ICIC Express Letters
Volume **10**, Number **3**, March **2016**

ICIC International ©2016 ISSN 1881-803X
pp. 555–562

AN IMPROVED GREY NEURAL NETWORK FOR URBAN AIR QUALITY FORECAST

Yuan Huang, Guoyan Huang, Wanchang Jiang and Bing Zhang

College of Information Science and Engineering
Yanshan University
No. 438, Hebei Avenue, Qinhuangdao 066004, P. R. China
757918272@qq.com

Received September 2015; accepted December 2015

Abstract. *Information about urban air quality is greatly important to protect human health and policy making of governments. In view of the fact that the urban air quality forecast is non-linear with data insufficiency and data fluctuation, an improved grey neural network model called IGNN model is proposed to improve the forecast accuracy of urban air quality. The background value of grey model and the data sample of neural network are optimized at the same time based on the analysis of the factors that can influence the forecast accuracy of grey neural network model. The experiment results show that IGNN model can be employed in air quality forecast with its less relative error and higher reliability, compared with grey model and traditional grey neural network model.*

Keywords: Air quality, Grey neural network, Optimizing, Forecast

1. **Introduction.** The continuous aggravation of urban air pollution has been a serious threat to human health. People are increasingly concerned with the information of air pollution, such as the concentrations of PM_{10} and SO_2. Researches and analysis on the current situation and change law of urban air quality, and scientific and reasonable forecast for the changing trend of air quality in the future have important guiding function to the evaluation, management and decision-making of urban air quality.

The numerical model and the statistical model are mainly adopted to forecast urban air quality [1, 2]. The numerical model applies the data obtained by the monitoring station to forecasting urban air quality. However, the monitoring station is expensive to construct and complex to operate. Besides, detailed data on pollutant emission inventories should be provided for the numerical model, and data acquisition is very difficult. Urban air quality forecast based on the statistical model requires relatively little detailed data, and it is inexpensive and easy to operate [3, 4, 5]. The statistical forecast models include artificial neural network models and grey models and so on. Baawain and Al-Serihi [6] proposed a rigorous method of preparing air quality data to achieve more accurate air pollution prediction models based on an artificial neural network. Pai et al. [7] utilized seven types of $GM(1,1)$ model to forecast hourly PM_{10} and $PM_{2.5}$ concentrations in Banciao city of Taiwan.

However, air quality forecast has the characteristics of data insufficiency, data fluctuation and nonlinearity. As for data fluctuation, fitting effect of grey model is poor, and only the latest several data have certain practical significance and the forecast precision. The nonlinear mapping of neural networks can make up for the defects. The neural networks are prone to in-sample over fitting, low convergence speed, and need a large number of measured data to track the change of the data for obtaining better forecast results. The grey model can use fewer samples for modeling and forecast, thereby making up for the deficiency of the neural network method. So the grey model and the neural network model are combined, which can increase the reliability and stability of the forecast results, avoid the information loss of single model and reduce the randomness. Lang et al. [8] used an organic grey neural network model to predict air pollution index. However, the grey neural network is rarely applied in researches on air quality forecast.

The Traditional Grey Neural Network (TGNN) model can be used to forecast air quality based on the combination of the advantages of the grey model and the artificial neural network model. However, the TGNN model has the disadvantages, such as improper background value setting of the grey model exerting influence on the forecast accuracy, and the insufficient data sample of the neural network model. The result lies in the reduction of the forecast accuracy and unstable forecast results. In order to improve the forecast accuracy, an improved grey neural network model called IGNN model is proposed in this paper, with the background value of grey model and the data samples of neural network optimized at the same time. The experimental results show that the IGNN model has higher forecast accuracy.

The rest of paper is organized as follows. Section 2 introduces the traditional grey neural network forecast model of air quality. An improved grey neutral network model called IGNN model is given in Section 3. Experimental results and analysis are given in Section 4. Section 5 concludes the paper.

2. **The Traditional Grey Neural Network Model.** Though grey neural network forecast has many kinds of combination method, the traditional serial grey neural network is chosen for its simplicity and generality. The traditional serial grey neural network includes two parts: GM$(1,1)$ model and Back-Propagation (BP) neural network model. The air pollutant concentration is used as the original sequence to establish GM(1,1) model and get the corresponding fitting values and forecast values. The fitting values and forecast values are used as the data samples of BP neural network and final air pollutant forecast concentrations are obtained through the trained neural network.

