

## Research Article

# A Hybrid Fuzzy Wavelet Neural Network Model with Self-Adapted Fuzzy $c$ -Means Clustering and Genetic Algorithm for Water Quality Prediction in Rivers

Mingzhi Huang<sup>1,2</sup>, Di Tian,<sup>2</sup> Hongbin Liu,<sup>3</sup> Chao Zhang,<sup>2</sup> Xiaohui Yi,<sup>1</sup> Jiannan Cai,<sup>4</sup> Jujun Ruan,<sup>5</sup> Tao Zhang,<sup>5</sup> Shaofei Kong,<sup>6</sup> and Guangguo Ying<sup>1</sup>

<sup>1</sup>Environmental Research Institute, Key Laboratory of Theoretical Chemistry of Environment Ministry of Education, South China Normal University, Guangzhou 510631, China

<sup>2</sup>School of Geography and Planning, Guangdong Provincial Key Laboratory of Urbanization and Geo-Simulation, Sun Yat-sen University, Guangzhou 510275, China

<sup>3</sup>Jiangsu Co-Innovation Center of Efficient Processing and Utilization of Forest Resources, Nanjing Forestry University, Nanjing 210037, China

<sup>4</sup>Zhongshan Environmental Monitoring Station, Zhongshan 528400, China

<sup>5</sup>School of Environmental Science and Engineering, Sun Yat-Sen University, Guangzhou 510275, China

<sup>6</sup>Department of Atmospheric Sciences, School of Environmental Studies, China University of Geosciences (Wuhan), Wuhan 430074, China

Correspondence should be addressed to Mingzhi Huang; [hzm2002xa@163.com](mailto:hzm2002xa@163.com), Jujun Ruan; [ruanjujun@mail.sysu.edu.cn](mailto:ruanjujun@mail.sysu.edu.cn), and Tao Zhang; [zhangt47@mail.sysu.edu.cn](mailto:zhangt47@mail.sysu.edu.cn)

Received 26 April 2018; Revised 18 July 2018; Accepted 13 August 2018; Published 12 December 2018

Academic Editor: Hassan Zargarzadeh

Copyright © 2018 Mingzhi Huang 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.

Water quality prediction is the basis of water environmental planning, evaluation, and management. In this work, a novel intelligent prediction model based on the fuzzy wavelet neural network (FWNN) including the neural network (NN), the fuzzy logic (FL), the wavelet transform (WT), and the genetic algorithm (GA) was proposed to simulate the nonlinearity of water quality parameters and water quality predictions. A self-adapted fuzzy  $c$ -means clustering was used to determine the number of fuzzy rules. A hybrid learning algorithm based on a genetic algorithm and gradient descent algorithm was employed to optimize the network parameters. Comparisons were made between the proposed FWNN model and the fuzzy neural network (FNN), the wavelet neural network (WNN), and the neural network (ANN). The results indicate that the FWNN made effective use of the self-adaptability of NN, the uncertainty capacity of FL, and the partial analysis ability of WT, so it could handle the fluctuation and the nonseasonal time series data of water quality, while exhibiting higher estimation accuracy and better robustness and achieving better performances for predicting water quality with high determination coefficients  $R^2$  over 0.90. The FWNN is feasible and reliable for simulating and predicting water quality in river.

## 1. Introduction

As environmental pollution becomes increasingly serious, water quality (WQ) has attracted extensive attention. Many countries have put forward a variety of pollution prevention measures to improve WQ and to protect the ecological water requirements [1]. Online analytic instruments include important information from the state of the river system, so

it can be utilized by operators to apply real-time management strategies and solve the problem of water pollution. However, due to the lack of suitable online analytic instrument (existing lag and instability) [2] for measuring chemical oxygen demand (COD) and turbidity (TU), the effective monitoring and control for WQ especially are hampered in the river system. Hence, in order to overcome these problems, it is necessary to develop a software sensor model to estimate

hard-to-measure variables from other online measurable variables. The modeling nonlinearity of WQ parameters and WQ prediction based on soft-measurement theory are important issues in water resource management and environmental engineering [3]. There are many mathematical modeling methods that have been used to forecast the long-term water demand [4, 5]. The traditional mathematical modeling technologies are restricted for applications of actual large-scale complex changing processes. The mathematical models could be set up, but their models are complex, so it is difficult to handle the nonlinear characteristic of water quality [6, 7].

In order to overcome the shortcomings of the traditional mathematical modeling methods, intelligence techniques, such as the artificial neural network (ANN), the fuzzy logic (FL), and the wavelet transform (WT), were developed to model the nonlinearity of the WQ management process [8–13]. The neural network (NN) is capable of modeling the complex relationships between the input and the output parameters without requiring a detailed mechanistic description of the phenomena [14], so it could be used to model the WQ management process. Faruk [15] used a backpropagation neural network to predict the WQ using 108-month WQ observational data, including water temperature, boron, and dissolved oxygen, in the Büyük Menderes River, Turkey. The results indicated that the model captures the nonlinear nature of the complex time series and produces more accurate predictions. Another continuous flow system performance was simulated using ANN by Ahmed [16] who developed a feedforward NN model and a radial basis function NN model to predict the dissolved oxygen from the biochemical oxygen demand (BOD) and the chemical oxygen demand (COD) in the Surma River. Although the ANN is successfully used for the modeling WQ management process, the ANN schemes still have a low convergence rate and a local minimum [17].

To overcome the shortcomings of the ANN, various types of novel integrated intelligent networks have been developed, such as the FNNs [2, 18–20] and the wavelet neural networks (WNNs) [21]. WNN is a new branch of ANN theory research that is constructed by taking the wavelet transform as the preprocessor of the ANN. Wavelet transform has a wide range of applications in the processing of complex systems due to its time-frequency localization characteristics. The WNN is better than the traditional ANN in modeling precise and convergence rates and learning memory ability and accuracy [22, 23]. On the basis of the various online input variables, an ensemble modeling approach based on the wavelet analysis model was designed to forecast reservoir inflow [24]. A wavelet-linear genetic programming model was proposed to monitor sodium ( $\text{Na}^+$ ) concentration in rivers [25]. The WNN has difficulty in overcoming the subjectivity of selecting fuzzy rules by human, which is one of the characteristics of the FL. Gharibi et al. developed a WQ index based on the FL for WQ assessment, particularly for the analysis of human drinking water [26]. The results indicated that the method solves the uncertainty problem of the change of the WQ. A WQ index using a fuzzy inference system has been successfully proposed, since

it has proven to be advantageous in addressing uncertainty problems [27].

A novel hybrid intelligent technique—fuzzy wavelet neural network (FWNN)—is proposed by combining NN, FL, and WT. The FWNN makes effective use of the self-adaptability of NN, the uncertainty capacity of FL, and the partial analysis ability of WT, so that the networks have better learning and higher speed and precision [28]. The FWNN combines the time-frequency localization ability of the wavelet, the fuzzy inference of the FL, and the education character of the ANN together, to improve its ability to reach the global best results. The FWNN offers a much more elegant and powerful way for the modeling process and simulation of the water management process than the traditional modeling approaches.

