Environmental Engineering and Management Journal

July 2019, Vol. 18, No. 7, 1575-1590
http://www.eemj.icpm.tuiasi.ro/; http://www.eemj.eu

"Gheorghe Asachi" Technical University of Iasi, Romania

STUDY ON PREDICTION MODEL OF SPACE-TIME DISTRIBUTION OF AIR POLLUTANTS BASED ON ARTIFICIAL NEURAL NETWORK

Zhuang Wu[1,2*], Jiaorong Fan[1], Ying Gao[1], Huayang Shang[1,2], Hongquan Song[3,4]

[1]*School of Information, Capital University of Economics and Business, Beijing 100070, China*
[2]*CTSC center, Information College, Capital University of Economics and Business, Beijing 100070, China*
[3]*Key Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions, Ministry of Education*
[4]*College of Environment and Planning, Henan University, Kaifeng, Henan 475004, China*

Abstract

Along with the development of industrialization and urbanization, the pollution of the atmosphere increases and the air quality decreases gradually, especially in the developed regions such as Beijing, Tianjin and Hebei in China. The frequent occurrence of haze leads to the decrease of visibility day by day, which causes great inconvenience to people's mind and body. Therefore, an in-depth study of regional weather quality impact factors can improve the air quality in the regions, which is of important significance to haze warning. This study mainly explores the main causes of haze formation and forecasts the haze, and puts forward some suggestions on haze control. This study proposes a space-time prediction model of air pollutants based on artificial neural network. The selected research regions are cities in Beijing-Tianjin-Hebei region. The data are real-time daily data of monitoring stations published by Beijing Environmental Protection Monitoring Center and China Meteorological Administration, and the air environment changes of 13 cities in Beijing-Tianjin-Hebei region in 2017 are studied. The impact factors include $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3. Through collecting and visualizing the data, evaluating with air quality index (AQI), this study designs, establishes, trains, and simulates the backpropagation neural network, adjusts the corresponding parameters, obtains the optimal model and the main research conclusion, as well as provides scientific reasonable suggestion.

Key words: air pollutants, backpropagation neural network, $PM_{2.5}$, space-time distribution

Received: September, 2018; Revised final: January, 2019; Accepted: April, 2019; Published in final edited form: April, 2019

1. Introduction

In developed regions such as Beijing, Tianjin and Hebei in China, haze pollution occurs persistently for a long time. Among them, Beijing is one of the most serious air pollution cities in China. Air pollution control is a problem to be solved, so the study on air pollution is the main problem discussed in the society nowadays (Singh et al., 2018; Song et al., 2012). There exists significant spatial gradient in pollution of fog and haze in Beijing and the surrounding areas, as indicated by progressive decrease from south to north which is mainly the result of mutual accumulation of regional pollution. The northern area is surrounded by mountains in and plenty of green vegetation can purify the air to a certain degree. On the other hand, as the southern area is adjacent to such severely polluted regions as Tianjin Municipality and Hebei Province, external influence is more significant (Chai et al., 2014; Wang et al., 2014). So, regional collaborative prevention and treatment is required to control the pollution of $PM_{2.5}$ in Beijing and special attention shall be paid to the transmission of pollutants from the southwest and the control of major sources.

$PM_{2.5}$ concentration fluctuates greatly in different reasons. As a whole, the concentration of $PM_{2.5}$ is relatively high in winter and low in summer with a U-shaped distribution all year around.

* Author to whom all correspondence should be addressed: e-mail: wuzhuang@cueb.edu.cn

According to "Measures of Beijing Municipality for Administration of Heat Supply and Use", the heating period of this Municipality is from November 15 of the current year to March 15 of the next year, so $PM_{2.5}$ concentration goes up in winter mainly because household heating increases and causes more particulate matter discharged into the air (Ji et al., 2012; Xiao et al., 2015). Earlier, the seasonal characteristics of spring drought, few rains and strong wind as well as unreasonable human activities in Beijing have frequently resulted in large-scale outbreak of sand storm, but in the experiment result of this paper, $PM_{2.5}$ concentration doesn't increase greatly in spring, suggesting that the sandstorm source control project in Beijing-Tianjin-Hebei Region has worked well. Besides, some research achievements have once pointed out that there is difference in pollutant concentration between weekdays and weekends (Blanchard and Tanenbaum, 2006; Motallebi et al., 2003), but according the experiment result of Mann-Whitney U test in this paper, the foregoing difference is not statistically significant and the credit should be given to the implementation of vehicle ban policy in China.

The daily change rule of $PM_{2.5}$ is subject to the height of boundary layer (Liu et al., 2015). According to the calculation result, the lowest concentration usually occurs during 16:00-18:00 and the highest during 9:00-11:00 every day. Generally speaking, the peak $PM_{2.5}$ concentration in the morning is caused by human activities while the fall of $PM_{2.5}$ concentration in the afternoon is because the increase of height of boundary layer has accelerated the diffusion of $PM_{2.5}$. In the night-time, the height of boundary layer drops again and human activities increase, so $PM_{2.5}$ concentration rises again. Therefore, $PM_{2.5}$ concentration every morning (7am-12am) is much lower than every night (7pm-6am). Especially in winter, heating increases from coal burning and so does the particulate matter released to the air. Besides, less solar radiation also moves the time when the height of boundary layer falls in advance, so the night-time $PM_{2.5}$ level is relatively high in winter (Guinot et al., 2006; Miao et al., 2009; Oprea et al., 2017). $PM_{2.5}$ concentration changes mainly due to the interaction between air pollutants and meteorological elements. In the entire diffusion process, average wind speed plays a decisive role and it affects the diffusion of CO and the occurrence of secondary reaction. For NO_2 and SO_2, average wind speed has a critical value. The concentration of NO_2 and SO_2 remains stable below the critical value and secondary reaction occurs when exceeding the critical value. For O_3, as it is mainly distributed in the stratosphere with a height of 10~50km, average wind speed has no effect on it, but it can be concluded that there is negative correlation between $PM_{2.5}$ concentration and O_3 concentration through the calculation result (Zhang et al., 2015a; Zhang et al., 2015b).

By studying the characteristics and impact factors of air pollution in Beijing-Tianjin-Hebei region of China, this study uses visualization software to analyse the mass concentration and space-time distribution of air pollutants, as well as adopts air quality index to evaluate and analyse. The BP neural network model is established to predict the space-time distribution of air pollutants. This study has certain scientific significance for air quality evaluation. The main contents of this study are as follows:

(1) To study the space-time distribution of haze in Beijing, Tianjin and Hebei, and to describe and analyse the process of spatial information changing with time. Firstly, ArcGIS software is used to vectorise meteorological monitoring point data to generate spatial point data. Then coordinate system is set to carry out coordinate transformation and projection system on the data, to conduct visualization processing and classification discussion on sample data through ArcGIS software, and to analyse the space-time distribution characteristics of haze in Beijing-Tianjin-Hebei region. Secondly, the pollution is studied from the aspect of air quality index, mainly including the annual average value distribution of air quality index and the average value distribution of each city. Finally, the sample data is divided into four seasons and the distribution of haze in Beijing-Tianjin-Hebei region in spring, summer, autumn and winter is studied.

(2) To study the environmental factors affecting haze in Beijing-Tianjin- Hebei region.

The selected air pollution factors mainly include $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3. The time-variation characteristics of these pollution factors are studied, and the inner link and the standard reaching situation of these pollution factors are studied and analyzed, standard reaching is a set of criteria for air quality evaluation. At the same time, AQI is used to evaluate the air quality in Beijing-Tianjin-Hebei region.

(3) BP neural network is used to predict the main impact factors of haze in time series

The time series data boast the characteristics that other data don't have, that is, the law of historical data can be applied to the future (Sumita et al., 2016). This study establishes the BP neural network prediction model by using its advantages (Lakshmipathi and Battula, 2018; Sánchez-Escalona and Góngora-Leyva, 2018). The change of haze weather is non-linear and the environmental impact factor is a very complex non-linear structure. BP neural network has powerful non-linear processing capability that can just make up for this deficiency. Therefore, BP neural network can depend on the inner link of the data to extract the corresponding characteristics to achieve the goal of modelling and forecasting haze weather (Ma et al., 2012). BP neural network consists of three layers. The variable of the input layer is 7 dimensions composed of 7 main impact factors. The output layer is the value of AQI. The number of nodes in the hidden layer is determined by experiments. The model is trained by the data to achieve good prediction results.

In this paper, section 1 is the introduction. Section 2 is about the research methods, including

source of data, neural networks, air quality index and so on. Section 3 is the evaluation of air quality in Beijing-Tianjin-Hebei region. Finally, section 4 is the conclusion.

2. Research data and research method

2.1. Source of data

The data used in this study, such as routine meteorological observation data, daily data of meteorological reports, mass concentration data of $PM_{2.5}$ and surface weather charts, come from the National Oceanic and Atmospheric Administration, NASA website and China Environmental Statistics Bureau. The monitoring data are processed by visual software ArcGIS. ArcGIS is a platform for organizations to create, manage, share, and analyse spatial data, a software suite that integrates many advanced GIS applications and has the ability of real-time analysis and processing of large data. We use it as a tool to design, share, manage and publish geographic information. It includes mapping, spatial analysis, complex spatial query, advanced data editing, distributed data management, batch spatial processing, implementation of spatial geometric integrity rules and so on. The air quality in Beijing-Tianjin-Hebei region is comprehensively evaluated. The different pollution levels show different depth of colors in the drawn graph. The air factor concentrations near Beijing-Tianjin-Hebei region are drawn in a table, introduced into MATLAB space, and normalized during the establishment of BP neural network model.

