# A Novel BP Neural Network with Wavelet Transform Inputs for Air Quality Index Prediction

**Zheng Fang**, **Lei Zhang and Yakun Huang**

School of Mechanical & Electrical Engineering, Xuzhou University of Technology, No. 2, Lishui Road, Yunlong District, Xuzhou City, Jiangsu Province, China.
Email: triple-stone@foxmail.com

**Abstract.** Air quality index prediction research is of great significance to people's life. Various prediction methods are used to pursue its accurate prediction. A novel BP neural network with wavelet transform inputs for air quality index prediction is proposed in this paper. The proposed method introduces wavelet transform in the front of the neural network. It decomposes the original time series signal into a linear superposition of the main signal and some noise signals. After establishing the predictive BP neural network for each decomposed signal, the air quality index is restored based on the linear additivity of wavelet decomposition. The simulation experiments show that the prediction accuracy of proposed method is high. The relative error of long-term prediction for 50 consecutive days is only 4.3994% for Xuzhou City's air quality index. The research work provides a useful reference for air quality index prediction.

## 1. Introduction

Air Quality Index (AQI) is an important indicator for evaluating air quality. Research on its prediction has been a hotspot. Various methods are used to pursue AQI's accurate prediction. Some of them relies on mining AQI historical data itself to predict, such as time series [1], gray model [2], Markov Model [3], and so on. The others need more data support besides AQI for data training and learning, such as neural network [4], random forest [5], principal component analysis [6] support vector machine [7] machine learning [8], etc. If there is no other data support except AQI, these methods' forecasting accuracy will be greatly reduced. For example, using BP neural network to predict AQI, if there is no supplementary data on major pollutant indicators, generally speaking, its prediction accuracy is unsatisfied. However, this is not necessarily the case. Previously, the authors of this paper only used AQI time series as the input of neural network, and used GA-BP neural network to predict, which received some good results [9]. And this paper will continue the previous study, discussing the AQI prediction problem caused by using time series directly as the input of BP neural network. In order to improve the prediction accuracy, wavelet transform will be introduced. The proposed method of this paper is shown in figure 1.
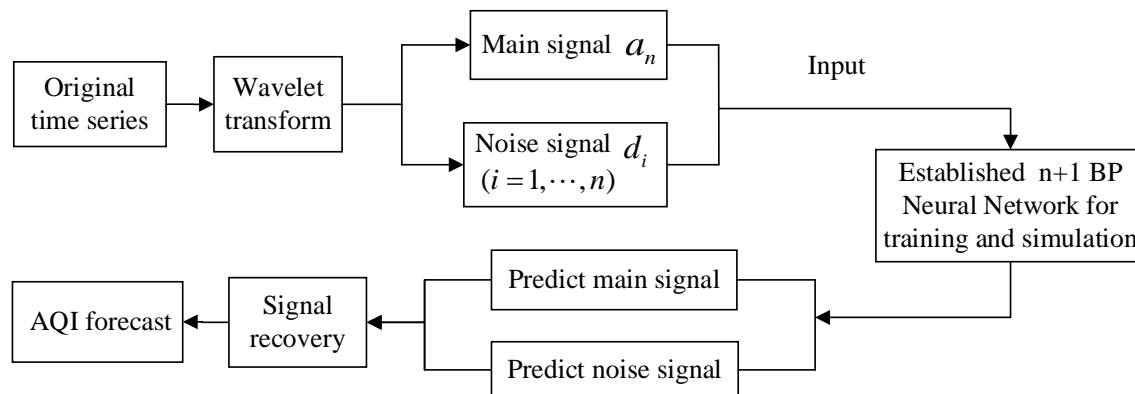
**Figure 1.** Research Route

As shown in figure 1, the proposed method can be described as follows: Firstly, the original time series which is regarded as unstable input signal is decomposed into a linear combination of the main signal $a_n$ and some noise signals $d_i(i=1, \ldots, n)$ by wavelet transform. Then, using $a_n$ and $d_i$ as sample training database, $n+1$ BP neural networks are established to predict main signal and noise signals respectively. Finally, the AQI time series is restored according to the linear combination, and its value is the predicted value. The proposed method does rely on AQI data itself without any auxiliary data, and it is different from the existing method of AQI prediction using wavelet neural network [10-13].

