

Research Article

Employing Artificial Neural Networks to Predict the Performance of Domestic Sewage Treatment Terminals in the Rural Region

Qiang Lin ¹, Ancheng Luo ¹, Yan Zhang ¹, Yunlong Wang ², Zhiwei Liang ¹,
and Ping Yuan ¹

¹Zhejiang University, College of Environmental & Resource Sciences, Hangzhou, China

²Zhejiang Academy of Agricultural Sciences, Hangzhou, China

Correspondence should be addressed to Zhiwei Liang; zhiweiliangzg@zju.edu.cn

Received 25 April 2021; Revised 27 August 2021; Accepted 15 October 2021; Published 23 December 2021

Academic Editor: José Francisco Gómez Aguilar

Copyright © 2021 Qiang Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Domestic sewage in rural regions is mainly treated by small-scale treatment terminals in China. The large quantities and high dispersion of these terminals render the chemical measurement of effluent to be a time and energy intensive work and further hinder the efficient surveillance of terminals' performance. After a thorough investigation of 136 operating terminals, this study successfully employs two artificial neural network (ANN) models to predict effluent total nitrogen (TN) and COD (R^2 both higher than 0.8) by setting some easily detectable parameters, e.g., pH and conductivity, as inputs. To prevent ANN models getting stuck on local optima and enhance the model performance, genetic algorithm (GA) and particle swarm optimization (PSO) are introduced into ANN, respectively. By comparison, ANN-PSO excels in modelling both TN and COD. The root mean square error (RMSE) and R^2 of ANN-PSO in modelling TN are 9.14 and 0.90, respectively, in the training stage, and 11.54 and 0.90, respectively, in the validation stage. The RMSE and R^2 of ANN-PSO in modelling COD are 22.10 and 0.90, respectively, in the training stage, and 26.57 and 0.85, respectively, in the validation stage. This is the first study to provide performance prediction models that are available for different terminals. Two established ANN-PSO models show great practical significance in monitoring huge amounts of terminals despite the slight sacrifice of models' accuracy caused by the great heterogeneity of different terminals.

1. Introduction

The economic boom and fast increase in the living standards of residents brings about the growing production of rural domestic sewage (RDS). It is estimated that, in China, the annual RDS discharge reaches up to 19.5 billion tons, which is about 63% of the urban domestic sewage [1]. In light of the large amounts of nutrients like organic matter or nitrogen contained in the RDS, either direct discharge or improper treatment of RDS will impose non-negligible threats to the receiving water [2]. In many developing countries, RDS has become the main source of pollution in the rural region [3, 4].

In the Zhejiang province, RDS is mainly treated by small-scale terminals with treatment capacities ranging from tons to dozens of tons. Traditional biological treatment

(A^2O) dominates the technology mainstream of these terminals with regard to its competitive edge in a low construction cost and energy demands. Whereas, the notorious problem of A^2O that the performance of the biological process is easily affected by the ambient environment has gradually stood out in recent years [5]. The approach of periodical manual sampling integrated with a traditional chemical test has been adopted as the main monitoring strategy to determine some important effluent index, like COD or total nitrogen (TN), by most regional governments and carried out for decades. However, the large quantities and high dispersion (sometimes, tens of thousands of terminals scatter throughout one city) of these terminals render the surveillance work to be a time and energy intensive work and require large capital investment.

1.1. Application of the ANN Model on Modelling the Water Quality. Machine learning (ML) methods provide some potential alternatives to control or simulate targets through examples or past experiences [6]. Among them, an artificial neural network (ANN) has become increasingly popular in the field of wastewater treatment and exhibited more excellent accuracy in modelling nonlinear targets like effluent water quality than many other ML methods [7]. For instance, Abyaneh [8] found ANN excelled higher accuracy and adequacy capacity in modelling BOD and COD of WWTP, compared with the multivariate linear regression method. In addition, in the study by Mahdiah et al. [9], ANN obtained the best prediction performance in accuracy relative to the extreme learning machine and support vector machine methods.

The backpropagation (BP) ANN model is one of the most studied ANN models, which can redistribute errors from the output to input layer by iterations in order to find the appropriate model parameters like weights and thresholds. Excellent self-learning and adaptability of BP ANN has already been reflected by the applications in multiple fields [10]. For instance, Antwi et al. [11] employed two BP ANN models in the prediction of ammonia and total nitrogen removal and demonstrated a good result ($R^2 > 0.98$). Likewise, Mandal et al. [12] also used BP ANN in simulating As(III) removal with R^2 above 0.97 for both training and validation processes. However, the drawbacks of BP ANN that it tends to be trapped by local optima due to severe initialization sensitivity are often put forward by researchers [13]. Apart from that, the high requirements for computational complexity and memory in some BP ANN intrinsic algorithms like Levenberg–Marquardt also deserve proper attention [14].

1.2. Application of Hybrid ANN on Modelling the Water Quality. Evolutionary algorithms, like particle swarm optimization (PSO) and genetic algorithm (GA), are often introduced into ANN as optimization strategies [15]. The principle of PSO is to globally search the solution space in order to select the most well-behaved particles [16]. It has edges in the low computing volume, strong memory ability for remembering the best position of each particles, and higher convergence characteristics as it only depends on the particle velocity to do the searching job [15, 17]. Improvement in the prediction accuracy of the PSO-based hybrid model has been documented in many previous studies. Mei et al. [18] introduced PSO into ANN in a electro-oxidation system and achieved accurate predictions with R^2 of 0.99 and 0.9944 for COD removal and total energy consumption, respectively. Khajeh et al. [19] validated the hybrid model, ANN-PSO, which was robust in modelling Mn(II) and Co(II) removal efficiency in adsorption (R^2 was 0.942 and 0.944 for Mn(II) and Co(II), respectively). GA is a metaheuristic algorithm inspired from the natural selection process [20]. It is suitable to search for a single and exclusive target and obtain satisfying performance with reduced complexity of ANN [15]. ANN-GA models have also shown to be superior than ANN in various fields. The

study of Azad et al. [21] showed that ANFIS models (adaptive neuro fuzzy inference system) only displayed good simulation in the training stage of modelling precipitation in the winter and spring, and the accuracy of models in the validation stage was very poor. ANFIS-GA made up for these shortcomings and achieved the purpose of optimization. Jalalkamali [22] reported that ANFIS-PSO and ANFIS-GA both exhibited excellent simulation of spatiotemporal groundwater quality, and the ANFIS-PSO model yielded better performance than ANFIS-GA.

1.3. Limitations of Current Cases Applied on Modelling the Effluent of RDS Terminals. Although many successful ANN cases have been applied to predict effluent quality of WWTP, two significant shortcomings are worth highlighting when these cases are extended to RDS: (1) database of the established model mainly comes from the historical data of a single target, like a specific WWTP. Great heterogeneity among terminals will inevitably challenge the availability of the model (established for a specific terminal) for other terminals, while constructing the model for each terminals would be too costly. (2) Inputs contained some parameters that are difficult or costly to be measured. In some cases, influent TN even served as inputs for effluent TN prediction [23].

This research is dedicated to finding a universal, practicable, and affordable monitoring approach for different terminals. To make the model applicable to as many terminals as possible, data from 136 operating terminals were collected. Then, ANN, ANN-GA, and ANN-PSO models are employed in this study to predict effluent TN and COD by setting some easily detectable and low cost parameters like pH and conductivity as inputs.

2. Methods and Materials

2.1. Investigation of Rural Domestic Sewage Terminal. Changxing is a county located in Huzhou City, Zhejiang Province, with a total area of 1430 sq. km. It has a subtropical monsoon climate, with an average annual temperature ranging from $14^\circ\text{C} \sim 22^\circ\text{C}$. According to the official data, there are more than 0.27 million residents living in the rural region. Domestic sewage in this region is mainly treated by small-scale A^2O treatment terminals. To have a full mastering of the current performance and preparing for the next round terminal upgrading, a survey was conducted from March to April, 2018. A total of 136 A^2O rural sewage treatment terminals were investigated.

2.2. Analysis of Water Quality and Selection of Inputs. Influent and effluent water samples were carefully collected at each terminal and stored in a -20° fridge until analysis. NH_4^+-N , TP, TN, and COD were determined by the HACH Kit (HACH, USA). Conductivity (DDSJ-308A, INESA, China), pH (HQ11d, HACH, USA), and turbidity (2100Q, HACH, USA) were measured by an online parameter. Pollutant removal efficiency is computed according to the following formula:

$$\text{Removal Efficiency} = \frac{(\text{Influent concentration} - \text{effluent concentration})}{\text{Influent concentration}}. \quad (1)$$

The parameters that are significantly correlated with effluent TN and COD are screened out through IBM SPSS statistics 24. Then, principal component analysis (PCA), subtractive clustering algorithm (SCA), and fuzzy c-means algorithm (FCM) are used in this study to further determine dimensions of inputs [24–26]. Initially, PCA is used to ensure the importance of inputs and minimize the redundancy problems caused by massive strongly intercorrelated data. Then, SCA and FCM are used to determine the number of clusters and clustering centers of outputs and inputs, respectively. Eventually, clustering centers of these inputs and outputs are treated by the Johnson Algorithm in the Rosetta Software to determine the input dimensions.

2.3. Methodology of ANN, ANN-GA, and ANN-PSO

2.3.1. ANN. Figure 1 shows the typical structure of classical ANN. Briefly, the ANN model consists of several layers, and according to their distinctive layers, neurons can be subdivided into input, hidden, and output neurons. Hidden layers, serving as feature detectors to introduce nonlinearity into the network, can be either single or multiarchitecture, depending on the case need. The construction of the ANN model includes training (input feed forward and error back propagation) and validation.

(i) *Input Feed Forward.* Simplified feed forward calculations are as follows [10]:

Hidden neurons receive signals from the input neurons through a set of specific weights, thresholds, and transferring functions as follows [10]:

$$b_j = F \left(\sum_{i=1}^m a_i \times w_{ij} + P_j \right). \quad (2)$$

Again, the signals are passed to the output neuron and form the final predicted values (output neurons) as follows [10]:

$$c' = F' \left(\sum_{j=1}^n v_j \times b_j + Q \right), \quad (3)$$

where a_j represents the value of the input neuron; b_j represents the value of the hidden neuron; w_{ija} and v_j are the weights between the input neuron a_i and hidden neuron b_{ij} and hidden neuron b_j and output neuron, respectively; P_j and Q are the connection thresholds of the hidden neuron and output neuron, respectively; F and F' means the transfer function from input neurons to hidden neurons and hidden neurons to output neurons, respectively. c' is the predicted value of effluent TN or COD concentration. Initially, w_{ij} , P_j , v_j , and Q are all randomly selected small values and will be readjusted in the latter feedback works.

