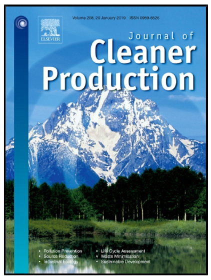
Accepted Manuscript



Using the geographical detector technique to explore the impact of socioeconomic factors on PM2.5 concentrations in China

Yueting Ding, Ming Zhang, Xiangyan Qian, Chengren Li, Sai Chen, Wenwen Wang

PII:	S0959-6526(18)33561-3
DOI:	10.1016/j.jclepro.2018.11.159
Reference:	JCLP 14918
To appear in:	*Journal of Cleaner Production*
Received Date:	29 July 2018
Accepted Date:	17 November 2018

Please cite this article as: Yueting Ding, Ming Zhang, Xiangyan Qian, Chengren Li, Sai Chen, Wenwen Wang, Using the geographical detector technique to explore the impact of socioeconomic factors on PM2.5 concentrations in China, *Journal of Cleaner Production* (2018), doi: 10.1016/j.jclepro.2018.11.159

Using the geographical detector technique to explore the impact of socioeconomic factors on PM2.5 concentrations in China

Yueting Ding [a], Ming Zhang [a], Xiangyan Qian [a], Chengren Li [c], Sai Chen [a], Wenwen Wang[b*]

[a] School of Management, China University of Mining and Technology, Xuzhou 221116, China

[b] School of Math, China University of Mining and Technology, Xuzhou 221116, China

[c] Shanghai Surveying & Mapping Institute, Wuning Rd 419, Shanghai 200063, China

Abstract: The purpose of this study is to explore the impact of socioeconomic factors on $PM_{2.5}$ concentrations and to provide insights into air quality improvement. We firstly studied the spatial autocorrelations of $PM_{2.5}$ concentrations using global Moran's *I* and Local Indicators of Spatial Association, then explained the spatial heterogeneity of regional $PM_{2.5}$ concentrations and identify the driving factors on $PM_{2.5}$ by geographical detector technique, using data extracted from satellite observations over the years from 2000 to 2015. The results showed that, the annual average $PM_{2.5}$ concentration in China ranged from 11.5 μg/m³ to 18.7 μg/m³ with an upward trend in general, while $PM_{2.5}$ pollutions were relatively serious in Beijing-Tianjin-Hebei region and Yangtze River Delta. Regional $PM_{2.5}$ concentrations showed significant global and local spatial autocorrelation. Regions of high $PM_{2.5}$ concentrations tend to cluster with regions of similar $PM_{2.5}$ concentrations. From a long-term perspective, population density has the greatest power of determinant on $PM_{2.5}$, followed by electricity consumption, industry structure, coal consumption, number of vehicles per capita and GDP per capita. Over the study period, the impact of population density revealed a trend to first rise and then fall, and the impacts of GDP per capita showed a slightly upward trend. The impact trends of number of vehicle per capita and industry structure presented to be fluctuated. The determinant power of coal consumption and electricity consumption had significant downward trends.

Keywords: PM2.5 concentrations; Socioeconomic factors; Spatial autocorrelation; Spatial heterogeneity; Geographical detector model

---

* Corresponding author. Tel./fax: t86 051683591280; zhangmingdlut@163.com(M.Zhang) , wangwenwencumt@163.com (WW,Wang).

Using the geographical detector technique to explore the impact of socioeconomic factors on $PM_{2.5}$ concentrations in China

Abstract: The purpose of this study is to explore the impact of socioeconomic factors on $PM_{2.5}$ concentrations and to provide insights into air quality improvement. We firstly studied the spatial autocorrelations of $PM_{2.5}$ concentrations using global Moran's *I* and Local Indicators of Spatial Association, then explained the spatial heterogeneity of regional $PM_{2.5}$ concentrations and identify the driving factors on $PM_{2.5}$ by geographical detector technique, using data extracted from satellite observations over the years from 2000 to 2015. The results showed that, the annual average $PM_{2.5}$ concentration in China ranged from 11.5 μg/m³ to 18.7 μg/m³ with an upward trend in general, while $PM_{2.5}$ pollutions were relatively serious in Beijing-Tianjin-Hebei region and Yangtze River Delta. Regional $PM_{2.5}$ concentrations showed significant global and local spatial autocorrelation. Regions of high $PM_{2.5}$ concentrations tend to cluster with regions of similar $PM_{2.5}$ concentrations. From a long-term perspective, population density has the greatest power of determinant on $PM_{2.5}$, followed by electricity consumption, industry structure, coal consumption, number of vehicles per capita and GDP per capita. Over the study period, the impact of population density revealed a trend to first rise and then fall, and the impacts of GDP per capita showed a slightly upward trend. The impact trends of number of vehicle per capita and industry structure presented to be fluctuated. The determinant power of coal consumption and electricity consumption had significant downward trends.

Keywords: $PM_{2.5}$ concentrations; Socioeconomic factors; Spatial autocorrelation; Spatial heterogeneity; Geographical detector model

Highlights:

The $PM_{2.5}$ pollution in China showed a significant upward trend from 2000 to 2007, and a flat and slightly downward trend from 2007 to 2015.

Global and local spatial autocorrelations of $PM_{2.5}$ pollution exist in China.

Population density has the highest power of determinant among 6 socioeconomic driving factors on $PM_{2.5}$ concentration.

Coal consumption and electricity consumption show significant downward trends of impact on $PM_{2.5}$ pollution.

1.Introduction

For the past few years, persistent air pollution in many countries has had a significant impact on the climate, human health and urban sustainable development (Fang et al., 2016; Lelieveld et al., 2015). Among the atmospheric pollutants, fine particulate matter $PM_{2.5}$ is the primary pollutant[1] with the greatest hazard (Li and Zhang, 2014). '$PM_{2.5}$' refers to any aerosol particles of 2.5 micrometers or smaller suspended in the air. Small in size, large in dispersion area, and easy to carry toxic substances, $PM_{2.5}$ can easily enter the human body and cause serious harm to human health (Madrigano et al., 2013; Sofowote et al., 2015), such as cardiovascular morbidity (Ostro et al., 2014), bronchitis and lung cancer (Chalbot et al., 2014), and mortality (Atkinson et al., 2014). Moreover, $PM_{2.5}$ can also cause serious environmental problems such as reduced air visibility and climate change (Brauer et al., 2012; Kan et al., 2012). These problems are more prominent in China due to rapid urbanization and industrialization (Matus et al., 2012), which results in mass population migration and increasing energy consumption (Lyu et al., 2016). As demonstrated in previous study, about 1.2 million to 1.6 million people die prematurely every year in China due to air pollution problems (Yang et al., 2013). Therefore, the problem of $PM_{2.5}$ has received increasing attention from scholars globally. A growing body of literature has focused on exploring the driving factors of $PM_{2.5}$ concentrations, finding that both natural conditions (Dayan et al., 2011;) and human activities (Lou et al., 2016) have significant impact on $PM_{2.5}$ concentrations.

In existing research, some scholars focus on the influence of natural factors on $PM_{2.5}$. Their studies indicate that natural conditions, such as temperature (Li et al., 2014), precipitation (Mazeikis, 2013), wind speed (Wang et al., 2016; Zhang et al., 2015), wind direction (Zhang and Cao, 2015), terrain (Vieira-Filho et al., 2015), etc., are important factors affecting the accumulation and diffusion of $PM_{2.5}$.For example, He et al. (2017) investigated the relationship between air pollution and meteorological conditions in major Chinese cities and found that, compared to coarse particle (PM_{10}), fine particle ($PM_{2.5}$) was easier to be affected by meteorological conditions, and meteorological conditions could explain more than 70% of the

[1] According to MEE of China, "primary pollutant" refers to the air pollutant of the highest individual Air Quality Index (AQI) among all the 6 air pollutants involved in the air quality assessment when the composite AQI is higher than 50.

variation of pollutant concentrations over China. Vakeva et al. (2000) and Hussein et al. (2006) both pointed out that urban temperature and local wind speeds have the greatest impact on $PM_{2.5}$ concentrations. However, since the formation and diffusion of air pollutants mainly take place within the planetary boundary layer (PBL) (Emeis et al., 2008), the structure of PBL is considered to be the most important natural factor on air pollutions, among others. Recent research investigated the linkage between the height/depth of PBL and air pollution in China based on sounding data, finding a significant anti-correlation between them (Miao and Liu, 2019). More specifically, a shallow PBL with low height is found to be responsible for severe air pollutions (Quan et al., 2014). Influenced by meteorological conditions, underlying surface and upper airflow velocity, the structure of PBL is more of a comprehensive reflection of natural factors affecting air pollutions.

Besides natural conditions, a growing number of scholars have explored the correlations between $PM_{2.5}$ concentrations and socioeconomic factors including population (Halkos and Paizanos, 2013), industrial structure (He, 2009), per capita GDP (Auffhammer and Carson, 2008), energy consumption (Wang et al., 2017; Xu et al., 2016), Vehicle population (Bozlaker et al., 2014) and so on, demonstrating that human activities are the fundamental causes of high $PM_{2.5}$ concentration (Han et al., 2016; Platt et al., 2014). For instance, based on the environmental Kuznets curves (EKC) hypothesis and econometric models, Wang et al. (2017) found that $PM_{2.5}$ pollutions in China were closely related to the urban population, per capita GDP, population density and fossil fuel combustion. Xu et al. (2016) used the Stochastic Impacts by Regression on Population, Affluence and Technology (STIRPAT) model and nonparametric additive regression models to examine the key driving forces of $PM_{2.5}$ pollutions in China, finding that the relationship between economic growth, urbanization, coal consumption, private vehicle and $PM_{2.5}$ concentration followed the inverted “U-shaped” pattern. By using the dynamic spatial panel model, Cheng et al. (2017) pointed out that the main causes of $PM_{2.5}$ in 285 cities in China were population size, the proportion of the second industry, the consumption of coal and the intensity of traffic.). In addition, since more power consumption requires more power supply (Guan et al., 2014), and coal-fired power plants would generate a lot of $PM_{2.5}$ during power generation (Gao et al., 2014), electricity consumption is a factor that cannot be ignored in the generation of $PM_{2.5}$.

Obviously, there has been a large amount of research on $PM_{2.5}$ pollutions, discovering the socioeconomic driving factors on $PM_{2.5}$. However, given these factors, which one or some factors have greater influence on $PM_{2.5}$ pollution compared to others? Moreover, do the influences of these factors remain unchanged over time?

In recently studies, many methods are used to study the driving factors governing $PM_{2.5}$ concentrations, such as hierarchical cluster analysis (Gao et al., 2014), dynamic factor analysis (DFA) (Yu et al., 2015), structural decomposition analysis (Djalalova et al., 2010; Guan et al., 2014), general econometrics models (Zhang et al., 2015). While time series and cross-sectional data are frequently used for analysis, the estimation results of panel data could be more optimized than that of time series and cross-sectional data (Pao and Tsai, 2010). Moreover, due to the high mobility of $PM_{2.5}$ and the first law of geography (Tobler, 1970), the concentration of $PM_{2.5}$ in China is considered to have significant spatial autocorrelations. Some scholars have adopted spatial econometric models such as dynamic spatial panel model (Cheng et al., 2017), geographic weighted regression (Hu et al., 2013), spatial Durbin model (Liu et al., 2017), spatial lag model (SLM) and spatial errors model (SEM) (Hao and Liu, 2016; Ma et al., 2016) for correlation analysis. Using the spatial statistical approaches including spatial interpolation and spatial regression based on ground-level $PM_{2.5}$ observations of 190 Chinese cities, Zhang et al. (2016) found that $PM_{2.5}$ concentration was positively related to population size, amount of atmospheric pollutants, and emissions from nearby cities, but inversely related to precipitation and wind speeds. By using spatial lag model and spatial error model, Hao and Liu (2016) found that the relationship between $PM_{2.5}$ concentrations and per capita GDP was inverted U-shaped. In addition, the vehicle population and the secondary industry have significant and positive influences on urban $PM_{2.5}$ concentrations. Despite the various methods applied in existing research, the geographical detector method (GDM) has been rarely used in the analysis of correlation between socioeconomic factors and $PM_{2.5}$ concentrations.

