J. Agr. Sci. Tech. (2013) Vol. 15: 343-356

Application of Artificial Neural Network in Environmental Water Quality Assessment

H. B. Chu[1], W. X. Lu[1*], and L. Zhang[2]

ABSTRACT

Water quality assessment provides a scientific basis for water resources development and management. This case study proposes a Factor analysis- Hopfield neural network model (FHNN) based on factor analysis method and Hopfield neural network method. The results showed that the factor analysis (FA) technique was introduced to identify important water quality parameters. Results revealed that biochemical oxygen demand, permanganate index, ammonia nitrogen, nitrogen, Cu, Zn and Pb were the most important parameters in assessing water quality variations of the study area. Considering these parameters, water samples of the sampling sites were classified as follows: six into Class III, eight into Class IV, and six into Class V. Afterwards, a water quality map was based on the results of water quality assessment by Factor analysis-Hopfield neural network model. It showed that the southwestern part of the study area had a generally optimum water quality, while in the northeastern part, the quality was seriously degraded. Factor Analysis-Hopfield Neural Network was much better than the Hopfield Neural Network in effectively reducing the degree of Hopfield neural network over-fitting caused by the inputs, thereby achieving more reasonable results. The comparisons with BPANN, fuzzy assessment method, and the Nemerow index method indicated that the FHNN model provided more reliable judgment and valuable information than the three other water quality classification methods.

Keywords: Eastern Liao river, Hopfield neural networks, Factor analysis, Water quality evaluation.

INTRODUCTON

Water resources management entails the development of appropriate quantities of water with an adequate quality (Fulazzaky *et al.*, 2010). It is required to ascertain the quality for various purposes such as drinking, agriculture, recreation, and industry (Khan *et al.*, 2003; Ali *et al.*, 2010; Prabu *et al.*, 2011). Many traditional approaches and techniques have been used in water quality assessment including multivariate statistical methods such as cluster analysis (CA), factor analysis (FA), principal component analysis (PCA) and discriminant analysis (DA). These have been applied to understand the water quality of different study areas and to identify the major factors affecting it. Also, water quality indices (WQI), proposed on the basis of comparison of the actual water quality parameters with the respective regulatory standards, have been used to summarize the large amounts of water quality data into a single number and identify its suitability (e.g., excellent, good, bad, etc.) for different purposes. In addition, remote sensing techniques can be used for water quality assessment (Bordalo *et al.*, 2006; Vignolo *et al.*, 2006; Hussain *et al.*, 2008; Chapagain *et al.*,2010; Boyacioglu 2010; Akbal *et al.*, 2011).

[1] Department of Hydrology and Water Resources, College of Environment and Resources, Jilin University, China.
[*] Corresponding authour; e-mail: luwennxi@jlu.edu.cu
[2] Department of Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, China.

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

During the last two decades, Artificial Neural Network (ANN) has seen an explosion of interest because it is an effective method for prediction, clustering and classification. This approach is becoming an effective and popular alternative for conventional methods (Yazdani *et al*., 2009; Cho *et al*., 2011; Ghasemloo *et al*., 2011). ANN constitutes intelligent bionic models and the nonlinear, large-scale, adaptive dynamics systems which consist of many interconnected neurons. ANN models have been widely applied to the water quality problems (Hornik, 1991; Lee *et al*., 1996; Capolo *et al*., 1999; Chen *et al*., 2001). In Turkey, a three-layer Levenberg–Marquardt feedforward neural network was used to model the eutrophication process in the water bodies (Karul *et al*., 2000). In Greece, ANN are used to forecast the maximum daily value of the European Regional Pollution Index as well as the number of consecutive hours of pollution during the day, 24 to 72 hours ahead (Moustris *et al*., 2010). In fact, water quality assessment using ANN is a typical pattern recognition problem that classifies water quality according to the standards. According to their network structure, ANNs can be divided into feedforward and the feedback networks, with Hopfield neural network (HNN) being a representative feedback network. Based on its associative memory function, HNN has been successfully introduced to the application of water quality assessment (Long *et al*., 2002).

Various water quality parameters have a mutual influence on a water body, but the degree of this influence varies and, therefore, water quality is difficult to evaluate when the number of parameters is large (Almeida *et al*., 2007). For a HNN model, not all existing factors are necessary to evaluate water quality. In fact, if these factors are not added optionally as input layers, other noise factors will be added into the model instead, decreasing the prediction ability. This situation is called overfitting. Factor analysis is a statistical analysis method which determines the main influencing parameters among many variables (Panchal *et al*., 2011; Ding *et al*., 2011). In research, factor analysis has been used along with a neural network model to solve the problem described above.

In recent years, one of the issues under discussion in neural network algorithm research has been the integration of multivariate statistical methods and Artificial Neural Network, as has been done in a wide variety of environmental applications (Ding *et al*., 2011). A biochemical oxygen demand (BOD) forecasting model in Brazil revealed that the best prediction performance was achieved when the data were processed using principal components analysis (PCA) before being fed to a backpropagation neural network (Oliveira-Esquerre *et al*., 2002).The farm smell forecasting model was constructed by the factor analysis and neural network method, and simulated the process of smell composition and occurrence patterns (Kevin *et al*., 2005). Kohonen neural network (KNN) and factor analysis were applied to regional geochemical pattern recognition for a Pb–Zn–Mo–Ag mining area in Qinghai Province, China. The results demonstrated that the approach effectively interpreted the geological significance of the factors, and also reduced the area of exploration targets (Sun *et al*., 2009). Until now, however, the factor analysis-Hopfield Neural Network (FHNN) combination has been rarely applied in practical situations.

The aims of the present study were (1) to solve the HNN over-fitting problem caused by the inputs, (2) to apply FHNN models to assess the water quality in the eastern Liao River region of Jilin Province, China and (3) to compare the advantages and disadvantages of the BP and HNN water quality assessment models.

