Desalination and Water Treatment
www.deswater.com

doi: 10.5004/dwt.2023.29320

286 (2023) 52–63
February

Enhance and improve modelling prediction by using an adaptive neuro-fuzzy inference system-based model to predict pollution removal efficacy in wastewater treatment plants

Hussein Y.H. Alnajjar*, Osman Üçüncü

Karadeniz Technical University, Trabzon, Turkey, Tel.: +90 5515876311; email: hyhn1989@hotmail.com (H.Y.H. Alnajjar), Tel.: +90 462 377 3142; email: oucuncu@ktu.edu.tr (O. Üçüncü)

Received 24 September 2022; Accepted 25 January 2023

ABSTRACT

An adaptive network-based fuzzy inference system (ANFIS) was used to create models for predicting the removal of biological oxygen demand (BOD), total nitrogen (TN), total phosphorus (TP), and total suspended solids (TSS) in a wastewater treatment plant treating process wastewaters. Temperature (*T*), hydraulic retention time, and dissolved oxygen were used as input variables for the BOD, TN, TP, and TSS models, using linear correlation matrices between input and output variables. The results show that the created system has provided reasonable forecasting and control performance. The minimum root mean square errors of 1.4816, 1.9558, 0.2299 and 0.4733 for effluent BOD, TN, TP and TSS could be achieved using ANFIS. The maximum *R*-square values for BOD, TN, TP and TSS were 0.9137, 0.9204, 0.9865 and 0.9231, respectively. ANFIS's architecture consists of both artificial neural networks and fuzzy logic including linguistic expression of membership functions and if–then rules, consequently it can overcome the limitations of traditional neural networks and increase the prediction performance.

Keywords: Adaptive network; Fuzzy inference; Neural networks; Wastewater treatment; Biological oxygen demand; Total nitrogen

1. Introduction

As the population grows and industries develop, wastewater treatment becomes increasingly important due to the increasing volume of wastewater generated by facilities each year. As a result, low-cost approaches that yield accurate results are required to aid in the prediction of treatment efficiency in wastewater treatment plants (WWTP). A number of complex and uncertain processes that are difficult to predict are involved in WWTP. The treatment plant's smooth and effective operation, on the other hand, is dependent on a suitable model capable of precisely capturing the system's dynamic character. The vast majority of previous models were used in industrial wastewater treatment plants. WWTP operation includes the physical, biological, and chemical characteristics 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, three different adaptive network-based fuzzy inference systems (ANFIS) and artificial neural networks (ANN) were used to forecast suspended solids (SS_{eff}) and chemical oxygen demand (COD_{eff}) in the effluent from a hospital wastewater treatment facility. In terms of effluent prediction, the results show that ANFIS statistically outperforms ANN [2]. The ANFIS model is used to forecast the pH quality of effluent. As a comparison, the artificial neural network is used [3].

* Corresponding author.

In another study, an adaptive neuro-fuzzy inference system (ANFIS) was described in another study to predict the effluent chemical oxygen demand load from a full-scale expanded granular sludge bed reactor (EGSBR) treating corn processing wastewater using five process variables: 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 drawn at random from the experimental domain. To validate the ANFIS-based predictions, various descriptive statistical measures such as root-mean-square error, index of agreement, a factor of two, fractional variance, the proportion of systematic error, and so on were used [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 the Tabriz wastewater treatment plant's effluent biological oxygen demand (BOD_{eff}) and chemical oxygen demand (COD_{eff}). Furthermore, the BO_{eff} and COD_{eff} parameters were predicted using the autoregressive integrated moving average (ARIMA) linear model to compare the abilities of linear and non-linear models in complicated process prediction [5]. In another research, ANFIS and generalized linear model (GLM) regression were used to identify the nonlinear system of an industrial wastewater treatment plant's activated sludge process. Predictive models of effluent chemical and 5-day biochemical oxygen demands were developed based on previously assessed inputs and outputs. From a set of candidates, the least absolute shrinkage and selection operator (LASSO) and a fuzzy brute force search were used to select the best combination of regressors for the GLMs and ANFIS models, respectively [6] code. Furthermore, ANFIS provides direct inverse control of the substrate in an activated sludge system. The performance of the proposed controller is demonstrated by tracking the substrate setpoints. The simulation results show that the proposed controller can effectively and precisely control the substrate concentration level. The proposed inverse controller could be a useful control method for the WWTP [3]. In another prior study, a full-scale aerobic biological wastewater treatment plant's removal efficiency of Kjeldahl nitrogen was evaluated using support vector machine (SVM) and adaptive neuro-fuzzy inference system (ANFIS) models. Input variables used during modeling include pH, COD, total solids (TS), free ammonia, ammonia nitrogen, and Kjeldahl nitrogen. The model work focused on developing an adaptive, functional, real-time, and alternative method for simulating Kjeldahl nitrogen removal efficiency [7]. Many studies have successfully employed ANFIS to increase the output of anaerobic digesters [8]. Also, for carbon and nitrogen removal, the ANFIS model was utilized. A feed-forward neural network is utilized as a comparison. The ANFIS model was found to have improved prediction power in all of the variables studied, including COD, suspended solids (SS), and ammonium nitrogen (NH_4–N) [9].

ANFIS model was created that could be used to calculate the effectiveness of pollutants being removed from the effluent of different primary and secondary treatment methods in a wastewater treatment plant (WWTP), including biological oxygen demand (BOD), total nitrogen (TN), total phosphorus (TP), and total suspended solids (TSS). Future WWTP design and pollutant removal efficiency forecasting might both be done using the ANFIS.

As a result, the primary purpose of this study is to apply, predict, and develop the pollutant removal efficiency for primary and biological treatment in WWTPs using the ANFIS model. For training, testing, and predictions, this modeling employed data from the urban coastal in Ordu City, Altnordu District – Durugöl Advanced Biological Wastewater Treatment Plant in Turkey, together with the conditions provided in the rules. The parameters studied were BOD, TN, TP, and TSS. For this study, wastewater samples were taken twice a month from the inlet and outlet of wastewater treatment plants in urban coastal towns in 2018. The data was standardized before running the simulation for prediction. The ANFIS model's output was compared to real-world training data and ANN data, and the error was minimized to produce the best operating points.

2. Materials and methods

Intelligent technology includes probabilistic reasoning, fuzzy logic, neural networks, and evolutionary computation. 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 [10].

A system that includes at least two intelligent technologies is referred to as a "Hybrid Intelligent System". Combining a neural network with a fuzzy system, for example, produces a hybrid neuro-fuzzy system. Soft computing (SC) is a new method for building hybrid intelligent systems that can reason and learn in an uncertain and imprecise environment. It combines probabilistic reasoning, fuzzy logic, neural networks, and evolutionary computation.

