Environmental Pollution 241 (2018) 475–483

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Quantifying the influence of natural and socioeconomic factors and their interactive impact on $PM_{2.5}$ pollution in China☆

Dongyang Yang [a], Xiaomin Wang [b], Jianhua Xu [a, *], Chengdong Xu [c, **], Debin Lu [a, d], Chao Ye [a], Zujing Wang [e], Ling Bai [f]

[a] School of Geographic Sciences & Institute of Eco-Chongming, East China Normal University, Shanghai 200241, China
[b] School of Geography, Beijing Normal University, Beijing 100875, China
[c] State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Science and Natural Resource Research, Chinese Academy of Sciences, Beijing 100101, China
[d] Department of Tourism and Geography, Tongren University, Tongren, Guizhou Province 554300, China
[e] School of Environmental Science and Engineering, Suzhou University of Science and Technology, Suzhou, Jiangsu Province 215009, China
[f] School of Economics and Management, Nanchang University, Nanchang 330031, China

A R T I C L E I N F O

Article history:
Received 7 February 2018
Received in revised form
14 March 2018
Accepted 14 May 2018

Keywords:
$PM_{2.5}$
Natural and socioeconomic factors
Interaction effect
GeogDetector

A B S T R A C T

$PM_{2.5}$ pollution is an environmental issue caused by multiple natural and socioeconomic factors, presenting with significant spatial disparities across mainland China. However, the determinant power of natural and socioeconomic factors and their interactive impact on $PM_{2.5}$ pollution is still unclear. In the study, the GeogDetector method was used to quantify nonlinear associations between $PM_{2.5}$ and potential factors. This study found that natural factors, including ecological environments and climate, were more influential than socioeconomic factors, and climate was the predominant factor ($q = 0.56$) in influencing $PM_{2.5}$ pollution. Among all interactions of the six influencing factors, the interaction of industry and climate had the largest influence ($q = 0.66$). Two recognized major contaminated areas were the Tarim Basin in the northwest region and the eastern plain region; the former was mainly influenced by the warm temperate arid climate and desert, and the latter was mainly influenced by the warm temperate semi-humid climate and multiple socioeconomic factors. The findings provided an interpretation of the influencing mechanisms of $PM_{2.5}$ pollution, which can contribute to more specific policies aimed at successful $PM_{2.5}$ pollution control and abatement.

1. Introduction

$PM_{2.5}$ pollution's negative effects on public health have burdened China severely in recent years (Butt et al., 2017; Lu et al., 2015). As a geographical phenomenon, $PM_{2.5}$ pollution is inevitably influenced by the comprehensive geographical environment, including natural and socioeconomic conditions (Guo et al., 2014; Timmermans et al., 2017; Wang et al., 2014). China is characterized as a vast territorial area, and the natural and socioeconomic conditions and their interactions differ widely across regions, simultaneously, $PM_{2.5}$ pollution presents geographically discrepant distribution in China (Ma et al., 2016).

Multidisciplinary researchers have conducted numerous studies on the sources and driving factors of $PM_{2.5}$. Researchers have recognized industry, fossil fuel combustion, motor vehicles, and building yards as general anthropogenic sources of $PM_{2.5}$ in many cities in the world (Chowdhury et al., 2007; de Miranda et al., 2012; Huang et al., 2014; Marcazzan et al., 2003; Timmermans et al., 2017). Social and economic activities can affect the generation of and changes in these sources. Thus, some researchers have examined the relationships between $PM_{2.5}$ pollution and socioeconomic factors and indicated that socioeconomic factors such as economic growth, urbanization, industrialization, and others drove increases in $PM_{2.5}$ concentrations (Guan et al., 2014; Hao and Liu, 2016; Li et al., 2016; Meng et al., 2015).

$PM_{2.5}$ is a type of air pollutant and the transformation, diffusion,

☆ This paper has been recommended for acceptance by Haidong Kan.