In this study, a hybrid intelligent software sensor model based on the FWNN is developed for the real-time estimation of the WQ of the Pearl River, China. A self-adapted fuzzy *c*-means clustering was used to withdraw the system's character and to optimize the network space. A hybrid learning algorithm was presented to further improve the neural network prediction capabilities. The FWNN model was developed to predict and estimate the WQ based on the available historical data, such as COD,  $\text{NH}_4^+$ -N, dissolved oxygen (DO), electrical conductivity (EC), pH, water temperature (WT), and turbidity (TU), which could increase the safety and improve the operational performance of WQ management.

## 2. Materials and Methods

**2.1. Study Area.** The Pearl River rises in the Yunnan Plateau, flows through the Guizhou, Guangxi, Guangdong, Hunan, and Jiangxi Provinces of China and through the northern part of Vietnam, and finally pours into the South China Sea. The Pearl River, including the Xijiang, Beijiang, and Dongjiang Rivers, is the third longest river in China, with an overall length of 2320 km and a river basin area of  $0.45 \times 10^6 \text{ km}^2$ . As the largest tributary of the Pearl River, the Xijiang River is “gold belt” with condensed resources in South China, which connects the coastal developed areas to the Southwest of China. The accuracy and the efficiency of the developed prediction model were tested using observed water quality data at the Hengsha station of the Xijiang River in Zhongshan City (China) (see Figure 1).

The Hengsha station with complete daily sampling was selected. Six important WQ variables, COD,  $\text{NH}_4^+$ -N, DO, EC, WT, pH, and TU, were analyzed. COD is one of the most important parameters for the degree of water pollution. It is commonly referred to as an important indicator of organic substances in water and is one of the most important parameters in water monitoring. The TU is considered an integrative parameter and is an important index in water. TU is closely related to other parameters associated with water quality, such as the COD, or the concentrations of different substances related to pollution (ammonium, sulphate, and nitrate). Due to the large effects of the other parameters, TU has become indicative of the comprehensive effect of other parameters. TU is an integrative parameter. The COD



FIGURE 1: Map of the Guangzhou section of the Pearl River under consideration.

and TU were selected as the prediction parameters. There were 340 samples collected from 2013 to 2014, of which 88 percent (300 sets of data) were used to train the model for predicting COD and TU level time series and 12 percent (40 sets of data) were used for model testing. DO, pH, EC, WT, and  $\text{NH}_4^+\text{-N}$  were used as the input values.

## 2.2. Genetic Algorithm Evolving Fuzzy Wavelet Neural Network (FWNN)

2.2.1. *Structure of the Proposed FWNN.* Figure 2 shows the basic structure of the FWNN. It was divided into five layers as follows [29]:

- (i) Layer 1 was the input layer, which was used to accept the present signal of the system. It consisted of the influencing factors  $(x_1, x_2, \dots, x_n)$  where the outputs were used as the input into the second layer. In this work, the input parameters were DO, pH, EC, WT, and  $\text{NH}_4^+\text{-N}$ ; then,  $n = 5$ .
- (ii) Layer 2 was the fuzzy layer, where the input characteristic variables were translated into fuzzy variables. In this work, the distribution of the WQ time series was close to a Gaussian function, so the membership function was composed of the Gaussian function [29], which made the network weight values have definite knowledge meaning. The outputs of the layer are shown below:

$$F_j(x_i) = \exp \left\{ -\frac{(x_i - c_{ij})^2}{2\sigma_{ij}^2} \right\}, \quad (1)$$

$$j = 1, 2, \dots, n, \quad i = 1, 2, \dots, m,$$

where  $i$  and  $j$  are the numbers of input variables and fuzzy linguistic terms, respectively, and  $c_{ij}$  and  $\sigma_{ij}$  are the center and the width parameters of the Gaussian function, respectively. A self-adapted fuzzy  $c$ -means clustering was employed to determine the number of clusters for the sample data and then to establish the number of rules (48 fuzzy rules). So, the number of nodes was  $48 \times 5$  in this layer.

- (iii) Layer 3 was termed the fuzzy rule layer, which was used to realize the logical inference based on the fuzzy rule. Each node at layer 3 corresponded to one fuzzy logic rule according to the multiplication. So, there were 48 nodes in layer 3. In this layer, the outputs were given as follows:

$$\mu_j(x) = \prod_{j=1}^n F_j(x_i), \quad i = 1, 2, \dots, m, \quad (2)$$

where  $n$  is the number of fuzzy rules.

- (iv) Layer 4 was termed the wavelet layer. In this layer, WNNs were utilized as the consequence of the network, where the wavelet function substituted for the effective function of the connotative layer and was used for data denoising transformation. The input in this layer was from the product of the output from layers 3 and 4. The output of the WNNs with the  $j$ th wavelet neuron was represented as

$$\begin{cases} wo_j = w_j \psi_j(z), \\ \psi_j(z) = \sum_{i=1}^m \frac{1}{\sqrt{|a_{ij}|}} (1 - z_{ij}^2) e^{z_{ij}^2/2}, \end{cases} \quad (3)$$

where  $z_{ij} = (x_i - b_{ij})/a_{ij}$ , in which  $a_{ij}$  and  $b_{ij}$  are the dilation and the translation parameters of the wavelet network, respectively, and  $w_j$  is the weight of the wavelet networks. The Mexican hat wavelet was the second derivative of the Gaussian function, which had excellent properties, particularly for the measurement of the instantaneous amplitude and frequency of oscillating signals. The Mexican hat wavelet was used as the mother wavelet, and the number of decomposition levels was the same as the number of fuzzy linguistic terms.

- (v) Layer 5 was the output layer. In this layer, the total output of the network  $y$  was as follows:

$$y_k = \frac{\sum_{j=1}^n \mu_j(x) wo_j}{\sum_{j=1}^n \mu_j(x)}. \quad (4)$$

In this work, the WQ indexes, COD, and TU were chosen as the outputs of the network.

The FWNN combines the time-frequency localization ability of the wavelet, the fuzzy inference, and the education character of the ANN together, where the ability to reach the global best results was improved [30].