MATERIALS AND METHODS

Assessment Criterion

The GB3838-2002 “Environmental quality standard for surface water” is employed as a comprehensive assessment criterion of the water quality, which classifies surface water

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

quality into five levels corresponding to surface water environmental functionality and protection targets:

Class I: Meet the water quality requirements for headwaters and National Nature Reserves.

Class II: Meet the water quality requirements for the first-level protection areas of drinking water reservoirs, rare aquatic creatures habitat, fish spawning ground, etc.

Class III: Meet the water quality requirements for the second-level protection areas of drinking water reservoirs, aquatics breeding area, swimming, and so on.

Class IV: Meet the water quality requirements for industrial and entertainment uses that do not involve direct contact with the human body.

Class V: Meet the water quality requirement for agricultural and ordinary landscape uses.

Class I represents the best water quality and Class V the worst, therefore, moving from Class I to Class V represents deterioration in water quality (GB3838-2002, China, 2002).

Experimental Data

In this research, the eastern Liao River was selected as the study area and the FHNN model, based on factor analysis and the Hopfield neural network, as the analytical tool for water quality assessment. The eastern Liao River, one of 19 major rivers that present a substantial mismatch between water supply and demand, is located in the western region of Jilin Province, China.

The scope of this work involved analyzing the samples for total nitrogen (TN), total phosphorus (TP), dissolved oxygen (DO), biochemical oxygen demand (BOD_5), permanganate index (COD_{Mn}), chemical oxygen demand (CODcr), ammonia nitrogen (NH_3–N), and chromium (Cr), arsenic (As), copper (Cu), zinc (Zn) and lead (Pb). Readings from the twenty monitoring sites during July, 2010 are shown in Figure 1. (Point1 is located at the junction of the eastern Liao River and the western Liao River).

METHODS

In the present study, the factor analysis method was used to determine the main parameters affecting water quality assessment, and these parameters were then used as input data sets for a Hopfield neural network to build a HNN water quality comprehensive assessment model. Then the BP and HNN water quality comprehensive assessment models were compared.

Introduction to Factor Analysis

Factor analysis is a statistical method which reduces dimensionality in the process of multivariate analysis. It also integrates a large number of variables, which have intricate relationships, into several significant factors and achieve a meaningful synthesis. These common factors explain the correlations among the observed variables.

If there are N number of water samples and each sample has P number of water quality parameters (x_1, x_2,... x_p), an $N\times P$ data matrix can be defined:

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

the observed variables are modeled as linear combinations of the common factors, plus the unique factors. In Factor Analysis model,

$$X = AF + \varepsilon \quad (1)$$

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

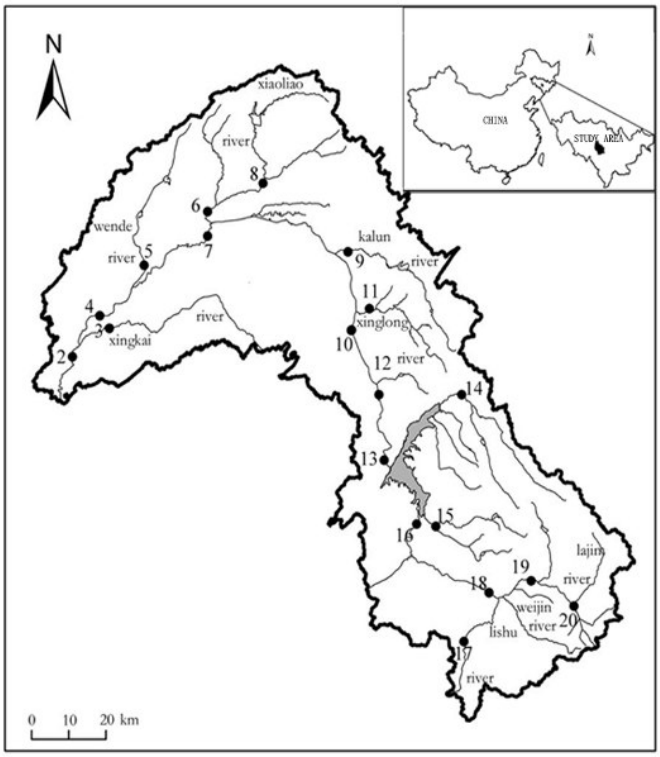

Figure 1. Map of the sampling points along Eastern Liao River (China).

Where,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix}$$ Factor loading matrix

$F = (F_{1,}\ F_{2,}\ \cdots, F_m)^T$ (Common factor matrix)

$\varepsilon = (\varepsilon_{1,}\ \varepsilon_{2,}\ \cdots, \varepsilon_m)^T$ (Unique factor matrix)

a_{ij} is the factor loading, which represents the related coefficient of the *i*-variable and the *j*-factor, F is the common factor, ε is the unique factor.

The common factors appear together in the expression of each original variable and contribute to the covariation among the variables. The unique factors are independently distributed error terms with zero mean and finite variance, and they are assumed to be orthogonal to each other, and don't contribute to the covariation between the variables. The minimum number of the common factors is determined according to the cumulative contribution rate (CCR), and it is possible to extract the information into a few common factors with the minimum loss. Through identifying *m* "common factors", the method displays optimally the differences among the *p* variables (*m*< *p*). Then, it needed to carry on a series of rotations on the factor loading matrix, finally, the factor scores were calculated. The specific steps about the factor analysis had been described in the study. (Ouyang, 2005; Astel *et al*., 2007).

Factor analysis (FA) techniques can identify important water quality parameters, but they cannot adequately assess the current level of water quality, furthermore, the public, managers, and policy makers require concise information about the water bodies. Therefore, HNN assessment model was used to complete the classification of water quality.