2.2. Back Propagation Neural Network

The basic idea of BP neural network algorithm is that the learning process consists of two processes: forward propagation of signals and back propagation of errors. In forward propagation, input samples are transferred from the input layer to the output layer after being processed by the hidden layer. When the difference between the output value and the target value is very large, the output error propagates in some form through the direction of the hidden layer, and the error is distributed to the neurons of each layer. The neurons of each layer use the error signal as the basis for modifying the weight value. The two processes of forward propagation and error back propagation should not be repeated until the error is very small or the pre-set iteration time and iteration times are reached. BP neural network is mature in network theory and performance. Its outstanding advantage is that it has strong non-linear mapping ability and flexible network structure. The number of middle layers and neurons in each layer of the network can be set arbitrarily according to the specific situation, and its performance varies with the difference of the structure.

2.3. Air Quality Index (AQI)

AQI, which is generally reflected by the cleanliness or pollution of the air, is an index to assess the effect of polluted air on health after people breathing polluted air for a period of time. The early API only contains 3 kinds of pollution indexes. With the complexity of air factor increasing, the evaluation index is changed to AQI in combination with the actual situation. The pollution factors $PM_{2.5}$, PM_{10}, SO_2, NO_2, CO and O_3 are study factors. Based on the analysis of the time-variation characteristics and standard reaching situation of each pollutant, air quality index is used to evaluate the air quality comprehensively. Air quality index can highlight the effect of single pollutant on the whole, that is, the air quality is determined by the improved air quality index corresponding to the concentration of a certain pollutant (Devi et al., 2016). For a certain pollutant P, the improved air quality index I_p corresponding to the mass concentration C_p can be calculated as follows Eq. (1), Air quality index and pollutant concentration limit is shown in the Table 1.

$$I_p = \frac{I_{ph} - I_{pl}}{C_{ph} - C_{pl}}(C_p - C_{pl}) + I_{pl} \tag{1}$$

Table 1. AQI and pollutant concentration limit

IAQI	*Concentration limits for pollutants*									
	SO_2 24-hour average/ ($\mu g/m^3$)	***SO_2 1-hour average/ ($\mu g/m^3$)***	***NO_2 24-hour average/ ($\mu g/m^3$)***	***NO_2 1-hour average/ ($\mu g/m^3$)***	***PM (with size ≤10um) 24-hour average/ ($\mu g/m^3$)***	***CO 24-hour average/ (mg/m³)***	***CO 1-hour average/ (mg/m³)***	***O3 1-hour average/ ($\mu g/m^3$)***	***O3 8-hour moving average/ ($\mu g/m^3$)***	***PM (with size ≤2.5um) 24-hour average/ ($\mu g/m^3$)***
0	0	0	0	0	0	0	0	0	0	0
50	50	150	40	100	50	2	5	160	100	35
100	150	500	80	200	150	4	10	200	160	75
150	475	650	180	700	250	14	35	300	215	115
200	800	800	280	1200	350	24	60	400	265	150
300	1600	(2)	565	2340	420	36	90	800	800	250
400	2100	(2)	750	3090	500	48	120	1000	(3)	350
500	2620	(2)	940	3840	600	60	150	1200	(3)	500

where, C_{ph} and C_{pl} indicates the maximum and minimum values of the concentration limits of pollutants close to C_p respectively, I_{ph} and I_{pl} are IAQI values corresponding to C_{ph} and C_{pl}. AQI value can be determined by the Eq. (2):

$$AQI = \max\{IAQI_1, IAQI_2, IAQI_3, \ldots, IAQI_n\} \quad (2)$$

where, if AQI is greater than 50, then the largest pollutant of IAQI is the primary pollutant. The grade and category of air quality is evaluated according to the value of AQI.

3. Evaluation of air quality in Beijing-Tianjin-Hebei region

3.1. Analysis of standard-reaching rate of concentration of main pollutants

(1) Standard-reaching rate of $PM_{2.5}$

The Fig. 1 shows the relationship between the standard-reaching rate of $PM_{2.5}$ and the annual average concentration.

$PM_{2.5}$ is a major pollutant affecting air quality. If its annual limit of secondary concentration is smaller than 35 μg/m³, it reaches the standard. Thus, the sample data is processed and screened out the data whose $PM_{2.5}$ concentration is smaller than 35μg /m³. Then it is divided by the total sample data of each city as the standard-reaching rate (Markowicz et al., 2016).

Fig. 1 shows that, the annual average concentration (μg/m³) of Chengde City, Qinhuangdao City and Zhangjiakou City in 2017 is lower. The lowest average annual concentration of Zhangjiakou is 32μg/m³. Only the standard-reaching rate of $PM_{2.5}$ in Chengde City and Zhangjiakou City in Beijing-Tianjin-Hebei region is higher than 60%, of which Zhangjiakou City has the highest standard-reaching rate of 71%. The rest of the cities are lower than 60%. In general, the standard-reaching rate of all cities in Beijing-Tianjin-Hebei region in 2017 is low, indicating that five sixths of cities are not at the standard level.

(2) Standard-reaching rate of PM_{10}

If the annual limit of secondary concentration is smaller than 70μg/m³, it reaches the standard. Thus, the sample data is processed and screened out the data whose PM_{10} concentration is smaller than 70μg /m³. Then it is divided by the total sample data of each city as the standard-reaching rate (Fang et al., 2016). The fig. 2 shows the relationship between the standard-reaching rate of PM_{10} and the annual average concentration.

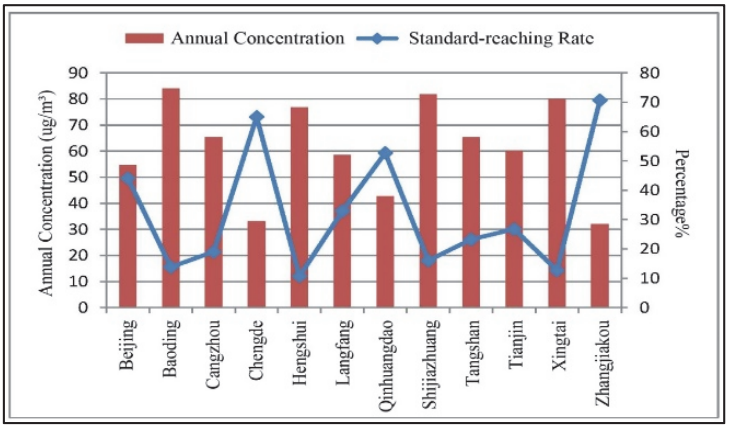


Fig. 1. Standard-reaching rate of $PM_{2.5}$

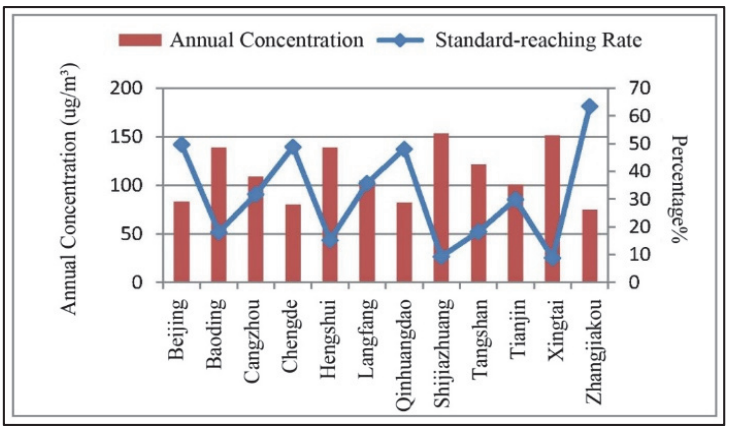


Fig. 2. Standard-reaching rate of PM_{10}

Although some cities, such as Beijing, Chengde City, Qinhuangdao City and Zhangjiakou City, don't have high concentrations of PM_{10} from the perspective of annual average, the other cities except Zhangjiakou City have low standard-reaching rate.

The standard-reaching rate of Zhangjiakou City is higher than 60% and the highest is 63.51%. In addition, the standard-reaching rate of all other cities is lower than 60%. Among them, the lowest rate of PM_{10} is Xingtai City whose standard-reaching rate is only 8.84%, and the standard-reaching rate of Shijiazhuang is 9.30%, which belongs to the low standard-reaching index. This also indicates that the content of PM_{10} in Beijing-Tianjin-Hebei region is very high in 2017.

(3) Standard-reaching rate of SO_2

The Fig. 3 shows the annual average and standard-reaching rate of SO_2.

When the standard-reaching rate of SO_2 is analyzed, the secondary concentration limit is used to screen the concentration limit of the sample data, which is no longer the same as that of the particulate matter. The secondary concentration limit of SO_2 is as high as 60μg /m³, and most of the concentrations of SO_2 are low and lower than the first concentration limit of 20μg /m³. Therefore, the first concentration limit value of SO_2 is analyzed in studying the standard-reaching rate of SO_2, and the calculation method is the same as above (Zhang et al., 2015c).

