

Research Article

A Long-Term Prediction Model of Beijing Haze Episodes Using Time Series Analysis

Xiaoping Yang,¹ Zhongxia Zhang,¹ Zhongqiu Zhang,² Liren Sun,¹ Cui Xu,¹ and Li Yu¹

¹School of Information, Renmin University of China, Beijing 100872, China

²School of Computer Science, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Li Yu; buaayuli@ruc.edu.cn

Received 20 April 2016; Revised 28 June 2016; Accepted 10 July 2016

Academic Editor: Francisco Martínez-Álvarez

Copyright © 2016 Xiaoping Yang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The rapid industrial development has led to the intermittent outbreak of pm2.5 or haze in developing countries, which has brought about great environmental issues, especially in big cities such as Beijing and New Delhi. We investigated the factors and mechanisms of haze change and present a long-term prediction model of Beijing haze episodes using time series analysis. We construct a dynamic structural measurement model of daily haze increment and reduce the model to a vector autoregressive model. Typical case studies on 886 continuous days indicate that our model performs very well on next day's Air Quality Index (AQI) prediction, and in severely polluted cases (AQI \geq 300) the accuracy rate of AQI prediction even reaches up to 87.8%. The experiment of one-week prediction shows that our model has excellent sensitivity when a sudden haze burst or dissipation happens, which results in good long-term stability on the accuracy of the next 3–7 days' AQI prediction.

1. Introduction

Industry of developing countries is mainly centralized around big cities, accompanied by a large population, consumption, and pollution. Together with Tianjin city and Hebei province, Northern China has become one of the most prosperous and polluted areas on Earth. By 2013, the transient population of Beijing was 37.5 million, and the intermittent outbreak of air pollution has greatly impacted every citizen's life: physiological diseases [1, 2], depression, and poor visibility in traffic [3, 4]. The main component of haze is pm2.5 (particulate matters less than 2.5 μm in aerodynamic diameter), and the concentration of pollution is described with Air Quality Index (AQI, the concentration of pm2.5). The Chinese Government began to monitor and record pm2.5 concentrations for major cities since 2013 [5]. According to the report of Quan et al. [6], the AQI reached 600 in Beijing during the haze event in January 2013. In recent years, more and more papers have referred to the haze episodes and the consequences in Northern China [7–11]. Researchers pointed out that, over the coming years, haze episodes would continue to burst frequently in Northern China [12].

This paper presents an AQI prediction model of Beijing based on time series analysis. We collected Beijing's AQI data of 29 continuous months since 2013 and constructed a dynamic structural prediction model. Statistical methods are used to obtain the maximum likelihood estimation of the prediction model. And both short-term and long-term experiments are carried out to test the accuracy and robustness of our model.

The remainder of this paper is organized as follows. In Section 2, we introduce recent related work. Section 3 presents our prediction model and proves our model to be a vector autoregressive model. Experiments and evaluations are reported in Section 4. We conclude the paper in Section 5 with future works.

2. Related Work

Generally, pm2.5, or haze, is born mainly through anthropogenic factors [13–16] and eliminated by natural diffusion. Several days after emission, secondary pm2.5 is produced through photochemical reactions among indiffusible pollutants. Secondary pm2.5 is the principal component in most

severe haze episodes in China [17]. A typical way of haze prediction is to use pollutant emission data (CO, SO₂, and NO_x) in the simulation [5, 18]. Huang et al. [14] analyzed the chemical compositions of pm2.5 and used chemical mass balance to identify the emission sources. Other more complex models are proposed to introduce the atmospheric features, chemistry components, and transport factors [15]. But the more common case is that pollutant emission data usually increase or decrease synchronously with AQI. Sun [19] took population, car ownership, and GDP into consideration and proposed a statistical index system of average annual haze episode days. They found that although most factors contribute to predicting pm2.5, the annual average of NO_x is negatively correlated with average severely polluted days. The paper [12] established a cubic exponential smoothing model by introducing dust emission into haze prediction. Liang et al. pointed out that there are various distribution and transmission patterns of pm2.5 [20]. In fact, Wang et al. mentioned that the government control policy should be considered in model simulations [9].

Many researches use backpropagation neural network as the simulation model [19, 21]. Statistical time series analysis is rarely used in haze prediction, so long-term haze prediction is difficult for current methods to accomplish [22]. Multiple linear regression models also perform well on daily scale prediction [23, 24]. However, the test data of existing researches is not ample; for example, [21] tested the prediction accuracy on only 3 days. Besides, Zhang et al. pointed out that pm2.5 accumulation in previous days significantly affects the present daily pm2.5 concentration, which should also be a concern in the modeling process [22].

Considering the above points, this paper presents a new AQI prediction model integrated with natural factor, humanity factor, and self-evolution factor.