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

Introduction to Hopfield Neural Network (HNN)

The HNN, proposed by J. J. Hopfield in 1982, is a recurrent single-layer interconnection neural network made up of a large number of neurons. Also, it has a symmetric connection structure because any two neurons are connected. HNNs have been successfully used to solve real-world problems, including assignment problems, scheduling problems, shortest-path problems, traveling salesman problems(TSPs), and vehicle routing problems (Hopfield, 1982). Based on the form of their output functions, Hopfield networks can be classified into one of two popular forms: discrete and continuous-time models. In this work, a discrete Hopfield neural network was chosen to classify water quality. The discrete Hopfield Neural Network (DHNN) is a single-layer and binary-type feedback neural network. All its nodes are connected to each other and the connection weights of each node accepts information feedback from the other nodes. Therefore, the output of any neurons is controlled by the other neurons, meaning that the output of each neuron can restrict that of the other neurons. As a result, each neuron has a threshold value to control the input of the noise (Yaleinoz *et al*., 2001). The network is shown in Figure 2.

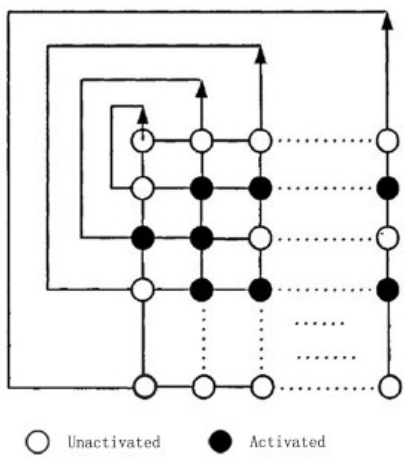

Figure 2. Structure of Hopfield neural network.

In the DHNN model, if $X = [x_1, x_2 \cdots x_n]^T$ is defined as the network state vector, the components are the output of *n* neurons and only take on values of -1 or 1.

$$X_i = \begin{cases} 1 & \sum_{i \neq i} W_{ij} x_j - \theta_i > 0 \\ -1 & \sum_{j \neq i} W_{ij} x_j - \theta_i \leq 0 \end{cases} \quad (2)$$

Where, W_{ij} is the matrix of the connection weights between units i and j; θ_i is the connection threshold of unit I; $W_{ij} = W_{ji}$, and $W_{ii} = 0$, that is, the DHNN symmetrical connection has no self-feedback.

Because a Hopfield neural network is a complex nonlinear dynamic system, its system stability is analyzed using *energy function*. The network state transitions from high energy to the minimum energy state and, finally, converges to the stable state of the system. When the function reaches steady state, an approximate solution of the problem has been obtained, and the local minimum of the energy function corresponds to the energy of the stored patterns. Therefore, the Hopfield neural network provides a good solution to the associative memory problem (Wen *et al*., 2009). The Liapunov function form can be written as follows:

$$E = -\frac{1}{2} \sum_i \sum_j W_{ij} X_i X_j + \sum_j \theta_j X_j \quad (3)$$

Where, X_i, X_j are the first two state variables, take on values of 1 or -1; W_{ij} is the matrix of the connection weights between units i and j; and θ_j is the connection threshold of unit j.

Modeling

In this paper, water quality data of July 2010 was used to extract four common factors according to the cumulative contribution rate and to compute the weight of the water quality parameters by the factor score coefficient and

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

the variance contribution rate. According to the weights of the parameters, the main parameters of water quality were determined as the input variables of and Hopfield neural network to carry out comprehensive assessment of water quality. Figure 3 shows the flow chart of the Factor analysis-Hopfield neural network model.

The basic steps of the new algorithm (FHNN) are as follows:

Step 1: Standardization of the original water quality data *X*.

Step 2: Calculation of the correlation coefficient matrix R.

Step 3: Calculation of the Eigen values of matrix *R*, denoted by $\lambda_1, \lambda_2 \ldots, \lambda_n$ and their arrangement as follows: $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_n$, then, solving the eigenvectors of matrix *R* (Table1).

Step 4: Computation of the variance contribution rate (VCR) of the *i*th Eigen value and the cumulative contribution rate (CCR) expressed as follows:

$$\mathrm{VCR}(\lambda_i) = \frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_n} \quad (4)$$

$$\mathrm{CCR}(k) = \mathrm{VCR}(\lambda_1) + \mathrm{VCR}(\lambda_2) + \ldots + \mathrm{VCR}(\lambda_k) = \frac{\lambda_1 + \lambda_2 + \cdots + \lambda_k}{\lambda_1 + \lambda_2 + \cdots + \lambda_k \cdots + \lambda_n} \quad (5)$$

Table1. The rate of factor contribution and the corresponding cumulative contribution.

	Eigen value	*VCR*	*CCR*
1	5.006	0.417	0.417
2	3.263	0.272	0.689
3	1.247	0.104	0.793
4	0.905	0.075	0.868

When CCR reached no less than 85%, the main common factors were determined.

Four main common factors (F1, F2, F3, and F4) were attained following the aforementioned steps. The cumulative contribution rate of variance reaches 86.8%, which can represent most of the information of the original data.

Step 5: Carrying the maximum orthogonal rotation on factor loading matrix, calculating the factor score, and then computing factor score coefficients β by regression analysis as follows (Table2).

$$F_i = \beta_{i1} x_1 + \beta_{i2} x_2 + \cdots + \beta_{ip} x_p \qquad i = 1, 2, \cdots, m \quad (6)$$

Where, x_1, x_2,... x_p are the water quality parameters, F_i is the *i*th common factor score, and β_{ij} is the factor score coefficients of the *i*th common factor and *j*th water quality parameters.