For the pollution of SO_2, the most serious city is Tangshan with the annual average of 38.7μg /m³, followed by Xingtai with the annual average of 38.00μg /m³. The lowest pollution concentration of SO_2 is Beijing. In 2017, the standard-reaching rate of SO_2 in Beijing, Chengde, Hengshui, Langfang and Zhangjiakou in Beijing-Tianjin-Hebei region is higher than 60%. Bejing has the highest rate of 93.24%. The rate of other cities is lower than 60%, and Tangshan has the lowest rate of 9.42%.

(4) Standard-reaching rate of NO_2

For the analysis of standard-reaching rate of NO_2 the first and secondary concentration limits of NO_2 are 40μg /m³, so we only need to screen out the concentration of NO_2 lower than 40μg /m³ for processing the sample data (Zhang et al., 2015b). The result of the analysis is shown in Fig. 4.

It can be seen from Fig.4 that in addition to Zhangjiakou and Chengde, the average annual concentration of NO_2 of other cities of Beijing-Tianjin-Hebei region has little difference and is between 40-50μg /m³. The lowest rate is only 11.63% in Tangshan, which indicates that the pollution of NO_2 is serious. Zhangjiakou City is "unique" with the highest standard rate of NO_2 reaching 91.09%.

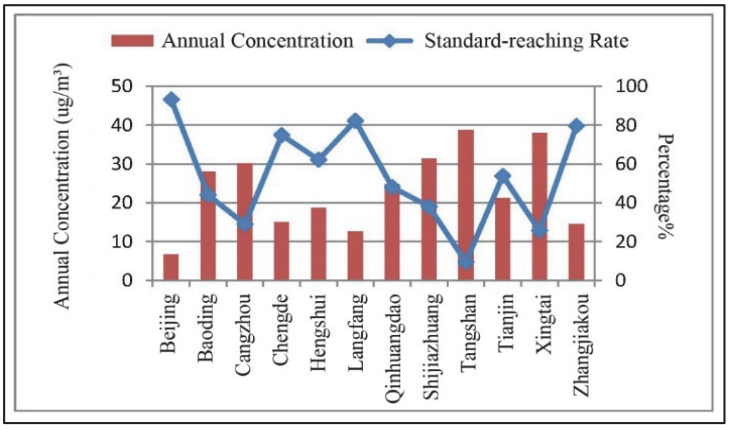


Fig. 3. Standard-reaching rate of SO_2

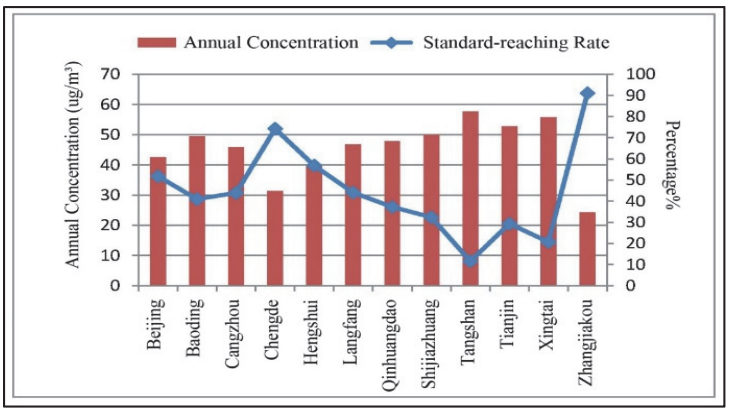


Fig. 4. Standard-reaching rate of NO_2

3.2. Air quality assessment of cities in Beijing-Tianjin-Hebei region

The air quality in Beijing-Tianjin-Hebei region is mainly assessed by AQI, that is, when AQI is less than or equal to 100, it achieves excellent and good rate. Fig. 5 is a statistical chart of excellent and good rate and pollution days in Beijing-Tianjin-Hebei region:

The excellent and good rate of air quality in Zhangjiakou, Chengde and Qinhuangdao is higher than 80%, and that of other cities is lower than 80%. The excellent and good rate of air quality of Shijiazhuang City, Handan City, Baoding City, Hengshui City and Xingtai City is even lower than 60%. This averages that these cities are exposed to varying degrees of air environmental pollution in half of the year. It can be seen from the distribution map of air quality levels in Beijing-Tianjin-Hebei region that Chengde City and Zhangjiakou City have the most "excellent" indexes, and the proportion of the days to reach the "good" level is 61%. In Baoding city, Handan city, Hengshui city, Shijiazhuang city and Xingtai city, the proportion of days with light pollution is higher than 25%, which is 29%, 39%, 36%, 32%, 39% respectively, and the proportion of days with moderate pollution and heavy pollution is higher than 10%. It averages that the pollution in these five cities is serious, and 13 cities in Beijing-Tianjin-Hebei region all have pollution problems of different degrees. The pollution exists for a long time and has a large distribution range. The distribution map of air quality levels in Beijing-Tianjin-Hebei region is shown in Fig. 6.

3.3. Change law of air quality

(1) Seasonal change law of air quality.

Seasonal change of air quality is obvious and regular. Generally, the concentration of $PM_{2.5}$ is the highest in winter and lowest in summer. The main reason for the high concentration of $PM_{2.5}$ in winter is that coal heating is not conducive to air diffusion, the secondary aerosol formation plays a decisive role too in cold season. Of course, traffic emissions can also be considerable contributors in all seasons of the year, and wood heating is less used in this area. In summer, the significant decrease of concentration of $PM_{2.5}$ is related to the absence of heating. In addition, the large-scale deposition of particulate matter and the oceanic monsoon climate have a great influence on air pollution (Gao and Zhang, 2014).

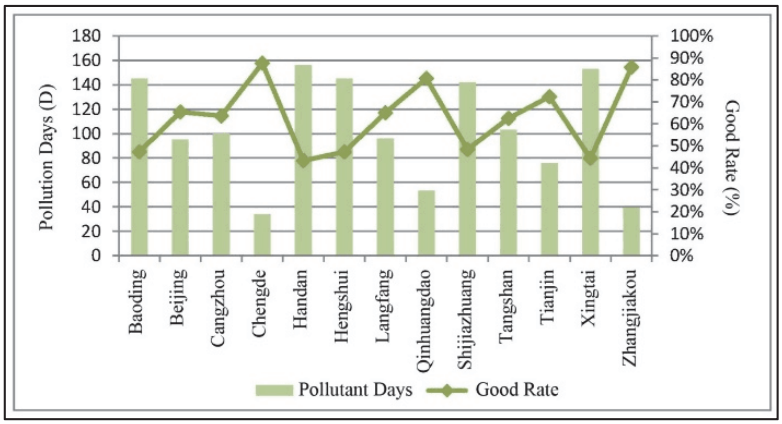


Fig. 5. Statistics of excellent and good rate and pollution days in Beijing-Tianjin-Hebei region

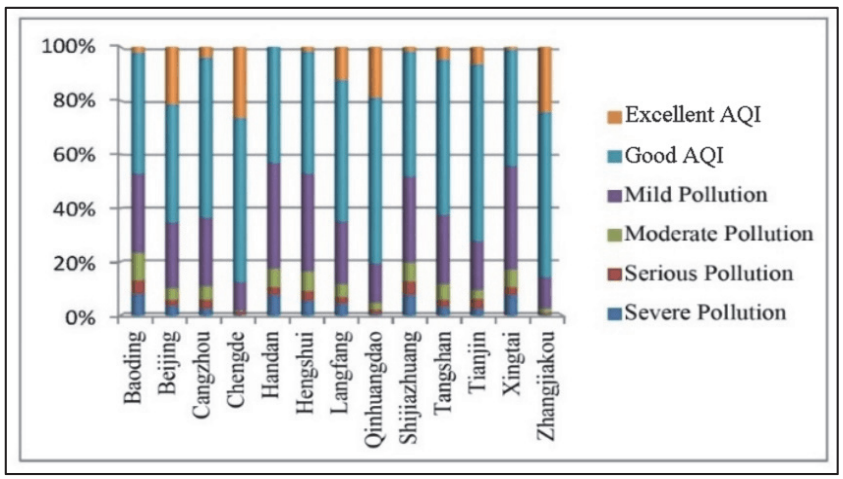


Fig. 6. Percent contribution of air quality levels in Beijing-Tianjin-Hebei region

The Beijing-Tianjin-Hebei region has the highest annual average concentration, and half of the 10 cities with the most serious haze in China are Baoding, Xingtai, Shijiazhuang, Handan and Hengshui. The climatic characteristics of Beijing-Tianjin-Hebei region are also one of the reasons for this phenomenon. Weak wind and lower atmospheric boundary layers make it easy for aerosols to accumulate (Che et al., 2014; Zhang et al., 2014). The Fig. 7 is seasonal spatial distribution map of $PM_{2.5}$.

(2) Monthly change of air quality

The highest AQI is from January to March and from October to December in Beijing-Tianjin-Hebei region, while the lowest AQI is from April to September. The peak is in January, February, November and December. The air quality is strongly associated with these months for they are the heating time. The Fig. 8 is the trend of monthly average concentration values of AQI.