3. The Prediction Model of Beijing's Daily AQI

3.1. The Parameters and Architecture of the Prediction Model.

The change of daily pm2.5 concentration depends on two factors: daily overall production of pm2.5 by human activities P_t and daily overall natural diffusion or overall natural accumulation of pm2.5 C_t . The production of haze depends a lot on the control policies of the government toward the emission of industry fuels I_t . The diffusion of haze mainly depends on the airflow W_t . Besides, complex chemical changes could occur between pm2.5 and other pollutants; thus, previous day's pm2.5 concentration also affects the AQI, which could be seen as the evolution result of previous day's pm2.5 and is represented by Y_t . Apparently, $P_t - C_t$ could be directly observed. P_t is generated by a semimanual method. P_t is mainly related to daily human activities, and we calculate P_t from AQI sequences of no less than five consecutive sunny and windless days. Special circumstances are also considered. In winter, P_t will be larger because the heating system is on. The car usage restrictions and temporary stoppage of factories during Beijing APEC 2014 are also taken into consideration. C_t is then calculated as $P_t - (P_t - C_t)$. Sometimes, C_t is greater than zero, which means pm2.5 accumulates because of nonhuman factors.

Thus, the daily net growth of pm2.5 ($P_t - C_t$) is a function of the evolution result Y_t , the industry control index I_t , and the forecast of wind power W_t . Consider this problem as a dynamic structural model, and our model can be described as

$$P_t - C_t = \beta_0 + \beta_1 Y_t + \beta_2 W_t + \beta_3 I_t + \beta_4 (P_{t-1} - C_{t-1}) + \mu_t^D. \quad (1)$$

Parameters β_1 , β_2 , and β_3 , respectively, represent the effect caused by the pm2.5 of the previous day, the wind power, and the industry control index. The net growth of previous day's pm2.5 partly affects present day's pm2.5 and partly affects the next day's pm2.5. The parameter β_4 represents this "partial adjustment." The disturbance μ_t^D represents other factors which affect present day's pm2.5.

3.2. Complexity Reduction of the Prediction Model. In order to facilitate the research and modeling process, we have proved that this model could be reduced to a vector autoregressive model.

Proposition 1. *Formula (1) is a vector autoregressive model.*

Proof. Assume that there exists sequence autocorrelation in formula (1). The autocorrelation is

$$\mu_t^D = \rho \mu_{t-1}^D + v_t^D \quad (2)$$

in which v_t^D is white noise. Here, we apply the Cochrane-Orcutt iteration to rewrite formula (2):

$$(1 - \rho L) \mu_t^D = v_t^D, \quad (3)$$

where L is the lag operator ($LV_t \equiv V_{t-1}$), which can convert the last phase to current value in a time series.

The next work is to find the most satisfying value of ρ through successive iteration method. Specifically, this method uses residual error to estimate the unknown ρ .

Assume that we use previous p days' AQI to predict present day's AQI. Multiply $(1 - \rho L)$ on both sides of formula (1); the expansion formula will be as follows:

$$\begin{aligned} P_t = & k_1 + \beta_{12}^0 C_t + \beta_{13}^0 I_t + \beta_{14}^0 Y_t + \beta_{15}^0 W_t + \beta_{11}^1 P_{t-1} \\ & + \beta_{12}^1 C_{t-1} + \beta_{13}^1 I_{t-1} + \beta_{14}^1 Y_{t-1} + \beta_{15}^1 W_{t-1} \\ & + \beta_{11}^2 P_{t-2} + \beta_{12}^2 C_{t-2} + \beta_{13}^2 I_{t-2} + \beta_{14}^2 Y_{t-2} \\ & + \beta_{15}^2 W_{t-2} + \cdots + \beta_{11}^p P_{t-p} + \beta_{11}^p C_{t-p} + \beta_{11}^p I_{t-p} \\ & + \beta_{11}^p Y_{t-p} + \beta_{11}^p W_{t-p} + v_t^D. \end{aligned} \quad (4)$$

In the substitution process, many assumptions are neglected. But the ordinary least square method (OLS estimation) should not be used in the estimation of formula (4), because OLS can only illustrate the relationship between daily pm2.5 production and the policy control index, the accumulation of history pm2.5, and the wind power. The net

growth of previous day's pm2.5 is only one reason of the correlation of these variables.

