the network generates outputs for these occurrences that it does not see. The accuracy values of the produced outputs provide information on the network’s learning. The better the results, the more effective the training. The “training set” is a sample set used in education, and the “test set” is a sample set used for testing (Öztemel, 2008).

We can define the ANN’s learning of the relationship in the data structure as the determination of the most appropriate values of the network weights with the help of the examples of the problem. For any weight (W);

$$W_{new} = W_i \mp \Delta W \quad (5)$$

The equation expresses how learning takes place mathematically. The ΔW in equation (5) is calculated according to a certain rule and gives the amount of change of the current weight values. The rules defined for determining ΔW are called “learning algorithms”. Many learning algorithms have been proposed to help find the best weight set (Chang, Chen, & Shieh, 2001).

Evaluation of predicting performance

To evaluate the predicting performance of ANN, correlation coefficient (R) and MSE were employed and described as:

$$R = \frac{\sum_{i=1}^n (obs_i - \overline{obs}) (\overline{pre} - pre_i)}{\sqrt{\sum_{i=1}^n (obs_i - \overline{obs})^2 \sum_{i=1}^n (\overline{pre} - pre_i)^2}} \quad (6)$$

$$MSE = \frac{1}{n} \sum_{i=1}^n (obs_i - pre_i)^2 \quad (7)$$

where obs_i is the observed value, pre_i is the prediction value, $\overline{obs}$ and $\overline{pre}$ are the average values of observed values and prediction values, respectively (Pai *et al.*, 2009).

Applying of ANN model on WWTP

The Durugöl Advanced Biological WWTP was controlled by an ANN that was put to the test. Additionally, the MATLAB App Designer model was developed and compared to the neural networks model using data from the Durugöl Advanced Biological WWTP processing facility.

After biological treatment, the Durugöl Advanced Biological WWTP has a capacity of 212,000 persons/day and was built using physical (coarse screen, fine screen and primary settler units) and advanced biological treatment projects (anaerobic tanks, aeration tanks and secondary settler). The year 2014, saw its commissioning as a discharge.

In order to protect the system, the grits that are present in the influent wastewater are removed from the grit chamber. During the primary treatment, a sizable portion of BOD, COD, SS and other pollutants are eliminated. As shown in Figure 2, the secondary treatment unit, which comprises anaerobic tanks, aeration tanks and a secondary settler receives effluent from the primary settler. The aeration tanks offer the ideal environment for the microorganisms needed to develop and produce sludge as they break down the residual dissolved organic contaminants in the wastewater. In the secondary settler, gravity sedimentation is used to separate the sludge from the cleaned water. To keep the microbe concentration high, some of the sludge is returned to the aeration unit, while the waste sludge is taken out and delivered to the sludge treatment plant. In order to apply control strategy or optimization approaches to the plant and improve treatment efficiency, a proper model may be helpful. In 2018, wastewater samples were taken twice a month from the Durugöl Advanced Biological WWTP intake and exit.



Figure 2.
(a) The interface of the MATLAB App Designer model for primary treatment (b) The interface of the MATLAB App Designer model for secondary treatment

Different types of primary treatment (two models) and secondary treatment (nine models) were the main processing techniques of the wastewater treatment plan.

Two types of primary treatment techniques were modeled, the first is a mechanical screen and primary sedimentation tank and the second is the mechanical screen, grit removal, grease trap and primary sedimentation tank.

As for secondary treatment, a comparison was made between nine different models that were clarified in [Table 1](#).

The MATLAB program was used to create a model that helps to control and predict the expected results of WWTPs (primary treatment and secondary treatment) using the App Designer command. Then, these results were compared with the WATER POLLUTION CONTROL REGULATION (Turkey) for treated wastewater to be drained into receiving water body. The model relies on a set of inputs within a given code to produce the outputs as shown in [Figures \(2a and b\)](#).

The inputs and outputs of the MATLAB model were used to then model the results obtained using the neural networks model, and [Table 2](#) shows part of the results of the App Designer model.

Among the most important inputs are the hydraulic retention time (HRT), temperature (T) and dissolved oxygen (DO). As for the output ratio, it is the removal efficiency of biological oxygen demand (BOD₅), TN, TP and TSS. The results of the App Designer models are shown in [Table 3](#).

The results obtained from the App Designer model were modeled using neural networks. On the other hand, the network was built with its three layers (inputs, hidden layers and outputs). The transition function is (in this study, the tangent sigmoid function) is selected.