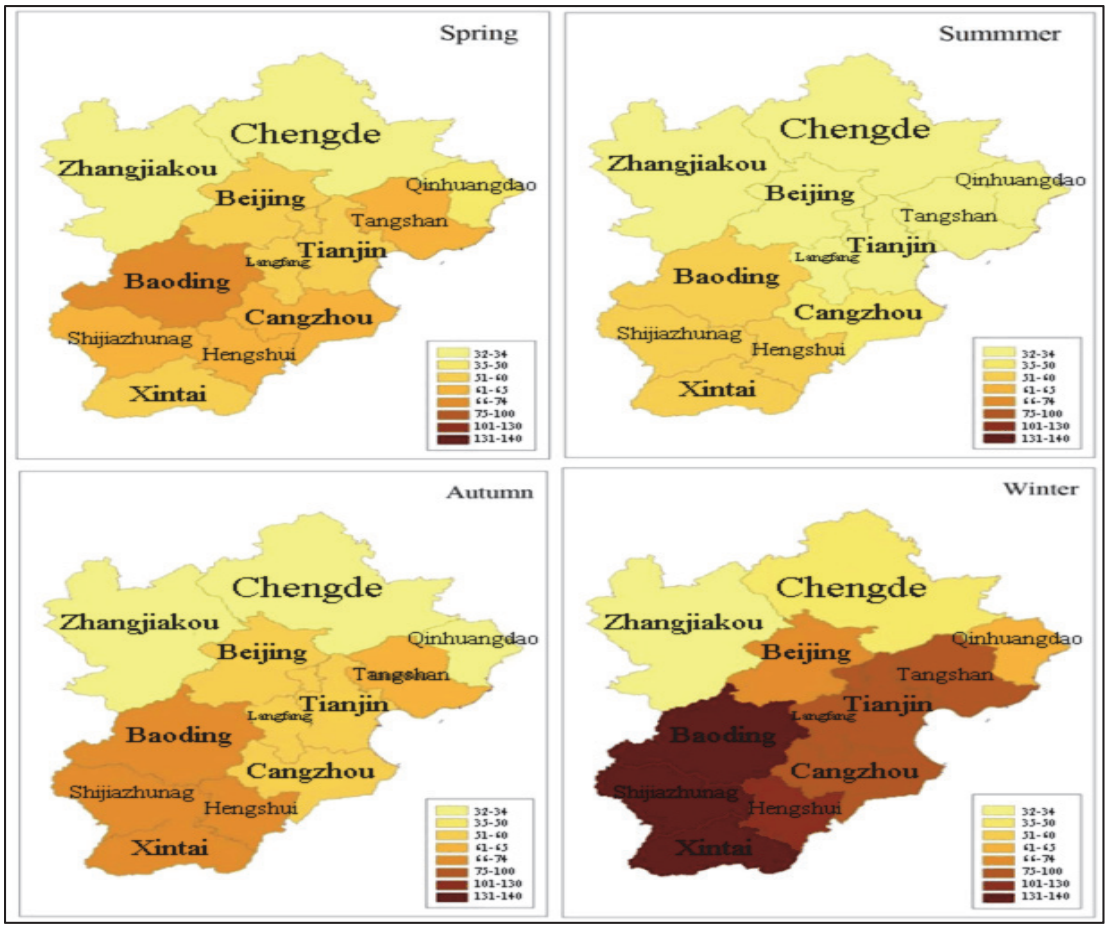


Fig. 7. Seasonal spatial distribution map of $PM_{2.5}$ in Beijing-Tianjin-Hebei Region

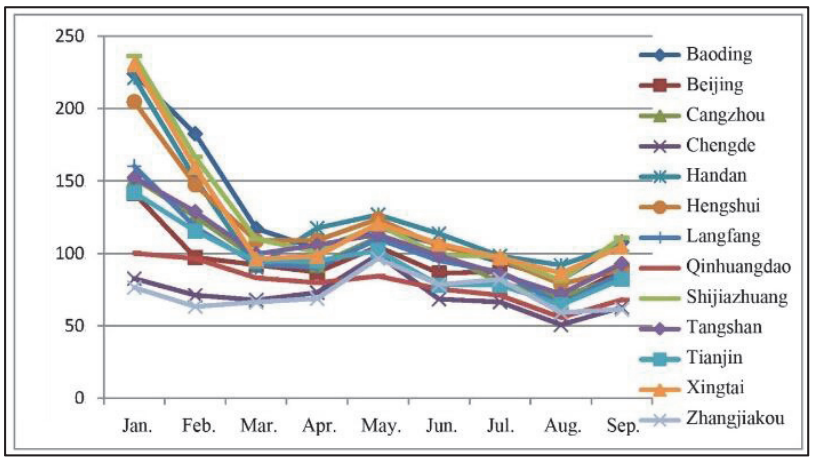
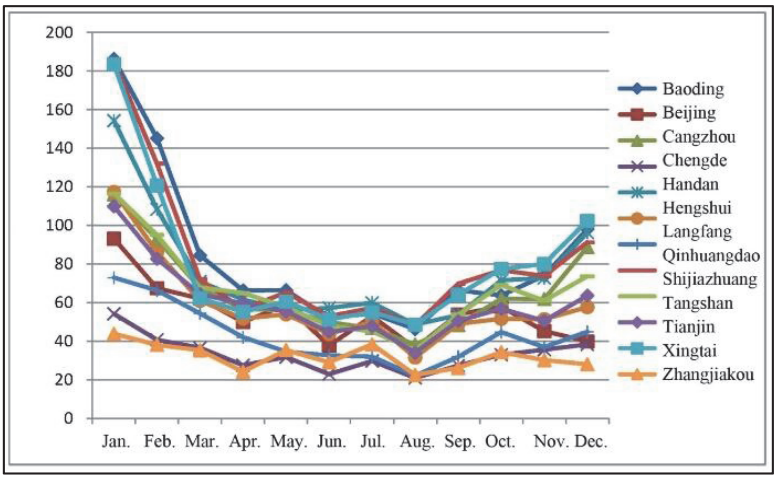


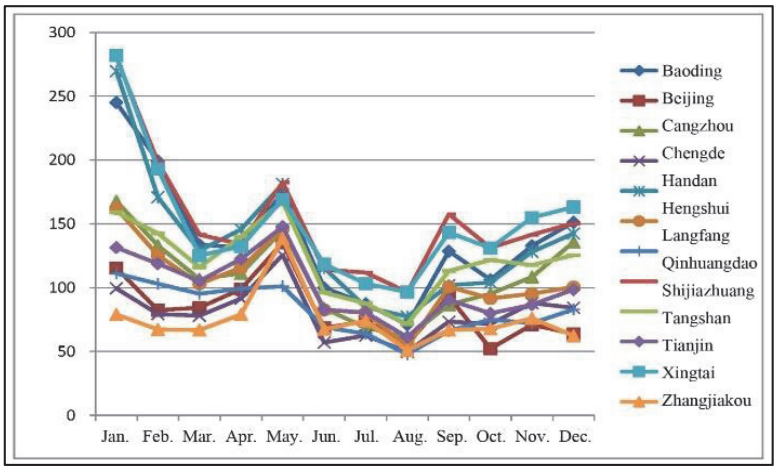
Fig. 8. Trend of monthly average concentration values of AQI

3.4. Monthly average concentration trend of major pollutants

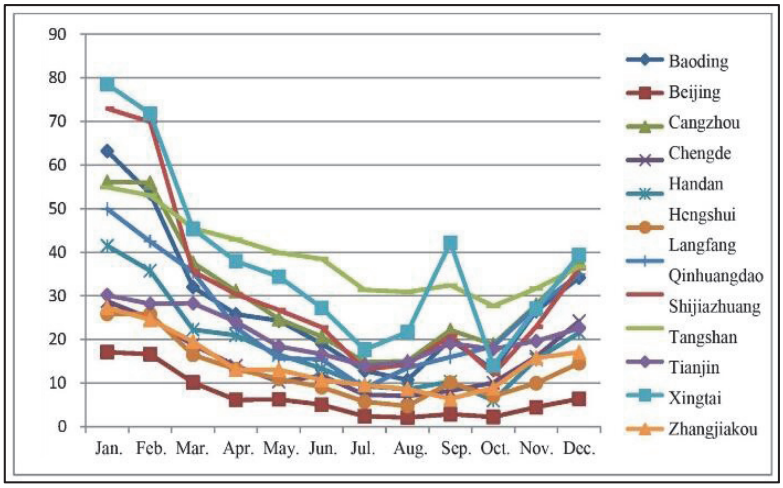
Monthly average concentration trend of $PM_{2.5}$ is shown on Fig. 9a, monthly average concentration trend of PM_{10} is shown on Fig. 9b, monthly average concentration trend of SO_2 is shown on Fig. 9c, monthly average concentration trend of NO_2 is shown on Fig. 9d, monthly average concentration trend of CO is shown on Fig. 9e, monthly average concentration trend of O_3 is shown on Fig. 9f.



