Mathematical and Computer Modelling 58 (2013) 807–813

Contents lists available at SciVerse ScienceDirect

Mathematical and Computer Modelling

journal homepage: www.elsevier.com/locate/mcm

Study of short-term water quality prediction model based on wavelet neural network

Longqin Xu [a,*], Shuangyin Liu [a,b,c]

[a] College of Information, Guangdong Ocean University, Zhanjiang, Guangdong 524025, China
[b] China-EU Center for ICT in agriculture, China Agricultural University, Beijing 100083, China
[c] Beijing Engineering Research Center for Agricultural Internet of Things, China Agricultural University, Beijing 100083, China

ARTICLE INFO

Article history:
Received 9 October 2012
Accepted 10 December 2012

Keywords:
Water quality prediction
Wavelet neural network
Wavelet analysis
BP neural network
Intensive pearl breeding

ABSTRACT

Improved water quality prediction accuracy and reduced computational complexity are vital for ensuring a precise control over the water quality in intensive pearl breeding. This paper combined the wavelet transform with the BP neural network to build the short-term wavelet neural network water quality prediction model. The proposed model was used to predict the water quality of intensive freshwater pearl breeding ponds in Duchang county, Jiangxi province, China. Compared with prediction results achieved by the BP neural network and the Elman neural network, the mean absolute percentage error dropped from 17.464% and 8.438%, respectively, to 3.822%. The results show that the wavelet neural network is superior to the BP neural network and the Elman neural network. Furthermore, the proposed model features a high learning speed, improved predict accuracy, and strong robustness. The model can predict water quality effectively and can meet the management requirements in intensive freshwater pearl breeding.

1. Introduction

Regulation of water quality in aquaculture is one of the most important tasks in intensive aquaculture management. Accurate prediction of water quality is the basis for decisions when controlling water quality, setting the aquaculture water plan, and dealing with water quality incidents. Therefore, the research on the various methods of prediction of intensive aquaculture water quality has important theoretical value and practical significance.

There already exist numerous water quality prediction methods. Water quality simulation method, regression analysis [1], gray theory [2], time series [3], fuzzy reasoning [4] and neural network method [5] are commonly used with positive results in sewage treatment, drinking water management, and various other fields. Wang Xupeng constructed the prediction of dissolved oxygen with seasonal ARIMA model in the Yuqiao Reservoir [6]. Guo Lianxi built the prediction model of dissolved oxygen in ponds using fuzzy neural network trained by fast particle swarm optimization algorithm [7]. Feng Minquan preprocessed water quality data with Rajda standards, automatically optimized BP network input nodes by the gray correlation degree, and put forward the new model combining the BP network with the Markov algorithm, effectively improving the prediction accuracy of water quality parameters [8]. Wang Ruimei established a dissolved oxygen prediction model in freshwater pond breeding based on fuzzy neural network systems [6]. Chen predicted pond water temperature in holothurian aquaculture using the RBF neural network [9]. Han put forward a water quality prediction model based on the adaptive RBF neural network model, which achieved notable results [10]. Mahapatra utilized the water quality prediction

* Corresponding author. Tel.: +86 759 2339 456.
E-mail address: xlqlw@126.com (L. Xu).

0895-7177/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.mcm.2012.12.023

of an Indian river with a series fuzzy inference system; compared to other methods its prediction accuracy was higher [11]. Faruk proposed the time series prediction model of water quality fused with ARIMA and neural network, and obtained better prediction results [3].

Intensive aquaculture water is easily affected both by ecological factors and by human activities which creates a complex, multivariable, nonlinear system having large delay and fuzzy uncertainty. At the same time, the traditional neural network methods make scientifically determining the network structure difficult, and the training parameters easily fall into the local minimum, which seriously affects the precision and reliability of the model. Therefore, when these methods are directly applied to intensive aquaculture water quality predicting, the results are often unsatisfactory.