* Corresponding author. School of Geographic Sciences & Institute of Eco-Chongming, East China Normal University, Shanghai 200241, China.
** Corresponding author. State Key Laboratory of Resources and Environmental Information System, Institute of Geographic Science and Natural Resource Research, Chinese Academy of Sciences, Beijing 100101, China.
E-mail addresses: Jhxu@geo.ecnu.edu.cn (J. Xu), xucd@lreis.ac.cn (C. Xu).

https://doi.org/10.1016/j.envpol.2018.05.043

and even the generation of areas severely polluted with $PM_{2.5}$ are inevitably affected by meteorological conditions. Generally, $PM_{2.5}$ pollution deteriorates and forms a thick haze in continuous and stable weather conditions (Green et al., 2015; Tao et al., 2014; Zhang and Cao, 2015). Long-term meteorological conditions are also related to the spatial and seasonal variability of $PM_{2.5}$, especially in China (Guo et al., 2017; Pant et al., 2015; Yang et al., 2017a; Zhang and Cao, 2015). The ecology or land use environment is the spatially common carrier of pollution sources and potential factors of $PM_{2.5}$; different ecological systems can also affect $PM_{2.5}$. Human settlements are undoubtedly concentrated primary and secondary anthropogenic emission sources (Cao et al., 2016; Guo et al., 2014; Zhang et al., 2015). Environmental chemists have also documented that fertilizer use on farmlands is the largest emission source of ammonia (NH_3), which mainly contributes to urban $PM_{2.5}$ by combining with urban emission pollutants, such as volatile organic compounds, NO_x, and SO_2(Gu et al., 2014; Wang et al., 2016a). Furthermore, air pollutants in city suburbs aggregate into the city and enhance $PM_{2.5}$ pollution under the urban heat islands effect (Aslam et al., 2017; Bloomer et al., 2009; Zhang et al., 2009). Other ecological systems also have different effects on $PM_{2.5}$; for example, desert areas are natural sources of sand and dust in $PM_{2.5}$ (Lu et al., 2017), and forests can absorb and purify $PM_{2.5}$ pollution (Nowak et al., 2014).

$PM_{2.5}$ is comprehensively affected by natural and socioeconomic factors, and the use of related factors have been highlighted with effectiveness in improving $PM_{2.5}$ mapping (Liu et al., 2018). However, the discrepant influences of the factors and their interactions have rarely been examined. Investigators are usually concerned with unilateral factors such as natural or socioeconomic factors, while a quantification of the influences of integrated natural and socioeconomic multi-factors has been overlooked. Traditional methods are disadvantaged in examining the interactions of factors influencing air pollution. The interaction of two factors can be multiple coupling forms in reality, while it is usually the product of two factors in traditional regression methods. Additionally, statistical models of interactions are usually created using local regions and are limited in reflecting the spatial global variability of influences (Pearce et al., 2011). Coefficients with spatial differences can be derived by using GIS-based regression methods or machine learning algorithms, but they have poor large-scale explanatory capability because of the existence of spatial stratified heterogeneity (Lin et al., 2014; Wang et al., 2016b).

China is experiencing tremendous economic growth and has increasingly become an important engine for world economic growth with its rapid urbanization and industrialization; it is also one of the main regions facing serious $PM_{2.5}$ pollution. Anthropogenic emissions drive the increases in $PM_{2.5}$, but not all concentrated areas of anthropogenic emissions are highly polluted because of the differences in natural conditions (Jin et al., 2017; Lu et al., 2017). Quantifying the influences of potential factors and their interactions are important for understanding the spatial pattern of $PM_{2.5}$ pollution and the driving mechanisms, which can contribute to successful policy making for controlling and reducing $PM_{2.5}$ pollution. This study aimed to quantitatively investigate the influence of potential factors, including industry, construction, traffic, coal combustion, ecological environments, and climate and their interactive effects on $PM_{2.5}$ pollution in China.

2. Materials and methods

2.1. $PM_{2.5}$ data

$PM_{2.5}$ data is a subset of a published global $PM_{2.5}$ concentration dataset for the study region in 2015 from the Atmospheric Composition Analysis Group at Dalhousie University (http://fizz.phys.dal.ca/~atmos/martin/?page_id=140) (Fig. 1). The spatial resolution of the raw data is $0.1° \times 0.1°$. This dataset has good accuracy with a high cross-validated R^2 of 0.81 (Van Donkelaar et al., 2016) and has been used in many research studies (Donkelaar et al., 2015; Lu et al., 2017; Mcguinn et al., 2017; Xie et al., 2016). The cross-validated R^2, between the estimated annual average $PM_{2.5}$ and observation values in 313 cities of China in 2015, is 0.72.