At the end of the information and fuzzy rule base interaction, the output unit generates variables. Fig. 1 depicts the general ANFIS process for developing the ANFIS prediction model.

2.1. 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. 2 [11]. The input unit contains the input variables, as well as any information about the input variables that will affect the scenario under investigation [12]. 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 [13]. 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.

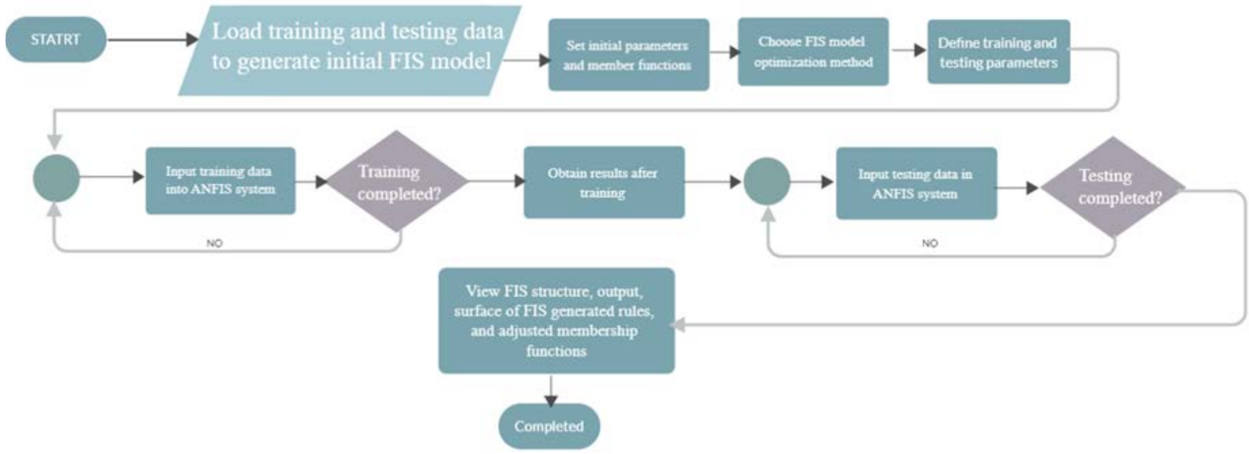

Fig. 1. Flow chart of ANFIS test step [8].

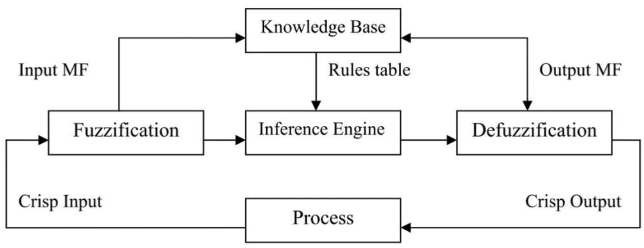

Fig. 2. Basic structure of the fuzzy logic controller.

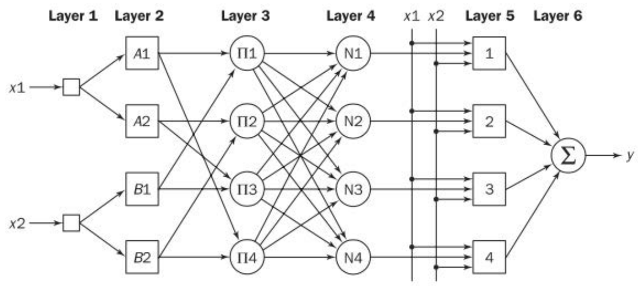

Fig. 3. ANFIS structure.

2.2. Model architecture and components

When a neural network and a fuzzy system are coupled, a powerful hybrid system capable of addressing complex problems is generated. The behavior of this hybrid system may be described in terms similar to human rules, making it a reliable tool for modelling non-linear functions [14]. ANFIS uses a hybrid learning technique to specify how the weights should be updated to reduce the error between the actual and desired output, while also altering the parameters and structure of the fuzzy inference system (FIS). The structure of ANFIS, which is a Sugeno fuzzy model, is shown in Fig. 3.

The Sugeno model's structure is constructed so that the input is mapped to the input membership function, the input membership function is then mapped to the rule, the rule is then mapped to the output membership function, and finally, the output membership function is mapped to the output. The system, therefore, requires five levels. A membership grade is generated by each node in the top layer. The firing strength of the rule is calculated by each node in the second layer. The firing strength of the *i*th rule about the sum of all firing strengths is calculated by each node in the third layer. The fourth layer's nodes are all adaptive nodes that correspond to the output membership functions. The node in the fifth layer gives the overall output [15].

ANFIS is an adaptable network that, like the Takagi–Sugeno fuzzy inference system, employs supervised learning as a learning mechanism [16]. The model's five primary components are inputs and outputs, database and pre-processor, fuzzy system generator, fuzzy inference system, and adaptive neural network [17]. In most cases, the input and output parameters are chosen or derived from the parameters of the system description. Model generation necessitates the use of a database and pre-processor, both of which include information regarding system performance. This data is often acquired by collecting data on parameters that the system continuously monitors. MATLAB R2021b [18] is a useful tool for this research, and it is used to generate system performance data. An adaptive network-based fuzzy inference system, as well as a Sugeno fuzzy inference system and associated adaptive networks, are used (ANFIS). The input and output variables are selected or generated from the variables commonly used to describe the system. The construction of a database containing system performance data is required for model building. Most of the time, it's made by pulling parameters from Durugöl Advanced Biological Wastewater Treatment Plant. The training database must be of good quality for the model to produce reliable information on the system. For the model to effectively define the system, the database must include adequate and reliable information about it. In contrast, a raw database is likely to contain some duplicated and inconsistent data. As a result, it's possible that the raw training database will need to be pretreated to eliminate duplicates and data conflicts. A fuzzy system generator is required since the ANFIS is often launched with a prototype fuzzy system. The software MATLAB R2021b (Matworks Inc.) provides this function.

The model was programmed in MATLAB R2021b by Jang [17], indicating that the language is suitable for model programming.

The model will be used to determine the relationship between the real data obtained from WWTP and the ANFIS model to achieve the lowest possible error.

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 (Wan et al. [20]), the following is an example of a rule set containing four fuzzy if–then rules:

- Rule 1: If x is A_1 and y is B_1 then $f_1 = p_1x + q_1y + r_1$
- Rule 2: If x is A_2 and y is B_2 then $f_2 = p_2x + q_2y + r_2$

where A_1, A_2, B_1 and B_2 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 Fig. 3, consists of five levels, each of which performs a different function in the ANFIS.