The government could make policies to control pm2.5 production of industry to obtain "satisfying" daily production of pm2.5; that is, I_t is an endogenous variable. And the policy control index depends on present day's and previous p days' accumulation of history pm2.5, the wind power, the daily production of pm2.5, and daily diffusion of pm2.5:

$$\begin{aligned} I_t = & k_3 + \beta_{31}^0 P_t + \beta_{32}^0 C_t + \beta_{33}^0 Y_t + \beta_{34}^0 W_t + \beta_{31}^1 P_{t-1} \\ & + \beta_{32}^1 C_{t-1} + \beta_{33}^1 Y_{t-1} + \beta_{34}^1 W_{t-1} + \beta_{35}^1 I_{t-1} \\ & + \beta_{31}^2 P_{t-2} + \beta_{32}^2 C_{t-2} + \beta_{33}^2 Y_{t-2} + \beta_{34}^2 W_{t-2} \quad (5) \\ & + \beta_{35}^2 I_{t-2} + \cdots + \beta_{31}^p P_{t-p} + \beta_{32}^p C_{t-p} + \beta_{33}^p Y_{t-p} \\ & + \beta_{34}^p W_{t-p} + \beta_{35}^p I_{t-p} + v_t^C, \end{aligned}$$

where v_t^C represents the influence brought about by other policies.

The net growths of previous days' pm2.5 and policy control index also have an effect on daily accumulation of pm2.5:

$$\begin{aligned} Y_t = & k_4 + \beta_{41}^0 P_t + \beta_{42}^0 C_t + \beta_{43}^0 I_t + \beta_{44}^0 W_t + \beta_{41}^1 P_{t-1} \\ & + \beta_{42}^1 C_{t-1} + \beta_{43}^1 I_{t-1} + \beta_{44}^1 W_{t-1} + \beta_{45}^1 Y_{t-1} \\ & + \beta_{41}^2 P_{t-2} + \beta_{42}^2 C_{t-2} + \beta_{43}^2 I_{t-2} + \beta_{44}^2 W_{t-2} \quad (6) \\ & + \beta_{45}^2 Y_{t-2} + \cdots + \beta_{41}^p P_{t-p} + \beta_{42}^p C_{t-p} + \beta_{43}^p I_{t-p} \\ & + \beta_{44}^p W_{t-p} + \beta_{45}^p Y_{t-p} + v_t^A, \end{aligned}$$

where v_t^A represents other factors that influence daily accumulation of pm2.5.

Analogized from formulas (4), (5), and (6), C_t and W_t can both be written in a similar form. Join formulas (4), (5), and (6) together, and rewrite them into vector form:

$$B_0 x_t = K + B_1 x_{t-1} + B_2 x_{t-2} + \cdots + B_p x_{t-p} + v_t \quad (7)$$

in which

$$\begin{aligned} x(t) = & (P_t, C_t, I_t, Y_t, W_t)^T, \\ v_t = & (v_t^D, v_t^S, v_t^C, v_t^A, v_t^H)^T, \\ K = & (k_1, k_2, k_3, k_4, k_5), \\ B_0 = & \begin{bmatrix} 1 & -\beta_{12}^0 & -\beta_{13}^0 & -\beta_{14}^0 & -\beta_{15}^0 \\ -\beta_{21}^0 & 1 & -\beta_{23}^0 & -\beta_{24}^0 & -\beta_{25}^0 \\ -\beta_{31}^0 & -\beta_{32}^0 & 1 & -\beta_{34}^0 & -\beta_{35}^0 \\ -\beta_{41}^0 & -\beta_{42}^0 & -\beta_{43}^0 & 1 & -\beta_{45}^0 \\ -\beta_{51}^0 & -\beta_{52}^0 & -\beta_{53}^0 & -\beta_{54}^0 & 1 \end{bmatrix}. \quad (8) \end{aligned}$$

In B_0 , the parameters in the 1st, 2nd, 3rd, 4th, and 5th row, respectively, relate P_t , C_t , I_t , Y_t , and W_t to the other variables. Every B_s is a 5×5 matrix. Premultiply formula (7) by B_0^{-1} (the inverse matrix of B_0):

$$x_t = c + \Psi_1 x_{t-1} + \Psi_2 x_{t-2} + \cdots + \Psi_p x_{t-p} + \varepsilon_t \quad (9)$$

in which

$$\begin{aligned} c = & B_0^{-1} K, \\ \Psi_s = & B_0^{-1} B_s, \quad (10) \\ \varepsilon_t = & B_0^{-1} v_t. \end{aligned}$$

This is the standard form of vector autoregressive model. So it is proved that our prediction model (formula (1)) is in fact a vector autoregressive model. \square

The regression parameters of our haze prediction model can be obtained as follows.