(ii) *Error Back Propagation.* The core of back propagation lies in redistribution of errors from the output layer to the former layer and readjustment of the parameters like weight and connection threshold accordingly. After certain iterations of back propagation, the error will be minimized, and the model will obtain a better fitness. In this study, the Levenberg–Marquardt Algorithm is adopted as the network training function for the update of previous parameters with regard to its fast computing speed and outstanding training ability. Models that are only established under the circumstances of the mean square error (MSE) is small enough [10],

$$\text{MSE} = \frac{1}{m} \sum_{k=1}^m (c'_k - c_k)^2, \quad (4)$$

where c' and c stand for the predicted value and measured value, respectively. m is the number of samples.

(iii) *Validation Procedure of Models.* Validation is the last important procedure to retest the reliability after model establishment. Subsequent model applications can be only carried out under the circumstances that the results of validation fit expectations.

2.3.2. ANN-GA. As aforementioned, to prevent the models trapped by local optima, GA and PSO are used for the selection of suitable initial weights and thresholds for ANN (Figure 2). The idea of GA was derived from the principles of natural selection and genetics. It treats the parameters (initial weights and thresholds) that need optimization as chromosomes. Chromosomes with high fitness will be selected, and others will be replaced by genetic propagation like crossover and mutation [28]. It is reported that GA is very good at global searching, independent of the initial value to achieve the convergence. However, compared to PSO, complicated processes like crossover and mutation will slow down the convergence rate of GA [15]. The brief methodology of GA can be made as previous studies and method description partly reproduces their wording [27]:

- (1) Start ANN and obtain the corresponding initial weight and threshold. These parameters are subsequently encoded into binary strings to form chromosomes.
- (2) Compute the fitness coefficient of each chromosome and retain the ones with high fitness.
- (3) Use crossover and mutation to treat rest chromosomes.

Crossover operator [27]:

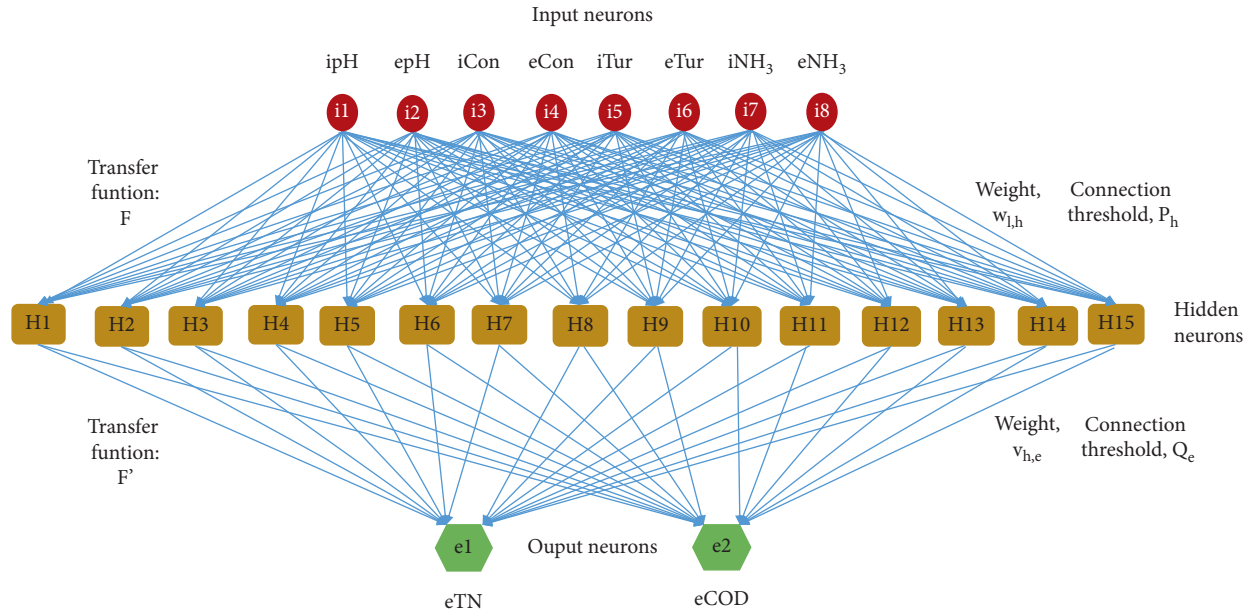


FIGURE 1: Typical structure of the ANN neural network.

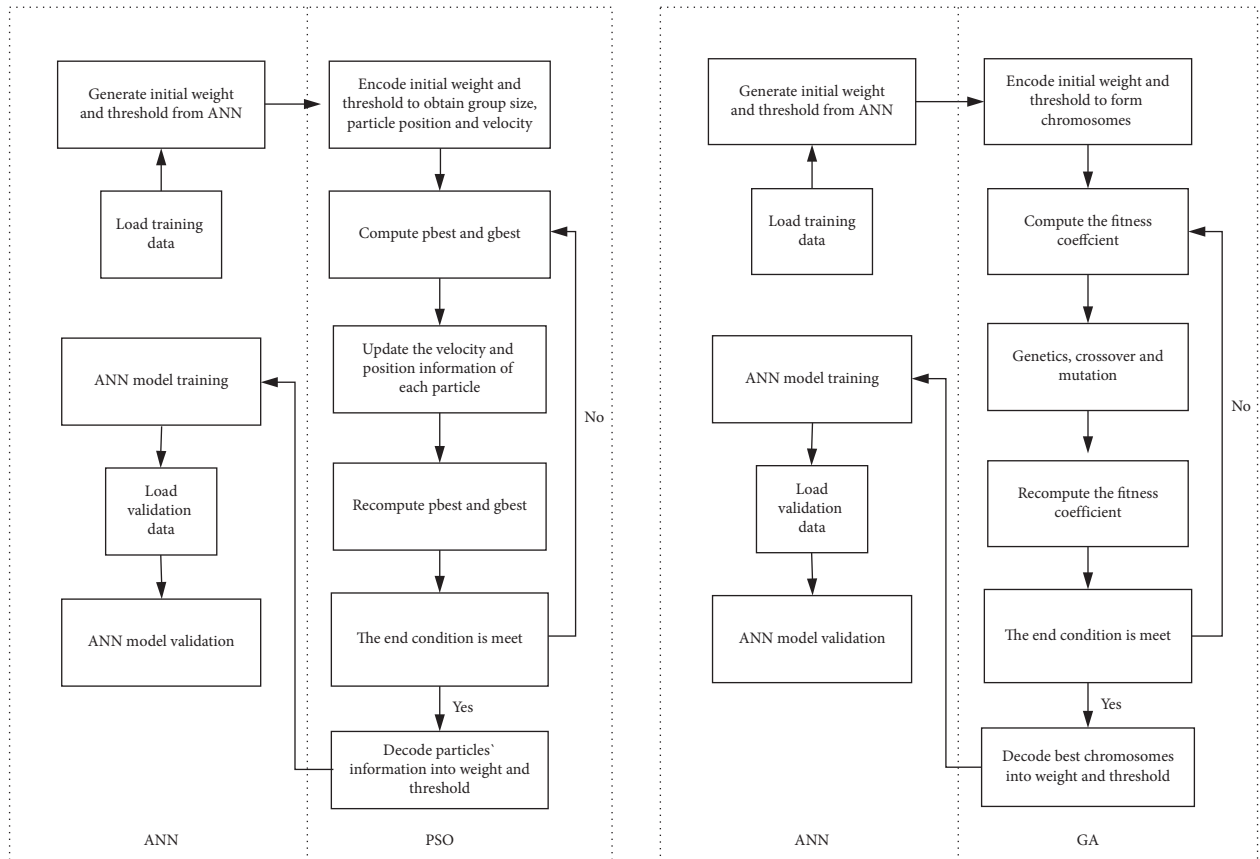


FIGURE 2: Flowchart of ANN-GA and ANN-PSO [27].

$$\begin{aligned} A_{KJ} &= A_{KJ}(1 - B) + A_{LJ}B, \\ A_{LJ} &= A_{LJ}(1 - B) + A_{KJ}B, \end{aligned} \quad (5)$$

where A_{KJ} and A_{LJ} are the K_{th} and the L_{th} chromosomes; B is the random value from 0 to 1.

Mutation operator [27]:

$$Q_{IJ} = \begin{cases} Q_{IJ} + (Q_{IJ} - Q_{\max}) \cdot R_2 \cdot \left(1 - \frac{g}{G_{\max}}\right), & \alpha > 0.5, \\ Q_{IJ} + (Q_{\min} - Q_{IJ}) \cdot R_2 \cdot \left(1 - \frac{g}{G_{\max}}\right), & \alpha \leq 0.5, \end{cases} \quad (6)$$

where Q_{IJ} is the J_{th} gene of the I_{th} chromosome; Q_{\max} and Q_{\min} are the maximum and minimum of gene Q_{IJ} ; g is the current iteration time; R_2 is a random number; G_{\max} is the maximum iteration time; and α is a random number ranging from 0 to 1.

- (4) Repeat step 2 until obtaining chromosomes with the best fitness after several iterations. Decode the chromosomes and replace the initial weights and threshold of the ANN model with these optimized ones.

2.3.3. ANN-PSO. PSO is a modern heuristic algorithm derived from natural foraging and swarming of birds or fish [17]. The bases of PSO is built on the team cooperation and information sharing [29]. The algorithm treats the parameters that need optimization (like initial weights and thresholds) as particles. Each particle represents an individual solution, and the swarms of particles show the whole solution space. The individual particle is not only aware of the position of itself and others, but also searches the solution space through its present velocity, previous experience, and the experience of its neighbour particles [16]. Hence, apart from fast convergence, PSO also has advantages in remembering particles' best location. However, as velocity, a key parameter for searching process, is lack of dynamic adjustment, PSO sometimes will lead to the consequences of difficult convergence and low convergence accuracy [15]. The following methodology of PSO-ANN has been obtained from previous studies, and the method description partly reproduces their wording [27].

- (1) Start ANN and obtain the corresponding initial weight and threshold. These parameters are subsequently encoded into particles of a group, and each particles get their corresponding position (e_p) and velocity (f_p) information [27],

$$\begin{aligned} e_p &= (e_{p1}, e_{p2}, e_{p3}, \dots, e_{ph}), \\ f_p &= (f_{p1}, f_{p2}, f_{p3}, \dots, f_{ph}), \end{aligned} \quad (7)$$

where h means the dimension of space.

- (2) Determine the fitness of each particle (p_{best}) and compare it to the best historical value of p_{best} .

- (3) Evaluate the overall fitness of the group (g_{best}) and compare it with the best historical value of the g_{best} .
- (4) Update the velocity and position information of each particle by the following formula [27]:

$$\begin{aligned} f_{ph} &= \omega \cdot f_{ph} + h_1 \cdot \text{Rand}_1 \cdot (p_{bestph} - e_{ph}) \\ &\quad + h_2 \cdot \text{Rand}_2 \cdot (g_{bestph} - e_{ph}), \\ e_{ph} &= e_{ph} + f_{ph}, \end{aligned} \quad (8)$$

where Rand_1 and Rand_2 are two uniform random functions, and h_1 and h_2 are the learn rates

- (5) Repeat step 2 until the particles with the best fitness after several iterations are obtained. Replace the initial weights and threshold of the ANN model with these optimized ones.

