Desalination and Water Treatment
www.deswater.com
doi: 10.5004/dwt.2023.29183

282 (2023) 96–106
January

Genetic algorithm optimized back propagation artificial neural network for a study on a wastewater treatment facility cost model

Gaiqiang Yang[a], Yunfei Xu[a], Lijuan Huo[a,*], Dongpeng Guo[a], Junwei Wang[a], Shuang Xia[a], Yahong Liu[a], Qi Liu[b]

[a]*School of Environment and Resources, Taiyuan University of Science and Technology, Taiyuan 030024, China, emails: ygq@tyust.edu.cn (G.Q. Yang), S20202311010@stu.tyust.edu.cn (Y.F. Xu), hlj@tyust.edu.cn (L.J. Huo), guodp@126.com (D.P. Guo), S20202301012@stu.tyust.edu.cn (J.W. Wang), S20190693@stu.tyust.edu.cn (S. Xia), S20202301010@stu.tyust.edu.cn (Y.H. Liu)*

[b]*Shanxi Geological Environment Monitoring and Ecological Restoration Center, Taiyuan 030024, China, email: 34130438@qq.com*

Received 7 July 2022; Accepted 3 December 2022

ABSTRACT

In this study, the genetic algorithm optimized back propagation artificial neural network (GA-BP-ANN) method is used to predict the cost of a wastewater treatment plant. With biological oxygen demand, design volume, catchment area and treatment degree as input data, the total cost and construction cost as output parameters, the cost of a wastewater treatment plant is simulated. Compared with the linear algorithm commonly used before, this method has the following advantages: (1) GA-BP-ANN is suitable for small sample analysis and can effectively improve the stability of data. (2) Remove the influence of subjectivity and provide better help for decision makers. The effectiveness and feasibility of this method are proved theoretically and verified by simulation. The results can provide guidance for the design and operation of sewage treatment plants.

Keywords: Genetic algorithm optimized back propagation artificial neural network; Genetic algorithm; Neural network; Wastewater treatment plant; Cost model

1. Introduction

In the world, wastewater affects the development of human beings. The amount of wastewater produced in China is among the top in the world, and the ecological and economic pressure brought by it is also very obvious. In China, the ecological environment damage caused by wastewater accounts for more than 1/4 of GDP [1], which brings great pressure to economic development and environmental restoration. Some scholars collected survey data of 467 sewage treatment plants and on-site reports of 38 sewage treatment plants [2], and found that due to the unclear investment budget and other issues in the initial stage of design, various indicators did not meet the ideal requirements during construction, for example, the length of drainage pipe network per capita was insufficient, only 0.85 m. With the process of urbanization and economic development, urban sewage discharge has continued to grow, from 33.18 billion tons in 2000 to 57.14 billion tons in 2020, according to the data published by the Ministry of Housing and Urban Rural Development of China [3]. In order to reduce the economic and energy consumption during the operation of the WWTP, we should plan in advance, select the appropriate cost model and optimize the construction of the sewage treatment plant.

The cost model not only reflects the economic characteristics well in the construction cost [4], but also makes a scientific and forecast in the operation cost [5], it is widely

* Corresponding author.

used in the optimal design of WWTPs and the pricing process of wastewater treatment charges. However, the design and operation of the WTTP is a complicated systematic project. The selection of wastewater treatment technology has obvious one sidedness, subjectivity and uncertainty [6], and involves many data parameters, such as inlet and outlet water flow, inlet and outlet water quality, treatment degree, service area population, catchment area and local economic development level, which are directly related to the construction and operation cost of the sewage treatment plant.

Therefore, scholars from all of the world are still exploring the cost model, and have also established some practical cost models, such as the urban wastewater treatment investment and operation cost function model based on the data of the first national pollution source census [7], comprehensive cost-benefit analysis and life cycle assessment [8], the industrial wastewater treatment investment and operation cost function considering the difference of wastewater treatment cost among sub industries [9]. However, at present, most cost models of WWTPs in the world are modeled in the form of linear function or power exponential function. This rigid limitation leads to the addition of too many subjective factors in the case of incomplete information collection, and there may be a large deviation between the predicted or optimized results and the actual cost [10]. At present, most WWTPs in China are facing transformation and upgrading, and the data collection method is still relatively backward. Therefore, it is very important to establish a model to simulate the cost of sewage treatment plants based on uncertainty.

In order to deal with these nonlinear relations and uncertain data more accurately, the combination of neural network and cost model can be used [11]. Some scholars have done sufficient research on neural network in predicting economic costs. Meng used genetic algorithm to optimize back propagation (BP) neural network to predict short-term economic scheduling of hydrothermal power generation system, which set an example for economic budget [12]. Neural network has the advantage of scientific and accurate when there is no clear linear relationship and the weight is uncertain [13]. When the data of the cost model of sewage treatment plant in the planning stage is not complete or the function relationship is not clear, it can solve the problem well [14]. However, due to the limitations of neural network itself, direct training of data may produce large errors, and it is easy to fall into local optimal solution. Genetic algorithm, based on natural selection and genetics, can effectively search the global optimal solution [15]. Combining genetic algorithm with neural network to optimize the weight and threshold of back propagation artificial neural network (BP-ANN) model can reduce the limitations of neural network and overcome its shortcomings of slow convergence speed and easy to fall into local optimal solution [16].

Therefore, this paper aims to develop an accurate cost model based on genetic algorithm optimized back propagation artificial neural network (GA-BP-ANN) to solve the early investment problem of sewage treatment plant, so as to accurately predict the total investment cost (total investment) and construction cost (plant construction cost). The input parameters of the neural network in GA-BP-ANN include the maximum inflow flow, treatment degree, inflow biological oxygen demand (BOD_5) concentration and catchment area and the output data are the total cost and construction cost.

2. Model establishment

2.1. Artificial neural network

Artificial neural network (ANN) is a nonparametric model, which forms a network through interconnected mathematical nodes or neurons to model complex functional relationships [17]. The mathematical structure, by simulating the calculation method of human brain and nervous system, realizes the approximation of any complex nonlinear process related to the input and output of any system [18]. At present, a variety of neural networks have been used to simulate the biological treatment process of wastewater, including back propagation artificial neural network (BP-ANN), radial basis function neural network, fuzzy neural network, echo state network and deep belief network [19]. Among them, BP-ANN involves a large number of mathematical operations. Its principle is to slightly adjust the neural network when errors are found, so as to continue to optimize the model [20]. As shown in Fig. 1, as the simplest neural network structure, it is composed of input layer, hidden layer and output layer. The construction parameters of sewage treatment plant are transmitted from the input layer to the hidden layer. After the parameters are adjusted, they are calculated according to the cost activation function, and then the cost results are output.