2.2. Potential driving factors, proxies, and data

Given that socioeconomic factors, including industry, construction, traffic, coal combustion, and natural factors, including ecological environment and climate, were related to aspects of the natural and anthropogenic sources, influences, and secondary influences of $PM_{2.5}$, we integrated these factors in this study. The corresponding proxies of these potential factors are shown in Fig. 2.

Socioeconomic data, including industrial output (IO), construction output (CO), and coal consumption (CC) in 2015 were obtained from the Provincial Statistical Yearbook on China National Knowledge Infrastructure (CNKI: http://data.cnki.net/Yearbook). The unit area mean values of the three types of socioeconomic data, which are the original values divided by the administrative area, were calculated and used. Data on roads were gathered from OpenStreetMap (https://www.openstreetmap.org). The roads included urban roads and multistage highways such as expressways; national, provincial, and county highways; and township roads. Road density (RD) was calculated using “line density” in spatial analysis tools in ArcGIS software. To match the following modeling needs, industrial output, construction output, and road density were discretized and classified using the geometrical interval method in ArcGIS according to their quantitative values, and coal consumption was classified using the natural breaks method (Fig. 3a–d).

Data on ecosystem types (ET) in 2015 were provided by the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences (http://www.resdc.cn). The spatial resolution of the raw data was $1\,km \times 1\,km$. To obtain the ecosystem types in each $10\,km \times 10\,km$ grid, we extracted the most common ecosystem type in each $10\,km \times 10\,km$ grid as the ecosystem type for each grid except human settlement. The area of human settlement was small compared with other ecosystem types, but human activity space and manmade pollutants, including $PM_{2.5}$, were not just confined to the spatial range of the human settlement. Thus, the ecosystem type where the total area of the human settlement was larger than $25\,km^2$ was considered a human settlement ecosystem (Fig. 3e).

China has diverse climatic conditions. Zheng et al. (2010) proposed a new scheme for generating China's climate regionalization (CR) based on daily meteorological observation data during 1971–2000. We regenerated the climate regionalization by vectorizing Zheng's scheme and merging some small partitions (Fig. 3f).

2.3. GeogDetector model

The GeogDetector model is a novel spatial variation analysis method, which can be used to explore the driving force of a responding variable under the assumption that the two variables are associated if their spatial stratified heterogeneity tend to be consistent (Wang et al., 2010; Wang and Hu, 2012). Compared to traditional linear models, GeogDetector is capable of handling categorical dependent variables, finding the dominant driving force, and investigating the interaction between two variables with no assumption of linearity and immunity to the colinearity of dependent variables.