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

$$O^1_{\mathrm{Ai}} = U_{\mathrm{Ai}}(x) i = 1,2 \tag{1}$$

$$O^1_{\mathrm{Bj}} = U_{\mathrm{Bj}}(x) j = 1,2 \tag{2}$$

where x and y are crisp inputs, and A_i and B_j are fuzzy sets characterized by appropriate MFs, which could be triangular, trapezoidal, Gaussian function, or other shapes, and A_i and B_j are fuzzy sets characterized by appropriate MFs, which could be triangular, trapezoidal, Gaussian function, or other shapes. The generalized bell-shaped MFs [Eqs. (3) and (4)] defined below are used in this investigation.

$$U_{\mathrm{Ai}}(x) = \frac{1}{1 + \left(\frac{x - c_i}{a_i}\right)^{2b_i}} \quad i = 1,2 \tag{3}$$

$$U_{\mathrm{Bj}}(x) = \frac{1}{1 + \left(\frac{x - c_j}{a_j}\right)^{2b_j}} \quad j = 1,2 \tag{4}$$

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 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:

$$O^2_{ij} = w_{ij} = U_{Ai}(x) U_{Bj}(y), i, j = 1,2 \tag{5}$$

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

$$O^3_{ij} = \overline{w_{ij}} = \frac{w_{ij}}{w_{11} + w_{12} + w_{21} + w_{22}}, i, j = 1,2 \tag{6}$$

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. (7) gives the outputs of this layer.

$$O^4_{ij} = \overline{w_{ij}} f_{ij} = \overline{w_{ij}}\left(p_{ij} + q_{ij}y + r_{ij}\right), \; i, j = 1,2 \tag{7}$$

Subsequent parameters refer to the parameters in this layer.

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

$$\begin{aligned} z = O^5_1 &= \sum_{i=1}^{2}\sum_{j=1}^{2} \overline{w_{ij}} f_{ij} = \sum_{i=1}^{2}\sum_{j=1}^{2} \overline{w_{ij}}\left(p_{ij}x + q_{ij}y + r_{ij}\right) \\ &= \sum\sum\left(\overline{w_{ij}}x\right)p_{ij} + \left(\overline{w_{ij}}y\right)q_{ij} + \left(\overline{w_{ij}}\right)r_{ij} \end{aligned} \tag{8}$$

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. Wan et al. [20] provides a detailed algorithm and mathematical basis for the hybrid learning approach.

The model is trained until results are produced with the least amount of inaccuracy. The settings for the training process must be chosen carefully to construct an ANFIS system for real-world challenges. Proper training and testing data sets are crucial. The testing data set will not validate the model if the datasets are improperly chosen. The model cannot capture any of the properties of the testing data if the testing data set is entirely different from the training data set.

The first epoch can thus attain the lowest testing error. The testing error lowers for the right data set while training continues up until a jump point. When the training goes above that level, overfitting happens. The optimization techniques are employed to gain knowledge of the training data. The membership parameters are changed as learning takes place. The two ANFIS parameter optimization techniques available in MATLAB are background and hybrid (default: mixed least squares and backpropagation) (backpropagation). As a training-stopping criterion, error tolerance—which is correlated with error size—is applied. Once the training data error is still within this tolerance, the training will end [5].

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 errors (RMSE), mean absolute percentage error (MAPE), and correlation coefficient (R) are utilized. The MSE performance index was established as follows:

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(\hat{y} - y\right)^2 \tag{9}$$

The RMSE performance index was defined as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\hat{y} - y\right)^2}{n}} \tag{10}$$

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

Mean absolute percentage error (MAPE):

$$\text{MAPE} = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \times 100 \tag{11}$$

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):

$$R = \frac{\sum_{t=1}^{n}\left(A_t - \bar{A}\right)\left(F_t - \bar{F}\right)}{\sqrt{\sum_{t=1}^{n}\left(A_t - \bar{A}\right)^2 \cdot \sum_{t=1}^{n}\left(F_t - \bar{F}\right)^2}} \tag{12}$$

MBE (mean bias error) = $\frac{1}{n}\sum_{i=1}^{n}\left(P_i - O_i\right)$ (11) where O_i is the observation value and P_i is the forecast value.

2.4. Plant description

Durugöl Advanced Biological Wastewater Treatment Plant construction has been designed according to physical (coarse screen, fine screen and primary settler units) and biological treatment project (anaerobic tanks, aeration tanks and secondary settler) and has a capacity of 212.000 person/d. Durugöl Advanced Biological Wastewater Treatment Plant is located in Ordu City, Turkey, whose location is given in Fig. 4. The treated wastewater is discharged to the Black Sea. It has been designed in 2 stages according to population growth.

- Phase 1 (2025): It will serve a population of 213,000 (34,000 m^3 flow rate).
- Phase 2 (2045): It will serve a population of 292,000 (43,000 m^3 flow rate).

The grits contained in the influent wastewater are removed in the grit chamber to avoid causing damage to the system. A large percentage of BOD, COD, SS and other pollutants are removed during the primary treatment. The effluent from the primary settler flows to the secondary treatment unit which consists of anaerobic tanks, aeration tanks and secondary settler as shown in Fig. 5. In the aeration tanks, favorable condition is provided for the microorganisms responsible for degrading the remaining dissolved organic pollutants in the wastewater to grow and form sludge. The sludge is separated from the treated water in the secondary settler by gravity sedimentation. A portion of the sludge is returned to the aeration unit to maintain

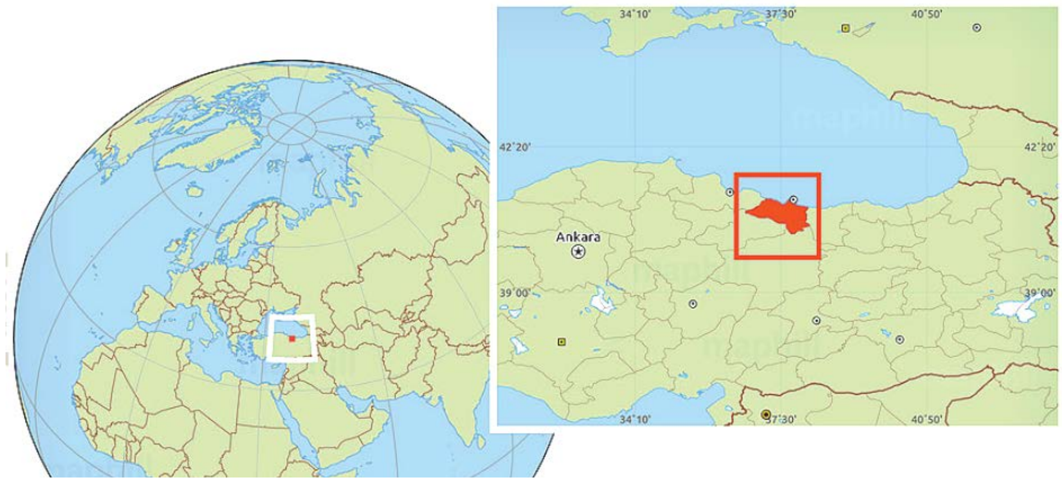

Fig. 4. Location of Durugöl Advanced Biological Wastewater Treatment Plant.