2.2.2. *Training Algorithm to Optimize the Proposed FWNN.* A novel hybrid algorithm that combines the good property of

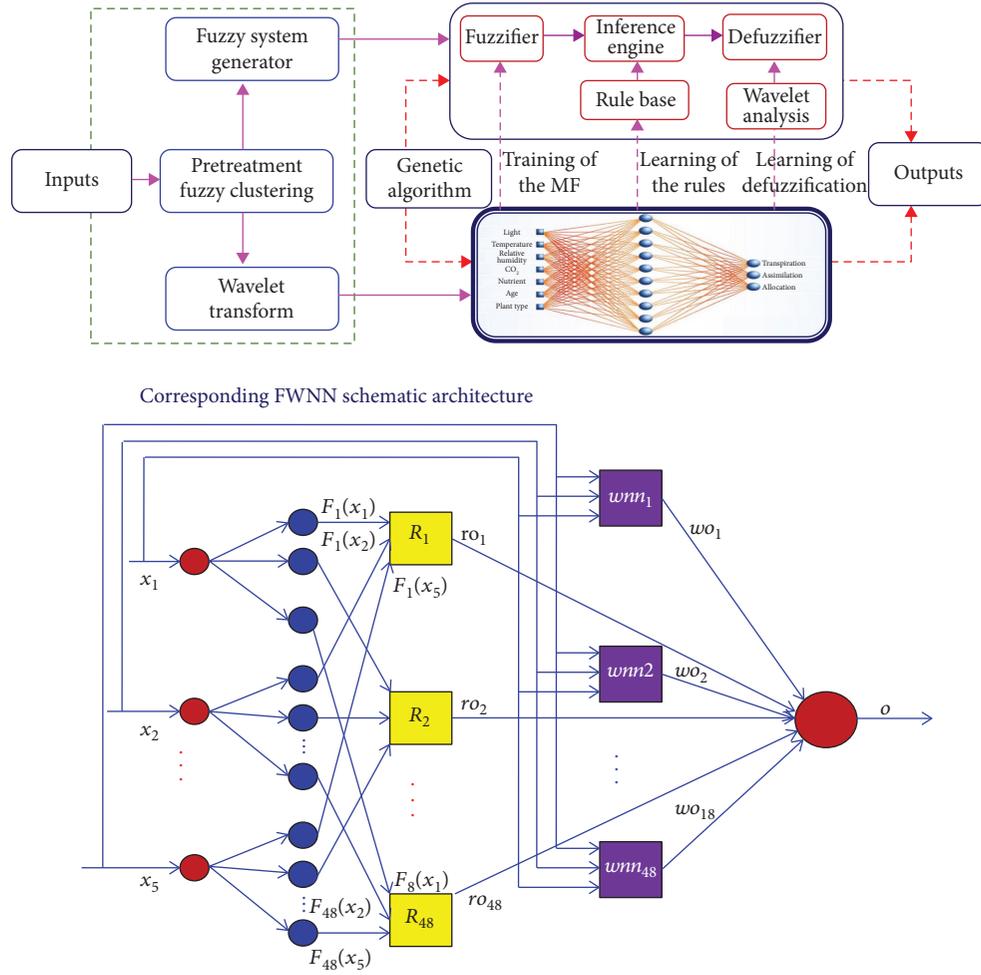


FIGURE 2: Schematic structure of FWNN model.

the global search of the genetic algorithm (GA) and the good property of the regional search of the gradient descent algorithm (GDA) was proposed to train the proposed FWNN and to optimize the network parameters. The center ( $c_{ij}$ ) and the width ( $\sigma_{ij}$ ) parameters of the Gaussian functions, dilation ( $a_{ij}$ ) and translation ( $b_{ij}$ ) parameters of the wavelet functions, and the weight ( $w_j$ ) of the wavelet networks were determined.

GA is a simple, but efficient, combinatorial optimization algorithm based on the principle of nature evolution [31]. GDA determined that it easily fell into the local optimum and was sensitive to the initial values. The GA initialized the network; then, the antecedent parameters and consequent parameters of the FWNN were simultaneously optimized with the GDA. In this work, an objective function was proposed to evaluate the minimum error ( $E$ ) between the desired values ( $y_d(L)$ ) and the output values of the FWNN ( $y_k$ ), which is given as follows:

$$E = \frac{1}{2} \sum_{k=1}^n (y_{dk} - y_k)^2. \quad (5)$$

The output of the FWNN according to the  $s$ -th chromosome with ( $y_k(L)$ ) was represented by the following equation:

$$\begin{aligned} y^s(L) &= \frac{\sum_{j=1}^n \mu_j^s(x_i(L)) \cdot wo_j^s}{\sum_{j=1}^n \mu_j^s(x_i(L))} \\ &= \frac{\sum_{j=1}^n wo_j^s \cdot \left[ \prod_i \exp \left\{ -\frac{(x_i(L) - c_{ij}^s)^2}{(\sigma_{ij}^s)^2} \right\} \right]}{\sum_{j=1}^n \left[ \prod_i \exp \left\{ -\frac{(x_i(L) - c_{ij}^s)^2}{(\sigma_{ij}^s)^2} \right\} \right]}. \end{aligned} \quad (6)$$

Here,

$$\begin{aligned} wo_j^s &= w_j^s \cdot \sum_{i=1}^m |a_{ij}^s|^{-1/2} \left( 1 - \left( \frac{x_i(L) - b_{ij}^s}{a_{ij}^s} \right)^2 \right) \\ &\cdot \exp \left( -0.5 \left( \frac{x_i(L) - b_{ij}^s}{a_{ij}^s} \right)^2 \right), \end{aligned} \quad (7)$$

where  $i$  and  $j$  are the numbers of input variables and wavelet neurons, respectively. The operation of the chromosome was the following real coding set:

$$Xn^s = [c_{ij}^s \ \sigma_{ij}^s \ a_{ij}^s \ b_{ij}^s \ w_j^s], \quad (8)$$

where

$$\begin{aligned} c_{ij}^s &= [c_{11}^s \ \dots \ c_{m1}^s \ \dots \ c_{1n}^s \ \dots \ c_{mn}^s], \\ \sigma_{ij}^s &= [\sigma_{11}^s \ \dots \ \sigma_{m1}^s \ \dots \ \sigma_{1n}^s \ \dots \ \sigma_{mn}^s], \\ a_{ij}^s &= [a_{11}^s \ \dots \ a_{m1}^s \ \dots \ a_{1n}^s \ \dots \ a_{mn}^s], \\ b_{ij}^s &= [b_{11}^s \ \dots \ b_{m1}^s \ \dots \ b_{1n}^s \ \dots \ b_{mn}^s], \\ w_j^s &= [w_1^s \ \dots \ w_n^s], \end{aligned} \quad (9)$$

where  $i$  and  $j$  are the numbers of input variables and fuzzy linguistic terms, respectively.

The initial parameters of the FWNN were optimized through selection, crossover, and mutation. The initial population size  $N$  pop was 100, crossover rat Pc was 0.7, and the interval of mutation Pm was 0.01.