Normally, a neural network consists of some basic nerve units (neurons) running in parallel, and the function of the network largely depends on the connections between these basic nerve units [21]. It is a single computing processor with two operators: summation, concatenation and transfer function [22,23]. The operators include weight and deviation. The transfer function determines the relationship between input and output, and adds nonlinearity and stability to the network [24]. A single output neural network with n hidden layer nodes can be described by Eq. (1) [25].

$$Y = \sum_{n=1}^{N} \omega_n \theta_n \left(X\right) \tag{1}$$

where X is the input parameter of the network, Y is the output result of the network, ω_n is the weight from the nth input layer to the output layer, $\theta_n(X)$ is the output value of the nth hidden layer, the calculation formula is:

$$\theta_n \left(X\right) = e^{\left(-\left\|x-\mu_n\right\|/\sigma_n^2\right)} \tag{2}$$

where μ_n is the center vector of the nth hidden node, x is Euclidean distance between $\left\|x-\mu_n\right\|$ and μ_n, σ_n is the radius of the nth hidden node.

The neural network developed in this paper consists of three layers, including an input layer composed of four neurons (including total inflow flow, treatment degree, BOD_5 and ponding area), a hidden layer composed of multiple neurons (these neurons can be changed in order to

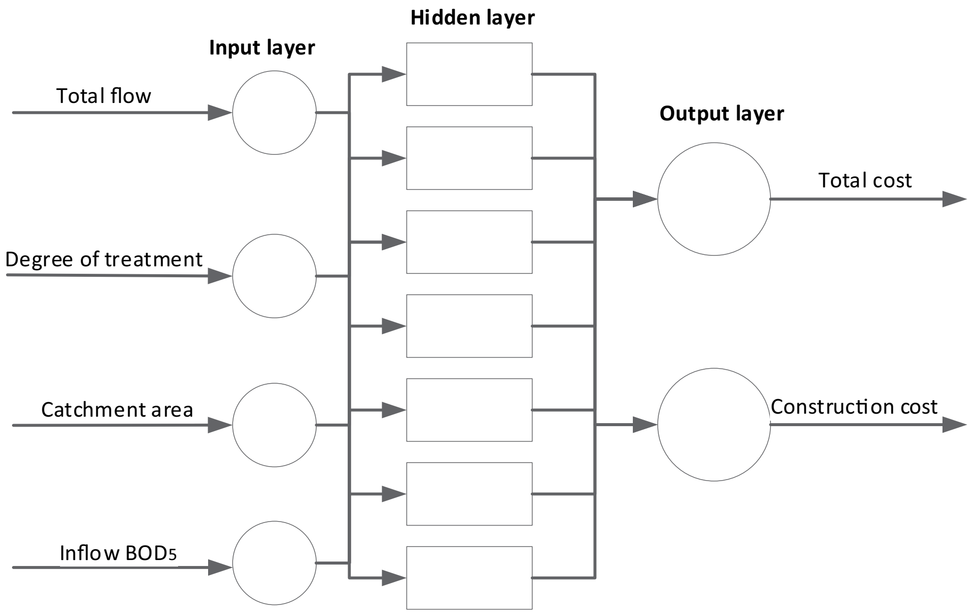

Fig. 1. Structure of neural network.

obtain the optimal model) and an output layer composed of output neurons (i.e., total cost and construction cost). The determination method of hidden layer function is determined by empirical formula [26].

$$m = \sqrt{m_1 + m_2} + a \tag{3}$$

where m is the number of hidden layer nodes, m_1 is the number of input layers, m_2 is the number of output layers, a is random number between 0 and 10.

Each network is trained until the network average root mean square error difference (RMSE) is the smallest and the variance R^2 is close to 1. The performance of the neural network model is measured between the network predicted value and the input value through R^2 and RMSE. The calculation method is shown in Eqs. (4) [27] and (5) [28].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i^* - y_p^i\right)^2}{\sum_{i=1}^{n}\left(y_i^* - \overline{y}\right)^2} \tag{4}$$

$$\text{RSME} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_p^i - y_i^*\right)^2}{n}} \tag{5}$$

where $\overline{y}$ is the average value of y_i^*, y_i^* and y_p^i are the predicted value and actual value of the nth target respectively.

The input and output range of neural network will be limited. The application of sample data needs data preprocessing before it can be applied to network training. When the network whose output is a single sigmoid function is selected, the output range is [0,1]. The middle region of the sigmoid function is sensitive to the change of input, and the two ends are slow to the change of input, so the input samples need to be compressed. Because different dimensions may produce different orders of magnitude, the input value can be transformed into a number between 0 and 1 for normalization. The commonly used proportional compression method can be used for compression [29], as shown in Eq. (6):

$$T = T_{\min} + \frac{T_{\max} - T_{\min}}{X_{\max} - X_{\min}}\left(X - X_{\min}\right) \tag{6}$$

where T is the transformed data, also known as target data, X is the original data (generally 0.2~0.8). $X_{\max}$ is the maximum value of the original data, $X_{\min}$ is the minimum value of the original data, $T_{\max}$ is the maximum value of the target data, and $T_{\min}$ is the minimum value of the target data.

2.2. Genetic algorithm

Genetic algorithm (GA) is an adaptive heuristic algorithm that guides parameter space coding through random technology. It is not only used to optimize the structure of large-scale neural network, but also used to extract features that can be used for identification and optimal control tasks. It has been widely used in the optimization of complex space in different scientific fields [30]. Genetic algorithm is mainly composed of initialization population, fitness function, selection function, crossover function, variant function and output. The best solution is obtained in all possible solutions by creating population, selection, crossover and mutation operators. The optimization methods include calculus, numerical method and random method [31].

At first, genetic algorithm will randomly generate a certain number of individuals, and generate new individuals through heredity, crossover and mutation. Then, according to individual fitness, the individuals with high fitness will be retained and the individuals with low fitness will be eliminated. Therefore, the new individuals inherit the excellent traits of the previous generation, which makes the algorithm evolve in the direction of the better solution of the problem, and finally find the optimal solution. It can be roughly divided into the following two stages [32].

2.2.1. Initialize population

The initial population is a certain number of optimal individuals generated from random numbers. The optimal individuals are selected in the iteration, and the final total number of individuals meets the needs of the initial population [33]. For the research of this paper, the order of each item is not limited, and the position can be changed from the data set.

2.2.2. Evolution process of genetic algorithm

The selection operation is to select some individuals in the population to reproduce a new generation. A single solution is selected through a fitness based process, where a more suitable solution is more likely to be selected (measured by the fitness function). There are different selection methods, such as random uniform, residual, uniform, linear shift, roulette and championship [34]. This paper adopts roulette selection method, and the selection probability of each element is directly proportional to its fitness value. Roulette is one of the most common selection methods, and it is also an effective selection method. The roulette selection method assigns a roulette to each individual, and the size of the roulette is proportional to the fitness of the individual. The more appropriate a component is, the larger the wheel blade it gets. In order to select an individual for selection, rotate the roulette wheel. At the place where the roulette stops, the corresponding sliced individual is caught as an individual living in future generations. The fitness function uses the inverse of the overall mean square deviation of the actual output and the predicted output plus one, as shown in Eq. (7).

$$f = \frac{1}{1 + \mathrm{mse}\left(Y - Y'\right)} \tag{7}$$

where Y is the actual output, Y' is the predicted output, and mse () is the mean square deviation of the data.