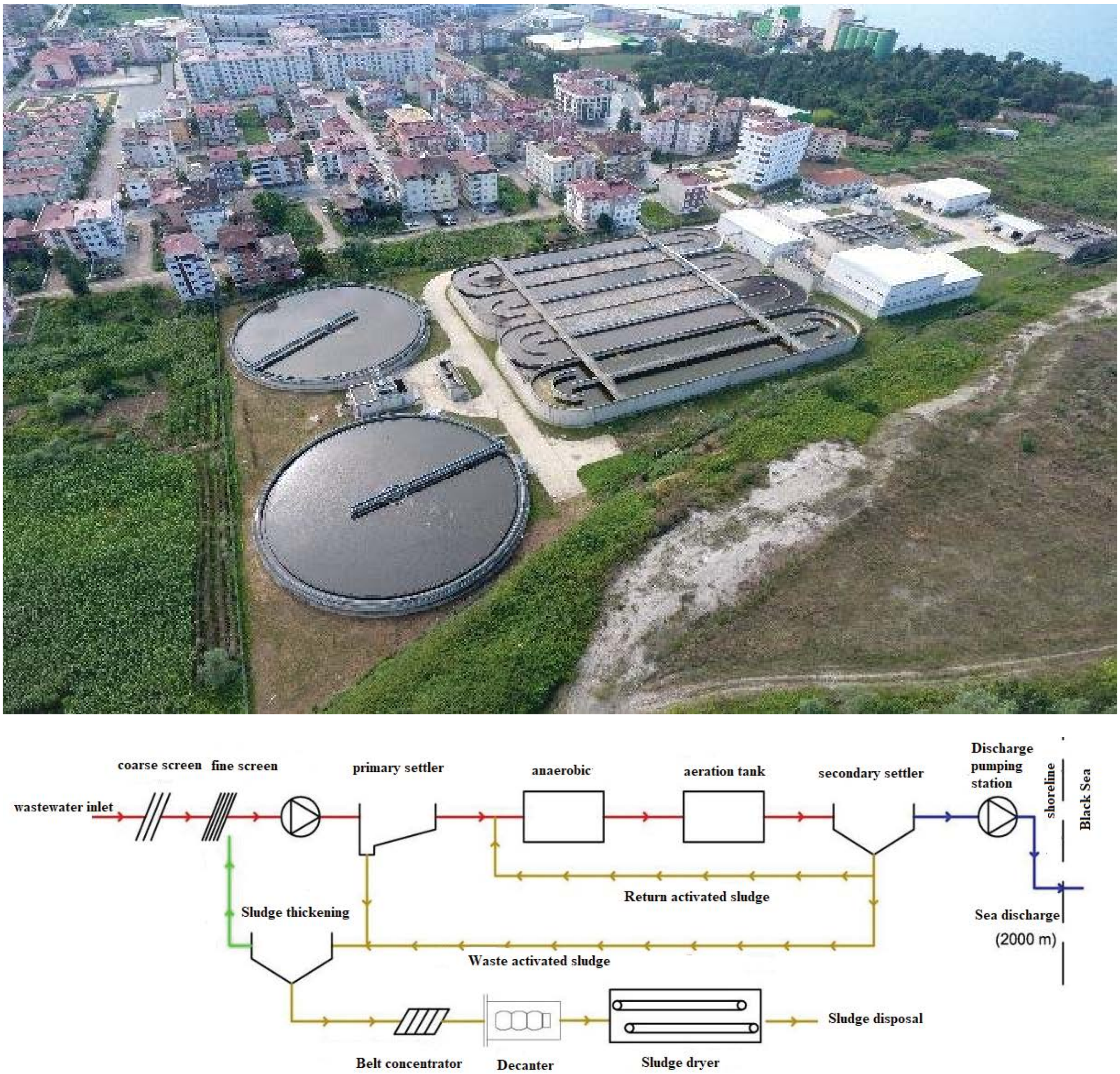

Fig. 5. Schematic of Durugöl Advanced Biological Wastewater Treatment processes.

the microorganisms' concentration, and the waste sludge is removed and transferred to the sludge treatment facility. A suitable model could be useful in the application of control strategy or optimization technique to the plant to increase the treatment efficiency.

2.5. Model implementation

The ANFIS 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 Durugöl Advanced Biological WWTP'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 WWTP data effluent parameters like output BOD, TN, TP, and TSS.

Fig. 6a and b show the topology of the ANFIS network that was employed. In the creation of a fuzzy system, ANFIS structures with varying input correlation (Fig. 6a) and consisted of five layers were established (Fig. 6b). The following are the meanings of each layer in Fig. 6b, 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: hydraulic retention time (HRT), *T*, and dissolved oxygen (DO) (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.

