Environ Res Tec, Vol. 5, Issue. 3, pp. 213–226, September 2022

Environmental Research and Technology
https://ert.yildiz.edu.tr - https://dergipark.org.tr/tr/pub/ert
DOI: https://doi.org/10.35208/ert.1106463

Research Article

Enhance modelling predicting for pollution removal in wastewater treatment plants by using an adaptive neuro-fuzzy inference system

Hussein Y. H. ALNAJJAR*, **Osman ÜÇÜNCÜ**

Department of Hydraulic, Karadeniz Technical University, Faculty of Civil Engineering, Trabzon, Türkiye

ARTICLE INFO

Article history
Received: 20 April 2022
Revised: 18 June 2022
Accepted: 15 July 2022

Key words:
ANFIS; Artificial neural networks; Biological oxygen demand; Biological treatment; Hydraulic retention time; Total suspended solids

ABSTRACT

Biological and physical treatment in wastewater treatment plants appears to be one of the most important variables in water quality management and planning. This crucial characteristic, on the other hand, is difficult to quantify and takes a long time to obtain precise results. Scientists have sought to devise several solutions to address these issues. Artificial intelligence models are one technique to monitor the pollutant parameters more consistently and economically at treatment plants and regulate these pollution elements during processing. This study proposes using an adaptive network-based fuzzy inference system (ANFIS) model to regulate primary and biological wastewater treatment and used it to model the nonlinear interactions between influent pollutant factors and effluent variables in a wastewater treatment facility. Models for the prediction of removal efficiency of biological oxygen demand (BOD), total nitrogen (TN), total phosphorus (TP), and total suspended solids (TSS) in a wastewater treatment plant were developed using ANFIS. Hydraulic retention time (HRT), temperature (T), and dissolved oxygen (DO) were input variables for BOD, TN, TP, and TSS models, as determined by linear correlation matrices between input and output variables. The findings reveal that the developed system is capable of accurately predicting and controlling outcomes. For BOD, TN, TP, and TSS, ANFIS was able to achieve minimum mean square errors of 0.1673, 0.0266, 0.0318, and 0.0523, respectively. The correlation coefficients for BOD, TN, TP, and TSS are all quite strong. In the wastewater treatment plant, ANFIS' prediction performance was satisfactory and the ANFIS model can be used to predict the efficiency of removing pollutants from wastewater.

Cite this article as: Alnajjar HYH, Üçüncü O. Enhance modelling predicting for pollution removal in wastewater treatment plants by using an adaptive neuro-fuzzy inference system. Environ Res Tec 2022;5:3:213–226.

INTRODUCTION

As the population grows and companies flourish, wastewater treatment becomes increasingly important due to the increased volume of wastewater generated by facilities each year. As a result, low-cost techniques that give accurate results for predicting treatment efficiency in wastewater treatment plants (WWTP) must be developed. WWTP entails a number of sophisticated and unpredictably unpredictable procedures. The treatment plant's smooth and effective operation, on the other hand, is dependent on a proper model capable of accurately representing the

***Corresponding author.**
*E-mail address: hyhn1989@hotmail.com

system's dynamic character. Previously, the majority of models were utilized in industrial wastewater treatment plants. WWTP operation includes physical, biological, and chemical features of wastewater streams, as well as biological and degrading mechanisms. Improved process control algorithms based on artificial intelligence (AI) technologies have received a lot of attention as a result of growing environmental and economic concerns [1].

According to the literature, suspended solids (SS_{eff}) and chemical oxygen demand (COD_{eff}) in the effluent from a hospital wastewater treatment facility were forecasted using three distinct adaptive neuro-fuzzy inference systems (ANFIS) and artificial neural networks (ANN) and in terms of effluent prediction, the results showed that ANFIS surpasses ANN statistically [2]. The ANFIS model is used to forecast effluent pH quality and artificial neural network is used as a comparison [3]. Another study used five process variables to predict the effluent chemical oxygen demand load from a full-scale expanded granular sludge bed reactor (EGSBR) treating corn processing wastewater, including influent chemical oxygen demand, influent flow rate, influent total Kjeldahl nitrogen, effluent volatile fatty acids, and effluent bicarbonate. The proposed ANFIS model was created using a hybrid learning approach, and its performance was assessed using a set of test data randomly selected from the experimental domain. The ANFIS-based predictions were validated using a variety of descriptive statistical metrics, including root-mean-square error, index of agreement, a factor of two, fractional variance, the proportion of systematic error, and so on [4]. Using daily data, feed-forward neural network (FFNN), support vector regression (SVR), and ANFIS black box artificial intelligence models (AI) were used to estimate effluent biological oxygen demand (BOD_{eff}) and chemical oxygen demand (COD_{eff}) of Tabriz wastewater treatment plant (WWTP). In addition, the BOD_{eff} and COD_{eff} parameters were predicted using the autoregressive integrated moving average (ARIMA) linear model to compare the linear and non-linear models' abilities in complicated process prediction [5]. In another research, the non-linear system of the activated sludge process in an industrial wastewater treatment plant was identified using the ANFIS and generalized linear model (GLM) regression. Predictive models of effluent chemical and 5-day biochemical oxygen demands were developed based on previously assessed inputs and outputs. From a list of possibilities, the least absolute shrinkage and selection operator (LASSO) and a fuzzy brute force search were employed to choose the best regressor combination for the GLMs and ANFIS models, respectively [6]. Furthermore, ANFIS allows direct inverse control of the substrate in an activated sludge system and the performance of the suggested controller is proven by tracking the substrate setpoints then the result of that reveal that the proposed controller can efficiently and precisely manage the substrate concentration level and the proposed inverse controller could be a beneficial control mechanism for the WWTP [7]. In another prior study support vector machine (SVM) and adaptive neuro-fuzzy inference system (ANFIS) models were used to evaluate the removal efficiency of Kjeldahl Nitrogen in a full-scale aerobic biological wastewater treatment facility and the input variables used in the modeling process include pH, COD, total solids (TS), free ammonia, ammonia nitrogen, and Kjeldahl nitrogen then the results of model development was provide an adaptable, functional, real-time, and alternate approach of replicating Kjeldahl nitrogen removal efficiency [8]. In another study, too, the use of successfully ANFIS modeling have employed to increase the output of anaerobic digesters [9]. The ANFIS model also was used to remove carbon and nitrogen. As a comparison, a feed-forward neural network is used. All of the variables investigated, including COD, suspended solids (SS), and ammonium nitrogen (NH4-N), were found to have increased prediction power using the ANFIS model [10].

According to prior research, most of them are focused on figuring out how to eliminate pollutants from wastewater during the biological treatment stage. Artificial intelligence models were used to analyze industrial and domestic wastewater treatment plants.

As a result, Biological and physical treatment in wastewater treatment plants appears to be one of the most important aspects of water quality management and planning. This crucial characteristic, on the other hand, is difficult to quantify and takes a long time to obtain precise results. Scientists have sought to devise a number of solutions to address these issues. Artificial intelligence models are one technique to monitor the pollutant parameters more consistently and economically at treatment plants and regulate these pollution elements during processing. Therefore, the fundamental goal of this research is to use the ANFIS model to apply, predict, and develop the pollutant removal efficiency for primary and biological treatment in WWTPs. This modeling study employed MATLAB APPDESIGNER model data for training, testing, and predictions. BOD, TN, TP, and TSS were the parameters investigated. Before running the simulation for prediction, the data was standardized. The output of the model was compared to actual training data and ANN data, and the error was minimized to produce the best operating points.

MATERIALS AND METHODS

Probabilistic reasoning, fuzzy logic, neural networks, and evolutionary computation are examples of intelligent technology. As can be observed, each of these technologies has its own set of benefits and drawbacks, and in many real-world applications, researchers will need to mix several intelligent technologies and learn from other sources. Hybrid intelligent systems have emerged as a result of the requirement for such a combination [11].