**2.2.3. Updating Parameters through the Gradient Descent Algorithm.** After the proposed FWNN was initialized by GA, the GDA was employed to update the parameters of the FWNN [29]. The center ( $c_{ij}$ ) and the width ( $\sigma_{ij}$ ) parameters of the Gaussian functions, the dilation ( $a_{ij}$ ) and translation ( $b_{ij}$ ) parameters of the wavelet functions, and the weight ( $w_j$ ) of the wavelet networks were adjusted according to the following function.

$$E = \frac{1}{2} [(y_d(t) - y(t))^2], \quad (10)$$

where  $y_d(t)$  and  $y(t)$  were the desired values and the output values of FWNN at time  $t$ , respectively. The parameter values of the FWNN were calculated by the following formulas:

$$w_j(t+1) = w_j(t) + \eta \frac{\partial E}{\partial w_j} + \xi (w_j(t) - w_j(t-1)), \quad (11)$$

$$a_{ij}(t+1) = a_{ij}(t) + \eta \frac{\partial E}{\partial a_{ij}} + \xi (a_{ij}(t) - a_{ij}(t-1)), \quad (12)$$

$$b_{ij}(t+1) = b_{ij}(t) + \eta \frac{\partial E}{\partial b_{ij}} + \xi (b_{ij}(t) - b_{ij}(t-1)), \quad (13)$$

$$\sigma_{ij}(t+1) = \sigma_{ij}(t) + \eta \frac{\partial E}{\partial \sigma_{ij}} + \xi (\sigma_{ij}(t) - \sigma_{ij}(t-1)), \quad (14)$$

$$c_{ij}(t+1) = c_{ij}(t) + \eta \frac{\partial E}{\partial c_{ij}} + \xi (c_{ij}(t) - c_{ij}(t-1)), \quad (15)$$

where  $\eta$  and  $\xi$  are the learning rate and the momentum factor of networks, respectively.

The values of derivatives in (11) and (15) were calculated by the following formulas:

$$\begin{aligned} \frac{\partial E}{\partial w_j} &= \frac{(y(t) - y^d(t)) \cdot \mu_j \cdot \psi(z_j)}{\sum_{j=1}^n \mu_j}, \\ \frac{\partial E}{\partial a_{ij}} &= \frac{\delta_j (3.5z_{ij}^2 - z_{ij}^4 - 0.5) e^{-z_{ij}^2/2}}{\sqrt{a_{ij}^3}}, \\ \frac{\partial E}{\partial b_{ij}} &= \frac{\delta_j (3z_{ij} - z_{ij}^3) e^{-z_{ij}^2/2}}{\sqrt{a_{ij}^3}}, \\ \frac{\partial E}{\partial c_{ij}} &= \sum_j \frac{\partial E}{\partial y} \frac{\partial y}{\partial \mu_j} \frac{\partial \mu_j}{\partial c_{ij}}, \\ \frac{\partial E}{\partial \sigma_{ij}} &= \sum_j \frac{\partial E}{\partial y} \frac{\partial y}{\partial \mu_j} \frac{\partial \mu_j}{\partial \sigma_{ij}}, \\ \frac{\partial E}{\partial y} &= y(t) - y^d(t), \\ \frac{\partial E}{\partial \mu_j} &= \frac{(w_j - y)}{\sum_{j=1}^n \mu_j}, \\ \frac{\partial E}{\partial c_{ij}} &= \sum_j \frac{\partial E}{\partial y} \frac{\partial y}{\partial \mu_j} \frac{\partial \mu_j}{\partial c_{ij}}, \\ \frac{\partial E}{\partial c_{ij}} &= \begin{cases} \mu_j(x_i) \frac{2(x_i - c_{ij})}{\sigma_{ij}^2}, & \text{if } i \text{ node is connected to rule node } j, \\ 0, & \text{otherwise,} \end{cases} \\ \frac{\partial E}{\partial \sigma_{ij}} &= \begin{cases} \mu_j(x_i) \frac{2(x_i - c_{ij})}{\sigma_{ij}^3} b, & \text{if } i \text{ node is connected to rule node } j, \\ 0, & \text{otherwise.} \end{cases} \end{aligned} \quad (16)$$

Here,

$$\delta_j = \frac{(y(t) - y^d(t)) \cdot \mu_j \cdot w_j}{\sum_{j=1}^n \mu_j}, \quad i = 1, \dots, m. \quad (17)$$

**2.3. Self-Adapted Fuzzy  $c$ -Means Clustering.** In order to optimize the FWNN's fuzzy rules automatically, a novel validity function,  $B(K)$ , was proposed to solve problems regarding the partial optimization and to determine the cluster number [32]. The numerator and the denominator of the validity function represented the sum of the distances between classes and the sum of the intradistances of all the clusters. Higher values for  $B(K)$  indicated better clustering results, and the best cluster number was obtained when  $B(K)$  reached its maximum value. The validity function was given as follows:

$$B(K) = \frac{\sum_{i=1}^K \sum_{j=1}^n u_{ij}^m \|v_i - \bar{x}\|^2 / (K-1)}{\sum_{i=1}^K \sum_{j=1}^n u_{ij}^m \|v_i - \bar{x}\|^2 / (n-K)}, \quad (18)$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^K \sum_{j=1}^n u_{ij}^m x_j,$$

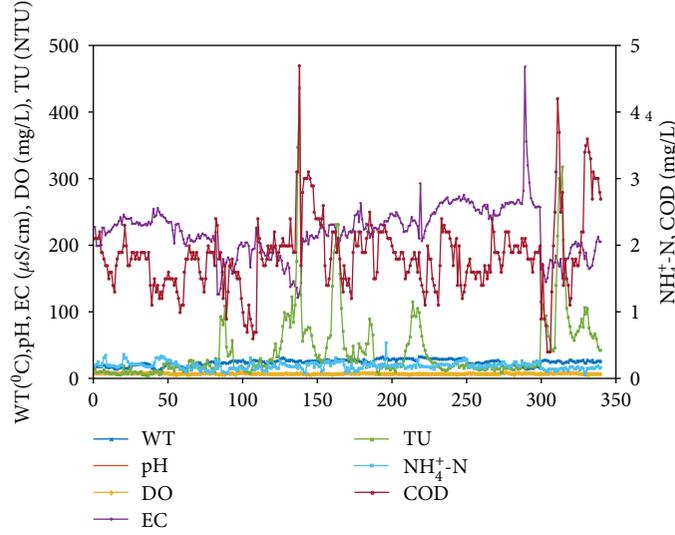


FIGURE 3: Variations of water quality parameters.

where  $K$  and  $n$  are the classification number and the object number in the calibration data.  $u_{ij}^m$  is the membership function value, and  $m$  is the weight coefficient that represented the fuzziness of the WQ classification. In which,  $\|\cdot\|$  represented the Euclidean distance measure.

The FCM clustering algorithm worked as an unsupervised classification method based on the minimization of a criterion function. The objective function was introduced based on the minimum square sum of the weighted Euclidean distances  $d_{ij}$ . The objective function was defined as follows:

$$J_m(U, V) = \sum_{i=1}^K \sum_{j=1}^n u_{ij}^m d_{ij}^2, \quad (19)$$

$$\sum_{i=1}^K u_{ij} = 1, \quad 1 \leq j \leq n, \quad (20)$$

where  $d_{ij}$  stands for the Euclidean distance between an object and a cluster, which was defined by

$$d_{ij} = \|x_j - v_i\|, \quad (21)$$

where  $x_j$  and  $v_i$  are the observed value and the cluster centroid, respectively.  $\|\cdot\|$  stands for the Euclidean distance measure. For the individual objects, (20) was used for the counting membership values. The membership function value  $u_{ij}$  to the  $i^{\text{th}}$  cluster at the time  $k$  was defined as follows:

$$u_{ij}^{(k)} = \frac{1}{\sum_{r=1}^c \left( d_{ij}^{(k)} / d_{rj}^{(k)} \right)^{2/(m-1)}}. \quad (22)$$

After computing the membership values for all the calibration objects, the cluster centers  $v_i$  were provided by the following equation:

$$v_i^{(k+1)} = \frac{\sum_{j=1}^n \left( u_{ij}^{(k)} \right)^m x_j}{\sum_{j=1}^n \left( u_{ij}^{(k)} \right)^m}. \quad (23)$$

The minimization of (19) originated after providing the initial values for the cluster centers. Equations (20)–(23) were repeated successively in each iteration step.