Wavelet neural network is a new mathematical modeling method combining wavelet transform with an artificial neural network. This new method combines the advantage of the localization character of wavelet transform and the self-learning ability of neural networks. The learning algorithm of the weights is easier than in conventional neural networks. The error function is a convex function and has nonlinear approximation, strong fault tolerance, fast convergence speed, and good prediction results [12,13]. This method has been effectively applied in signal processing, predictive control, fault diagnosis, and pattern recognition.

Based on these previous applications, this paper proposes a short-term aquaculture water quality prediction model based on the wavelet neural network to improve the prediction accuracy. The model was applied to predict the water quality of intensive freshwater pearl breeding ponds in Duchang County in China's Jiangxi Province. The results show that the wavelet neural network is faster, more precise, and more robust than the BP neural network or the Elman neural network, and can effectively predict water quality in intensive freshwater pearl cultures.

2. Wavelet neural network prediction model

2.1. Wavelet analysis

Wavelet analysis is a new branch of mathematics that has different resolutions in different positions of the time–frequency plane. It is a multi-resolution signal analysis method, which provides a powerful tool for non-stationary signal analysis and processing. Its essence is to express a function through the telescopic or translation of basic wavelet function $\psi(x)$. Related concepts are outlined as follows [14]:

Definition 1. Set $\psi(t)$ belongs to function space $L^2\,(R)$ if

$$\int_{-\infty}^{+\infty} \psi(t)dt = 0 \tag{1}$$

so that $\psi(T)$ is regarded as a mother wavelet and a wavelet.

Definition 2. The translation and expansion of the mother wavelet $\psi(T)$ can produce the wavelet basis function. The formula is

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right) \tag{2}$$

where $a > 0$ is the scale parameter and $b \in R$ is the translation parameter.

2.2. Wavelet neural network structural design

Wavelet neural network is based on the topology and structure of the BP neural network. The wavelet basis function is the incentive function of the neurons. By using the advantage of the wavelet signal analysis and combining the function of the neural network training and prediction, the signal is transmitted forward and the error is propagated backward, so as to achieve a more accurate predictive value signal. In view of the multiple input and single output in aquaculture water quality prediction, this study adopted a three-layer network structure: the input layer's m nodes, the hidden layer's n nodes, and the output layer's 1 node. The network structure is shown in Fig. 1.

The wavelet neural network model above can be expressed by the following mathematical formula [14]:

$$\hat{y} = \sum_{j=1}^{n} w_i \psi_j \left(\sum_{k=1}^{m} \frac{x_k \times U_{kj} - b_j}{a_j} \right). \tag{3}$$

In the formula, $x = (x_1, x_2, \ldots, x_m)^T$ is the input vector; $\hat{y}$ is the predicted output value; U_{kj} is the connection weight from the input layer kth node to the hidden layer jth node; ψ_j is the activation function of the hidden layer jth neuron; w_j is the layer weight from the hidden layer jth nodes to the output; a_j is the expansion parameter of wavelet function; and b_j is the translation parameter of wavelet function.

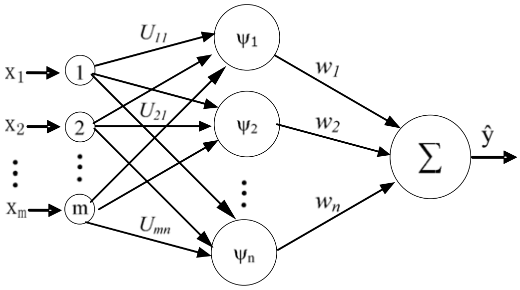

Fig. 1. Wavelet neural network structural diagram.