Let

$$-\Gamma = [K \ B_1 \ B_2 \ \cdots \ B_p], \quad (11)$$

$$y_t = [1 \ x_{t-1} \ x_{t-2} \ \cdots \ x_{t-p}]^T.$$

The dynamic structural system (formula (7)) is

$$B_0 x_t = -\Gamma x_t + v_t. \quad (12)$$

Assume that the disturbance terms are not sequence correlated or correlated to each other, which means

$$E(v_t v_\tau^T) = \begin{cases} D, & t = \tau, \\ 0, & t \neq \tau. \end{cases} \quad (13)$$

D is a main diagonal matrix. Formula (12) could be written as

$$x_t = \Phi^T y_t + \varepsilon_t \quad (14)$$

in which

$$\begin{aligned} \Phi^T = & B_0^{-1} - \Gamma, \\ \varepsilon_t = & B_0^{-1} v_t. \end{aligned} \quad (15)$$

Let Ω be the variance-covariance matrix of ε_t :

$$\Omega = E(\varepsilon_t \varepsilon_t^T) = B_0^{-1} E(v_t v_t^T) (B_0^{-1})^T = B_0^{-1} D (B_0^{-1})^T. \quad (16)$$

Suppose B_0 is a lower triangular matrix, in which all main diagonal elements are assigned 1, and D is a main diagonal matrix. The parameters (B_0, Γ, D) can be obtained through the maximum likelihood estimation of complete information. The maximum likelihood estimation of Ω can be obtained by the variance-covariance matrix of regression residual.

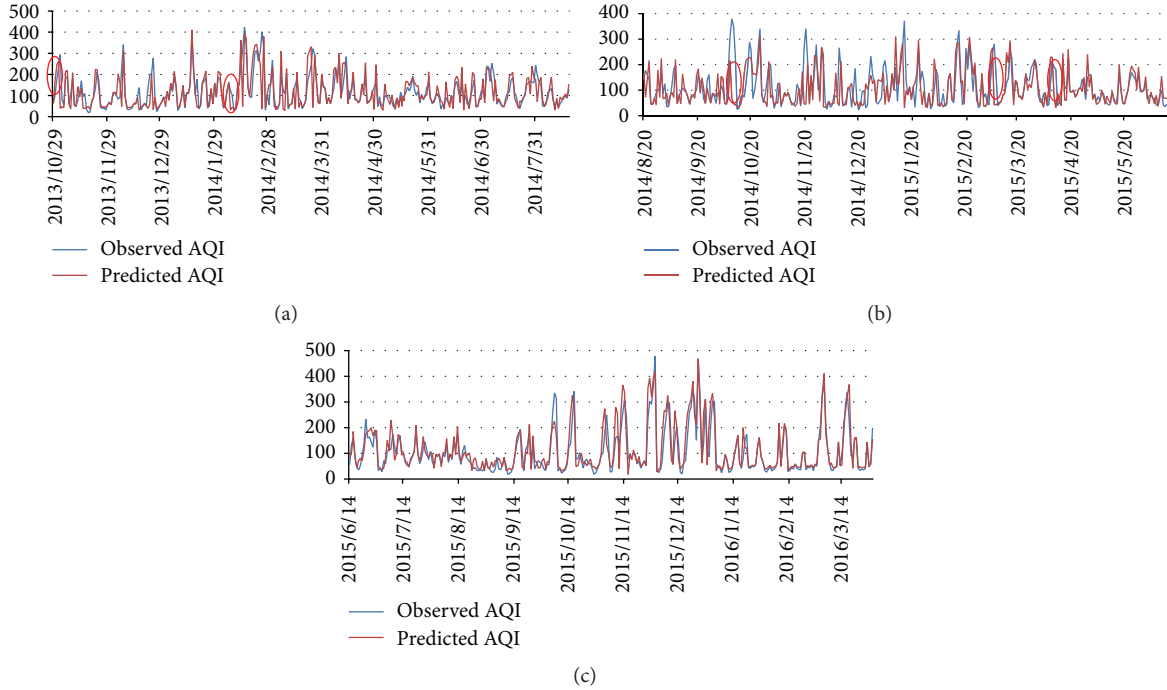


FIGURE 1: (a, b, c) Next day's AQI prediction on 886 continuous days.

Finally, \widehat{B}_0^{-1} and D are calculated through triangular decomposition of $\widehat{\Omega}$; thus, Γ can be evaluated.

Above all, the prediction model of Beijing AQI has considered factors including industry emission and policy control, together with the chemical changes of previous days' pollution accumulation and the diffusion conditions. This model also takes the correlations between these factors into consideration and introduces time series haze features into the dynamic structural model. The policy control index is simulated by the record of 4 severe haze episodes during this period. The diffusion is evaluated by weather record of daily wind power.

4. Model Evaluations

We collected the daily AQI and daily weather information from 28 Oct. 2013 to 31 Mar. 2016. This complete sequence is used to test the accuracy of the prediction model. The next day's AQI prediction experiment (Section 4.1) and long-term AQI prediction experiment (Section 4.2) are both implemented. The next day's AQI prediction is evaluated from two perspectives: the accuracy of daily prediction and the accuracy of statistical analysis.