### 3. Results and Discussion

**3.1. Data Collection and Preprocessing.** The data between 2013 and 2014 was obtained from the China National Environmental Monitoring Centre and used to develop the software sensor, as seen in Figure 3. The 340 samples were collected from the Hengsha section of the Pearl River to form daily composite samples for analysis. Among the total numbers of data, the numbers for training and testing (predicting) were 300 and 40.

A database that contains system performance information is a prerequisite for the model development. Generally, it is developed by regularly collecting monitored parameters. The quality of the training database was critical for the model to produce correct information about the system. The database should contain adequate and correct information in the system for accurately describing the process. It remains common for a raw database to contain some redundant or conflicting data. Thus, it is necessary to pretreat the raw database by removing redundancies and resolving conflicts in the data.

**3.2. FWNN Development.** In this study, the performance of the WQ software sensor system for forecasting the COD and the TU was simulated using the FWNN model. DO, pH, EC, WT, and  $\text{NH}_4^+-\text{N}$  were the inputs of the models, and the output variables were COD and TU. Through analyzing the data set with the self-adapted fuzzy  $c$ -means clustering, 48 fuzzy rule clusters were obtained. The self-adapted fuzzy  $c$ -means cluster algorithm that preprocessed

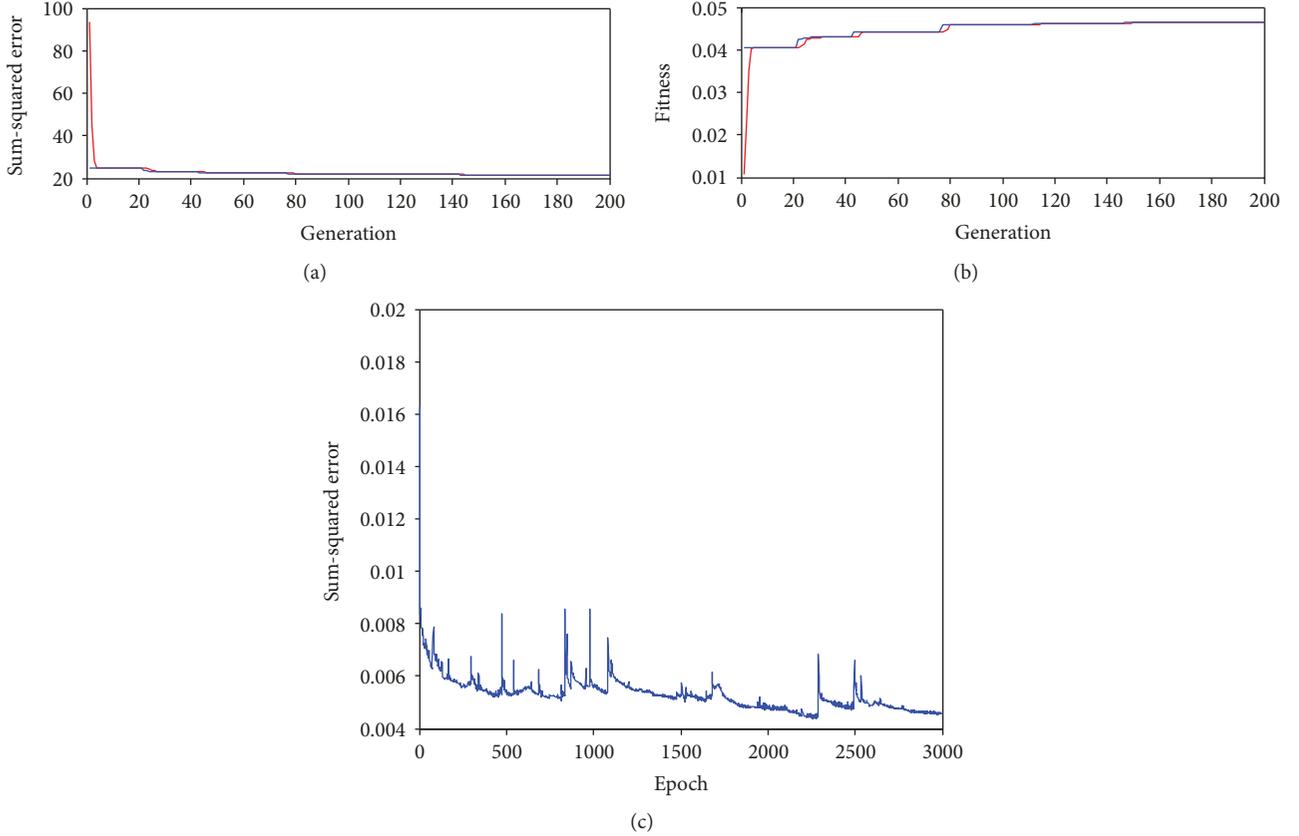


FIGURE 4: Training performance of FWNN based on hybrid GA-GDA algorithms.

the raw training data was required to prepare a concise database and to serve as an actual training database. The software sensor system contained two FWNN models, the *FWNNCOD* and the *FWNNTU*, for predicting the COD and the TU. Each model had its own rule base, but they share the same input. The model was used for predicting the COD and the TU. The structure of the FWNN software sensor system is shown in Figure 2.

After the initial structure and the parameters of the FWNN model were determined, a hybrid learning algorithm that integrated the GA and the GDA was applied to train and optimize the network parameters. After the structure and the parameters of the FWNN were optimized by the GA, the GDA was employed to update the parameters of the network.

**3.3. Simulation of the Hybrid FWNN Model.** Two forecasting models based on the FWNN were implemented onto MATLAB. The initial population size  $N_{pop}$  was 100, the crossover rate  $P_c$  was 0.7, the interval of mutation  $P_m$  was 0.01, the maximum generation number was 200, the learning rate  $\eta$  was 0.01, and the momentum factor  $\xi$  was 0.5. Figure 4 shows the training process of the FWNN (take *FWNNCOD* for example). Figure 3 shows that the hybrid algorithm had a rapid convergence ability and it rapidly met the target error. The centers and the widths of the membership functions of the fuzzification layer ( $c_{ij}$  and  $\sigma_{ij}$ , respectively), the dilation and translation of the wavelet functions of the wavelet layer ( $a_{ij}$  and  $b_{ij}$ , respectively), and the weight of the wavelet

networks ( $w_j$ ) were decided, as shown in Table S1–S4 (Supplementary Information).

