MDPI

Article

Application of a New Architecture Neural Network in Determination of Flocculant Dosing for Better Controlling Drinking Water Quality

Huihao Luo [1], Xiaoshang Li [2], Fang Yuan [2], Cheng Yuan [2], Wei Huang [2], Qiannan Ji [2], Xifeng Wang [1], Binzhi Liu [3,*] and Guocheng Zhu [1,*]

[1] College of Civil Engineering, Hunan University of Science and Technology, Xiangtan 411201, China
[2] Xiangtan Middle Ring Water Business Limited Corporation, Xiangtan 411201, China
[3] School of Civil and Transportation Engineering, Guangdong University of Technology, Guangzhou 510006, China

* Correspondence: lbz@gdutedu.cn (B.L.); zgc945ahhn@163.com (G.Z.); Tel./Fax: +86-0731-5829-0269 (G.Z.)

Abstract: In drinking water plants, accurate control of flocculation dosing not only improves the level of operation automation, thus reducing the chemical cost, but also strengthens the monitoring of pollutants in the whole water system. In this study, we used feedforward signal and feedback signal data to establish a back-propagation (BP) model for the prediction of flocculant dosing. We examined the effect of the particle swarm optimization (PSO) algorithm and data type on the simulation performance of the model. The results showed that the parameters, such as the learning factor, population size, and number of generations, significantly affected the simulation. The best optimization conditions were attained at a learning factor of 1.4, population size of 20, 20 generations, 8 feedforward signals and 1 feedback signal as input data, 6 hidden layer nodes, and 1 output node. The coefficient of determination (R^2) between the predicted and measured values was 0.68, and the root mean square error (RMSE) was lower than 20%, showing a good prediction result. Weak time-delay data enhanced the model accuracy, which increased the R^2 to 0.73. Overall, with the hybridized data, PSO, and weak time-delay data, the new architecture neural network was able to predict flocculant dosing.

Keywords: drinking water; flocculation; neural network; BP; PSO

1. Introduction

The demand for clean drinking water is increasing in tandem with cities' and populations' rapid growth. However, the quality of the raw water that produces it is gradually deteriorating, necessitating complex water quality control [1], which poses great challenges to drinking water treatment. Coagulation is the primary operation unit in drinking water treatment, which is dependent on the action of the coagulant. Traditional coagulation dosing methods have placed significant strain on drinking water treatment plants in China and many other developing countries [2,3]. In 2017, the World Health Organization (WHO) proposed that excessive chemical use will cause a bad taste and sediment accumulation in distribution systems [4]. Depending on the characteristics of the raw water, hourly dosing control has become a critical development need. Manual coagulation dosing has limited treatment capability, and this dosing method is neither economical nor environmentally friendly [5]. As a result, a cost-effective method must be found. In addition, drinking water plants have received wide attention regarding water pollution, in order to enhance the understanding of water quality and water management. However, external disturbances to the coagulation system include strong nonlinearity and time delay; furthermore, water plant workers lack a thorough understanding of controlling interference factors. It seems

difficult to use a traditional numerical model to address these issues [6,7]. Therefore, there is a growing demand for the application of neural network models.

The neural network model is a bionic model that mimics the information processing mechanism of the human brain [8–13], which is created through the processes of training, validation, and testing on a dataset. The model has self-learning, self-adaption, and fault tolerance abilities, and it can completely approach any complex nonlinear relationship, which is not easy for traditional methods. Therefore, it has been widely utilized in the field of water quality and water resources monitoring [14–18]. Although machine learning models have been used in water plants, most water plants still use operator experience to determine flocculation dosing [19,20]. Establishing an effective dosing model will support the development of automatic dosing systems toward full automation and greatly reduce the intensity of manual operation.

In 1960, a neural network model was applied to remediate water pollution occurring in a U.K. sewage plant in Norway and the Shafdan sewage plant in Israel [21–23]. It has also been widely used in many aspects in other countries and regions, which include the prediction of water quality in sewage plants [24], discharge water quality [25], runoff [26], and sedimentation rate [27]. In 2008, 19 of the 27 provincial capital cities in China applied a neural network model to the prediction of operating parameters in the field of drinking water treatment. Different types of models have been utilized, such as the back-propagation (BP) [28], fuzzy [29], general regression (GRNN) [30], radial basis function (RBF) neul [31], and multilayer perceptron (MLP) [32] neural network models. In general, the simulation effect using laboratory data is better than that using industry data. For example, a turbidity prediction using the MLP neural network model attained a coefficient of determination (R^2) of 0.96 [32]; Du et al. [33] applied a neural network model to monitor sewage treatment in a laboratory, for which the R^2 reached 0.99. C. W. Baxter et al. [34] used a full-scale artificial neural network to improve coagulation in removing natural organic matter (NOM) for the Rossdale water treatment plant, where the R^2 was as high as 0.71. A. Najah et al. [35] used an MLP-NN to predict the total dissolved solids in the Johor River Basin, and the R^2 for the tributary was only 0.58. In industry, time-delay, nonlinearity, and multiple influencing factors enable flocculation dosing, and addressing these issues is becoming more complex; thus, enhancing simulation performance has received wide attention. In addition, most of the research on neural network models has focused on its architecture adjustment and algorithm development, whereas creating input data to enhance their performance has rarely been reported.

In order to strengthen the management of water treatment facilities, an effective back-propagation model was established in this study to predict flocculation dosing. The model was improved by particle swarm optimization (PSO), which was carried out using scientific software, MATLAB 2010b. We examined the effects of those parameters that affected prediction performance, including the learning factor, population size, number of generations, and data type.

2. Materials and Methods

2.1. Proposed Architecture of Neural Network

The feedforward and feedback signals are used to control flocculant dosing (see Figure 1). Among them, we needed to test feedforward signals to create a model, and we needed to test feedback signals (effluent quality) to estimate whether effluent quality meets requirement. The flow chart is shown in Figure 1.
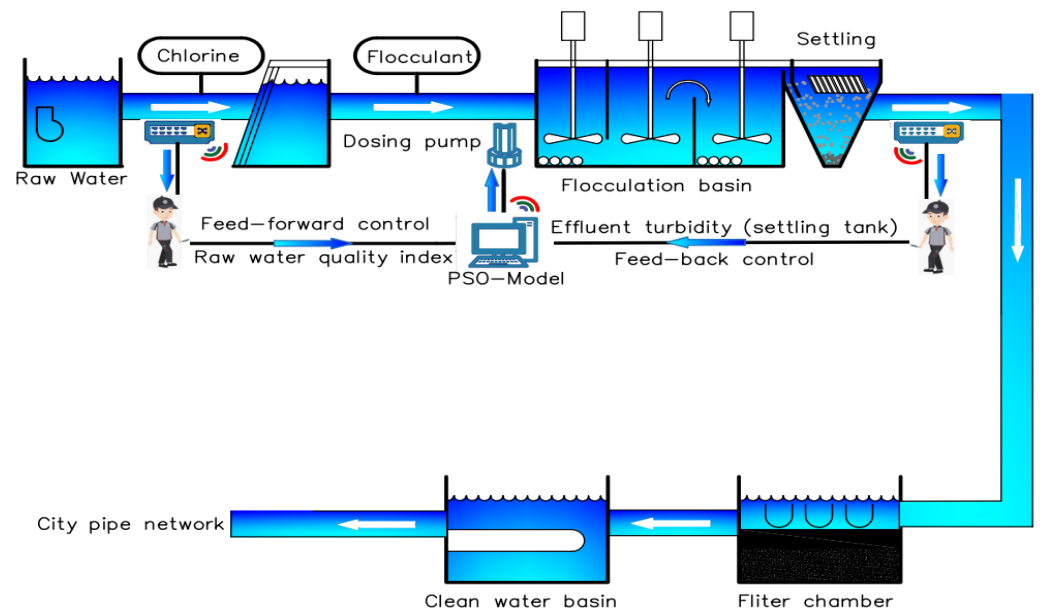
Figure 1. A controlling diagram of flocculant dosing in drinking water plant.