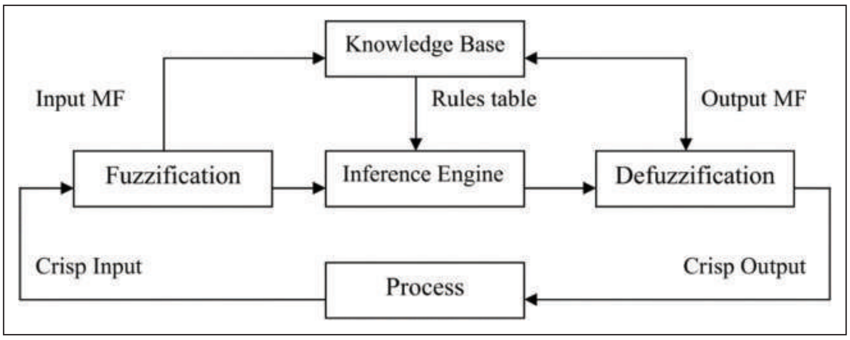

Figure 1. The fuzzy logic controller's basic structure [12].

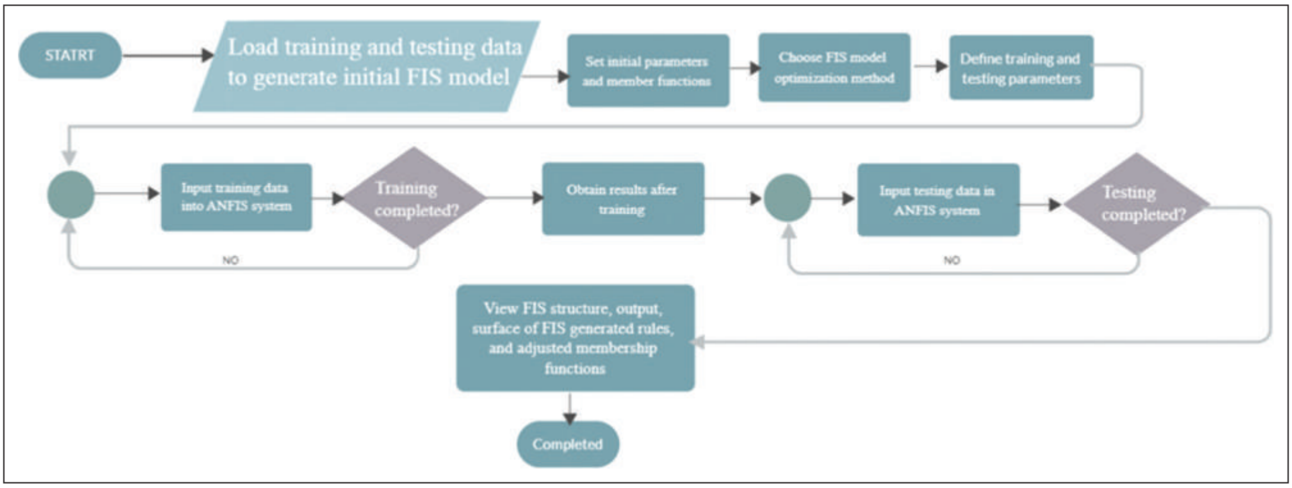

Figure 2. Flow chart of ANFIS test step [9].

The term "hybrid intelligent system" refers to a system that incorporates at least two intelligent technologies. Combining a neural network with a fuzzy system, for example, produces a hybrid neuro-fuzzy system.

Soft Computing (SC), an emerging technique to constructing hybrid intelligent systems capable of reasoning and learning in an uncertain and imprecise environment, is based on a combination of probabilistic reasoning, fuzzy logic, neural networks, and evolutionary computation.

Fuzzy Logic and Fuzzy Inference System

Fuzzification, fuzzy rule base, fuzzy output engine, and defuzzification are the four steps of a fuzzy system, as depicted in (Fig. 1) [12]. The input unit contains the input variables, as well as any information about the input variables that will affect the scenario under investigation [13]. The information with respect to the input variables is generally referred to as a database. The variables in the input can be numerical or textual [14]. Fuzzification is a method of assigning numerical values to linguistic adjectives and calculating the number of membership functions in fuzzy system sets. The fuzzy rule base is made up of all logical rules that connect the input and output variables, as well as any possible intermediary connections. The input variables are converted to their appropriate outputs by the fuzzy output engine. This is accomplished by considering the numerous relationships established in the fuzzy rule base. Finally, defuzzification is the process of converting the fuzzy system's language outputs into numerical values.

At the end of the information and fuzzy rule base interaction, the output unit generates variables. (Fig. 2) depicts the general ANFIS process for developing the ANFIS prediction model.

Model Architecture and Components

When a neural network is combined with a fuzzy system, a strong hybrid system capable of tackling complicated issues is created. This hybrid system's behavior may be defined in terms comparable to human rules, making it an accurate tool for simulating non-linear functions [15]. ANFIS employs a hybrid learning technique that specifies how the weights should be updated to reduce the error between the actual and desired output, adjusting the fuzzy inference system's parameters and structure in the process (FIS). (Fig. 3) depicts the structure of ANFIS, which is a Sugeno fuzzy model.

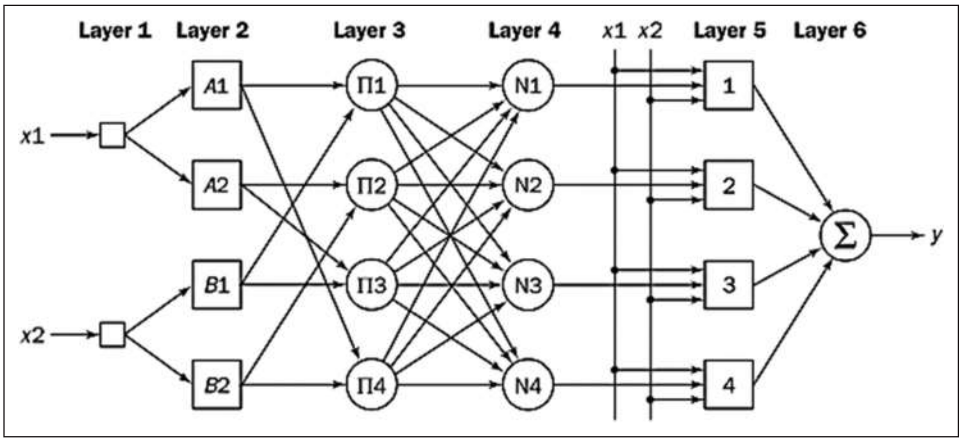

Figure 3. ANFIS structure.

ANFIS is an adaptive network that uses supervised learning on the learning method, similar to the Takagi-Sugeno fuzzy inference system [16]. Inputs and outputs, database and pre-processor, fuzzy system generator, fuzzy inference system, and adaptive neural network are the five major components of the model [17]. In most cases, the input and output parameters are chosen or derived from the system description parameters. The database and pre-processor are required for model creation and contain information about system performance. This information is normally gathered by collecting data on parameters that the system monitors on a regular basis. MATLAB is regarded a good tool for this study and is utilized to create system performance information.

A Sugeno fuzzy inference system and associated adaptive networks, as well as an ana adaptive network-based fuzzy inference system, are used (ANFIS). The input and output variables are chosen or generated from the variables that are typically used to describe the system. Model development necessitates the creation of a database containing system performance data. In most cases, it is created by gathering parameters from the APPDESIGNER model. For the model to produce accurate information on the system, the training database must be of high quality. The database must include sufficient and reliable information on the system for the model to accurately characterize it. A raw database, on the other hand, is likely to contain some duplicated and contradictory data. As a result, the raw training database may need to be pretreated to reduce duplicates and resolve data conflicts. Because the ANFIS is normally launched with a prototype fuzzy system, a fuzzy system generator is required. This function is provided by the software MATLAB (Matworks Inc.). Jang [17] utilized MATLAB to program the model, demonstrating that the language is adequate for model programming.