The difference between the actual and desired output values is calculated, and the network model's link weights are adjusted based on the results. The creation of the network,

Table 1.
Modeled primary and
secondary treatment
techniques

System number	Processing	Techniques
1	Primary Treatment	Mechanical screen + primary sedimentation tank
2	Primary Treatment	Mechanical screen + Grit removal + Grease trap + primary sedimentation tank
3	Secondary Treatment	Facultative pond + Secondary sedimentation tank
4	Secondary Treatment	Anaerobic ponds + Facultative Pond + Secondary sedimentation tank
5	Secondary Treatment	Aerobic ponds(Partial Mixing) + Facultative Pond + Secondary sedimentation tank
6	Secondary Treatment	Aerobic ponds (Complete Mixing) + Facultative Pond + Secondary sedimentation tank
7	Secondary Treatment	Anaerobic ponds + Secondary sedimentation tank
8	Secondary Treatment	Aerobic ponds (Partial Mixing) + Secondary sedimentation tank
9	Secondary Treatment	Aerobic ponds (Complete Mixing) + Secondary sedimentation tank
10	Secondary Treatment	Anaerobic ponds + Aerobic ponds(Partial Mixing) + Secondary sedimentation tank
11	Secondary Treatment	Anaerobic ponds + Aerobic ponds(Complete Mixing) + Secondary sedimentation tank

Table 2.
Example of the results
of the App
Designer model

#	Inputs			Outputs		
	HRT	T	BOD	TN	TP	TSS
1	30.00	−5.00	15.00	1.00	1.00	10.00
2	37.39	−4.00	15.87	1.74	1.52	10.87
3	44.78	−3.00	16.74	2.48	2.04	11.74
4	52.17	−2.00	17.61	3.22	2.57	12.61
5	59.57	−1.00	18.48	3.96	3.09	13.48
6	66.96	0.00	19.35	4.70	3.61	14.35
7	74.35	−5.00	20.22	5.43	4.13	15.22
8	81.74	−4.00	21.09	6.17	4.65	16.09
9	89.13	−3.00	21.96	6.91	5.17	16.96
10	96.52	−2.00	22.83	7.65	5.70	17.83
11	103.91	−1.00	23.70	8.39	6.22	18.70
12	111.30	0.00	24.57	9.13	6.74	19.57
13	118.70	−5.00	25.43	9.87	7.26	20.43
14	126.09	−4.00	26.30	10.61	7.78	21.30
15	133.48	−3.00	27.17	11.35	8.30	22.17
16	140.87	−2.00	28.04	12.09	8.83	23.04
17	148.26	−1.00	28.91	12.83	9.35	23.91
18	155.65	0.00	29.78	13.57	9.87	24.78
19	163.04	−5.00	30.65	14.30	10.39	25.65
20	170.43	−4.00	31.52	15.04	10.91	26.52
21	177.83	−3.00	32.39	15.78	11.43	27.39
22	185.22	−2.00	33.26	16.52	11.96	28.26
23	192.61	−1.00	34.13	17.26	12.48	29.13
24	200.00	0.00	35.00	18.00	13.00	30.00
25	30.00	1.00	35.87	18.74	13.52	30.87
26	32.88	2.00	36.02	18.90	13.68	31.02

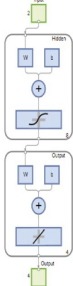
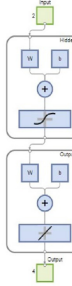
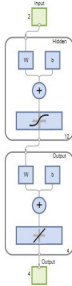
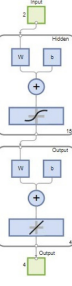
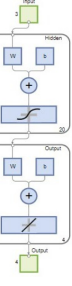
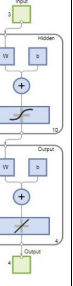
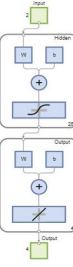
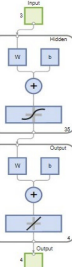
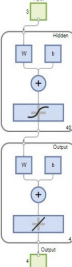
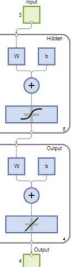
Table 3.
A summary of the
results obtained from
the MATLAB App
Designer model

System number	Removal efficiency (%)							
	BOD		TN		TP		TSS	
	From	To	From	To	From	To	From	To
1	2.70	33.00	0.80	27.00	1.50	32.00	4.50	60.00
2	2.70	34.50	0.80	29.50	1.50	34.50	6.00	65.00
3	10.00	85.00	4.00	59.95	1.00	35.00	10.00	80.00
4	15.00	85.00	1.00	60.00	1.00	35.00	10.00	80.00
5	5.00	85.00	1.00	31.50	1.00	35.00	50.00	80.00
6	10.00	88.00	3.00	35.00	5.00	40.00	50.00	80.00
7	5.00	75.00	3.00	60.00	1.00	35.00	40.00	90.00
8	6.00	80.00	1.50	30.00	1.00	35.00	50.00	80.00
9	6.00	80.00	1.00	30.00	5.00	40.00	50.00	87.00
10	5.00	85.00	5.00	60.00	3.00	35.00	65.00	90.00
11	9.00	90.00	6.00	65.00	3.00	40.00	65.00	90.00

which starts with the connections of the output layers and concludes with the connections of the input layers, realizes the return passage resulting from the weights of the connections.

The data in the data set are divided into three portions at random: training, validity and test sets. The data from the training set is utilized to train the network. The validity set is used in conjunction with a classifier's weights. The validity set is used to determine how many hidden units are present in an ANN. The test set is used to assess the training's effectiveness. As stated in Table 4, 70% of the observation data is allocated to the training set, 15% to the validity set and 15% to the test set.