It comprises four modules: factor detector, interaction detector,

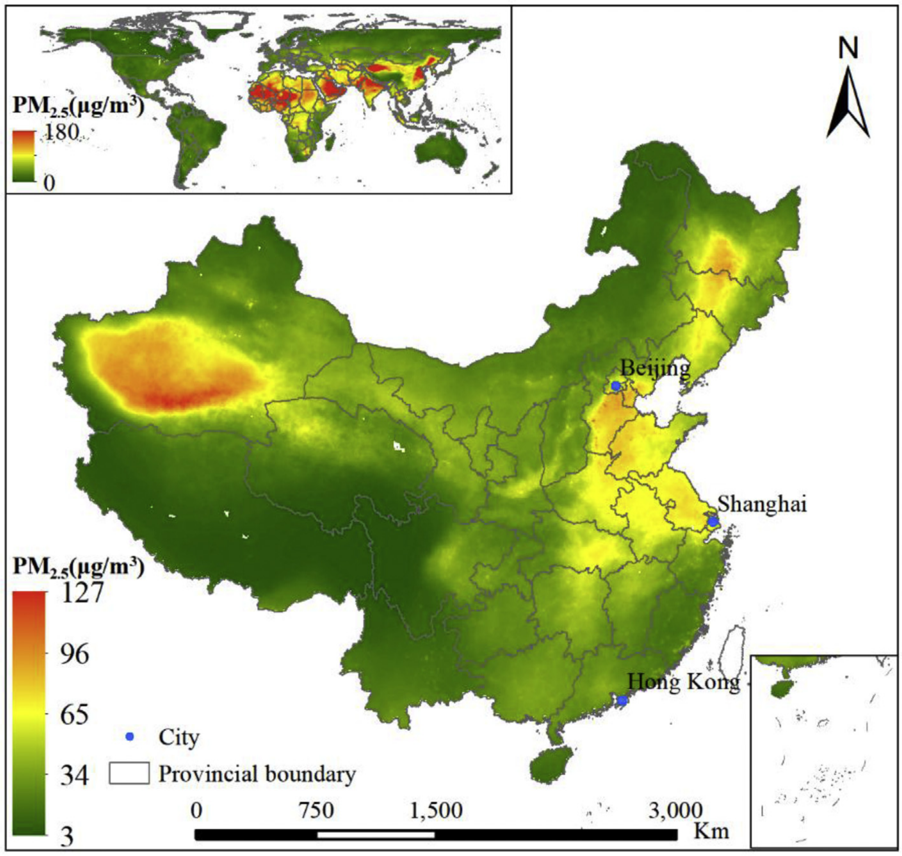

Fig. 1. Spatial pattern of $PM_{2.5}$ concentrations in 2015 for mainland China.

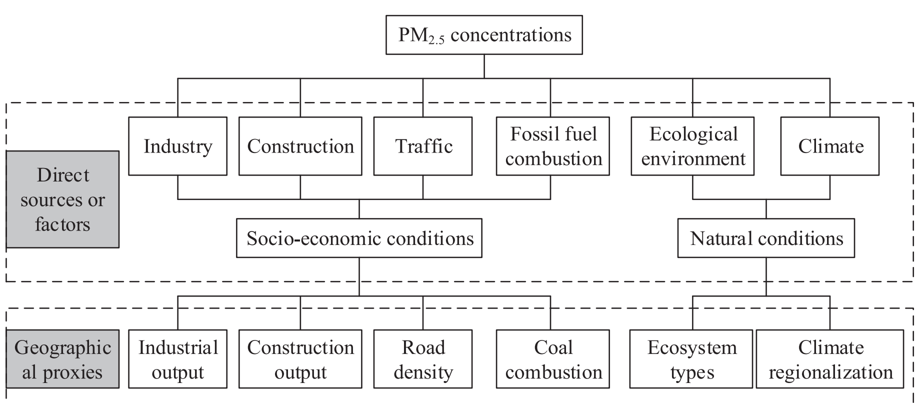

Fig. 2. Factors and their proxies.

risk detector, and ecological detector (Wang and Hu, 2012). The factor detector uses a q value to quantify the influences of factors (Xs) on Y; q is given by following formula:

$$q = 1 - \frac{\sum_{h=1}^{L} N_h \sigma_h^2}{N\sigma^2}$$

$$SSW = \sum_{h=1}^{L} N_h \sigma_h^2, \quad SST = N\sigma^2$$

where $h = 1, \ldots, L$ is the number of subregions (or subclasses) of factors X and $N_{h,}$ and N donates the number of samples in subregion h and the total number of samples over the whole study region, respectively; σ_h^2 and σ^2 denotes the variance of samples in subregion h and the global variance of Y over the entire study region. SSW and SST are the within sum of squares and the total sum of squares, respectively. The value of q ranges from 0 to 1. The larger the q value, the stronger the influence of variable X on Y.

The interaction detector can examine the interaction of different factors (Xs) and reveals whether the interaction of factors ($X1$ and $X2$) weaken or enhance the influence on Y or whether they are independent in influencing Y. The q value of factors $X1$ and $X2$ calculated from a factor detector can be marked as $q\,(X1)$ and $q\,(X2)$. A new factor layer and subregions can be generated by overlaying the factor layer $X1$ and $X2$ spatially; it can be marked as $X1 \cap X2$ and