In this study, we used a BP neural network to create the feedforward control model, which included three layers: the input, hidden, and output layers. The main difference between the BP neural network and the traditional model was that two signals and target set values were introduced into the model. Due to some shortcomings, the model was optimized by the particle swarm optimization algorithm, which is discussed in Section 2.2. To predict desired flocculant dosing, we took the feedforward and feedback signals as the input layer values and flocculant dosage as the output layer. A simple scheme to describe the model's architecture is shown in Figure 2. Once the feedforward signal is put into the model and the feedback signal is set as the target value, the flocculant dosing required to reach the target water quality can be calculated.
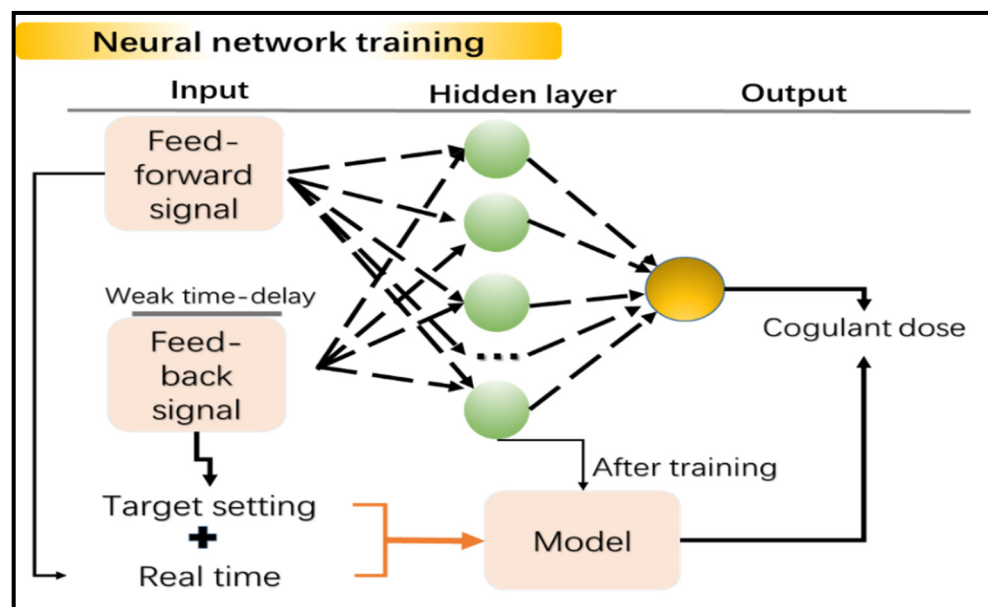

Figure 2. Proposed architecture of neural network model for flocculant dosing.

2.2. Structure Optimization

Some issues, such as falling into local minima, occurred due to the sluggish convergence of the BP neural network [36,37]. We had to find an effective strategy to avoid these

issues, thus improving model performance. Like birds searching for a good route to travel while foraging, we introduced particle swarm optimization (PSO) algorithm. The PSO algorithm adjusts the particle velocity and spatial position to tackle nonlinear problems [38,39], which aims to obtain the best initial weight and thresholds for the model.

Figure 3 shows a calculation diagram of the PSO algorithm. PSO assumes that a population in a D dimensional space is made up of n individual particles. The weights and thresholds of the neural network were joined together as a particle. The position of the ith particle is denoted by $X_i = (X_{i1}, X_{i2}, \ldots, X_{id})$, and the particle speed that corresponds to position i is denoted by $V_i = (V_{i1}, V_{i2}, \ldots, V_{in})$. The optimal position for one particle is represented by $S_i = (S_{i1}, S_{i2}, \ldots, S_{in})$. For all particles, it is expressed by $S_{gd} = (S_{g1}, S_{g2}, \ldots, S_{gn})$. During each iteration process, the particle adjusts its position according to the fitness variation of the current X_i, in order to obtain updated S_i and S_g. The new speed and location are calculated according to Equation (1) [40].

$$V_{id} = WV_{id} + c_1 r_1 (S_{id} - X_{id}) + c_1 r_1 \left(S_{gd} - X_{gd} \right)$$
$$X_{id} = X_{id} + V_{id} \tag{1}$$

where d $(1, 2, \ldots, D)$ is the dimension; W is the weight, which we fixed at one in this study; r_1 and r_2 are random numbers in $[0, 1]$; and c_1 and c_2 are learning factors, usually in $[0,2]$, the values of which were equal in this study.
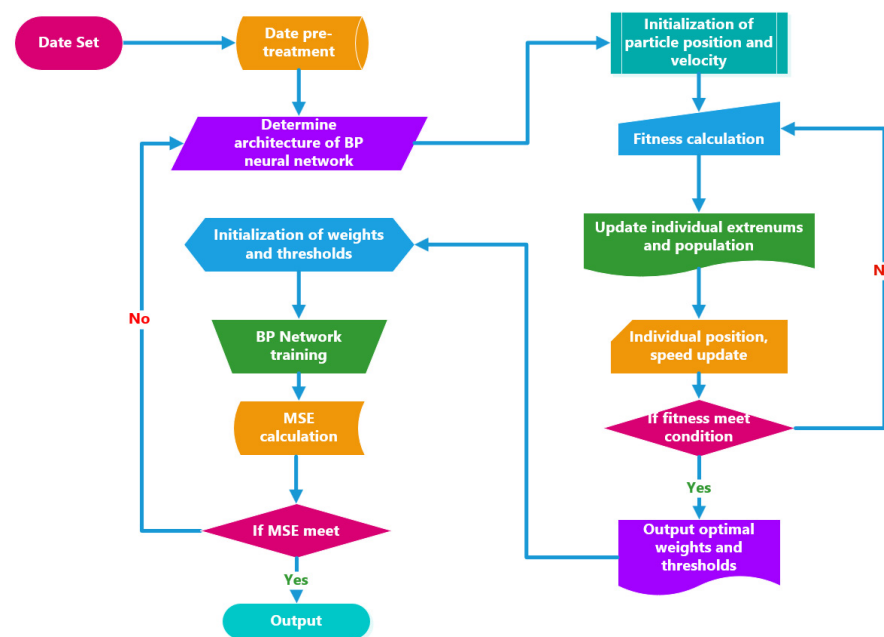

Figure 3. A simple scheme for BP optimization through the PSO algorithm.