Step 6: Counting the weight of all

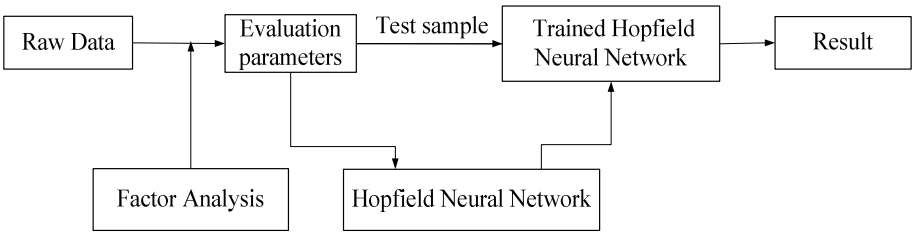

Figure 3. Flow chart of Factor analysis-Hopfield neural network model.

Table 2. The component score coefficient matrix of the first four factors.

	β_1	β_2	β_3	β_4	β_5	β_6	β_7	β_8	β_9	β_{10}	β_{11}	β_{12}
F1	-0.37	0.18	-1.06	-0.10	0.08	-0.08	0.398	0.36	0.66	-3.69	1.91	2.46
F2	0.36	-0.05	0.54	-0.32	-0.13	0.59	-0.54	-0.21	-0.79	0.83	-0.78	0.47
F3	0.08	0.87	-0.24	0.34	0.08	0.05	-0.17	0.06	-0.18	0.22	-0.02	0.01
F4	0.12	0.10	0.41	0.03	0.82	0.528	0.17	-0.33	-0.24	0.60	-0.16	0.03

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

assessment parameters as follows:

$$W_j = \frac{\sum_{i=1}^{m} \beta_{ij} e_i}{\sum_{j=1}^{p} \sum_{i=1}^{m} \beta_{ij} e_i} \qquad (e_i = \frac{\lambda_i}{\sum_{i=1}^{m} \lambda_i}) \tag{7}$$

Where, i is the number of the eigenvalues, i= 1, 2, 3, 4; j is the number of the parameters, 1, 2, …12; m is the number of selected common factors, m= 4; p is the number of water quality parameters, p= 12; β_{ij} is the factor score coefficients of the ith common factor and jth water quality parameters; W_j is the weight of jth water quality parameters.

Step 7: According to the results presented in Table 3, which shows the weight of the water quality parameters in descending order, the cumulative weight of the seven water quality parameters is 85.58%. Therefore, seven parameters can be chosen from the twelve original parameters, including biochemical oxygen demand, permanganate index, and ammonia nitrogen, copper, zinc, nitrogen, and lead content.

Step 8: The 5×7 neurons Hopfield neural network model for water quality assessment was defined, which used the water quality standards (Table 4) as the training samples (GB3838-2002), as follows:

Defined the network memory model, that is, the pre-storage model and the memory patterns were obtained and encoded with using values of 1 and -1. This memory model is illustrated in Figure 4. Called the *newhop* function of MATLAB to define a discrete Hopfield neural network, and to obtain the weight matrix and the threshold vector through the training.

Set the original data coding modes as the initial state and called the *sim* function of MATLAB to achieve network converge by iteration. Once the network was stable, the

Table 3. The weight of evaluation parameters.

	TN	TP	NH_4	COD_{Mn}	As	DO
W_j	0.105	0.046	0.136	0.108	0.010	0.028
	BOD_5	Cr	CODcr	Cu	Zn	Pb
W_j	0.088	0.039	0.023	0.156	0.127	0.134

Table 4. Environmental quality standard for surface water (Unit: mg l^{-1}).

	I [a]	II	III	IV	V
BOD_5	3	3	4	6	10
COD_{Mn}	2	4	6	10	15
NH_3-N	0.015	0.5	1	1.5	2
Cu	0.01	1	1	1	1
Zn	0.05	1	1	2	2
Pb	0.01	0.01	0.05	0.05	0.1
TN	0.2	0.5	1	1.5	2

[a] I- V means the types of water quality.

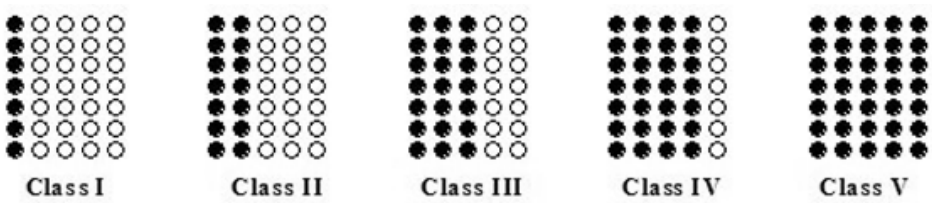

Figure 4. The standard model of water quality(I- V means the types of water quality)

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

results could be performed.

Step 9: Used the trained network to test the samples, and the water quality assessment results at the measurement point.

RESULTS AND DISCUSSION

Seven parameters were selected from the 12 original parameters by Factor analysis, which indicated that water quality of the study area was mainly affected by heavy metal (Zn, Pb, Cu), ammonia nitrogen, and organic pollution. And the FHNN model was used to assess water quality level at the twenty monitoring sites in the eastern Liao River, China. The results of this assessment are described in Figure 5.

In the latest Environmental Status Bulletin issued by the Jilin Provincial Government, water quality assessment values in the Erlongshan reservoir (13th sample) and the town of Liaoyuan (20th sample) were classified into Class IV and III respectively, but these water quality assessment values were classified into Class V and III, respectively, by the FHNN. However, the Hopfield network alone classified water quality at these locations as Class II and I. The assessment results of FHNN were almost the same as the latest Environmental Status Bulletin of 2010, but the results from the HNN alone were smaller. KongJia town (8th sample) was classified into Class V by the FHNN and Class II by the HNN alone, however, because of the direct effects of industrial wastewater pollution from Gongzhuling town, the water quality of KongJia town was not good and did not meet the requirements of Class II for drinking water reservoirs or rare aquatic creatures habitat, but it meets the requirements of Class V for agricultural and ordinary landscape uses. Thus, the classification by the FHNN was in line with the actual situation. From Table 5, it is apparent that the values from the HNN alone are smaller, because of overfitting and the addition of noise factors to the model. The FHNN is applicable not only for determination of quantitative water quality parameters, but also for qualitative parameters. The network design is simple, and the assessment process is intuitive and more stable than the conventional