2.3.4. Modeling Performance Criteria. The root mean square error (RMSE), the coefficient of determination (R^2), mean absolute percentage error (MAPE), and nash sutcliffe efficiency coefficient (NSEC) are the four criteria to evaluate model precision from different aspects [30, 31],

$$\text{RMSE} = \sqrt{\frac{\sum_{k=1}^m (c'_k - c_k)^2}{m}},$$

$$R^2 = \left[\frac{m \sum_{k=1}^m c_k c'_k - (\sum_{k=1}^m c_k)(\sum_{k=1}^m c'_k)}{\sqrt{[m \sum_{k=1}^m c_k^2 - (\sum_{k=1}^m c_k)^2] \times [m \sum_{k=1}^m c_k'^2 - (\sum_{k=1}^m c'_k)^2]}} \right]^2,$$

$$\text{MAPE} = \frac{1}{m} \left| \sum_{k=1}^m \frac{c_k - c'_k}{c_k} \right| \times 100, \quad (9)$$

$$\text{NSEC} = \left(\frac{\sum_{k=1}^m (c'_k - c_k)^2}{\sum_{k=1}^m (c_k - \bar{c})^2} \right), \quad (10)$$

where \bar{c} stands for the measured value.

2.3.5. Index Contribution and Sensibility Analysis. For a better description of the contribution from each input parameter within models, the importance of each input parameter is computed by the subsequent formula from the perspective of the weights of the input neurons [32],

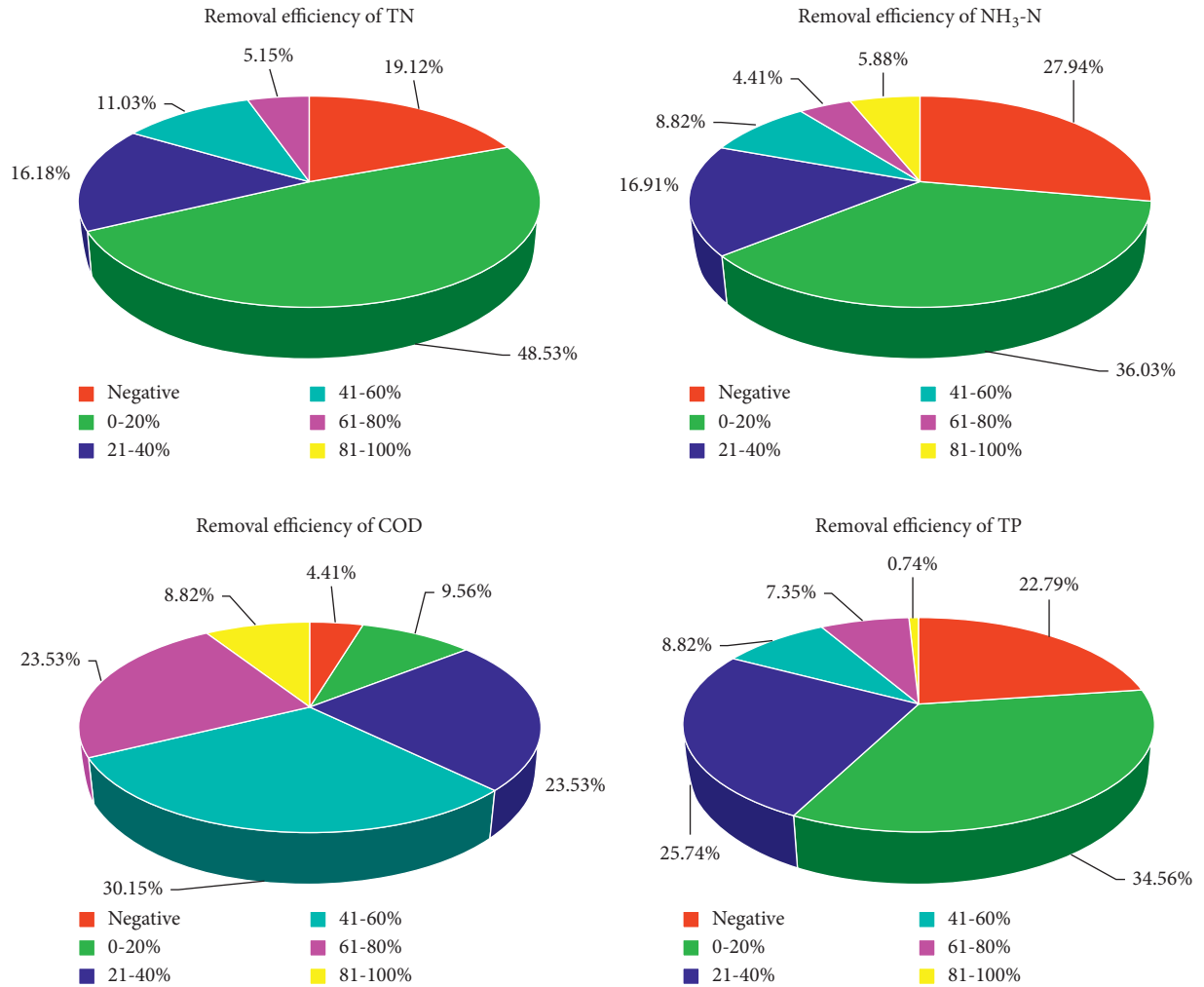
$$C_i = \frac{\sum_{j=1}^{n_h} \text{ABS}(w_{ij})}{\sum_{l=1}^{n_v} (\sum_{j=1}^{n_h} \text{ABS}(w_{lj}))}, \quad (11)$$

where C_i stands for the contribution index of the input i ; n_h stands for the number of hidden neurons; n_v stands for the number of input variables; w_{ij} stands for the weight of the input layer to the hidden layer; and ABS represents the absolute value of function.

The Morris screening method is used to identify the sensibility of the model to each input from the perspective of the prediction outcome [33]. Briefly, the sensitivity of a certain input parameter will be evaluated by increasing or

TABLE 1: Influent and effluent parameter of the rural domestic sewage terminal.

Parameter	Min		Max		Average	
	Influent	Effluent	Influent	Effluent	Influent	Effluent
pH	6.2	6.3	8.56	8.67	6.88	6.96
Conductivity ($\mu\text{s}/\text{cm}$)	248	50.6	1479	1502	972.78	948.7
Turbidity (NTU)	4	1.71	435	98.6	52.55	22.46
NH ₃ -N (mg/L)	4.4	1	131	156	53.41	45.98
TN (mg/L)	22	11	130	141	68.32	57.92
TP (mg/L)	0.3	0.37	11.03	11	5.187	4.248
COD (mg/L)	27	3	659	335	208.92	103.43

FIGURE 3: Removal efficiency of the terminals for TN, NH₃-N, COD, and TP.

decreasing its value by 10% and keeping others intact and seeing how the model will react to the change [33],

$$d_{i,k} = \frac{\text{input}_{b,k}(c_k'' - c_k')}{(\text{input}_{pc,k} - \text{input}_{b,k})c_k''} \quad (12)$$

$$\mu_i = \frac{1}{m} \sum_{k=1}^m d_{i,k},$$

where input_b refers to the original input value; input_{pc} refers to the proportional change in the original input value; c' is the original outcome of the model; c'' is the model reaction to the corresponding changes of input c ; and μ_i is the sensibility index of each input.

All the aforementioned processes are performed in IBM SPSS statistics 24, Matlab R2017b, Excel 2016, and AutoCAD 2019.

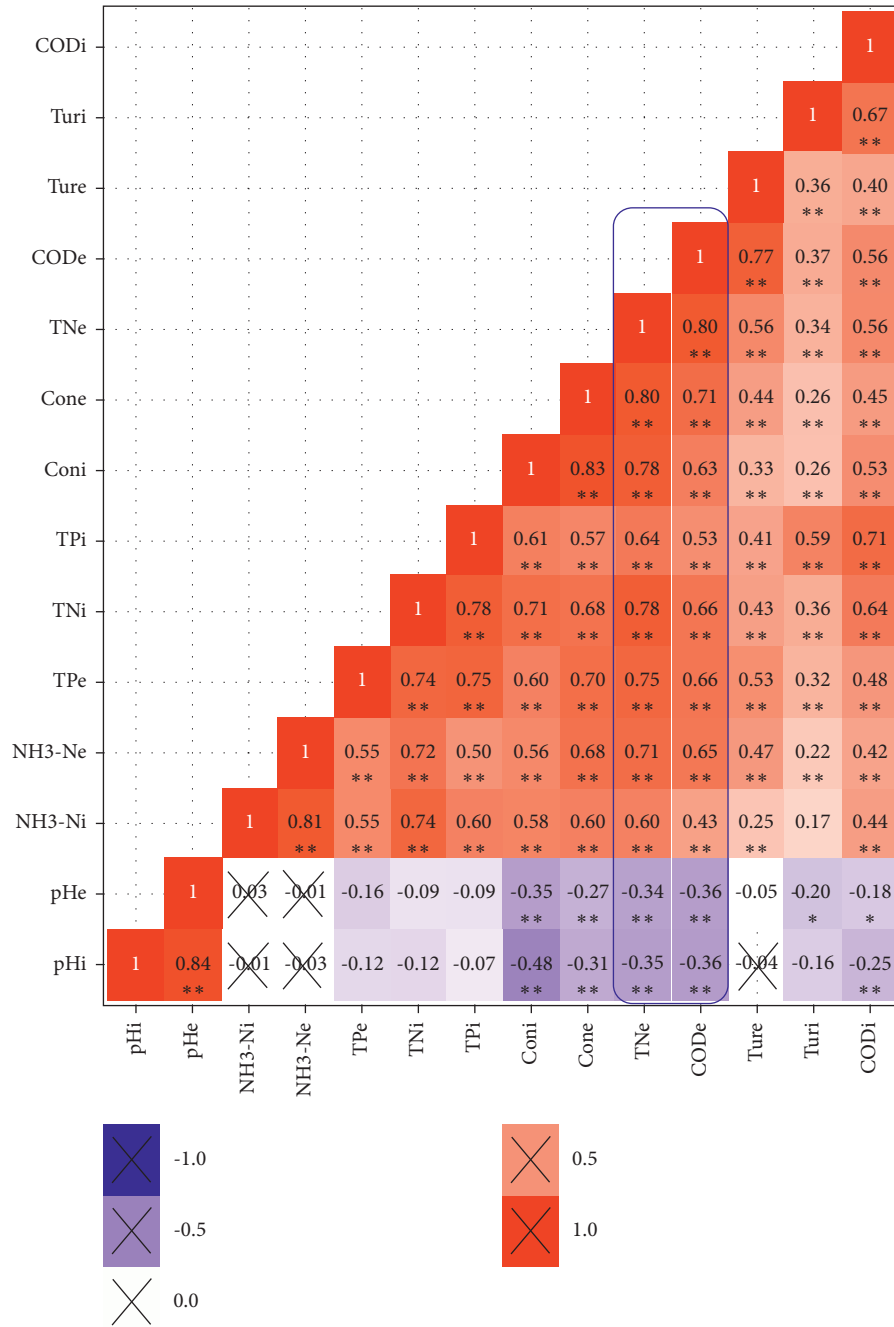


FIGURE 4: Correlation coefficient (R^2) matrix of influent and effluent parameters (number with the single-asterisk indicates significant correlation, and two-asterisks indicate extremely significant correlation).