4.1. Next Day's AQI Prediction. The next day's weather forecast information is applied in next day's AQI prediction. The observed and predicted daily mean AQI in Beijing are illustrated in Figure 1. The simulation result shows that the predicted value matched the observed value very well on the whole sequence of 886 days. Sometimes, there is severe deviation from the observed value; for example, on 19 Feb. 2014, the observed AQI was 89, while our model gives a prediction of 209, with an offset of 135%. But the fact is,

in the afternoon of 19 Feb., the wind of Beijing suddenly changed from northeasterly to southwesterly, and by 19:00 the AQI has reached already up to 170, which could be interpreted as our model successfully forecasted a severe haze outbreak several hours in advance; in the coming 7 days, the average daily AQI of Beijing is 305. The occasional occurrence of this "foreseeing" phenomenon is caused by coarse time granularity (daily), and this phenomenon is marked with red ellipse in Figure 1. These marks indicate that our model could "foresee" the sharp change of both outbreaks and diffusions. Most haze outbreaks and diffusions could be accurately simulated; some could be foreseen but could never be delayed.

Figure 2(a) is a scatter diagram of daily AQI, including both observed value and predicted value. Most points lie close to $y = x$ (the red line). But some points lie in a queue at the bottom part, which means the observed AQI exceeds 200, while the predicted value is less than 50. There are altogether 15 such outliers, 7 of which "foresee" haze diffusion, while the other 8 bug points could not be well interpreted. All the 15 points are checked and listed in Table 1. "✓" means a "foreseeing" phenomenon, and "?" represents bug points. Figure 2(b) is a scatter diagram of annual AQI (sum of daily AQI in a certain year). Our data covers only 2 months of 2013 and 3 months of 2016, so, in this diagram, these 2 points lie in the lower left corner.

The pie chart in Figure 3 shows the distribution of prediction accuracy. The deviation of predicted and observed AQI is obtained through the following formula:

$$\text{Dev}_t = \frac{|\text{PredictedValue} - \text{ObservedValue}|}{\text{ObservedValue}} \times 100\%. \quad (17)$$

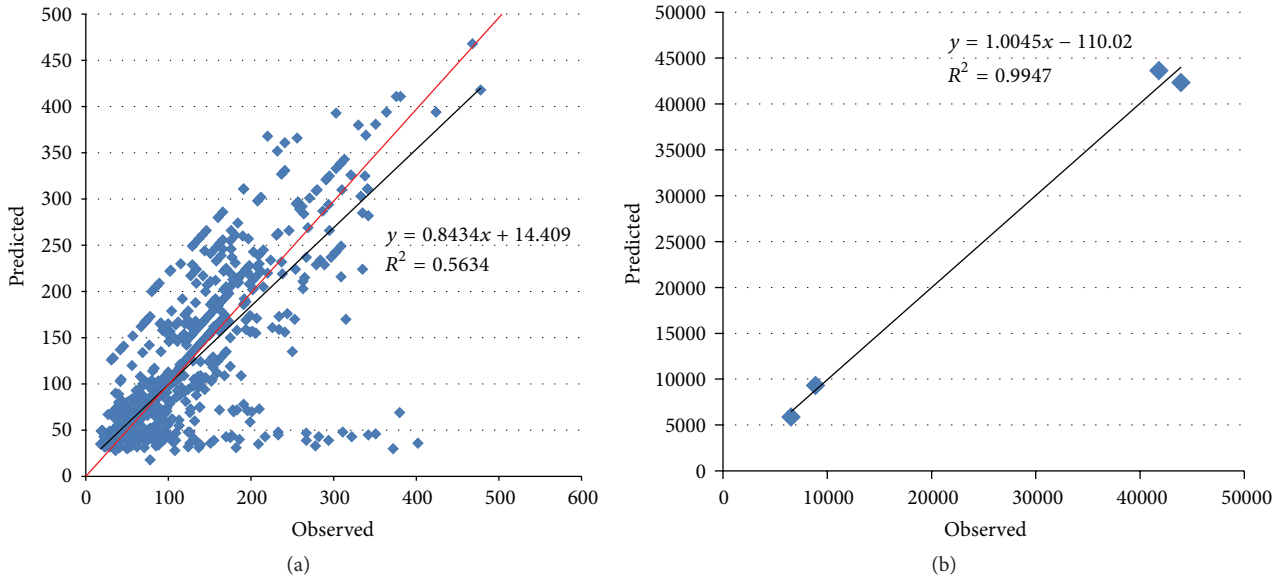


FIGURE 2: (a) Daily AQI of the 886 days. (b) Annual AQI from 2013 to 2016.

TABLE 1: All the 15 outliers in Figure 2(a).