As shown in Figures 5 and 6, the forecasting results of the software sensor model based on the FWNN for testing datasets were demonstrated. The observed values were consistent with predictions made by the forecasting models. In order to evaluate the forecast accuracy of the proposed FWNN model, various performance indexes were used to assess the stability of the predicted values, which included the coefficient of determination ( $R^2$ ), correlation coefficient ( $R$ ), root mean square errors (RMSE), mean square error (MSE), and mean absolute percentage error (MAPE). As shown in Table 1, the performance indexes of the proposed FWNN models were acquired for testing datasets.

Table 1 shows that  $R$  values of 0.9499 and 0.9678 for the COD and the TU, respectively, were achieved by using the FWNN. The  $R^2$  values for the COD and the TU were 0.9023 and 0.9366, respectively. MAPEs for the COD and the TU were 19.2030 and 26.3203, respectively. The RMSE values were 0.3127 and 4.5452 for the COD and the TU, respectively, via the FWNN. These results showed that the proposed FWNN model achieved a satisfactory performance for WQ prediction.

As seen in Table 1, the proposed FWNN model demonstrated a very satisfactory performance on the prediction of the WQ with high determination coefficient values ( $R^2$ ) of about 0.90 and 0.94 for COD and TU, respectively. The high value of the determination coefficient indicated that only

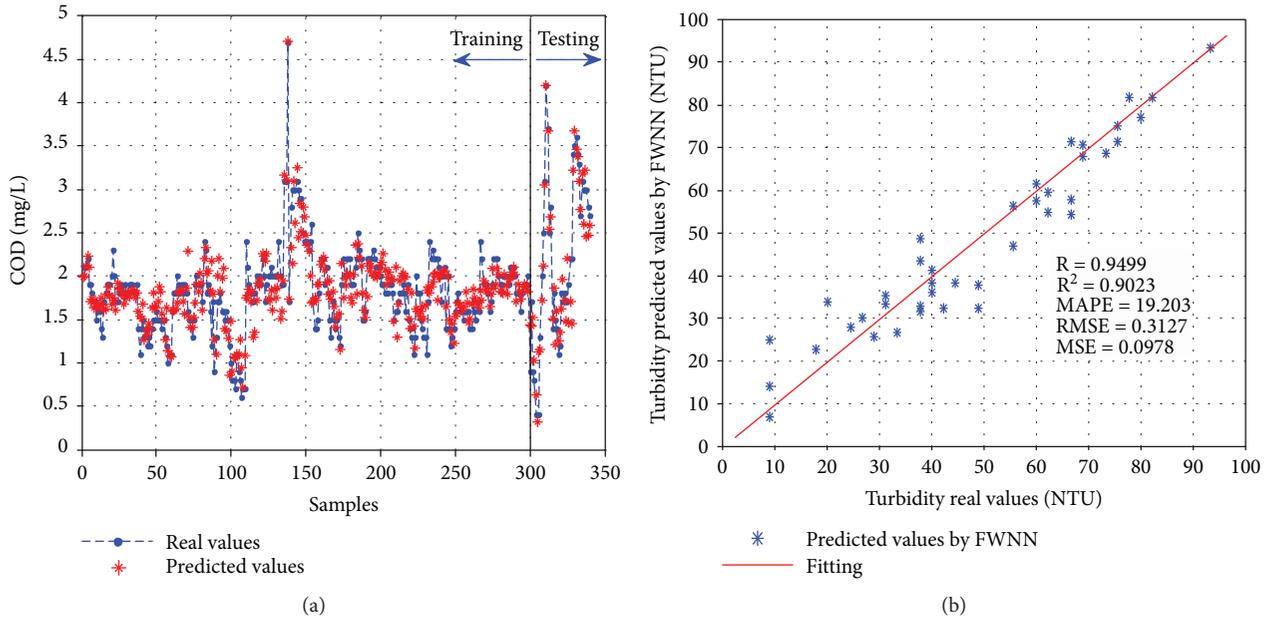


FIGURE 5: FWNN performance of the estimation of water quality for COD.

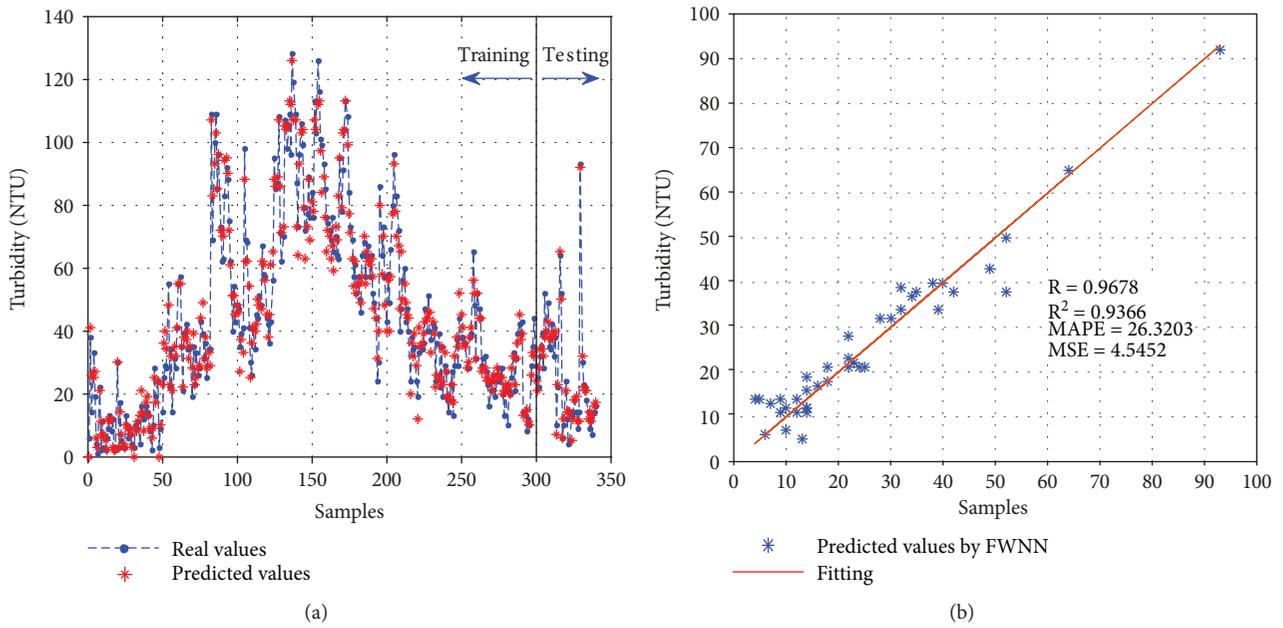


FIGURE 6: FWNN performance of the estimation of water quality for turbidity.

TABLE 1: Performances of FWNN, WNN, and NN in modeling water quality management.