3. Results and Discussion

3.1. Performance of Rural Domestic Sewage Terminal. Seven vital water parameters were measured and listed in Table 1. Indeed, the average $\text{NH}_3\text{-N}$, TN, TP, and COD concentrations reached up to 53.41 mg/L, 68.32 mg/L, 5.19 mg/L, and 208.92 mg/L, respectively, in the influent, which can be bracketed with or even higher than the pollutant load in some WWTPs [34]. The average $\text{NH}_3\text{-N}$ concentration is very close to the average TN concentration, implying that ammonia nitrogen dominates the nitrogen

form in the RDS. Besides that, substantial differences are demonstrated among influents from different terminals. Discrepant regional customs and dilution effect from various factors like rainfall contribute to these differences.

Figure 3 shows that the terminals had relatively limited power for pollutant removal. The average removal efficiency of turbidity, $\text{NH}_3\text{-N}$, TN, TP, and COD were only 11.18%, 16.09%, 13.31%, and 46.39%, respectively. Negative removal efficiency of these pollutants occasionally occurred on some terminals due to factors like releasing of bulking sludge [35]. Similarly, Yu et al. [36] identified about 29% of RDS

TABLE 2: The results of PCA analysis.

Components	Characteristic value	Contribution rate, %	Accumulated contribution rate, %
1	3.582	44.772	44.77
2	1.878	23.478	68.25
3	1.036	12.946	81.2
4	0.624	7.797	88.99
5	0.426	5.326	94.32
6	0.211	2.638	96.96
7	0.146	1.823	98.78
8	0.098	1.22	100

TABLE 3: Clustering centers of inputs and outputs.

Parameter	Cluster centers
Effluent TN (mg/L)	59
Effluent COD (mg/L)	128
Influent pH	6.64
Effluent pH	6.75
Influent ammonia concentration (mg/L)	56.5
Effluent ammonia concentration (mg/L)	53.5
Influent conductivity (us/cm)	976
Effluent conductivity (us/cm)	928
Influent turbidity (NTU)	39.9
Effluent turbidity (NTU)	10.9

TABLE 4: Parameters setting in ANN, ANN-GA, and ANN-PSO.

Model	Common setting for ANN	Others
ANN		None
ANN-GA	Inputnum = 8 (number of input neurons); Hiddennum = 15 (number of hidden neurons); Outputnum = 1 (number of output neurons); Net.trainParam.epochs = 20000; Net.trainParam.lr = 0.01 (learning rate); Net.trainParam.goal = 0.000004 (training goal); Purelin, transfer function from input layer to hidden layer; Tansig, transfer function from hidden layer to output layer;	Maxgen = 100 (number of Iterations); Sizepop = 10 (size of population); Pcross = 0.7 (possibility of crossover); Pmutation = 0.1 (possibility of mutation) Maxgen = 100 (number of Iterations); Sizepop = 10 (size of population); Vmax = 1 (maximum speed); Vmin = -1 (minimum speed); Popmax = 5 (maximum population); Popmin = -5 (minimum population)
ANN-PSO		

terminals in Jiaying (another city in Zhejiang), which were in the ineffective operation. The following reasons are speculated for unsatisfied performance: (1) unstability of biochemical reaction; (2) relatively limited maintenance in light of the massive amounts of terminals; (3) traditional chemical measurement cannot satisfy the need of real-time assessment as it requires intensive time for the digestion of pollutants [37]. Failure to evaluate the performance of terminals on time will let the problematic terminals fall into a worse situation. Every year, local governments have to bear great financial burdens and put considerable amount of economic and human resources into more surveillance. Finding an easier and quicker monitoring approach is an urgent desire at present.

3.2. Selection of Input Parameters. Significant correlations between some critical water parameters have been referred in the previous studies. Some easily detectable parameters can serve as rough surrogates for pollutant concentration or problems during the operation. For instance, the study of Yu et al. [36] demonstrated that conductivity was significantly correlated with TN, NH_4^+ -N, TP, and COD within both the influent and effluent. Thus, a low correlation between conductivity and TN might imply the leakage of sewer transporting system. Analogously, strong correlations were found between turbidity and parameters like TN and COD [38]. The study of Slaets et al. [39] showed that turbidity is a reliable and cost-effective predictor variable for the linear mixed model developed to account for TN. Apart from

TABLE 5: Weights and thresholds of ANN, ANN-GA, and ANN-PSO in modelling TN and COD.

ANN in modelling TN												ANN in modelling COD											
W _{ij}						V _j						W _{ij}						V _j					
-1.49	-0.23	1.07	0.62	-1.09	1.60	-1.65	-2.44	-0.47	4.06	-1.20	-1.94	-0.54	2.22	-0.21	1.37	0.44	-1.67	0.30	4.42				
-0.47	-0.63	0.20	0.22	-1.83	-1.35	-2.37	1.98	0.90	3.40	-0.53	-2.31	-0.81	0.69	-1.95	1.24	-1.96	0.14	0.58	3.32				
0.99	1.54	3.66	-1.44	2.00	1.23	-1.54	-1.28	2.17	-3.11	-0.66	-1.21	-0.68	-0.54	0.76	-3.24	0.68	-2.42	-1.17	2.32				
1.80	-0.77	0.96	0.01	2.50	-0.67	-1.05	-2.19	0.89	-2.80	-0.04	-1.81	-2.77	1.02	1.52	-0.53	0.16	0.59	-0.15	-2.50				
1.22	2.14	-2.10	0.11	0.86	-0.72	1.83	0.98	-0.09	-1.70	-0.80	-2.65	1.39	1.82	-0.12	1.96	-0.23	-0.43	0.42	-0.44				
2.79	3.51	-3.27	-0.01	0.22	0.65	-0.05	0.53	-0.66	2.01	-0.74	0.20	-2.75	-0.12	1.77	-0.35	-2.35	1.77	0.39	1.16				
0.46	1.44	-1.14	0.18	1.53	-0.57	2.67	-1.60	-0.23	-1.24	-0.45	-1.34	-0.59	-2.22	0.37	-1.68	-1.97	1.14	-0.27	1.71				
-1.77	-1.51	1.64	-0.06	1.07	-2.49	2.29	-0.80	-0.45	0.01	-0.76	-3.04	1.53	-0.06	-0.11	-0.04	-1.07	1.84	0.22	0.24				
1.37	-1.17	0.43	-3.65	0.21	-0.09	0.10	-3.38	-0.59	-0.77	0.11	0.57	-0.16	-0.70	0.04	2.75	-0.93	2.46	-0.87	0.34				
0.76	1.44	2.72	1.39	-1.68	1.34	-1.36	-1.37	-0.26	1.50	2.45	1.44	-0.05	1.39	0.24	-2.12	1.80	-0.54	-0.38	0.65				
0.40	1.38	2.64	-0.97	1.20	3.19	2.09	0.16	0.46	-0.42	1.40	2.91	1.89	1.06	1.80	-0.94	1.47	-1.44	0.66	1.13				
-0.84	0.38	-1.88	-0.78	3.66	-0.83	0.45	1.46	-1.28	-2.79	0.14	1.81	2.12	0.35	0.64	-3.39	0.69	-2.42	-0.81	-2.26				
-0.51	1.44	1.71	0.20	-2.23	-1.17	2.24	0.32	0.21	-2.39	0.69	2.21	1.47	0.28	-0.71	0.68	2.01	0.53	0.63	3.42				
-0.68	1.32	2.60	-0.41	-0.77	-1.78	-1.71	0.27	-1.11	-5.16	-2.02	0.09	-1.85	-0.38	1.14	0.05	-2.40	1.03	-0.59	-3.49				
1.80	1.53	2.55	-0.41	2.84	-1.70	1.51	-0.92	0.08	2.98	-2.12	-0.22	-0.89	0.40	2.20	-0.64	2.18	1.36	-0.78	-3.80				

ANN-GA in modelling TN												ANN-GA in modelling COD											
W _{ij}						V _j						W _{ij}						V _j					
-3.60	1.62	0.98	-1.10	-1.75	1.78	2.33	1.26	-0.23	2.56	1.59	1.46	-1.22	-1.86	1.78	0.26	2.05	0.63	0.18	0.07				
0.57	-0.13	-0.93	-2.53	-3.16	-0.34	-0.19	1.87	-1.03	1.11	2.07	-1.35	1.43	1.09	3.27	1.88	1.26	-0.31	0.79	-2.75				
-0.89	0.20	1.35	-3.21	0.63	-1.97	-1.72	2.21	-1.26	0.82	1.72	-2.45	-3.49	-3.35	-0.75	-1.86	0.12	0.76	0.15	0.97				
-0.16	0.41	1.00	1.23	3.26	-1.15	-1.95	0.99	-1.40	-2.54	2.14	2.28	-2.46	-0.11	3.13	-0.03	2.57	1.19	1.20	-2.37				
-1.38	-2.33	-0.15	-1.83	-0.67	2.24	1.96	3.06	-0.75	-1.69	-2.59	-2.06	2.62	-0.20	-2.37	0.96	-0.07	-0.25	0.06	0.98				
0.94	-0.04	3.55	1.30	0.72	1.40	-2.60	-2.51	1.00	-2.58	-1.03	0.50	2.57	1.78	-0.03	-1.23	3.10	-1.41	0.09	1.81				
1.28	0.02	0.82	0.97	0.90	-1.64	0.26	-2.58	-1.26	2.46	-1.52	-2.76	2.57	-0.27	2.31	2.60	2.86	0.37	-0.11	-0.17				
0.01	-2.37	-2.19	-1.18	-1.47	-0.95	-2.45	0.47	-0.21	1.16	2.20	2.49	-1.95	-2.66	1.13	-0.19	0.03	-2.43	0.00	1.39				
-2.62	1.99	2.79	-2.27	0.36	2.96	-1.62	-0.98	-0.69	0.10	0.26	0.10	-2.60	-2.35	-3.06	-1.70	-1.56	1.37	-0.10	-0.56				
-1.55	0.60	1.45	-1.35	-0.44	0.65	-0.52	2.22	1.98	-0.40	-2.15	2.88	0.48	-3.24	-2.39	-1.43	1.67	-1.90	-0.06	-0.39				
0.08	-0.83	-2.87	-0.65	-1.44	1.67	1.56	1.64	-1.05	0.13	1.80	2.33	-0.26	-2.18	-0.26	-2.65	-1.50	-2.28	-0.28	-0.91				
2.38	-2.28	0.06	0.08	-2.66	-2.67	0.57	-2.40	0.88	2.17	2.93	-1.94	-2.73	1.13	1.03	-3.28	1.03	-1.35	-0.32	0.69				
-0.70	-1.00	-2.90	-2.30	0.28	-0.93	0.95	2.86	1.40	1.32	-2.88	-0.83	0.39	0.30	-2.63	-0.88	-2.41	-1.40	0.67	1.16				
0.62	-0.41	-1.33	-0.02	1.61	0.02	2.64	-0.74	0.36	-2.86	-0.22	-0.99	2.24	-0.13	-0.10	0.39	-2.19	-0.26	0.40	1.54				
-0.58	-1.75	0.82	0.36	-1.44	0.45	1.08	-3.23	-0.22	0.26	-2.33	-2.18	-2.88	1.53	-3.61	1.57	0.15	-2.42	0.01	1.01				

TABLE 5: Continued.