Once the criteria condition is satisfied, the iteration is terminated. The criteria condition is calculated by a fitness function, which is expressed by Equation (2) [40].

$$F = \sum_{i=1}^{N} (Y_O - Y_p) \tag{2}$$

where F is the fitness value; N is the number of samples; Y_o is the predicted value; Y_p is the observed value.

2.3. Sampling

The feedforward and feedback signal data were collected from one drinking water plant in Xiangtan, Hunan province, China. The sampling duration was fixed between April and December of 2021. The monthly collecting period was 10–16 days, 8 h every day, in 1 h intervals. The raw water parameters, including temperature (°C), pH, TDS, total

phosphorus (g/L), UV254 (cm^{-1}), flow rate (L/h), and settling tank water turbidity (NTU), were collected. The coagulant (polyaluminum chloride) dosage (L/h) was calculated using the effective flocculation component (Al_2O_3).

2.4. Data Pretreatment

Both training samples and test samples were randomly selected from the total sample at a fixed ratio of 8:2. In order to increase the accuracy, convergence, and consistency of the model and reduce the influence of differences in dimension size among samples, sample data were normalized between −1 and 1, which we calculated by Equation (1) [40].

$$y = \frac{(y_{max} - y_{min})(x - x_{min})}{(x_{max} - x_{min})} + y_{min} \tag{3}$$

where y_{min} is −1; y_{max} is 1; x is a specified variable value; x_{min} is the minimum value of the specified variable, x; and x_{max} is the maximum value of the specified variable, x. When $x_{max} = x_{min}$ or both of them are infinite, $y = x$.

2.5. Accuracy

The criteria indexes for evaluating the model performance were as follows: the coefficient of determination (R^2), probability value (*p*-value) of the results of an independent sample t-test, root mean square error (RMSE) as well as its percent (RMSE%), percent bias (PBIAS), model efficiency (EF), and index of agreement (d), which are calculated using the following equations [41]:

$$R^2 = \frac{\sum_i^N \left(P_i - \bar{O} \right)^2}{\sum_i^N \left(O_i - \bar{O} \right)^2} \tag{4}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^N (P_i - O_i)^2} \tag{5}$$

$$RMSE(\%) = RMSE/\bar{O} \tag{6}$$

$$PBIAS = \sum_{i=1}^N (O_i - P_i)^2 / \sum_{i=1}^N O_i \times 100\% \tag{7}$$

$$EF = 1 - \sum_{i=1}^N (P_i O_i)^2 / \sum_{i=1}^N (\bar{O} - O_i)^2 \tag{8}$$

$$d = \left[\sum_{i=1}^N (P_i - O_i)^2 / \sum_{i=1}^N (|P_i'| + |Q_i'|)^2 \right], (0 \le d \le 1) \tag{9}$$

where N is the number of measured data; P_i and O_i are the predicted value and the measured value, respectively; and $\bar{O}$ is the average observed value.

R^2 is used to demonstrate the relationship between measured and simulated values. If R^2 is close to one, the simulation value better agrees with the measurements [42]. The root mean square error (RMSE) evaluates the model's prediction error. The RMSE% calculates the consistency between the measured value and the simulation value. The model is considered excellent if the RMSE% is less than 10, good if the RMSE% is less than 20, general if the RMSE% is greater than 30, and poor if the RMSE% is greater than 30 [43]. The percent bias (PBIAS) is used to determine whether the predicted value is greater or less than the measured value on average. The best PBIAS value is 0. If the PBIAS is positive, the model tends to underestimate; otherwise, the model tends to overestimate [16]. The EF is used to estimate model performance through a comparison between the measured and simulated values. If EF is positive, it indicates that the simulation value is more reliable

than the mean of the measured value s. If EF is close to zero, it means that the mean value of the measurements is more reliable than that of the simulation [44]. d is used to calculate the fitting effect, the value of which is in the range of 0 to 1. If d is close to one, it indicates that the value of the simulation is more consistent with the value of the measurements, indicating few simulation errors [45].

3. Results and Discussion

In this study, we investigated the effect of the parameters that affected simulation performance, such as the learning factor, number of generations, and size of the population. The input and output data consisted of 586 training samples and 147 testing samples. These input and output data were collected at the same time points (called time-delay data).

3.1. Effect of Learning Factor

The effect of the learning factor on the simulation was investigated under these simulation conditions: 8 input factors, 6 hidden layer nodes, 1 output layer node, 20 generations, population size of 5, and learning factors of 0.175–5.6. Figure 4 shows the results of the learning factor effect on the prediction of flocculant dosing, the prediction accuracy of which was evaluated by measuring the R^2 between the measured and simulated values.
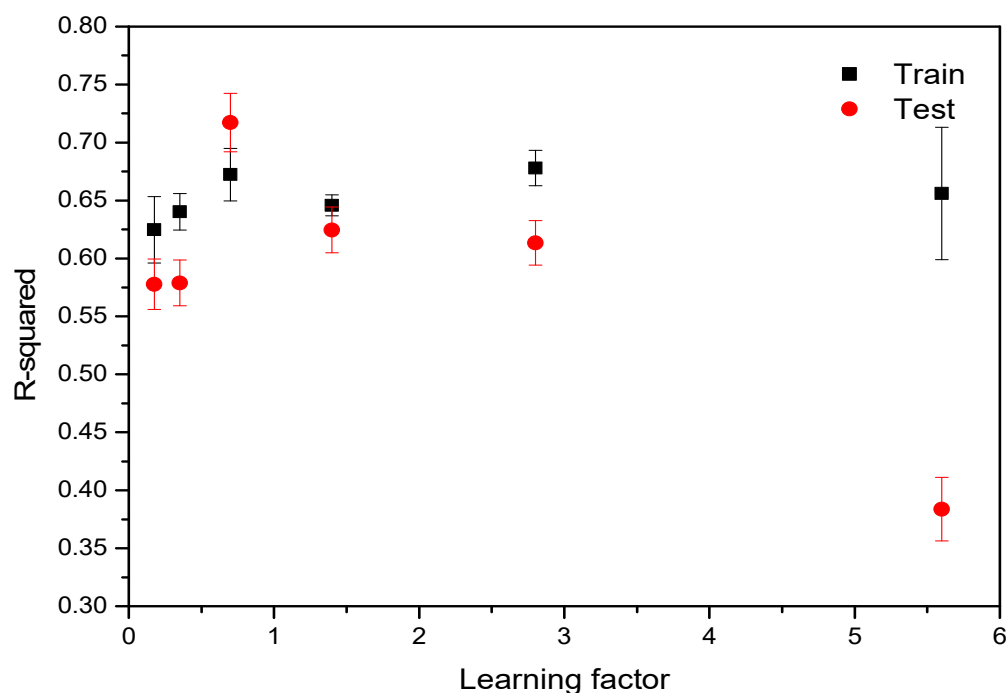

Figure 4. Effect of learning factor on variations in R^2 between the measurements and simulated values.