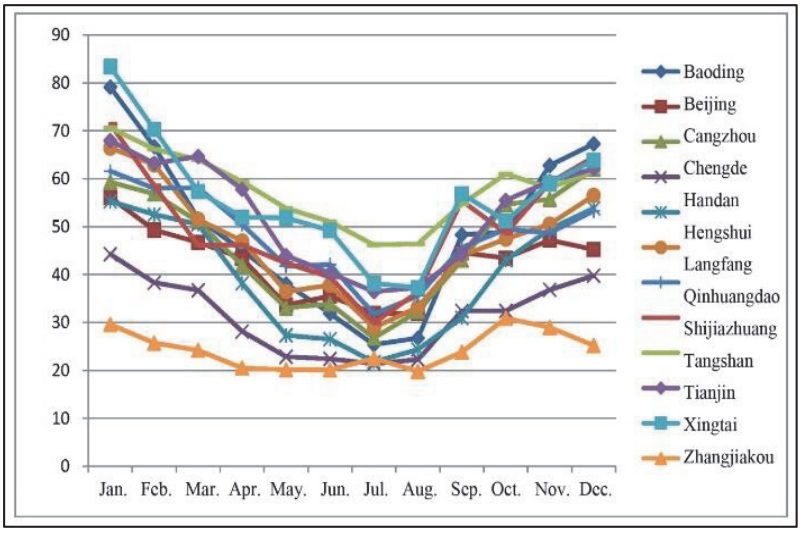
(a)



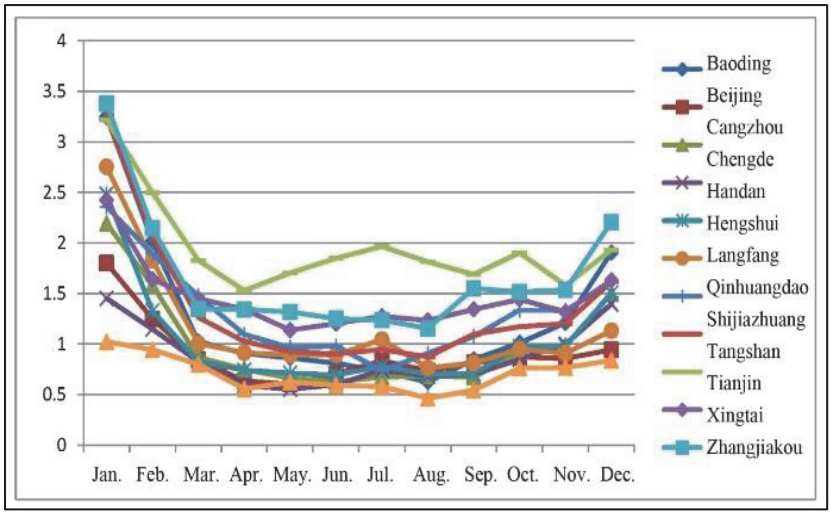
(b)



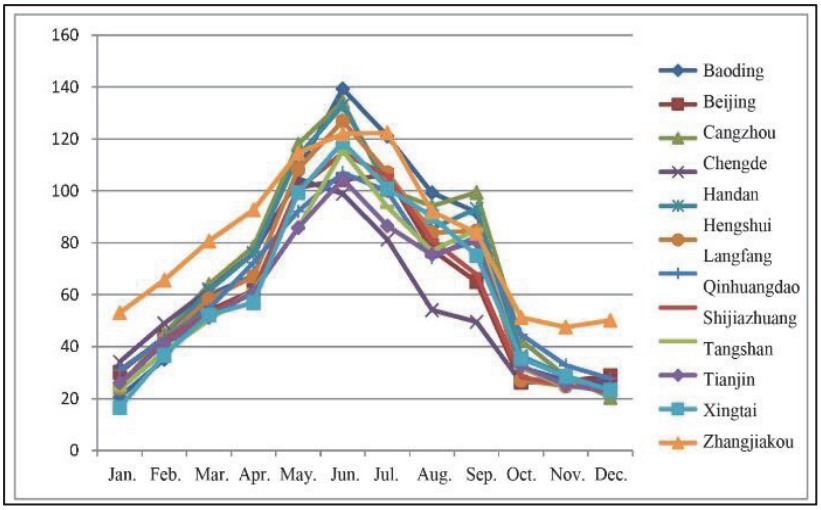
(c)



(d)



(e)



(f)

Fig. 9. Monthly average concentration trend of major pollutants: (a) Change trend of monthly average concentration value of $PM_{2.5}$; (b) Change trend of monthly average concentration value of PM_{10}; (c) Change trend of monthly average concentration value of SO_2; (d) Change trend of monthly average concentration value of NO_2; (e) Change trend of monthly average concentration value of CO; (f) Change trend of monthly average concentration value of O_3

3.5. $PM_{2.5}$ Prediction Model

This study divides the data of 2017 into four stages: spring (from January to March), summer (from April to June), autumn (from July to September) and winter (from October to December). The prediction study of $PM_{2.5}$ is carried out in spring, summer, autumn and winter and the data is trained by BP neural network model. The convergence speed and convergence precision, as well as the fitting degree of the prediction are analysed.

The result of BP neural network prediction model is analysed from the aspects: (1) error value Mean-Square Error (MSE) of BP neural network, in which MSE after simulation is listed and compared; (2) recombination coefficient of the output data of BP neural network; (3) analysis of the prediction result, namely, draw the prediction result and real result into a graph for comparison (Wang et al., 2013; Li et al., 2015).

3.6. $PM_{2.5}$ prediction model in spring

The MSE output from BP neural network is the main evaluation of network training performance index, and the expected value of error accuracy of the model is set to 0.001. The result of network training produced by the spring model is shown in the Fig. 10a, the minimum output error value of BP neural network is 0.0026852, and the number of iterations is 20, the minimum output error value after simulation is 0.002303, and the number of iterations is only 1. The recombination coefficient of the output data of BP neural network in spring is shown in the Fig. 10b, it represents the consistency between predicted results and target expectations.

The predicted value of the trained BP neural network is compared with the real value in the Fig. 10c, It is easy to see that the BP neural network has high accuracy in the prediction of $PM_{2.5}$.

3.7. $PM_{2.5}$ prediction model in summer

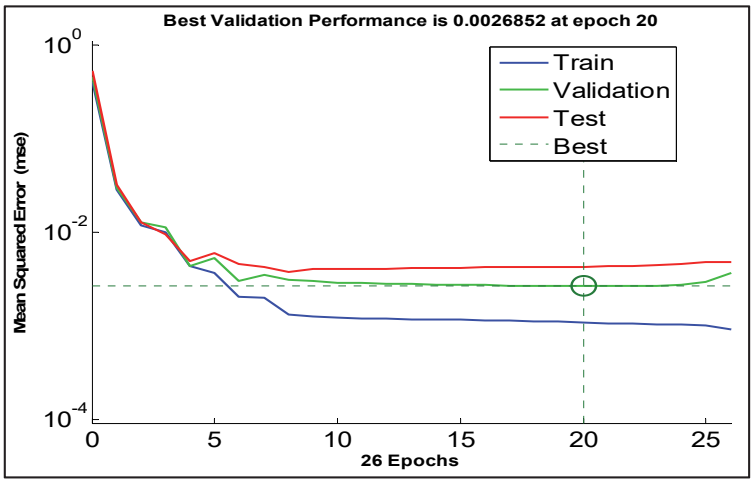
The result of network training produced by the summer model is shown in the Fig. 11a, the minimum output error value of BP neural network is 0.0071208, and the number of iterations is 9, the minimum output error value after simulation is 0.0041481, and the number of iterations is only 0. The recombination coefficient of the output data of BP neural network in summer is shown in the Fig. 11b. The comparison of real value and predicted value in summer is shown in the Fig. 11c.