Table 4.
The number of
observation data and a
diagram of the neural
network of the
systems, showing the
number of inputs,
outputs and hidden
layers

System number	1	2	3	4	5	6
Observations	120	120	144	144	576	576
illustration figures						
System number	7	8	9	10	11	
Observations	205	576	576	576	576	
illustration figures						

Results and discussion

Based on the data generated by the App Designer model using MATLAB software (MathWorks Inc., USA); an ANN model was created to estimate the final removal efficiency of BOD, TN, TP and TSS. The first model used a three-layer feed-forward neural network (2–8–4).

Table 5.
The accuracy of the
ANN models

System number	Observations	ANN architecture (number of neurons in layers)	Number of iterations (epoch)	Coefficient of determination (<i>R</i>)	Mean squared error (MSE)
1	120	(2, 8, 4)	36	0.9983	0.5594
2	120	(2, 10, 4)	9	0.9922	3.8031
3	144	(2, 12, 4)	27	0.9927	5.5298
4	144	(2, 15, 4)	17	0.9909	5.9620
5	576	(3, 20, 4)	49	0.9981	2.3641
6	576	(3, 10, 4)	26	0.9972	3.1428
7	205	(2, 25, 4)	20	0.9978	2.4704
8	576	(3, 30, 4)	110	0.9998	0.2767
9	576	(3, 35, 4)	197	0.9999	0.1260
10	576	(3, 40, 4)	174	1.0000	0.0337
11	576	(3, 5, 4)	165	0.9944	9.5476

The input layer contains two neurons, the hidden layer contains eight neurons and the output layer contains four neurons.

The input neurons represented App Designer model variables (such as DO concentration, temperature and HRT), whereas the output neurons indicated BOD, TN, TP and TSS elimination efficiency. To acquire the minimum value of mean square error, the number of

ANN for
pollution
prevention in
WWTPs

619

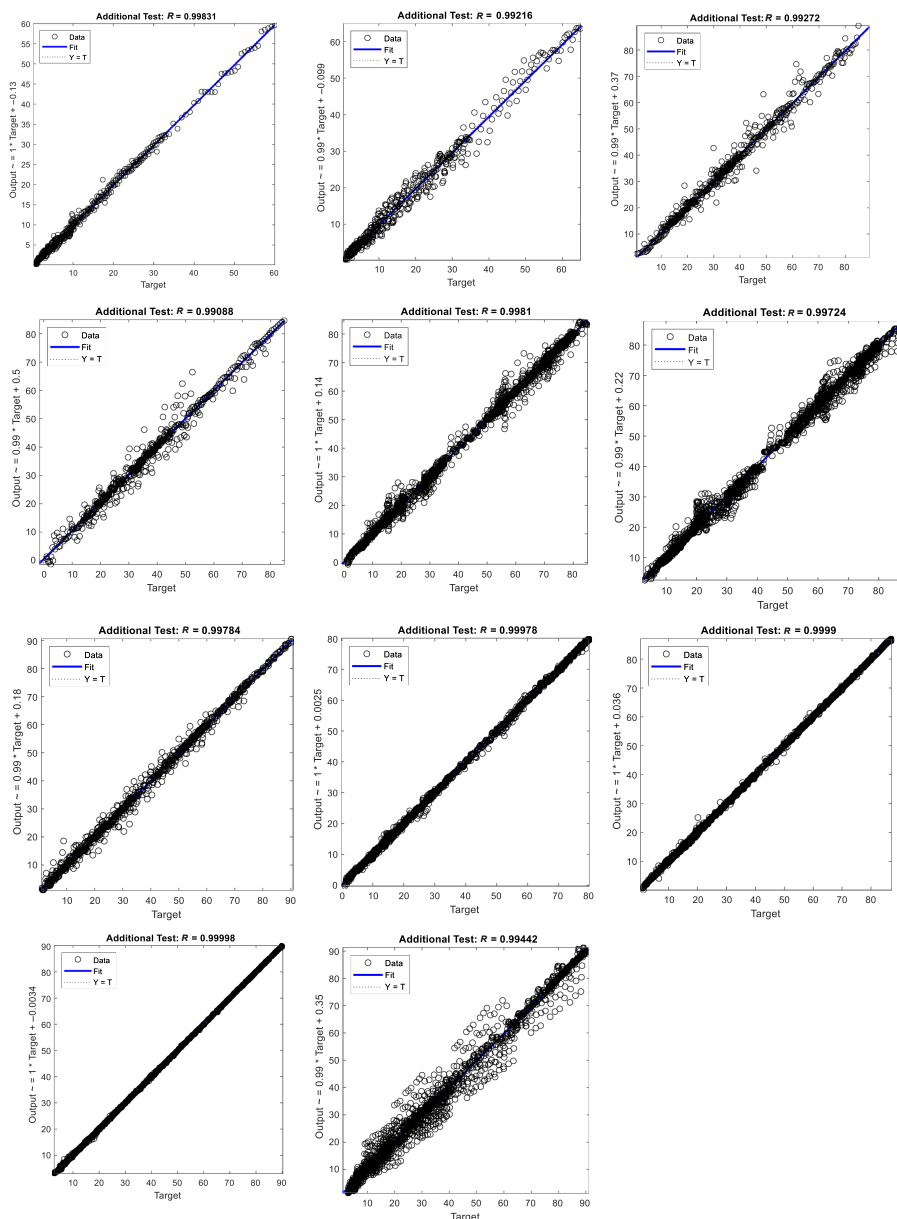


Figure 3.
The distribution of the
App Designer model
and predicted values
for ANN model

hidden neurons was determined using the trial-and-error method (MSE). The hyperbolic tangent sigmoid (tanh) function was selected for the hidden neurons in the ANN model because it performed well in regression when compared to the sigmoid function. In this scenario, the output neuron was given an identity (purelin) function.