In figure 1, several issues need to be discussed deeply. The remainder of this paper will clarify them in four sections. Section 2 shows the selection of wavelet function and the results of wavelet decomposition. Section 3 introduces the establishment of BP network, and a simulation case study is discussed in section 4. Finally, a summary is provided in section 5. The AQI data shown in figure 2 in this paper is from December 2013 to November 2018 in Xuzhou City, Jiangsu Province, China. A total of 1814 AQI data are obtained from the website of Jiangsu Meteorological Bureau.
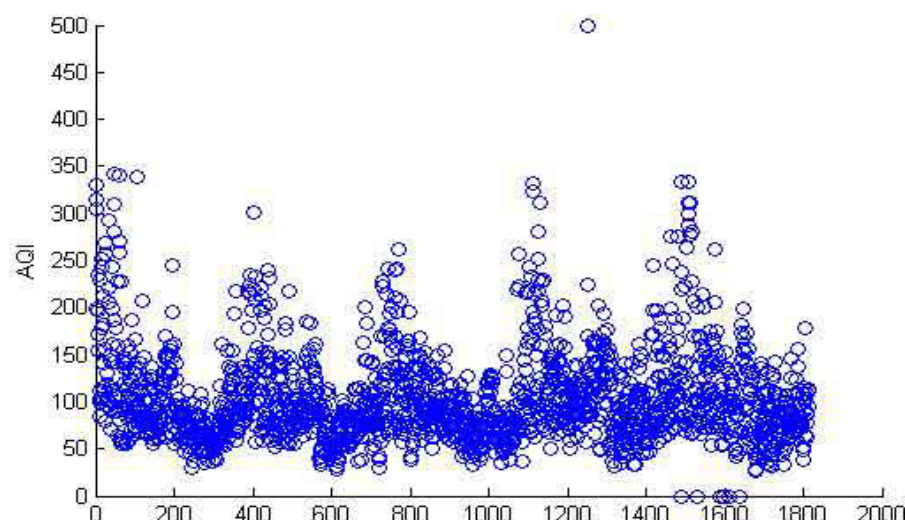


**Figure 2.** AQI Data of Xuzhou from December 2013 to November 2018

## 2. Wavelet Transform
The original AQI time series shown in figure 2 could be regarded as a group of unstable signals. Following the wavelet transform theory, it could be decomposed as a linear superposition of the main

signal and some noise signals. The decomposition process could be illustrated in figure 3.
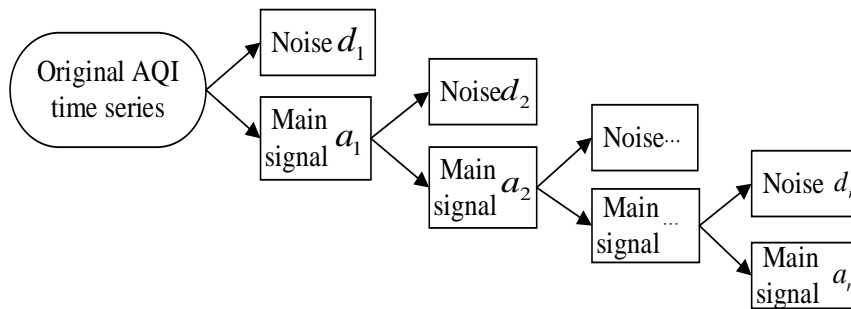


**Figure 3.** Wavelet Decomposition Schematic

From figure 3, the original AQI time series by wavelet decomposition can be expressed as follows,

$$\text{AQI} = a_n + \sum_{i=1}^{n} d_i \text{ ,} \qquad (1)$$

where $a_n$ is the main signal and $d_i$ ($i = 1, …, n$) are noise signals, i.e., AQI is a linear superposition of the $a_n$ and $d_i$ ($i = 1, …, n$).