The purpose of mutation in genetic algorithm is to generate new individuals by changing all or part of the genes of the selected individuals in the population, so as to avoid the algorithm falling into local solutions. Mutation operator is one of the strategies used to ensure the variability within the population and design space exploration. Mutation is applied to the offspring produced by crossover, and its mutation probability is usually assigned a lower value.

2.3. Genetic algorithm optimization of back propagation neural network

In order to overcome the limitations of ANN and avoid poor training effect and falling into local optimal solution, this paper combines neural network and genetic algorithm, uses the variability of genetic algorithm to improve the global search ability, and optimizes back propagation artificial neural network based on genetic algorithm to apply it to the cost model research of sewage treatment plant. The specific flow of GA-BP-ANN is shown in Fig. 2. Its structure is similar to the process of neural network prediction

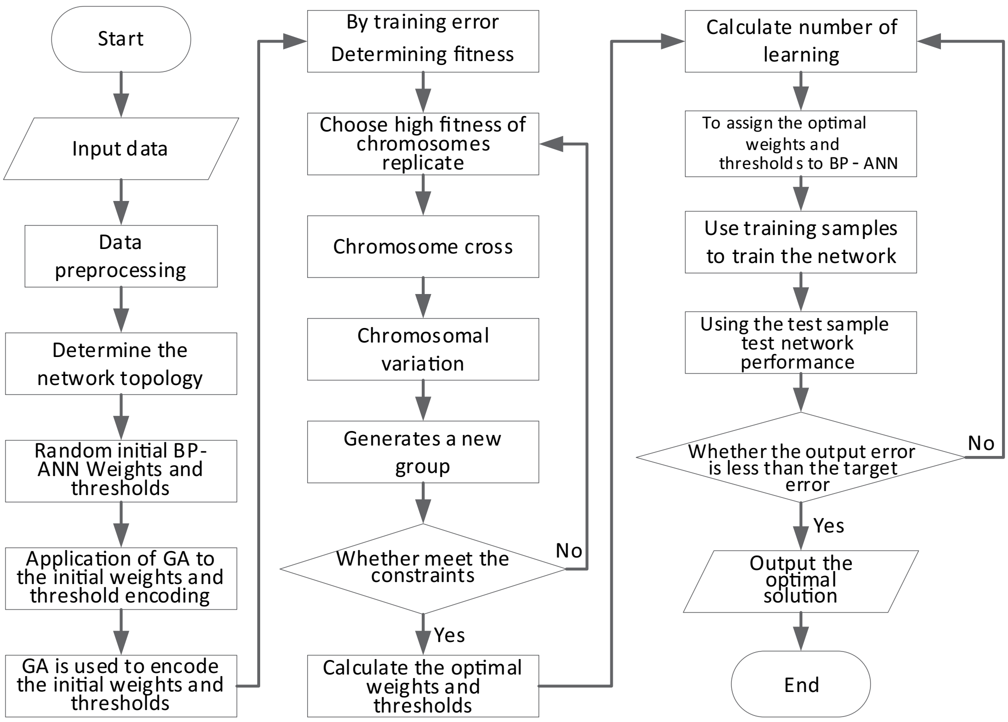

Fig. 2. GA-BP-ANN flowchart.

method. First, determine the network topology. Second, genetic algorithm is introduced into the training process of the network to calculate the weight and threshold. Then sample training is carried out and the optimal solution is output.

3. Case study

3.1. Data sources

The model is built according to the statistical data of 26 sewage treatment plants in Taiwan, China province [35], as shown in Supplemental Material S1. The basic statistical data include design water volume A1, treatment degree A2 (primary, secondary and advanced treatment), influent BOD_5 concentration A3 and ponding area A4; The cost obtained from the basic data includes the total cost Y1 and the construction cost Y2.

The establishment of a linear model usually requires a good degree of correlation between various data [36], so it is difficult to collect the data required by the traditional cost model of sewage treatment plant, while the neural network can predict the total cost and construction cost of sewage treatment plant according to the intrinsic correlation of data itself. Can be applied to sewage treatment plant data with the same requirements at different times.

The correlation analysis of the collected data shows that the correlation between the data is not strong, as shown by the scatter location of the correlation data distribution graph and the R^2 of each two sets of data. As shown in Fig. 3, the closer the variance is to 1, the better the correlation is. The best R^2 is 0.70098 in the correlation analysis of catchment area and total flow, and the worst is –0.00796 in the correlation analysis of inlet BOD_5 and catchment area. The R^2 of other data are all below 0.35 and far less than 1. Because the correlation between input data is not strong, so the linear algorithm cannot be used to establish a good prediction model, but the genetic algorithm is used to optimize the neural network to avoid this problem.

3.2. Model parameter setting

The basic parameters of the neural network are set as follows: the number of epochs are 4000 times, the rate of training is 0.25, the target accuracy is 0.0001, and the momentum constant is 0.7. The momentum adaptive adjustment network with a 4-7-2 three-layer structure is adopted. The transfer function from the input layer to the hidden layer is Log-sigmoid function, the linear function from the hidden layer to the output layer is 'purelin', and the faster 'train1m' is selected as the training function. In the input samples, 25 groups of data are randomly selected as the training data, and the remaining group of data is used to test the network performance. During modeling, the best network model is selected according to the minimum error. According to the parameter setting of genetic algorithm, the training results show that the fitness training tends to a gentle optimal value after 768 times, as shown in Fig. 4. After multiple verification, the population size is set as