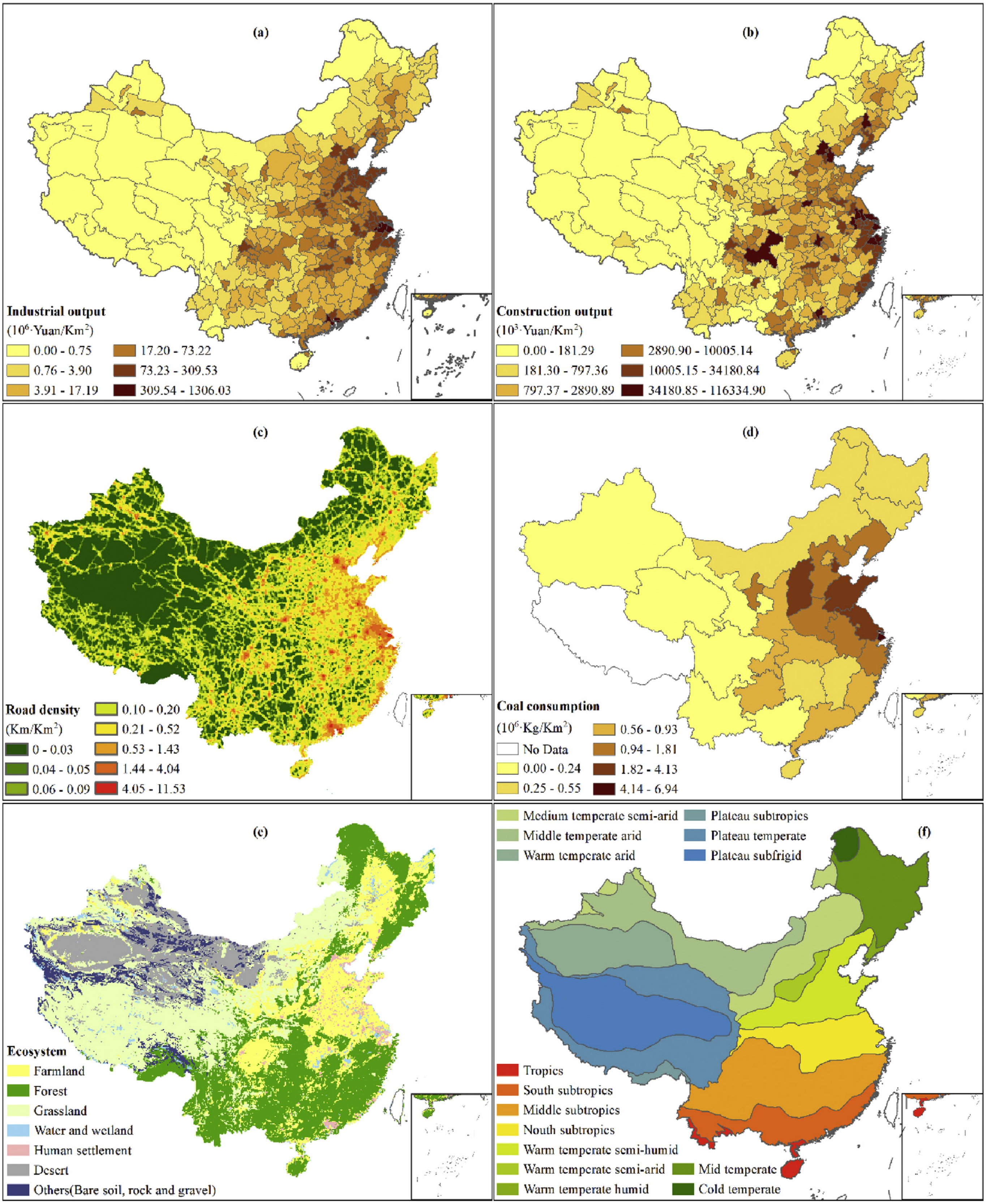

Fig. 3. Spatial distribution of six underlying factors: (a) industrial output, (b) construction output, (c) road density, (d) coal consumption, (e) ecosystem type, and (e) climate regionalization.

∩ denotes the intersection between the factor layer $X1$ and $X2$. Then, the q value of $X1 \cap X2$, represented as $q\ (X1 \cap X2)$, can be calculated. The interactive relationship can be interpreted as five categories by comparing the interactive q value of the two factors and the q value of each of the two factors (Wang and Hu, 2012; Wu et al., 2017). The five categories are presented in Table 1.

The spatial difference of an influencing factor should have different influences on Y in different regions. The risk detector can determine whether there is a significant difference in influence on Y in two subregions via a t-test. Its formula is:

$$t_{\overline{y}_{h=1}\overline{y}_{h=2}} = \frac{\overline{Y}_{h=1} - \overline{Y}_{h=2}}{\left[\frac{Var(\overline{Y}_{h=1})}{n_{h=1}} + \frac{Var(\overline{Y}_{h=2})}{n_{h=2}}\right]^{1/2}}$$

where $\overline{Y}_h$ represents the average of Y in the subregion h; n_h is the size of samples in subregion h, and Var is variance.