Figure 4 shows the effect of learning factor on the variations in R^2. The results showed that the variations in R^2 for the training results were essentially the same as those in the test results. Except for the effect at a learning factor of 0.7, the R^2 in the training results was better than the test results. The R^2 in training results gradually increased as the learning factor increased, then stabilized and finally decreased. Increasing the learning factor did not improve simulation accuracy, but it caused serious over-fitting. When the learning factor was 1.4, the ratio of the training accuracy to the test accuracy (R^2_{train}/R^2_{test}) was close, and the model was neither over- nor under-fit. However, with a higher learning rate, such as 5.6, a higher ratio occurred over one such as $R^2_{train}/R^2_{test} = 1.7$. Therefore, we selected a learning rate of 1.4 as the optimized value to further examine the influence of PSO's parameters on simulation performance.

3.2. Effect of Generations

The effect of the number of generations on the simulation was investigated under these simulation conditions: 8 input factors, 6 hidden layer nodes, 1 output node, learning factor of 1.4, population size of 5, and 5–160 generations. Figure 5 shows the results of the generations' effect on the prediction of flocculant dosing, the prediction accuracy of which was evaluated by measuring the R^2 between the measurements and simulated values.
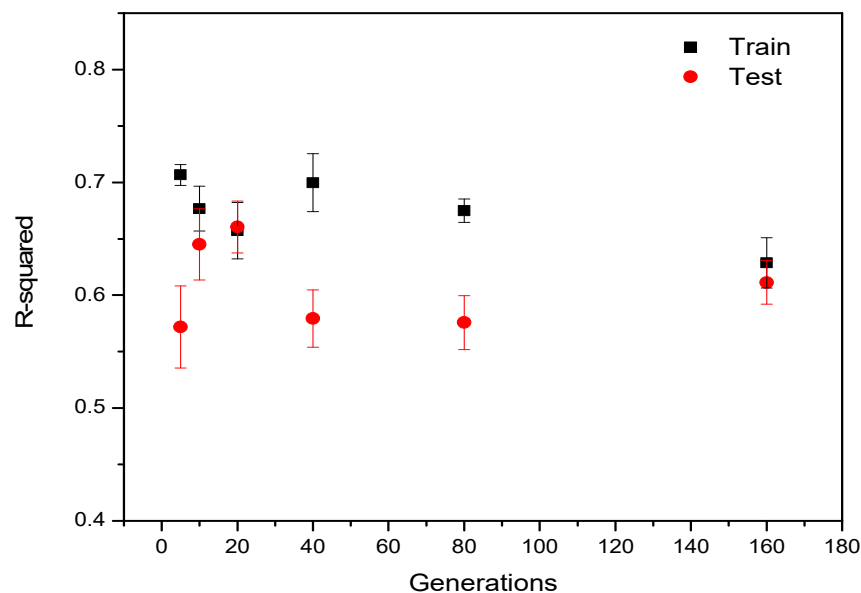

Figure 5. Effect of the number of generations on the variations in R^2 between the measurements and simulation.

Figure 5 shows the effect of the number of generations on the variations in R^2 between the measurements and simulation. It showed that the variations in R^2 in the training results basically followed those in the testing results as well. However, the R^2 in the training results was higher than that in the testing results. With the increase in the number of generations, the R^2 in the training results decreased first, then increased, and finally decreased. The increase in the number of generations was not conducive to the enhancement in model accuracy. When the number of generations was 20, the ratio of the training accuracy to the test accuracy (R^2_{train}/R^2_{test}) was close, and over- or under-fitting did not appear. However, when the number of generations was too high (e.g., 80 for $R^2_{train}/R^2_{test} = 1.2$) or too low (e.g., 5 for $R^2_{train}/R^2_{test} = 1.24$), the R^2_{train}/R^2_{test} ratio was significantly higher than one, and serious over-fitting occurred. The optimal number of the generations was fixed at 20 in this study.

3.3. Effect of Population Size

The effect of the number of the population size on the simulation performance was investigated under these simulation conditions: 8 input factors, 6 hidden layer nodes, 1 output node, learning factor of 1.4, 20 evolutions, 20 generations, and population size of 5–80. Figure 6 shows the results of the population size effect on prediction of flocculant dosing, the prediction accuracy of which was evaluated by measuring the R^2 between the measurements and the simulation values.

Figure 6 shows the effect of the population size on the variations in R^2 between the measurements and the simulation. In both the training and testing results, the variations in R^2 were similar. The R^2 had better results in the training than that in the test as a whole. The R^2 in the training results decreased first and subsequently increased with the increase in the population size. The increased population size did not result in a higher R^2. With a population size of 20, the ratio of the training accuracy to the test accuracy (R^2_{train}/R^2_{test}) was closer, and over- or under-fitting did not appear. At a population size

of 40, the R^2_{train}/R^2_{test} ratio over one generated serious over-fitting. It was better to select a population size of 20 as the optimal value for the simulation.
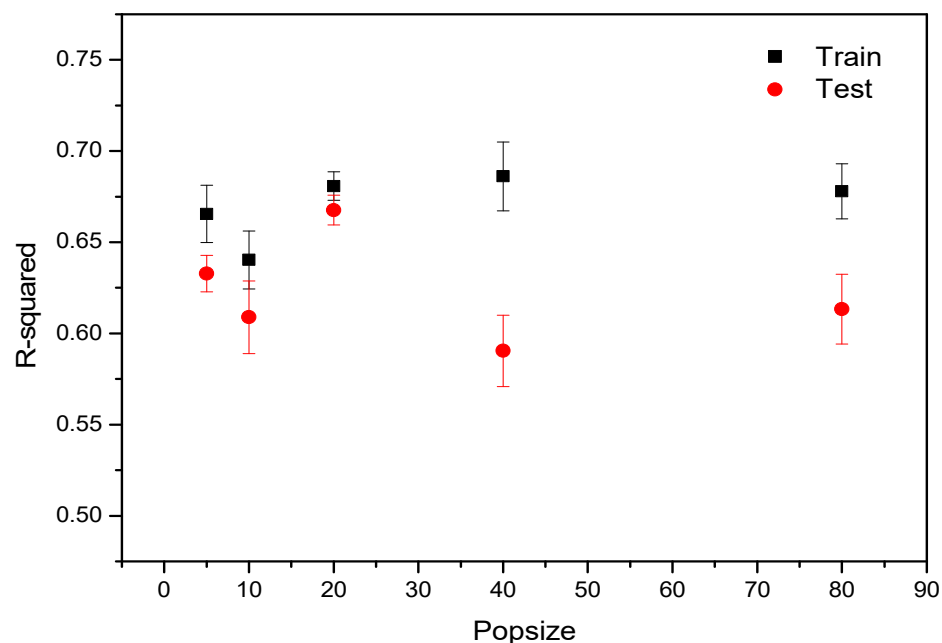

Figure 6. Effect of population size on the variations in R^2 between the measurements and simulation.

3.4. Effect of Weak Time-Delay

Flocculation processes always need a certain time to complete; therefore, an effluent test has to be conducted after the completion of flocculation. However, those effluent quality parameters (denoted time-delay data in this study) tested at the time that flocculant is added do not reflect the real flocculation result of the added flocculant. This is called the time-lag effect of flocculation.