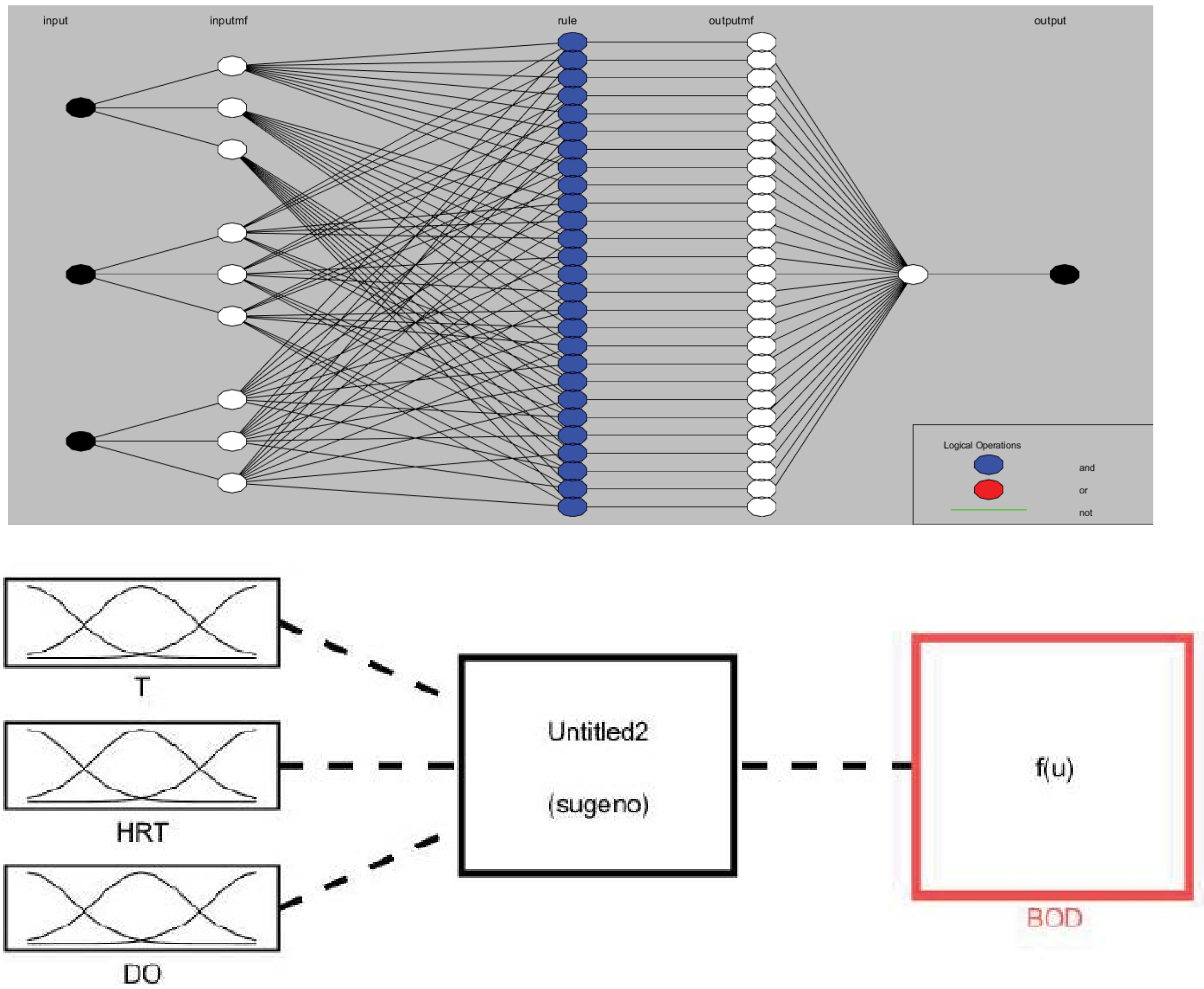

Fig. 6. Schematic diagram of (a) ANFIS models with all input variables and (b) input–output mapping structure of ANFIS models with input variables.

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 x_i computed using soft-min or product. Layer of the output membership function: In the function layer, each weighs the result of its linear regression f_i, resulting in the rule output.

Normalization layer:

Each x_i is scaled into the normalization layer normalization.

Normalization is performed with the equation:

$$x_{\text{norm}} = \left(x_{\text{value}} - x_{\text{min}}\right)\left(x_{\text{max}} - x_{\text{min}}\right) \quad (13)$$

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

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

The measured data for WWTP of the ANFIS and ANN models is shown in Table 1.

3. Results and discussion

As a modeling method, the ANFIS tool from MATLAB R2021's Fuzzy Logic Toolbox (The MathWorks, Inc.) was employed. To build the fuzzy rule basis sets, subtractive clustering and grid division methods of fuzzy inference systems were used. To attain the minimal RMSE and maximum R^2, a maximum of 100 epochs were used for training. Predicted values and observed data were evaluated using the following statistical measures (RMSE and R^2), which have been used by numerous studies to quantify model performance [21].

In this study, the ANFIS modelling has been applied to predict some important parameters in WWTP. In the modelling, influent parameters such as influent T, HRT and DO were used as input parameters to predict the effluent removal efficiency of BOD, TN, TP and TSS. Then the results of ANFIS model were compared with the results of the ANN model. This data set was divided randomly into two subsets for training and for testing purposes. More data were used in the training phase because ANFIS is more adapted nonlinear functional dependency between input

Table 1
Measured data for WWTP of the ANFIS and ANN models

Months	T (°C)	HRT (h)	DO (mg/L)	BOD (mg/L)			TN (mg/L)			TP (mg/L)			TSS (mg/L)		
				Input	Output	Efficiency	Input	Output	Efficiency	Input	Output	Efficiency	Input	Output	Efficiency
Jan.	8.00	4.80	3.80	88.30	21.60	75.54	30.30	8.70	71.29	5.80	0.40	93.10	223.00	16.00	92.83
Feb.	7.00	5.00	4.10	105.20	14.90	85.84	34.50	9.50	72.46	4.50	0.30	93.33	217.50	9.50	95.63
Mar.	9.00	4.60	3.60	139.90	17.90	87.21	34.50	7.80	77.39	4.50	0.40	91.11	201.50	9.50	95.29
Apr.	12.00	4.40	2.80	104.30	18.60	82.17	36.00	7.40	79.44	3.90	0.30	92.31	141.50	10.50	92.58
May	16.00	4.00	2.50	89.10	16.70	81.26	34.50	6.40	81.45	4.90	0.40	91.84	190.00	7.50	96.05
Jun.	21.00	3.80	2.20	124.70	16.00	87.17	31.00	7.20	76.77	4.00	0.20	95.00	204.00	7.00	96.57
Jul.	23.00	3.50	2.00	176.90	14.60	91.75	33.50	7.20	78.51	3.40	0.20	94.12	130.00	9.00	93.08
Aug.	24.00	3.20	1.80	134.70	15.20	88.72	23.50	6.30	73.19	3.10	0.30	90.32	171.50	10.00	94.17
Sep.	21.00	3.70	2.40	132.90	18.60	86.00	30.00	5.50	81.67	4.00	0.30	92.50	181.50	4.50	97.52
Oct.	17.00	4.10	2.60	135.60	13.20	90.27	34.50	8.40	75.65	3.90	0.40	89.74	216.50	7.50	96.54
Nov.	12.00	4.30	2.90	94.40	10.70	88.67	24.50	9.40	61.63	3.50	0.40	88.57	201.50	9.50	95.29
Dec.	9.00	4.50	3.50	97.10	15.90	83.63	22.50	9.20	59.11	3.90	0.30	92.31	157.50	5.50	96.51
Min.	7.00	3.20	1.80	88.30	10.70	75.54	22.50	5.50	59.11	3.10	0.20	88.57	130.00	4.50	92.58
Max.	24.00	5.00	4.10	176.90	21.60	91.75	36.00	9.50	81.67	5.80	0.40	95.00	223.00	16.00	97.52
Mean	14.92	4.16	2.85	118.59	16.16	85.68	30.78	7.75	74.05	4.12	0.33	92.02	186.33	8.83	95.17
Standard deviation	6.22	0.542	0.7441	26.74	2.8353	4.46829	4.785	1.312	7.205628	0.73	0.075	1.84368	30.55	2.918	1.644507

and output variables. 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 premise 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. 7a). 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. 8 shows an