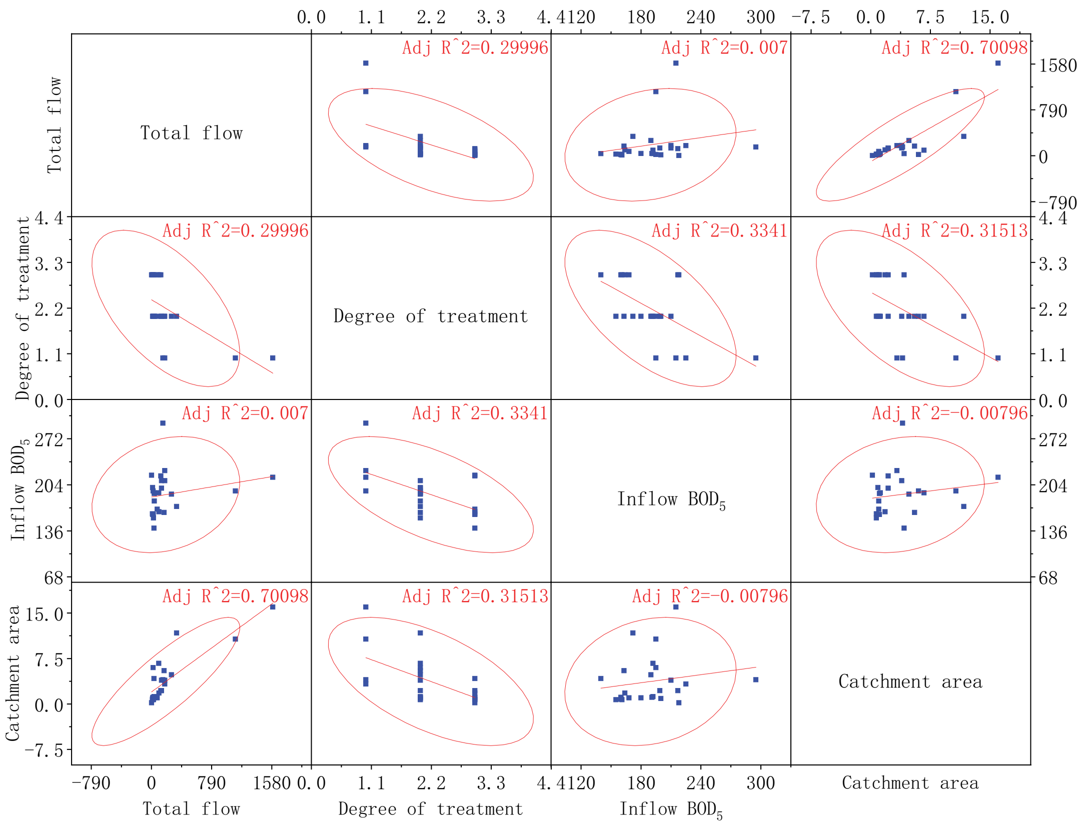

Fig. 3. Input data correlation diagram.

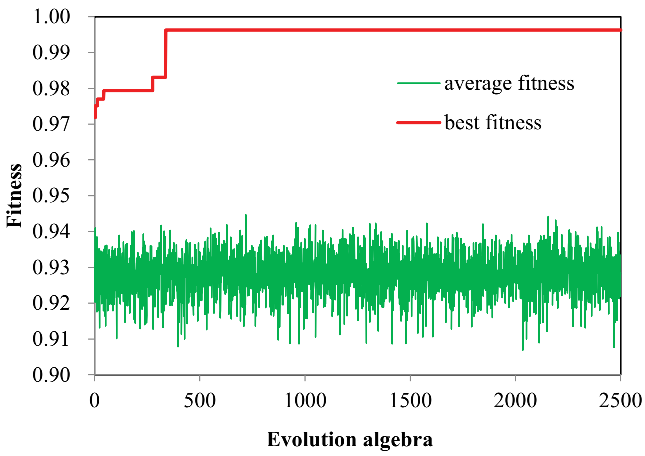

Fig. 4. Optimal fitness curve.

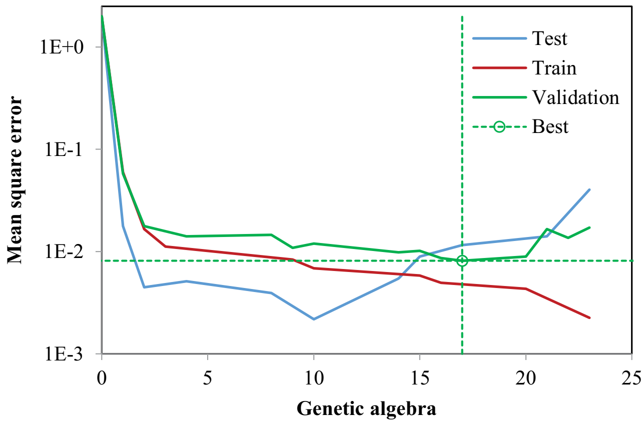

Fig. 5. Neural network training optimal mean variance algebra display.

25 times, the number of iterations is 2,500 times, the crossover probability is 0.75 and the mutation probability is 0.01.

4. Results and discussion

4.1. Model performance evaluation

Use the leave one method for data processing [37], and randomly select one group each time to test the rest of the network for training. As shown in Fig. 5, when the genetic algebra is 17 generations, the verification sequence has the best performance, and its mean square deviation is 0.813×10^{-2}. When the genetic algebra is more than 3 generations, the mean square deviation of the training sequence decreases steadily, and the mean square deviation is less than 0.01 from 10 generations, and the mean square deviation of the test and validation series is also basically maintained at around 0.01. The mean square error results show that the neural network has good performance.

Fig. 6 shows the training and prediction cost under the GA-BP-ANN method, in which Fig. 6a and b show the simulation effect of training data set of total cost and construction cost respectively. The training data R^2 of the total cost is 0.986 and the training data R^2 of the construction cost is 0.957, which are close to 1. It can be seen that GA-BP-ANN has a very good training effect.

4.2. Comparative analysis of models

In the past, many scholars studied the optimization of neural network by genetic algorithm. Lira et al. [16] applied genetic algorithm to optimize the model of neural network for NO_x hydrodynamics. Guan et al. [38] applied genetic algorithm optimized neural network to the data mining and design of electromagnetic characteristics of Co/FeSi filling coating, and Shin et al. applied genetic algorithm to the prediction of natural gas NO_x [39]. These studies made a comparison before and after the application of genetic algorithm. It is found that genetic algorithm plays an excellent role in optimizing the weight and threshold value of neural network transmission process, which makes the prediction result more accurate and the model more practical.