Most neural networks used time-delay data as the input, so do not consider the impact of the time delay. Because varying the learning factor, number of generations, and population size further increased the simulation accuracy, we tried to reduce the impact to increase simulation accuracy. According to engineering experience, the real flocculation result appears in one hour. Therefore, we carried out a simulation using raw water parameters and effluent quality parameters after flocculation for one hour. This is called a weak time-delay simulation. There were a total of 447 training samples and 113 test samples used in this study.

3.4.1. Result of Weak Time-Delay Data Training

A comparison was made between the weak time-delay simulation and the time-delay simulation. Those conditions for simulation that were used in this study included 8 input variables, 6 hidden layer nodes, 1 output node, a learning factor of 1.4, 20 generations, and a population size 20. The results are shown in Figure 7.

Figure 7a,b shows that with time-delay signal data, the R^2 values were 0.68 for training and 0.67 for testing. Their *p* values were 0.827 and 0.819, respectively. This demonstrated that there was a nonsignificant difference between the simulated and measured values. Similar variations between them were also examined (see Figure 7e,f). With weak time-delay signal data, the R^2 increased to 0.73 (see Figure 7c,d), and the *p* values were 0.855 and 0.856, respectively. Additionally, there was no significant difference occurring between them, and their variation trend was nearly the same (see Figure 7g,h). However, the accuracy was enhanced, as indicated by the R^2.
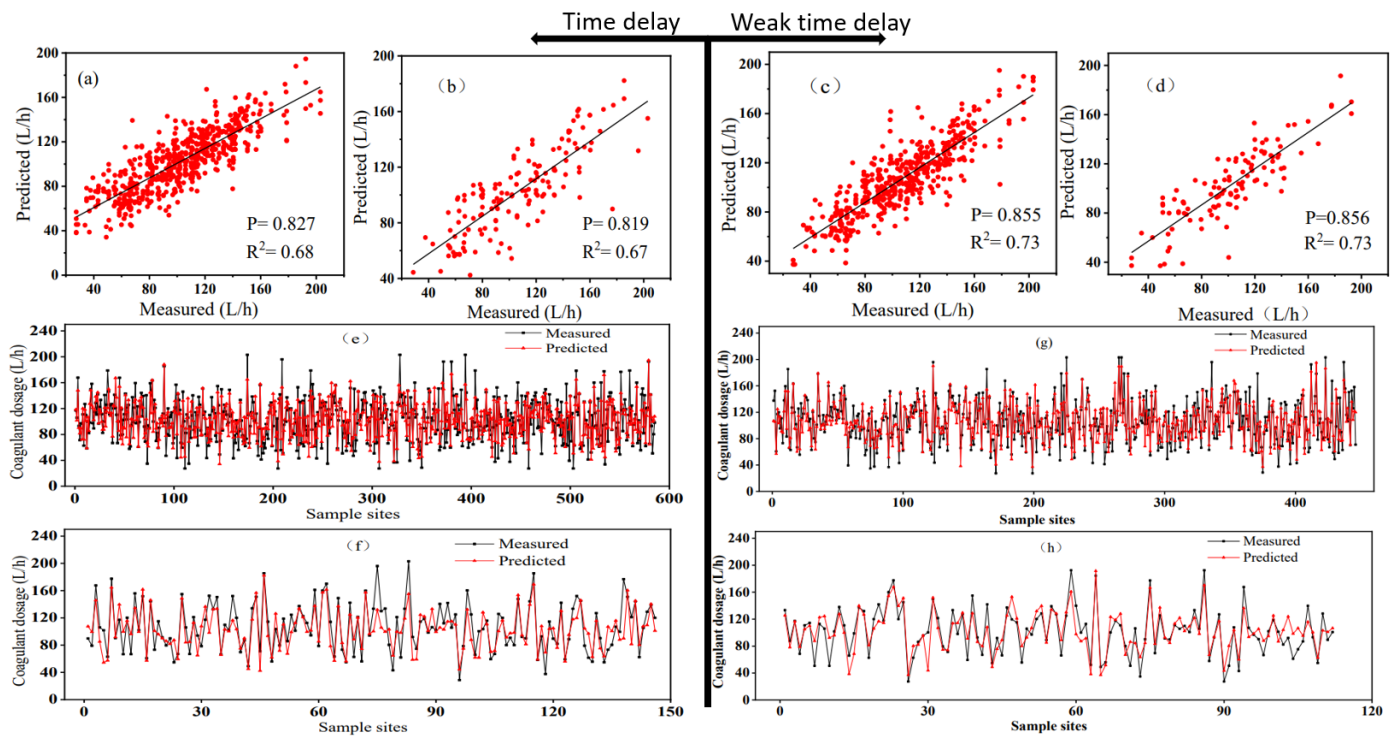
Figure 7. Correlation plots of simulated and measured values of (**a**,**b**) time-delay signal training and testing and (**c**,**d**) weak time-delay signal training and testing; distribution plots of simulated and measured values of (**e**,**f**) time-delay signal training and testing and (**g**,**h**) weak time-delay signal training and testing.

More evaluation indicators were compared between the two simulations. The results are shown in Table 1. Table 1 shows that their PBIAS values were zero, indicating that the average trends in their predicted and measured values were neither high nor low, and their EF values were positive, indicating that the simulation values were more reliable than the mean value of the measurements. The simulation with weak time-delay data showed a lower RMSE (around 18.15) and RMSE% (<18%) values, the d value of which was closer to one, indicating that it was better than the time-delay simulation. The main reason for the improvement was attributed to the time-delay effect. Therefore, using weak time-delay data to reflect flocculation results was better, and it had a good result in the simulation of flocculant dosing.

Table 1. Evaluation indicators of weak time-delay data training and time-delay data training.

Type	R^2	RMSE	RMSE%	PBIAS	EF	d
Weak time-delay train	0.73	17.97	16.90	0.00	0.73	0.91
Weak time-delay test	0.73	18.15	17.76	0.00	0.73	0.91
Time-delay train	0.68	19.08	18.58	0.00	0.68	0.89
Time-delay test	0.67	21.10	19.73	0.00	0.65	0.89

3.4.2. Validation

With the weak time-delay data, we examined the validation results of the model as indicated by mean squared normalized error (MSE). The simulation results are shown in Figure 8. Figure 8 shows that the variations in MSE were nearly the same. There were no significant differences among training, testing, and validation. Over- and under-fitting did occur in the simulation. The best validation performance was achieved at epoch 11. It was feasible for us to use the model to predict flocculant dosing.
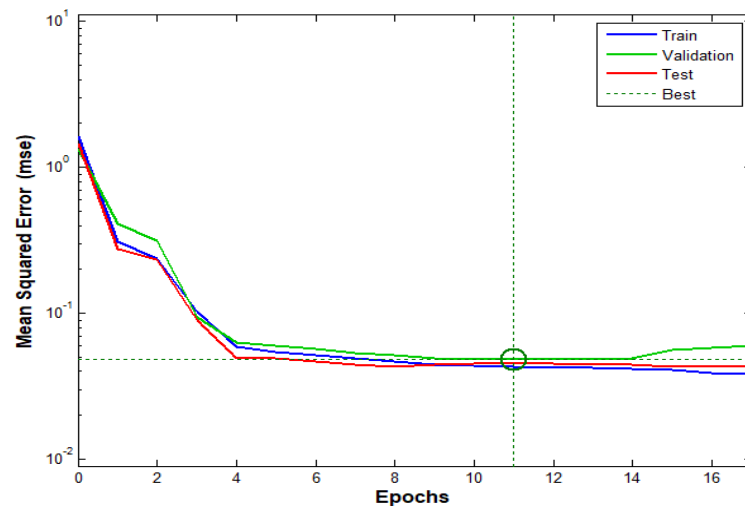
Figure 8. Validation performance of the model.