Date of outlier	Label
Nov. 2, 2013	✓
Dec. 7, 2013	?
Dec. 25, 2013	✓
Feb. 14, 2014	?
Feb. 25, 2014	?
Mar. 26, 2014	?
Oct. 10, 2014	✓
Oct. 11, 2014	✓
Nov. 19, 2014	?
Nov. 20, 2014	?
Nov. 30, 2014	✓
Dec. 9, 2014	?
Jan. 4, 2015	?
Jan. 15, 2015	✓
Mar. 7, 2015	✓

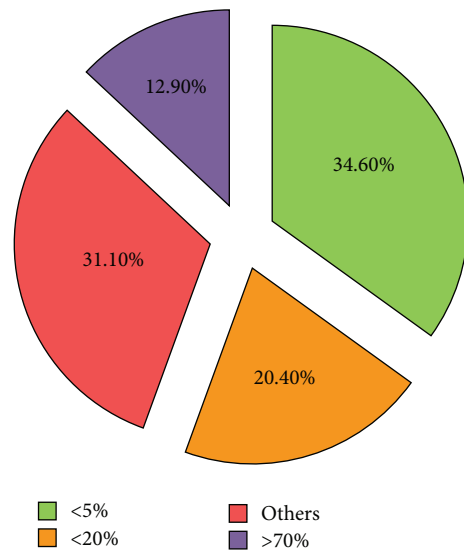


FIGURE 3: The deviation of predicted and observed AQI.

Figure 3 shows that 55% predictions match the observed values very well (<20% deviation). The purple part is mainly caused by the “foreseeing” phenomenon. Most samples of the red part come from less-polluted days. For example, on 12 Jan. 2016, the AQI prediction is 40 while the observed AQI is 29, which makes a deviation of 37.9%. In fact, statistics also indicate that our model performs better in worse air conditions (Figure 4). A sample is correctly predicted if the deviation of a sample is less than 20% or the predicted air quality level matches the observed level.

4.2. Long-Term AQI Prediction. In the long-term prediction, we use history haze data sequence and weather forecast information to predict the next 7 days’ AQI. A sample is

correctly predicted if the deviation of a sample is less than 20% or the predicted air quality level matches the observed level. From 26 Dec. 2015 to 31 Mar. 2016, we predict the AQI in the next 7 days and check the accuracy of n -day predictions. Figure 5 shows the accuracy of long-term prediction in the 91 days’ experiment. Figure 5 shows that the accuracy stays stable on the next 3, 4, 5, 6, and 7 days’ AQI prediction, which indicates that our model has excellent robustness on the task of long-term prediction. The next day’s prediction accuracy surprisingly reaches 79.1%, which is far better than the experiment in Section 4.1. The reason is that, during the 91 days, 6 haze episodes attacked Beijing. These frequent attacks did contribute a lot to the overall performance because our model is very sensitive to sudden changes of AQI, including

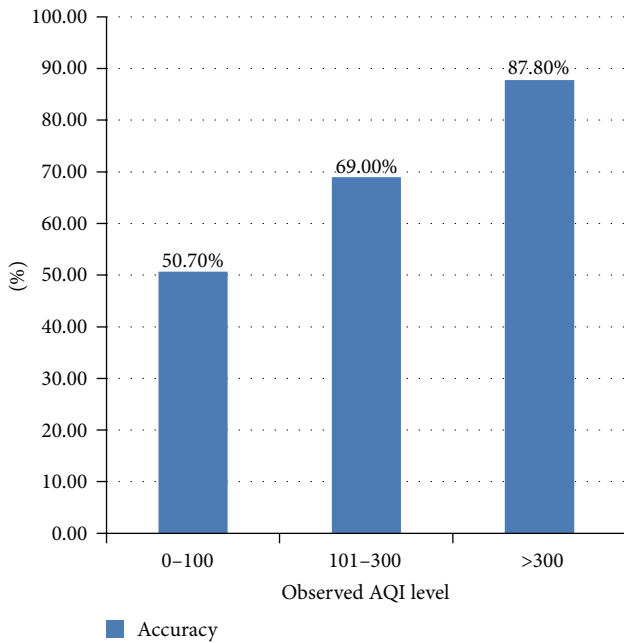


FIGURE 4: Prediction accuracies of different air qualities.