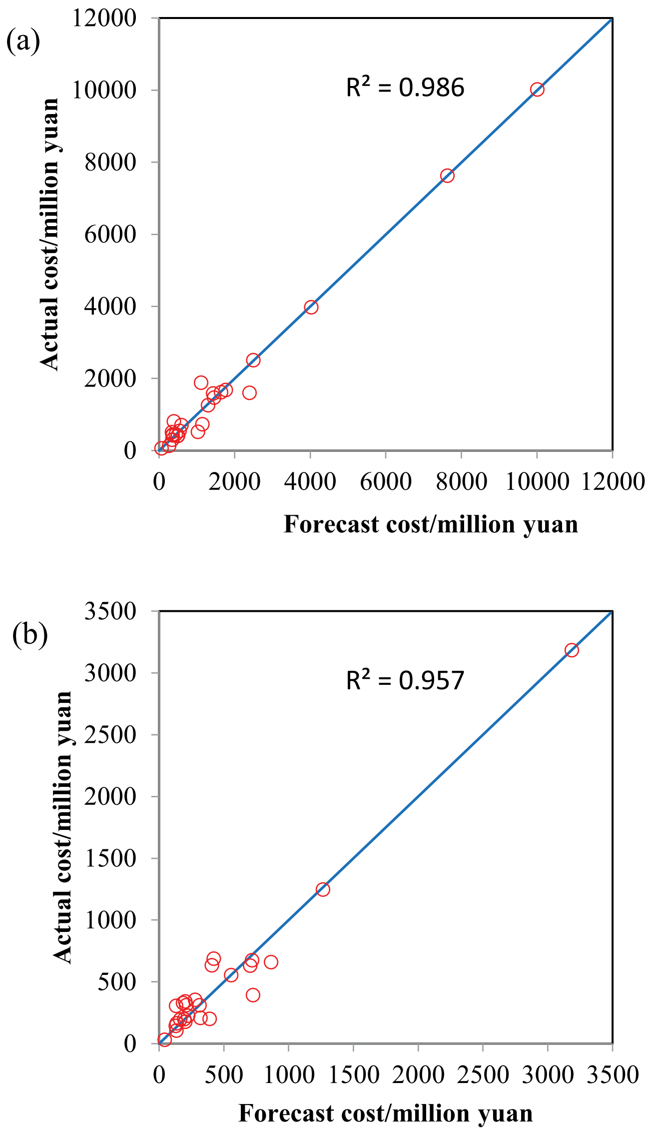

Fig. 6. GA-BP-ANN training effect comparison: (a) total cost and (b) construction cost.

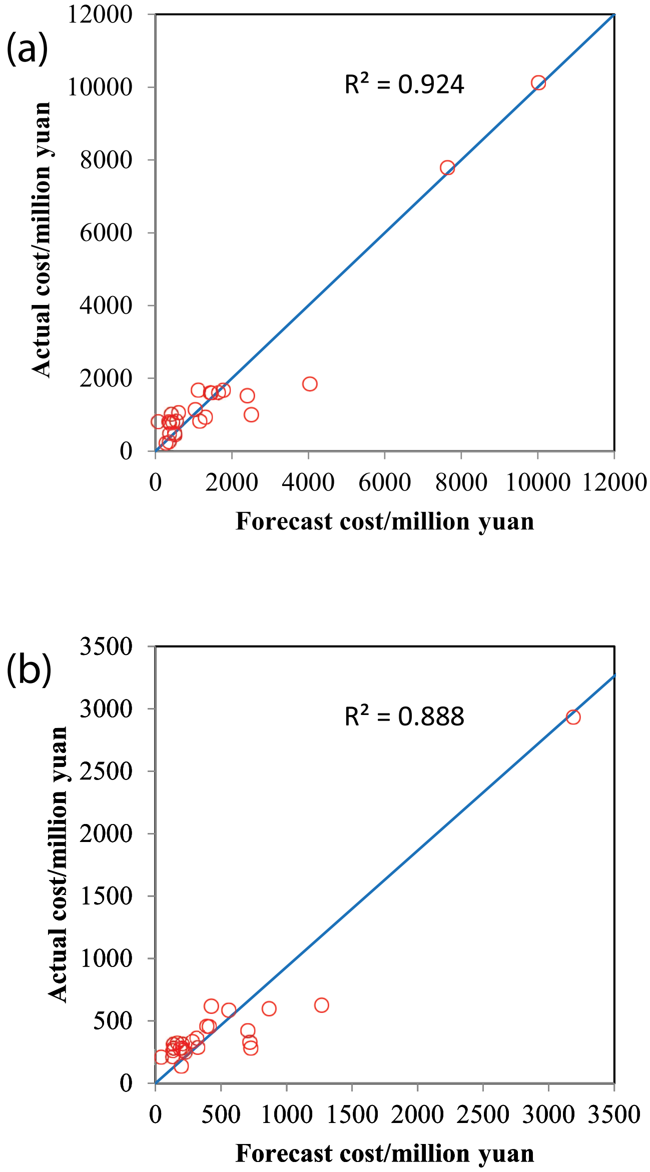

Fig. 7. BP-ANN training effect comparison: (a) total cost and (b) construction cost.

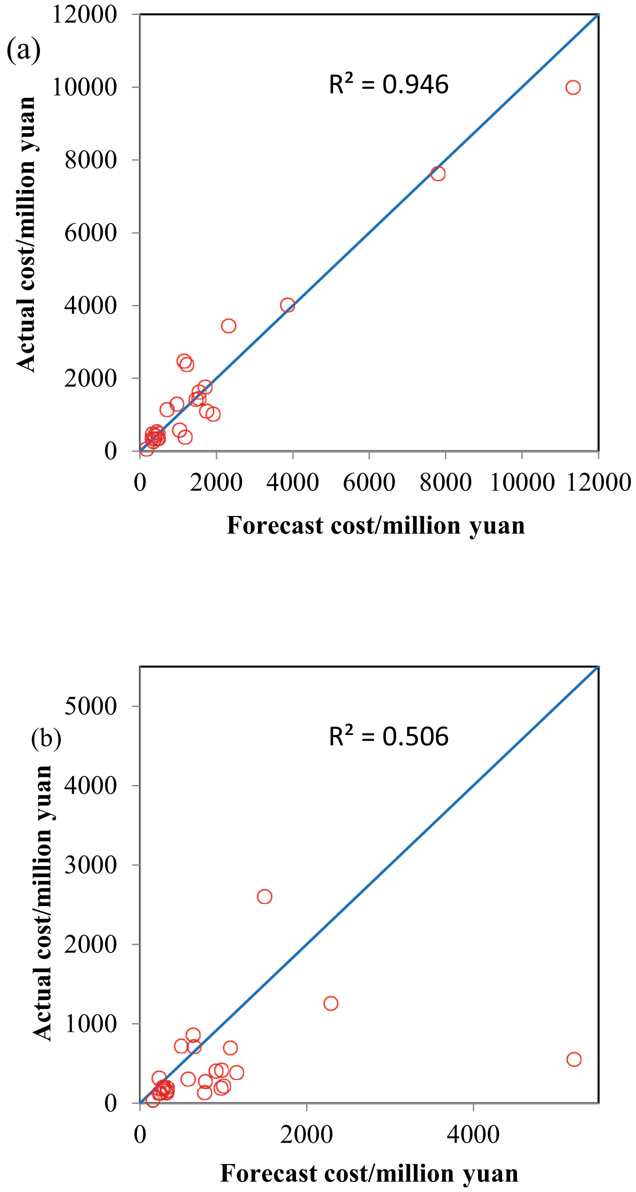

Fig. 8. FLR simulation effect comparison: (a) total cost and (b) construction cost.

In order to verify the effectiveness and superiority of GA-BP-ANN, the same training data set is used and the BP-ANN method is used to train the data set. The results are shown in Fig. 7. However, by comparing Figs. 6 and 7 it can be seen that the training data R^2 of GA-BP-ANN is greater than that of BP-ANN. Therefore, from the experimental data, the training effect of GA-BP-ANN is significantly better than that of BP-ANN.