Proposed by Wang et al (2010), the basic idea of GDM is that, if the values of two variables tend to share the similar spatial distribution, these two variables are spatially correlated. While relatively new, this method has some significant advantages: first, no linear assumptions are required in the analysis of dependent and independent variables; second, interactive influence of two independent variables on the dependent variable can be detected; third, any potential factors can be included in the analysis without having to consider the problem of multiple collinearity. Though first applied in the research of geographical issues, GDM has later been used in the study of human health (Wang and Hu, 2012), housing price (Wang et al., 2017) aeolian desertification (Du et al., 2016) and air pollutions (Zhou et al., 2018). However, research on the correlation between $PM_{2.5}$ pollutions and socioeconomic factors using GDM have focused on either single year (Zhou et al., 2018) or small scale of area (Lou et al., 2016). Given this, we use GDM in this paper to explore the correlations between $PM_{2.5}$ concentrations and socioeconomic factors for a long period of time from 2010 to 2015 and on a national scale of China. Moreover, since each time the GDM analysis is performed, it is performed based on a set of cross-sectional data, e.g. data of a single year. Therefore, in this paper, based on data from 2000 to 2015, we can obtain the impact of socioeconomic factors on $PM_{2.5}$ concentration for each year separately and thus the trend of impact for each factor over the study period.

The main goal of this paper is to identify the impact of driving factors on $PM_{2.5}$ concentration in China. As such, we began by using ArcGIS software to extract $PM_{2.5}$ concentration data from satellite observations over the period of 2000 to 2015. Secondly, driving factors were selected based on former studies. Thirdly, the temporal-spatial patterns of $PM_{2.5}$ concentrations were explored by spatial autocorrelation analysis. Finally, the power of determinants for each driving factor of $PM_{2.5}$ were investigated based on geographical detector method, and the trend of determinant power of each factor was analyzed. The main conclusions could be beneficial for making further endeavor to improve atmospheric environmental quality.

The remainder of the paper is organized as follows. Section 2 describes the data and presents the methods used in this paper. Section 3 presents the results of analysis and our main findings. Finally, Section 4 concludes and proposes policy recommendations.

2.Material and methodology

2.1 Data and data sources

2.1.1 Data of annual average $PM_{2.5}$ concentrations

In existing research, data of $PM_{2.5}$ concentrations were mostly derived from urban ground air quality monitoring sites (Djalalova et al., 2010; Zhou et al., 2018) or emission inventory database

(Guan et al., 2014; Li et al., 2018). However, the latest version of Ambient Air Quality Standards (GB 3095-2012) in China was revised and implemented in 2012, and since then the readings of $PM_{2.5}$ concentrations were included and recorded for the first time. Therefore, considering the attainability of data, the $PM_{2.5}$ concentrations in this paper were calculated based on global satellite observations provided by Socioeconomic Data and Applications Center (SEDAC[2]), which presented as annual global surface of concentrations (micrograms per cubic meter, $\mu g/m^3$) of mineral dust and sea-salt filtered $PM_{2.5}$ over the period of 1998-2016 (van Donkelaar et al., 2018). The annual estimates were generated following a geographically weighted regression technique (van Donkelaar et al., 2016).

Global in scope as the original dataset is, a vector layer of China was used as the clip extent to extract the study area, covering all 34 provinces, autonomous regions and municipalities in China. The time range in this study is from 2000 to 2015, considering the availability of socioeconomic driving factors data. While the original raster grids had a grid resolution of 0.01°×0.01°, resampling was applied to decrease the resolution to 0.18°×0.18° to reduce the points to 29,426, for the reason of data rows limitation (32,767) in GeoDetector program. Then the raster grid files were transformed into point files by extracting the value of each point. For each year, the $PM_{2.5}$ concentrations were recorded in the point file with each point containing a value of $PM_{2.5}$ concentration. All processes were realized in ArcGIS 10.2 software. Fig. 1 shows the maps representing ground-level annual average $PM_{2.5}$ concentration of the year 2000, 2005, 2010, and 2015, respectively. As shown in the figure legend, the spectrum of $PM_{2.5}$ concentrations range from a minimum of 0 $\mu g/m^3$ and maximum 100.2 $\mu g/m^3$. Notably, compared to the year 2010, the Northeast China region including Liaoning, Jilin and Heilongjiang shows an obvious and sharp increase of $PM_{2.5}$ concentrations, while the $PM_{2.5}$ pollutions in central and western regions of China shows distinct improvements.

Fig. 2 shows the trends of the average $PM_{2.5}$ concentrations in different regions of China from 2000 to 2015. It can be seen that among the three major economic growth poles, the $PM_{2.5}$

[2] SEDAC is a data center in Earth Observing System Data and Information System (EOSDIS) of the U.S. National Aeronautics and Space Administration (NASA), hosted by the Center for International Earth Science Information Network (CIESIN) within the Earth Institute at Columbia University.

pollution situation in the Beijing-Tianjin-Hebei region is relatively the most serious. During the study period, the highest annual average $PM_{2.5}$ concentration in the Beijing-Tianjin-Hebei region reached 53.2 μg/m³, which is significantly higher than the national average. The $PM_{2.5}$ pollution in the Yangtze River Delta region is also serious, with the highest annual average of 48.8μg/m³. The situation in the Pearl River Delta is relatively better, but it is still about twice the national average. Overall, annual average $PM_{2.5}$ concentration in three major economic growth poles reached their highest level around 2007 and began to decline slowly before the year 2013. As we can see from Fig. 2, the subsequent sharp increase in the year 2013 made the overall trend appear to be on the rise again.

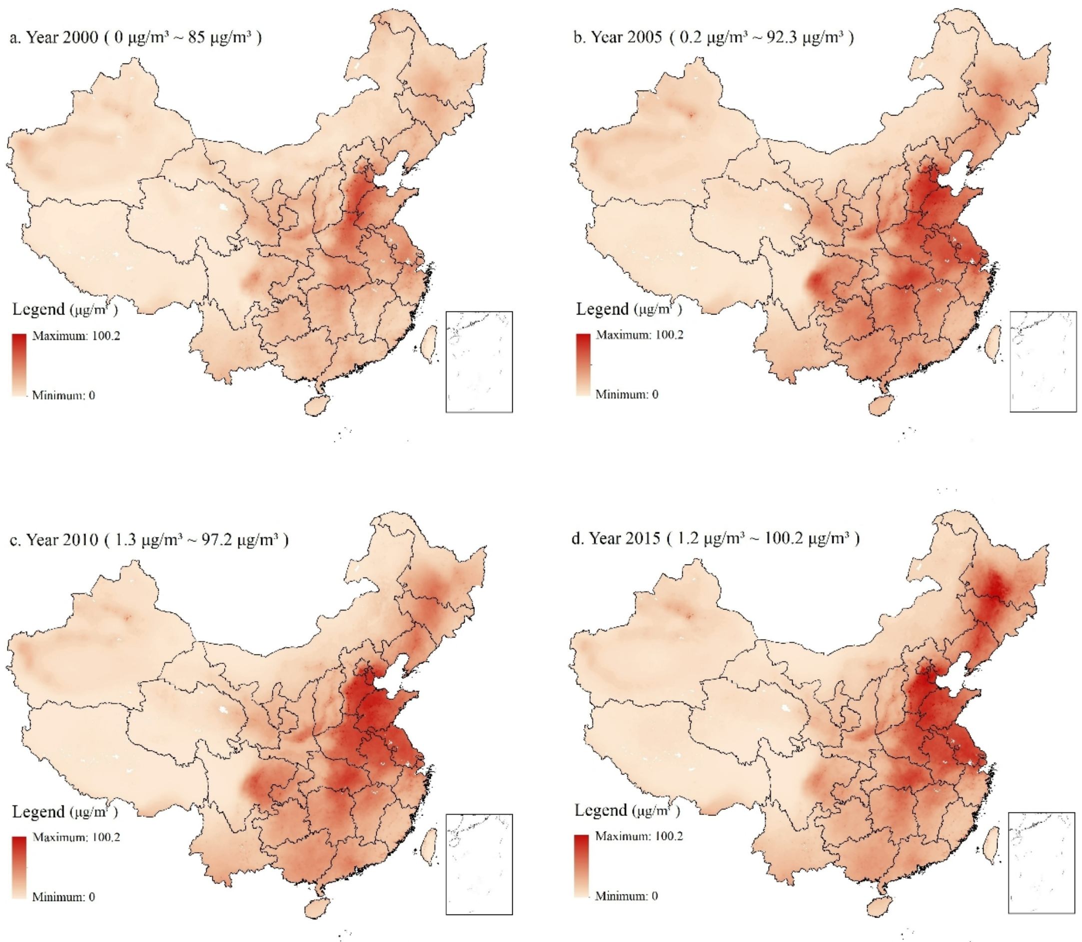


Fig. 1 Spatial distributions of annual average $PM_{2.5}$ concentration in the year 2000, 2005, 2010 and 2015

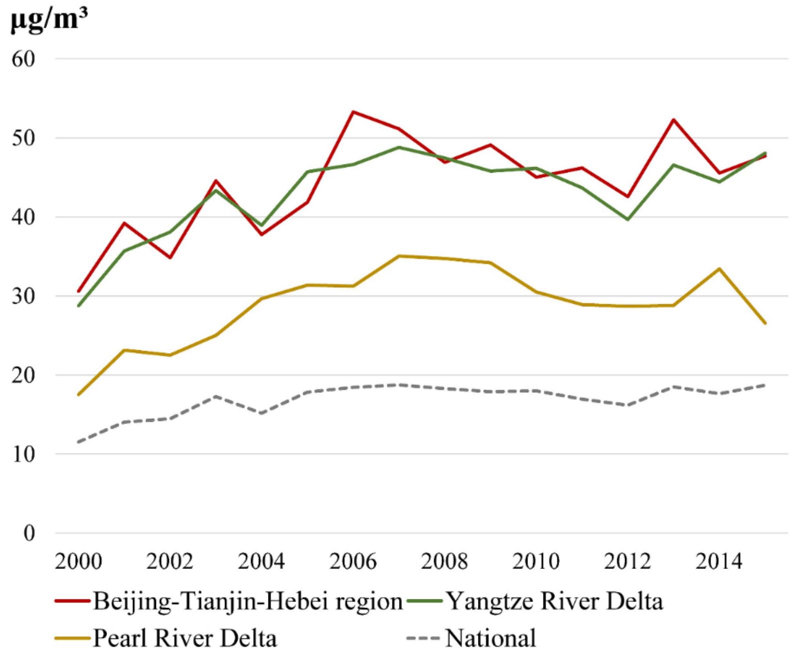


Fig. 2. Average $PM_{2.5}$ concentrations of Beijing-Tianjin-Hebei region, Yangtze River Delta, Pearl River Delta and nationwide from 2000 to 2015

2.1.2 Data of indicators for driving factors

Based on former researchers, the explanatory variables we selected are population density (PD), GDP per capita (GDPPC), number of vehicles per capita (NVPC), industry structure (IS), coal consumption (CC) and electricity consumption (EC).