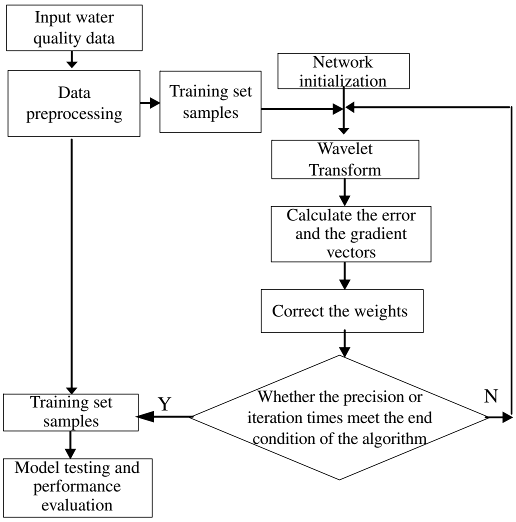

Fig. 2. Prediction flow chart of wavelet neural network.

Whether the choice of wavelet base function is appropriate will directly affect the prediction accuracy and generalization ability, but there is no mature theory to guide the plan. Of the broader Mexico hat wavelet, Morlet wavelet, orthogonal wavelet, and Gauss spline wavelet, Morlet wavelet has the smallest error and the best computational stability (Abiyev, 2012); therefore, this study used Morlet wavelet as the wavelet function model. The formula is given below:

$$\psi = \cos(1.75x) \times \exp\left(-\frac{x^2}{2}\right). \quad (4)$$

2.3. Wavelet neural network training algorithm

The parameters for training of the wavelet neural network model are connection weights U_{kj} and w_j. The optimized process of the expansion parameters a_j and translation parameters b_j is divided into two stages [15]: (1) forward propagation process and (2) back propagation process. In the forward propagation process, the output layer is calculated step by step from the input layer according to input samples. In back propagation, correct weights are calculated backwards from the output layer of the network. The two processes alternate repeatedly until meeting the prediction accuracy of the network. The training and correcting algorithm steps of the network parameters are as follows and as shown in Fig. 2.

Step 1: *Network initialization.* The connection weights U_{ij} and w_j of the network are initialized randomly, setting up the expansion parameters a_j and translation parameters b_j of the wavelet function as well as the network learning rate η, the error threshold ε, and maximum iterations T.

Step 2: *Data preprocessing.* Data preprocessing and the samples are divided into training sets and test sets. The training samples are used for training the network; the test sets are used for testing algorithm performance.

Step 3: *Calculating error and gradient vectors.* Firstly the training samples are input into the network; next the predictive output is calculated. Then the error and the gradient vectors of the output of the network and the expected output are calculated and its target error functions defined as follows:

$$E = \frac{1}{2}\sum_{k}^{m} \hat{y}(k) - y(k)^2. \tag{5}$$

The gradient vectors are calculated by the predicted and targeted error $\Delta U_{k,j}\,(i+1)$, $\Delta w_j\,(i+1)$, $\Delta a_j\,(i+1)$, and $\Delta b_j\,(i+1)$. Its expressions are as follows:

$$\Delta U_{k,j}(i+1) = -\eta\left(\frac{\partial E}{\partial \Delta U_{k,j}(i)}\right) \tag{6}$$

$$\Delta w_j(i+1) = -\eta\left(\frac{\partial E}{\partial \Delta w_j(i)}\right) \tag{7}$$

$$\Delta a_j(i+1) = -\eta\left(\frac{\partial E}{\partial \Delta a_j(i)}\right) \tag{8}$$

$$\Delta b_j(i+1) = -\eta\left(\frac{\partial E}{\partial \Delta b_j(i)}\right). \tag{9}$$

In the formulas, $\hat{y}(k)$ and $y(k)$ are the predicted output value and the desired output values respectively of the wavelet neural network, while η is the learning rate of the momentum.

