Causes of PM2.5 and its mathematical model based on carbon analysis

Accepted 29th January, 2018

ABSTRACT

In recent years, haze problem has been a serious issue in central and eastern regions of China. It is very important to comprehensively understand the haze phenomenon in the central and eastern part of China by analyzing the mechanism of PM2.5 from the perspective of physical process and establishing the mathematical model of PM2.5 microparticle forming physical process. In this paper, the mathematical model of TC-PM2.5 causal relationship is proposed by the cause of PM2.5 physical formation process. The strong correlation between Tc and PM2.5 concentration has been confirmed by PM2.5 content and the impact of the factors of the scatter plot. By comparing with the results of multiple linear regression models, it is found that the mathematical model of PM2.5 is better. The causal relationship between TC-PM2.5 proposed in the analysis of PM2.5 physical formation process is proved by data translation analysis. In the PM2.5 physical formation process, the effect of TC is amplified, while the magnification is about 5 times and carbon particles play a true "source" role in PM2.5 formation.

Keywords: PM2.5, TC, cause analysis mathematical model, multiple linear regression models, micro structure.

INTRODUCTION

In recent years, haze weather has often occurred in Beijing and central and eastern regions of China. This is another major air pollution incident after 1873 London haze incident and photochemical smoke event in Los Angeles in the 1950s has caused the community great concern as regards the air quality and environmental pollution issues. Compared with Europe and the United States and other countries, China haze structure and generation mechanism are both common and special. To study the mechanism of the haze in China, it is necessary to start from the particularity, analyze the deep-seated causes, and then put forward the coping strategies, which is the most important problem to be solved at present.

High concentration of fine particulate matter (PM2.5) pollution is considered to be the root cause of the formation of haze, mainly by the direct discharge of primary particles and gaseous pollutants through the transformation - aging process gradually formed. At present, the research on origin and cause of PM2.5 particulate matter are mainly by the "attribution analysis" (source analysis) and chemical reaction process methods.

Attribution analysis is an analysis of the formation of environmental pollution by environmental scientists in the 1960s. It is based on the principle of mass balance or statistical method to obtain the contribution of different sources to the particulate matter by analyzing the composition of the sample of the source sample and the receptor (atmospheric particulate matter), which has played a great role in solving the atmospheric pollution, especially the primary emission pollution (John and Judith, 2001; Guo-Liang et al., 2009). However, due to the limitations of the traditional attribution analysis theory, mainly for a source of emissions, the secondary source pollutant resolution is poor. A large number of recent studies have shown that PM secondary fine particulate matter concentration is the main reason for the frequent haze in China (Yang et al., 2015; Sun et al., 2013; Yele et al., 2016, 2015). Its formation contains a complex physical and
chemical process; the traditional attribution source analysis is difficult to objectively describe the true cause of PM2.5. Studying the chemical reaction and its catalytic mechanism in the formation of PM2.5 has become an important choice for environmental scientists.

He (2014) proposed a new Chinese haze mechanism that the SO$_2$ conversion to sulphate granules was promoted under the condition of mineral particulate matter and NO$_2$ coexisting condition by observing the data in the Beijing-Tianjin-Hebei region of 5 times heavy haze and laboratory smoke box simulation experiment in 2013. After more scholars (Wang et al., 2014; Cheng et al., 2016; Zheng et al., 2015a, 2015b) from the observation data, model simulation and other aspects of further exploration, verification, the new haze chemical mechanism in China of a unique high-mineral dust surface, high humidity conditions and NO$_2$ as the oxidant sulfate heterogeneous conversion was formation. For heterogeneous conversion, in addition to chemical reactions, the effective reaction surface area (particle size and surface structure), particulate water content, PM2.5 hygroscopicity and other physical environment is also critical and provides the necessary conditions for the heterogeneous reaction (Liu et al., 2015; Ye and Chen, 2013; Xuexi et al., 2017). Porous, small particle size provides more reaction interface to promote the reaction rate.

Li et al. (2010) studied the heterogeneous conversion of NO$_2$ to calcium carbonate. It was found that the conversion rate of NO$_2$ under high humidity was about 6 times higher than that in dry condition. The current research on PM is mainly focused on the chemical process and ignores the "chemical reaction of the micro-physical environment," the key issue.