Definitions, unit of measurement and data sources of the variables above are presented in Table 1. All data were collected from the year 2000 to 2015.

Table 1. Descriptions of the indicators for driving factors.

Variables	Definitions	Units	Data sources
PD	Regional resident population density	100 people/km^2	China statistical yearbook
GDPPC	Regional gross domestic product	yuan	China statistical yearbook
NVPC	Possession of civil vehicles per capita	vehicles	China statistical yearbook
IS	Percent of value added of industry in GDP	percent	China industry statistical yearbook
CC	Regional consumption of coal	10^4 tons	China energy statistical yearbook
EC	Regional consumption of electricity	100 million kwh	China energy statistical yearbook

2.2 Methods

2.2.1 Global spatial autocorrelation analysis

According to Tobler's First Law of Geography, everything is related to everything else, but near things are more related to each other (Tobler, 1970). To find out whether the $PM_{2.5}$ concentrations have an impact on neighboring regions, global Moran's I (Moran, 1950), which was invented by

Patrick Moran in 1950, was calculated to examine the spatial autocorrelation patterns of $PM_{2.5}$ concentration. Formula for calculating global Moran's I is expressed as follows:

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{S^2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \quad (1)$$

$$S^2 = \sum_{i=1}^{n}(y_i - \bar{y})^2 \Big/ n \quad (2)$$

where n is the number of sample regions, which should be at least 30 to obtain reliable results (Mitchel and Esri, 2005); y_i and y_j are the regional average $PM_{2.5}$ concentrations of region i and j, respectively; $\bar{y}$ is the average $PM_{2.5}$ concentrations of all regions; w_{ij} is the spatial weight matrix, which can be both contiguity-based and distance-based. In this paper, we adopt the contiguity-based spatial weight using *queen* criterion[3], for which the setting principle is as follows:

$$w_{ij} = \begin{cases} 1 & \text{when region } i \text{ and } j \text{ are adjacent;} \\ 0 & \text{when region } i \text{ and } j \text{ are not adjacent;} \\ 0 & \text{when } i = j. \end{cases} \quad (3)$$

Usually, though not always, the value of Moran's I ranges from -1 to 1, where positive correlations of air pollution exist among cities when I is positive, and negative correlations when I is negative. No correlations exist when I is 0. The results of Moran's I are tested using standardized statistic z to test the existence of spatial autocorrelation between regions:

$$z = \frac{I - \mathrm{E}[I]}{\sqrt{\mathrm{VAR}(I)}} = \frac{I - \mathrm{E}[I]}{\mathrm{E}[I^2] - \mathrm{E}[I]^2} \quad (4)$$

where:

$$\mathrm{E}[I] = -1/(n-1) \quad (5)$$

The statistic z is calculated based on a random permutation procedure under 999 Monte-Carlo runs in this paper. Significant level of statistic z is determined by p-value, which is acceptable when $p < 0.01$ as we set in this paper.

[3] The *queen* criterion determines neighboring units as those that have any point in common, including both common boundaries and common corners.

2.2.2 Local spatial autocorrelation analysis

While global statistics such as global Moran's *I* are frequently used in the analysis of spatial autocorrelation, stationarity over space could be unreliable as the number of spatial observations increases. To test local instabilities in overall spatial association, Local Indicators of Spatial Association (LISA) is introduced in this paper. Here we use local Moran's *I*(Anselin, 1995), of which the formula is as follows:

$$I_i = \frac{n(y_i - \bar{y})\sum_{j=1}^{n} w_{ij}(y_j - y)}{\sum_{i=1}^{n}(y_i - \bar{y})^2} = \frac{(y_i - \bar{y})\sum_{j=1}^{n} w_{ij}(y_j - \bar{y})}{S^2} \qquad (6)$$

where y_i, y_j, $\bar{y}$, w_{ij} and S are the same as in Equ. (1) and Equ. (2).

The results of local Moran's I are tested using z-score, which is computed as:

$$z_{I_i} = \frac{I_i - \mathrm{E}[I_i]}{\sqrt{\mathrm{E}[I_i^2] - \mathrm{E}[I_i]^2}} \qquad (6)$$

where:

$$\mathrm{E}[I_i] = -\frac{\sum_{i \neq j, j=1}^{n} w_{i,j}}{n-1} \qquad (7)$$

When $I_i > 0$, positive correlations exist between region i and adjacent regions with similar $PM_{2.5}$ concentrations, and spatial clusters exists when z-score is statistically significant at 5% level ($p<0.05$), namely High-High cluster or Low-Low cluster; when $I_i < 0$, negative correlations exist between region i and adjacent regions with dissimilar $PM_{2.5}$ concentrations, and spatial outliers exist when z-score is statistically significant, namely High-Low outlier or Low-High outlier.

Both global and local autocorrelation analysis are completed in GeoDa software (Anselin, 2005) and all data (in the form of ERIS shapefiles) used in the analysis are preprocessed in the ArcGIS 10.2 software.

2.2.3 Geographical detector model (GDM)

Although the study of spatial autocorrelation breaks through the hypothetical conditions under which the data are independent and identically distributed (IID), the problem of spatial stratified heterogeneity has arisen gradually. Proposed by Wang et al. (2010), geographical detector model

(GDM) is a newly developed method to identify the pattern of spatial stratified heterogeneity and the factors responsible for the risk (Wang et al., 2010). The basic idea is that if a factor (e.g. population density) takes on a similar spatial distribution to that of the risk (e.g. $PM_{2.5}$ concentration), this factor does contribute to the certain risk. The power of determinant of influencing factor x on the risk can be examined as follows:

$$q = 1 - \frac{\sum_{i=1}^{n} N_i \sigma_i^2}{N\sigma^2} = 1 - \frac{\sum_{i=1}^{n} N_i \sum_{h=1}^{m} \left[\left(y_i - \overline{y_i} \right)^2 / N_{ih} \right]}{N\sigma^2} \qquad (8)$$

where q is the power of determinant; N is the number of units in the whole study region; based on spatial stratified heterogeneity, the whole region is classified into n sub-regions, donated by i=1,2,…,n; N_i is the number of sub-regions; σ^2 and σ_i^2 are the variances of the whole study region and sub-region i, respectively; y_i is the risk observations within sub-region i; $\overline{y_i}$ is the mean value of risk observations within sub-region i; N_{ih} is the number of observations y_i in sub-region i. As shown in Fig. 3, the study region, where $PM_{2.5}$ concentrations are recorded in raster grid cells, would be transformed into dot files, each dot containing a $PM_{2.5}$ concentration value. Then the dependent variable x and independent variable y in the study region are separated into 2 layers. In the x layer, the whole region is classified into n sub-regions, according to geographical factors (e.g. population density). The strata of y is obtained by overlaying y layer with x layer, from which we can calculate q. Usually, $q \in [0,1]$, which is the power of determinant assessing the relationship between y and x. The closer q is to 1, the greater the influence x is on y. For instance, if q value is 0.5, it means that x can explain 50% of y.

It's worth mentioning that, based on the idea of Analysis of Variance (ANOVA), GDM has no linear assumption on variables, which means that the multicollinearity of input factors can be eliminated and ignored. Therefore, adding new factors or excluding existing factors does not affect the results of other factors, which is one of the advantages of GDM. The analysis of GDM in completed in GeoDetector program, which can be obtained at http://www.geodetector.org (Wang and Xu, 2012).

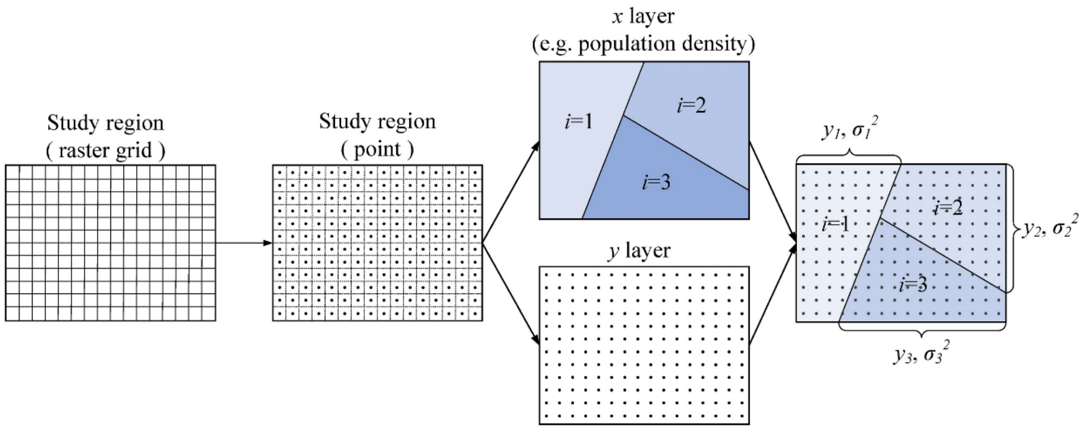


Fig. 3. The principle of geographical detector models

2.2.4 Natural Breaks classification method

The GDM has been proved advantageous as no linear hypothesis is needed in the analysis, because the dependent variables are categorial rather than numerical. Therefore, numerical classification technique should be applied to convert numerical variables to categorial ones. Here we used the Natural Breaks classification method (Jenks, 1967) as the classification method. Natural Breaks is a data classification method to optimize the arrangement of a set of value into "natural" classes. The basic idea of Natural Breaks is to seek to minimize each class's average deviation from the class mean, and to maximize each class's deviation from the means of other classes. Given the number of classes, threshold values for classification would be determined according to the algorithm of Natural Breaks. Here we classify the values of each dependent variable into 5 levels for each year, by the means of ArcGIS software.

Due to space limitation, Table 2 provides an example of threshold values in classifications for the year 2015. In Table 2, square brackets "[" and "]" mean boundary values included, while round brackets "(" means boundary values excluded. Spatial distributions of classifications for 6 dependent variables in 2015 are shown in Fig. 4.