In order to achieve the lowest possible error, the model will be used to determine the relationship between the APPDESIGNER MATLAB model and the ANFIS model.

ANFIS is a multilayer feed-forward network that maps inputs into outputs with the use of neural network learning techniques and fuzzy reasoning. It's an adaptable neural network-based fuzzy inference system (FIS). The architecture of a typical ANFIS for the first order Sugeno fuzzy model, with two inputs, two rules, and one output (MFs). For a first order Sugeno fuzzy model [18], the following is an example of a rule set containing four fuzzy if–then rules:

Rule 1: If x is A1 and y is B1 then $f_1 = p_1x + q_1y + r_1$

Rule 2: If x is A2 and y is B2 then $f_2 = p_2x + q_2y + r_2$

where A1, A2, B1 and B2 are the MFs for the inputs x and y, respectively, p_{ij}, q_{ij} and r_{ij} (i,j =1,2) are consequent parameters [19].

The architecture of a typical ANFIS, as shown in Figure 3, consists of five levels, each of which performs a different function in the ANFIS and is described below.

Layer 1: This layer's nodes are all adaptive nodes. They assign membership scores to the inputs. This layer's outputs are determined by

$$O_{Ai}^1 = U_{Ai}(x) \quad i = 1,2$$
$$O_{Bj}^1 = U_{Bj}(x) \quad j = 1,2 \tag{1}$$

where x and y are crisp inputs, and Ai and Bj are fuzzy sets characterized by appropriate MFs, which could be triangular, trapezoidal, Gaussian function, or other shapes, and Ai and Bj are fuzzy sets characterized by appropriate MFs, which could be triangular, trapezoidal, Gaussian function, or other shapes. The generalized bell-shaped MFs (Eq. (2)) defined below are used in this investigation.

$$U_{Ai}(x) = \frac{1}{1 + (\frac{x - c_i}{a_i})^{2b_i}} \quad i = 1,2$$
$$U_{Bj}(x) = \frac{1}{1 + (\frac{x - c_j}{a_i})^{2b_j}} \quad j = 1,2 \tag{2}$$

where {a_i, b_i, c_i} and {a_j, b_j, c_j} are the parameters of the MFs, governing the bell-shaped functions. Parameters

Table 1. Modeled primary and secondary treatment techniques

System number	Processing	Inputs	Outbuts				Processing
			a	b	c	d	
ANFIS1	primary treatment	HRT, T	BOD	TN	TP	TSS	Mechanical screen + primary sedimentation tank
ANFIS 2	primary treatment	HRT, T	BOD	TN	TP	TSS	Mechanical screen + Grit removal + Grease trap + primary sedimentation tank
ANFIS 3	Secondary Treatment	HRT, T	BOD	TN	TP	TSS	Facultative pond+ Secondary sedimentation tank
ANFIS 4	Secondary Treatment	HRT, T	BOD	TN	TP	TSS	Anaerobic ponds + Facultative Pond+ Secondary sedimentation tank
ANFIS 5	Secondary Treatment	HRT, T, DO	BOD	TN	TP	TSS	Aerobic ponds (Partial Mixing) + Facultative Pond+ Secondary sedimentation tank
ANFIS 6	Secondary Treatment	HRT, T, DO	BOD	TN	TP	TSS	Aerobic ponds (Complete Mixing) + Facultative Pond+ Secondary sedimentation tank
ANFIS 7	Secondary Treatment	HRT, T	BOD	TN	TP	TSS	Anaerobic ponds+ Secondary sedimentation tank
ANFIS 8	Secondary Treatment	HRT, T, DO	BOD	TN	TP	TSS	Aerobic ponds (Partial Mixing) + Secondary sedimentation tank
ANFIS 9	Secondary Treatment	HRT, T, DO	BOD	TN	TP	TSS	Aerobic ponds (Complete Mixing) + Secondary sedimentation tank
ANFIS 10	Secondary Treatment	HRT, T, DO	BOD	TN	TP	TSS	Anaerobic ponds + Aerobic ponds (Partial Mixing) + Secondary sedimentation tank
ANFIS 11	Secondary Treatment	HRT, T, DO	BOD	TN	TP	TSS	Anaerobic ponds + Aerobic ponds (Complete Mixing) + Secondary sedimentation tank

in this layer are referred to as premise parameters or antecedent parameters.

Layer 2: The nodes in this layer are fixed nodes with the number 2 next to them, indicating that they act as a simple multiplier. This layer's outputs are expressed as

$$O_{ij}^{2} = w_{ij} = U_{Ai}(x)U_{Bj}(y), \;\; i,j = 1,2 \tag{3}$$

which represents the firing strength of each rule. The degree to which the antecedent element of the rule is satisfied is referred to as the firing strength.

Layer 3: The nodes in this layer are also fixed nodes with the label, indicating that they play a role in network normalization. This layer's outputs can be expressed as

$$O_{ij}^{3} = \overline{w_{ij}} = \frac{w_{ij}}{w_{11}+w_{12}+w_{21}+w_{22}}, \quad i,j = 1,2 \tag{4}$$

which are called normalized firing strengths.

Layer 4: The output of each node in this layer is just the product of the normalized firing strength and a first-order polynomial (for a first order Sugeno model). As a result, Eq. (5) gives the outputs of this layer.

$$O_{ij}^{4} = \overline{w_{ij}}\, f_{ij} = \overline{w_{ij}}\left(p_{ij} + q_{ij}y + r_{ij}\right), \quad i,j = 1,2 \tag{5}$$

Subsequent parameters refer to the parameters in this layer.

Layer 5: This layer's single node is a fixed node labelled ∑ that computes the total output as the sum of all incoming signals, i.e.,

$$z = O_1^5 = \sum_{i=1}^{2}\sum_{j=1}^{2}\overline{w_{ij}}\, f_{ij} = \sum_{i=1}^{2}\sum_{j=1}^{2}\overline{w_{ij}}\left(p_{ij}x + q_{ij}y + r_{ij}\right)$$
$$= \sum\sum(\overline{w_{ij}}\,x)p_{ij} + (\overline{w_{ij}}y)q_{ij} + (\overline{w_{ij}}\,)r_{ij} \tag{6}$$

When the values of the premise parameters are fixed, the result is a linear combination of the subsequent parameters. The ANFIS design may be seen to have two adaptive layers: Layers 1 and 4. to the input MFs. Layer 4 has modifiable parameters $\{p_{ij}, q_{ij}, r_{ij}\}$ Layer 1 has modifiable parameters $\{a_i, b_i, c_i\}$ and $\{a_j, b_j, c_j\}$ related pertaining to the first-order polynomial. The learning algorithm for this ANFIS architecture's task is to tune all the changeable parameters to match the training data in the ANFIS output. The hybrid learning algorithm is a two-step procedure for learning or altering certain adjustable parameters. The premise parameters are held constant in the forward pass of the hybrid learning algorithm, node outputs advance till Layer 4, and the subsequent parameters are determined using the least squares approach. The subsequent parameters are held constant in the backward pass, the error signals flow backward, and the premise parameters are updated using the gradient descent algorithm. Jang provides a detailed algorithm and mathematical basis for the hybrid learning approach [18].