Under normal circumstances, the formation of haze particles need to produce the relevant chemical reaction of the micro-physical environment which will experience the cohesion of nuclear particles, condensed particles and droplets particles, three key links. In this study, the correlation and causality of the key components in the three states of "condensed nucleus", "condensed state" and "droplet state" were analyzed to determine the main cause of haze in Beijing. From the point of view of condensed nucleus, condensed state and droplet state, the condensed nucleus plays the most important role. The condensation of water in condensed particles plays an important role, and the influence of humidity is obvious. A large number of water-soluble substances, such as sulfate, nitrate and ammonium ions, etc., their content and the amount of pollutants is associated with the atmosphere. In the case of relatively stable levels of contaminants in the atmosphere, the content of these components in PM2.5 is largely dependent on the amount of condensed particles (Ye and Chen, 2013). Based on this magnitude of relationship, and the logical causal relationship they reflect, the condensed kernel is the most critical variable that determines the number of PM2.5, while the material components of the agglomerated nuclear particles are mainly composed of elemental carbon and organic carbon. Reasonably, it is possible to establish the cause analysis model of PM2.5 from the correlation between PM2.5 and its total carbon content.

In the process of the formation of haze particles, for the relatively high humidity situation, the condensed core can be soil dust, coal dust and other relatively small particles. Such as London's coastal climate and high humidity characteristics, making the presence of solid particles in the air can easily become the formation of water droplets of condensed nuclei. In the case of low relative humidity, it is necessary to coagulate the nucleus with strong water absorption, which can absorb the gaseous water molecules in the air to form microscopic water droplets and provide the "reaction dish" for the subsequent chemical reaction process of the formation of haze particles. The analysis of the photochemical contamination events in Los Angeles, for example, shows that the various types of hydrocarbons that are not fully combusted in the vehicle exhaust are the main source of this condensed nucleus. In this sense, it constitutes the absorption of other chemical pollutants in the atmosphere of the "reaction dish" from the condensate to the condensed particles.

Through the "micro-physical environment of chemical reaction" analysis, because most of the eastern part of China is a dry climate, it is not easy to directly form micro-water droplets, the need for condensed nucleus has strong water absorption in order to further the formation of haze, and this still did not form a targeted study. Therefore, it is important to analyze the mechanism of PM2.5 in China from the perspective of physical process and establish the mathematical model of PM2.5 micro-particles to form physical processes, which is of great significance to comprehensively understand the haze phenomenon in eastern China.

In this paper, the haze of Beijing area is taken as the research object and the correlation between the total content of PM2.5 and the key elements is analyzed specifically for the formation process. The mathematical model of PM2.5 genetic analysis is established to study on the mechanism of PM2.5 genesis, formation mechanism and its cohesive nuclei, in order to effectively inhibit the production of PM2.5 to provide theoretical guidance.

**RESULTS AND DISCUSSION**

**Correlation between PM2.5 and total carbon content and causal analysis**

The main components of the condensed nuclei in PM2.5 particles are elemental carbon and organic carbon, which are commonly referred to as "total carbon" in environmental analysis and are indicated by TC in this paper. In order to distinguish it from the concept of "carbon emission" in environmental science and engineering, this carbon content
In order to determine the correlation between PM2.5 content and various key factors, this project analyzed the PM2.5, TC, humidity, temperature and atmospheric pressure data of a monitoring point in Beijing from January, 1st 2017 to April, 10th 2017. The main influencing factor analysis can be achieved by scatter plot. Figure 1 shows the PM2.5 content and the impact of the scatter plot. It can be seen from the scatter plot that TC and PM2.5 content have a strong correlation, and have a statistically positive correlation, humidity; PM2.5 have a certain degree of correlation, while temperature and atmospheric pressure observation points are scattered, indicating that the correlation value of PM2.5 is not significant, not as an influencing factor. The aforementioned analysis shows that this paper has the rationality of TC as the condensed nucleus in the analysis part of PM2.5 physical formation process.

Accordingly, the quantitative relationship between the PM2.5 content in air and the various carbon particle contents can be analyzed objectively. As shown in Figure 2, the PM2.5 content, elemental carbon (EC) content, organic carbon (OC) and the total carbon content (TC) were taken as the ordinate at the sampling time as the abscissa. It can be seen from Figure 3 that the elemental carbon (EC) content, organic carbon (OC) and total carbon content (TC) have very high agreement with PM2.5 content. Observation of the details of the figure, PM2.5 content and total carbon content between the highest consistencies can be found. (Figure 2, due to the use of equipment failure, part of the collected data is blank) can be specifically for the PM2.5 content and TC content between the data correlation analysis and it is expected to find more in line with the actual situation of the law of change.