2.1. **GM(1,1) model.** The traditional GM(1,1) model is shown as follows:

Step 1: Denote the original sequence as:

$$X^{(0)} = \left\{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(N)\right\} \tag{1}$$

The sequence $x^{(0)}$ for 1-AGO (Accumulated Generating Operation) can be:

$$X^{(1)} = \left\{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(N)\right\} \tag{2}$$

where $x^{(1)}(k) = \sum\limits_{i=1}^{k} x^{(0)}(i)$, $k = 1, 2, \cdots, N$.

Step 2: Build up constant differential equation:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \tag{3}$$

where a and u are respectively called the developing coefficient and the control variable.

The Ordinary Least Square (OLS) method is utilized to calculate a and u:

$$\hat{U} = [a \quad u]^T = (B^TB)^{-1}B^TY \tag{4}$$

$$B = \begin{bmatrix} -z_1^{(1)}(2) & 1 \\ -z_1^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z_1^{(1)}(N) & 1 \end{bmatrix} \quad Y = \left[x_1^{(0)}(2)\ x_1^{(0)}(3) \cdots x_1^{(0)}(N)\right]^T \tag{5}$$

where $z^{(1)}(k+1)$ is called background value

$$z^{(1)}(k+1) = \frac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k+1)\right), \quad k = 1, 2, \cdots, N-1 \tag{6}$$

Step 3: Obtain the discrete form of first-order grey differential equation. The solution of $x^{(1)}$ is:

$$\hat{x}^{(1)}(k+1) = \left[x^{(1)}(1) - \frac{u}{a}\right] e^{-ak} + \frac{u}{a} \tag{7}$$

Thus, the forecast values can be derived by one inverse accumulated generating operation $(1 - IAGO)$:

$$\hat{x}^{(0)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) = (1 - e^{a}) \left(x^{(0)}(1) - \frac{u}{a}\right) e^{-ak} \tag{8}$$

2.2. **BP neural network model.** The Back-Propagation (BP) neural network model is a method for training the weights in a multi-layer feed-forward neural network. The algorithm can be decomposed in the following steps: feed-forward computation, back propagation to the output layer, back propagation to the hidden layer and weight updates.

In the grey neural network model, the fitting values of GM(1,1) model are used as the input sample of the BP neural network, and the corresponding real values as the output sample. The appropriate network structure is used for training to achieve the expected error target, and then the forecast values of GM(1,1) as test sample are input into the trained neural network for simulation and forecast. The BP neural network is formally defined as follows:

$$O = \psi\left(\sum_{i=1}^{q} \omega_{ki}\phi\left(\sum_{j=1}^{M} \omega_{ij} + \theta_i\right) + a\right) \tag{9}$$

where O denotes the outputs of output layer, $\psi(x)$ is the activation function of output layer, ω_{ki} denotes the weight between the k-th note of output layer and the i-th note of input layer, $i = 1, 2, \cdots, q$, $\phi(x)$ is the activation function of hidden layers, ω_{ij} denotes the weight between the i-th note of hidden layer and the j-th note of input layer, $j = 1, 2, \cdots, M$, θ_i denotes the threshold of the i-th note of hidden layers, and a denotes the threshold of output layer.

3. **The Improved Grey Neural Network.** In grey neural network forecast model, the selection of background value of GM$(1,1)$ model affects the forecast accuracy of grey neural network model. The fitting values and forecast values of GM$(1,1)$ model as the data sample of BP neural network are insufficient, which leads to unstable final forecast results of grey neural network, and unsatisfactory forecast effect. So the background value of grey model and the data samples of neural network are optimized at the same time to improve the forecast accuracy of grey neural network model.

3.1. **The improvement of the background value.** According to Equation (7), the fitting accuracy and forecast accuracy of GM$(1,1)$ model depend on parameters a and u. And the values of a and u depend on the construction of the background value, so the background value $z^{(1)}(k+1)$ becomes the key to directly affecting the fitting and forecast values of GM$(1,1)$ model.

Figure 1 shows that $z^{(1)}(k+1)$ is obtained by solving definite integrals of both sides of Equation (3) in the range $[k, k+1]$, namely the area of trapezoid $ABCD$. However, the actual area of the corresponding section of the curve $x^{(1)}(t)$ is always less than the area of trapezoidal $ABCD$, and there is area difference ΔS in the range $[k, k+1]$. When the sequence data changes gently, the difference between $x^{(1)}(k)$ and $x^{(1)}(k+1)$ is very small and ΔS is also small. However, when the change of sequence data is intensive, the difference between $x^{(1)}(k)$ and $x^{(1)}(k+1)$ becomes larger and the ΔS also increases, thus leading to the great reduce of forecast accuracy.