3.8. $PM_{2.5}$ prediction model in autumn

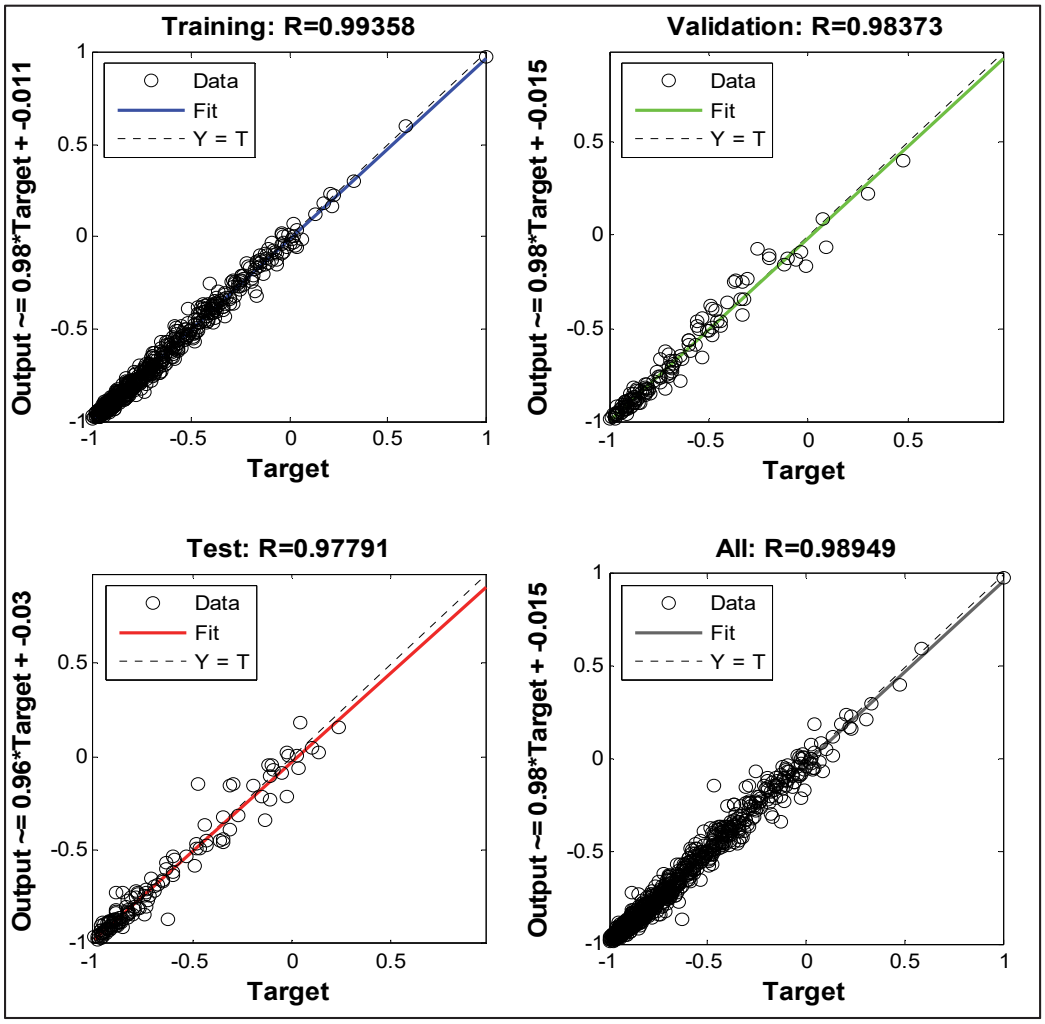
The result of network training produced by the autumn model is shown in the Fig. 12a. The minimum output error value of BP neural network is 0.0084171, and the number of iterations is 11, the minimum output error value after simulation is 0.005480, and the number of iterations is only 0. The recombination coefficient of the output data of BP neural network in autumn is shown in the Fig. 12b. The comparison of real value and predicted value in autumn is shown in the Fig. 12c.

3.9. $PM_{2.5}$ prediction model in winter

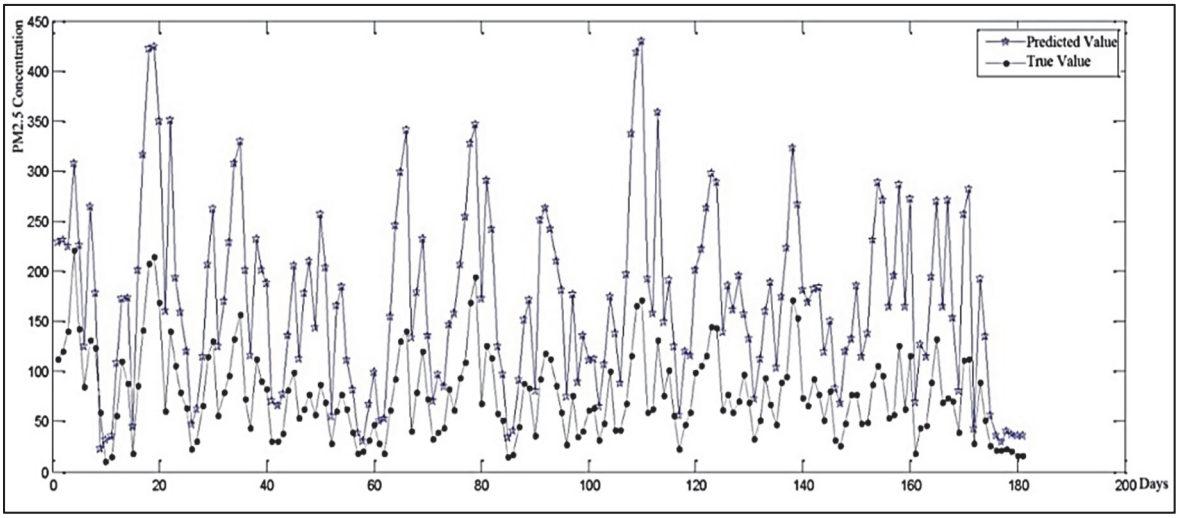
The result of network training produced by the winter model is shown in the Fig. 13a. The minimum output error value of BP neural network is 0.0094386, and the number of iterations is 4, the minimum output error value after simulation is 0.0054804, and the number of iterations is only 0. The recombination coefficient of the output data of BP neural network in winter is shown in the Fig. 13b. The comparison of real value and predicted value in winter is shown in the Fig. 13c.



(a)

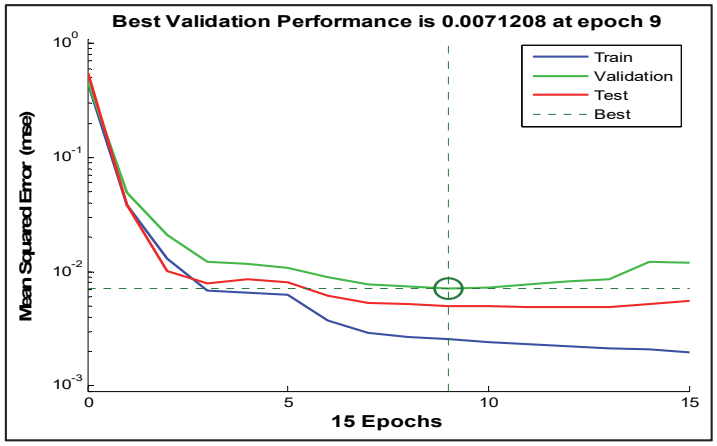


(b)

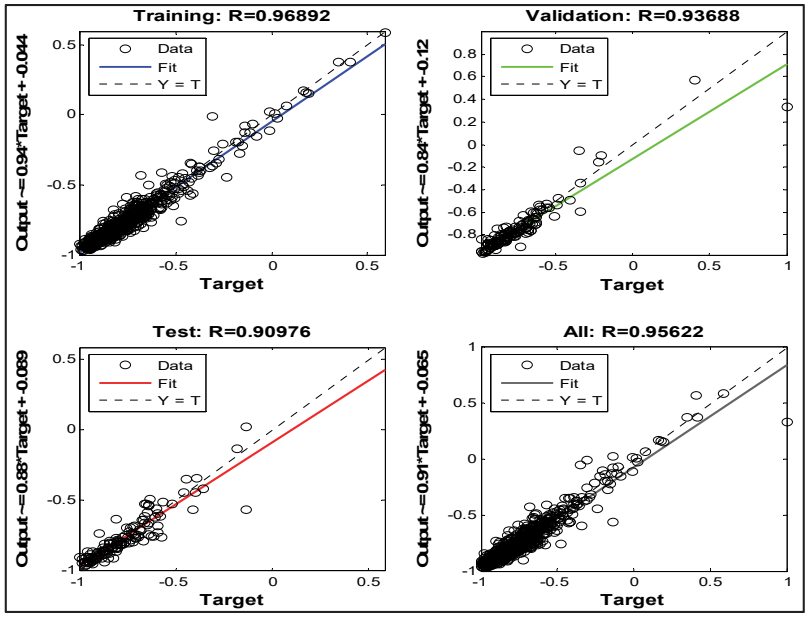


(c)

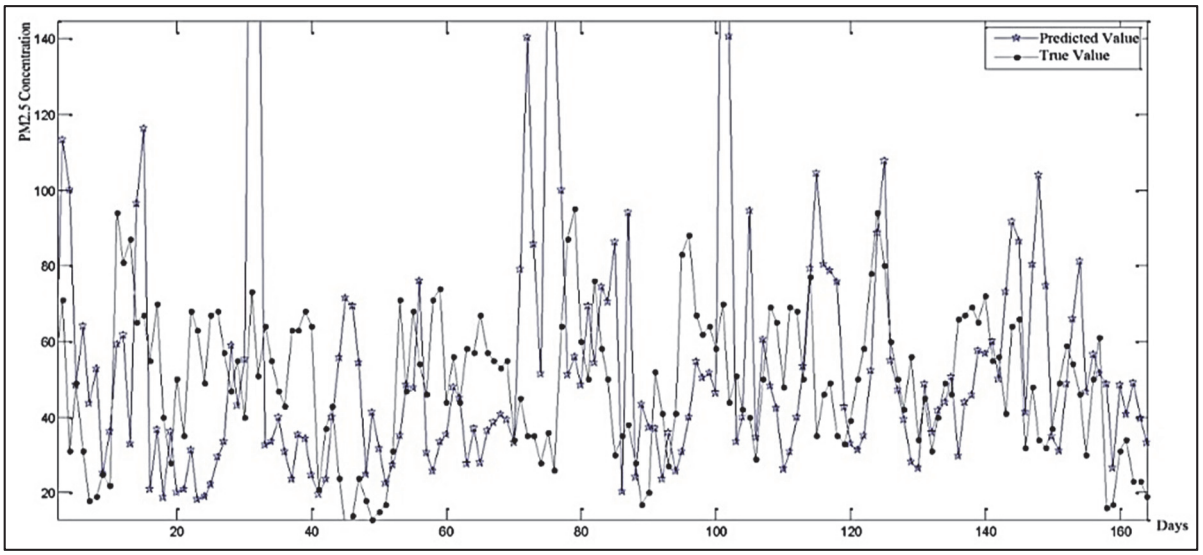
Fig. 10. $PM_{2.5}$ prediction model in spring: (a) MSE of spring model; (b) Recombination coefficient of BP neural network in spring; (c) Comparison of real value and predicted value in spring