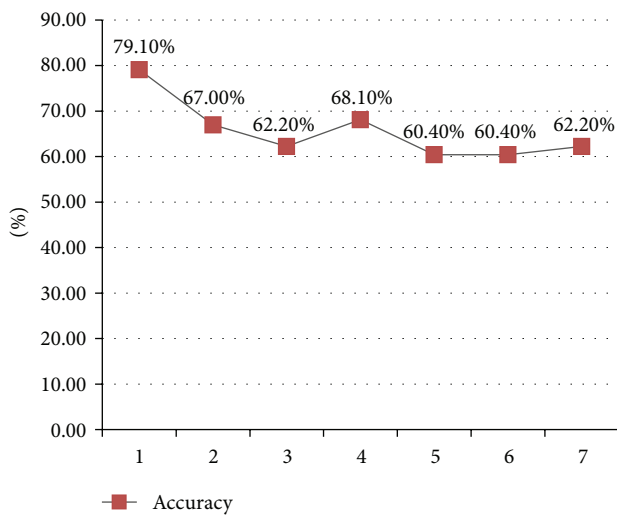


FIGURE 5: The accuracy of long-term AQI prediction.

outbreaks and diffusions (Section 4.1; Figure 4). Figures 6 and 7 show several haze episodes during the 91 days. Both figures show a pm2.5 change process of more than 2 weeks. Figure 6 also shows a “foreseeing” phenomenon caused by coarse time granularity, marked by a red ellipse.

5. Conclusion and Future Work

This paper presented a dynamic structural model to predict Beijing’s daily AQI. This model integrated natural factor, humanity factor, and self-evolution factor into the time series model. This dynamic structural measurement model of daily haze increment is proven to be a vector autoregressive model. Experiments reflected two highlights of this model. First,

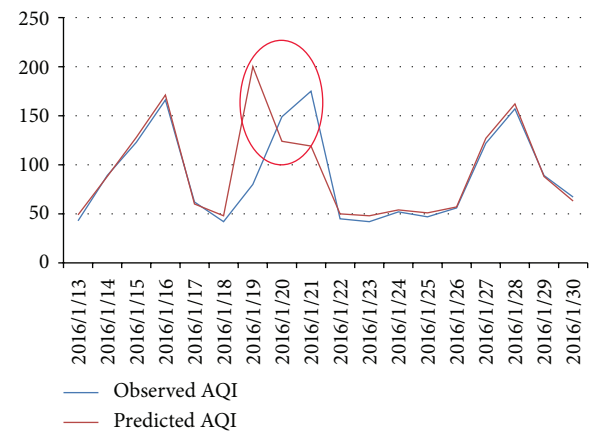


FIGURE 6: Three haze episodes in Jan. 2016.

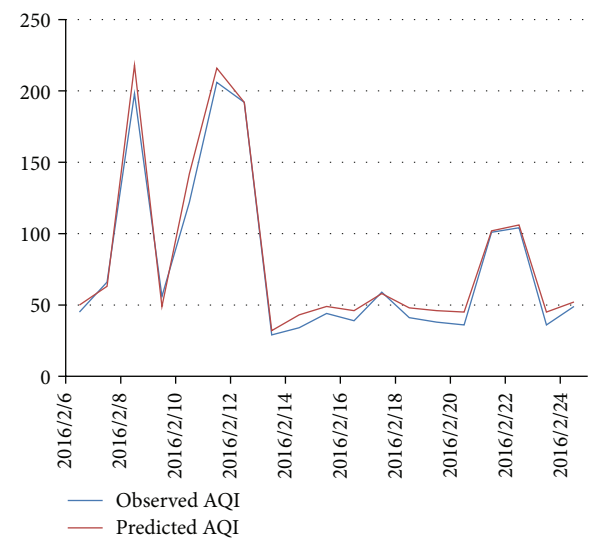


FIGURE 7: Three haze episodes in Feb. 2016.

our model is very sensitive to and performs very well on predicting sudden changes of AQI, including both outbreaks and diffusions. Second, the model has great robustness on the task of long-term AQI prediction. Lastly, limited by the coarse time granularity, our model sometimes “foresees” but never delays or misses any sudden changes of haze.

Many researchers use simple backpropagation neural network to accomplish nonlinear prediction models. But since methods based on time series are proven to be effective in haze prediction modeling, we believe that recurrent neural networks give better performances in such a prediction task. Although the related factors are limited in existing models, the overfitting problem should still be concerned, because, in long-term prediction, a deviation could spread and be exaggerated in the following days’ predictions.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This research was supported by the National Natural Science Foundation of China under Grant 71271209, Beijing Municipal Natural Science Foundation under Grant 4132052, and Humanity and Social Science Youth Foundation of Ministry of Education of China under Grant I1YJC630268.