Further analysis of the relationship between PM2.5 content and carbon particle content in air and the linear regression equation can be used to obtain the data graph of Figure 3. Through the quantitative relationship shown in Figure 3, it can be found that there is a high correlation between the PM2.5 content and the carbon content in the air at the monitoring point; the correlation coefficient is about 85%, while the PM2.5 content is about 5 times of the TC content, or TC content of about 20% of the total PM2.5. PM2.5 content and TC content showed a very high correlation, it can be expected that there is a necessary causal relationship between them. However, this correlation analysis cannot give an answer as to whether the change in

Figure 1: PM2.5 content and the impact factor of the scatter plot.
PM2.5 content led to changes in TC content, or changes in TC content led to changes in PM2.5 content.

The project will be through the time axis to shift the data method to find the time sequence of the both changes. First, PM2.5 data and TC data are normalized, and then divided into four steps for analysis; in the first step, the sampled data in Figure 2 are subjected to multiple linear regression analysis. That is, the number of PM2.5 and TC using the same time data for linear regression analysis and the correlation was 0.823 (Figure 4). In the second step, the TC data in the sampled data in Figure 2 is shifted by one hour and subjected to multiple linear regression analysis. That is, the use of TC delayed one hour of data for linear regression analysis and the correlation was 0.859 (Figure 5). In the third step, the TC data in the sampled data in Figure 2 is shifted by two hours, and the blank data is removed and subjected to multiple linear regression analysis. That is, the use of TC delayed two hours of data for linear regression analysis and the correlation was 0.811 (Figure 6). In the fourth step, the TC data in the sampled data in Figure 2 is...
shifted by three hours, and the blank data is removed and subjected to multiple linear regression analysis. That is, the use of TC delayed three hours of data for linear regression analysis and the correlation was 0.753 (Figure 7). Through the translation of the time series, it can be found that the PM2.5 data has the highest correlation with the TC
data after an hour shift. That is, PM2.5 content with the TC content changes and the lag time is about an hour. Through the logical causality judgment, PM2.5 content changes must be the result of TC content changes.

Prediction and analysis of PM2.5 mass concentration based on regression analysis model

In general, PM2.5 aerosols contain four substances: carbon particles, water, secondary inorganic salts and mineral dust. Among them, the carbon particles mainly refers to the particles in the elemental carbon and organic carbon in general and its content with TC (Total carbon) shows the water for the air in the gaseous water molecules are adsorbed to the particles on the liquid presence of water molecules, it is absorbed PM2.5 chemical pollutants in the "container", its content with W (Water) shows the secondary inorganic salt mainly refers to inorganic salts or acidic substances such as sulfuric acid, nitrogen oxides, ammonia gas and the like which are adsorbed to the water contained in PM2.5 and formed by chemical reaction such as sulfate, nitrate and ammonium salts; its content with C (Chemical) shows mineral dust mainly refers to the soil dust, industrial process of mineral dust, plant dust and other fine particles, the content of D (Dusty) said.

Based on the results of the previous analysis, the haze particle formation process can be assumed, as shown in Figure 8. Because most of the eastern part of China is a dry climate, it is necessary for cohesive nucleus to have a strong water absorption, in order to further the formation of haze objects through the observation found that China's eastern haze PM2.5 and the total carbon content is high, hence, the eastern part of China area PM2.5 formed by the pollution phenomenon should be called "haze". Therefore, it can be inferred from the physical process point of view that PM2.5 formation process will be carbon particles for the aggregation of nucleus adsorption of water in the air, and then adsorption of chemical pollutants in the air through a certain chemical reaction and become PM2.5 gas sol.

This study was designed to analyze the correlation between the total content of PM2.5 and the total carbon content of the key elements, thus, establishing the mathematical model of PM2.5 genetic analysis. If the mass of PM2.5 in unit volume is expressed by S, then:

\[ S = Tc + W + C + D \]  

(1)

Where TC is for the total carbon, W for the adsorption of water, C for the chemical pollutants and D is for fine particles dust.