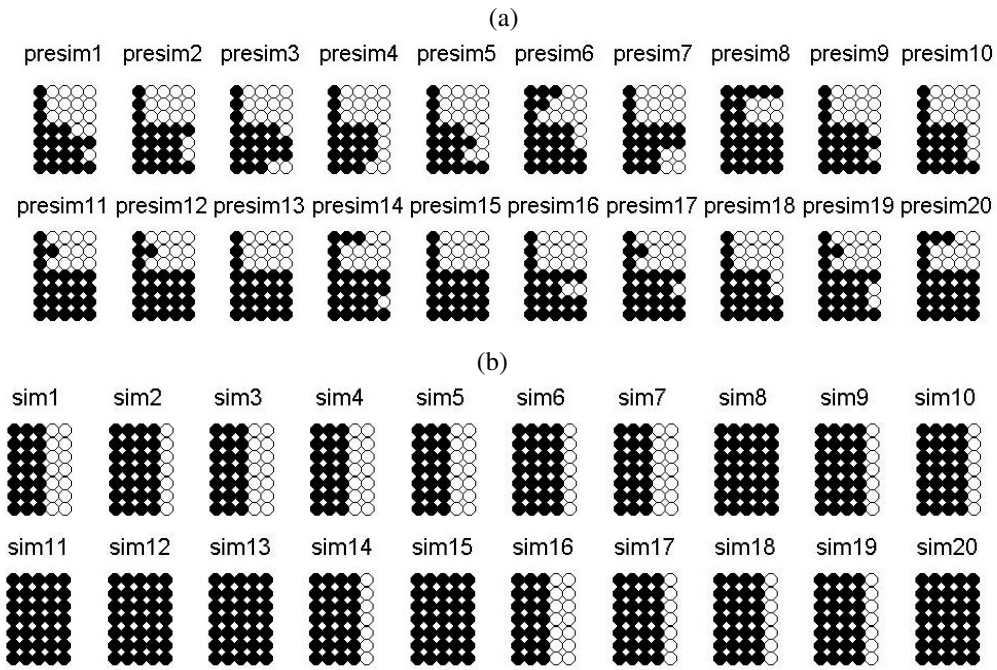

Figure 5. (a): The pre-results for each sampling station (FHNN) **(b):** Water quality classifications for each sampling station (FHNN).

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

Table 5. Comparison of evaluation results in three methods.

	1[a]	2	3	4	5	6	7	8	9	10
HNN	I	I	I	I	I	II	I	II	II	I
FHNN	III	III	III	III	III	IV	III	V	IV	IV
	11	12	13	14	15	16	17	18	19	20
HNN	II	I	II	III	III	III	II	I	II	III
FHNN	V	V	V	IV	V	III	IV	IV	IV	V

[a] 1-20 is the number of the sampling station.

methods. For all these reasons, the algorithm is obviously superior to the HNN alone.

The FHNN model combines the characteristics of factor analysis to identify major water quality parameters with the nonlinear calculation characteristics of the Hopfield neural network. This approach not only compensates, to a certain extent, for the deficiencies of factor analysis in water quality classification in practical application, but also excludes the impact of certain data and effectively reduces the degree of over-fitting caused by the inputs of the HNN. In addition, the FHNN improves network recognition accuracy, simplifies the neural network structure, speeds up network convergence, and reduces the running time.

A water quality map is shown in Figure 6, which was prepared based on the results of water quality comprehensive assessment by the combined factor analysis-Hopfield neural network model. ArcGIS with an interpolation technique (ordinary kriging) was used for the spatial distribution of water quality. The water quality map revealed that six sampling sites (30% of the twenty sampling sites) were in Class II, eight sampling sites (40% of the twenty sampling sites) were in Class IV, and six sampling sites (30% of the twenty sampling sites) were in Class V. Figure 6 shows that the southwestern part of the study area has a generally optimum water quality, while in the northeastern part, the quality is seriously degraded.

BPANN (Back Propagation Artificial Neural Networks) and Hopfield Neural Networks are commonly used to assess water quality. In this research, the three-layer Back Propagation Artificial Neural Network (BPANN) was used. From the above paragraph, seven nodes

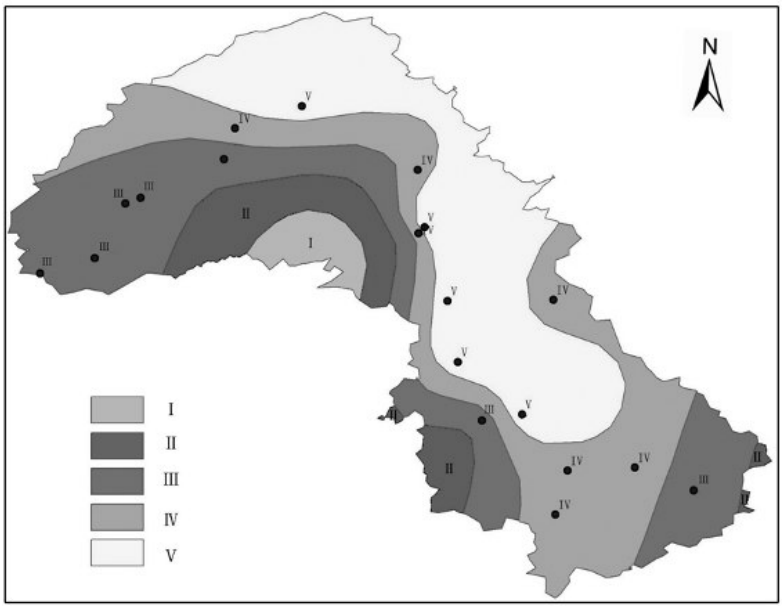

Figure 6. Distribution of Hopfield Quality evaluation results.