Baskar et al. [40] successfully applied fuzzy linear regression to rice quality assessment, Pandelara et al. [41] applied it to the relationship analysis between electricity consumption and GDP, and Prieto et al. [42] applied it to the prediction of the functional life of cultural heritage. These studies show that fuzzy linear regression has been very mature and applied to the simulation prediction in various fields. However, the disadvantages of fuzzy linear regression analysis (FLR) are also obvious, for it is very sensitive to outliers [43]. If there are too many outliers, the error in the prediction results will increase. Hence in order to compare and analyze the difference between this method and the traditional linear algorithm, Wen et al. used FLR method to compare the data set simulation, as shown in Fig. 8. The results show that the simulation effect of the total cost of FLR method is good, R^2 is 0.946, but the simulation effect of construction cost is poor, R^2 is only 0.506. The error between the simulation results and the actual data is large, because there are many linear relationships between various factors. The more data, the greater the amount of calculation, and there are inevitably subjective factors in the selection and calculation of ambiguity.

Compare the error distribution of GA-BP-ANN and FLR, as shown in Fig. 9. It can be seen that the error value of

GA-BP-ANN is generally less than that of FLR, and the data is more concentrated around 0 error, especially in the error comparison of construction cost in Fig. 9b. On the whole, individual data of GA-BP-ANN error is greater than FLR error, which is due to the small number of data groups for training and the data is not de-noised. In addition, the fault tolerance of GA-BP-ANN is greater than FLR. For the quantity with total cost less than 400 million, GA-BP-ANN is 20 groups, accounting for 77% of the total quantity; FLR was 16 groups, accounting for 61% of the total. The comparison error of construction cost is only the error result of GA-BP-ANN, which obeys the normal distribution, and the result of FLR shows obvious skewness. The error of construction cost is within 200 million, and GA-BP-ANN is 21 groups, accounting for 81% of the total; FLR was 11 groups, accounting for 52% of the total.

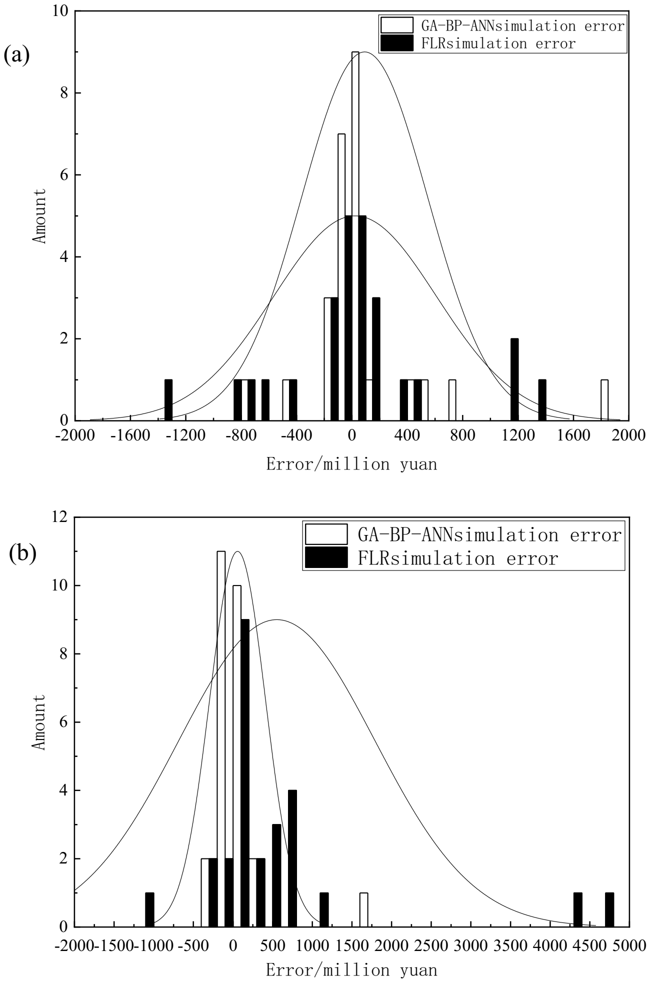

Fig. 9. Model error distribution: (a) total cost and (b) construction cost.

The R^2 and RMSE of the three models are listed in Table 1. It can be seen that the maximum R^2 of GA-BP-ANN simulation result is 0.986 of the total cost training set and the minimum is 0.782 of the total construction cost data set. The maximum of RMSE is 455.63 of the total cost data set, and the minimum is 129.45 of the construction cost training set. Under the same conditions, it is obvious that the RMSE and R^2 of GA-BP-ANN are less than BP-ANN and FLR.

On the whole, GA-BP-ANN is better than BP-ANN and FLR, and GA-BP-ANN removes subjective factors and only considers the correlation of the data itself. The application of genetic algorithm reduces the possibility of neural network falling into local optimal solution and obtains the best initial threshold and weight. When the functional relationship is clear and the amount of data is too small, FLR can be considered. If the data volume is large enough and the linear relationship is not obvious, GA-BP-ANN can be used to avoid errors caused by subjective factors and unclear internal function relationship of the data.

5. Conclusion

In this paper, a GA-BP-ANN model is developed and used to study the sewage treatment fee model. In the process of neural network training, the input parameters include design water volume, treatment degree, influent BOD_5 concentration and ponding area. Genetic algorithm is used to optimize the neuron, function, weight and threshold of back-propagation neural network. The training results show that the performance of this model is good. The minimum R^2 of the total cost and construction cost of sewage treatment plant is 0.782 and the maximum is 0.986. The application of genetic algorithm can improve the accuracy of neural network, make the prediction results more accurate, and avoid the training process falling into local optimal solution. Through the comparative analysis of the models, the simulation effect of GA-BP-ANN is better than BP-ANN and FLR in RMSE and R^2. GA-BP-ANN

Table 1
Comparison of model simulation effects

Model Name	Cost categories	RMSE		R^2	
		Training set	Total set	Training set	Total set
GA-BP-ANN	Total cost	279.31	455.63	0.986	0.962
	Construction costs	129.45	349.82	0.957	0.782
BP-ANN	Total cost	644.37	718.04	0.924	0.904
	Construction costs	217.79	461.83	0.888	0.632
FLR	Total cost	–	577.58	–	0.946
	Construction costs	–	1,333.06	–	0.506

is more suitable for the prediction of data when the internal function is unclear.

GA-BP-ANN can be used to predict the required cost for the construction or upgrading of sewage treatment plant, so that decision makers can have an intuitive data reference. However, GA-BP-ANN needs a large amount of field data to train and test the network, and it may face the impact of excessive noise for a large amount of data. Therefore, it can be considered to smooth the data curve through cluster analysis, so as to be better used for network training and testing. When it is difficult to collect data, we can refer to GA-BP-ANN and apply genetic algorithm to optimize small sample data to improve the stability of data.

Acknowledgments

This work was supported by National Natural Science Foundation of China (No. 51709195, 41807130, U21A20524), Fundamental Research Program of Shanxi Province (No. 20210302123213) and Taiyuan University of Science and Technology Joint Training Base for Graduate Students (No. JD2022018).