	ANN-PSO in modelling TN							ANN-PSO in modelling COD											
	W_{ij}							W_{ij}											
					V_j	P_j	Q					V_j	P_j	Q					
1.79	-0.10	0.74	1.93	0.17	0.84	-4.76	2.35	0.23	-3.41	0.00	-0.87	2.81	-0.48	-0.97	-1.51	1.72	-1.02	0.53	0.18
1.09	1.87	-0.19	-0.84	0.52	-1.65	4.39	1.91	0.12	0.14	-1.98	-3.06	3.02	2.25	-0.47	-2.81	-0.30	1.29	-0.03	-1.67
-0.58	-0.38	1.41	1.08	1.37	-2.12	-2.30	-1.79	-0.20	-3.28	0.15	0.54	-0.68	-1.06	-0.76	2.01	-2.59	-0.30	0.03	-1.60
-0.30	-2.75	-1.99	2.57	-2.93	0.13	-1.78	-2.24	0.27	-0.55	2.84	1.13	-0.71	1.00	0.44	0.73	-1.35	-0.50	0.17	1.10
-0.74	3.69	2.14	-0.35	-0.59	-1.62	0.93	-1.75	2.14	3.73	0.38	2.30	1.62	-0.13	-0.66	0.74	-0.80	1.97	0.10	1.94
0.27	0.82	-4.79	0.03	-0.88	1.91	-1.67	1.49	-0.01	0.85	-2.56	0.11	-0.32	0.33	0.89	-0.73	-2.31	0.04	-0.39	2.04
3.58	-3.08	-0.98	-0.23	2.68	-3.23	1.06	-3.40	-0.08	1.28	-0.48	-1.47	1.32	-2.14	-0.98	-1.38	-0.95	3.04	-0.21	-0.40
0.04	0.87	1.75	3.90	3.39	0.32	-0.47	-0.08	0.25	-0.95	1.43	0.55	-1.70	-0.98	-0.79	0.87	-1.36	2.72	0.43	-1.00
-1.02	0.46	-2.06	-3.30	3.72	-1.78	-1.80	2.03	-0.02	1.58	-2.44	-1.82	-0.06	-1.51	-0.35	1.64	1.47	0.05	-0.11	-2.22
-1.03	1.41	-0.05	-0.69	-3.22	-0.09	-0.58	3.08	-0.25	0.12	1.20	1.32	0.96	-3.06	0.67	2.04	0.96	-1.76	-0.23	-1.05
-4.77	1.84	-0.02	-0.09	3.69	-0.17	-2.13	0.91	0.20	2.86	-1.24	0.83	-0.55	1.92	-0.20	-0.42	0.86	0.32	0.13	-0.76
-1.16	2.98	-2.13	-0.08	1.79	-0.42	0.26	-0.89	-0.26	1.61	0.26	1.13	-0.08	0.05	0.07	-1.10	0.85	-1.44	-0.89	-0.29
-0.03	1.43	2.14	-0.01	2.19	4.94	0.45	-1.44	-0.02	1.95	-1.74	1.19	1.14	3.42	-0.97	0.80	1.87	-1.67	0.09	1.59
3.16	-1.63	1.23	3.20	-0.93	-3.10	-1.58	-0.70	-3.85	2.80	1.76	3.19	0.83	0.21	0.15	0.66	0.39	-0.86	-0.22	0.62
2.20	-1.14	-0.84	-1.53	-2.88	-1.22	0.89	3.13	0.22	-2.73	0.30	-0.62	-1.86	2.82	-0.78	1.29	-0.96	0.80	-0.33	0.44

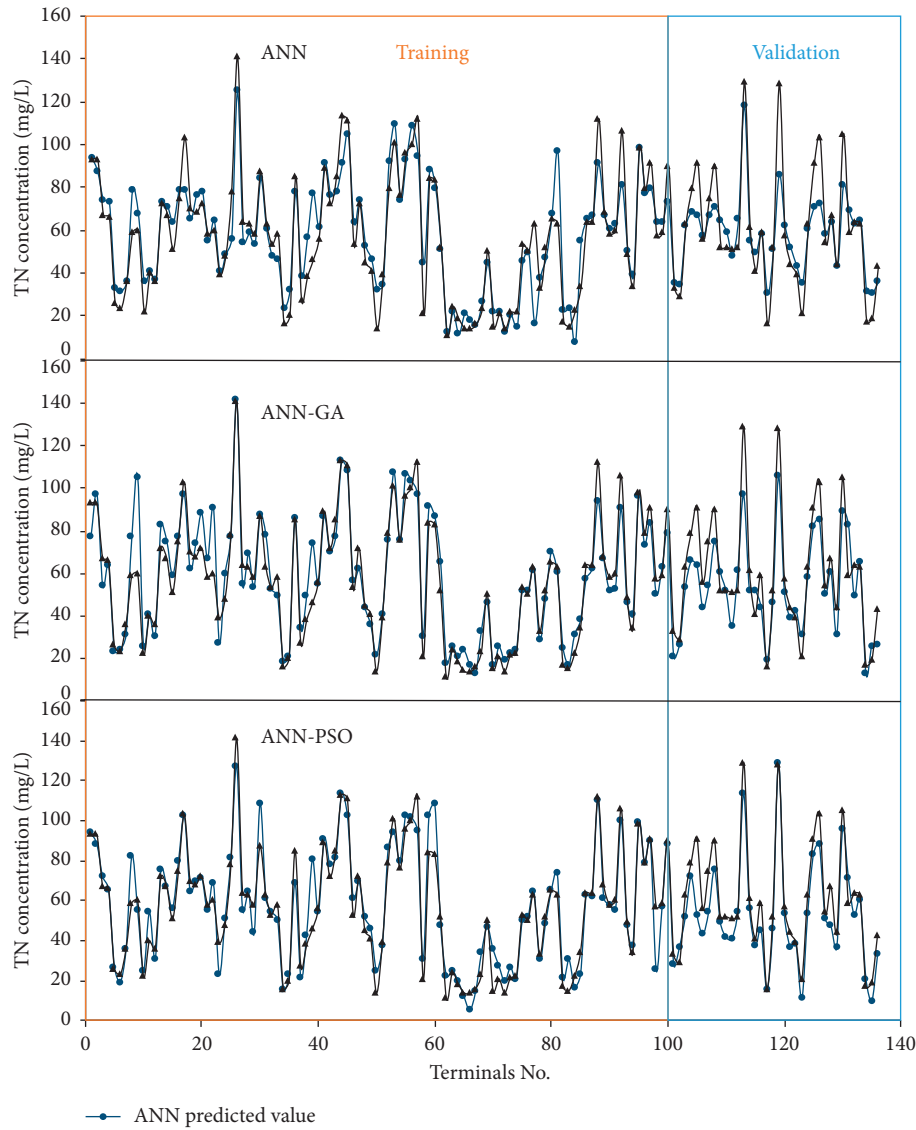


FIGURE 5: Performance of TN prediction.

conductivity and turbidity, pH had also presented a weak correlation with TN and was served as input in the ANN model to predict TN [40]. Figure 4 shows that these rules are also applicable in the field of RDS. Effluent TN of RDS displays strong correlations with influent conductivity, effluent conductivity, influent ammonia, effluent ammonia, effluent turbidity, respectively, and a weak correlation with influent pH, effluent pH, and effluent turbidity, respectively. Effluent COD of RDS exhibits strong correlations with effluent turbidity, effluent ammonia, influent conductivity, and effluent conductivity, respectively, and a weak correlation with influent ammonia, influent turbidity, influent pH, and effluent pH, respectively. Remarkably, R^2 between effluent TN and effluent conductivity can reach 0.80, indicating nitrogen might be mainly presented in the dissolved ammonia form. The high R^2 (0.77) between effluent COD

and effluent turbidity implies that particle pollutants play an important role in the effluent COD.

Correlation analyses indicate that ANN models can be developed to account for effluent TN and COD with these easily detectable parameters (pH, turbidity, conductivity, and ammonia of influent and effluent) as inputs. The results of PCA (Table 2.) show that the first principle can explain 44.77% of all variance, and first four components contain 88.99% of variance. Generally, the overall data can be characterized by components that explain more than 85% of variance [41]. SCA, FCM, and Johnson Algorithm are subsequently used to determine the dimension of inputs. Clustering centers of all parameters are shown in Table 3, and the final results of the Johnson Algorithm show that pH, turbidity, ammonia concentration, and conductivity of both influent and effluent can all act as inputs.

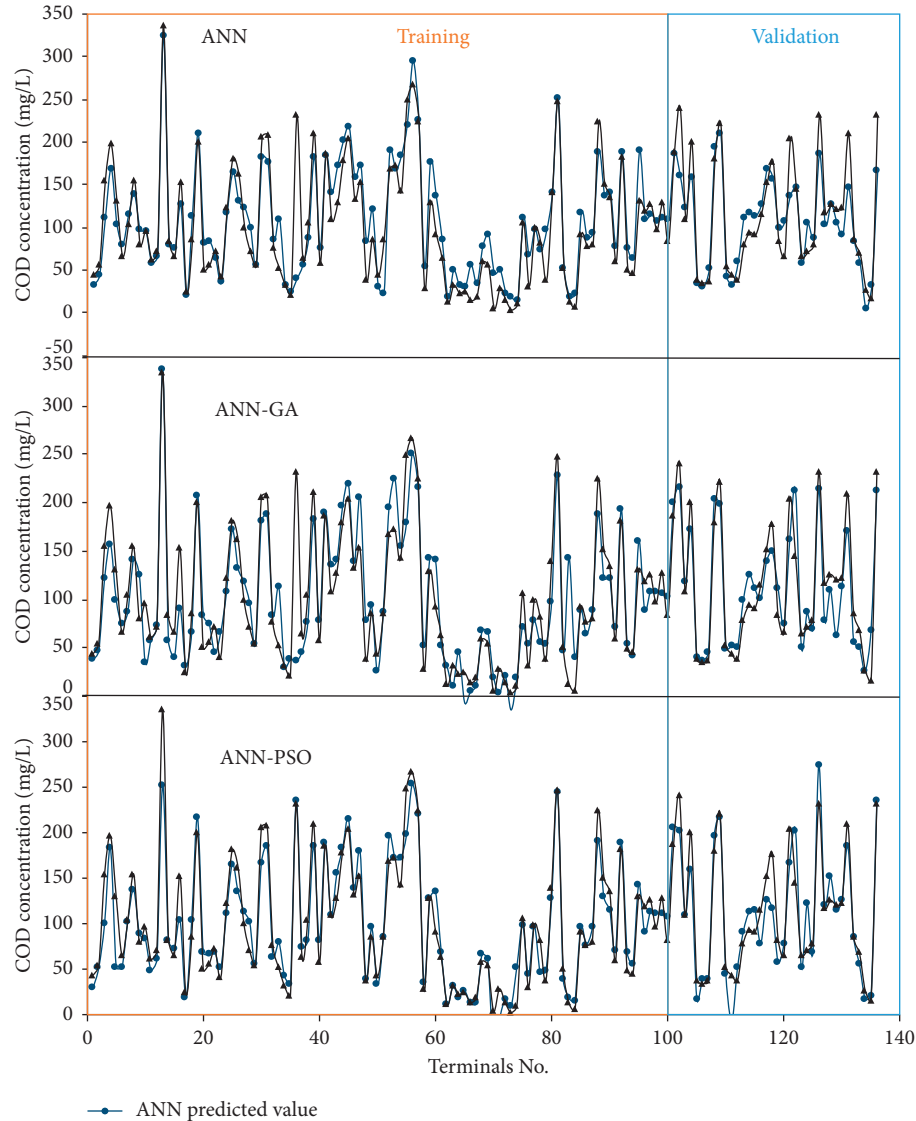


FIGURE 6: Performance of COD prediction.