3.4.3. Variations in PSO's Fitness and Accuracy

As demonstrated in the previous section, the result for the weak time-delay simulation was better than that of the time-delay simulation. The fitness value with the weak time-delay data was significantly lower than that of the time-delay data (see Figure 9a), which showed that those simulations with the two kinds of data were different, and the weak time-delay data could be better applied to the parameter optimization by PSO. In general, using more data is more conducive to simulation. Although the amount of weak delay data was small, they still produced better results. This indicated that the weak time-delay data were different from the time-delay data, which better reflected the real system.
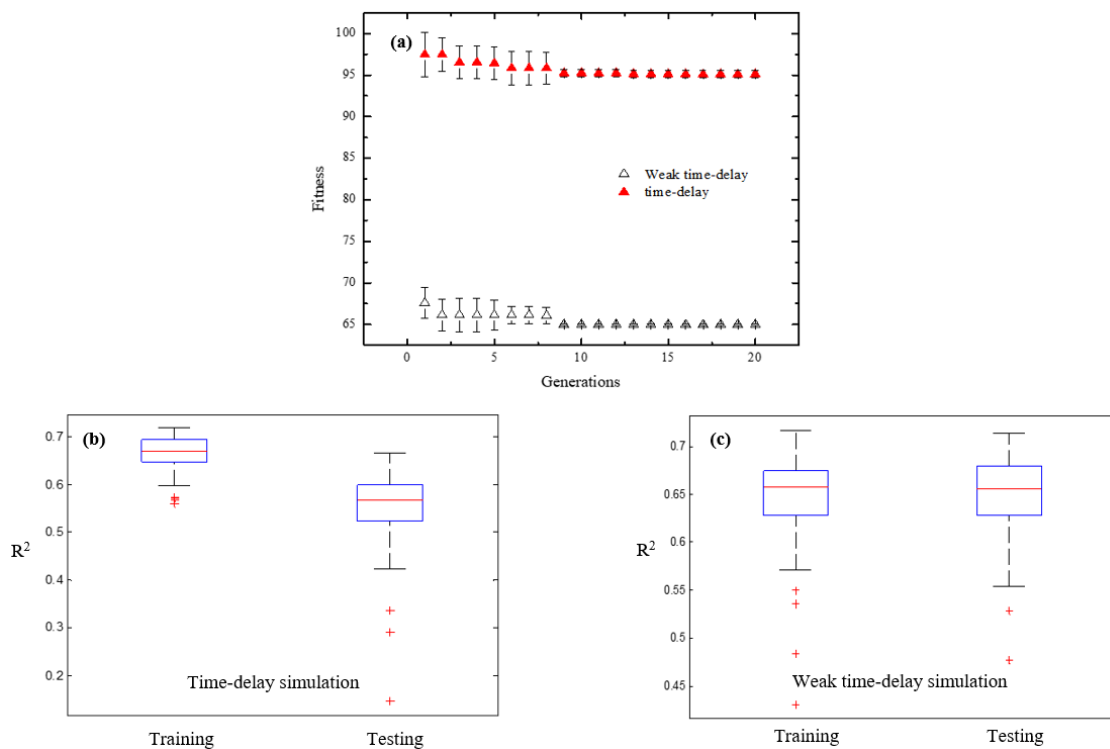

Figure 9. (**a**) The fitness of particles affected by the time-delay data and the weak time signals by varying the number of the generations, and the box plots of the variation in R^2 with (**b**) time-delay data and (**c**) weak time-delay data.

In addition, we performed 100 training repetitions and found that the results of R^2 with the weak time-delay data better agreed with the testing data, as indicated by the box plots in Figure 9b,c. The chart for the 100 training repetitions indicates that the training results were more consistent with the testing result with the weak time-delay data.

3.4.4. Sensitivity Analysis

Coagulation is affected by various factors, such as temperature, pH, and turbidity. Therefore, we examined the sensitivity of the model to these input variables using the Olden algorithm. The Olden algorithm determines the importance of the contribution of the factors to the flocculant dosing via the weight of the neural network model [46], which is expressed using the following equation.

$$S_i = \frac{\sum_{k=1}^{Y} w_{ik} v_k}{\sum_{i=1}^{X} \sum_{k=1}^{Y} w_{ik} v_k} \tag{10}$$

where S_i denotes the ith input neuron's sensitivity; w_{ik} denotes the weight of the connection between the ith input neuron and the kth hidden layer neuron; and v_k denotes the weight of the connection between the k_{th} hidden layer neuron and the output neurons. The number of neurons in the input layer is denoted by X, whereas the number of neurons in the hidden layer is denoted by Y.

Figure 10 shows the results of the sensitivity of the model to the input variables. It shows that the raw water turbidity, flow rate, temperature, and TDS were more important to the model than other factors, especially the raw water turbidity. Coagulation in a drinking water plant is not completely the same as that in a laboratory. Some factors, such as pH, are very stable. The range of adjustment in the laboratory is wider than that in the water plant, so the impact of these factors in a drinking water plant is relatively weaker. Flow rate, temperature, and turbidity were the most important factors, which strongly contributed to the model. Usually, researchers pay more attention to the laboratory scenario and ignore actual engineering situations. According to the layout of the laboratory, it may increase costs. We cannot rely solely on laboratories to analyze the situations of water plants. Based on the research results of this study, we better understand the importance of these factors and adjusting flocculant dosing to reduce the cost. Therefore, these results have reference value for creating models.
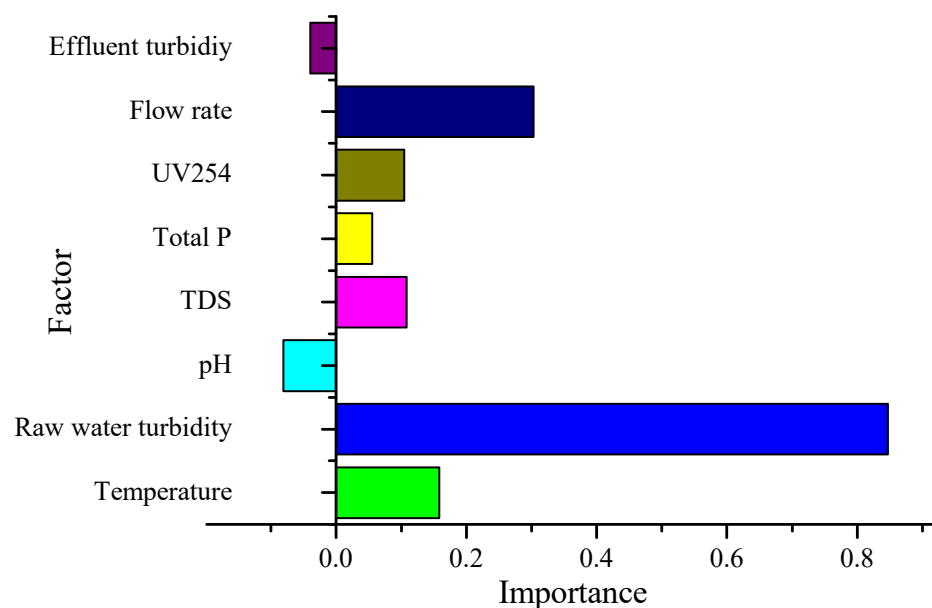

Figure 10. The results of sensitivity analysis of the model to input variables.