Table 2. Threshold values of dependent variables in 2015

threshold	PD	GDPPC	NVPC	IS	CC	EC
level 1	[0.027, 1.441]	[26165, 31999]	[0.074, 0.085]	[19.74, 23.65]	[1071.92, 4728.13]	[40.53, 658]
level 2	(1.441, 3.962]	(31999, 43805]	(0.085, 0.102]	(23.65, 40.69]	(4728.13, 9805.31]	(658, 1334.32]
level 3	(3.962, 7.844]	(43805, 52321]	(0.102, 0.136]	(40.69, 45.96]	(9805.31, 13826.07]	(1334.32, 2160.34]
level 4	(7.844, 13.42]	(52321, 77644]	(0.136, 0.177]	(45.96, 48.42]	(13826.07, 23719.94]	(2160.34, 3553.9]
level 5	(13.42, 32.919]	(77644, 107960]	(0.177, 0.246]	(48.42, 50.48]	(23719.94, 40926.94]	(3553.9, 5310.69]

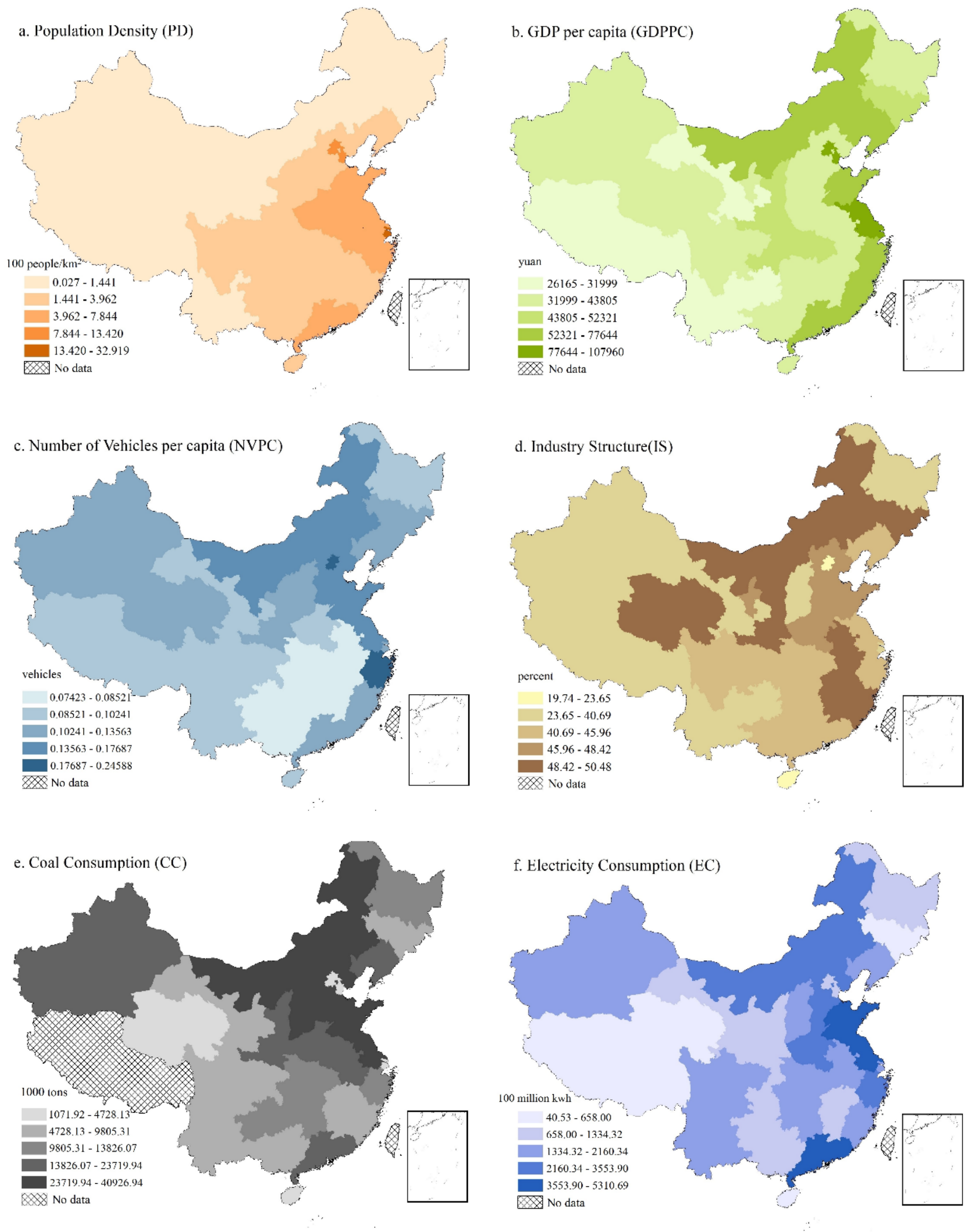


Fig. 4. Spatial distributions of classifications for 6 dependent variables in 2015

3.Results and discussions

3.1 Temporal-spatial patterns of $PM_{2.5}$ concentrations

3.1.1 Global spatial autocorrelation analysis

As shown in Table 3, the global Moran's *I* of $PM_{2.5}$ concentrations for the year 2000-2015 are all positive and the *z*-score all statistically significant at the 1% level (p<0.01). Positive Moran's *I* indicates that spatial agglomeration exists among regional $PM_{2.5}$ concentrations. Regions with high $PM_{2.5}$ concentration values tend to cluster near other high value regions, and low values cluster near other low ones. Note that the Moran's *I* values appear in a narrow band around 0.55 for the year 2000-2015, with maximum value of 0.587 and minimum 0.530, which reveals a relatively high level of stability in spatial autocorrelation within the study period. With help of GeoDa software, the Moran's *I* scatterplots of regional $PM_{2.5}$ concentration for the year 2000-2015 were drawn. Due to space limitation, only 4 years are reported among others. In the scatterplots (Fig. 5), the horizontal axis represents the standardized regional $PM_{2.5}$ concentrations, while the vertical axis represents standardized lagged regional $PM_{2.5}$ concentrations. As shown in Fig. 5, most of the dots appear in the first and third quadrants, meaning that the positive spatial autocorrelations of $PM_{2.5}$ concentration exist in most regions of China during the study period. From the internal point distribution of the scatterplots, the positive spatial correlations of $PM_{2.5}$ has long-term stability (Anselin, 1995).

Table 3. Global Moran's *I* of $PM_{2.5}$ concentrations for the year 2000-2015

Year	Moran's *I*	*z*-score	Year	Moran's *I*	*z*-score
2000	0.530	4.942***	2008	0.566	5.275***
2001	0.560	5.437***	2009	0.544	5.094***
2002	0.557	5.027***	2010	0.544	4.796***
2003	0.587	5.276***	2011	0.573	5.487***
2004	0.572	5.171***	2012	0.558	5.172***
2005	0.544	4.963***	2013	0.564	5.281***
2006	0.556	5.503***	2014	0.558	5.331***
2007	0.586	5.578***	2015	0.560	5.176***

*** The 1% level of significance ($p < 0.01$)

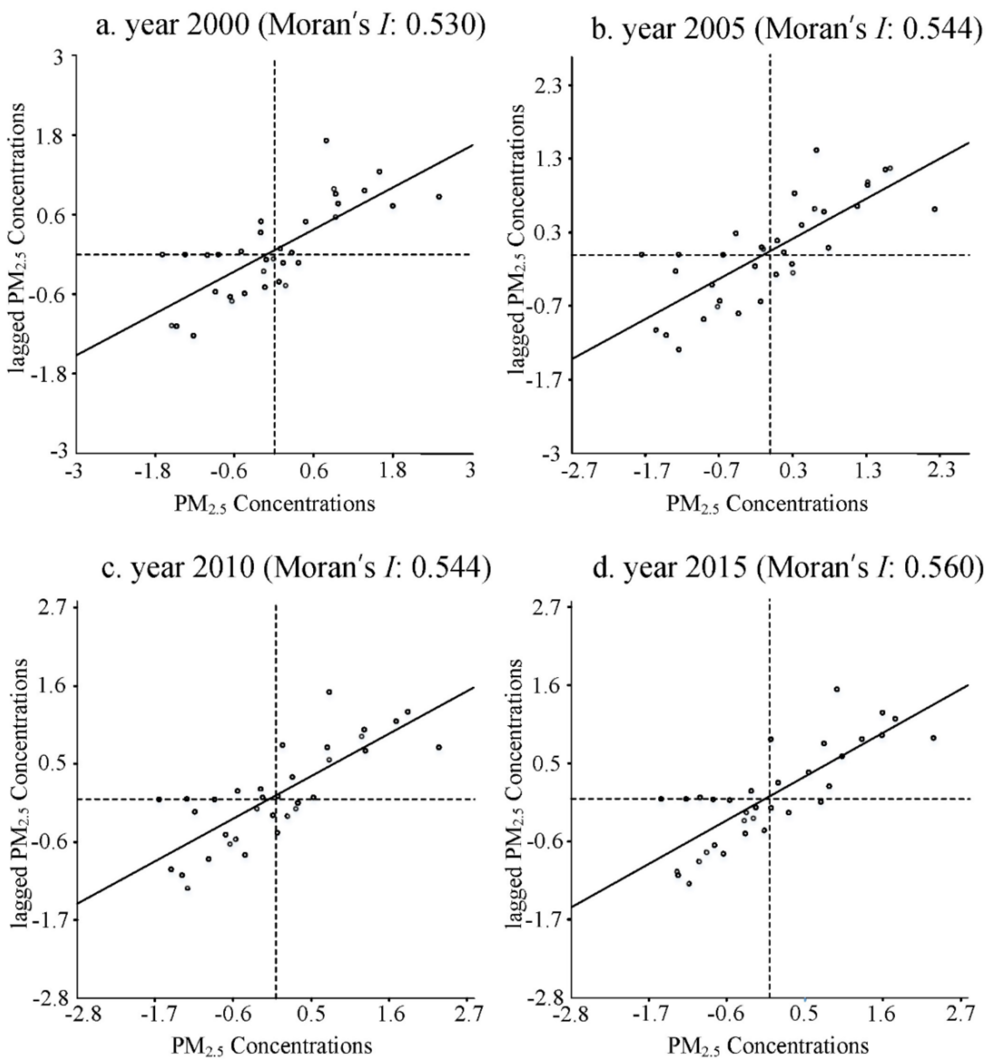


Fig. 5. Global Moran's I scatterplots of $PM_{2.5}$ concentrations in 2000, 2005, 2010 and 2015

3.1.2 Cluster and outlier analysis (Local Moran's I)

As we can see from Fig. 5, a few dots appear in the second and fourth quadrants, which reveals local patterns of association in certain regions. Therefore, local Moran's I was calculated to identify spatial clusters or outliers of regional $PM_{2.5}$ concentrations. As previously mentioned, a local Moran's I value, a z-score, a p-value, and a code representing the cluster type for each region would be calculated to identify spatial clusters or outliers.

Given the space limitation, we present the calculation results of all regions in the year 2015 as an example (Table 4), where "HH" means High-High cluster and "LL" Low-Low. As shown in Table 4, the z-score of 11 regions are statistically significant at 5% level (p<0.05), among which 6 regions are High-High cluster and 5 regions Low-Low. No High-Low outliers nor Low-High outliers exist between any regions, which reveals that no regions tend to cluster with regions of dissimilar $PM_{2.5}$ pollution.