As an example, for the ANN model a total of 144 sets of 100 data were used to train and test the model. Specifically, 22 sets were used for training, and 22 sets for testing the neural

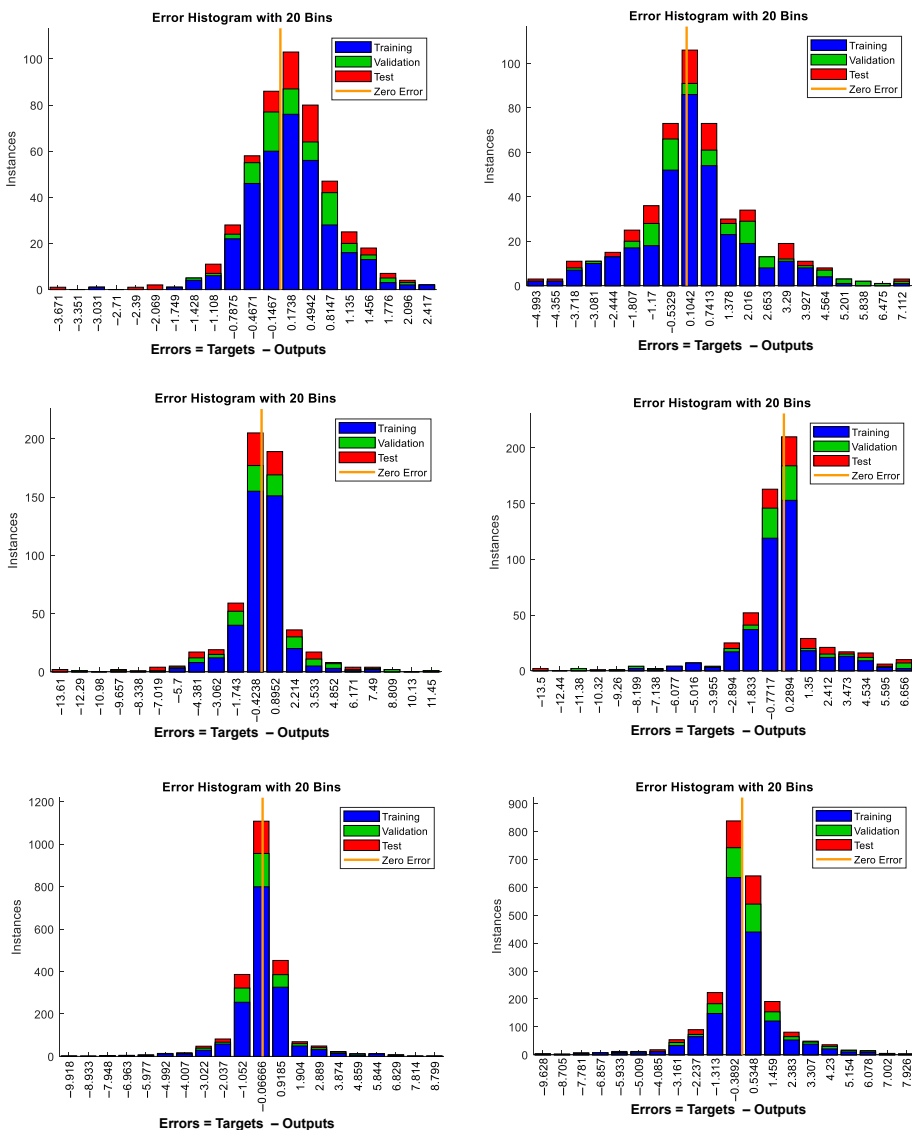


Figure 4.
Error histogram for
ANN models to get
zero error

(continued)