(a)

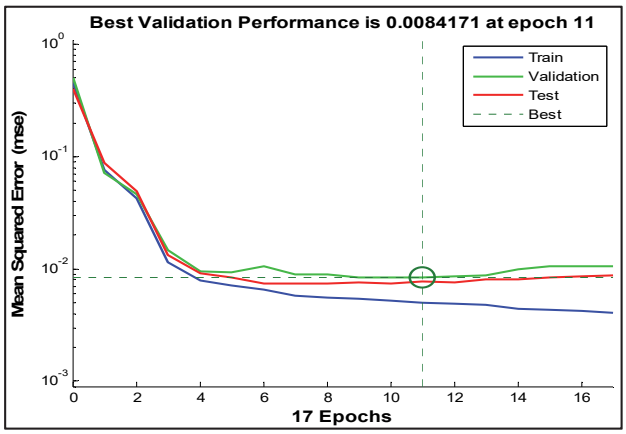


(b)

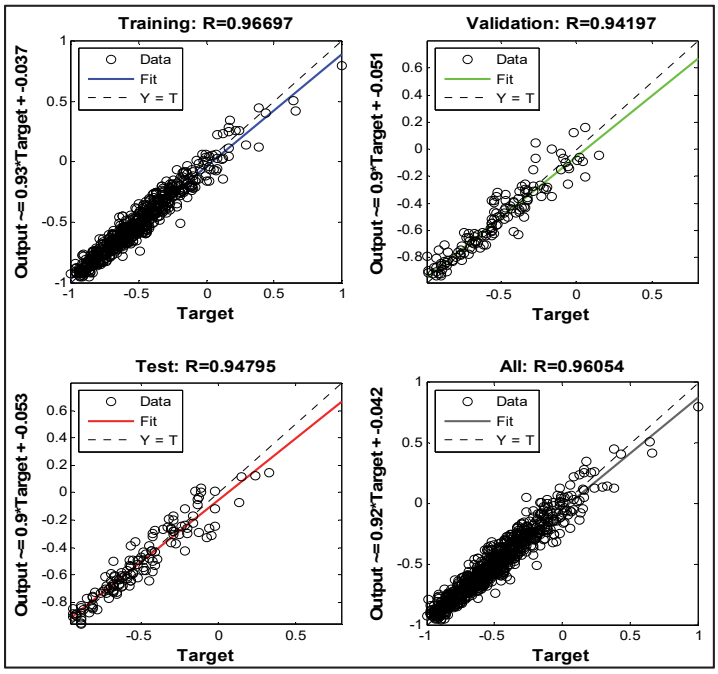


(c)

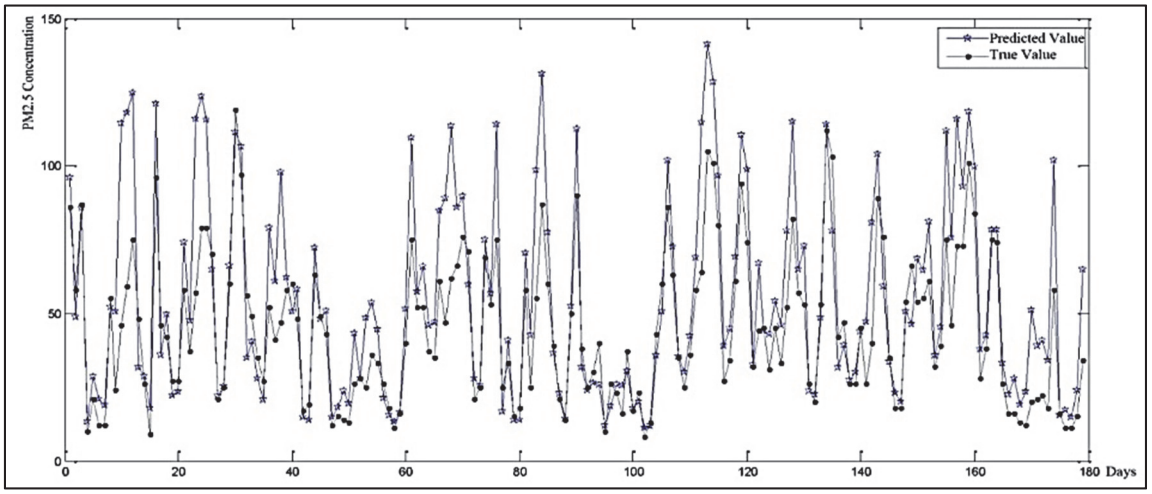
Fig. 11. $PM_{2.5}$ prediction model in summer: (a) MSE of summer model, (b) Recombination coefficient of BP neural network in summer, (c) Comparison of real value and predicted value in summer



(a) MSE of autumn model

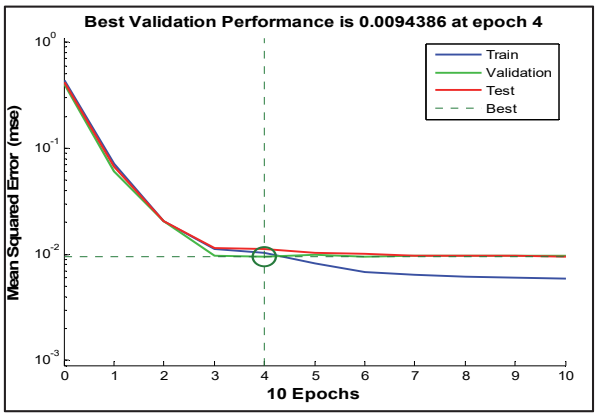


(b)

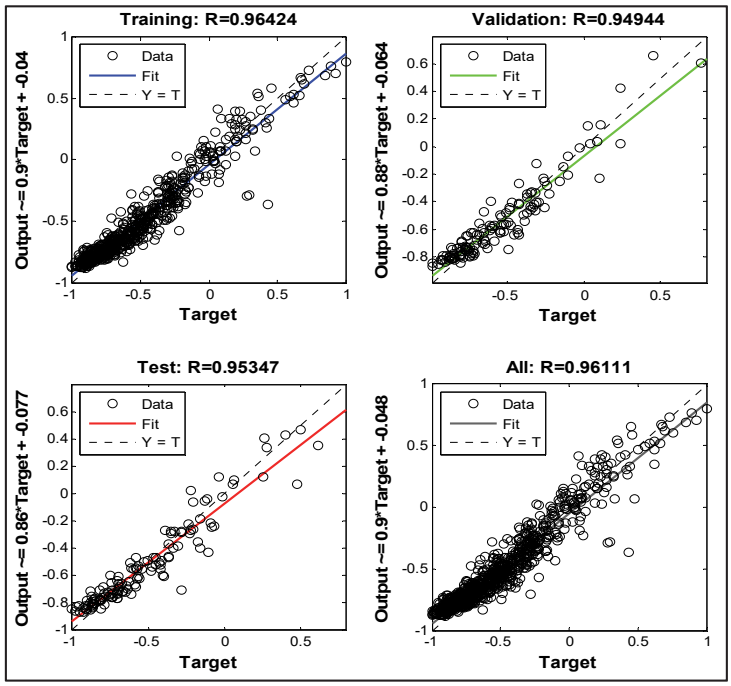


(c)

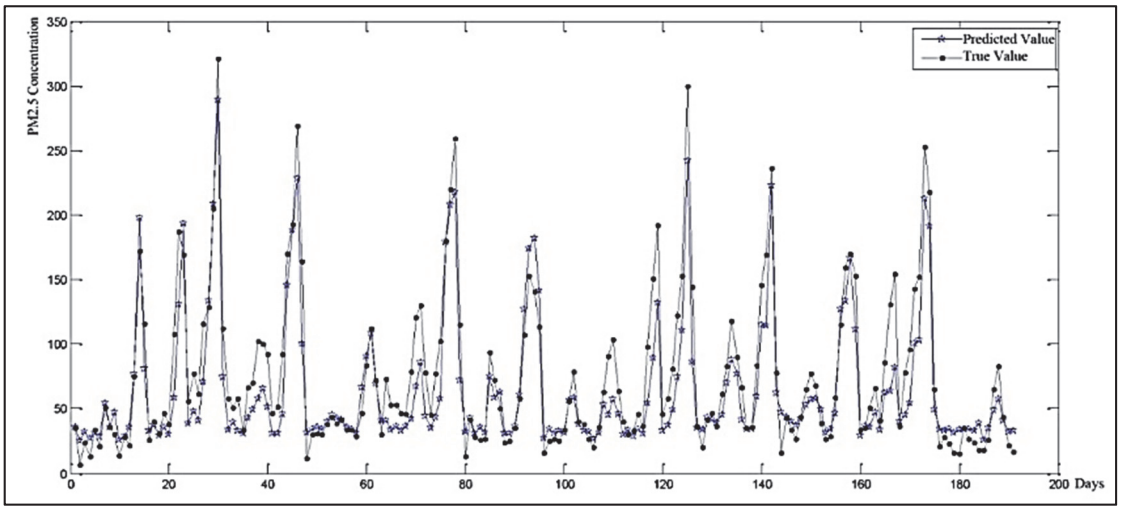
Fig. 12. $PM_{2.5}$ prediction model in autumn: (a) MSE of autumn model, (b) Recombination coefficient of BP neural network in autum, (c) Comparison of real value and predicted value in autumn