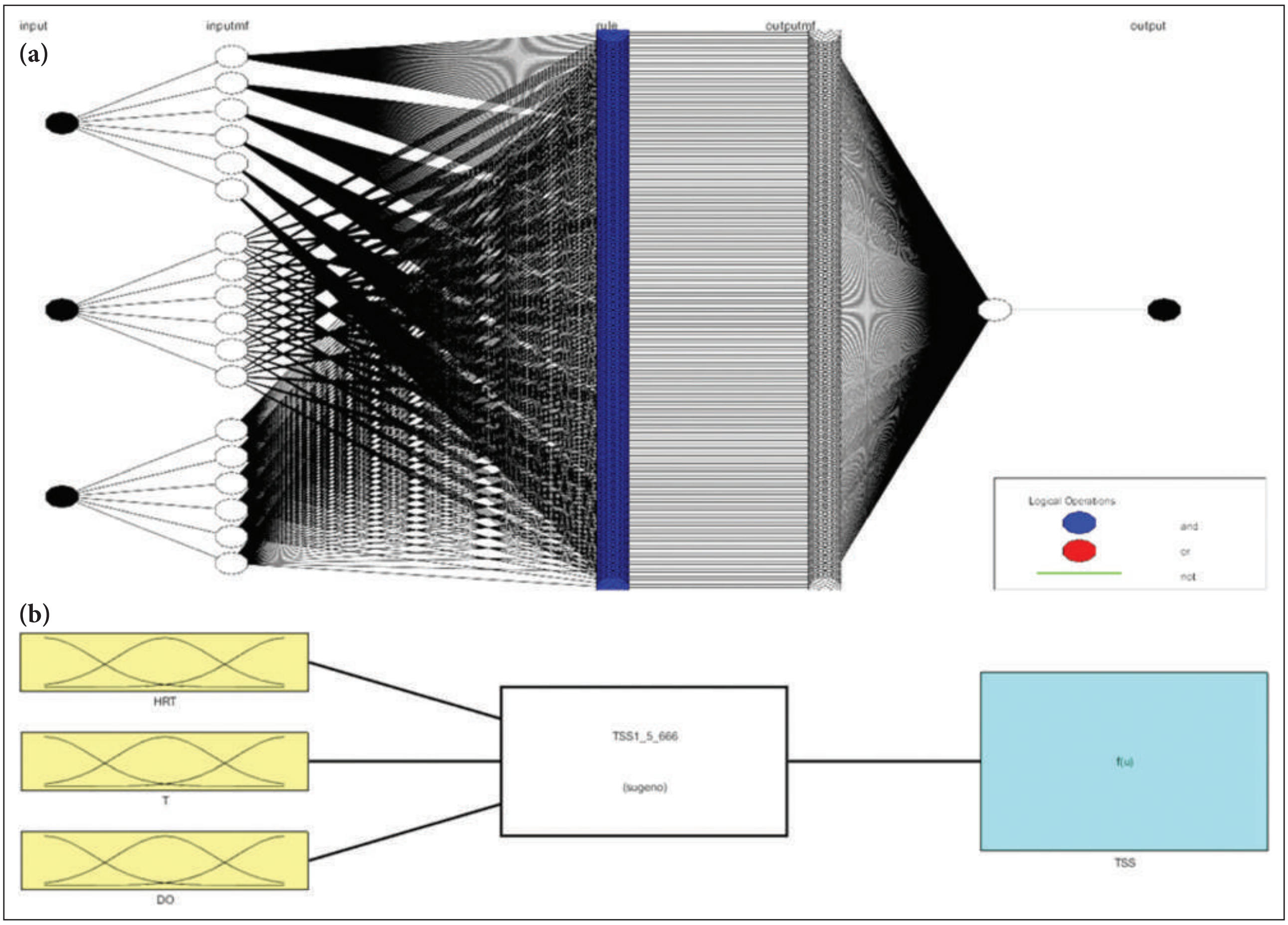

Figure 4. Schematic diagram of **(a)** ANFIS models with all input variables and **(b)** input–output mapping structure of ANFIS models with input variables.

To evaluate the prediction power of ANFIS and ANN trained by each data set, performance indices such as mean square error (MSE), root mean square normalized error (RMSE), mean absolute percentage error (MAPE), and correlation coefficient (R) are utilized. The MSE performance index was established as follows:

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

The RMSE performance index was defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}-y)^2}{n}} \quad (8)$$

where y is the measured values, $\hat{y}$ the corresponding predicted values and n is the number of samples.

Mean absolute percentage error (MAPE):

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_t-F_t}{A_t}\right| \times 100 \quad (9)$$

Where $\bar{A} = \frac{1}{N}\sum_{t=1}^{N} A_t$ and $\bar{F} = \frac{1}{N}\sum_{t=1}^{N} F_t$ are the average values of A_t and F_t over the training or testing dataset. The smaller RMSE and MAPE mean better performance.

Correlation coefficient (R):

$$R = \frac{\sum_{t=1}^{n}(A_t-\bar{A})(F_t-\bar{F})}{\sqrt{\sum_{t=1}^{n}(A_t-\bar{A})^2 . \sum_{t=1}^{n}(F_t-\bar{F})^2}} \quad (10)$$

The Plant Description

The ANFIS model was tested as an artificial intelligence model to operate a MATLAB-modeled wastewater treatment system.

The wastewater treatment plan's major processing techniques were several types of primary treatment (two models) and secondary treatment (nine models).

Mechanical screen and primary sedimentation tank are one sort of primary treatment procedure, while mechanical screen, grit removal, grease trap, and primary sedimentation tank are the other.

In terms of secondary treatment, nine different models were compared and clarified in Table 1. The inputs to the ANFIS1 (a) model, for example, are HRT and T, and the output is BOD.

The influent variables include hydraulic retention time (HRT), temperature (T) and dissolved oxygen (DO) and the effluent variables include the removal efficiency of bio-

Table 2. The detailed information of the models

Model	Training data	Testing data	Number of input MF for the following outputs				Rules for the following outputs			
			BOD	TN	TP	TSS	BOD	TN	TP	TSS
ANFIS1	120	18	46	86	86	84	24	48	48	32
ANFIS2	120	18	86	108	106	106	48	80	60	60
ANFIS3	144	25	88	1010	99	1010	64	100	81	100
ANFIS4	144	22	106	88	66	98	60	64	36	72
ANFIS5	576	84	866	888	1086	666	288	512	480	216
ANFIS6	576	93	686	868	888	1086	288	384	512	480
ANFIS7	205	41	86	88	108	88	48	64	80	64
ANFIS8	576	107	688	787	677	856	384	392	294	240
ANFIS9	576	107	865	888	777	699	240	512	343	486
ANFIS10	576	105	878	988	789	567	448	576	504	210
ANFIS11	576	98	668	985	867	644	288	360	336	96

logical oxygen demand (BOD5), total nitrogen (TN), total phosphorous (TP) and total suspended solids (TSS).

Model Implementation

The ANFIS (Adaptive Neuro-Fuzzy Inference System) editor of the Fuzzy toolbox in MATLAB was used to create a model in Sugeno structure (R2021 version, The MathWorks Inc., USA). The membership functions were extracted from the APPDESIGNER system's data set, which had been standardized and divided into training and testing data. The model's parameters were estimated using a hybrid learning method, and the model was validated using APPDESIGNER model data effluent parameters like output BOD, TN, TP, and TSS.

Figure 4a, b shows the topology of the ANFIS network that was employed. In the creation of a fuzzy system, eleven ANFIS structures with varying input correlation (Fig. 4a) and consisted of five layers were established (Fig. 4b). The following are the meanings of each layer in (Fig. 4b), as well as their counterpart in the ANFIS structures:

Input layer: In the ANFIS inputs layer, state variables are nodes: There are three input variables in total: HRT, T, and DO are all acronyms for hormone replacement therapy (from the influent) Layer with the membership function: Each state variable's term sets are nodes in the ANFIS values layer, which compute the membership value.