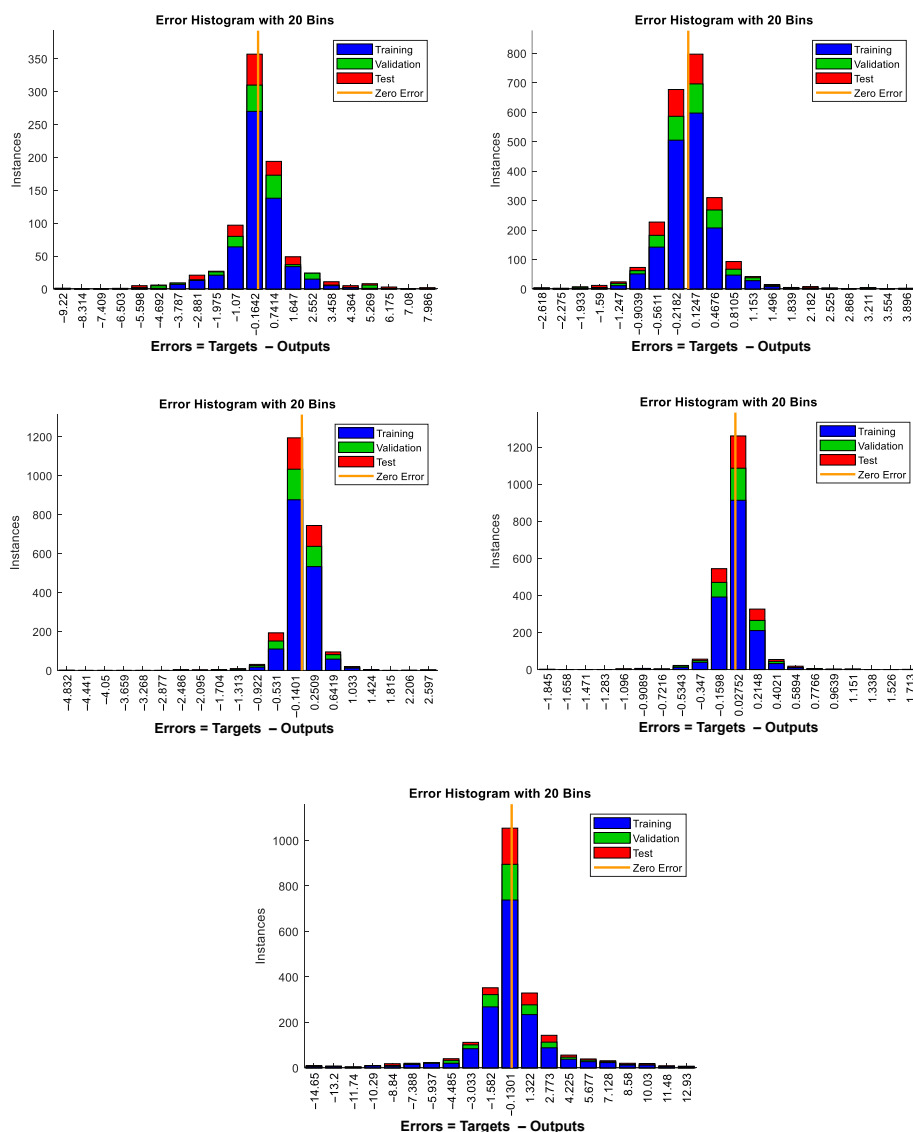


Figure 4

network. The input and output data were standardized between -1 and 1 before the model was trained, and the model was trained using the Bayesian regularization (BR) procedure. The outlier datasets produced good generating outcomes.

The R -value (Correlation coefficient) of 0.9909 and the MSE of 5.962 indicate that the model is accurate. The findings for all ANN model systems are shown in Table 5. Figure 3 depicts the App Designer model's distribution as well as the ANN model's anticipated values. Additionally, Figure 3 shows all R values are more than 0.98 which means compatibility

between ANN and App Designer models. Then as a result these models can be used for managing and controlling in Durugöl Advanced Biological WWTP. Figure 4 shows the error bars for all of the experimental data.

As shown in Figure 4 height of the bar in the bar plot means how many data points are near the bin value. Nonetheless, errors are close to zero meaning the output is close to the target value, and the NN models well. The error values are based on the bin number set up in the histogram function (i.e. histogram (data, 20)), it can be changed if you use a different bin value. You can try bin numbers with 50 or 100. Then the error values are the closest number to 0 among those bin numbers. Because an error equal to zero means the outputs are close to the target, model performance is good if the bin bars close to zero are taller than higher error numbers.

The ANN model was tested after it had been trained. When the test set estimates (BOD, TN, TP and TSS) were compared to the values obtained from the MATLAB App Designer model, it was discovered that the ANN estimates produced results that were very near to those observed. Figure 5 shows that the projected values are quite similar to the observed values, their trends are nearly identical, and the figure also indicates that the model can be used to manage the WWTP. In addition, Figure 6 depicts the ANN model's best validation performance. When the comparisons are made, it becomes clear that the ANN model and the MATLAB App Designer model produce extremely similar outcomes.

The ANN model was evaluated after training. It was found that the ANN estimates yielded outcomes that were extremely close to those observed when the test set estimates (BOD, TN, TP and TSS) were compared to the values acquired from the Ordu WWTP. The best validation performance of the ANN model was also shown in Figure 6. It is evident from the comparisons that the results produced by the ANN model and the observed values are comparable.