In order to decompose AQI into linear combinations shown in equation (1), the selection of wavelet functions is a very important work. Different wavelet functions will result in different decomposition results which will have a great impact on BP neural network prediction. However, for the selection of wavelet functions, there are still no unified selection principles today. We used the common wavelet functions to make lots of tests in Matlab 2014a based on data from figure 2. The following results are obtained.

For *Haar* wavelet function, it is discontinuous and has no smoothness. When it is used to decompose the signal of figure 2, the decomposition signal is seriously distorted. The *dbN* wavelet function is asymmetric. When it is used to decompose the signal of figure2, the signal distortion is also very obvious, especially in the edge. Therefore, *Haar* and *dbN* cannot be used as wavelet function in this situation. Although *coifN* wavelet function has good symmetry, its performance decreases with the increase of its support length, and the computational cost decreases with the increase of its computational efficiency. However, it is inconsistent with the noise signal integrity in our proposed model. Hence, it is not suitable either. *Morl* function and *Mexh* function are not taken into consideration due to their insufficiency in compact support, scale function and orthogonality. Compared with *dbN*, *symN* function has the advantages of approximate symmetry and compact orthogonality. It is consistent with our application. Therefore, we use *symN* wavelet function to perform wavelet decomposition.

On the choice of *N* value, the symmetry of *symN* is not good as *N*<4. When *N*>4, the function distribution is too dense to decompose. Hence, *N*=4 is proposed. For the layer of wavelet decomposition, the tests are carried out in Matlab 2014a. When layer s < 7, the main signal's decomposition is insufficient. When layer s > 7, the main signal is over-decomposed and becomes a square wave. Therefore, the decomposition layer s = 7 is determined. After above analyses and tests, the original signal is decomposed by *sym*4 wavelet function. The decomposed results are shown in figure 4.
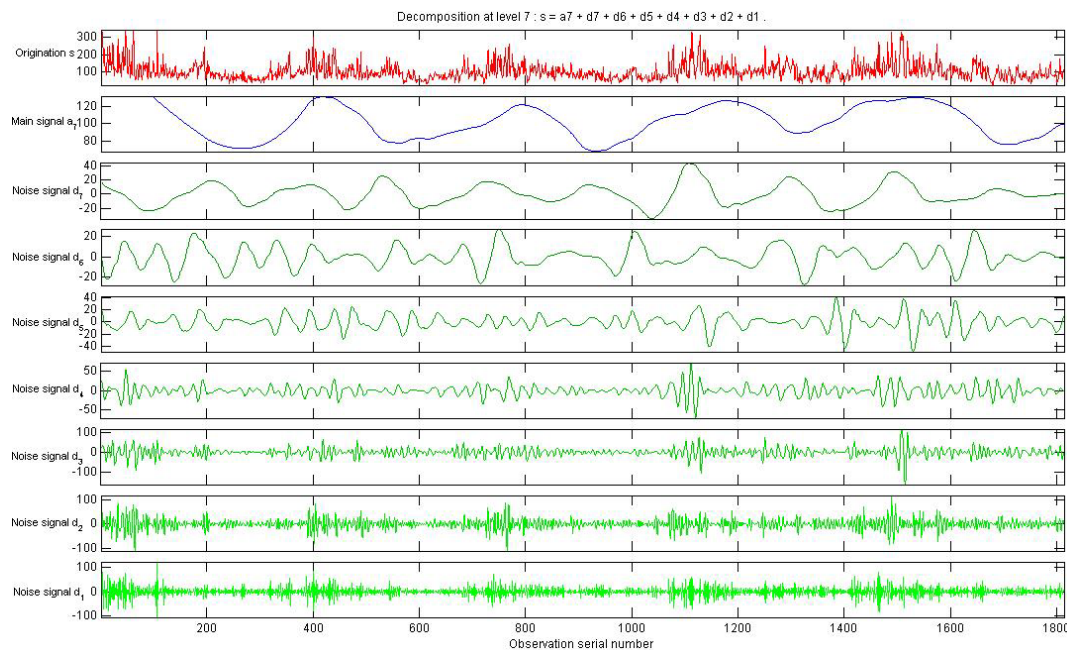
**Figure 4.** AQI Wavelet Decomposition Signal Diagram