Step 4: *Weight modification.* The back propagation of error corrects the weights of the wavelet neural network and the parameters of the wavelet function [15]. The formulas are as follows:

$$U_{k,j}(i+1) = U_{k,j}^{(i)} + \Delta U_{k,j}(i+1) + \eta(U_{k,j}(i) - U_{k,j}(i-1)) \tag{10}$$

$$w_j(i+1) = w_j^{(i)} + \Delta w_j(i+1) + \eta(w_j(i) - w_j(i-1)) \tag{11}$$

$$a_j(i+1) = a_j(i) + \Delta a_j(i+1) + \eta(a_j(i) - a_j(i-1)) \tag{12}$$

$$b_j(i+1) = b_j(i) + \Delta b_j(i+1) + \eta(b_j(i) - b_j(i-1)). \tag{13}$$

Step 5: *Judgment of whether the end conditions are met.* First the algorithm judges whether the targeted error function is less than the predetermined threshold $\varepsilon(\varepsilon > 0)$ or exceeds the maximum iterations. When $E < \varepsilon$, network training is stopped; otherwise, the algorithm returns to step 3 to continue training.

3. Water quality prediction based on wavelet neural network

3.1. Research object and data source

The research object was the freshwater pearl aquaculture pond water quality index of Jishan Lake in Duchang County, Jiangxi Province. The researchers took the ecological environment monitoring data of the mussel aquaculture pond as research samples; each sample included solar radiation, water temperature, dissolved oxygen, pH, humidity, and wind speed. The sampling period was from July 21 to July 27, 2010. Data were collected once every 60 min for a total of 168 samples; 144 samples were extracted as training sets; 24 samples from the final day of sampling, including the quantitative predicts of dissolved oxygen concentrations, were used as test sets. The variation curves of the original data are shown in Fig. 3.

In order to reduce the influence of the prediction performance due to the different dimensions of sample data, the sample data were normalized to improve the prediction accuracy according to the following formula:

$$x' = \frac{x - x_{\max}}{x_{\max} - x_{\min}} \tag{14}$$

where $x_{\min}$ and $x_{\max}$ are the maximum and minimum values of the raw data, respectively, x is the data before normalization, and x' is the data after normalization.

3.2. Experimental platform of the algorithm and parameter initialization

The algorithm was realized by Matlab7.13 programming language, in which the structure of the wavelet neural network and the BP neural network were 6–4–1 and the structure of the Elman neural network was 6–4–1. The initialized connection weights, wavelet function expansion parameters, and translation parameters were generated randomly from $[-\frac{5}{\sqrt{6}}, \frac{5}{\sqrt{6}}]$. The maximum iteration was 500, the threshold of error precision ε was 0.005, and the learning rate η was 0.3. The first half hour of the dissolved oxygen, pH, temperature, air humidity, wind speed, and solar radiation levels were used as inputs;

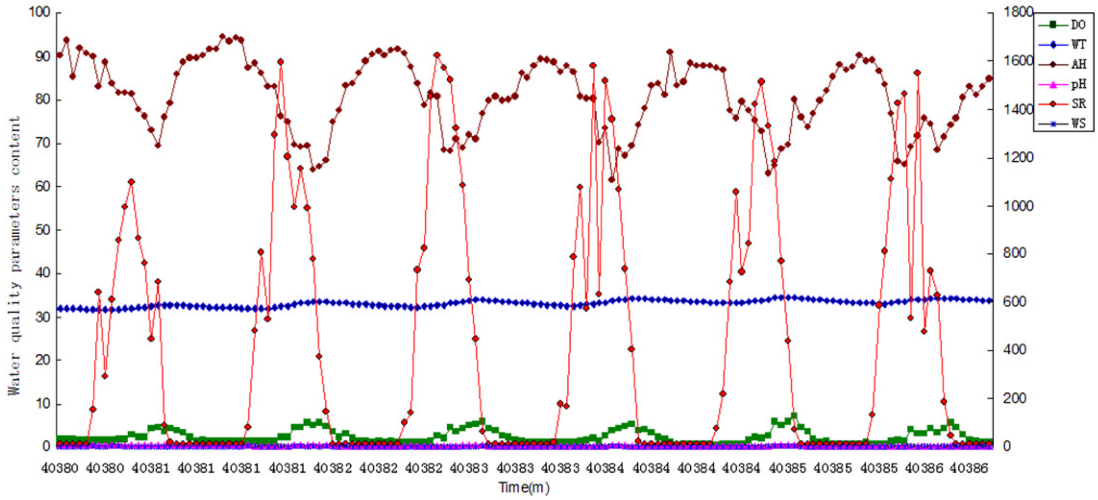

Fig. 3. Variation curves of the original water quality data.