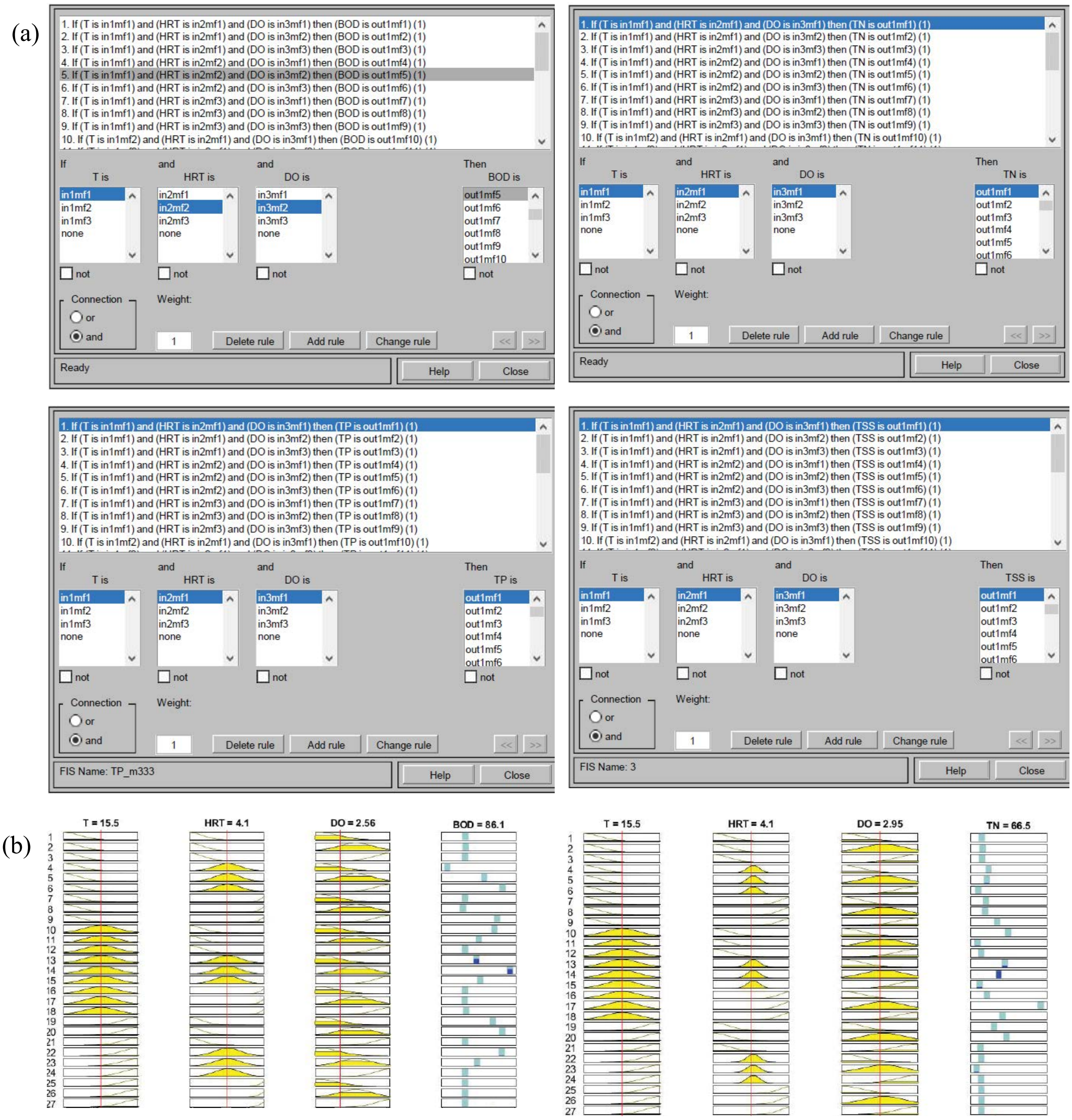

Fig. 7. (a) Rule editor of MATLAB R2021b Fuzzy Logic Toolbox and (b) rule viewer screen to obtain defuzzified results.

