Forecast of Water Quality along the Luanhe River Line Based on BP Neural Network

Hong Lin¹, Yuanbo Kang², Danyang Wang¹, Zeyu Lu¹, Wei Tian¹ and Shuqian WANG¹*

¹School of water conservancy and hydroelectric power, Hebei University of Engineering, Handan, Hebei, 056000, China
²Yellow River International Engineering Consulting (Henan) Co., Ltd. Zhengzhou, Henan, 450000, China
* Shuqian WANG’s e-mail: 320338462@qq.com

Abstract. In order to understand the changes of water quality in the water, the paper uses the improved BP network of LM algorithm to learn and train the data, and implements the neural network model on the MATLAB platform, and uses the processed samples for the established BP neural network. Learning training, in order to prevent some neurons from reaching the state of supersaturation. By normalizing the data, the trained network is used to predict the water quality indicators of the Luanhe River Line. The results show that it is feasible to evaluate the water quality along the sputum by BP neural network model. The model has strong learning, association and fault tolerance functions. The analysis results and process are close to the thinking process and analysis method of the human brain, which makes the water quality evaluation. The accuracy of the results is greatly improved.

1. Introduction

The project of introducing and sinking into Tianjin is a cross-basin large-scale water diversion project to solve urban water use in Tianjin, with a total length of 234 kilometers. The project was officially started on May 11, 1982, and was completed on September 11, 1983[1]. It completely ended the history of drinking salt water and bitter water in Tianjin residents. The quality of domestic water reached the standard of drinking water and became one of the best drinking water quality cities in the country[2]. As of September 2013, the project has been through water for 30 years, with a total of 15.9 billion m³ of water from the Weihe River and a safety water supply of 22.8 billion m³. With the rapid development of Tianjin’s economy and the development and opening up of Binhai New Area, the water supply scope of the water supply is expanding from the central urban area and the built-up area of Binhai New Area to the economic development hotspots such as Zhongxin Ecological City, Central Fishing Port and Nangang Industrial Park. It reached 2 million m³. Water quality pollution indicators along the drainage line are selected as examples to carry out water quality prediction, so that the water quality changes of the water diversion and drainage can be better understood to achieve reasonable control and early prevention of water pollution damage[3].

Artificial neural network is an algorithmic mathematical model that mimics the behavioral characteristics of animal neural networks and performs distributed parallel information processing[4]. This kind of network relies on the complexity of the system to adjust the relationship between a large number of internal nodes to achieve the purpose of processing information. Xie Wuming et al. used
BP neural network and genetic algorithm to predict the power consumption of sewage treatment plants, which has obvious advantages for energy-saving optimization. Wang Fuzhen used artificial neural network combined with SWAT model to predict runoff, and the prediction results were in good agreement. Li Zhixin and Lai Zhiqin used the BP neural network prediction model to study the annual runoff variation. The results show that the model has a good forecasting effect and more accurately expresses the mapping relationship between the annual runoff forecasting factor and the forecasting object. In addition, artificial neural network is also more common in mountain flood forecasting. Li Hongyan and others used the improved BP algorithm to make the artificial neural network flood forecasting model significantly improve the forecasting accuracy of flood peaks, thus ensuring the reliability of flood peak forecasting. In recent years, artificial neural network algorithms are also widely used in water quality prediction. This paper carries out sampling monitoring according to each monitoring section among 1987-2007. According to the pollution situation and reservoir function in the region, the main indicators for water quality analysis and monitoring are: permanganate index, biochemical oxygen demand on the 5th, ammonia nitrogen, dissolved oxygen.

2. Improvement of BP algorithm

The BP algorithm theory has the advantages of reliability, rigorous derivation process, high precision and good versatility. However, the standard BP algorithm has the following disadvantages: slow convergence speed; easy to fall into local minimum; difficult to determine hidden layer and hidden layer nodes Number. In practical applications, the BP algorithm is difficult to perform, so many improved algorithms have emerged. This paper intends to use L-M learning rules and momentum methods as the training and learning functions of neural networks.