After decomposition, the modeling work could be started. It will be demonstrated in the next section.

## 3. Establishment of BP Neural Network

### 3.1 Determination of the Number of Input Nodes in BP Neural Network

From figure 4, one group of main signals and seven groups of noise signals are obtained. Hence, eight BP neural networks are needed to be established according to figure 1. Since the data itself is used as input, it is crucial to determine the number of input nodes for each BP neural sub-network.

In this paper, the optimal number of input nodes is confirmed by calculating Mean Square Error (MSE) value and the Correlation Test (CT) value of test samples. Taking the noise signal $d_1$ in figure 4 as an example, calculating MSE and CT values of 13 input nodes, respectively in Matlab 2014a (Here, 13 is an approximate number, and more input nodes could be taken into account). The results are shown in the following table.

**Table 1.** Test Results of Different Input Nodes of Noise Signals $d_1$

| The number of input node | MSE | CT | The number of input node | MSE | CT |
|:---:|:---:|:---:|:---:|:---:|:---:|
| 2 | 0.104 | 0.81366 | 8 | 0.000998 | 0.98355 |
| 3 | 0.0036 | 0.93959 | 9 | 0.00096 | 0.98436 |
| 4 | 0.0016 | 0.97362 | 10 | 0.00092 | 0.98497 |
| 5 | 0.0013 | 0.97926 | 11 | 0.000994 | 0.98362 |
| 6 | 0.000992 | 0.98364 | 12 | 0.000998 | 0.98347 |
| 7 | 0.000985 | 0.98389 | 13 | 0.000994 | 0.98355 |

From table 1, When the number of input node is 10, MSE is the smallest and CT is the largest among 13 input nodes. This shows that at this time the performance of the BP neural network is better. Hence, 10 is determined to be optimal for $d_1$. That is to say, 10 consecutive AQI values are input into the BP neural network of $d_1$, and the 11st AQI value is used as label data. By analogy, the BP neural network of $d_1$ has 1804 teams of data. Following this method, the number of input nodes for $d_i(i=2,\ldots,7)$ and $a_7$ is determined as shown in table 2.

**Table 2.** The Number of Input Nodes for each Signal BP Neural Network

| Signal Sequence | Number of input nodes | Data teams | Signal Sequence | Number of input nodes | Data teams |
|---|---|---|---|---|---|
| $a_7$ | 3 | 1811 | $d_4$ | 5 | 1809 |
| $d_1$ | 10 | 1804 | $d_5$ | 5 | 1809 |
| $d_2$ | 10 | 1804 | $d_6$ | 5 | 1809 |
| $d_3$ | 5 | 1809 | $d_7$ | 5 | 1809 |

*3.2 Determination of the Number of Hidden Layer Nodes*
At present, for BP neural network, the determination of the number of hidden layer nodes usually depends on empirical formula. Based on empirical formula and MSE of network loss function which is provided by computer under setting different hidden layer nodes, the hidden layer nodes for each BP neutral sub-network are determined as follows.