For each input variable:

Membership: triangle mf or gauss mf Membership number.

Rules layer: Each rule in the fuzzy class is a node in the ANFIS rules layer, with the rule matching factor xi computed using soft-min or product. Layer of the output membership function: In the function layer, each weighs the result of its linear regression fi, resulting in the rule output.

Table 3. Excel data from APPDESIGNER model

	Inputs		Outputs			
HRT	T	DO	BOD	TN	TP	TSS
30.00	-5.00	0.10	5.00	1.00	1.00	50.00
37.39	-4.00	0.31	5.65	1.17	1.11	50.43
44.78	-3.00	0.53	6.30	1.35	1.22	50.87
52.17	-2.00	0.74	6.96	1.52	1.33	51.30
59.57	-1.00	0.95	7.61	1.70	1.43	51.74
66.96	0.00	1.17	8.26	1.87	1.54	52.17
30.00	-5.00	1.38	8.91	2.04	1.65	52.61
37.39	-4.00	1.59	9.57	2.22	1.76	53.04
44.78	-3.00	1.80	10.22	2.39	1.87	53.48
52.17	-2.00	2.02	10.87	2.57	1.98	53.91
59.57	-1.00	2.23	11.52	2.74	2.09	54.35
66.96	0.00	2.44	12.17	2.91	2.20	54.78
30.00	-5.00	2.66	12.83	3.09	2.30	55.22
37.39	-4.00	2.87	13.48	3.26	2.41	55.65
44.78	-3.00	3.08	14.13	3.43	2.52	56.09
52.17	-2.00	3.30	14.78	3.61	2.63	56.52
59.57	-1.00	3.51	15.43	3.78	2.74	56.96
66.96	0.00	3.72	16.09	3.96	2.85	57.39
30.00	-5.00	3.93	16.74	4.13	2.96	57.83
37.39	-4.00	4.15	17.39	4.30	3.07	58.26
44.78	-3.00	4.36	18.04	4.48	3.17	58.70
52.17	-2.00	4.57	18.70	4.65	3.28	59.13
59.57	-1.00	4.79	19.35	4.83	3.39	59.57
66.96	0.00	5.00	20.00	5.00	3.50	60.00
74.35	-5.00	0.10	20.65	5.17	3.61	60.43
81.74	-4.00	0.31	21.08	5.30	3.70	60.85

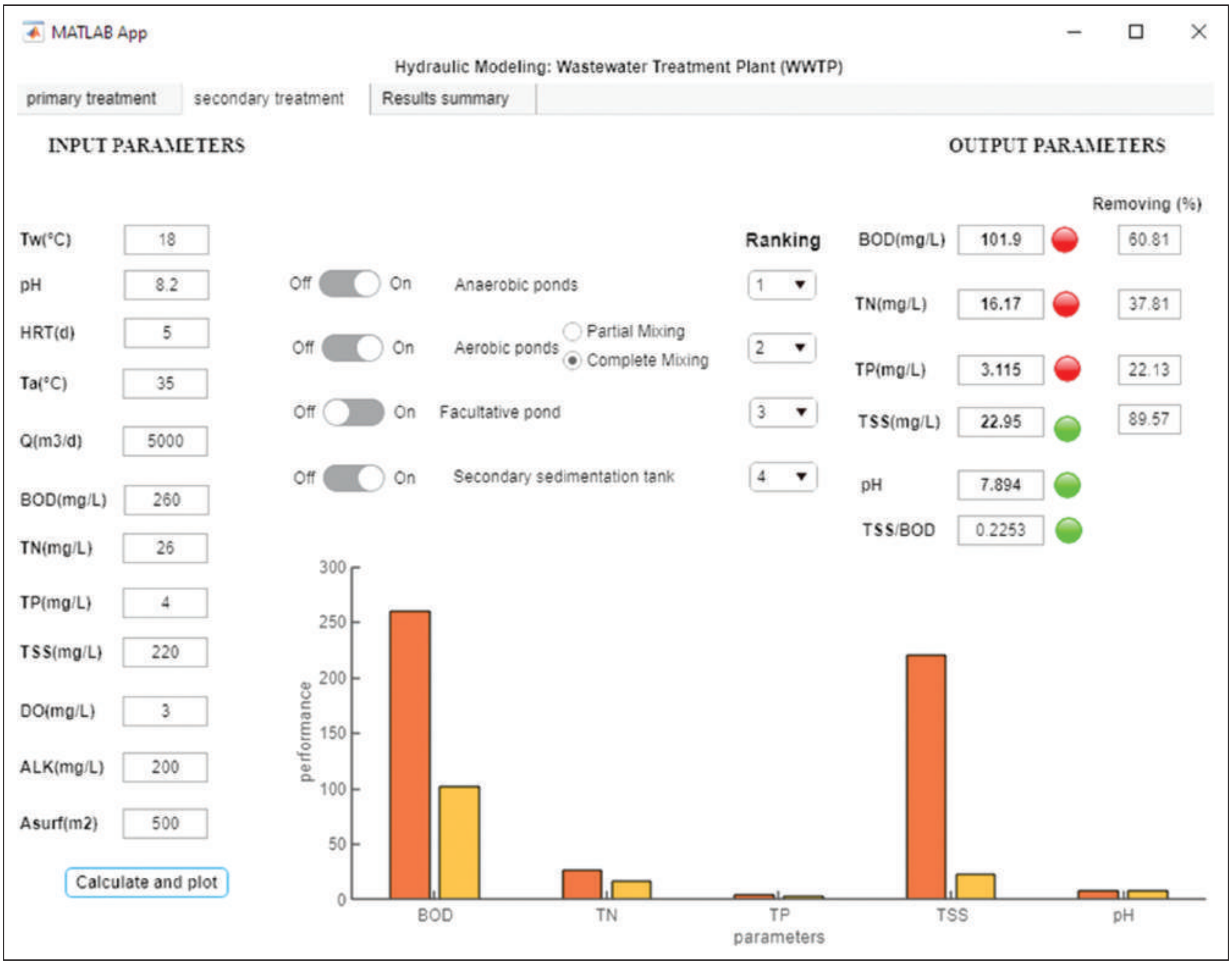

Figure 5. APPDESIGNER model interface.

Normalization layer:

Each xi is scaled into the normalization layer Normalization

Normalization is performed with the equation:

$$x_{norm} = (x_{value} - x_{min})(x_{max} - x_{min}) \quad (11)$$

Output layer: Each rule output is added to the output layer.

Outputs: BOD, TN, TP and TSS (effluent).

Results and Discussion

As shown in (Fig. 5), the data from the APPDESIGNER model was used to create eleven different ANFIS models. As an example of data, the data generated from the APPDESIGNER model has been organized in tables for usage in ANFIS, as seen in Table 3. The effluent BOD, TN, TP, and TSS were monitored in the system as indicators of treatment performance and stability using the ANFIS models applied to the APPDESIGNER model.

The ANFIS model in this paper was created using Matlab's fuzzy function. The data was first examined using Matlab's fuzzy subtractive clustering tool, and the cluster centers were determined. The initializing parameters were determined using the cluster centers, which indicate the initial value of premise parameters.