3.3. ANN Prediction Performance. Fan et al. [15] concluded 44 studies that used ANN to model and optimize pollutant removal processes. In this review, most studies used about 60% to 80% of data as the training database. Accordingly, this study uses data from 100 terminals as the training database (73.53% of total), then the rest data from 36 terminals are applied to validate the performance of the model. A trial and error approach is used in this study to determine the number of hidden neurons [42, 43]. Since the standard multilayer feedforward network with one hidden layer has been considered as a universal approximator, analogously, this study also configures all models with only one hidden layer [44, 45]. Eventually, ANN, ANN-GA, and ANN-PSO models all contain three distinctive layers. A total of 8 neurons, including pH, conductivity, turbidity, and ammonia concentration of both influent and effluent, are set in the input layer, and 15 neurons are set in the hidden layer. The preset parameters, weights, and thresholds of models can be found in Tables 4 and 5.

The prediction performance of the three models for TN and COD can be seen in Figures 5 and 6, respectively. The prediction curves of the three models not only acquire the knowledge base of these terminals, but also closely capture the fluctuation trend of the true curves. As shown in Figure 7, the linear fit for ANN-PSO curves is closest to the reference line (100% accuracy), followed by the linear fit for ANN-GA curves and finally ANN curves, demonstrating ANN-PSO yields the best predicting performance for both TN and COD [27]. Table 6 shows ANN-PSO also obtained the most reliable performance in terms of the model error. The R^2 , RMSE, and MAPE of ANN-PSO in modelling TN are 0.90, 9.14, and 16.19%, respectively, in training, and 0.90, 11.54, and 16.79%, respectively, in validation. In terms of COD prediction, R^2 , RMSE, and MAPE of ANN-PSO are 0.90, 22.10, and 34.57%, respectively, in training, and 0.85, 26.57, and 22.30%, respectively, in validation. Considering that ANN-PSO models possess higher R^2 and lower RMSE than ANN models, ANN-PSO models neither get into

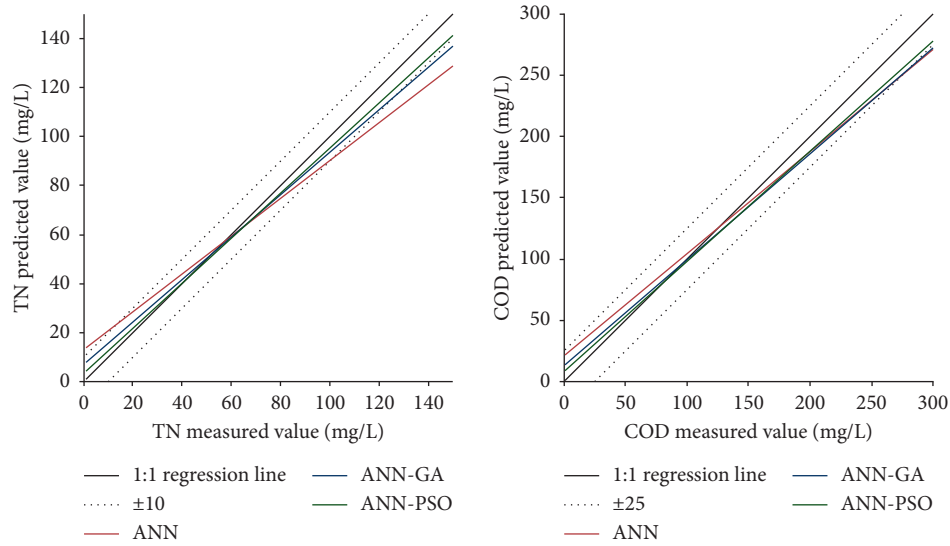


FIGURE 7: Linear regression curve of ANN, ANN-GA, and ANN-PSO models.

TABLE 6: Error analysis about ANN, ANN-GA, and ANN-PSO.

Model	Input parameters	Predicted parameters	R^2	RMSE	MAPE	NSEC
ANN	Influent pH, effluent pH, influent conductivity, effluent conductivity, influent turbidity, effluent turbidity, influent $\text{NH}_3\text{-N}$, effluent $\text{NH}_3\text{-N}$.	Effluent TN	0.84 for training; 0.85 for validation	11.40 for training; 13.49 for validation	18.99% for training; 19.93% for validation	0.95 for training; 0.95 for validation
		Effluent COD	0.81 for training; 0.82 for validation	30.88 for training; 30.00 for validation	45.42% for training; 24.54% for validation	0.80 for training; 0.80 for validation
ANN-GA		Effluent TN	0.90 for training; 0.84 for validation	9.42 for training; 13.56 for validation	14.90% for training; 20.09% for validation	0.97 for training; 0.95 for validation
		Effluent COD	0.76 for training; 0.85 for validation	34.31 for training; 26.25 for validation	53.92% for training; 26.91% for validation	0.76 for training; 0.84 for validation
ANN-PSO		Effluent TN	0.90 for training; 0.90 for validation	9.14 for training; 11.54 for validation	16.19% for training; 16.79% for validation	0.97 for training; 0.97 for validation
		Effluent COD	0.90 for training; 0.85 for validation	22.10 for training; 26.57 for validation	34.57% for training; 22.30% for validation	0.90 for training; 0.84 for validation

overfitting nor underfitting after optimization. In addition, this study uses NSEC to evaluate the predictive power of models. Theoretically, NSEC ranges from $-\infty$ to 1. 0 indicates that the prediction performance of the model is close to the mean of the measured value; in the other words, the overall result is credible and 1 indicates that the model is in the perfect prediction. The closer NSEC is to 1, the more accuracy models can reach [30]. The NSEC of ANN-PSO in modelling TN are both 0.97 for training and validation and NSEC of ANN-PSO in modelling COD are 0.89 and 0.84, respectively, for training and validation, showing strong prediction power of ANN-PSO. Except the accuracy advantage, ANN-PSO shows superiority in terms of computational time. It takes ANN-PSO less than 1 min for 100 iterations of model convergence, while it takes ANN-GA about 6 min to do the same work.

3.4. Contributions and Sensibility Analysis of Each Input.

Contributions of inputs in the ANN-PSO models are calculated in Figure 8. In the ANN-PSO modelling TN, the indices range from 10.16% to 15.84% among parameters. Influent turbidity makes the biggest contribution to TN prediction. Although the ANN model is often regarded as a black box, lacking a direct mechanism to demystify the interrelationship between neurons, the contribution results strongly suggests that inputs like influent turbidity play more important roles than others in the ANN-PSO modelling TN [32, 46]. While, in the ANN-PSO modelling COD, the contributions of inputs range from 6.43% to 15.00%. Inputs like effluent conductivity and pH significantly participate in the COD prediction.

Morris screening is used to identify the sensibility of each input for ANN-PSO models (Table 7) [33]. Accordingly, the

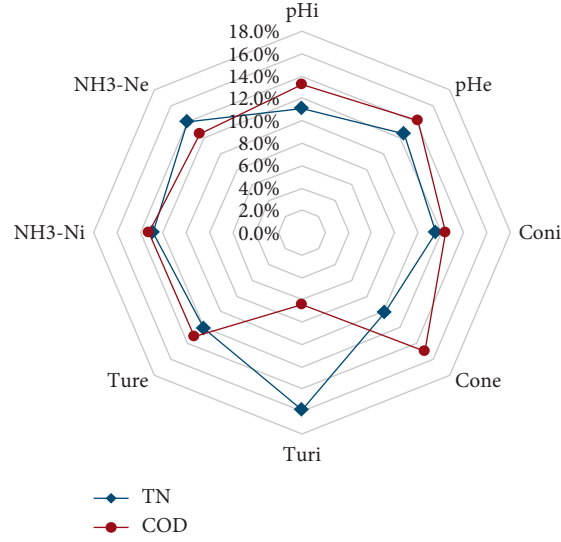


FIGURE 8: Contribution analysis of ANN-PSO in modelling TN and COD.

TABLE 7: Sensibility analysis of ANN-PSO in the modelling of TN and COD.

TN	Upper 10%	Lower 10%	Rank
e_pH	3.08	4.49	1
i_PHi	2.34	2.71	2
i_Con	0.64	0.65	3
e_Con	0.58	0.65	4
e_NH3	0.41	0.38	5
i_NH3	0.33	0.29	6
e_Tur	0.28	0.14	7
i_Tur	0.1	0.1	8
COD			
e_pH	5.57	6.17	1
i_PHi	3.01	2.68	2
i_Con	1.33	1.45	3
e_Con	1.18	1.18	4
i_NH3	0.57	0.56	5
e_NH3	0.41	0.38	6
e_Tur	0.35	0.35	7
i_Tur	0.11	0.11	8

sensibility index (μ_i) higher than 1 implies that the outcomes of the model exhibit more drastic changes than the corresponding changes of inputs. Therefore, in the ANN-PSO modelling TN, only effluent and influent pH cause larger change to the model. While, in the ANN-PSO modelling COD, not only effluent and influent pH, but also influent and effluent conductivity yield μ_i higher than 1. The sensibility results show that, in both two models, effluent and influent pH are the most sensible inputs, the second most sensible inputs are influent and effluent conductivity and $\text{NH}_3\text{-N}$, the less sensible inputs are effluent and influent turbidity.

3.5. Advantages, Limitations, and Recommendation for Future Works. Table 8 summarizes some previous successful studies. By comparison, R^2 of the two ANN-PSO models in this study (0.85 to 0.90) are at a similar level to that in the previous studies (about 0.70 to 0.99). One shortcoming of this study lies in our relatively high RMSE. The sharp fluctuation of TN and COD cannot be ignored for this issue. For example, effluent COD mostly fluctuated within 10 to 60 mg/L in the study of Luo et al. [23]. In contrast, the range is magnified to 3–335 mg/L in this study. Great heterogeneity among these terminals will inevitably introduce new errors into the models and make the models slightly lose

TABLE 8: ANN application cases in the previous works.