the subsequent dissolved oxygen predictive values were used as the outputs. According to the steps of the wavelet neural network training algorithm, after the prediction model is properly trained, predict of the dissolved oxygen concentration of the freshwater pearl aquaculture pond was possible.

3.3. Prediction error analysis

The prediction error was used to calculate the absolute percentage error (APE) and the mean absolute percentage error (MAPE), which were defined as follows:

$$\text{APE} = \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (15)$$

$$\text{MAPE} = \frac{1}{N}\sum_{t=1}^{N} \frac{|y_i - \hat{y}_i|}{y_i} \quad (16)$$

where y_i is the true value, $\hat{y}_i$ the predicted value, and N the number of testing samples. Maximum and minimum absolute percentage errors were APE_{max} and APE_{min}.

In order to test the algorithm performance, the researchers compared the wavelet neural network algorithm with the common BP neural network model and the Elman neural network model. The algorithm and structure of the three models were identical without wavelet transform. The prediction curves generated based on 48 sets of data collected on July 27, 2010, are shown in Fig. 4, while the water quality prediction results are listed in Table 1.

Fig. 4 shows the water quality prediction based on the wavelet neural network algorithm, which can better fit the complex nonlinear relationship between the ecological environment factors and dissolved oxygen; furthermore, the prediction value and real value fitting curves are better than the prediction results provided by the Elman neural network and the standard BP neural network. The improved prediction results of the wavelet neural network algorithm correspond to the real freshwater pearl aquaculture water quality.

The prediction accuracy was improved greatly compared to the standard BP neural network and the Elman neural network models. As can be seen from Table 1, the mean absolute percentage error dropped from 17.464 and 8.438 to 3.822; the maximum absolute error (APE_{max}) dropped from 40.715 and 38.601 to 9.668; and the minimum absolute percentage error (APE_{min}) dropped from 1.432 and 0.468 to 0.026.

This study combined the technique of wavelet transform and BP neural network, establishing the wavelet neural network prediction model. This network learning model realized fewer tuning parameters, wavelet basis function with compact support, very slight neuronal interactions, high learning speed, no local minimum points, and suitable fitting effect results. The overall effect in predicting water quality in freshwater pearl breeding ponds demonstrated obvious improvement, thus offering a new method for intensive freshwater pearl farming water quality prediction.

4. Conclusions

Water quality prediction plays an important role in control, management, and planning of aquaculture. The BP neural network and Elman neural network prediction models have many weaknesses, so we combined the theory of wavelet