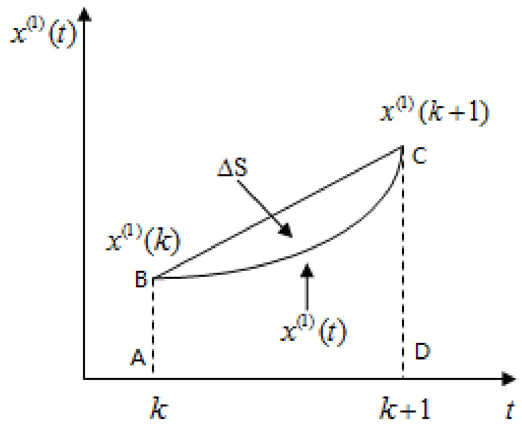

FIGURE 1. Principle diagram of the background value

According to the analysis on the impact factors of background values, a novel background value is given as Equation (10) to improve forecast accuracy.

$$z^{(1)}(k+1) = \frac{n+3}{2n}x^{(1)}(k) + \frac{n-3}{2n}x^{(1)}(k+1) \tag{10}$$

where n is the number of elements in the original sequence, $n = 1, 2, 3, \cdots$.

3.2. **The improvement of data sample of BP neural network.** In order to get better forecast results, a large amount of data is required to train the BP neural network, but fewer data samples of the BP neural network in the grey neural network model affect the final forecast accuracy. So the forms of input, output and test data sample of BP neural network are improved to increase the number of data samples. The detailed steps as follows.

Firstly, the fitting values $\hat{x}^{(0)}(k)$ of GM(1,1) are used as input data of BP neural network, the forecast values $\hat{x}^{(0)\prime}(k)$ as test data and the real values of air pollutant concentration $X^{(0)} = x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(N)$ as output data.

Secondly, $\hat{x}^{(0)}(k), \hat{x}^{(0)}(k+1), \cdots, \hat{x}^{(0)}(k+m-1)$ is used as the k-th column of input data, and the corresponding real values $x^{(0)}(k+m-1)$ as the k-th data of output data, and among them $k+m-1 \le N$, $k = 1, 2, \cdots, N-m+1$. The constructed new forms of input data and output data are expressed as follows:

$$I = \begin{bmatrix} \hat{x}^{(0)}(1) & \cdots & \hat{x}^{(0)}(k) & \cdots \\ \hat{x}^{(0)}(2) & \cdots & \hat{x}^{(0)}(k+1) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \hat{x}^{(0)}(m) & \cdots & \hat{x}^{(0)}(k+m-1) & \cdots \end{bmatrix} \quad O = [x^{(0)}(k) \cdots x^{(0)}(k+m-1) \cdots]$$

Finally, the number of elements predicted in the grey model is used as the number of columns in the matrix of test data. The number of rows is m, and the form of test data is constructed according to the constructed method of input data, denoted as T.

3.3. **The forecast process of the IGNN model.** The forecast process of the IGNN model is presented in Algorithm 1.

Step 1: The annual concentration of an air pollutant is selected as the original sequence of grey model. The background value is optimized by using Equation (10) to replace Equation (6) to improve the forecast accuracy of grey model in grey neural network model and get the fitting values and forecast values.

Step 2: The fitting values are used as input data, the forecast values as test data and the real values as output data. According to the introduction of Section 3.2, the new input, output and test data of BP neural network are respectively constructed.

Algorithm 1: CRF

Input: The annual concentration of an air pollutant, $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(N)\}$

Output: The forecast values of air pollutant concentration, $y^{(0)} = \{y^{(0)}(1), y^{(0)}(2), \cdots, y^{(0)}(N)\}$

1. Do
2. Define the novel background value
3. Apply the improved GM$(1,1)$ model to generating fitting values and forecast values
4. Use the fitting values as input data, the forecast values as test data and the real values as output data. Construct the new forms of input, output and test data
5. Input the new input data, test data and output data into the BP neural network
6. Return $y^{(0)} = \{y^{(0)}(1), y^{(0)}(2), \cdots, y^{(0)}(N)\}$

Step 3: The improved BP neural network is trained with the new input and output data to achieve the expected error objective. Then the new test data are input into trained BP neural network to forecast the air pollutant concentrations.

Thus, the background value of grey model and the data sample of BP neural network model are optimized at the same time in the IGNN model to improve the forecast accuracy.

4. Experiments.

4.1. **Datasets.** The air pollutants impacting air quality in Beijing are mainly PM_{10} and SO_2. Therefore, we respectively forecast the annual concentrations of PM_{10} and SO_2 of Beijing based on the environment quality bulletin from 2004 to 2014 issued by Beijing Municipal Environmental Protection Bureau. The datasets are detailed in Table 1.