Table 4. Local Moran's *I* of 34 regions in the year 2015

Region	Regional average $PM_{2.5}$ concentration (μg/m^3)	Local Moran's *I*	*z*-score	*p*-value	Cluster type
Beijing	48.918	2.936	2.196**	0.028	HH
Tianjin	75.632	3.982	2.962***	0.003	HH
Hebei	45.295	4.230	1.892	0.058	
Inner Mongolia	11.120	-0.095	0.060	0.952	
Liaoning	46.652	0.432	0.318	0.751	
Jilin	44.327	-0.078	0.008	0.994	
Heilongjiang	30.693	-0.006	0.040	0.968	
Shanghai	61.244	2.824	2.114**	0.035	HH
Zhejiang	30.494	0.062	0.104	0.917	
Jiangsu	65.003	7.970	4.328***	0.000	HH
Anhui	55.756	6.613	3.071***	0.002	HH
Fujian	19.079	0.031	0.074	0.941	
Jiangxi	32.571	0.166	0.157	0.875	
Shandong	61.658	7.795	4.235***	0.000	HH
Henan	50.345	3.697	1.753	0.080	
Hubei	40.975	1.171	0.611	0.541	
Hunan	35.490	-0.043	0.063	0.950	
Guangdong	26.453	0.111	0.128	0.898	
Guangxi	28.738	0.077	0.106	0.915	
Hainan	14.623	-	-	-	
Chongqing	25.849	0.171	0.157	0.875	
Sichuan	12.769	4.683	2.085**	0.037	LL
Guizhou	23.508	0.576	0.354	0.723	
Yunnan	14.960	2.022	1.147	0.252	
Tibet	4.617	5.320	2.911***	0.004	LL
Shaanxi	23.141	0.863	0.449	0.654	
Gansu	10.745	5.237	2.449**	0.014	LL
Qinghai	4.966	5.546	3.032***	0.002	LL
Ningxia	17.488	1.503	0.969	0.333	
Xinjiang	7.900	4.063	2.525**	0.012	LL
Shanxi	25.196	-0.118	0.002	0.999	
Taiwan	6.971	-	-	-	
Hong Kong	23.700	0.063	0.095	0.924	
Macau	-	-	-	-	

** The 5% level of significance ($p < 0.05$)

*** The 1% level of significance ($p < 0.01$)

Representative examples of local Moran's *I* scatterplots in 2000 and 2015 are presented in Fig. 6, where orange area corresponds to "HH" and blue area "LL". As shown in Fig. 6, High-High clusters are mainly distributed in Beijing, Shandong, Jiangsu and so on, while Low-Low cluster mainly distributed in Xinjiang, Tibet, Qinghai and so on. During the study period 2000-2015, frequencies of being cluster are counted and presented in Fig. 7, where Jiangsu, Shandong, Anhui and Tianjin manifest as HH in every year, while Xinjiang, Tibet and Qinghai manifest as LL in every year. In general, it appears that high $PM_{2.5}$ concentration clusters mainly distribute in the Beijing-Tianjin-Hebei region and the Yangtze River Delta region with long-term stability.

From the results above, it can be seen that the spillover effect of $PM_{2.5}$ pollution does exist. Despite stricter environmental regulations and lower secondary industry proportion, Beijing manifest as HH in most of years over the study period (12 out of 16). This fact shows that short-distance industrial transfer does not lead to a complete improvement of the atmospheric environment. In fact, industrial transfer is regarded by many environmental economists as one of the important reasons for the improvement of environmental quality in developed countries, and may also be the real motivation behind the EKC hypothesis, in which environmental quality would first deteriorate and then improve along with the economic growth (Stern, 2004). However, due to the spatial spillover effects of air pollution, Beijing's industrial transfer to neighboring regions does not lead to improvements in air quality, nor can it obtain all the benefits of strict environmental regulations (Fredriksson and Millimet, 2002).

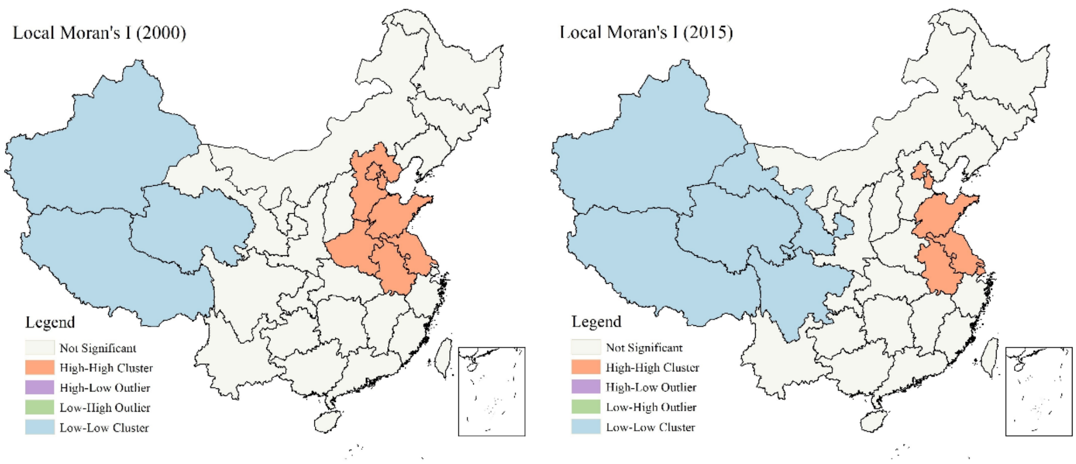


Fig. 6. Local Moran's I scatterplots in 2000 and 2015

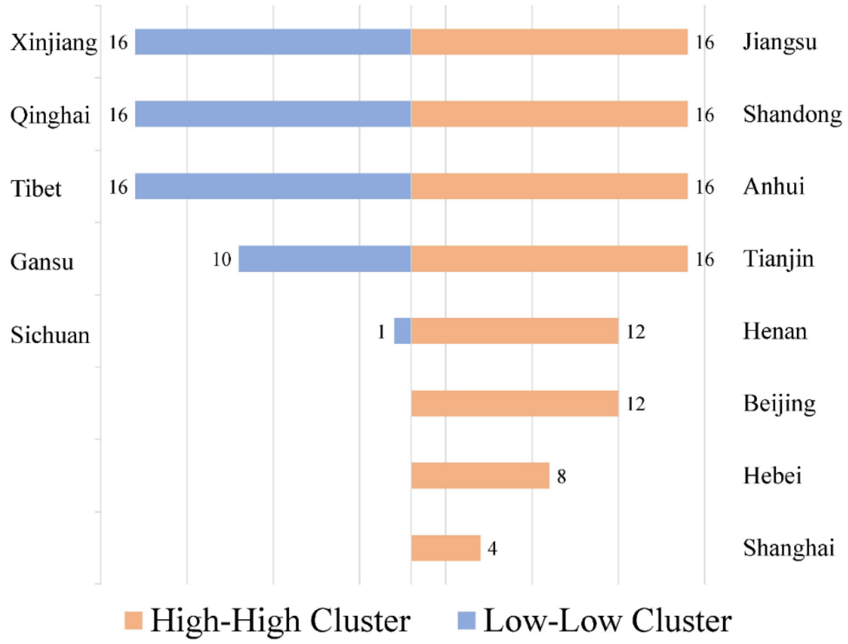


Fig. 7. Frequencies of being HH or LL during the year 2000 to 2015

3.2 Driving factors on $PM_{2.5}$ concentrations

As described in the previous section, positive spatial autocorrelations exist mainly in the relatively developed regions. The agglomeration pattern of regional $PM_{2.5}$ concentration was influenced by many factors. In order to identify the driving factors and their power of determinant on $PM_{2.5}$ pollution, geographical detector model was employed in this section. Here we examined the effect of 6 socioeconomic indicators on $PM_{2.5}$ concentration by the means of GeoDetector program. Definitions of 6 indicators are shown in Table 1. Since the variables of driving factors used in GDM must be categorial variables, the Natural Breaks classification method is applied to convert the original dependent variables from numerical variables to categorial ones, including PD, GDPPC, NVPC, IS, CC and EC, during the study period of 2000 to 2015. The power of determinant values (q) were then calculated using GDM (Table 5), and all the results of q values are significant at 1% level (p<0.01). Consider the results of the year 2015, based on the power of determinant on spatial heterogeneity, the driving factors can be ranked as follow: PD > IS > GDPPC > EC > CC > NVPC.

Population density, compared to other driving factors, contributed a remarkably prominent impact on $PM_{2.5}$ pollution in China, generating a q value from 0.425 to 0.658. Existing studies have demonstrated that anthropogenic emissions were the key factors which significantly give rise to $PM_{2.5}$ concentrations (Karagulian et al., 2015; Lou et al., 2016). Anthropogenic emissions are also the essentially and originally causes of secondary aerosols, which are the most abundant source of $PM_{2.5}$(Liang et al., 2016). Salim and Shafiei (2014) found that high population density lead to more non-renewable energy consumption (Salim and Shafiei, 2014), which is generally believed to be a major source of $PM_{2.5}$. This idea is also supported by the relatively high q values of coal consumption, which indicate significant correlations between coal consumption (CC) and $PM_{2.5}$ concentrations.

As shown in Table 5, GDP per capita has a q value ranging from 0.042 to 0.210, revealing a rather significant impact on $PM_{2.5}$ concentration. Represented by GDP per capita, economic development is usually believed to have non-ignorable effect on environmental degradation. In the famous theoretical hypothesis of Environmental Kuznets Curve (EKC), economic development is regarded to have an inverted U-shaped relationship with environmental degradation, as indicated by Simon Kuznets (Kuznets, 1955). However, as to whether EKC really exists, different studies have reached different conclusions. For example, by detailed review of 35 literatures on EKC, Kaika et al. (2013) found that conflicting results were reached depending on different analysis (cross-country or time-series) and on different periods under analysis, including positive, inverted U-shaped or non-significant relationships between economic growth and environmental degradation.

The q values of NVPC indicate the impact of vehicle emissions on $PM_{2.5}$ concentrations, which in Table 5 shows a range from 0.139 to 0.376, meaning that NVPC can explain about 13.9% to 37.6% of the annual average $PM_{2.5}$ concentration. As revealed in the China Vehicle Environmental Management Annual Report (2018), vehicle emissions accounts for about 10% to 30% of $PM_{2.5}$ sources, varies according to different cities. Although the contribution of $PM_{2.5}$ concentrations in most cities is dominated by coal combustion, in some cities, vehicle emissions have become the primary source of $PM_{2.5}$ in 2017 (MEE, 2017).

The factor EC generates a q value from 0.209 to 0.573, indicating a significantly high determinant power of electricity consumption on $PM_{2.5}$ pollutions. As a matter of fact, up to 70 % of electricity consumption happen in the secondary industry, while the rest happen in agriculture, service sector and residents. Since electricity consumptions such like agricultural irrigations or household electrical appliances have almost little impact on air quality. Therefore, we can assume that industrial electricity consumption has important influence on $PM_{2.5}$ concentration. Interestingly, q values of industry structure (IS) are obviously lower than those of electricity consumption (EC), which result in the inconsistence between strong influence of industrial electricity consumption and smaller influence of industry structure. One possible explanation could be that: in this paper, IS is measured by the percentage of value added of industry in GDP, which can probably not best reflect the environmental impact of industrial activities. In other words, compared to industrial electricity consumptions, spatial distributions of industrial economic output have lower similarity to the spatial distribution of $PM_{2.5}$ concentrations. With energy efficient technologies applied, environmental impact of economic activities become smaller (Kaika and Zervas, 2013).

Table 5. Power of determinant value (q) of each driving factor from 2000 to 2015

	PD	GDPPC	NVPC	IS	CC	EC
2000	0.475	0.067	0.252	0.309	0.377	0.473
2001	0.548	0.165	0.271	0.294	0.337	0.383
2002	0.565	0.098	0.175	0.298	0.380	0.449
2003	0.523	0.078	0.144	0.360	0.417	0.573
2004	0.658	0.115	0.272	0.242	0.318	0.462
2005	0.593	0.184	0.245	0.333	0.314	0.450
2006	0.586	0.080	0.139	0.386	0.288	0.366
2007	0.615	0.063	0.193	0.376	0.296	0.375
2008	0.590	0.042	0.183	0.229	0.252	0.434
2009	0.568	0.141	0.248	0.305	0.244	0.384
2010	0.562	0.146	0.269	0.199	0.294	0.387
2011	0.574	0.202	0.268	0.341	0.362	0.415
2012	0.604	0.121	0.162	0.239	0.229	0.284
2013	0.548	0.169	0.166	0.376	0.184	0.378
2014	0.575	0.113	0.219	0.376	0.179	0.423
2015	0.425	0.210	0.147	0.318	0.184	0.209

All the results of q value are significant at 1% level.