4. Conclusions

There are various difficulties experienced in the control of flocculant dosing, including the interference of multiple factors and the time-delay of flocculation. Using an intelligent control model may improve the accuracy of flocculant dosing, thus avoiding those difficulties. In this study, we created a BP neural network model and used PSO and weak time-delay data to improve the model for prediction. The main conclusions are as follows: it was effective to use hybridized feedforward and feedback signals as input data to create the model; adjusting the learning factor, number of generations, and population size produced good results, including an R^2 up to 0.68 and an RMSE between 18% and 20%; weak time-delay data had a better effect on the simulation, which increased the R^2 to 0.73 and reduced the RMSE to lower than 18%. These results are helpful for establishing an effective neural network model and improving water plant management. It is extremely rare to improve the model performance through data type. This study proved its effectiveness, and in future work, we will strengthen the use of weak-delay data and pay more attention to research on the role of data type.

Author Contributions: Conceptualization and supervision, G.Z.; formal analysis; B.L.; investigation, H.L.; resources, F.Y., X.L., C.Y., W.H., Q.J. and X.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data cannot be made publicly available; readers should contact the corresponding author for details.

Conflicts of Interest: The authors declare there is no conflict of interest.

References

1. Zhang, Y.; Gao, X.; Smith, K.; Inial, G.; Liu, S.; Conil, L.B.; Pan, B.K. Integrating water quality and operation into prediction of water production in drinking water treatment plants by genetic algorithm enhanced artificial neural network. *Water Res.* **2019**, *164*, 114888. [CrossRef] [PubMed]
2. Dayarathne, H.N.P.; Angove, M.J.; Aryal, R.; Abuel-Naga, H.; Mainali, B. Removal of natural organic matter from source water: Review on coagulants, dual coagulation, alternative coagulants, and mechanisms. *J. Water Process Eng.* **2021**, *40*, 101820. [CrossRef]
3. Li, X.; Cheng, Z.; Yu, Q.; Bai, Y.; Li, C. Water-Quality Prediction Using Multimodal Support Vector Regression: Case Study of Jialing River, China. *J. Environ. Eng.* **2017**, *143*, 04017070. [CrossRef]
4. Choo, G.; Oh, J.-E. Seasonal occurrence and removal of organophosphate esters in conventional and advanced drinking water treatment plants. *Water Res.* **2020**, *186*, 116359. [CrossRef]
5. Guo, H.; Jeong, K.; Lim, J.; Jo, J.; Kim, Y.M.; Park, J.-P.; Kim, J.H.; Cho, K.H. Prediction of effluent concentration in a wastewater treatment plant using machine learning models. *J. Environ. Sci.* **2015**, *32*, 90–101. [CrossRef]
6. Bai, Y.; Li, C. Daily natural gas consumption forecasting based on a structure-calibrated support vector regression approach. *Energy Build.* **2016**, *127*, 571–579. [CrossRef]
7. Sezen, C.; Bezak, N.; Bai, Y.; Šraj, M. Hydrological modelling of karst catchment using lumped conceptual and data mining models. *J. Hydrol.* **2019**, *576*, 98–110. [CrossRef]
8. Xiao, X.; Mo, H.; Zhang, Y.; Shan, G. Meta-ANN—A dynamic artificial neural network refined by meta-learning for Short-Term Load Forecasting. *Energy* **2022**, *246*, 123418. [CrossRef]
9. Bai, Y.; Sun, Z.; Zeng, B. A multi-pattern deep fusion model for short-term bus passenger flow forecasting. *Appl. Soft Comput.* **2017**, *58*, 669–680. [CrossRef]
10. Hassan, W.H.; Hussein, H.H.; Alshammari, M.H.; Jalal, H.K.; Rasheed, S.E. Evaluation of gene expression programming and artificial neural networks in PyTorch for the prediction of local scour depth around a bridge pier. *Results Eng.* **2022**, *13*, 100353. [CrossRef]
11. Hassan, W.H.; Attea, Z.H.; Mohammed, S.S. Optimum layout design of sewer networks by hybrid genetic algorithm. *J. Appl. Water Eng. Res.* **2020**, *8*, 2324–9676. [CrossRef]

12. Jalal, H.K.; Hassan, W.H. Effect of bridge pier shape on depth of scour[C], IOP Conference Series: Materials Science and En-gineering. *IOP Publ.* **2020**, *671*, 012001. [CrossRef]

13. Odili, J.B.; Noraziah, A.; Babalola, A.E. A new fitness function for tuning parameters of Peripheral Integral Derivative Controllers. *ICT Express* 2021, *in press*. [CrossRef]

14. Zhu, G.C.; Xiong, N.N.; Wang, C.; Li, Z.; Hursthouse, A.S. Application of a new HMW framework derived ANN model for optimization of aquatic dissolved organic matter removal by coagulation. *Chemosphere* **2021**, *262*, 127723. [CrossRef]

15. Marzouk, M.; Elkadi, M. Estimating water treatment plants costs using factor analysis and artificial neural networks. *J. Clean. Prod.* **2016**, *112*, 4540–4549. [CrossRef]

16. Li, K.; Li, L.; Qin, J.; Liu, X. A facile method to enhance UV stability of PBIA fibers with intense fluorescence emission by forming complex with hydrogen chloride on the fibers surface. *Polym. Degrad. Stab.* **2016**, *128*, 278–285. [CrossRef]

17. Godo-Pla, L.; Emiliano, P.; Valero, F.; Poch, M.; Sin, G.; Monclús, H. Predicting the oxidant demand in full-scale drinking water treatment using an artificial neural network: Uncertainty and sensitivity analysis. *Process Saf. Environ. Prot.* **2019**, *125*, 317–327. [CrossRef]

18. Li, L.; Rong, S.; Wang, R.; Yu, S. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review. *Chem. Eng. J.* **2021**, *40*, 126673. [CrossRef]

19. Agudosi, E.S.; Abdullah, E.C.; Mubarak, N.M.; Khalid, M.; Pudza, M.Y.; Agudosi, N.P.; Abutu, E.D. Pilot study of in-line: Continuous flocculation water treatment plant. *J. Environ. Chem. Eng.* **2018**, *6*, 7185–7191. [CrossRef]