(a)



(b)



(c)

Fig. 13. $PM_{2.5}$ prediction model in winter: (a) MSE of winter model, (b) Recombination coefficient of BP neural network in winter, (c) Comparison of real value and predicted value in winter

4. Conclusions

This study applies BP neural network to the air quality prediction in Beijing-Tianjin-Hebei region in China, and carries out simulation test to improve the accuracy of the model. The daily data of 13 cities in Beijing-Tianjin-Hebei region are collected and sorted and each sample contains the influence of multiple air factors; The sample data is evaluated by AQI, and the space-time distribution is analyzed by ArcGIS software, and the space-time distribution of air factors is also analyzed; Matlab software is used to establish the BP neural network model. In the establishment of the model, it is necessary to divide, import and transpose the sample data, as well as to normalize the input data and inversely normalize the predicted result data; The error and fitting degree of the BP neural network are analyzed, and the model is confirmed to be the optimal model.

This study draws conclusions as follows:

(1) AQI can be used to evaluate and analyze the space-time distribution of air data directly. (2) In the study of standard-reaching rate of major pollution factors, Zhangjiakou City has the best standard-reaching rate and the lowest annual average concentration among the 13 cities in Beijing-Tianjin-Hebei region, which indicates that air quality of Zhangjiakou City is the best among the 13 cities. (3) The BP neural network model has high fitting degree and prediction accuracy in the air pollution prediction experiment in Beijing-Tianjin-Hebei region in 2017. The error data of predicted value and real value are analyzed and the accuracy rate is about 75%.

In the research process of using BP neural network model to predict haze data, some research achievements have been obtained through the collection of sample data and the support of theory. However, due to the limited time and research conditions, as well as the complexity of impact factors of air quality, the research contents in this study are not rich, which needs perfect and in-depth study in the future. Further studies are as follows:

(1) More and more complete data are needed to support the model. (2) The main model used in this study is BP neural network, and the data should be experimented with different models to get the optimal model, which can better support the theoretical results.

(3) In the aspect of data processing, with the increase of data, different methods should be selected and combined to make the accuracy of the model more accurate.

Acknowledgements

This research was supported by the National Social Science Foundation of China (No.16BGL145). This research was jointly supported by the National Natural Science Foundation of China (No. 71371128).

References

Blanchard C.L., Tanenbaum S., (2006), Weekday/Weekend differences in ambient air pollutant concentrations in atlanta and the south-eastern United States, *Air Repair*, **56**, 271-283.

Chai F., Gao J., Chen Z., Wang S., Zhang Y., Zhang J., Zhang H., Yun Y., Ren C., (2014), Spatial and temporal variation of particulate matter and gaseous pollutants in 26 cities in China, *Journal of Environmental Sciences*, **26**, 75-82.

Che H., Xia X., Zhu J., Li Z., Dubovik O., (2014), Column aerosol optical properties and aerosol radiative forcing during a serious haze-fog month over North China Plain in 2013 based on ground-based sunphotometer measurements, *Atmospheric Chemistry and Physics*, **14**, 2125-2138.

Devi K.K., Kaskaoutis D.G., San L.H., (2016), Overview of atmospheric aerosol studies in Malaysia: Known and unknown, *Atmospheric Research*, **182**, 302-318.

Fang X., Zou B., Liu X., (2016), Satellite-based ground $PM_{2.5}$ estimation using timely structure adaptive modeling, *Remote Sensing of Environment*, **186**, 152-163.

Gao Y., Zhang M.G., (2014), Numerical simulation of a heavy fog-haze episode over the North China Plain in January 2013, *Climatic and Environmental Research*, **19**, 140-152.

Guinot B., Roger J.C., Cachier H., Pucai W., Jianhui B., Tong Y., (2006), Impact of vertical atmospheric structure on Beijing aerosol distribution, *Atmospheric Environment*, **40**, 5167-5180.

Ji D., Wang Y., Wang L., Chen L., Hu B., Tang G., Xin J., Song T., Wen T., Sun Y., Pan Y., Liu Z., (2012), Analysis of heavy pollution episodes in selected cities of northern China, *Atmospheric Environment*, **50**, 338-348.

Lakshmipathi A.N., Battula B.P., (2018), Deep convolutional neural networks for product recommendation, *Ingénierie des Systèmes d'Information*, **23**, 161-172.

Li Y., Cheng N., Zhang D., (2015), PM(2.5) background concentration at different directions in Beijing in 2013, *Environmental Science*, **36**, 4331-4339.

Liu Z., Hu B., Wang L., Wu F., Gao W., Wang Y., (2015), Seasonal and diurnal variation in particulate matter (PM 10, and PM 2.5) at an urban site of Beijing: analyses from a 9-year study, *Environmental Science & Pollution Research*, **22**, 627-642.

Ma Z.Q., Zhao X.J., Meng W., Meng Y.J., He D., Liu H.Y., (2012), Comparison of influence of fog and haze on visibility in Beijing, *Research of Environmental Sciences*, **25**, 1208-1214.

Markowicz K.M., Pakszys P., Ritter C., (2016), Impact of North American intense fires on aerosol optical properties measured over the European Arctic in July 2015, *Journal of Geophysical Research-Atmospheres*, **121**, 14487-14512.

Miao S., Chen F., LeMone M.A., Tewari M., Li Q., Wang Y., (2009), An observational and modelling study of characteristics of urban heat island and boundary layer structures in Beijing, *Journal of Applied Meteorology & Climatology*, **48**, 484-501.

Motallebi N., Tran H., Croes B.E., Larsen L.C., (2003), Day-of-week patterns of particulate matter and its chemical components at selected sites in California, *Air Repair*, **53**, 876-888.

Oprea M., Dunea D., Liu H.Y., (2017), Development of a knowledge based system for analyzing particulate matter air pollution effects on human health, *Environmental Engineering and Management Journal*, **16**, 669-676.

Sánchez-Escalona A.A., Góngora-Leyva E., (2018), Artificial neural network modeling of hydrogen sulphide gas coolers ensuring extrapolation capability, *Mathematical Modelling of Engineering Problems*, **5**, 348-356.

Singh B.K., Singh A.K., Singh V.K., (2018), Exposure assessment of traffic-related air pollution on human health - a case study of a metropolitan city, *Environmental Engineering and Management Journal*, **17**, 335-342.

Song J., Cheng T., Xie Z.Q., Miao Q., (2012), Impact on spatio-temporal variation of fog and haze days due to rapid urbanization in Jiangsu, *Scientia Meteorologica Sinica*, **32**, 275-281.

Sumita K., Ribu C., Sahidul I., (2016), Regional simulation of aerosol radiative effects and their influence on rainfall over India using WRFChem model, *Atmospheric Research*, **182**, 232-242.

Wang Q., Chen X., He G.-l., Lin S.-l., Liu Z., Xu D.-q., (2013), Study on characteristics of elements in $PM_{2.5}$ during haze-fog weather in winter in urban Beijing, *Spectroscopy and Spectral Analysis*, **33**, 1441-1445.

Wang Y., Ying Q., Hu J., Zang H., (2014), Spatial and temporal variations of six criteria air pollutants in 31 provincial capital cities in China during 2013-2014, *Environment International*, **73**, 413-422.

Xiao Q., Ma Z., Li S., Liu Y., (2015), The impact of winter heating on air pollution in China, *PLOS ONE*, **10**, e0117311.

Zhang H., Wang Y., Hu J., Ying Q., Hu X.M., (2015), Relationships between meteorological parameters and criteria air pollutants in three megacities in China, *Environmental Research*, **140**, 242-254.

Zhang Q., Quan J., Tie X., Li X., Liu Q., Gao Y., Zhao D., (2015a), Effects of meteorology and secondary particle formation on visibility during heavy haze events in Beijing, China, *Science of the Total Environment*, **12**, 502-578.

Zhang Q., Duan F.K., He K.B., (2015b), Organic nitrogen in $PM_{2.5}$ in Beijing, *Frontiers of Environmental Science & Engineering*, **9**, 1004-1014.

Zhang Y.W., Zhang X.Y., Zhang Y.M., (2015c), Significant concentration changes of chemical components of PM1 in the Yangtze River Delta area of China and the implications for the formation mechanism of heavy haze-fog pollution, *Science of the Total Environment*, **538**, 7-15.

Zhang X.H., Liu L.Y., Chen X.H., (2014), Characteristics of fog/haze and its influencing factors in Changsha, *Chinese Journal of Environmental Engineering*, **8**, 3361-3366.