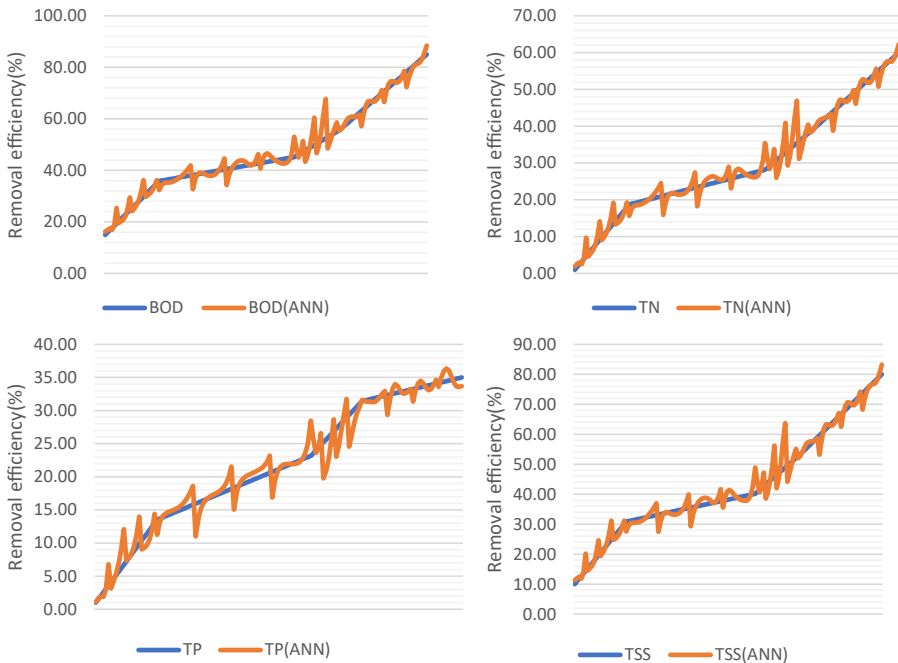


Figure 5.
Comparison between
the removal efficiency
of BOD, TN, TP and
TSS for the MATLAB
App Designer model
and the ANN model
(model no. 4 as an
example)

Conclusions

The complicated interactions between the wastewater influent parameters and BOD, TN, TP and TSS removal in the treated effluent were investigated using ANN models. The results in this paper show that ANN can be used to mimic the dynamic behavior of the WWTP. Not only can ANN properly forecast process (removal efficiency) behavior, but it can also provide insight into the dynamic behavior of a partially known WWTP system by allowing AI models to extract

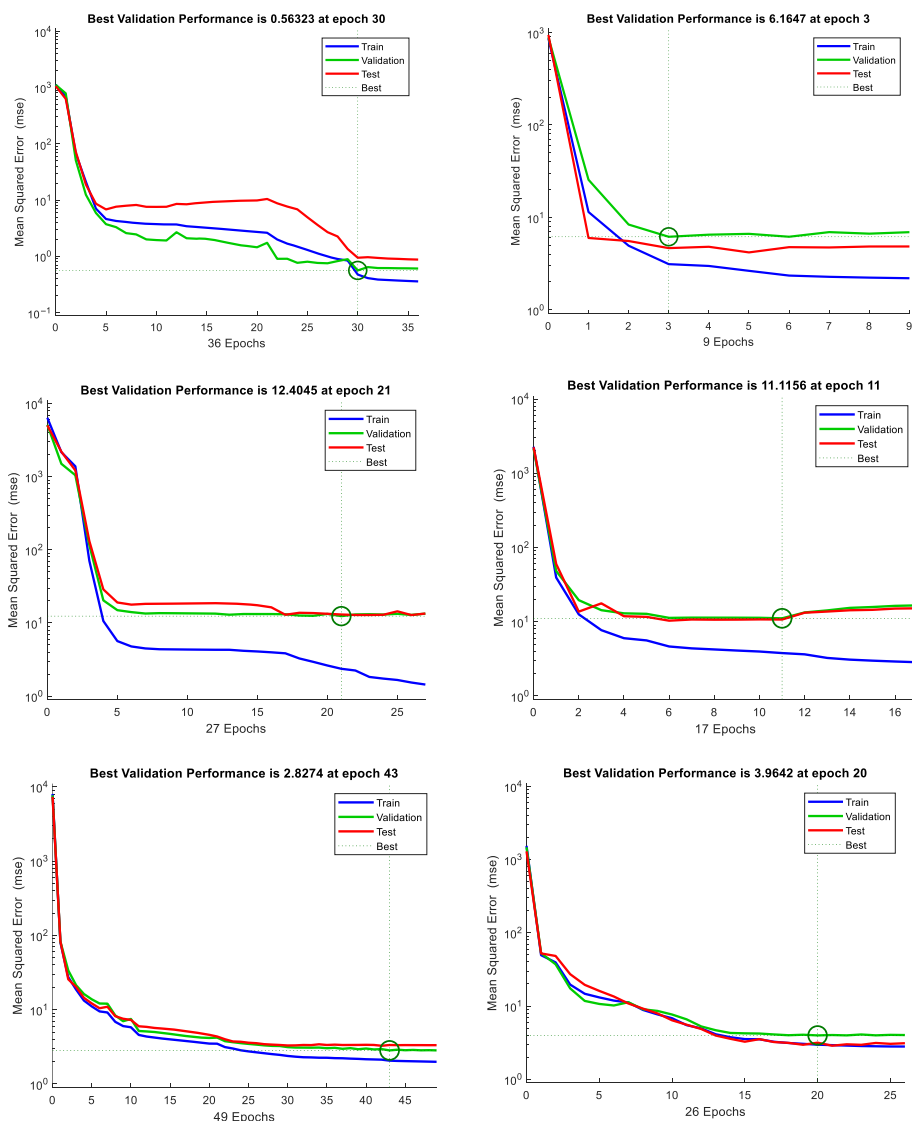


Figure 6.
The best validation
performance for the
ANN model

(continued)