Application scenario	Model	Input parameters	Predicted parameters	R^2	Error	Authors
WWTP	Adaptive neuro fuzzy inference system	Influent pH, influent temperature, influent SS, and influent COD.	Effluent SS	0.88 for training; 0.89 for validation	1.87 for training; 0.71 for validation**	[34]
			Effluent COD	0.79 for training; 0.83 for validation	3.83 for training; 2.47 for validation**	
			Effluent pH	0.93 for training; 0.93 for validation	0.07 for training; 0.05 for validation****	
Fenton reactor	Artificial neural network	Reaction time, pH, antibiotic concentration, H_2O_2 , etc	COD removal efficiency	0.997	0.000376*	[47]
WWTP	Artificial neural network-feed forward back propagation	COD, BOD, TSS	COD	0.93 for training; 0.86 for validation	—	[48]
Anoxic/oxic reactor	Artificial neural network	Influent COD, reflux ratio, carbon-nitrogen ratio	Effluent COD	0.98596	2.82*	[49]
	Artificial neural network-genetic algorithm			0.99476	1.12*	
The upflow anaerobic filter reactor	Multilayer perceptron neural network	Influent chemical oxygen demand, hydraulic retention time, and influent cyanide concentration	Effluent COD	0.983 for training; 0.876 for validation	104.75 for training; 98.35 for validation**	[31]
	Radial basis neural network			0.995 for training; 0.708 for validation	56.81 for training; 157.95 for validation**	
	Regression neural network			0.951 for training; 0.751 for validation	175.19 for training; 140.51 for validation**	
Lakes	Artificial neural network	NH4-N, secchi depth	TN	0.72 for training; 0.69 for validation	0.15 for training; 0.14 for validation**	[50]
Partial-nitrification/anammox reactor	Back propagation artificial neural network	Influent NO_3^- , NO_2^- , pH and effluent NO_3^- , NO_2^- ,	TN removal efficiency	0.889	0.091*	[51]
WWTP	Fuzzy rough-back propagation	Return ratio, sludge volume, MLSS,	Effluent COD	0.921	2.86**	[23]
	Principal component analysis-back propagation	DO, Q	Effluent TN	0.882	1.35**	
		(flow), effluent suspended solid,	Effluent COD	0.886	3.15**	
	Back propagation-neural networks	influent TP, influent TN, influent COD,	Effluent TN	0.862	1.52**	
		influent ammonia concentraion.	Effluent COD	0.813	3.46**	
			Effluent TN	0.822	1.97 **	

In the error column, numbers with one asterisks is MSE, and numbers with two asterisks is RMSE.

their edges in precision. However, compared with previous studies, two ANN-PSO models in this study are both available for different terminals and do not require historical data from terminals, which obviously save a lot of time and energy and be more practical even at the cost of sacrificing certain degree of precision. Another special advantage of this study is that the inputs are easier to be obtained and can all be measured by electrodes (NH_3-N was measured by

traditional chemical methods in this study, but it can be also measured by using the ammonia gas sensing electrode [52]).

Based on the above findings, this study has the following two recommendations for future works: 1. As shown in Figure 9, use electrodes to collect input data and realize remote online prediction of effluent water quality based on ANN-PSO, which has not been done before. 2. Since biological treatment is greatly influenced by procedure

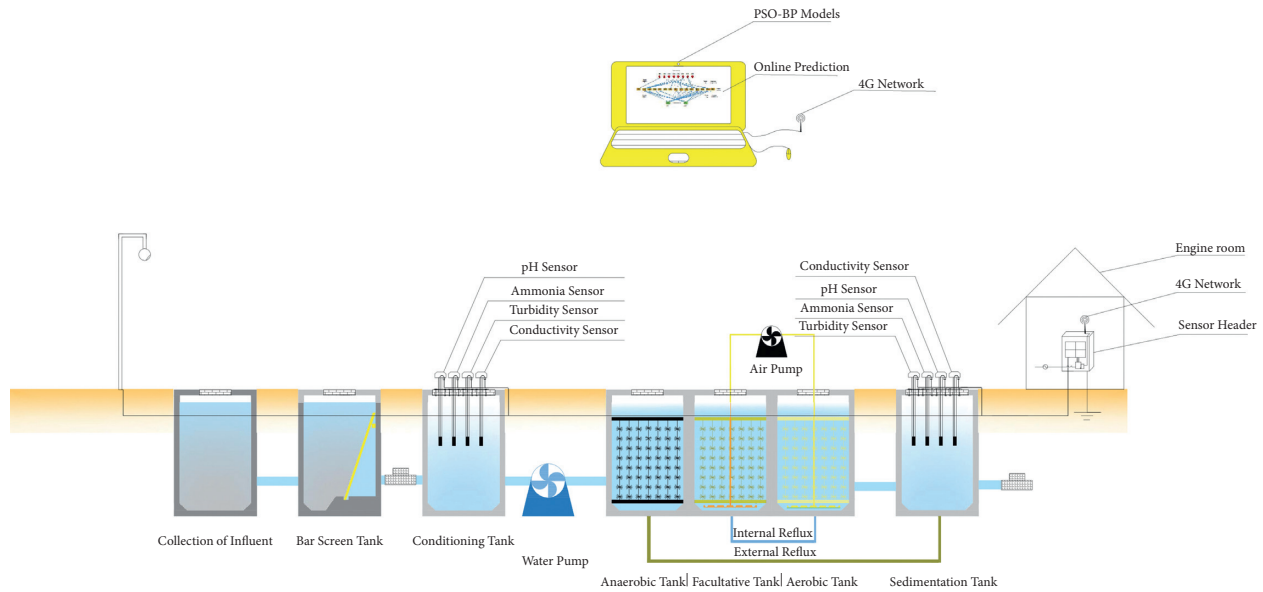


FIGURE 9: Online prediction of TN and COD based on ANN-PSO models.

variables, like DO of aerobic tank [5], future studies can try to use some procedure variables as inputs for the improvement of model accuracy.

4. Conclusion

Complicated influent situation and unsatisfying treatment performance of large numbers of rural domestic sewage terminals highlight the urgent need to find a quicker and simpler effluent measurement. Significant correlations are found between some easily detectable parameters (e.g., conductivity and turbidity) and effluent TN and COD, which triggers the idea of using these easily detectable parameters as inputs to predict effluent TN and COD in the ANN models. The results turn out that the ANN models can successfully simulate the effluent TN and COD with R^2 both higher than 0.8. Then, GA and PSO are used as two optimization strategies to improve the ANN performance. By comparison, ANN-PSO yields the better prediction capacity for both TN and COD. R^2 and RMSE of ANN-PSO on modelling TN are 0.90 and 9.14, respectively, in the training, 0.90 and 11.54, respectively, in the validation. R^2 and RMSE of ANN-PSO on modelling COD are 0.90 and 22.10, respectively, in the training, 0.85 and 26.57, respectively, in the validation. Contribution analysis shows that influent turbidity and effluent conductivity make the biggest contribution to ANN-PSO on modelling TN and COD, respectively. Sensibility analysis shows that effluent and influent pH are the two most sensible inputs for both two models. In the end, considering that all inputs can be detected by the electrodes, this study also proposes an ANN-PSO-based remote online water quality monitoring approach.

Abbreviations

RDS: Rural domestic sewage
TN: Total nitrogen

ML: Machine learning
ANN: Artificial neural network
BP: Back propagation
GA: Genetic algorithm
PSO: Particle swarm optimization
RMSE: Root mean square error
MAPE: Mean absolute percentage error
NSEC: Nash sutcliffe efficiency coefficient.

Data Availability

Publication of data required the permission from all team members. The data are not published for now and will be available later.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors' Contributions

Lin Qiang, Luo Ancheng, and Zhang Yan investigated the study; developed the methodology; wrote the original draft, and reviewed and edited the article; provided the software; and performed data analysis and curation. Wang Yunlong investigated the study; developed the methodology; and supervised the study, and reviewed and edited the article. Liang Zhiwei developed the methodology; supervised the study, and reviewed and edited the article. Yuan Ping performed data analysis and curation. Lin Qiang, Luo Ancheng, and Zhang Yan contributed equally to this work.

Acknowledgments

This work was mainly supported by the Research and Demonstration of Integrated Rural Domestic Wastewater Reclamation Technology, China (2019YFC0408803).

Meanwhile, the Project was also supported by the Scientific Research Fund of Zhejiang Provincial Education Department (188310-542122/002/013).