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

were defined as the input and one node as the output layer. After repeated tests, the number of hidden layer nodes was three, that is, the final networks were inclusive of seven input nodes, three hidden nodes, and one output node. Provided the initialization learning rate was 0.7 (Fu, 1995), the training accuracy was 1×10^{-6}, and the maximum epochs was 1000. The sum of squared errors calculated for the training or test subsets was chosen as the convergence criteria. After 164 training repetitions, the error was 9.846×10^{-7}, which was less than the training accuracy (1×10^{-6}). The error variance is shown in Figure 7.

In Table 6, BP networks and Hopfield networks are further analyzed and compared with respect to their structures, learning rules, stability, applications, and other aspects.

The HNN is essentially a fixed network, and the weights of the whole neural network do not always change. However, for each input and output data, the BP network needs to adjust its weights. This overall approach will lead to slow learning, which is time consuming. It will take 30 minutes to implement water quality assessment by BP networks, however, HNN can implement water quality assessment in a few seconds, reducing execution time substantially.

The HNN and BPANN procedures were each run 100 times to investigate the operational stability of the network, and the results of this assessment are shown in Table 7. It is apparent that the BPANN assessment results fluctuate over a certain range, while the Hopfield network achieved 100% accuracy. It can, therefore, be concluded that the Hopfield neural network is more stable than the BP neural network.

In this research, the BPANN, the HNN, a fuzzy assessment method, and the Nemerow index method were used to assess the water quality of 20 samples from the eastern Liao River and the results were compared. The assessment criterion was the GB3838-2002 standard "Environmental quality standard for surface water". For the sake of comparison, all the four assessment methods used the seven factors described above.

Figure 8 shows that the variability exhibited

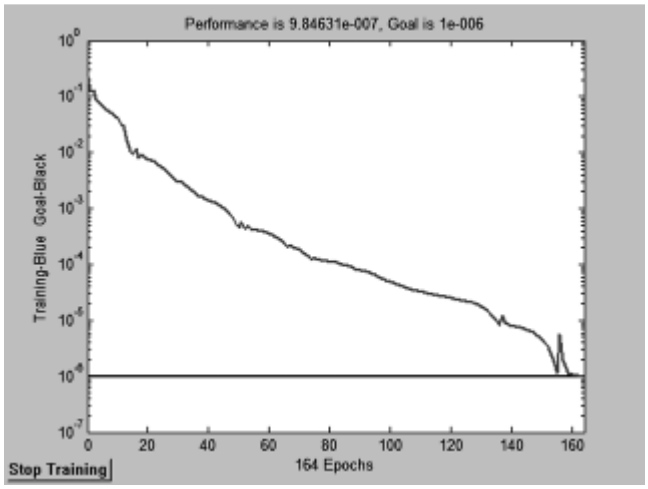

Figure 7. Variations in error function during training of the BPANN.

Table7. Evaluation Results of BP and Hopfield network (Unit: Times).

sampling station	HNN				BPNN			
	I	II	III	result	I	II	III	result
1	100			I	10	85	5	II
2		100		II	4	90	6	II
3			100	III	3	5	92	III

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

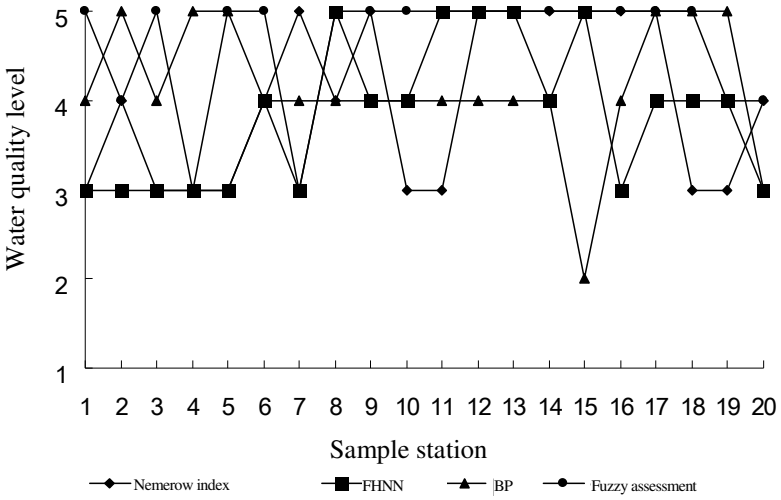

Figure 8. Water Quality Evaluation results of four methods.

(Class I-1, Class II-2, Class III-3, Class IV-4, Class V-5)

by the Nemerow index method and the fuzzy pattern recognition method was greater than that of the other methods because these two methods were too sensitive to larger values of pollution. The degree of approximation to the other assessment results achieved by the HNN was better than that obtained by the BP network, and water quality assessment results by the HNN and BP network had a greater deviation for samples 2, 5, and 15. In fact, quality of water samples 2 and 5 meet the requirements of Class III for the second-level protection areas of drinking water reservoirs or swimming, and sample 15 meets the requirements of Class V for agricultural and ordinary landscape uses. Therefore, the HNN was adapted to classify water quality for different uses and purposes.

CONCLUSIONS

This study suggests that FA techniques are useful tools for identification of the important surface water quality monitoring parameters. The weighting order of water quality monitoring parameters revealed that biochemical oxygen demand, permanganate index, ammonia nitrogen, and nitrogen, Cu, Zn, and Pb were the most important parameters in assessing water quality variations of the study area. The southwest of the study area had a generally optimum water quality, while waters in the northeast of the study area were seriously polluted. Therefore, more attention should be given to the water quality in this area and effective measures should be taken to control pollution.