References

[1] C.H. Liu, W. Cai, M.Y. Zhai, Z. Guang, C.X. Zhang, Z.G. Jiang, Decoupling of wastewater eco-environmental damage and China's economic development, Sci. Total Environ., 789 (2021) 147980, doi: 10.1016/j.scitotenv.2021.147980.
[2] J.B. Zhang, H.C. Wang, Y.T. Shao, G.H. Liu, L. Qi, W.Y. Dang, J.L. Yuan, Y.H. Li, Z.H. Xia, Analysis on common problems of the wastewater treatment industry in urban China, Chemosphere, 291 (2021) 132875, doi: 10.1016/j.chemosphere.2021.132875.
[3] Ministry of Housing and Urban Rural Development of the People's Republic of China, Statistical Yearbook of Urban and Rural Construction 2020, China Scientific Research Press, Beijing, China, 2021.
[4] K.S.V.S. Pujitha, K. Venkatesh, Forecasting the construction cost by using unit based estimation model, Mater. Today Proc., 33 (2020) 613–619.
[5] H.L. Jiang, H.X. Zhang, X.H. Shi, Refined production simulation and operation cost evaluation for power system with high proportion of renewable energy, Energy Rep., 8 (2022) 108–118.
[6] Y. Xu, H.Y. Li, Y. Li, X.Q. Zheng, C.X. Zhang, P.Z. Chen, Q. Li, L. Tan, Systematically assess the advancing and limiting factors of using the multi-soil-layering system for treating rural sewage in China: from the economic, social, and environmental perspectives, J. Environ. Manage., 312 (2022) 114912, doi: 10.1016/j.jenvman.2022.114912.
[7] Y.K. Yang, K.R. Kim, R.R. Kou, Y.P. Li, J. Fu, L. Zhao, H.B. Liu, Prediction of effluent quality in a wastewater treatment plant by dynamic neural network modeling, Process Saf. Environ. Prot., 158 (2022) 515–524.
[8] Y.J. Xie, L. Zeng, P. Wang, X.D. Wu, T.J. Feng, Water cost for water purification: renewability assessment of a typical wastewater treatment plant in China, J. Cleaner Prod., 349 (2022) 131474, doi: 10.1016/j.jclepro.2022.131474.
[9] J. Yeo, S.S. Chopra, D. von Eiff, S. Jeong, L. Zhang, A.K. An, An integrated techno-economic analysis on wastewater reclamation in Hong Kong: a comprehensive cost – benefit analysis with life cycle assessment, J. Cleaner Prod., 357 (2022) 131838, doi: 10.1016/j.jclepro.2022.131838.
[10] D. Torregrossa, U. Leopold, F. Hernandez-Sancho, J. Hansen, Machine learning for energy cost modelling in wastewater treatment plants, J. Environ. Manage., 223 (2018) 1061–1067.
[11] X.L. Zheng, H. Nguyen, X.N. Bui, Exploring the relation between production factors, ore grades, and life of mine for forecasting mining capital cost through a novel cascade forward neural network-based salp swarm optimization model, Resour. Policy, 74 (2021) 102300, doi: 10.1016/j.resourpol.2021.102300.
[12] X.P. Men, H.G. Zhan, W.Y. Tan, A hybrid method of GA and BP for short-term economic dispatch of hydrothermal power systems, Math. Comput. Simul., 51 (2000) 341–348.
[13] G. Durai, R. Ramsenthil, M. Dilipkumar, Artificial neural network based modeling for the degradation of tannery wastewater in sequential batch reactor, Desal. Water Treat., 251 (2022) 27–34.
[14] R. Saravanathamizhan, K.H. Vardhan, D.G. Prakash, N. Balasubramanian, RSM and ANN modeling for electro-oxidation of simulated wastewater using CSTER, Desal. Water Treat., 55 (2015) 1445–1452.
[15] L. Godo-Pla, P. Emiliano, F. Valero, M. Poch, G. Sin, H. Monclus, Predicting the oxidant demand in full-scale drinking water treatment using an artificial neural network: uncertainty and sensitivity analysis, Process Saf. Environ. Prot., 125 (2019) 317–327.
[16] J.O.B. Lira, H.G. Riella, N. Padoin, C. Soares, Computational fluid dynamics (CFD), artificial neural network (ANN) and genetic algorithm (GA) as a hybrid method for the analysis and optimization of micro-photocatalytic reactors: NO_x abatement as a case study, Chem. Eng. J., 431 (2022) 133771, doi: 10.1016/j.cej.2021.133771.
[17] E. Maleki, N. Maleki, Artificial neural network modeling of Pt/C cathode degradation in PEM fuel cells, J. Electron. Mater., 45 (2016) 3822–3834.
[18] N. Maleki, E. Maleki, Modeling of Cathode Pt/C Electrocatalyst Degradation and Performance of a PEMFC Using Artificial Neural Network, ICEMIS '15: Proceedings of The International Conference on Engineering & MIS 2015, 2015, pp. 1–8, doi: 10.1145/2832987.2833000.
[19] E. Szatyowicz, I. Skoczko, P. Ofman, Using artificial neural networks for modeling wastewater treatment in small wastewater treatment plant, Rocz. Ochr. Sr., 18 (2016) 493–506.
[20] G.M. Wang, Q.S. Jia, M.C. Zhou, J. Bi, J.F. Qiao, A. Abusorrah, Artificial neural networks for water quality soft-sensing in wastewater treatment: a review, Artif. Intell. Rev., 55 (2022) 565–587.
[21] E. Malekia, M.J. Mirzaali, M. Guagliano, S. Bagherifard, Analyzing the mechano-bactericidal effect of nano-patterned surfaces on different bacteria species, Surf. Coat. Technol., 408 (2021) 126782, doi: 10.1016/j.surfcoat.2020.126782.
[22] S. Amin, Backpropagation – Artificial Neural Network (BP-ANN): understanding gender characteristics of older driver accidents in West Midlands of United Kingdom, Saf. Sci., 122 (2020) 104539, doi: 10.1016/j.ssci.2019.104539.
[23] M. Khayet, C. Cojocaru, Artificial neural network modeling and optimization of desalination by air gap membrane distillation, Sep. Purif. Technol., 86 (2012) 171–182.
[24] N. Malek, S. Kashanian, E. Maleki, M. Nazari, A novel enzyme based biosensor for catechol detection in water samples using artificial neural network, Biochem. Eng. J., 128 (2017) 1–11.
[25] M. Bagheri, S.A. Mirbagheri, Z. Bagheri, A.M. Kamarkhani, Modeling and optimization of activated sludge bulking for a real wastewater treatment plant using hybrid artificial neural networks-genetic algorithm approach, Process Saf. Environ. Prot., 95 (2015) 12–25.
[26] B.W. Zhang, S.M. Guo, H. Jin, Production forecast analysis of BP neural network based on Yimin lignite supercritical water gasification experiment results, Energy, 246 (2022) 123306, doi: 10.1016/j.energy.2022.123306.
[27] E. Maleki, Artificial neural networks application for modeling of friction stir welding effects on mechanical properties of 7075-T6 aluminum alloy, IOP Conf. Ser.: Mater. Sci. Eng., 103 (2015) 012034, doi: 10.1088/1757-899X/103/1/012034.
[28] E. Maleki, O. Unal, K.R Kashyzadeh, Fatigue behavior prediction and analysis of shot peened mild carbon steels, Int. J. Fatigue, 116 (2018) 48–67.
[29] D.C. Jin, P. Wang, Z.H. Bai, X.X. Wang, H. Peng, R. Qi, Z.S. Yu, G.Q. Zhuang, Analysis of bacterial community in bulking sludge using culture-dependent and -independent approaches, J. Environ. Sci., 23 (2011) 1880–1887.