To identify a suitable ANFIS model, the types and numbers of MFs in ANFIS were investigated, including Gaussian, generalized bell-shaped, triangular, and trapezoidal-shaped functions, as well as the parameters. The values of RMSE and R between the model output values and observed values were used as selection criteria for the optimal final architecture. All ANFIS models with generalized bell-shaped MFs for each input variable showed the best results with diverse input variables. BOD, TN, TP, and TSS were all predicted using these models. Thus, monitoring the BOD, TN, TP, and TSS dynamics for the l wastewater treatment process, which was optimized by trial and error during the training phase, was adequate. The hybrid approach was used to train the network after selecting the initial value of the premise parameter and the design of the predictive model. The network's prem-

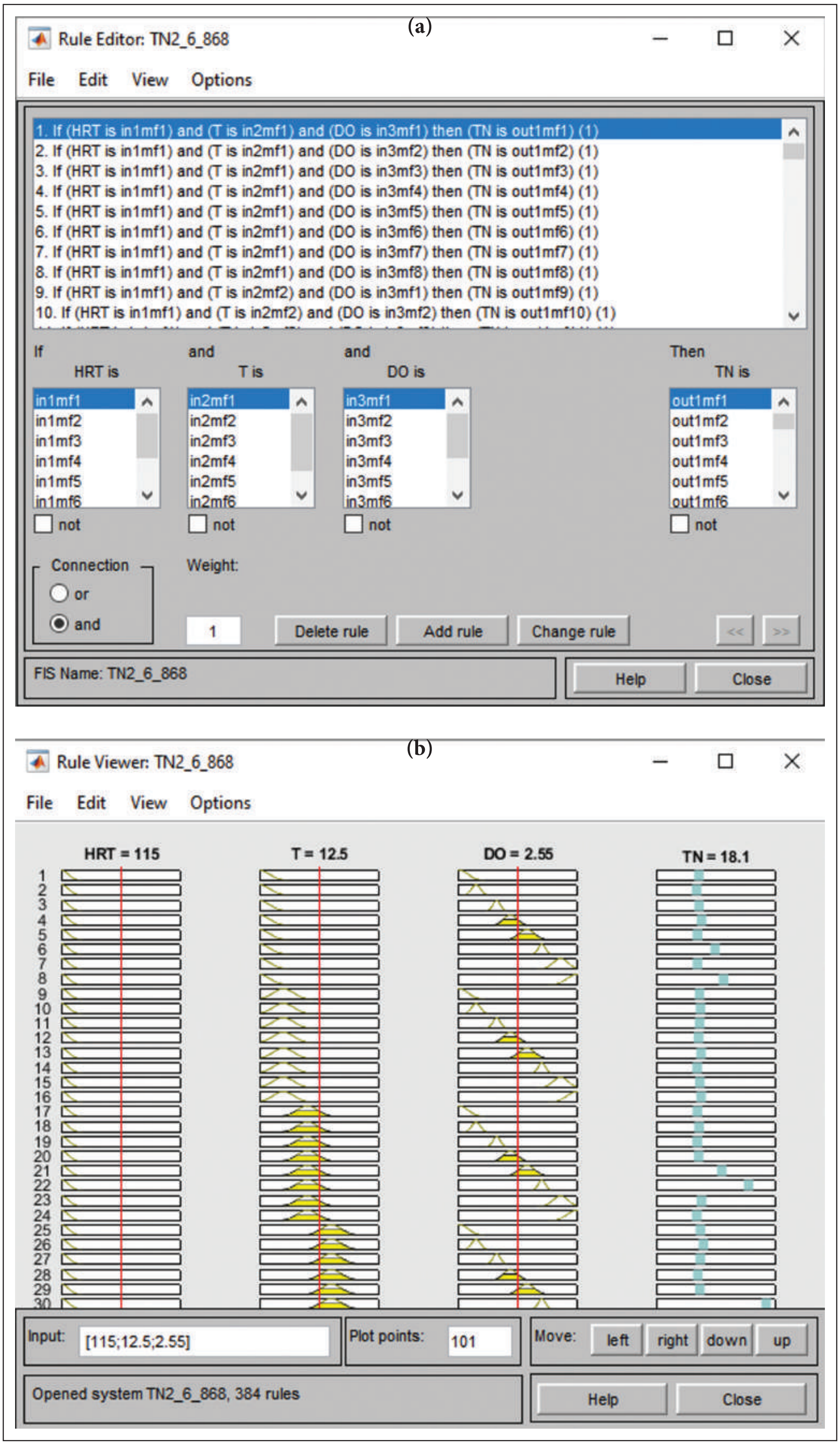

Figure 6. **(a)** Rule editor of Matlab fuzzy logic toolbox **(b)**Rule viewer screen to obtain defuzzified.

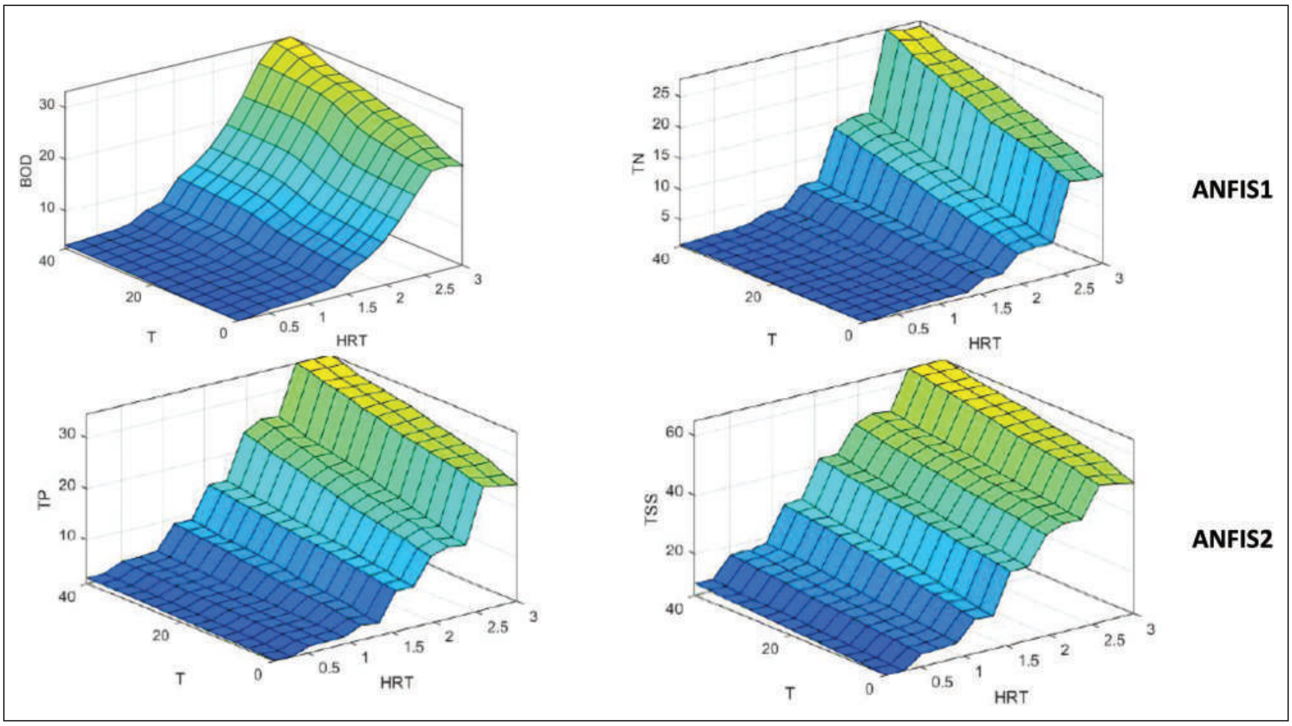

Figure 7. 3D response surface graph.

ise and associated parameters were then trimmed. After obtaining the premise parameter, membership functions for the variables were drawn.

Following the training of the model, inference was done using fuzzy language rules (Fig. 6a). After the network had been trained, those rules were obtained. In terms of comparing output values to input values, several additional heuristic criteria were also introduced. Defuzzified findings and graphical outputs can also be generated. (Fig. 7) shows an example of a Surface Viewer screen generated by the Fuzzy Logic Toolbox. Variable outcomes can be plotted and compared in two or three dimensions. According to the mass center of variables, (Fig. 6b) displays the outcomes of applied rules and their related outputs. Defuzzified values for output variables can be determined manually using the interface by changing input values. The Rule Viewer can produce a variety of output values depending on the input data. Using the interface to acquire defuzzified output values for all of the genuine input values is not flexible. As a result, a program using Matlab codes is built to drive defuzzified output outcomes in line with real-world input values.