The FHNN model has successfully solved the overfitting problem caused by the presence of large number of correlated water-quality parameters in the inputs to the ANN water quality assessment model. It is obvious that the FHNN is much better than the Hopfield neural network alone in effectively reducing the degree of Hopfield neural network over-fitting. It also excludes the impact of certain data to improve network recognition accuracy and to make the classification results more reasonable. At the same time, the FHNN also simplifies neural-network structure, speeds up network convergence, and reduces execution time. All of the four water quality assessment models can be used to classify water quality in the study area. However, the results of this study suggest that the FHNN model is more reliable and objective than the other three methods.

In summary, the FHNN can be considered as a comprehensive tool for assessing the

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

quality of water for different uses because it offers the best guidance on how to allocate rationally the use and development of water resources in the study area. It should be noted that water quality in certain regions of the study area i.e. the eastern Liao River, is seriously degraded and, therefore, proper measures need to be taken such as controlling point sources of pollution and treatment of all wastewater before discharge.

Based on the aforementioned results, two issues are proposed for further study. The first is to investigate other models that might reduce network dimensionality and simplify the network architecture. Examples of such models are PCA (principal component analysis) and PLS (partial least squares). Another central and important issue for further study is to solve the problems caused by the HNN itself including difficulties in training, the large number of local minima in the error surface, and, sometimes, the difficulty in adapting the networks to new data.

ACKNOWLEDGEMENTS

This work was supported by "The National Critical Patented Projects in the Control and Management of Polluted Water Bodies" (2009ZX07526-006-04-01), "The National Natural Science Funds" (41072171) and Key Laboratory of Groundwater Resources and Environment, Ministry of Education, Jilin University, Changchun, China.

REFERENCES

1. Akbal, F., Gürel, L., Bahadır, T., Güler, İ., Bakan, G. and Büyükgüngör, H. 2011. Multivariate Statistical Techniques for the Assessment of Surface Water Quality at the Mid-Black Sea Coast of Turkey. *Water Air Soil Pollu.*, **216:** 21–37.
2. Ali, E., Melike, G., Alpaslan, E., Cigdem, T., Aysegul, U., Dursun, Z. S., Aysegul, T. and Izzet, O. 2010. Water Quality Assessment and Meta Model Development in Melen Watershed–Turkey. *J. Environ. Manage.*, **91(7):** 1526-1545.
3. Almeida,C. A., Quintar, S., González, P. and Mallea, M. A. 2007. Influence of Urbanization and Tourist Activities on the Water Quality of the Potrero de los Funes River (San Luis-Argentina). *Environ. Moni. Assess.*, **133:** 459–465.
4. Astel, A., Tsakovski, S., Barbieri, P. and Simeonov, V. 2007. Comparison of Self-Organizing Approach with Cluster and Principal Component Analysis for Large Environmental Data Sets. *Water Res.*, **41:** 4566–4578.
5. Bordalo, A. A., Teixeira, R. and Wiebe, W. J. 2006. A Water Quality Index Applied to an International Shared River Basin: The Case of the Douro River. *Environ. Manage.*, **38:** 910–920.
6. Boyacioglu, H. 2010. Utilization of the Water Quality Index Methods A Classification Tool. *Environ. Moni. Assess.*, **167:** 115–124.
7. Capolo, M., Andreussi, P. and Soldati, A. 1999. River Filled Forecasting with a Neural Network Model. *Water Resource Res.*, **35:** 1191–1197.
8. Chapagain, S. K., Pandey, V. P., Shrestha, S., Nakamura T. and Kazama, F. 2010. Assessment of Deep Groundwater Quality in Kathmandu Valley Using Multivariate Statistical Techniques. *Water Air Soil Pollu.*, **210:** 277–288.
9. Chen, R. M. and Huang, Y. M. 2001. Competitive Neural Network to Solve Scheduling Problems. *Neurocomputing*, **37:** 177–196.
10. Cho, K. H., Sthiannopkao, S., Pachepsky, Y. A., Kim, K. W. and Kim, J. H. 2011. Prediction of Contamination Potential of Groundwater Arsenic in Cambodia, Laos, and Thailand Using Artificial Neural Network. *Water Res.*, **45(17):** 5535-5544.
11. Ding, S. F., Jia, W. K., Su, C. Y., Zhang L. W. and Liu, L. L. 2011. Research of Neural Network Algorithm Based on Factor Analysis and Cluster Analysis. *Neural Compu. Appli.*, **20(2):** 297-302.
12. Fu, L. 1995. Neural Networks in Computer Intelligence. McGraw-Hill, New York.
13. Fulazzaky, M. A., Seong, T. W. and Masirin, M. I. M. 2010. Assessment of Water Quality Status for the Selangor River in Malaysia. *Water Air Soil Pollu.*, **205:** 63–77.