[30] L.L. Xu, Research on computer interactive optimization design of power system based on genetic algorithm, Energy Rep., 7 (2021) 1–13.

[31] F. Momenbeik, M. Roosta, A.A. Nikoukar, Simultaneous microemulsion liquid chromatographic analysis of fat-soluble vitamins in pharmaceutical formulations: optimization using genetic algorithm, J. Chromatogr. A, 1217 (2010) 3770–3773.

[32] V. Ganesan, M. Sobhana, G. Anuradha, P. Yellamma, O.R. Devi, K.B. Prakash, J. Naren, Quantum inspired meta-heuristic approach for optimization of genetic algorithm, Comput. Electr. Eng., 94 (2021) 107356, doi: 10.1016/j.compeleceng.2021.107356.

[33] B. Liu, S.C. Li, L.C. Nie, J. Wang, X. L, Q.S. Zhang, 3D resistivity inversion using an improved Genetic Algorithm based on control method of mutation direction, J. Appl. Geophys., 87 (2012) 1–8.

[34] U. Kilic, K. Ayan, U. Arifoglu, Optimizing reactive power flow of HVDC systems using genetic algorithm, Int. J. Electr. Power Energy Syst., 55 (2014) 1–12.

[35] C.G. Wen, C.S. Lee, Development of a cost function for wastewater treatment systems with fuzzy regression, Fuzzy Sets Syst., 106 (1999) 143–153.

[36] C.-Q. Guo, Y.-L. Cui, Improved solute transport and pollutant degradation model of free water surface constructed wetlands considering significant linear correlation between model parameters, Bioresour. Technol., 327 (2021) 124817, doi: 10.1016/j.biortech.2021.124817.

[37] R.J. Brook, M. Stone, F.Y. Chan, L.K. Chan, Cross-validatory graduation, Insur. Math. Econ., 7 (1988) 59–66.

[38] Z.J. Guan, R. Li, J.T. Jiang, B. Song, Y.X. Gong, L. Zhen, Data mining and design of electromagnetic properties of Co/FeSi filled coatings based on genetic algorithms optimized artificial neural networks (GA-ANN), Composites, Part B, 226 (2021) 109383, doi: 10.1016/j.compositesb.2021.109383.

[39] Y. Shin, Z. Kim, J. Yu, G. Kim, S. Hwang, Development of NO_x reduction system utilizing artificial neural network (ANN) and genetic algorithm (GA), J. Cleaner Prod., 232 (2019) 1418–1429.

[40] C. Baskar, N. Nesakuma, J.B.B. Rayappan, M. Doraipandian, A framework for analysing E-Nose data based on fuzzy set multiple linear regression: paddy quality assessment, Sens. Actuators, A, 267 (2017) 200–209.

[41] D. Pandelara, W. Kristjanpoller, K. Michell, M.C. Minutolo, A fuzzy regression causality approach to analyze relationship between electrical consumption and GDP, Energy, 239 (2022) 122459, doi: 10.1016/j.energy.2021.122459.

[42] A.J. Prieto, A. Silva, J. de Brito, J.M. Macias-Bernal, F.J. Alejandre, Multiple linear regression and fuzzy logic models applied to the functional service life prediction of cultural heritage, J. Cult. Heritage, 27 (2017) 20–35.

[43] M.M. Nasrabadi, E. Nasrabadi, A.R. Nasrabady, Fuzzy linear regression analysis: a multi-objective programming approach, Appl. Math. Comput., 163 (2005) 245–251.

Appendix

The statistical data of 26 sewage treatment plants is listed in Table S1. Total investment costs and construction costs are converted from Taiwan Dollar to Chinese Yuan based on exchange rate 0.219.

Table S1
Statistics of sewage treatment plants in Taiwan [35]

Planning region	Design flow rate (10 m^3/d)	Treatment degree	Influent BOD_5 (mg/L)	Collection area (10^3 ha)	Total investment cost (million yuan)	Construction cost (million yuan)
Kaohsiung city						
- Nantzu area	175	Primary	225	3.3	1,636.149	224.694
- Kaohsiung downtown	1,103	Primary	195	10.7	7,632.369	556.26
- Linhai area	151	Primary	295	4	1,457.007	195.129
Taipei suburban	1,594	Primary	215	16	10,008.3	3,186.45
Keehmg city	122	Advanced	217	2.2	2,491.125	136.656
Juifang county	16	Secondary	200	0.9	489.027	318.864
Hsinchu city	21	Secondary	195	6	2,393.67	863.955
Chupei city	38	Secondary	191	1.1	498.663	203.013
Kanping fiver basin						
- Pingtung area	99	Advanced	164	1.8	1,298.889	718.539
- Chimei area	13	Advanced	161	0.7	361.788	132.057
Chunan & Toufang cities	133	Secondary	199	2.2	395.733	278.349
Hualien city	165	Secondary	163	5.5	1,024.482	391.134
Wu river basin						
- Nantou area	26	Secondary	155	0.7	267.618	128.991
- Tsaotun area	36	Secondary	192	1.2	356.751	163.812
Panchiao & Hsintien cities	95	Secondary	192	6.7	1,115.148	421.137
Tungkang river basin						
- Chaochou area	23	Advanced	160	0.8	443.475	210.459
- Neipu county	22	Advanced	160	1.1	541.587	202.137
- Tungkang city	17	Advanced	160	1	338.136	186.15
Tanshui city	75	Advanced	168	1	1,149.75	724.89
Erhjen river basin						
- Yungkang city	176	Secondary	210	3.9	1,765.359	702.114
- Jente area	132	Secondary	210	3.9	1,430.727	406.683
Taoyuan city	263	Secondary	190	4.8	3,457.134	2,609.604
Chungli city	331	Secondary	172	11.7	4,029.162	1,264.506
Peikang city	34	Advanced	140	4.2	591.957	311.418
Taitung city	38	Secondary	180	1	346.896	132.714
Kengting area	2	Advanced	218	0.2	67.89	43.362