RESULTS

The influence of the ANFIS model inputs (temperature, dissolved oxygen, and hydraulic retention time) on the model outputs is also shown in Figure 7 (BOD, TN, TP and TSS). The ANFIS 1 model, for example, indicates that raising the temperature and lengthening the hydraulic retention time improves the efficiency of pollutant removal in wastewater.

All R-square and RMSE values for the removal efficiency of BOD, TN, TP and TSS are also shown in Table 4. When training, R value was 0.9782 using ANFIS but when validating, R value was 0.9888 using ANFIS.

When training and validating, the RMSE values for ANFIS2 was 0.28 for BOD and 0.0266 using ANFIS was lower than that of 1.7289 and 1.6172 using ANN. The RMSE value of 0.0318 using ANFIS was also lower than that of 1.7398 using ANN when predicting for TN. Figure 8 show the training and predicting results using ANFIS and APDESIGNER model.

The architecture of ANFIS combines ANN and fuzzy logic, as well as linguistic expressions of MFs and if–then rules, to overcome the limitations of traditional neural networks, such as the risk of becoming trapped in a local minimum and model architecture selection, and to improve predicting performance. As a result, ANFIS is an excellent alternative for simulating wastewater treatment performance. Furthermore, ANN is a black box in nature, with difficult to interpret links between inputs and outputs, whereas ANFIS is clear, with simple to understand and interpret if–then rules. The ANFIS model's prediction performance in the wastewater treatment plant was excellent, and the ANFIS model may be used to estimate the efficiency of eliminating contaminants

Table 4. Determination of the appropriate ANFIS, APPDESIGNER and ANN models

System number	RMSE				R-square value(APPDESIGNER-ANFIS)							
					BOD		TN		TP		TSS	
	BOD	TN	TP	TSS	Training	Testing	Training	Testing	Training	Testing	Training	Testing
ANFIS1	1.3838	0.0995	0.2020	0.3414	0.9782	0.9888	0.9998	0.9999	0.9995	0.9999	0.9996	0.9996
ANN1	1.0476	0.9313	1.0402	2.1390								
ANFIS2	0.2800	0.0266	0.0318	0.1009	0.9992	0.9995	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
ANN2	1.7289	1.6172	1.7398	2.8112								
ANFIS3	1.2242	0.3460	0.4186	0.6890	0.9968	0.9992	0.9995	0.9998	0.998	0.9967	0.9984	0.9998
ANN3	3.1565	2.3511	1.4402	2.9541								
ANFIS4	1.2582	1.1756	1.1975	1.8257	0.9944	0.9957	0.9935	0.9828	0.9837	0.9906	0.9886	0.9721
ANN4	2.8542	2.3733	1.7433	2.8592								
ANFIS5	1.4918	0.1276	0.1784	0.9143	0.9925	0.9911	0.9998	0.9992	0.9996	0.9984	0.9898	0.9907
ANN5	0.9486	0.5106	0.6192	0.6485								
ANFIS6	2.4545	0.4452	0.1838	0.1289	0.9777	0.972	0.9967	0.9978	0.9996	0.9991	0.9998	0.9995
ANN6	2.2790	1.2215	1.7855	1.4524								
ANFIS7	1.5691	1.3102	0.4659	0.1036	0.9947	0.9973	0.9948	0.996	0.9979	0.998	0.9999	0.9999
ANN7	2.4422	2.0268	1.2311	0.6645								
ANFIS8	0.9822	0.8183	1.0491	0.0650	0.9979	0.9977	0.9901	0.9854	0.9879	0.9873	0.9999	0.9999
ANN8	3.5613	1.3684	2.0934	0.4360								
ANFIS9	1.9230	0.1842	0.6866	0.0326	0.9919	0.9899	0.9995	0.9994	0.9948	0.9953	0.9999	0.9999
ANN9	3.0028	1.1767	1.8458	0.3803								
ANFIS10	0.8467	0.4560	0.1174	0.0523	0.9984	0.998	0.9991	0.9968	0.9998	0.9998	0.9999	0.9999
ANN10	0.1673	0.1656	0.1835	0.0722								
ANFIS11	1.3892	1.3254	0.8686	0.1254	0.996	0.994	0.9942	0.9926	0.9943	0.9947	0.9997	0.9995
ANN11	3.9526	3.5289	2.5361	1.0930								

from wastewater. As a result, based on the artificial intelligence model, it is possible to build physical and biological treatment units in wastewater treatment plants, reducing the high costs and time necessary for wastewater treatment plant design.

CONCLUSIONS

Eleven models based on adaptive neuro-fuzzy inference system (ANFIS) were constructed in this research to predict biological oxygen demand (BOD), total nitrogen (TN), total phosphorous (TP), and total suspended solids (TSS) removal efficiency for a primary and biological wastewater treatment process. The developed models were trained and tested using data from the APPDESIGNER Matlab model for BOD, TN, TP, and TSS. For comparison, the ANN was also used. The neural network models generated good estimations for the BOD, TN, TP, and TSS data sets, which span a wide range of data for training and testing. ANFIS was able to anticipate the variation in removal efficiency based on the findings. Minimum root means square errors (RMSEs) of 0.1673, 0.0266, 0.0318, and 0.0523 were also attained for BOD, TN, TP, and TSS, respectively. For the BOD, TN, TP, and TSS data sets, which span a wide range of data for training and testing, the neural network models generated good estimations. Overall, the results showed that the simulated removal efficiency of BOD, TN, TP, and TSS closely matched observed concentrations, as seen by the low RMSE and very high R values. Given the high level of complexity in the wastewater treatment process, the significant amount of variable information dispersed over the dataset, and the wide concentration ranges, ANFIS models' excellent prediction results for both effluent parameters are particularly relevant. As a result, the ANFIS modelling approach could serve as a generic foundation for modelling different treatment procedures. Furthermore, the ANFIS modelling approach could be used to anticipate and control the performance of treatment processes in treatment plants.