20. Katrivesis, F.K.; Karela, A.D.; Papadakis, V.G.; Paraskeva, C.A. Revisiting of coagulation-flocculation processes in the production of potable water. *J. Water Process Eng.* **2019**, *27*, 193–204. [CrossRef]

21. Alharbi, M.; Hong, P.-Y.; Laleg-Kirati, T.-M. Sliding window neural network bafsed sensing of bacteria in wastewater treatment plants. *J. Process Control* **2022**, *110*, 35–44. [CrossRef]

22. Weil, M.; Mandelboim, M.; Mendelson, E.; Manor, Y.; Shulman, L.; Ram, D.; Barkai, G.; Shemer, Y.; Wolf, D.; Kraoz, Z.; et al. Human enterovirus D68 in clinical and sewage samples in Israel. *J. Clin. Virol.* **2017**, *86*, 52–55. [CrossRef] [PubMed]

23. Newhart, K.B.; Holloway, R.W.; Hering, A.S.; Cath, T.Y. Data-driven performance analyses of wastewater treatment plants: A review. *Water Res.* **2017**, *157*, 498–513. [CrossRef]

24. Wang, J.H.; Zhao, X.L.; Guo, Z.W.; Yan, P.; Gao, X.; Shen, Y.; Chen, Y.P. A full-view management method based on artificial neural networks for energy and material-savings in wastewater treatment plants. *Environ. Res.* **2022**, *211*, 113054. [CrossRef]

25. García-Alba, J.; Bárcena, J.F.; Ugarteburu, C.; García, A. Artificial neural networks as emulators of process-based models to analyse bathing water quality in estuaries. *Water Res.* **2019**, *150*, 283–295. [CrossRef]

26. Sharma, S.K.; Tiwari, K.N. Bootstrap based artificial neural network (BANN) analysis for hierarchical prediction of monthly runoff in Upper Damodar Valley Catchment. *J. Hydrol.* **2009**, *374*, 209–222. [CrossRef]

27. Lin, D.; Hu, L.; Bradford, S.A.; Zhang, X.; Lo, I.M.C. Prediction of collector contact efficiency for colloid transport in porous media using Pore-Network and Neural-Network models. *Sep. Purif. Technol.* **2022**, *290*, 120846. [CrossRef]

28. Onukwuli, O.D.; Nnaji, P.C.; Menkiti, M.C.; Anadebe, V.C.; Oke, E.O.; Ude, C.N.; Ude, C.J.; Okafor, N.A. Dual-purpose optimization of dye-polluted wastewater decontamination using bio-coagulants from multiple processing techniques via neural intelligence algorithm and response surface methodology. *J. Taiwan Inst. Chem. Eng.* **2021**, *125*, 372–386. [CrossRef]

29. Huang, M.; Ma, Y.; Wan, J.; Wang, Y. Simulation of a paper mill wastewater treatment using a fuzzy neural network. *Expert Syst. Appl.* **2009**, *36*, 5064–5070. [CrossRef]

30. Zhang, X.; He, X.; Wei, M.; Li, F.; Hou, P.; Zhang, C. Magnetic flocculation treatment of coal mine water and a comparison of water quality prediction algorithms. *Mine Water Environ.* **2019**, *38*, 391–401. [CrossRef]

31. Zheng, H.; Zhu, G.; Jiang, S.; Tshukudu, T.; Xiang, X.; Zhang, P.; He, Q. Investigations of coagulation–flocculation process by performance optimization, model prediction and fractal structure of flocs. *Desalination* **2011**, *269*, 148–156. [CrossRef]

32. Zangooei, H.; Asadollahfardi, G.; Delnavaz, M. Prediction of coagulation and flocculation processes using ANN models and fuzzy regression. *Water Sci. Technol.* **2016**, *74*, 1296–1311. [CrossRef] [PubMed]

33. Du, J.; Shang, X.; Shi, J.; Guan, Y. Removal of chromium from industrial wastewater by magnetic flocculation treatment: Experimental studies and PSO-BP modelling. *J. Water Process Eng.* **2022**, *47*, 102822. [CrossRef]

34. Baxter, C.W.; Stanley, S.J.; Zhang, Q. Development of a full-scale artificial neural network model for the removal of natural organic matter by enhanced coagulation. *Aqua* **1999**, *48*, 129–136. [CrossRef]

35. Najah, A.; El-Shafie, A.; Karim, O.A.; El-Shafie Amr, H. Application of artificial neural networks for water quality prediction. *Neural Comput. Appl.* **2013**, *22*, 187–201. [CrossRef]

36. Liu, H. Optimal selection of control parameters for automatic machining based on BP neural network. *Energy Rep.* **2022**, *8*, 7016–7024. [CrossRef]

37. Xing, J.; Luo, K.; Pitsch, H.; Wang, H.; Bai, Y.; Zhao, C.; Fan, J. Predicting kinetic parameters for coal devolatilization by means of Artificial Neural Networks. *Proc. Combust. Inst.* **2019**, *37*, 2943–2950. [CrossRef]

38. Zhou, Z.; Gong, H.; You, J.; Liu, S.; He, J. Research on compression deformation behavior of aging AA6082 aluminum alloy based on strain compensation constitutive equation and PSO-BP network model. *Mater. Today Commun.* **2021**, *28*, 102507. [CrossRef]

39. Zou, X.F.; Hu, Y.J.; Long, X.B.; Huang, L.Y. Prediction and optimization of phosphorus content in electroless plating of Cr12MoV die steel based on PSO-BP model. *Surf. Interfaces* **2020**, *18*, 100443. [CrossRef]

40. Li, S.; Fan, Z. Evaluation of urban green space landscape planning scheme based on PSO-BP neural network model. *Alex. Eng. J.*
     2022, *61*, 7141–7153. [CrossRef]
41. Kuzenkov, O.; Kuzenkova, G. Identification of the Fitness Function using Neural Networks. *Procedia Comput. Sci.* **2020**, *169*,
     692–697. [CrossRef]
42. Edelmann, D.; Móri, T.F.; Székely, G.J. On relationships between the Pearson and the distance correlation coefficients. *Stat. Probab.
     Lett.* **2021**, *169*, 108960. [CrossRef]
43. Karunasingha, D.S.K. Root mean square error or mean absolute error? Use their ratio as well. *Inf. Sci.* **2022**, *585*, 609–629.
     [CrossRef]
44. Al-Swaidani, A.M.; Khwies, W.T.; Al-Baly, M.; Lala, T. Development of multiple linear regression, artificial neural networks and
     fuzzy logic models to predict the efficiency factor and durability indicator of nano natural pozzolana as cement additive. *J. Build.
     Eng.* **2022**, *52*, 104475. [CrossRef]
45. Zhao, S.; Xu, W.; Chen, L. The modeling and products prediction for biomass oxidative pyrolysis based on PSO-ANN method:
     An artificial intelligence algorithm approach. *Fuel* **2022**, *312*, 122966. [CrossRef]
46. Olden, J.D.; Joy, M.K.; Death, R.G. An accurate comparison of methods for quantifying variable importance in artificial neural
     networks using simulated data. *Ecol. Model.* **2004**, *178*, 389–397. [CrossRef]