References

- [1] F. Cheng, Z. Dai, S. Shen, S. Wang, and X. Lu, "Characteristics of rural domestic wastewater with source separation," *Water Science and Technology*, vol. 83, no. 1, pp. 233–246, 2021.
- [2] L. Wang, Y. Zhang, X. Luo, J. Zhang, and Z. Zheng, "Effects of earthworms and substrate on diversity and abundance of denitrifying genes (*nir S* and *nir K*) and denitrifying rate during rural domestic wastewater treatment," *Bioresource Technology*, vol. 212, pp. 174–181, 2016.
- [3] T. Wang, B. Zhu, and M. Zhou, "Ecological ditch system for nutrient removal of rural domestic sewage in the hilly area of the central Sichuan Basin, China," *Journal of Hydrology*, vol. 570, pp. 839–849, 2019.
- [4] Y. M. Zhang, G. Huang, H. W. Lu, and L. He, "Planning of water resources management and pollution control for Heshui River watershed, China: a full credibility-constrained programming approach," *The Science of the Total Environment*, vol. 524, pp. 280–289, 2015.
- [5] S. Li, X. Fei, Y. Chi, X. Jiao, and L. Wang, "Integrated temperature and DO effect on the lab scale A2O process: performance, kinetics and microbial community," *International Biodeterioration & Biodegradation*, vol. 133, pp. 170–179, 2018.
- [6] J. Schmidt, M. R. G. Marques, S. Botti, and M. A. L. Marques, "Recent advances and applications of machine learning in solid-state materials science," *NPJ Computational Materials*, vol. 5, no. 1, 2019.
- [7] L. Zhao, T. Dai, Z. Qiao, P. Sun, J. Hao, and Y. Yang, "Application of artificial intelligence to wastewater treatment: a bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse," *Process Safety and Environmental Protection*, vol. 133, pp. 169–182, 2020.
- [8] H. Zare Abyaneh, "Evaluation of multivariate linear regression and artificial neural networks in prediction of water quality parameters," *Journal of Environmental Health Science and Engineering*, vol. 12, no. 1, 2014.
- [9] U. Mahdiyah, M. I. Irawan, and E. M. Imah, "Study comparison backpropagation, support vector machine, and extreme learning machine for bioinformatics data," *Jurnal Ilmu Komputer dan Informasi*, vol. 8, no. 1, pp. 53–59, 2015.
- [10] S. Ding, C. Su, and J. Yu, "An optimizing BP neural network algorithm based on genetic algorithm," *Artificial Intelligence Review*, vol. 36, no. 2, pp. 153–162, 2011.
- [11] P. Antwi, D. Zhang, L. Xiao et al., "Modeling the performance of Single-stage Nitrogen removal using Anammox and Partial nitrification (SNAP) process with backpropagation neural network and response surface methodology," *The Science of the Total Environment*, vol. 690, pp. 108–120, 2019.
- [12] S. Mandal, S. S. Mahapatra, M. K. Sahu, and R. K. Patel, "Artificial neural network modelling of As(III) removal from water by novel hybrid material," *Process Safety and Environmental Protection*, vol. 93, pp. 249–264, 2015.
- [13] M. Gong, J. Liu, A. K. Qin, K. Zhao, and K. C. Tan, "Evolving deep neural networks via cooperative coevolution with backpropagation," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 420–434, 2021.
- [14] F. M. Dias, A. Antunes, J. Vieira, and A. M. Mota, "Implementing the levenberg-marquardt algorithm on-line: a sliding window approach with early stopping," *IFAC Proceedings Volumes*, vol. 37, no. 16, pp. 49–54, 2004.
- [15] M. Fan, J. Hu, R. Cao, W. Ruan, and X. Wei, "A review on experimental design for pollutants removal in water treatment with the aid of artificial intelligence," *Chemosphere*, vol. 200, pp. 330–343, 2018.
- [16] S. Rana, S. Jasola, and R. Kumar, "A boundary restricted adaptive particle swarm optimization for data clustering," *International Journal of Machine Learning and Cybernetics*, vol. 4, no. 4, pp. 391–400, 2013.
- [17] M. J. Mahmoodabadi, Z. Salahshoor Mottaghi, and A. Bagheri, "HEPSO: high exploration particle swarm optimization," *Information Sciences*, vol. 273, pp. 101–111, 2014.
- [18] Y. Mei, J. Yang, Y. Lu et al., "BP-ANN model coupled with particle swarm optimization for the efficient prediction of 2-chlorophenol removal in an electro-oxidation system," *International Journal of Environmental Research and Public Health*, vol. 16, no. 14, p. 2454, 2019.
- [19] M. Khajeh, A. Sarafraz-Yazdi, and A. F. Moghadam, "Modeling of solid-phase tea waste extraction for the removal of manganese and cobalt from water samples by using PSO-artificial neural network and response surface methodology," *Arabian Journal of Chemistry*, vol. 10, pp. S1663–S1673, 2017.
- [20] P. E. Poh, D. Gouwanda, Y. Mohan, A. A. Gopalai, and H. M. Tan, "Optimization of wastewater anaerobic digestion using mechanistic and meta-heuristic methods: current limitations and future opportunities," *Water Conservation Science and Engineering*, vol. 1, no. 1, pp. 1–20, 2016.
- [21] A. Azad, S. Farzin, H. Sanikhani, H. Karami, O. Kisi, and V. P. Singh, "Approaches for optimizing the performance of adaptive neuro-fuzzy inference system and least-squares support vector machine in precipitation modeling," *Journal of Hydrologic Engineering*, vol. 26, 2021.
- [22] A. Jalalkamali, "Using of hybrid fuzzy models to predict spatiotemporal groundwater quality parameters," *Earth Science India*, vol. 8, no. 4, pp. 885–894, 2015.
- [23] F. Luo, R. Yu, Y. Xu, and Y. Li, "Effluent quality prediction of wastewater treatment plant based on fuzzy-rough sets and artificial neural networks," in *Proceedings of the 2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, pp. 47–51, Tianjin, China, August 2009.
- [24] B. S. Ahn, S. S. Cho, and C. Y. Kim, "The integrated methodology of rough set theory and artificial neural network for business failure prediction," *Expert Systems with Applications*, vol. 18, no. 2, pp. 65–74, 2000.
- [25] Y. Qiang, Z. Dongxu, and T. Feng, "An initialization method for fuzzy C-means algorithm using subtractive clustering," in *Proceedings of the 2010 Third International Conference on Intelligent Networks and Intelligent Systems*, pp. 393–396, Thessalonika, Greece, November 2010.
- [26] J. T. Vogelstein, E. W. Bridgeford, M. Tang et al., "Supervised dimensionality reduction for big data," *Nature Communications*, vol. 12, no. 1, 2021.
- [27] C. Zhu, J. Zhang, Y. Liu, D. Ma, M. Li, and B. Xiang, "Comparison of GA-BP and PSO-BP neural network models with initial BP model for rainfall-induced landslides risk assessment in regional scale: a case study in Sichuan, China," *Natural Hazards*, vol. 100, no. 1, pp. 173–204, 2020.
- [28] C. Li, Z. Yang, H. Yan, and T. Wang, "The application and research of the GA-BP neural network algorithm in the MBR membrane fouling," *Abstract and Applied Analysis*, vol. 2014, pp. 1–8, Article ID 673156, 2014.
- [29] R.-B. Chen, S.-P. Chang, W. Wang, H.-C. Tung, and W. K. Wong, "Minimax optimal designs via particle swarm

- optimization methods,” *Statistics and Computing*, vol. 25, no. 5, pp. 975–988, 2015.
- [30] S. L. Wong, K. K. W. Wan, and T. N. T. Lam, “Artificial neural networks for energy analysis of office buildings with daylighting,” *Applied Energy*, vol. 87, no. 2, pp. 551–557, 2010.
- [31] T. Yilmaz, G. Seckin, and A. Yuceer, “Modeling of effluent COD in UAF reactor treating cyanide containing wastewater using artificial neural network approaches,” *Advances in Engineering Software*, vol. 41, no. 7–8, pp. 1005–1010, 2010.
- [32] R. H. McArthur and R. C. Andrews, “Development of artificial neural networks based confidence intervals and response surfaces for the optimization of coagulation performance,” *Water Supply*, vol. 15, no. 5, pp. 1079–1087, 2015.
- [33] C.-s. Zhan, X.-m. Song, J. Xia, and C. Tong, “An efficient integrated approach for global sensitivity analysis of hydrological model parameters,” *Environmental Modelling & Software*, vol. 41, pp. 39–52, 2013.
- [34] T. Y. Pai, P. Y. Yang, S. C. Wang et al., “Predicting effluent from the wastewater treatment plant of industrial park based on fuzzy network and influent quality,” *Applied Mathematical Modelling*, vol. 35, no. 8, pp. 3674–3684, 2011.
- [35] H.-G. Han, L.-X. Dong, and J.-F. Qiao, “Data-knowledge-driven diagnosis method for sludge bulking of wastewater treatment process,” *Journal of Process Control*, vol. 98, pp. 106–115, 2021.
- [36] Q. Yu, R. Liu, J. Chen, and L. Chen, “Electrical conductivity in rural domestic sewage: an indication for comprehensive concentrations of influent pollutants and the effectiveness of treatment facilities,” *International Biodeterioration & Biodegradation*, vol. 143, Article ID 104719, 2019.
- [37] APHA, “Standard methods for the examination of water and wastewater,” in A. P. H. A, A.P.H. Association, Washington, DC, USA, 16th edition, 1985.
- [38] Y. Liu, L. Hou, W. Bian, B. Zhou, D. Liang, and J. Li, “Turbidity in combined sewer sewage: an identification of stormwater detention tanks,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 9, p. 3053, 2020.
- [39] J. I. F. Slaets, P. Schmitter, T. Hilger et al., “A turbidity-based method to continuously monitor sediment, carbon and nitrogen flows in mountainous watersheds,” *Journal of Hydrology*, vol. 513, pp. 45–57, 2014.
- [40] F. Bagherzadeh, M.-J. Mehrani, M. Basirifard, and J. Roostaei, “Comparative study on total nitrogen prediction in wastewater treatment plant and effect of various feature selection methods on machine learning algorithms performance,” *Journal of Water Process Engineering*, vol. 41, Article ID 102033, 2021.
- [41] R. Noori, A. R. Karbassi, A. Moghaddamnia et al., “Assessment of input variables determination on the SVM model performance using PCA, Gamma test, and forward selection techniques for monthly stream flow prediction,” *Journal of Hydrology*, vol. 401, no. 3–4, pp. 177–189, 2011.
- [42] I. Ebtehaj and H. Bonakdari, “Evaluation of sediment transport in sewer using artificial neural network,” *Engineering Applications of Computational Fluid Mechanics*, vol. 7, no. 3, pp. 382–392, 2013.
- [43] M. M. Hamed, M. G. Khalafallah, and E. A. Hassanien, “Prediction of wastewater treatment plant performance using artificial neural networks,” *Environmental Modelling & Software*, vol. 19, no. 10, pp. 919–928, 2004.
- [44] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.
- [45] Y. Zhang and B. Pan, “Modeling batch and column phosphate removal by hydrated ferric oxide-based nanocomposite using response surface methodology and artificial neural network,” *Chemical Engineering Journal*, vol. 249, pp. 111–120, 2014.
- [46] F. S. Mjalli, S. Al-Asheh, and H. E. Alfadala, “Use of artificial neural network black-box modeling for the prediction of wastewater treatment plants performance,” *Journal of Environmental Management*, vol. 83, no. 3, pp. 329–338, 2007.
- [47] E. S. Elmolla, M. Chaudhuri, and M. M. Eltoukhy, “The use of artificial neural network (ANN) for modeling of COD removal from antibiotic aqueous solution by the Fenton process,” *Journal of Hazardous Materials*, vol. 179, no. 1–3, pp. 127–134, 2010.
- [48] M. S. Nasr, M. A. E. Moustafa, H. A. E. Seif, and G. El Kobrosy, “Application of Artificial Neural Network (ANN) for the prediction of EL-AGAMY wastewater treatment plant performance-Egypt,” *Alexandria Engineering Journal*, vol. 51, no. 1, pp. 37–43, 2012.
- [49] Y. Ma, M. Huang, J. Wan, K. Hu, Y. Wang, and H. Zhang, “Hybrid artificial neural network genetic algorithm technique for modeling chemical oxygen demand removal in anoxic/oxic process,” *Journal of Environmental Science and Health, Part A*, vol. 46, no. 6, pp. 574–580, 2011.
- [50] S. Huo, Z. He, J. Su, B. Xi, and C. Zhu, “Using artificial neural network models for eutrophication prediction,” *Procedia Environmental Sciences*, vol. 18, pp. 310–316, 2013.
- [51] P. Antwi, D. Zhang, W. Luo et al., “Performance, microbial community evolution and neural network modeling of single-stage nitrogen removal by partial-nitrification/anammox process,” *Bioresource Technology*, vol. 284, pp. 359–372, 2019.
- [52] L. R. McKenzie and P. N. W. Young, “Determination of ammonia-, nitrate- and organic nitrogen in water and waste water with an ammonia gas-sensing electrode,” *The Analyst*, vol. 100, no. 1194, pp. 620–628, 1975.