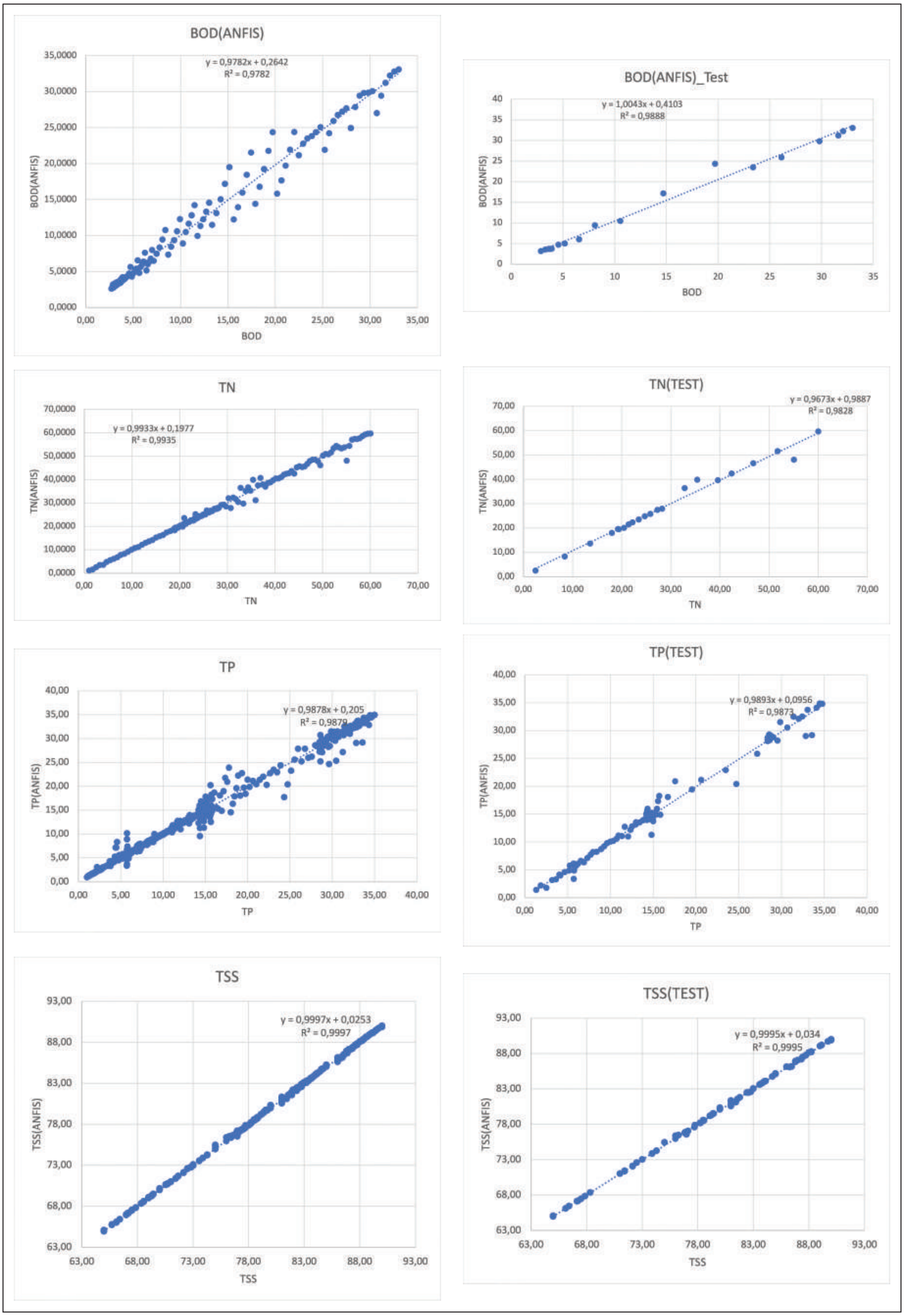

Figure 8. Prediction results of BOD, TN, TP and TSS for ANFIS model.

The proposed ANFIS outperformed Artificial neural network (ANN) in terms of performance and generalization ability. The RMSE and R^2 values for forecasting the removal efficiency of BOD, TN, TN, and TSS using ANFIS were considerably improved. Overall, the findings suggest that ANFS can be used to predict pollutant removal mechanisms in wastewater treatment systems.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

REFERENCES

[1] H. Y. H. Alnajjar and O. Üçüncü, "Using of a fuzzy logic as one of the artificial intelligence models to increase the efficiency of the biological treatment ponds in wastewater treatment plants," International Journal of Environmental Pollution and Environmental Modelling, Vol. 4(2), pp. 85–94, 2021.

[2] T. Y. Pai, T. J. Wan, S. T. Hsu, T. C. Chang, Y. P. Tsai, C. Y. Lin, H. C. Su, and L. F. Yu, "Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent," Computers & Chemical Engineering, Vol. 33(7), pp. 1272–1278, 2009. [CrossRef]

[3] M. S. Gaya, N. A. Wahab, Y. M. Sam, and S. I. Samsuddin, "ANFIS based effluent pH quality prediction model for an activated sludge process," Advanced Materials Research, Vol. 845, pp. 538–542, 2014. [CrossRef]

[4] K. Yetilmezsoy, H. Ozgun, R. K. Dereli, M. E. Ersahin, and I. Ozturk, "Adaptive neuro-fuzzy inference-based modeling of a full-scale expanded granular sludge bed reactor treating corn processing wastewater," Journal of Intelligent & Fuzzy Systems, Vol. 28(4), pp. 1601–1616, 2015. [CrossRef]

[5] V. Nourani, P. Asghari, and E. Sharghi, "Artificial intelligence based ensemble modeling of wastewater treatment plant using jittered data," Journal of Cleaner Production, Vol. 291, Article 125772, 2021. [CrossRef]

[6] D. O. Araromi, O. T. Majekodunmi, J. A. Adeniran, and T. O. Salawudeen, "Modeling of an activated sludge process for effluent prediction—a comparative study using ANFIS and GLM regression," Environmental Monitoring and Assessment, Vol. 190(9), 2018. [CrossRef]

[7] M. S. Gaya, N. Abdul Wahab, Y. M. Sam, S. I. Samsudin, and I. W. Jamaludin, "ANFIS direct inverse control of substrate in an activated sludge wastewater treatment system," Applied Mechanics and Materials, Vol. 554, pp. 246–250, 2014. [CrossRef]

[8] D. S. Manu and A. K. Thalla, "Artificial intelligence models for predicting the performance of biological wastewater treatment plant in the removal of Kjeldahl Nitrogen from wastewater," Applied Water Science, Vol. 7(7), pp. 3783–3791, 2017. [CrossRef]

[9] E. Hong, A. M. Yeneneh, T. K. Sen, H. M. Ang, and A. Kayaalp, "ANFIS based Modelling of dewatering performance and polymer dose optimization in a wastewater treatment plant," Journal of Environmental Chemical Engineering, Vol. 6(2), pp. 1957–1968, 2018. [CrossRef]

[10] M. S. Gaya, N. A. Wahab, Y. M. Sam, A. N. Anuar, and S. I. Samsuddin, "ANFIS modelling of carbon removal in domestic wastewater treatment plant," Applied Mechanics and Materials, Vol. 372, pp. 597–601, 2013. [CrossRef]

[11] M. Negnevitsky, "Artificial Intelligence A Guide to Intelligent Systems," (Second ed.). Vol. 123. Pearson, 2005.

[12] S. Akkurt, G. Tayfur, and S. Can, "Fuzzy logic model for the prediction of cement compressive strength," Cement and Concrete Research, Vol. 34(8), pp. 1429–1433, 2004. [CrossRef]

[13] F. I. Turkdogan-Aydinol and K. Yetilmezsoy, "A fuzzy-logic-based model to predict biogas and methane production rates in a pilot-scale mesophilic UASB reactor treating molasses wastewater," Journal of Hazardous Materials, Vol. 182(1–3), pp. 460–471, 2010. [CrossRef]

[14] D. Erdirencelebi and S. Yalpir, "Adaptive network fuzzy inference system modeling for the input selection and prediction of anaerobic digestion effluent quality," Applied Mathematical Modelling, Vol. 35(8), pp. 3821–3832, 2011. [CrossRef]

[15] Z. Hu, Y. V. Bodyanskiy, and O. K. Tyshchenko, Self-Learning and Adaptive Algorithms for Business Applications, Emarald Publishing, 2019. [CrossRef]

[16] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its applications to modeling and control," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-15(1), pp. 116–132, 1985. [CrossRef]

[17] J. R. Jang, "ANFIS : Adaptive-ne twork-based fuzzy inference system," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 23(3), 1993. [CrossRef]

[18] J. Wan, M. Huang, Y. Ma, W. Guo, Y. Wang, H. Zhang, W. Li, and X. Sun, "Prediction of effluent quality of a paper mill wastewater treatment using an adaptive network-based fuzzy inference system," Applied Soft Computing, Vol. 11(3), pp. 3238–3246, 2011. [CrossRef]

[19] Y. M. Wang, and T. M. S. Elhag, "An adaptive neuro-fuzzy inference system for bridge risk assessment," Expert Systems with Applications, Vol. 34(4), pp. 3099–3106, 2008. [CrossRef]