References

- [1] L. Gao, G. Jia, R. Zhang et al., “Visual range trends in the Yangtze River Delta Region of China, 1981–2005,” *Journal of the Air and Waste Management Association*, vol. 61, no. 8, pp. 843–849, 2011.
- [2] Q. Liu, J. Baumgartner, Y. Zhang, and J. J. Schauer, “Source apportionment of Beijing air pollution during a severe winter haze event and associated pro-inflammatory responses in lung epithelial cells,” *Atmospheric Environment*, vol. 126, pp. 28–35, 2016.
- [3] X. Tie, G. P. Brasseur, C. Zhao et al., “Chemical characterization of air pollution in Eastern China and the Eastern United States,” *Atmospheric Environment*, vol. 40, no. 14, pp. 2607–2625, 2006.
- [4] J. Deng, T. Wang, Z. Jiang et al., “Characterization of visibility and its affecting factors over Nanjing, China,” *Atmospheric Research*, vol. 101, no. 3, pp. 681–691, 2011.
- [5] MEPC (Ministry of Environmental Protection of China), *China Environment Yearbook*, China Environment Yearbook Press, Beijing, China, 2014.
- [6] J. Quan, X. Tie, Q. Zhang et al., “Characteristics of heavy aerosol pollution during the 2012–2013 winter in Beijing, China,” *Atmospheric Environment*, vol. 88, pp. 83–89, 2014.
- [7] X. J. Zhao, P. S. Zhao, J. Xu et al., “Analysis of a winter regional haze event and its formation mechanism in the North China Plain,” *Atmospheric Chemistry & Physics*, vol. 13, no. 11, pp. 5685–5696, 2013.
- [8] H. Wang, J. An, L. Shen et al., “Mechanism for the formation and microphysical characteristics of submicron aerosol during heavy haze pollution episode in the Yangtze River Delta, China,” *Science of the Total Environment*, vol. 490, pp. 501–508, 2014.
- [9] Y. Wang, L. Li, C. Chen et al., “Source apportionment of fine particulate matter during autumn haze episodes in Shanghai, China,” *Journal of Geophysical Research Atmospheres*, vol. 119, no. 4, pp. 1903–1914, 2014.
- [10] D. Ji, L. Li, Y. Wang et al., “The heaviest particulate air-pollution episodes occurred in northern China in January, 2013: insights gained from observation,” *Atmospheric Environment*, vol. 92, pp. 546–556, 2014.
- [11] J. Hu, Y. Wang, Q. Ying, and H. Zhang, “Spatial and temporal variability of PM_{2.5} and PM₁₀ over the North China Plain and the Yangtze River Delta, China,” *Atmospheric Environment*, vol. 95, pp. 598–609, 2014.
- [12] Q. Hou and H. Yang, “Analysis and forecasting of haze weather based on the cubic exponential smoothing model,” *Environmental Protection Science*, vol. 6, pp. 73–77, 2014.
- [13] S. Guo, M. Hu, M. L. Zamora et al., “Elucidating severe urban haze formation in China,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 49, pp. 17373–17378, 2014.
- [14] R.-J. Huang, Y. Zhang, C. Bozzetti et al., “High secondary aerosol contribution to particulate pollution during haze events in China,” *Nature*, vol. 514, no. 7521, pp. 218–222, 2015.
- [15] M. Vieno, M. Heal R, S. Hallsworth et al., “The role of long-range transport and domestic emissions in determining atmospheric secondary inorganic particle concentrations across the UK,” *Atmospheric Chemistry & Physics*, vol. 14, pp. 8435–8447, 2014.
- [16] G. Kiesewetter, J. Borkenkleefeld, W. Schöpp et al., “Modelling street level PM₁₀ concentrations across Europe: source apportionment and possible futures,” *Atmospheric Chemistry & Physics*, vol. 14, pp. 18315–18354, 2014.
- [17] R. Zhang, L. Wang, A. F. Khalizov et al., “Formation of nanoparticles of blue haze enhanced by anthropogenic pollution,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 106, no. 42, pp. 17650–17654, 2009.
- [18] J. M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, Mass, USA, 2nd edition, 2010.
- [19] Y. Sun, *Research on the Forecast of Haze Days Based on the Principal Component-BP Neural Network Model*, Jiangsu University, 2015.
- [20] X. Liang, T. Zou, B. Guo et al., “Assessing Beijing’s PM_{2.5} pollution: severity, weather impact, APEC and winter heating,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2182, 2015.
- [21] H. Ai and Y. Shi, “Study on prediction of haze based on BP neural network,” *Computer Simulation*, vol. 32, pp. 402–405, 2015.
- [22] X. Zhang, Y. Wu, and B. Gu, “Characterization of haze episodes and factors contributing to their formation using a panel model,” *Chemosphere*, vol. 149, pp. 320–327, 2016.
- [23] L. Li, J. Qian, C.-Q. Ou, Y.-X. Zhou, C. Guo, and Y. Guo, “Spatial and temporal analysis of Air Pollution Index and its timescale-dependent relationship with meteorological factors in Guangzhou, China, 2001–2011,” *Environmental Pollution*, vol. 190, pp. 75–81, 2014.
- [24] G. Tian, Z. Qiao, and X. Xu, “Characteristics of particulate matter (PM₁₀) and its relationship with meteorological factors during 2001–2012 in Beijing,” *Environmental Pollution*, vol. 192, pp. 266–274, 2014.