**Table 3.** Hidden Layer Node Settings

| Signal Sequence | Number of Hidden Layer Nodes | Signal Sequence | Number of Hidden Layer Nodes |
|---|---|---|---|
| $a_7$ | 35 | $d_4$ | 50 |
| $d_1$ | 30 | $d_5$ | 50 |
| $d_2$ | 30 | $d_6$ | 50 |
| $d_3$ | 35 | $d_7$ | 50 |

After determining the number of input nodes and hidden layer nodes, BP neural network could be constructed by the neural network toolbox of Matlab 2014a. In next section, the simulation results of figure 2 using the proposed method will be given and discussed.

**4. Simulation Case Study**

*4.1 Simulation Tests of Network Performance*
Eight BP neural sub-networks were established respectively by computer. The default training accuracy is 0 and the default training function is *trainlm* in the neural network toolbox of Matlab 2014a. The first 1764 data of 1814 data were used as training set and the last 50 data were used as test set. When the simulation results of eight sub-networks are obtained, the AQI simulation results of the whole system can be calculated according to formula (1). The root mean square error (RMSE) between the predicted value and the real value of the test set and 50-day prediction accuracy were calculated for eight sub-networks and whole system, respectively. They are shown as table 4.
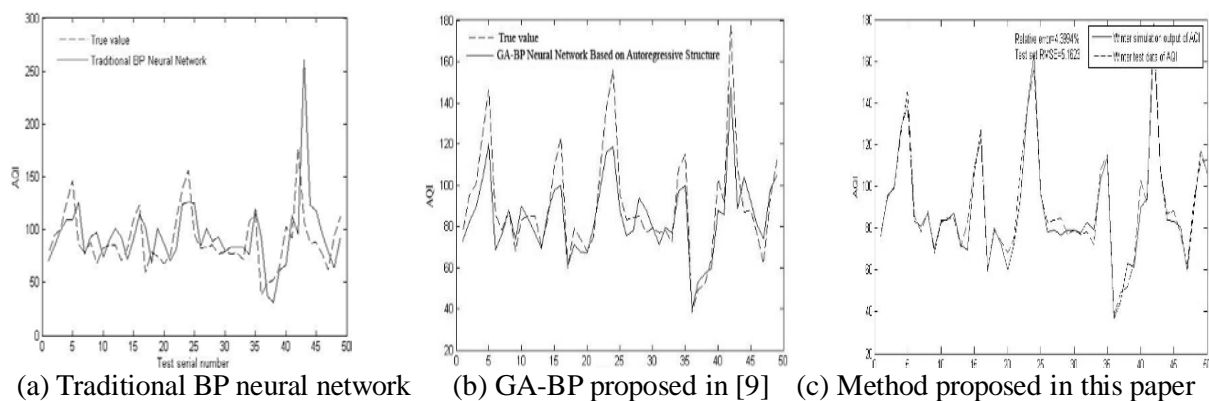
**Table 4.** Prediction Error of each Group of Signals by Simulation

| Signal sequence | RMSE | 50-day prediction accuracy | System MSE |
|---|---|---|---|
| $a_7$ | 0.1302 | 0.1301% | 0.0007784 |
| $d_1$ | 3.3265 | 4.8766% | 0.0009819 |
| $d_2$ | 3.1169 | 5.2634% | 0.0009860 |
| $d_3$ | 3.1013 | 9.8701% | 0.0009815 |
| $d_4$ | 1.2647 | 6.1459% | 0.0006421 |
| $d_5$ | 0.2556 | 4.5628% | 0.0001189 |
| $d_6$ | 0.0993 | 3.5435% | 0.0001106 |
| $d_7$ | 0.0406 | 2.3040% | 0.0003459 |
| Predictive value of AQI simulation | 5.1623 | 4.3994% | 0.0006186 |

The first eight rows of table 4 are the simulation test performance of eight sub-networks. What the last row of table 4 shows is the whole system's performance. From table 4, the RMSE values of eight sub-networks are about 3 or less, whereas the whole system's is 5.2, which indicates that the predicted values deviate very small from the real values. The whole system's 50-day prediction accuracy is about 4.4% and MSE value is almost 0, which illustrate that the restored prediction system has a good accuracy.