To obtain a more intuitive understanding of q values in all years, the results were illustrated in the box-whisker plot (Fig. 8), from which we can know: in the study period, the mean of q values for population density is relatively the highest among others, following electricity consumption, industry structure, coal consumption, number of vehicles per capita and GDP per capita. With the smallest interquartile range, PD has an influence of long-term stability on $PM_{2.5}$ pollution, which means that population density can predominantly explain the spatial heterogeneity of $PM_{2.5}$

pollution stably. EC and GDPPC also have relatively narrow interquartile ranges. There are relatively high dispersions in data of factors like NVPC, IS and CC, revealing unstable or changing patterns of trend for them on $PM_{2.5}$ pollution.

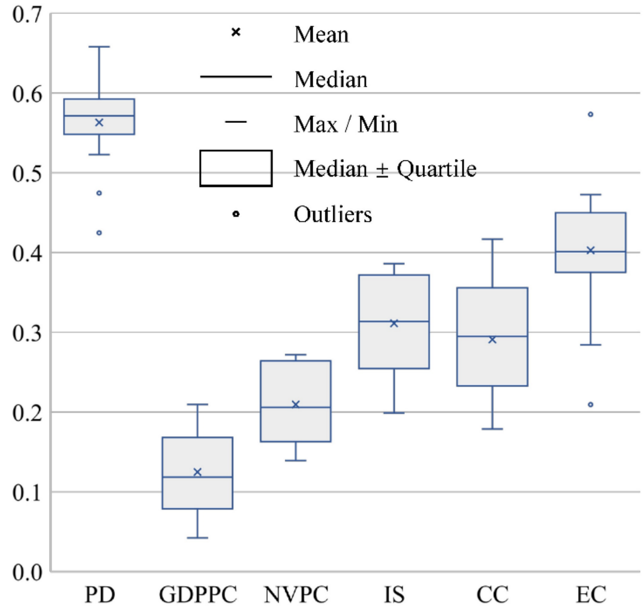


Fig. 8. Box-whisker plot of q values for each driving factor over 2000 to 2015

To explore the changing patterns of these driving factors, the results of q values were also illustrated in scatter plots (Fig. 9). The year trend lines were pictured by quadratic polynomials, with R-squared labeled near the lines. As shown in Fig. 9, the trend of PD is basically stable over the study period, revealing a gentle trend to rise first and then fall. Long term uptrend exists in both GDPPC. The upward trend of GDPPC indicates an increasing power of determinant on $PM_{2.5}$ concentrations, meaning that the relationship between GDPPC and $PM_{2.5}$ has an increasing marginal effect.

Interestingly, though relatively high at the very beginning, the q values of CC and EC show remarkable decline over the years, which means the impact of these two factors on $PM_{2.5}$ concentrations have weakened gradually. One possible explanation could be the implementation of air pollutant standard policies. Since thermal power generation accounts for up to 70% of the country's total power generation, and thermal power coal consumption accounts for 50% of total coal consumption, the policies for thermal power plants can greatly affect the environmental impact of coal consumption. The Emission Standard of Air Pollutants for Thermal Power Plants (GB13223-2011) (MEE, 2011) was first published by the Chinese Ministry of Environmental

Protection in 1991 and then revised in 1996, 2003 and 2011, with increasingly stringent emission standards. In order to meet stringent air emission standards, for example, post-combustion technologies or emission monitoring systems have been implemented to lower the impact of coal consumption on $PM_{2.5}$ concentrations. Similarly, in China, industrial electricity consumption accounts for 70% to 75% of total electricity consumption (2000-2016). Therefore, the *q* values of EC largely reflect the impact of industrial activities on $PM_{2.5}$ concentrations. From this we can argue that, compared to value added of industry in GDP, industrial coal consumption and industrial electricity consumption can better explain the environmental press of industrial activity.

As to NVPC and IS, the R^2 values of trend lines are 0.0253 and 0.0636, respectively, meaning that nonnegligible fluctuations of determinant power exist for these two indicators. Therefore, the trends of NVPC and IS are not that clear as with the other factors. Nevertheless, it is notable that, in spite of the increasing ownership of vehicles and the elevating contribution of nitrate which mainly from vehicle emission, the determinant power of NVPC did not reveal steady upward trend along the study period. The possible reason for this might have been the reduction in air pollutant emissions from vehicles, influenced by air pollution control policies.

According to existing research, ions in the air contributing (approximately 30%) to $PM_{2.5}$, mainly include nitrate ions and sulphate ions, of which nitrate ions are mostly formed by nitric oxide (NO_x) emitted from vehicles. However, despite the increasing ownership of vehicles, air pollutant emissions from vehicle have been declining recently. For example, vehicle emission of NO_x was firstly included and recorded in 2012, from when the emissions of NO_x have reduced from 5.829 million tons to 5.328 million tons in 2017 (MEE, 2018), which would reduce the impact of vehicles emissions on $PM_{2.5}$ concentrations through reduced nitrate formations. Moreover, series of stricter standards for new vehicle production and vehicle fuels have been adopted ever since the year 2000. For instance, the National I Standard, National II Standard, National III Standard and National IV Standard for light gasoline vehicles were enforced in 2001, 2005, 2008 and 2011, respectively (MEE, 2018), making vehicle emissions meeting more and more strict emission standards. Other favorable policies include elimination of older cars which are heavy polluting vehicles, and increasing promotion of new energy vehicles.

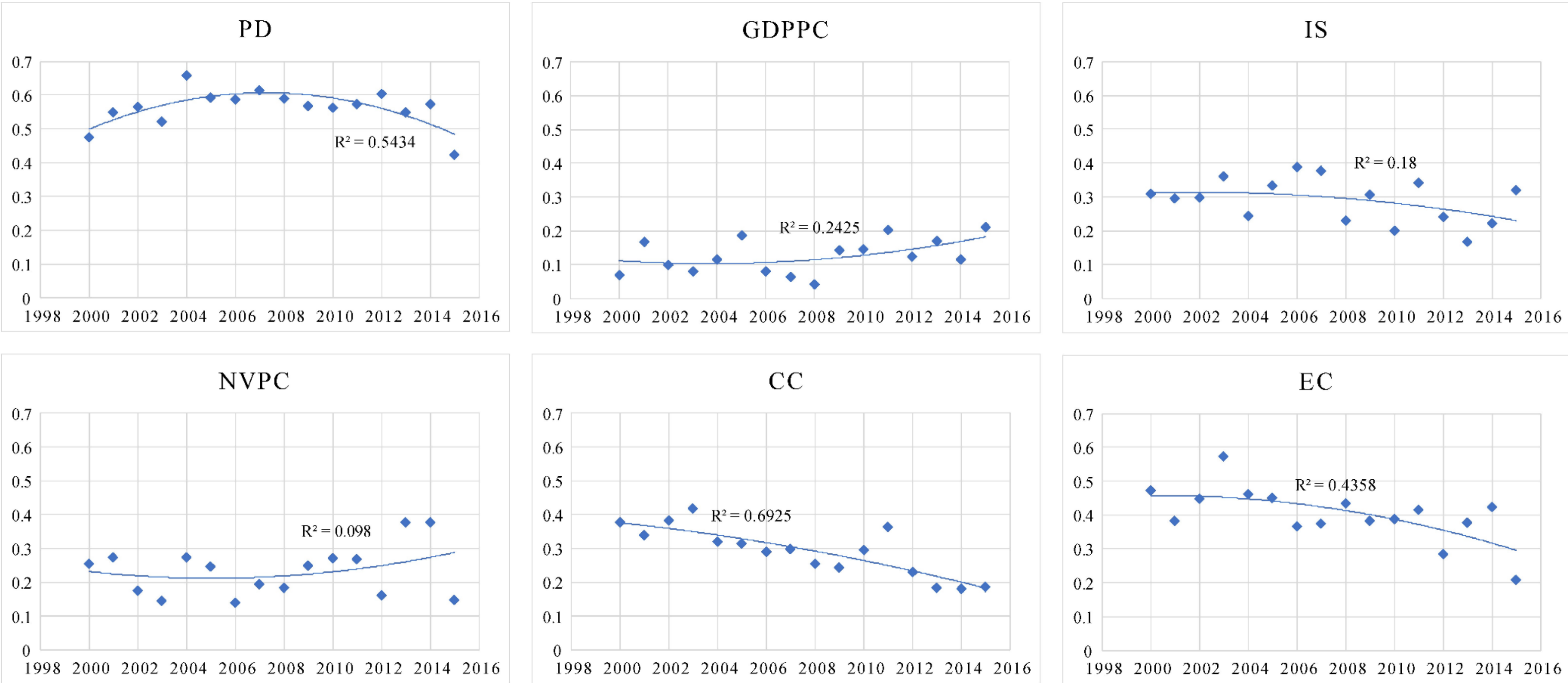


Fig. 9. Scatter plots with trend lines of *q* values for each driving factor over 2000 to 2015

Moreover, up on closer inspection, we can find an interesting fact that the q values of GDPPC, IS, NVPC, CC and EC all show an obvious decrease in the year 2012 compared to the previous year. Combined with the trend of average $PM_{2.5}$ concentrations in China (Fig. 2), in which the average $PM_{2.5}$ concentration of 2012 shows a relatively lowest level among recently years, we believe that, there exist other potential factors making significant impact on $PM_{2.5}$ concentrations. Specifically, the impact of policy factors such as stricter environmental regulations, has made the impact of the drivers examined in this paper less effective. For instance, the Ambient Air Quality Standards (GB3095-2012) was implemented in the year 2012, in which the readings of $PM_{2.5}$ were included and monitored for the first time in China (MEE, 2012). Another example is, with the goal of improving air quality, the Ministry of Environmental Protection of China has signed the "Responsibility Letter for Air Pollution Prevention and Control Targets" with 31 provinces, municipalities and autonomous regions in China in the year 2014, clearly specifying the targets of decreasing the average annual concentration of $PM_{2.5}$ in each region. The implementation of these regulations has aroused the attention of both the government and the public to the environmental hazards of $PM_{2.5}$, and may also greatly promote the governance of $PM_{2.5}$ and the improvement of air quality.

4.Conclusions and policy recommendations

In this paper, we firstly studied the temporal-spatial patterns of $PM_{2.5}$ concentrations using global Moran's I and LISA. Based on the results of spatial autocorrelation analysis, geographical detector method was applied to explain the spatial heterogeneity of regional $PM_{2.5}$ concentrations in China over the study period from 2010 to 2015 and identify the driving factors and their impact on $PM_{2.5}$ pollution. The main results are as follows:

1) $PM_{2.5}$ concentrations in China showed a notably rise from 2000 to 2007 and a basically flat and slightly downward trend from 2007 to 2015. Among the three major economic growth poles, the $PM_{2.5}$ situations in the Beijing-Tianjin-Hebei region (ranging from 30.6 μg/m³ to 53.3 μg/m³) and the Yangtze River Delta region (ranging from 28.8 μg/m³ to 48.8 μg/m³) are relatively serious, with the annual average reaching the highest both in 2007. The annual average concentration of $PM_{2.5}$ in the Pearl River Delta region is relatively low (ranging from 17.5 μg/m³ to 35.0 μg/m³), but it is still about twice the national average (ranging from 11.5 μg/m³ to 18.7 μg/m³).