	Model (COD)				Model (TU)			
	FWNN	FNN	WNN	NN	FWNN	FNN	WNN	NN
R	0.9499	0.8129	0.7733	0.5639	0.9678	0.8502	0.8332	0.6690
R <sup>2</sup>	0.9023	0.6608	0.598	0.318	0.9366	0.7228	0.6942	0.4476
MAPE	19.2030	26.2873	46.9316	34.0226	26.3203	29.6884	41.3267	66.0219
RMSE	0.3127	0.6035	0.6404	0.9198	4.5452	10.8572	11.9338	14.1188
MSE	0.0978	0.3643	0.4102	0.8461	20.6585	117.878	142.4146	199.3415

10% and 6% of the total variations for COD and TU, respectively, were not explained by the proposed FWNN model. According to the high  $R$  for the COD and the TU, the predicted values of the new model were closer to the measured ones. According to the values of other descriptive performance indexes (MAPE, RMSE, and MSE), the developed FWNN model showed a superior prediction performance. There was a small deviation produced by the developed FWNN model. Using the WQ prediction model, the unfavorable influence of the river water function by a user's water drainage was estimated. This was attributed to the severe changes of concentration of the COD and the TU in the Zhongshan region.

**3.4. Comparisons with FNN, WNN, and NN.** The developed FWNN model were compared with the FNN, the WNN, and the NN models to demonstrate the correctness, the efficiency, and the advantages of the hybrid network. The FWNN model had smaller RMSE (or MSE) and MAPE and higher  $R^2$  and  $R$  values, as seen in Table 1. When the FWNN model was used for predicting the COD in rivers (example being the COD),  $R$ ,  $R^2$ , MAPE, RMSE, and MSE values were 0.9499, 0.9023, 19.203, 0.3127, and 0.0978, respectively. When the FNN, the WNN, and the NN models were used to predict the COD in rivers, the  $R$  values were 0.8129, 0.7733, and 0.5639, respectively. The  $R^2$  values were 0.6608, 0.598, and 0.3180, respectively. The MAPE values were 26.2873, 46.9316, and 34.0226, respectively. The RMSE values were 0.6025, 0.6404, and 0.9198, respectively. The MSE values were 0.3643, 0.4102, and 0.8461, respectively.

The NN model was unable to extract the nonstationarity from the data set. The NN model demonstrated their weakness to extract the nonlinearity when the length of the training data set was the same. This showed that the NN model had a limited capability to extract the nonstationarity from the data sets. The performances of the WNN and the FNN models were better than that of the NN model. This was because the wavelet transformation decomposed components of time series data extracted different time-varying components, which could be representative of the sum of the subprocesses associated with the original time series data set. These different components facilitated the ability of the NN model that used the WT and the fuzzy inferences to extract nonlinearity and nonstationarity, making its performance superior to the NN model developed using raw data sets. The performance of the FWNN model, which included the NN, the FL, the WT, and the GA, was found to be the best compared to those of the remaining models.

The FWNN model achieved better performances than the FNN, the WNN, and the NN models, which illustrated that the FWNN model for predicting the WQ was more accurate than the FNN, the WNN, and the NN models. The results clearly indicated that the FWNN model had a high ability to extract the dynamic behavior and the complex interrelationships from various WQ variables. The results of this study suggested that the FWNN model had a high ability to extract the dynamic changes of the water resource management system.

Considering the high level of complexity in WQ management, there was a large quantity of variable information spread in the dataset and the wide concentration ranges, such as good prediction performance of the FWNN model for the parameters. The FWNN was a good choice for modeling the WQ management. The simulated models, based on the FWNN model, can be effectively applied to WQ management in order to cope with water quality variations. The results indicated that after hybrid learning, the proposed FWNN could perform better in the WQ management process than the FNN, the WNN, and the NN. With the environmental standards maintained, the FWNN model could effectively achieve both environmental and economic objectives of the WQ management in real time.

## 4. Conclusion

The WQ prediction is an important means for understanding water pollution tendencies. This study presented a WQ model based on FWNN for estimating for the WQ data sets. The performance of the FWNN model was compared to the performances of the three different models: the traditional ANN, the WNN, and the FNN models. The proposed WQ prediction model, based on the FWNN, produced better performance than the other three models. The descriptive performance indexes proposed by the FWNN model indicated that the FWNN could handle the severely fluctuating time series data of the better WQ accuracy. The proposed hybrid approach provided an effective and useful tool for modeling the WQ time series, which enables engineers to monitor various WQ parameters for improving WQ management.

## Data Availability

The data used to support the findings of this study have not been made available because the authors are asked to sign a confidentiality agreement with the Chinese government which provides the basis of the data.

## Conflicts of Interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This research was supported by the National Natural Science Foundation of China (nos. 51208206 and 51210013), Guangdong Province Natural Science Foundation (no. 2016A030306033), Guangdong Provincial Science and Technology Plan Project Foundation (no. 2014A020216007), Technological Innovation Young Talents of Guangdong Special Support Plan (no. 2014TQ01Z530), Pearl River Nova Program of Guangzhou (no. 201506010058), and Fundamental Research Funds for the Central Universities (no. 310829161005). The authors are thankful to the anonymous reviewers for their insightful comments and suggestions.

## Supplementary Materials

(1) The parameters of FWNNTU were determined: the center and the width parameters of the membership functions ( $c_{ij}$  and  $\sigma_{ij}$ , respectively), the dilation and the translation parameters of the wavelet functions ( $a_{ij}$  and  $b_{ij}$ , respectively), and the weight of the wavelet networks ( $w_j$ ) were drawn, as shown in Table S1–S4. (2) The main code of the hybrid algorithm was added. (*Supplementary Materials*)