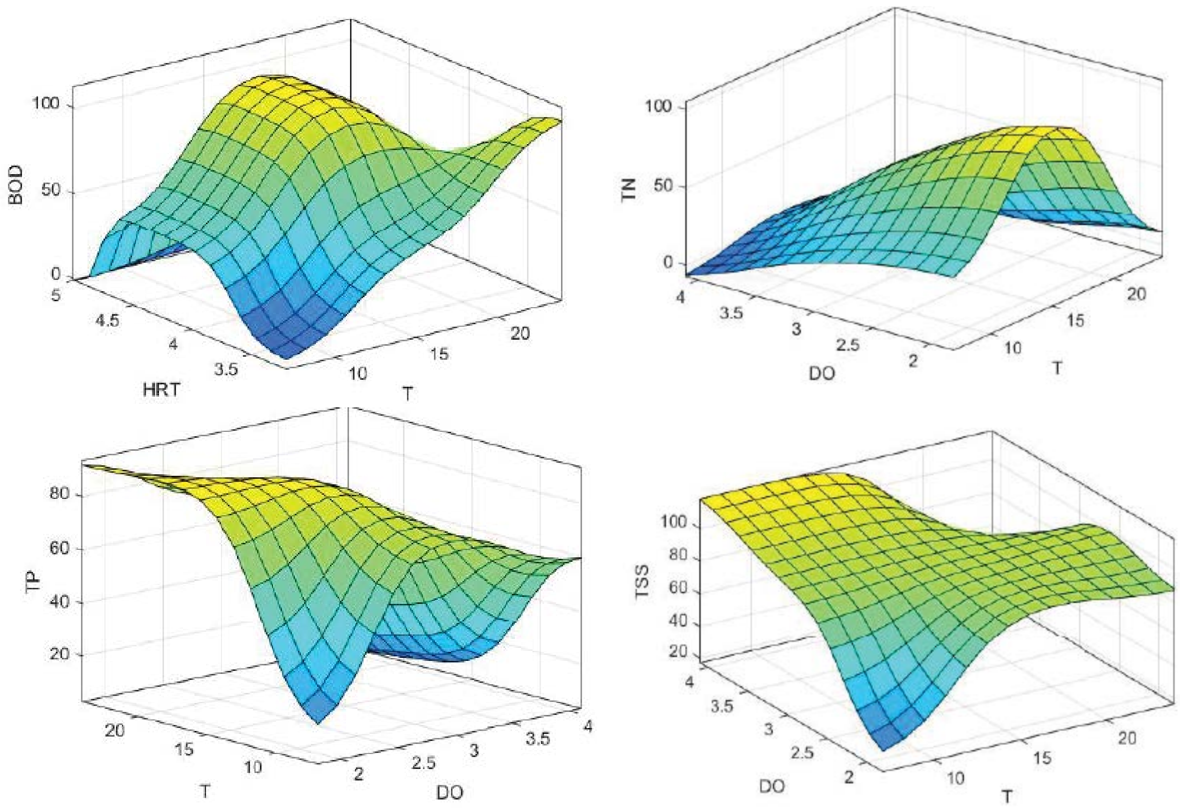

Fig. 8. 3D response surface graph.

Table 2
Determination of the appropriate ANFIS and ANN models

Outputs	Training data	Testing data	Number of input MF	Number of rules	ANFIS			ANN-ANFIS		
					R^2	RMSE	MBE	R^2	RMSE	MBE
BOD	12	4	3.3	27	0.9137	1.4816	–0.192	0.5116	2.4068	0.083
TN	12	4	3.3	27	0.9204	1.9558	0.148	0.7264	3.7375	0.07
TP	12	4	3.3	27	0.9865	0.2299	–0.005	0.7461	0.8372	–0.056
TSS	12	4	3.3	27	0.9231	0.4733	0.045	0.4731	0.9819	–0.143

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. 7b 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.

All *R*-square and RMSE values for the removal efficiency of BOD, TN, TP and TSS are also shown in Table 2. When training, *R*-square value was 0.9137 using ANFIS But when comparing the ANFIS and ANN, the value of R^2 was 0.5116 for BOD, which indicates that the efficiency of the ANFIS model is higher in predicting the efficiency of pollutant removal in wastewater treatment plants (more details in Table 2).

The influence of the ANFIS model inputs (temperature, dissolved oxygen, and hydraulic retention time) on the model outputs is also shown in Fig. 8 (BOD, TN, TP and TSS).

The ANFIS model, for example, indicates that raising the temperature and lengthening the hydraulic retention time improves the efficiency of pollutant removal in wastewater.

The RMSE value of 1.9558 using ANFIS was also lower than that of 3.7375 using ANN when predicting for TN. Fig. 9 shows the training and predicting results using ANFIS and ANN models.

To overcome the constraints of standard neural networks, such as the risk of becoming trapped in a local minimum and model architecture selection, and to increase predicting performance, ANFIS' architecture incorporates

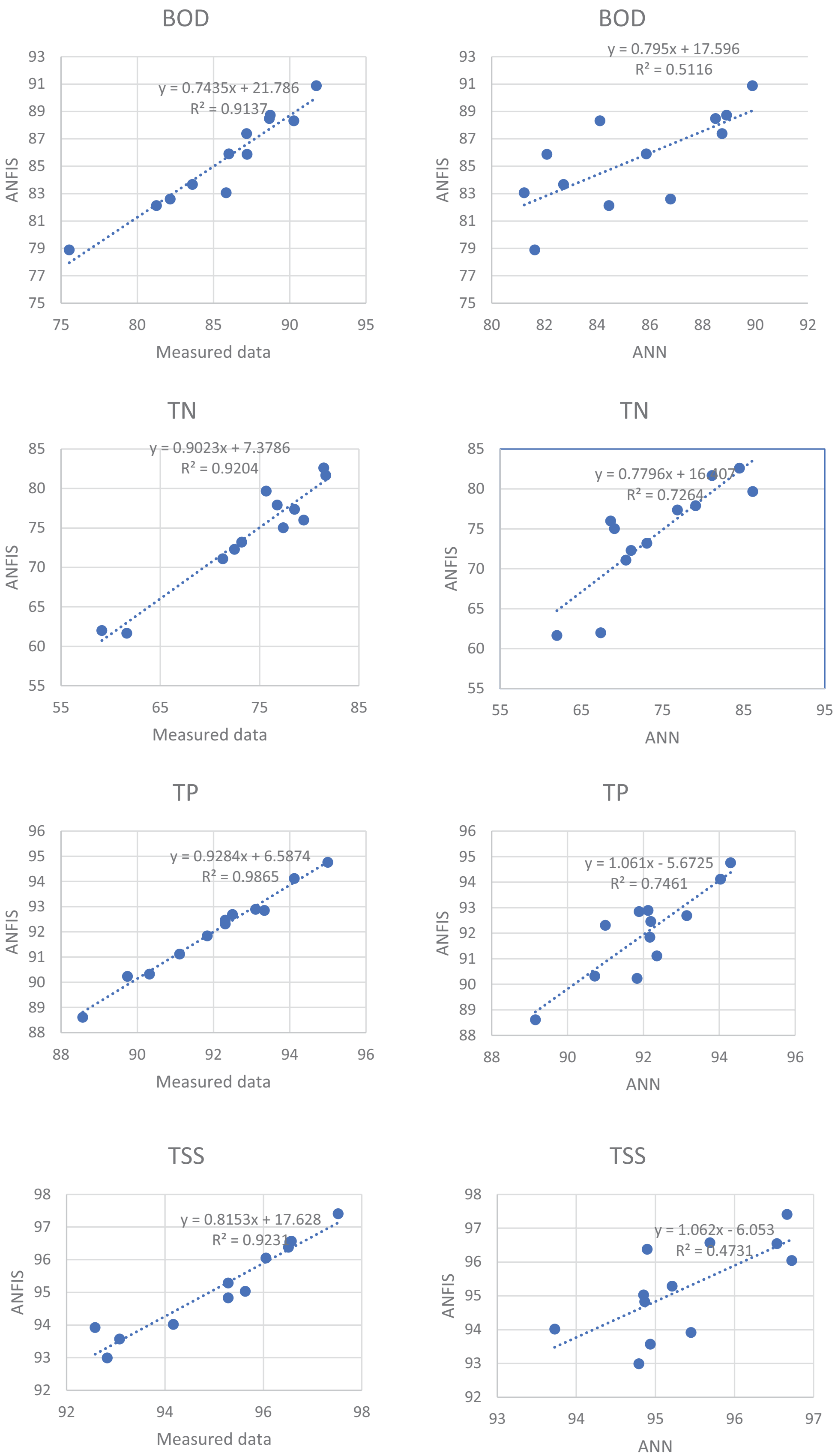

Fig. 9. Prediction results of BOD, TN, TP and TSS for the ANFIS and ANN models.