TABLE 1. Annual concentrations of PM_{10} and SO_2 of Beijing (2004-2014)

	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
PM_{10} (ug/m^3)	149	142	151	148	122	121	121	114	109	108	116
SO_2 (ug/m^3)	55	50	53	47	36	34	32	28	28	27	22

4.2. **The establishment of IGNN model.** We take PM_{10} as an example to introduce the modeling process of the improved gray neural network model, and the modeling process of SO_2 is the same as that of PM_{10}.

Step 1: The annual concentrations of PM_{10} from 2004 to 2011 are used as the original sequence to establish the improved GM$(1,1)$ model. The annual concentrations of PM_{10} from 2012 to 2014 are forecasted by the improved grey model. We can get the fitting values and forecast values.

Step 2: The input data sample, output data sample and test data samples are constructed. After repeated comparison, we set m as 4. We get 4×5 input sample matrix I (Made up of fitting values of PM_{10} from 2004 to 2011), 1×5 output sample matrix O (Made up of real values of PM_{10} from 2007 to 2011) and 4×3 test sample matrix T (Made up of fitting values of PM_{10} from 2009 to 2011 and forecast values of PM_{10} from 2012 to 2014). The detailed data samples as follows:

$$I = \begin{bmatrix} 149 & 150 & 143 & 137 & 131 \\ 150 & 143 & 137 & 131 & 125 \\ 143 & 137 & 131 & 125 & 120 \\ 137 & 131 & 125 & 120 & 114 \end{bmatrix} \quad O = \begin{bmatrix} 148 & 122 & 121 & 121 & 114 \end{bmatrix}$$

$$T = \begin{bmatrix} 125 & 120 & 144 \\ 120 & 114 & 109 \\ 114 & 109 & 104 \\ 109 & 104 & 99 \end{bmatrix}$$

Step 3: The improved BP neural network is trained with input data I and output data O to achieve the expected error objective. Then the test data T are input into trained BP neural network to forecast the annual concentrations of PM_{10} from 2012 to 2014.

4.3. **The forecast results and analysis.** With improved grey neural network used to forecast the air pollution concentrations in Beijing, the BP neural network adopts Levenberg-Marquardt algorithm. We set a linear function (purelin) for the neurons in the input layer and a sigmoid function for those in the hidden and output layers. The forecast results of Improved Grey Neural Network (IGNN) model are respectively compared with those of the Traditional Grey Neural Network (TGNN) model and GM$(1,1)$ model. The simulation forecast results and relative errors of three kinds of air quality forecast models are respectively shown in Table 2 and Table 3.

TABLE 2. Forecast results and relative error of PM_{10} (ug/m^3)

Year	Real value	PM_{10}					
		GM$(1,1)$		TGNN$(1,1)$		IGNN$(1,1)$	
		Forecast value	Relative error	Forecast value	Relative error	Forecast value	Relative error
2012	109	109	0	112	2.8%	109	0
2013	108	104	3.7%	112	3.7%	109	0.9%
2014	116	100	13.8%	110	5.2%	112	3.4%

TABLE 3. Forecast results and relative error of SO_2 (ug/m^3)

Year	Real value	SO_2					
		GM$(1,1)$		TGNN$(1,1)$		IGNN$(1,1)$	
		Forecast value	Relative error	Forecast value	Relative error	Forecast value	Relative error
2012	28	25	10.7%	29	3.6%	27	3.6%
2013	27	23	14.8%	28	3.7%	25	7.4%
2014	22	21	4.5%	26	18.2%	22	0

Table 2 and Table 3 present that the forecast accuracy of the Improved Grey Neural Network (IGNN) is significantly improved compared with that of GM$(1,1)$ and the Traditional Grey Neural Network (TGNN) model. For PM_{10}, the average relative error of the TGNN model is 3.9% and that of IGNN model is only 1.4%. For SO_2, the average relative error of the TGNN model is 8.5% and that of IGNN model is 3.7%. The forecast accuracy of IGNN model improves by 4.8%, compared with that of the TGNN model. So using the IGNN model to forecast the air quality can significantly improve the forecast accuracy.

The improved grey neural network has higher prediction accuracy, which is suitable for the urban air quality forecast. The IGNN model is applied to forecasting the annual concentrations of PM_{10} and SO_2 of Beijing from 2015 to 2019.