2) Regional $PM_{2.5}$ concentrations revealed significant global and local spatial autocorrelation. The global Moran's *I* of $PM_{2.5}$ concentrations over the study period indicates a stable and high level of spatial autocorrelations. With the help of LISA, we found that High-High cluster mainly distributed in relatively developed regions such as Beijing, Shandong, Jiangsu, and Shanghai, while Low-Low cluster mostly distributed in western China such as Xinjiang, Tibet and Qinghai. No sign of significant High-Low outlier or Low-High outlier exist.

3) By the means of geographical detector, all the 6 driving factors explored in this paper have significant impact on $PM_{2.5}$ concentration. From a long-term perspective, population density, among others, has the greatest power of determinant on $PM_{2.5}$ pollution, followed by electricity consumption, industry structure, coal consumption, number of vehicles per capita and GDP per capita. The impact of coal consumption and electricity consumption has a significant downward trend along the study period, while the impact of population density show a trend to first rise and then fall.

4) From the results of GDM analysis, we found that the impact of most factors (PD, NVPC, CC and EC) showed downward trends. Even though the trend of NVPC had fluctuations, it had no dramatical raise during the study period. Therefore, consider the totally upward trend of PM2.5 pollution during the study period, there might be other factors influencing PM2.5 concentration with increasing power of determinant, which leaves us questions about the impacts of other factors on $PM_{2.5}$ pollution. Moreover, despite the power of determinant of certain factors, in GDM analysis, the influence directions cannot be told, which is also a problem to be overcome in future research.

Based upon the findings above, we propose some policy recommendations as follows.

1) Due to the significant spatial autocorrelations between regions, the government should pay attention to the importance of regional joint governance mechanisms in the $PM_{2.5}$ governance process. Considering the spatial spillover effect of air pollution, air quality improvement brought by short-distance industrial transfer can only be short-term and temporary. In addition, if regional linkages of environmental regulation are not emphasized, implementation of strict environmental regulations in a separate region cannot bring the full benefits to this region.

2) Strict emission standards and exhaust gas purification technologies turned out effective and

need further implementations. Even though the total coal consumption in China increased year by year during the study period, the $PM_{2.5}$ pollution situation was basically flat and slightly decreased from 2006 to 2012. Specifically, though the proportion of lignite imports rose from 0.9% in 2004 to 12.3% in 2010, the increasing consumption of lignite did not result in increasing impact on $PM_{2.5}$ pollution, consider the combustion of lignite would cause heavy pollutant emission. In fact, according to the analysis results above, q values of coal consumption have been declining significantly during these years. This reveals the importance of clean utilization of coal, namely, strict emission standards, advanced exhaust gas purification technologies, and a sound supervision system for coal consumptions. As is widely believed, the high proportion of coal consumption in the total energy structure in China would not change in the near future, thus, clean utilization of coal should be enhanced and continued.

Acknowledgements

The authors gratefully acknowledge the financial support from the Fundamental Research Funds for the Central Universities (2015XKMS069), the Natural Science Foundation of Jiangsu Province (BK20160246), and National Natural Science Foundation of China (71804182). We also would like to thank the anonymous referees for their helpful suggestions and corrections on the earlier draft of our paper, and upon which we have improved the content.

Appendix. Supplementary data

Supplementary data related to this article can be found online at:

http://sedac.ciesin.columbia.edu/data/set/sdei-global-annual-gwr-pm2-5-modis-misr-seawifs-aod/data-download

References

Anselin, L., 1995, Local indicator of spatial association-LISA. Geographical Analysis 27, 91-115

Anselin, L., 2005. Exploring spatial data with GeoDa™: a workbook., Centre for Spatially Integrated Social Science, Spatial Analysis Laboratory, Department of Geography, University of Illinois, Urbana Champaign.

Atkinson, R.W., Kang, S., Anderson, H.R., Mills, I.C., Walton, H.A., 2014, Epidemiological time series studies of $PM_{2.5}$ and daily mortality and hospital admissions: a systematic review and meta-analysis. Thorax 69, 660-665. https://doi.org/10.1136/thoraxjnl-2013-204492

Auffhammer, M., Carson, R.T., 2008, Forecasting the path of China's CO_2 emissions using province-level information. Journal of Environmental Economics and Management 55, 229-247. https://doi.org/10.1016/j.jeem.2007.10.002

Bozlaker, A., Spada, N.J., Fraser, M.P., Chellam, S., 2014, Elemental characterization of $PM_{2.5}$ and PM_{10} emitted from light duty vehicles in the Washburn Tunnel of Houston, Texas: Release of rhodium, palladium, and platinum. Environ. Sci. Technol. 48, 54-62. https://doi.org/10.1021/es4031003

Brauer, M., Amann, M., Burnett, R.T., Cohen, A., Dentener, F., Ezzati, M., Henderson, S.B., Krzyzanowski, M., Martin, R.V., Van Dingenen, R., van Donkelaar, A., Thurston, G.D., 2012, Exposure assessment for estimation of the global burden of disease attributable to outdoor air pollution. Environ. Sci. Technol. 46, 652-660. https://doi.org/10.1021/es2025752

Chalbot, M.G., Jones, T.A., Kavouras, I.G., 2014, Trends of non-accidental, cardiovascular, stroke and lung cancer mortality in Arkansas are associated with ambient $PM_{2.5}$ reductions. Int J. Env Res Pub He 11, 7442-7455. https://doi.org/10.3390/ijerph110707442

Cheng, Z.H., Li, L.S., Liu, J., 2017, Identifying the spatial effects and driving factors of urban $PM_{2.5}$ pollution in China. Ecol. Indic. 82, 61-75. https://doi.org/10.1016/j.ecolind.2017.06.043

Dayan, U., Erel, Y., Shpund, J., Kordova, L., Wanger, A., Schauer, J.J., 2011, The impact of local sources and meteorological factors on nitrogen oxide and particulate matter concentrations: A case study of the Day of Atonement in Israel. Atmos. Environ. 45, 3325-3332. https://doi.org/10.1016/j.atmosenv.2011.02.017

Djalalova, I., Wilczak, J., McKeen, S., Grell, G., Peckham, S., Pagowski, M., DelleMonache, L., McQueen, J., Tang, Y., Lee, P., 2010, Ensemble and bias-correction techniques for air quality model forecasts of surface O_3 and $PM_{2.5}$ during the TEXAQS-II experiment of 2006. Atmos. Environ. 44, 455-467. https://doi.org/10.1016/j.atmosenv.2009.11.007

Du, Z.Q., Xu, X.M., Zhang, H., Wu, Z.T., Liu, Y., 2016, Geographical detector-based identification of the impact of major determinants on aeolian desertification risk. Plos One 11. https://doi.org/10.1371/journal.pone.0151331

Emeis, S., Schäfer, K., Münkel, C., 2008, Surface-based remote sensing of the mixing-layer height - a review. Meteorol. Z. 17, 621-630. https://doi.org/10.1127/0941-2948/2008/0312

Fang, C.L., Liu, H.M., Li, G.D., 2016, International progress and evaluation on interactive coupling effects between urbanization and the eco-environment. J. Geogr Sci 26, 1081-1116. https://doi.org/10.1007/s11442-016-1317-9

Fredriksson, P.G., Millimet, D.L., 2002, Strategic interaction and the determination of environmental policy across U.S. states. Journal of Urban Economics 51, 101-122. https://doi.org/10.1006/juec.2001.2239.

Gao, L.J., Tian, Y.Z., Zhang, C.Y., Shi, G.L., Hao, H.Z., Zeng, F., Shi, C.L., Zhang, M.G., Feng, Y.C., Li, X., 2014, Local and long-range transport influences on $PM_{2.5}$ at a cities-cluster in northern China, during summer 2008. Particuology 13, 66-72. https://doi.org/10.1016/j.partic.2013.06.006

Halkos, G.E., Paizanos, E.A., 2013, The effect of government expenditure on the environment: An empirical investigation. Ecol. Econ. 91, 48-56. https://doi.org/10.1016/j.ecolecon.2013.04.002

Han, L.J., Zhou, W.Q., Pickett, S.T.A., Li, W.F., Li, L., 2016, An optimum city size? The scaling relationship for urban population and fine particulate ($PM_{2.5}$) concentration. Environ. Pollut. 208, 96-101. https://doi.org/10.1016/j.envpol.2015.08.039

Hao, Y., Liu, Y.M., 2016, The influential factors of urban $PM_{2.5}$ concentrations in China: a spatial econometric analysis. J. Clean. Prod. 112, 1443-1453. https://doi.org/10.1016/j.jclepro.2015.05.005

He, J., 2009, China's industrial SO2 emissions and its economic determinants: EKC's reduced vs. structural model and the role of international trade. Environment and Development Economics 14, 227. https://doi.org/10.1017/S1355770X0800452X

He, J.J., Gong, S.L., Yu, Y., Yu, L.J., Wu, L., Mao, H.J., Song, C.B., Zhao, S.P., Liu, H.L., Li, X.Y., Li, R.P., 2017, Air pollution characteristics and their relation to meteorological conditions during 2014–2015 in major Chinese cities. Environ. Pollut. 223, 484-496. https://doi.org/10.1016/j.envpol.2017.01.050

Hu, X.F., Waller, L.A., Al-Hamdan, M.Z., Crosson, W.L., Estes, M.G., Estes, S.M., Quattrochi, D.A., Sarnat, J.A., Liu, Y., 2013, Estimating ground-level $PM_{2.5}$ concentrations in the southeastern U.S. using geographically weighted regression. Environ. Res. 121, 1-10.

https://doi.org/10.1016/j.envres.2012.11.003

Hussein, T., Karppinen, A., Kukkonen, J., Härkönen, J., Aalto, P.P., Hämeri, K., Kerminen, V., Kulmala, M., 2006, Meteorological dependence of size-fractionated number concentrations of urban aerosol particles. Atmos. Environ. 40, 1427-1440. https://doi.org/10.1016/j.atmosenv.2005.10.061

Jenks, G.F., 1967. The data model concept in statistical mapping., International Yearbook of Cartography, pp. 186-190

Kaika, D., Zervas, E., 2013, The Environmental Kuznets Curve (EKC) theory—Part A: Concept, causes and the CO_2 emissions case. Energ. Policy 62, 1392-1402. https://doi.org/10.1016/j.enpol.2013.07.131.