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

14. GB3838. 2002. *Environmental Quality Standard for Surface Water*. Ministry of Environmental Protection of the People's Republic of China, China.
15. Ghasemloo, N., Mobasheri, M. R. and Rezaei, Y. 2011. Vegetation Species Determination Using Spectral Characteristics and Artificial Neural Network (SCANN). *J. Agr. Sci. Tech.*, **13(Suppl):** 1223-1232.
16. Hopfield, J. J. 1982. Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proceedings of the National Academy of Sciences of the United States of America,* **79:** 2554–2558.
17. Hornik K. 1991. Approximation Capability of Multi-layer Feedforward Networks. *Neural Networks*. **4:** 241–257.
18. Hussain, M., Ahmed, S. M. and Abderrahman, W. 2008. Cluster Analysis and Quality Assessment of Logged Water at an Irrigation Project, Eastern Saudi Arabia. *J. Environ. Manage.,* **86(1):** 297–307.
19. Karul, C., Soyupak, S., Cilesiz, A. F., Akbay, N. and Germen, E. 2000. Case Studies on the Use of Neural Networks in Eutrophication Modeling. *Ecological Modelling,* **134:** 145–152.
20. Kevin, R. J., Simon, X. Y. and Roger, R. H. 2005. Pork Farm Odor Modeling Using Multiple-Component Multiple-Factor Analysis and Neural Networks. *Appl. Soft Computing,* **6:** 53–61.
21. Khan, F., Husain, T. and Lumb, A. 2003. Water Quality Evaluation and Trend Analysis in Selected Watersheds of the Atlantic Region of Canada. *Environ. Monit. Assess.,* **88:** 221–242.
22. Lee, K. C., Han, I. and Kwon, Y. S. 1996. Hybrid Neural Network Models for Bankruptcy Predictions. *Decision Support Sys.*, **18:** 63–72.
23. Long, T. R., Guo, J. S. and Huo, G. Y. 2002. Hopfield Network Models of Water Quality Evaluation. *J. Chongqing Jianzhu Univ.,* **24(2):** 57-60.
24. Moustris, K. P., Ziomas, I. C. and Paliatsos, A. G. 2010. 3-day-ahead Forecasting of Regional Pollution Index for the Pollutants NO_2, CO, SO_2, and O_3 Using Artificial Neural Networks in Athens, Greece. *Water Air Soil Pollu.*, **209:** 29–43.
25. Oliveira-Esquerre, K. P., Mori, M. and Bruns, R. E. 2002. Simulation of an Industrial Wastewater Treatment Plant Using Artificial Neural Networks and Principal Components Analysis. *Brazil. Jour. Chem. Engin.*, **21:** 81–87.
26. Ouyang, Y. 2005. Evaluation of River Water Quality Monitoring Stations by Principal Component Analysis. *Water Research*, **39:** 2621–2635.
27. Panchal, G., Ganatra, A., Kosta, Y. P. and Panchal, D. 2011. Behaviour Analysis of Multilayer Perceptrons with Multiple Hidden Neurons and Hidden Layers. *Inter. Jour. Compu. Theory Engin.,* **3(2):** 332-337.
28. Prabu, P. C., Wondimu, L. and Tesso, M. 2011. Assessment of Water Quality of Huluka and Alaltu Rivers of Ambo, Ethiopia. *J. Agr. Scie. Tech.*, **13(1):** 131-138.
29. Sun, X., Deng, J., Gong, Q. J., Wang. Q. F., Yang, L. Q. and Zhao, Z. Y. 2009. Kohonen Neural Network and Factor Analysis Based Approach to Geochemical Data Pattern Recognition. *J. Geochem. Explor.,* **103(1):** 6-16.
30. Vignolo, A., Pochettino, A. and Cicerone, D. 2006. Water Quality Assessment Using Remote Sensing Techniques: Medrano Creek, Argentina. *J. Environ. Manage.,* **81(4):** 429-433.
31. Wen, U. P., Lan, K. M. and Shih, H. S. 2009. A review of Hopfield Neural Networks for Solving Mathematical Programming Problems. *Euro. Jour. Operat. Res.*, **198:** 675–687.
32. Yaleinoz, T., Cory, B. J. and Short, M. J. 2001. Hopfield Neural Network Approaches to Economic Dispatch Problems. *Inter. Jour. Elec. Power Energy Sys.,* **23:** 435–442.
33. Yazdani, M. R., Saghafian, B., Mahdian, M. H. and Soltani, S. 2009. Monthly Runoff Estimation Using Artificial Neural Networks. *J. Agr. Sci. Tech.,* **11(3):** 355-362.

[DOR: 20.1001.1.16807073.2013.15.2.5.4]

کاربرد شبکه عصبی مصنوعی در ارزیابی کیفیت آب در محیط زیست

ح. ب. چو، و. ژ. لو، و ل. ژانگ

چکیده

ارزیابی کیفیت آب مبنایی علمی برای برنامه های توسعه و مدیریت منابع آب به دست میدهد.برای این منظور،مطالعه حاضر روشی شامل کاربرد ترکیبی روش تحلیل فاکتوری(Factor analysis) ومدل شبکه عصبی هاپفیلد(Hopfield neural network) پیشنهاد می کند. در این مطالعه روش تحلیل فاکتوری برای شناسایی پارامترهای اصلی موثر در کیفیت آب به کار رفت. نتایج نشان داد که نیاز زیستی به اکسیژن (BOD) ، شاخص پرمنگنات، نیتروژن آمونیاکی، نیتروژن، مس، روی، و سرب مهمترین پارامترها در ارزیابی تغییرات کیفیت آب در منطقه مطالعه بودند. با در نظر گرفتن این پارامترها، آب مناطق نمونه برداری شده به صورت زیر طبقه بندی شدند: ۶ مورد در کلاس III ، ۸ مورد در کلاس IV، و۶ مورد در کلاس V. بر اساس این ارزیابی ها، یک نقشه کیفیت آب بر اساس نتایج روش ترکیبی تحلیل فاکتوری و مدل شبکه عصبی هاپفیلد تهیه شد. این نقشه نشان داد که بخشهای جنوب غربی منطقه مطالعه به طورکلی کیفیت آب بهینه ای داشتند در حالیکه در بخشهای شمال شرقی منطقه کیفیت آب شدیدا کاهش یافته بود. روش ترکیبی ارزیابی کیفیت آب از روش مدل شبکه عصبی هاپفیلد بسیار بهتر بود و نتایج قابل قبولتری به دست داد که علت آن کاهش موثر مسایل ناشی از درونداد داده ها در روش هاپفیلد بود. مقایسه روش ترکیبی با روشهایBPANN و روش فازی ارزیابی، وروش شاخص نمرو (Nemerow index) نشان داد که مدل ترکیبی FHNN قضاوتی قابل اعتمادتر و اطلاعاتی با ارزش تر از سه مدل دیگر طبقه بندی کیفی آب به دست می دهد.

[DOR: 20.1001.1.16807073.2013.15.2.5.4]