Figure 2 shows the annual concentrations of PM_{10} and SO_2 are the gradual downward trend in the next few years. In 2019, the concentration of PM_{10} is 100ug/m^3 and the concentration of SO_2 is 16ug/m^3. In comparison with the concentrations of PM_{10} and

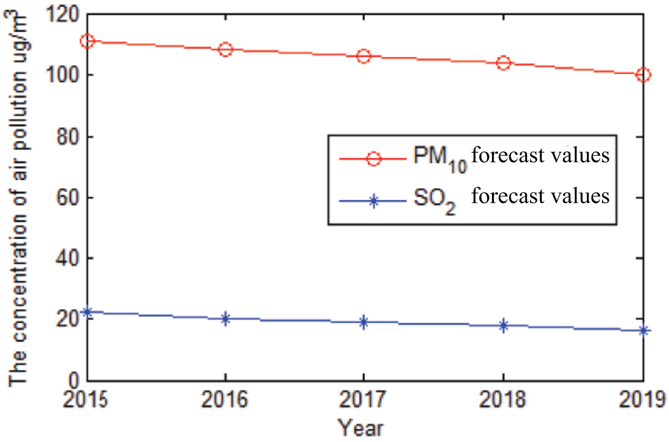

Figure 2. Forecast results of the air pollutant concentrations

SO_2 in 2014, the concentrations of PM_{10} and SO_2 in 2019 decline with respective descent rate of 13.8% and 27.3%. The concentrations of PM_{10} far exceed the secondary annual concentration thresholds of PM_{10} of the ambient air quality standard of China (70ug/m^3), which suggests that although the air quality of Beijing has improved and the prevention and control of air pollution treatment measures have achieved a certain effect, more governance is required for PM_{10} pollution.

5. **Conclusions.** In order to improve the urban air quality forecast accuracy, an Improved Grey Neural Network (IGNN) model is proposed for forecasting urban air quality on the basis of the simultaneous optimizations of the background value of grey model and the data samples of neural network, and takes Beijing as an example for experiments and analysis. Through the comparison of the forecast results of PM_{10} and SO_2 concentrations, the average relative error of IGNN model is respectively 6.3% lower than GM$(1,1)$ and 3.1% lower than TGNN model. With ideal forecast effect, IGNN model meets the requirements for air quality forecast. The air quality of Beijing before 2019 is forecasted by IGNN model, the forecast results of which show that by 2019 the overall air quality of Beijing will have been improved. In the future, we would like to apply IGNN model to more cities, further improve the forecast accuracy and study the root causes of air pollution.

Acknowledgment. This work is supported by the National Natural Science Foundation of China under Grant No. 61170190 and No. 61472341, and the Natural Science Foundation of Hebei Province P. R. China under Grant No. F2012203062, No. F2013203324 and No. F2014203152. Authors also gratefully acknowledge the helpful comments and suggestions of reviewers, which have improved the presentation.

REFERENCES

[1] Y. Zhang, M. Bocquet, V. Mallet et al., Real-time air quality forecasting, part I: History, techniques, and current status, *Atmospheric Environment*, vol.60, pp.632-655, 2012.

[2] Y. Zhang, M. Bocquet, V. Mallet et al., Real-time air quality forecasting, part II: State of the science, current research needs, and future prospects, *Atmospheric Environment*, vol.60, pp.656-676, 2012.

[3] W. G. Cobourn, An enhanced $PM_{2.5}$ air quality forecast model based on nonlinear regression and back-trajectory concentrations, *Atmospheric Environment*, vol.44, pp.3015-3023, 2010.

[4] P. Wang, Y. Liu, Z. Qin and G. Zhang, A novel hybrid forecasting model for PM_{10} and SO_2 daily concentrations, *Science of the Total Environment*, vol.505, pp.1202-1212, 2015.

[5] X. Feng, Q. Li, Y. Zhu et al., Artificial neural networks forecasting of $PM_{2.5}$ pollution using air mass trajectory based geographic model and wavelet transformation, *Atmospheric Environment*, vol.107, pp.118-128, 2015.

[6] M. S. Baawain and A. S. Al-Serihi, Systematic approach for the prediction of ground-level air pollution using an artificial neural network, *Aerosol and Air Quality Research*, vol.14, pp.124-134, 2014.
[7] T.-Y. Pai, C.-L. Ho, S.-W. Chen et al., Using seven types of $\mathrm{GM}(1,1)$ model to forecast hourly particulate matter concentration in banciao city of Taiwan, *Water, Air, and Soil Pollution*, vol.217, pp.25-33, 2011.
[8] J. Lang, X. Su and X. Zhou, Prediction for air pollution index based on organic grey neural network model, *Journal of Harbin Institute of Technology*, vol.36, pp.1598-1602, 2004.