Kan, H.D., Chen, R.J., Tong, S.L., 2012, Ambient air pollution, climate change, and population health in China. Environ. Int. 42, 10-19. https://doi.org/10.1016/j.envint.2011.03.003

Karagulian, F., Belis, C.A., Dora, C.F.C., Prüss-Ustün, A.M., Bonjour, S., Adair-Rohani, H., Amann, M., 2015, Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level. Atmos. Environ. 120, 475-483. https://doi.org/10.1016/j.atmosenv.2015.08.087

Kuznets, S., 1955, Economic growth and income inequality. American Economic Review 45, 1-28

Lelieveld, J., Evans, J.S., Fnais, M., Giannadaki, D., Pozzer, A., 2015, The contribution of outdoor air pollution sources to premature mortality on a global scale. Nature 525, 367-371. https://doi.org/10.1038/nature15371

Li, L., Qian, J., Ou, C.Q., Zhou, Y.X., Guo, C., Guo, Y.M., 2014, Spatial and temporal analysis of Air Pollution Index and its timescale-dependent relationship with meteorological factors in Guangzhou, China, 2001–2011. Environ. Pollut. 190, 75-81. https://doi.org/10.1016/j.envpol.2014.03.020

Li, M., Li, C., Zhang, M., 2018, Exploring the spatial spillover effects of industrialization and urbanization factors on pollutants emissions in China's Huang-Huai-Hai region. J. Clean. Prod. 195, 154-162. https://doi.org/10.1016/j.jclepro.2018.05.186

Li, M.N., Zhang, L.L., 2014, Haze in China: Current and future challenges. Environ. Pollut. 189, 85-86. https://doi.org/10.1016/j.envpol.2014.02.024

Liang, C.S., Duan, F.K., He, K.B., Ma, Y.L., 2016, Review on recent progress in observations, source identifications and countermeasures of $PM_{2.5}$. Environ. Int. 86, 150-170. https://doi.org/10.1016/j.envint.2015.10.016

Liu, H.M., Fang, C.L., Zhang, X.L., Wang, Z.Y., Bao, C., Li, F.Z., 2017, The effect of natural and anthropogenic factors on haze pollution in Chinese cities: A spatial econometrics approach. J. Clean. Prod. 165, 323-333. https://doi.org/10.1016/j.jclepro.2017.07.127

Lou, C.R., Liu, H.Y., Li, Y.F., Li, Y.L., 2016, Socioeconomic drivers of $PM_{2.5}$ in the accumulation phase of air pollution episodes in the Yangtze River Delta of China. Int J. Env Res Pub He 13. https://doi.org/10.3390/ijerph13100928

Lyu, W.N., Li, Y., Guan, D.B., Zhao, H.Y., Zhang, Q., Liu, Z., 2016, Driving forces of Chinese primary air pollution emissions: an index decomposition analysis. J. Clean. Prod. 133, 136-144. https://doi.org/10.1016/j.jclepro.2016.04.093

Ma, Y.R., Ji, Q., Fan, Y., 2016, Spatial linkage analysis of the impact of regional economic activities on $PM_{2.5}$ pollution in China. J. Clean. Prod. 139, 1157-1167. https://doi.org/10.1016/j.jclepro.2016.08.152

Madrigano, J., Kloog, I., Goldberg, R., Coull, B.A., Mittleman, M.A., Schwartz, J., 2013, Long-term exposure to $PM_{2.5}$ and incidence of acute myocardial infarction. Environ. Health Persp. 121, 192-196. https://doi.org/10.1289/ehp.1205284

Matus, K., Nam, K.M., Selin, N.E., Lamsal, L.N., Reilly, J.M., Paltsev, S., 2012, Health damages from air pollution in China. Global Environmental Change 22, 55-66. https://doi.org/10.1016/j.gloenvcha.2011.08.006

Mazeikis, A., 2013, Urbanization influence on meteorological parameters of air pollution: Vilnius case study. Baltica 26, 51-56. https://doi.org/10.5200/baltica.2013.26.06

MEE, Ministry of Ecology and Environment of People's Republic of China, 2011, Emission standard of air pollutants for thermal power plants, (GB13223-2011), Beijing, China, 2011/07/29. http://kjs.mee.gov.cn/hjbhbz/bzwb/dqhjbh/dqgdwrywrwpfbz/201109/t20110921_217534.shtml

MEE, Ministry of Ecology and Environment of People's Republic of China, 2012, Ambient air quality standards, (GB3095-2012), Beijing, China, 2012/02/29.

http://www.mee.gov.cn/gkml/hbb/bgg/201203/t20120302_224145.htm

MEE, Ministry of Ecology and Environment of People's Republic of China, 2018, China vehicle environmental management annual report (2018), (accessed date: Sep. 12, 2018). http://www.mee.gov.cn/gkml/sthjbgw/qt/201806/t20180601_442293.htm

Miao, Y., Liu, S., 2019, Linkages between aerosol pollution and planetary boundary layer structure in China. Sci. Total Environ. 650, 288-296. https://doi.org/10.1016/j.scitotenv.2018.09.032

Mitchel, A., Esri, 2005, The ESRI Guide to GIS analysis, Volume 2: Spatial measurements and statistics, (accessed date: Sep. 12, 2018). http://agris.fao.org/openagris/search.do?recordID =SO2007100043.

Moran, P.A.P., 1950, Notes on continuous stochastic phenomena. Biometrika 37, 17-23

Ostro, B., Malig, B., Broadwin, R., Basu, R., Gold, E.B., Bromberger, J.T., Derby, C., Feinstein, S., Greendale, G.A., Jackson, E.A., Kravitz, H.M., Matthews, K.A., Sternfeld, B., Tomey, K., Green, R.R., Green, R., 2014, Chronic $PM_{2.5}$ exposure and inflammation: Determining sensitive subgroups in mid-life women. Environ. Res. 132, 168-175. https://doi.org/10.1016/j.envres.2014.03.042

Pao, H.T., Tsai, C.M., 2010, CO_2 emissions, energy consumption and economic growth in BRIC countries. Energ. Policy 38, 7850-7860. https://doi.org/10.1016/j.enpol.2010.08.045

Platt, S.M., Haddad, I.E., Pieber, S.M., Huang, R.J., Zardini, A.A., Clairotte, M., Suarez-Bertoa, R., Barmet, P., Pfaffenberger, L., Wolf, R., Slowik, J.G., Fuller, S.J., Kalberer, M., Chirico, R., Dommen, J., Astorga, C., Zimmermann, R., Marchand, N., Hellebust, S., Temime-Roussel, B., Baltensperger, U., Prévôt, A.S.H., 2014, Two-stroke scooters are a dominant source of air pollution in many cities. Nat Commun 5. https://doi.org/10.1038/ncomms4749

Quan, J., Tie, X., Zhang, Q., Liu, Q., Li, X., Gao, Y., Zhao, D., 2014, Characteristics of heavy aerosol pollution during the 2012–2013 winter in Beijing, China. Atmos. Environ. 88, 83-89. https://doi.org/10.1016/j.atmosenv.2014.01.058

Salim, R.A., Shafiei, S., 2014, Urbanization and renewable and non-renewable energy consumption in OECD countries: An empirical analysis. Economic Modelling 38, 581-591. https://doi.org/10.1016/j.econmod.2014.02.008

Sofowote, U.M., Su, Y., Dabek-Zlotorzynska, E., Rastogi, A.K., Brook, J., Hopke, P.K., 2015, Constraining the factor analytical solutions obtained from multiple-year receptor modeling of ambient $PM_{2.5}$ data from five speciation sites in Ontario, Canada. Atmos. Environ. 108, 151-157. https://doi.org/10.1016/j.atmosenv.2015.02.045

Stern, D.I., 2004, The rise and fall of the Environmental Kuznets Curve. World Development 32, 1419-1439. https://doi.org/10.1016/j.worlddev.2004.03.004

Tobler, W.R., 1970, A computer movie simulating urban growth in the Detroit region. Economic geography 46, 234-240

Vakeva, M., Hameri, K., Puhakka, T., Nilsson, E.D., Hohti, H., Makela, J.M., 2000, Effects of meteorological processes on aerosol particle size distribution in an urban background area. J. Geophys. Res.-Atmos. 105, 9807-9821. https://doi.org/10.1029/1999JD901143

van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., Lyapustin, A., Sayer, A.M., Winker, D.M., 2016, Global estimates of fine particulate matter using a combined geophysical-statistical method with information from satellites. Environ. Sci. Technol. 50, 3762. https://doi.org/10.1021/acs.est.5b05833

van Donkelaar, A., Martin, R.V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., Lyapustin, A., Sayer, A.M., Winker, D.M., 2018. Global annual $PM_{2.5}$ grids from MODIS, MISR and SeaWiFS Aerosol Optical Depth (AOD) with GWR, 1998-2016., NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY

Vieira-Filho, M.S., Lehmann, C., Fornaro, A., 2015, Influence of local sources and topography on air quality and rainwater composition in Cubatão and São Paulo, Brazil. Atmos. Environ. 101, 200-208. https://doi.org/10.1016/j.atmosenv.2014.11.025

Wang, J.F., Hu, Y., 2012, Environmental health risk detection with GeoDetector. Environ. Modell. Softw. 33, 114-115. https://doi.org/10.1016/j.envsoft.2012.01.015

Wang, J.F., Li, X.H., Christakos, G., Liao, Y.L., Zhang, T., Xue, G., Zheng, X.Y., 2010, Geographical detectors-based health risk assessment and its application in the neural tube defects study of the Heshun Region, China. Int. J. Geogr. Inf. Sci. 24, 107-127. https://doi.org/10.1080/13658810802443457

Wang, J.F., Xu, C.D., 2012, GeoDetector, State Key Laboratory of Resources and Environmental Information System; Institute of Geographic Sciences and Natural Resources Research; Chinese

Academy of Sciences, Beijing, 100101, PR China.

Wang, X.Y., Wang, K.C., Su, L.Y., 2016, Contribution of atmospheric diffusion conditions to the recent improvement in air quality in China. Sci Rep-Uk 6. https://doi.org/10.1038/srep36404

Wang, Y., Wang, S.J., Li, G.D., Zhang, H.O., Jin, L.X., Su, Y.X., Wu, K.M., 2017, Identifying the determinants of housing prices in China using spatial regression and the geographical detector technique. Applied Geography 79, 26-36. https://doi.org/10.1016/j.apgeog.2016.12.003

Xu, B., Luo, L.Q., Lin, B.Q., 2016, A dynamic analysis of air pollution emissions in China: Evidence from nonparametric additive regression models. Ecol. Indic. 63, 346-358. https://doi.org/10.1016/j.ecolind.2015.11.012

Yang, G.H., Wang, Y., Zeng, Y.X., Gao, G.F., Liang, X.F., Zhou, M.G., Wan, X., Yu, S.C., Jiang, Y.H., Naghavi, M., Vos, T., Wang, H.D., Lopez, A.D., Murray, C.J.L., 2013, Rapid health transition in China, 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet 381, 1987-2015. https://doi.org/10.1016/S0140-6736(13)61097-1

Yu, H.L., Lin, Y.C., Kuo, Y.M., 2015, A time series analysis of multiple ambient pollutants to investigate the underlying air pollution dynamics and interactions. Chemosphre 134, 571-580. https://doi.org/10.1016/j.chemosphere.2014.12.007

Zhang, Y.L., Cao, F., 2015, Fine particulate matter ($PM_{2.5}$) in China at a city level. Sci Rep-Uk 5. https://doi.org/10.1038/srep14884

Zhang, Z.Y., Zhang, X.L., Gong, D.Y., Quan, W.J., Zhao, X.J., Ma, Z.Q., Kim, S.J., 2015, Evolution of surface O_3 and $PM_{2.5}$ concentrations and their relationships with meteorological conditions over the last decade in Beijing. Atmos. Environ. 108, 67-75. https://doi.org/10.1016/j.atmosenv.2015.02.071

Zhou, C.S., Chen, J., Wang, S.J., 2018, Examining the effects of socioeconomic development on fine particulate matter ($PM_{2.5}$) in China's cities using spatial regression and the geographical detector technique. Sci. Total Environ. 619-620, 436-445. https://doi.org/10.1016/j.scitotenv.2017.11.124