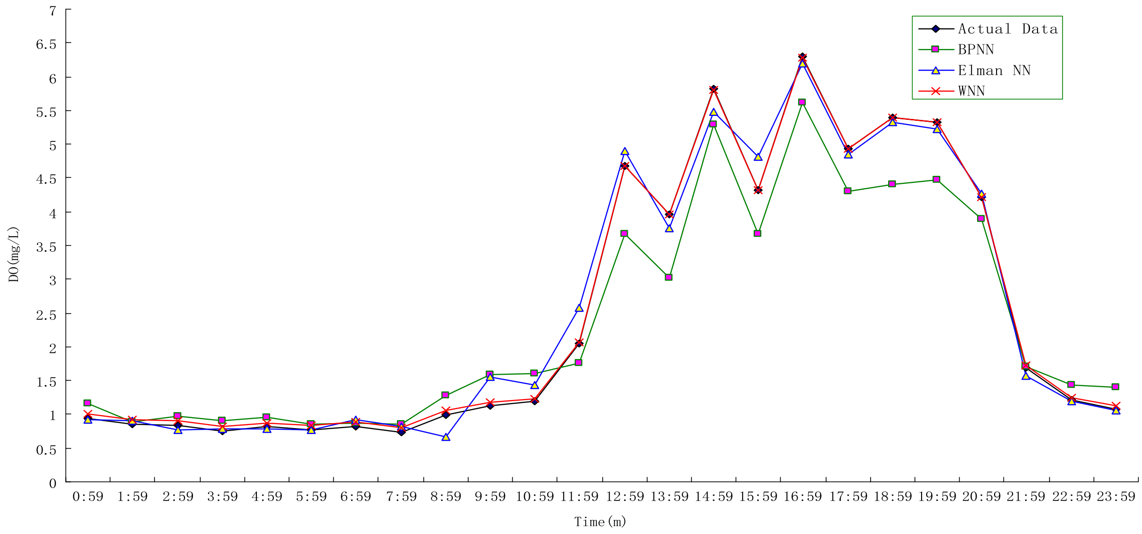

Fig. 4. Comparison of predictions by Elman NN, standard BPNN and WNN.

Table 1
Prediction result comparisons for July 27, 2010.

Time	Real value of DO (mg/L)	BPNN		Elman NN		WNN	
		Predicted value of DO (mg/L)	APE (%)	Predicted value of DO (mg/L)	APE (%)	Predicted value of DO (mg/L)	APE (%)
0:59	0.941	1.155	22.713	0.929	1.257	1.003	6.598
1:59	0.854	0.896	4.841	0.897	5.014	0.916	7.254
2:59	0.833	0.966	15.974	0.761	8.592	0.898	7.822
3:59	0.751	0.900	19.845	0.780	3.853	0.815	8.530
4:59	0.817	0.956	16.930	0.779	4.622	0.877	7.382
5:59	0.766	0.850	10.968	0.770	0.545	0.830	8.399
6:59	0.821	0.873	6.397	0.930	13.366	0.884	7.740
7:59	0.730	0.855	17.081	0.828	13.332	0.801	9.668
8:59	0.994	1.277	28.397	0.673	32.333	1.057	6.310
9:59	1.125	1.582	40.715	1.559	38.601	1.177	4.679
10:59	1.190	1.596	34.205	1.432	20.356	1.237	4.013
11:59	2.051	1.753	14.528	2.584	25.964	2.073	1.042
12:59	4.676	3.676	21.375	4.903	4.864	4.670	0.122
13:59	3.958	3.027	23.515	3.750	5.246	3.953	0.106
14:59	5.829	5.286	9.330	5.479	6.020	5.811	0.308
15:59	4.326	3.679	14.954	4.820	11.438	4.317	0.204
16:59	6.298	5.611	10.906	6.194	1.652	6.278	0.320
17:59	4.942	4.300	12.985	4.851	1.834	4.937	0.095
18:59	5.400	4.398	18.554	5.332	1.268	5.402	0.026
19:59	5.323	4.475	15.936	5.219	1.950	5.328	0.084
20:59	4.216	3.890	7.724	4.270	1.274	4.218	0.043
21:59	1.689	1.713	1.432	1.574	6.811	1.730	2.453
22:59	1.206	1.442	19.504	1.201	0.468	1.254	3.990
23:59	1.080	1.408	30.338	1.063	1.594	1.129	4.521

transform with the BP neural network. The improved model integrated the function of self-learning, adaptive and nonlinear approximation technology. Based on these features, this paper proposed the short-term aquaculture water quality prediction model based on wavelet neural network to improve the prediction accuracy. The model was applied to predict the water quality of intensive freshwater pearl breeding ponds in Duchang County, Jiangxi Province. The proposed wavelet neural network proved to be faster, more precise, and more robust than the BP neural network and the Elman neural network, and was shown to be able to predict effectively the water quality in intensive freshwater pearl breeding. Algorithm identification accuracy was greater than 90%. The wavelet neural network also has higher prediction precision and stronger learning and generalization ability compared with the traditional BP neural network and the Elman neural network. The proposed model can meet the management requirements in intensive freshwater pearl breeding. Because water quality is heavily affected by hydrological and meteorological factors, establishing different predictive models according to weather conditions and combining those prediction models to improve the prediction accuracy is suggested as topics for further research.