2.1 Improving the BP Algorithm by Momentum Method

The standard BP algorithm is essentially a simple steepest descent static optimization method. When W(K) is corrected, it is only corrected according to the negative gradient direction of step K, without considering the accumulated experience, that is, the previous time. The direction of the gradient, which often causes the learning process to oscillate and converge slowly. The specific method of the momentum weight adjustment algorithm is to add a part of the last weight adjustment amount to the weight adjustment amount calculated according to the current error, as the actual weight adjustment amount of this time, namely:

\[ \Delta W(n) = -\eta \nabla E(n) + \alpha \Delta W(n-1) \]

Where: \( \alpha \) is a momentum coefficient, usually 0 < \( \alpha \) < 0.9; \( \eta \) - learning rate, ranging between 0.001 and 10. The momentum factor added by this method is actually equivalent to the damping term, which reduces the oscillation tendency during the learning process, thereby improving the convergence. The momentum method reduces the sensitivity of the network to the local details of the error surface, effectively suppressing the network from falling into local minima.

2.2 L-M learning rules

The L-M (Levenberg-Marquardt) algorithm is much faster than the aforementioned BP algorithms using the gradient descent method, but for complex problems, this method requires considerable storage space. The weight adjustment rate of the L-M (Levenberg-Marquardt) optimization method is selected as:

\[ \Delta W = (J^T J + \mu I)^{-1} \cdot J^T e \]

Where: \( e \)—error vector; \( J \)—Jacobian matrix of network error versus weight derivative; \( \mu \)—scalar, when \( \mu \) is large, the upper formula is close to the gradient method, and when \( \mu \) is small, the upper formula becomes Gauss-Newton method, in this method, \( \mu \) is also adaptively adjusted.
3. Training and results of BP neural network

A multi-layer feedforward BP network is implemented by means of the MATLAB neural network toolbox. The actual output value of the neural network is related to the input value and each weight and threshold. In order to match the actual output value with the expected output value of the network, the network is trained with a set of samples containing a certain number of learning samples and a corresponding set of expected output values. The measured sample data from 1987 to 2007 was used during training.

3.1 Model training

According to the available data, four pollution indicators such as permanganate index, five-day biochemical oxygen demand, ammonia nitrogen and dissolved oxygen were used as input samples of the network. The output samples of the network also use these four indicators as output samples of the network. The annual average of 1987-2007 was used as a sample set. Take 20% of the sample set for confirmation and 20% for the test set. Train the network for 20,000 times, the target error is $10^{-6}$, and the training function we use trainlm. After verification, the number of nodes in the hidden layer unit is the best. After the training is completed, the output is simulated with the sim function and compared with the target output to verify the performance of the network. The hidden layer and the output layer respectively use the "tansig" and "purelin" conversion functions, and the calculation ends when the error reaches the set learning performance target value ($10^{-6}$).

Since the network's transfer function (sigmoid function) has a large gradient in the [0,1] interval, the general network training transforms the sample data into this region, and in order to prevent some neurons from reaching the supersaturation state, that is, the data Perform normalization. The BP network structure diagram is shown in Figure 1.

This paper uses the neural network training toolbox provided by matlab to further analyze the network training results. The neural network training toolbox uses a linear regression method to analyze the relationship between the network output and the target output, that is, the rate of change of the output of the network relative to the target output, thereby evaluating the network training results. Figures a, b, c and d below represent the Yujiao Reservoir Station Center (a), Shaheqiao Station (b), Linhe Bridge Station (c) and Guoheqiao Station (d).

![Figure 1. BP network structure](image)
Figure 2. Training and detection curve of BP neural network model

Figure 3. Training state diagram
Figure 4. Correlation between predicted and measured values

Figure 5. Relative error graph of predicted value and actual value, order of magnitude 10
3.2 Analysis of training results

It can be seen from Figure 2 that the network constructed by Yuqiao Reservoir Station Center (a), Shaheqiao Station (b), Linhe Bridge Station (c) and Guoheqiao Station (d) are 163 steps, 96 steps, 2098 steps respectively. And 39 steps, the network error meets the learning requirements, and the learning rate is ideal. As can be seen from Fig. 4, R represents the correlation coefficient between the network output and the target output. The closer R is to 1, the closer the output is to the target output, and the better the network performance. Here R=1. It can be seen from Fig. 6 that the predicted value is basically consistent with the measured value, and the prediction effect is good.