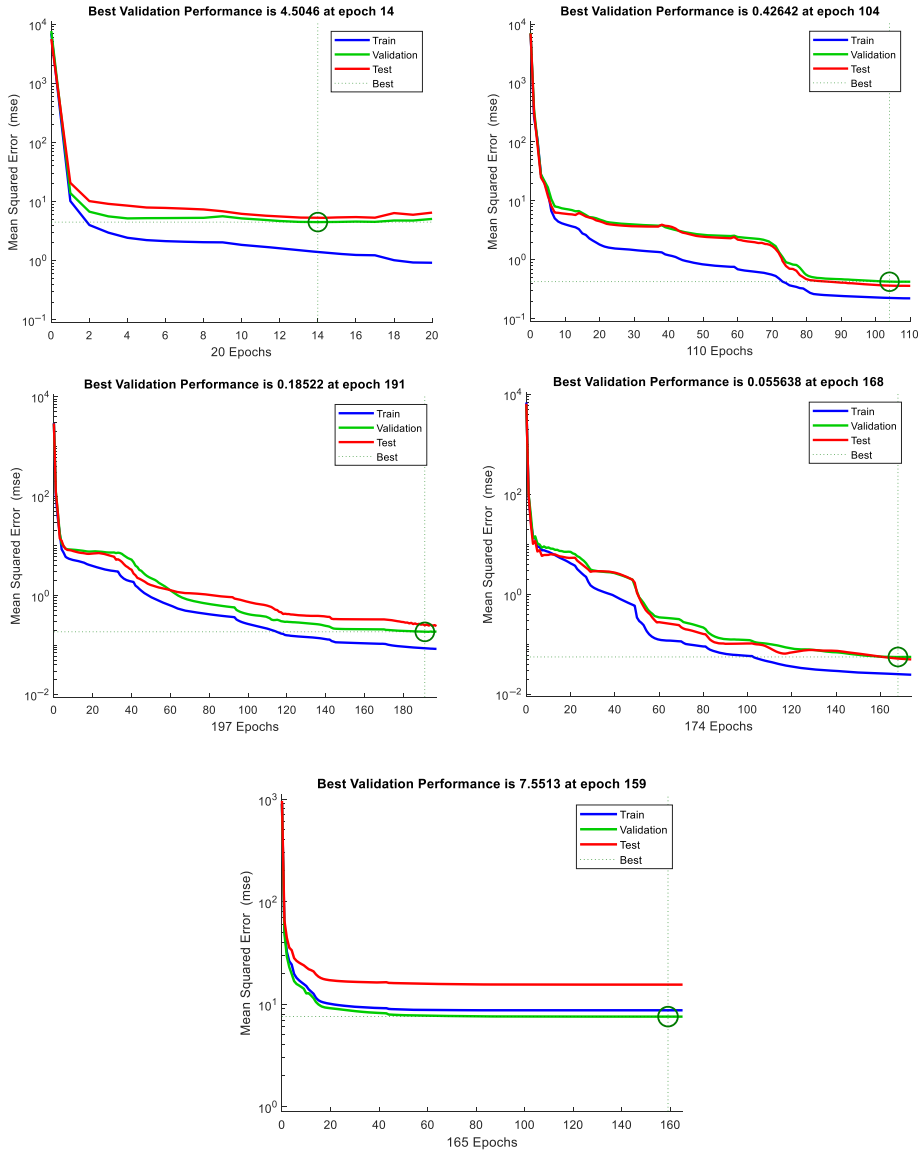


Figure 6

knowledge. Utilizing three input parameters, including temperature, dissolved oxygen and hydraulic retention time, the suggested model was trained, validated and tested. Two statistical measures are employed to assess the performance of the model: the correlation coefficient ($R^2 = 0.9909$) and mean square error ($MSE = 5.962$) indicate that the model is accurate. The results of the ANN model are distributed in a manner that is exactly consistent with the distribution of the observed data, demonstrating the effectiveness with which the ANN model can predict the removal effectiveness of the pollutants in primary and secondary treatment.

Additionally, the ANN technique is cost-effective, allowing us to construct prototype models swiftly and cheaply for the complicated industrial system. ANN also helps us construct correct models in less time, even if we only have a limited amount of experience and knowledge. The process expert can lead the ANN process to derive more sophisticated and accurate AI models by incorporating the expertise of the human expert and altering the functional sets. Future automatic real-time control systems for WWTP may use ANNs due to their capacity to forecast and quickly respond to changes in the status of dynamic processes.