ANN and fuzzy logic, as well as linguistic expressions of MFs and if–then rules. As a result, ANFIS is a fantastic tool for simulating wastewater treatment efficiency. ANN is also a black box in nature, with difficult to grasp relationships between inputs and outputs, but ANFIS is transparent, with simple to understand and interpret if–then rules.

4. Conclusions

ANFIS was used to predict the removal efficiency of BOD, TN, TP and TSS S from Durugöl Advanced Biological Wastewater Treatment Plant in Turkey. The ANN was also adopted for comparison. According to the results, ANFIS could predict the effluent variation. The minimum RMSEs of 1.4816, 1.9558, 0.2299 and 0.4733 for effluent BOD, TN, TP and TSS could be achieved using ANFIS. The maximum *R*-square values for BOD, TN, TP and TSS were 0.9137, 0.9204, 0.9865 and 0.9231, respectively. It also revealed that the influent indices could be applied to the prediction of effluent quality. After obtaining good predicting results using ANFIS, it is advised that the ANFIS be employed as the objective function or constraints in optimization for the optimal design or operation in future studies. Given the high level of complexity in the wastewater treatment process, the significant amount of variable information dispersed across the dataset, and the wide concentration ranges, ANFIS models' excellent prediction results for both effluent parameters are particularly relevant. As a result, the ANFIS modeling approach could serve as a generic foundation for modeling different treatment procedures. Furthermore, the ANFIS modeling approach could be used to anticipate and control the performance of treatment processes in treatment plants.

References

[1] H.Y.H. Alnajjar, 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, Int. J. Environ. Pollut. Environ. Modell., 4 (2021) 85–94.

[2] T.Y. Pai, T.J. Wan, S.T. Hsu, T.C. Chang, Y.P. Tsai, C.Y. Lin, H.C. Su, L.F. Yu, Using fuzzy inference system to improve neural network for predicting hospital wastewater treatment plant effluent, Comput. Chem. Eng., 33 (2009) 1272–1278.

[3] M.S. Gaya, N.A. Wahab, Y.M. Sam, S.I. Samsuddin, ANFIS-based effluent pH quality prediction model for an activated sludge process, Adv. Mater. Res., 845 (2014) 538–542.

[4] K. Yetilmezsoy, H. Ozgun, R.K. Dereli, M.E. Ersahin, I. Ozturk, Adaptive neuro-fuzzy inference-based modeling of a full-scale expanded granular sludge bed reactor treating corn processing wastewater, J. Intell. Fuzzy Syst., 28 (2015) 1601–1616.

[5] V. Nourani, P. Asghari, E. Sharghi, Artificial intelligence based ensemble modeling of wastewater treatment plant using jittered data, J. Cleaner Prod., 291 (2021) 125772, doi: 10.1016/j.jclepro.2020.125772.

[6] D.O. Araromi, O.T. Majekodunmi, J.A. Adeniran, T.O. Salawudeen, Modeling of an activated sludge process for effluent prediction—a comparative study using ANFIS and GLM regression, Environ. Monit. Assess., 190 (2018) 495, doi: 10.1007/s10661-018-6878-x.

[7] D.S. Manu, A.K. Thalla, Artificial intelligence models for predicting the performance of biological wastewater treatment plant in the removal of Kjeldahl nitrogen from wastewater, Appl. Water Sci., 7 (2017) 3783–3791.

[8] E. Hong, A.M. Yeneneh, T.K. Sen, H.M. Ang, A. Kayaalp, ANFIS based modelling of dewatering performance and polymer dose optimization in a wastewater treatment plant, J. Environ. Chem. Eng., 6 (2018) 1957–1968.

[9] M.S. Gaya, N.A. Wahab, Y.M. Sam, A.N. Anuar, S.I. Samsuddin, ANFIS modelling of carbon removal in domestic wastewater treatment plant, Appl. Mech. Mater., 372 (2013) 597–601.

[10] M. Negnevitsky, Artificial Intelligence A Guide to Intelligent Systems, 2nd ed., Vol. 123, Pearson Education, England, 2005.

[11] S. Akkurt, G. Tayfur, S. Can, Fuzzy logic model for the prediction of cement compressive strength, Cem. Concr. Res., 34 (2004) 1429–1433.

[12] F.I. Turkdogan-Aydinol, K. Yetilmezsoy, A fuzzy-logic-based model to predict biogas and methane production rates in a pilot-scale mesophilic UASB reactor treating molasses wastewater, J. Hazard. Mater., 182 (2010) 460–471.

[13] D. Erdirencelebi, S. Yalpir, Adaptive network fuzzy inference system modeling for the input selection and prediction of anaerobic digestion effluent quality, Appl. Math. Modell., 35 (2011) 3821–3832.

[14] Z. Hu, Y.V. Bodyanskiy, O.K. Tyshchenko, Self-Learning and Adaptive Algorithms for Business Applications: A Guide to Adaptive Neuro-fuzzy Systems for Fuzzy Clustering under Uncertainty Conditions, No. 2019, Emerald Publishing Limited, United Kingdom, 2019.

[15] V. Vaidhehi, The role of dataset in training ANFIS system for course advisor, Int. J. Innov. Res. Adv. Eng., 1 (2014) 2349–2163.

[16] T. Takagi, M. Sugeno, Fuzzy identification of systems and its applications to modeling and control, IEEE Trans. Syst. Man Cybern., SMC-15 (1985) 116–132.

[17] J.-S.R. Jang, ANFIS: adaptive-network-based fuzzy inference system, IEEE Trans. Syst. Man Cybern., 23 (1993) 665–685.

[18] MATLAB, The MathWorks Inc. Version R2022b, The MathWorks Inc., The MathWorks, Inc., United States, 2022. Available at: https://matlab.mathworks.com

[19] Y.-M. Wang, T.M.S. Elhag, An adaptive neuro-fuzzy inference system for bridge risk assessment, Expert Syst. Appl., 34 (2008) 3099–3106.

[20] J. Wan, M. Huang, Y. Ma, W. Guo, Y. Wang, H. Zhang, W. Li, X. Sun, Prediction of effluent quality of a paper mill wastewater treatment using an adaptive network-based fuzzy inference system, Appl. Soft Comput., 11 (2011) 3238–3246.

[21] Z. Cheng, X. Li, Y. Bai, C. Li, Multi-scale fuzzy inference system for influent characteristic prediction of wastewater treatment, Clean - Soil, Air, Water, 46 (2018) 1–11.