The water content W in PM2.5 depends directly on the
content of condensed nucleus, which is mainly composed of carbon particles, and the chemical content C is directly dependent on the water content W. Thus, Equation (1) can be expressed as:

\[
S = Tc + E_1 \times Tc + E_2 \times W + D
\]

\[
= (1 + E_1 + E_2) \times Tc + D
\]

Among them:

\[
E = E_1 + E_1 \times E_2
\]

E1 is a humidity-related parameter; E2 is a parameter with the atmosphere of SO2, nitrogen oxides, ammonia, catalytic components and other related parameters. Considering that the chemical contaminants in the atmosphere are relatively stable over a period of time, it can be assumed that E is considered only as a coefficient related to atmospheric relative humidity. The relationship between E and humidity is non-linear, but when the relative humidity is between 35 and 95%, respectively, then, their relationship can be expressed as a linear relationship given as:

\[
E \approx a \times Rh + b
\]

Where Rh is the relative humidity. When the dust content is very small, formula (2) can be simplified as:

\[
S \approx E \times Tc
\]

The formula 3 into the formula (2), that is, PM2.5 genetic analysis of the "three parameters model" is given as:

\[
S \approx (a \times Rh + b) \times Tc + D
\]

Or simplified as a "two-parameter model" given as:

\[
S \approx (a \times Rh + b) \times Tc
\]

This paper compares the results of multiple linear regression and data analysis based on the mathematical model set by Equation (7).

Multiple linear regression prediction has the characteristics of simple model, accurate prediction result and strong model explanation ability, and has been widely used in model prediction. However, in the haze cause analysis, the situation encountered not only reflects a certain number of relationships, but also need to reflect the number of physical and chemical relations. On the basis of multiple linear regression analysis, this project presents a mathematical model which reflects the relationship between quantities and reflects the specific physical and chemical relationship. In order to test the reliability and validity of the parametric model obtained by PM2.5 genetic analysis, the fitting results and the multiple linear regression fitting were compared.

First, the correlation between TC and humidity and PM2.5 concentration was obtained by the analysis of scatter plot. Therefore, we used multivariate linear regression method to fit the relationship between TC, humidity and PM2.5 concentration. Before the fitting, the feature data is normalized, and for the absence of data in the case of individual elements, the total sample is deleted by deletion in the regression and both passed the significance test.

Figure 9 shows the plot of the measured value and the model fit value. Compared with the curve (black line) of the fitting curve (red line) and the measured mass concentration, it can be seen that the constructed multiple linear regression model can fit the change of PM2.5 mass concentration and the correlation is 0.82306.

Secondly, we use this parameter to fit the same data in the parameter model obtained by PM2.5 genetic analysis and the correlation is 0.83426. By comparing the correlation, it is found that the parameter model obtained by the PM2.5 genetic analysis is better than the multiple linear regression models considering the same TC and humidity correlation.
As shown in Figure 10, 83.4% of the data source changes can be explained. The results of this fitting have been very good considering the complexity of the sample quantity and the change in the mass concentration of fine particles. In addition, for multiple sets of data multiple regression analysis, the relative humidity is expressed in decimal, while a, b values are about 3. Therefore, when the relative humidity is about 0.3, PM2.5 concentration of TC concentration is about 4 times; when the relative humidity is 0.9 or so, PM2.5 concentration of TC concentration is about 6 times; under normal circumstances, PM2.5 is about 5 times the average concentration of TC (corresponding to the case of relative humidity of 0.7).

The main reason for the relative fitting effect of the parameter model obtained by PM2.5 data analysis is that the carbon particles are aggregated according to the cause of the PM2.5 physical formation process and the water content in PM2.5 is directly dependent on the carbon content of aggregated nuclei, dominated by carbon particles, which not only makes the mathematical model of PM2.5 relatively stable and reliable, but also gives a clear physical meaning, which is helpful to the precise positioning of PM2.5 source.

CONCLUSION

Based on the analysis of PM2.5 physical formation process, this paper presents the mathematical model of PM2.5 causality of TC-PM2.5 causality. The results of the scattergram analysis of PM2.5 content and the influencing factors show the strong correlation between TC and PM2.5 concentration. Compared with the results of multivariate linear regression model, although the fitting degree of multiple linear regression model is very high, the mathematical model of PM2.5 is better and the fitting degree is as high as 83.4%. Thereafter, we also made the TC 1 h data fitting, and found that the mathematical effect of the PM2.5 formation model was further improved, which fully proved that the PM-PM2.5 was proposed in the PM2.5 physical formation process causal relationship. The analysis shows that the effect of TC is magnified during the physical formation of PM2.5, and the magnification is about 5 times. PM2.5 is 5 times the mass of TC and carbon particles in the formation of PM2.5 play a real "source" role.

REFERENCES