### 4.2 Comparison of Simulation Prediction

In order to demonstrate the superiority of our proposed method, the traditional BP neural network and GA-BP neural network proposed in reference [9] are input into the time series shown in figure 2. The predictive results are shown in figure 5 and the predictive performances are shown in table 5.



(a) Traditional BP neural network      (b) GA-BP proposed in [9]    (c) Method proposed in this paper

**Figure 5.** Predictive Contrast Diagram

**Table 5.** Performance Comparison of three Prediction Methods

| Prediction System | System MSE | Prediction Relative Error | Test Set RMSE |
|---|---|---|---|
| Traditional BP neural network | 0.0621 | 36.6628% | 42.3130 |
| GA-BP neural network proposed in reference [9] | 0.0462 | 9.1375% | 11.4722 |
| BP neural network established in this paper | 0.0006186 | 4.3994% | 5.1623 |

From figure 5 and table 5, it is obvious that the method proposed in this paper has best accuracy for AQI prediction.

## 5. Conclusion

The BP neural network with wavelet transform inputs is an improvement to the traditional BP neural network. It is suitable for the situation where time autoregressive data are input into BP neural network. In future, we will conduct a comparative study between this network and time series model to increase the understanding of this network.

## 6. Acknowlegement

## 7. Reference

[1]     Li Weidong, Li Li, Xu Yan. The Concentration Research of PM2.5 in Beijing with Time Series Analysis. Operational Research Transactions, 2018, 22 (2):115-126.
[2]     Zhao Meiling, Xue Rui. Application of Grey Theory in Ambient Air Quality Trend Analysis. Environment and Development, 2013, 29 (2): 76-79.
[3]     Yang Jinwei, Sun Baolei. The Prediction of Air Pollutant Concentration in Pingdingshan Based on Grey Markov Model. Practice and Recognition of Mathematics, 2014, 44(2): 64-70.
[4]     Du Xiaoliang. Application of Artificial Neural Network in Air Quality Prediction. Shandong University, 2009.
[5]     Lin Kaichun, Shao Fengjing. Research on Air Quality Prediction Based on Random Forest and Neural Network. Journal of Qingdao University (Engineering Edition), 2018, 33 (02): 32-36.
[6]     Wei Zhen, Lin Fang. Air Quality Evaluation of Cities Based on Cluster Analysis and Principal Component Analysis. Journal of Huaiyin Institute of Technology, 2018, 27 (03):86-96.
[7]     Guo Fei, Xie Liyong. Air Quality Index Prediction based on Meteorological Factors and Improved Support Vector Machine. Environmental Engineering, 2017, 35 (10): 151-155.
[8]     Yin Qi. Air Quality Classification and Prediction based on Machine Learning. Zhongbei University, 2018.
[9]     Fang Zheng, Zhang Lei, Wang Yuqin, et al. AQI prediction based on Autoregressive GA-BP neural network. Environmental Monitoring and Warning, 2019, 11 (02): 22-25.
[10]    Wang Xiaohan, Zhang Yanping, Zhao Shu, et al. Air quality prediction based on dynamic granularity wavelet neural network. Computer Engineering and Application, 2013, 49 (6): 221-224.
[11]    Dunea D, Pohoata A, Iordache S. Using wavelet–feedforward neural networks to improve air pollution forecasting in urban environments. Environmental Monitoring and Assessment, 2015, 187(7): 477-493.
[12]    Bai Y, Li Y, Wang X, et al. Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions. Atmospheric Pollution Research, 2016,7 (6): 557-566.
[13]    Morabito F C, Versaci M. Wavelet neural network processing of urban air pollution. International Joint Conference on Neural Networks. IEEE, 2002.