The parameter values (annual average) along the enthalpy predicted according to this model are shown in Table 1.

| Station name            | years | Dissolved oxygen | Ammonia nitrogen | Permanganate index | Five-day biochemical oxygen demand |
|-------------------------|-------|------------------|------------------|--------------------|-------------------------------------|
| Yuqiao Reservoir Center | 2010  | 10.5             | 0.18             | 4.9                | 1.4                                 |
|                         | 2015  | 8.9              | 0.14             | 3.2                | 1.1                                 |
|                         | 2020  | 10.6             | 0.29             | 3                  | 1.6                                 |
| Shahe Bridge            | 2010  | 11.4             | 0.59             | 2.3                | 3.4                                 |
|                         | 2015  | 11.4             | 0.25             | 4.4                | 1.8                                 |
|                         | 2020  | 9.4              | 0.42             | 2.3                | 3.6                                 |
| Linhe Bridge            | 2010  | 8.2              | 0.36             | 2.2                | 1.1                                 |
|                         | 2015  | 11.7             | 0.4              | 1.3                | 3.4                                 |
|                         | 2020  | 9.4              | 0.4              | 1.5                | 1.3                                 |
| Guoheqiao Station      | 2010  | 10.8             | 0.33             | 2.5                | 4                                   |
|                         | 2015  | 11.4             | 0.36             | 2.6                | 1.6                                 |
4. Conclusion

- Based on the measured data along the enthalpy, the BP neural network improved by LM algorithm was used to predict the permanganate index, the five-day biochemical oxygen demand, ammonia nitrogen and dissolved oxygen. The prediction results show that the model has a good pan. The ability is stable and reliable, which can provide reference for the practical application of water quality prediction.
- BP neural network can accurately express the mapping relationship between annual runoff forecasting factor and object. The trainlm training function can further improve the generalization ability and forecasting accuracy of the model.
- Through programming calculations, MATLAB's powerful and efficient scientific computing function makes the prediction of water quality using neural network simpler in MATLAB.

Acknowledgment

Fund: Study on Plant Configuration and Water Purification Technology of Typical Wetland in the Lower Reaches of Fuyang River (1723209055-4)

References

[1] Yang Limei(2007) Evaluation of Tianjin Water Resources Allocation Scheme. Tianjin University.
[2] Zhang Hui(2007) Research on the optimal allocation of water resources in Tianjin. Tianjin University.
[3] Zhang Penghe, Wei Xiaoxi, Yang Zhengjian, et al(2017) Absorption of phytoplankton by bio-soils from water conservancy projects. Hydroelectric Power, 43(7): 1-4, 61.
[4] Lei Lei(2007) Application of artificial neural network in air pollution prediction. Beijing University of Technology.
[5] Xie Wuming, Li Jun, Zhou Fengping, et al(2018) Power Consumption Prediction of Wastewater Treatment Plant Based on BP Neural Network and Genetic Algorithm. Hydropower Science, 36(8): 202-204.
[6] Wang Fuzhen(2017) Application of artificial neural network combined with SWAT model in river runoff prediction[J]. Water Resources Development and Management, (8): 67-70.
[7] Li Zhixin, Lai Zhiqin(2018) Research on BP neural network prediction model for annual runoff variation. Hydropower Science, 36(7): 10-12.
[8] Li Hongyan, Liu Hanbing, Yuan Ximin, et al(2002) Artificial neural network peak identification theory and its application in flood forecasting. Journal of Hydraulic Engineering, (6): 15-20.
[9] Liang Nan(2007) Water quality prediction based on artificial neural network and MATLAB implementation. Chang’an University.
[10] Ni Shenhui, Bai Yuhui(2000) Application of BP Neural Network Model in Groundwater Quality Evaluation. Systems Engineering - Theory & Practice, (8): 124-127.