Acknowledgments

The authors would like to thank native English speaking expert Schiller Laurie Elaine, for polishing our paper. This paper was supported by the Guangdong Science and Technology Program Project (2012A020200008, 2011B040200034), Guangdong's Natural Science Foundation Project (S2012010008261), and the Zhanjiang Science and Technology Program Project (2010C3113011).

References

[1] J. Abaurrea, J. Asín, A.C. Cebrián, M.A. García-Vera, Trend analysis of water quality series based on regression models with correlated errors, Journal of Hydrology 400 (3) (2011) 341–352.
[2] X.Y. Wang, X.P. Zhao, Z.W. Liu, S.Q. Dong, Research on lake eutrophication predictioning methods based on grey theory, Computer Simulation 28 (1) (2011) 17–19.
[3] D.Ö. Faruk, A hybrid neural network and ARIMA model for water quality time series prediction, Engineering Applications of Artificial Intelligence 23 (4) (2010) 586–594.
[4] S.S. Mahapatra, S.K. Nanda, B.K. Panigrahy, A cascaded fuzzy inference system for Indian river water quality prediction, Advances in Engineering Software 42 (10) (2011) 787–796.
[5] J. Eynard, S. Grieu, M. Polit, Wavelet-based multi-resolution analysis and artificial neural networks for predictioning temperature and thermal power consumption, Engineering Applications of Artificial Intelligence 24 (3) (2011) 501–516.
[6] X.P. Wang, B.P. Qin, Y.S. Li, Application of seasonal ARIMA model in prediction DO of yuqiao reservoir, Northwest Water Resources & Water Engineering 21 (2) (2010) 39–41.
[7] L.X. Guo, C.H. Deng, Prediction model for dissolved oxygen in fish pond based on fuzzy neural network, Journal of Fisheries of China 30 (2) (2006) 225–229.
[8] M.Q. Feng, X.P. Xing, P.S. Xue, Water quality prediction using BP network and Markov model: based on grey incidence analysis, Journal of Natural Disasters 20 (5) (2011) 169–175.
[9] R.M. Wang, Z.T. Fu, Y.Y. He, Prediction model of dissolved oxygen fuzzy system in aquaculture pond based on neural network, Agricultural Science & Technology 11 (8) (2010) 14–18.
[10] H.G. Han, Q.L. Chen, J.F. Qiao, An efficient self-organizing RBF neural network for water quality prediction, Neural Networks 24 (7) (2011) 717–725.
[11] S.S. Mahapatra, S.K. Nanda, B.K. Panigrahy, A cascaded fuzzy inference system for Indian river water quality prediction, Advances in Engineering Software 42 (10) (2011) 787–796.
[12] R.H. Abiyev, Fuzzy wavelet neural network based on fuzzy clustering and gradient techniques for time series prediction, Neural Computing & Applications 20 (2) (2011) 249–259.
[13] G. Inoussa, P. Hui, W. Jun, Nonlinear time series modeling and prediction using functional weights wavelet neural network-based state-dependent AR model, Neurocomputing 86 (1) (2012) 59–74.
[14] X.D. Xiang, The application research of non-linear predictioning based on wavelet neural network predictioning method, Modern Management Science 6 (2) (2007) 32–33.
[15] S. Yilmaz, Y. Oysal, Fuzzy wavelet neural network models for prediction and identification of dynamical systems, IEEE Transactions on Neural Networks 21 (10) (2010) 1599–1609.