## References

- [1] J. Zheng, J. Ruan, L. Dong, T. Zhang, M. Huang, and Z. Xu, "Hollow aluminum particle in eddy current separation of recovering waste toner cartridges," *ACS Sustainable Chemistry and Engineering*, vol. 5, no. 1, pp. 161–167, 2017.
- [2] M. Huang, Y. Ma, J. Wan, and X. Chen, "A sensor-software based on a genetic algorithm-based neural fuzzy system for modeling and simulating a wastewater treatment process," *Applied Soft Computing*, vol. 27, pp. 1–10, 2015.
- [3] N. M. Gazzaz, M. K. Yusoff, A. Z. Aris, H. Juahir, and M. F. Ramli, "Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors," *Marine Pollution Bulletin*, vol. 64, no. 11, pp. 2409–2420, 2012.
- [4] M. Huang, J. Q. Wan, Y. W. Ma, W. J. Li, X. F. Sun, and Y. Wan, "A fast predicting neural fuzzy model for on-line estimation of nutrient dynamics in an anoxic/oxic process," *Bioresource Technology*, vol. 101, no. 6, pp. 1642–1651, 2010.
- [5] H. Boyacioglu, "Development of a water quality index based on a European classification scheme," *Water SA*, vol. 33, no. 1, pp. 101–106, 2009.
- [6] C. L. Wu and K. W. Chau, "Mathematical model of water quality rehabilitation with rainwater utilisation: a case study at Haigang," *International Journal of Environment and Pollution*, vol. 28, no. 3/4, pp. 534–545, 2006.
- [7] Y. Jiang and Z. Nan, "Assessment of river water quality using uncertainly mathematical model: a case study of Yellow River, China," in *2012 International Symposium on Geomatics for Integrated Water Resource Management*, pp. 1–4, Lanzhou, China, October 2012.
- [8] B. Khalil, T. B. M. J. Ouarda, and A. St-Hilaire, "Estimation of water quality characteristics at ungauged sites using artificial neural networks and canonical correlation analysis," *Journal of Hydrology*, vol. 405, no. 3-4, pp. 277–287, 2011.
- [9] C. L. Wu, K. W. Chau, and Y. S. Li, "Predicting monthly streamflow using data-driven models coupled with data-preprocessing techniques," *Water Resources Research*, vol. 45, no. 8, pp. 2263–2289, 2009.
- [10] H. R. Maier, A. Jain, G. C. Dandy, and K. P. Sudheer, "Methods used for the development of neural networks for the prediction of water resource variables in river systems: current status and future directions," *Environmental Modelling and Software*, vol. 25, no. 8, pp. 891–909, 2010.
- [11] Y. Wang, T. Zheng, Y. Zhao et al., "Monthly water quality forecasting and uncertainty assessment via bootstrapped wavelet neural networks under missing data for Harbin, China," *Environmental Science and Pollution Research*, vol. 20, no. 12, pp. 8909–8923, 2013.
- [12] S. E. Kim and I. W. Seo, "Artificial neural network ensemble modeling with conjunctive data clustering for water quality prediction in rivers," *Journal of Hydro-Environment Research*, vol. 9, no. 3, pp. 325–339, 2015.
- [13] M. J. Alizadeh and M. R. Kavianpour, "Development of wavelet-ANN models to predict water quality parameters in Hilo Bay, Pacific Ocean," *Marine Pollution Bulletin*, vol. 98, no. 1-2, pp. 171–178, 2015.
- [14] N. Nasr, H. Hafez, M. H. el Naggat, and G. Nakhla, "Application of artificial neural networks for modeling of biohydrogen production," *International Journal of Hydrogen Energy*, vol. 38, no. 8, pp. 3189–3195, 2013.
- [15] D. Ö. Faruk, "A hybrid neural network and ARIMA model for water quality time series prediction," *Engineering Applications of Artificial Intelligence*, vol. 23, no. 4, pp. 586–594, 2010.
- [16] A. A. M. Ahmed, "Prediction of dissolved oxygen in Surma River by biochemical oxygen demand and chemical oxygen demand using the artificial neural networks (ANNs)," *Journal of King Saud University-Engineering Sciences*, vol. 29, no. 2, pp. 151–158, 2017.
- [17] M. R. Mustafa, R. B. Rezaur, S. Saiedi, and M. H. Isa, "River suspended sediment prediction using various multilayer perceptron neural network training algorithms—a case study in Malaysia," *Water Resources Management*, vol. 26, no. 7, pp. 1879–1897, 2012.
- [18] Z. Xiao, L. Peng, Y. Chen, H. Liu, J. Wang, and Y. Nie, "The dissolved oxygen prediction method based on neural network," *Complexity*, vol. 2017, Article ID 4967870, 6 pages, 2017.
- [19] J. Ruan, C. Zhang, Y. Li et al., "Improving the efficiency of dissolved oxygen control using an on-line control system based on a genetic algorithm evolving FWNNTU software sensor," *Journal of Environmental Management*, vol. 187, pp. 550–559, 2017.
- [20] M. C. Valverde, E. Araujo, and H. Campos Velho, "Neural network and fuzzy logic statistical downscaling of atmospheric circulation-type specific weather pattern for rainfall forecasting," *Applied Soft Computing*, vol. 22, pp. 681–694, 2014.
- [21] S. A. M. Dehghan, M. Danesh, F. Sheikholeslam, and M. Zekri, "Adaptive force-environment estimator for manipulators based on adaptive wavelet neural network," *Applied Soft Computing*, vol. 28, pp. 527–540, 2015.
- [22] M. Mohammadi, A. Danandeh, H. Nasir Aghdam, and N. Ojaroudi, "Wavelet neural network based on islanding detection via inverter-based DG," *Complexity*, vol. 21, no. 2, 324 pages, 2015.
- [23] F. Evrendilek and N. Karakaya, "Monitoring diel dissolved oxygen dynamics through integrating wavelet denoising and temporal neural networks," *Environmental Monitoring and Assessment*, vol. 186, no. 3, pp. 1583–1591, 2014.
- [24] S. Kumar, M. K. Tiwari, C. Chatterjee, and A. Mishra, "Reservoir inflow forecasting using ensemble models based on neural networks, wavelet analysis and bootstrap method," *Water Resources Management*, vol. 29, no. 13, pp. 4863–4883, 2015.
- [25] M. Ravansalar, T. Rajaei, and M. Zounemat-Kermani, "A wavelet-linear genetic programming model for sodium ( $\text{Na}^+$ ) concentration forecasting in rivers," *Journal of Hydrology*, vol. 537, pp. 398–407, 2016.
- [26] H. Gharibi, A. H. Mahvi, R. Nabizadeh, H. Arabalibeik, M. Yunesian, and M. H. Sowlat, "A novel approach in water quality assessment based on fuzzy logic," *Journal of Environmental Management*, vol. 112, pp. 87–95, 2012.

- [27] S. R. M. M. Roveda, A. P. M. Bondanca, J. G. S. Silva, J. A. F. Roveda, and A. H. Rosa, "Development of a water quality index using a fuzzy logic: a case study for the Sorocaba River," in *International Conference on Fuzzy Systems*, pp. 1–5, Barcelona, Spain, July 2010.
- [28] H. Loussifi, K. Nouri, and N. Benhadj Braiek, "A new efficient hybrid intelligent method for nonlinear dynamical systems identification: the wavelet kernel fuzzy neural network," *Communications in Nonlinear Science and Numerical Simulation*, vol. 32, pp. 10–30, 2016.
- [29] Z. Liu, Q. Wu, Y. Zhang, Y. Wang, and C. L. P. Chen, "Adaptive fuzzy wavelet neural network filter for hand tremor canceling in microsurgery," *Applied Soft Computing*, vol. 11, no. 8, pp. 5315–5329, 2011.
- [30] M. Huang, T. Zhang, J. Ruan, and X. Chen, "A new efficient hybrid intelligent model for biodegradation process of DMP with fuzzy wavelet neural networks," *Scientific Reports*, vol. 7, no. 1, article 41239, 2017.
- [31] M. Huang, J. Wan, Y. Ma et al., "A GA-based neural fuzzy system for modeling a paper mill wastewater treatment process," *Industrial and Engineering Chemistry Research*, vol. 50, no. 23, pp. 13500–13507, 2011.
- [32] M. Huang, J. Wan, K. Hu, Y. Ma, and Y. Wang, "Enhancing dissolved oxygen control using an on-line hybrid fuzzy-neural soft-sensing model-based control system in an anaerobic/anoxic/oxic process," *Journal of Industrial Microbiology and Biotechnology*, vol. 40, no. 12, pp. 1393–1401, 2013.