References

- Agatonovic-Kustrin, S., & Beresford, R. (2000). Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of Pharmaceutical and Biomedical Analysis*, 22(5), 717–727. doi: [10.1016/S0731-7085\(99\)00272-1](https://doi.org/10.1016/S0731-7085(99)00272-1).
- Akdağ, R., & Karahan, M. (2014). *Yapay sinir ağları yöntemiyle diyarbakır ili kent merkezi içme suyu talep tahmini uygulaması* [Ph.D. thesis]. Dicle Üniversitesi Sosyal Bilimler Enstitüsü, Turkey.
- Akratos, C. S., Papaspyros, J. N. E., & Tsihrintzis, V. A. (2008). An artificial neural network model and design equations for BOD and COD removal prediction in horizontal subsurface flow constructed wetlands. *Chemical Engineering Journal*, 143(1-3), 96–110. doi: [10.1016/j.cej.2007.12.029](https://doi.org/10.1016/j.cej.2007.12.029).
- Cardoso, J. B., de Almeida, J. R., Dias, J. M., & Coelho, P. G. (2008). Structural reliability analysis using Monte Carlo simulation and neural networks. *Advances in Engineering Software*, 39(6), 505–513. doi: [10.1016/j.advengsoft.2007.03.015](https://doi.org/10.1016/j.advengsoft.2007.03.015).
- Chang, N. B., Chen, W. C., & Shieh, W. K. (2001). Optimal control of wastewater treatment plants via integrated neural network and genetic algorithms. *Civil Engineering and Environmental Systems*, 18(1), 1–17. doi: [10.1080/02630250108970290](https://doi.org/10.1080/02630250108970290).
- Chen, Y., Song, L., Liu, Y., Yang, L., & Li, D. (2020). A review of the artificial neural network models for water quality prediction. *Applied Sciences (Switzerland)*, 10(17). doi: [10.3390/app10175776](https://doi.org/10.3390/app10175776).
- Jana, D. K., Bhunia, P., Das Adhikary, S., & Bej, B. (2022). Optimization of effluents using artificial neural network and support vector regression in detergent industrial wastewater treatment. *Cleaner Chemical Engineering*, 3(June), 100039. doi: [10.1016/j.clce.2022.100039](https://doi.org/10.1016/j.clce.2022.100039).
- Khatri, N., Khatri, K. K., & Sharma, A. (2019). Prediction of effluent quality in ICEAS-sequential batch reactor using feedforward artificial neural network. *Water Science and Technology*, 80(2), 213–222. doi: [10.2166/wst.2019.257](https://doi.org/10.2166/wst.2019.257).
- Khatri, N., Khatri, K. K., & 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. *Journal of Water Process Engineering*, 37(March), 101477. doi: [10.1016/j.jwpe.2020.101477](https://doi.org/10.1016/j.jwpe.2020.101477).
- Küçükocaoglu, G., Keskin Benli, Y., & Kucuksozen, C. (2005). Detecting the manipulation of financial information by using artificial neural network models. *Istanbul Stock Exchange Review*, July 2000, 9(36), 1–23.
- Negnevitsky, M. (2005). Artificial intelligence A guide to intelligent systems. In *British Library Cataloguing* (2nd ed., Vol. 123). doi: [10.1016/j.poly.2016.11.012](https://doi.org/10.1016/j.poly.2016.11.012).
- Niu, G., Yi, X., Chen, C., Li, X., Han, D., Yan, B., . . . Ying, G. (2020). A novel effluent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment. *Journal of Cleaner Production*, 265, 121787. doi: [10.1016/j.jclepro.2020.121787](https://doi.org/10.1016/j.jclepro.2020.121787).
- Öztemel, E. (2008). Yapay sinir ağları. In *Papatya Yayıncılık*. Available from: http://papatyabilim.com.tr/PDF/yapay_sinir_aglari.pdf
- Pai, T. Y., Wan, T. J., Hsu, S. T., Chang, T. C., Tsai, Y. P., Lin, C. Y., . . . Yu, L. F. (2009). Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent. *Computers and Chemical Engineering*, 33(7), 1272–1278. doi: [10.1016/j.compchemeng.2009.02.004](https://doi.org/10.1016/j.compchemeng.2009.02.004).

- Qiao, J., Wang, G., Li, W., & Li, X. (2018). A deep belief network with PLSR for nonlinear system modeling. *Neural Networks*, 104, 68–79. doi: [10.1016/j.neunet.2017.10.006](https://doi.org/10.1016/j.neunet.2017.10.006).
- Sakiewicz, P., Piotrowski, K., Ober, J., & Karwot, J. (2020). Innovative artificial neural network approach for integrated biogas – wastewater treatment system modelling: Effect of plant operating parameters on process intensification. *Renewable and Sustainable Energy Reviews*, 124. doi: [10.1016/j.rser.2020.109784](https://doi.org/10.1016/j.rser.2020.109784).
- Shin, Y., Kim, Z., Yu, J., Kim, G., & Hwang, S. (2019). Development of NOx reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA). *Journal of Cleaner Production*, 232, 1418–1429. doi: [10.1016/j.jclepro.2019.05.276](https://doi.org/10.1016/j.jclepro.2019.05.276).
- Šiljić, A., Antanasijević, D., Perić-Grujić, A., Ristić, M., & Pocajt, V. (2014). Artificial neural network modelling of biological oxygen demand in rivers at the national level with input selection based on Monte Carlo simulations. *Science and Pollution Research*, 22(March 2015), 4230–4241. doi: [10.1007/s11356-014-3669-y](https://doi.org/10.1007/s11356-014-3669-y).
- Yurtoğlu, H. (2005). *Yapay Sinir Ağları Metodolojisi ile Öngörü Modellemesi: Bazı Makroekonomik Değişkenler için Türkiye Örneği* (p. 104). Ankara: DPT.

Corresponding author

Hussein Y.H. Alnajjar can be contacted at: hyhn1989@hotmail.com

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgroupublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com