The ecological detector is used to compare whether $X1$ has a significantly greater influence or contribution than $X2$. It is measured using the statistics F:

$$F = \frac{N_{X1}(N_{X2} - 1)SSW_{X1}}{N_{X2}(N_{X1} - 1)SSW_{X2}}$$

$$SSW_X = \sum_{h=1}^{L1} N_h \sigma_h^2, SST_{X2} = \sum_{h=1}^{L2} N_h \sigma_h^2$$

where N_{X1} and N_{X2} represent the number of samples of the two factors $X1$ and $X2$, respectively; SSW_{X1} and SSW_{X2} are the within sum of squares in the subregion generated by factor layers $X1$ and $X2$, respectively. $L1$ and $L2$ represent the number of subregions of $X1$ and $X2$, respectively. The null hypothesis is defined as H_0: $SSW_{X1} = SSW_{X2}$. The rejected H_0 at the significance level α indicates that it is statistically significant.

In this research, the geographical detector model is used to examine the influence of six potential influencing factors, including industry, construction industry, traffic, coal combustion, ecological environment, and climate and their interaction effects on $PM_{2.5}$ concentrations.

3. Results

3.1. The influence of potential driving factors on $PM_{2.5}$ concentrations

The influence (q values) of the six driving factors on $PM_{2.5}$ concentrations were calculated using the factor detector and are presented in Table 2. Among the six potential driving factors, CR had the greatest influence on the pattern of $PM_{2.5}$ concentrations ($q = 0.56$), followed by another natural factor, ET ($q = 0.30$). Compared with natural factors, socioeconomic factors presented smaller influences on $PM_{2.5}$ concentrations. Among the socioeconomic factors, IO presented as having a dominant influence ($q = 0.14$) on $PM_{2.5}$ concentrations, while RD presented as having the least influence ($q = 0.10$).

Table 1
The interactive categories of two factors and the interactive relationship.

Description	Interaction
$q\ (X1 \cap X2) < \text{Min}\ (q\ (X1), q\ (X2))$	Weaken; univariate
$\text{Min}\ (q\ (X1), q\ (X2)) < q\ (X1 \cap X2) < \text{Max}\ (q\ (X1), q\ (X2))$	Weaken; univariate
$q\ (X1 \cap X2) > \text{Max}\ (q\ (X1), q\ (X2))$	Enhanced, bivariate
$q\ (X1 \cap X2) = q\ (X1) + q\ (X2)$	Independent
$q\ (X1 \cap X2) > q\ (X1) + q\ (X2)$	Nonlinearly enhance

Table 2
The influences of factors in driving $PM_{2.5}$ concentrations.

Factors		q value
Socioeconomic factors	IO ($10^6 \cdot$Yuan/Km^2)	0.14***
	CO($10^3 \cdot$Yuan/Km^2)	0.12***
	RD (Km/K m^2)	0.10***
	CC($10^6 \cdot$Kg/Km^2)	0.13***
Natural factors	ET	0.30***
	CR	0.56***

Note: ***donates that q value is significant at the 0.001 level ($p < 0.001$).

3.2. The interaction effects of factors on $PM_{2.5}$ concentrations (interaction detector)

A total of 15 pairs of interactions between the 6 factors were detected using the interaction detector. The interactive q value of each pair of factors was found to be more than both the two factors' q values but was less than the sum of the two factors' q values. Thus, the interactive relationship between each pair of factors was bivariate enhanced each other in influencing $PM_{2.5}$ concentrations. Fig. 4 shows the specific comparisons between the interactive q value and the two factors' q values. Among the interactions of socioeconomic factors, q (IO∩CC) was the maximum (0.21), indicating that the interaction between IO and CC was strongest. Also, the interaction was strongest between IO and ET among the interactions between socioeconomic factors and ET, and the interaction between IO and CR was strongest among all the interactions between socioeconomic factors and natural factors. Additionally, the interaction between natural factors, ET and CR, was not significantly enhanced compared with both the original q values of ET and CR.

3.3. The leading impact areas (subregions) of factors in influencing $PM_{2.5}$ concentrations

The average $PM_{2.5}$ concentrations in all subregions of the six factors were calculated and the significance of influence differences was captured using the risk detector. The socioeconomic factors were divided according to their numerical values from small to large. $PM_{2.5}$ concentrations in all subregions of the socioeconomic factors presented increasing trends along with increases in these factors generally (Fig. 5), indicating that larger the values of these

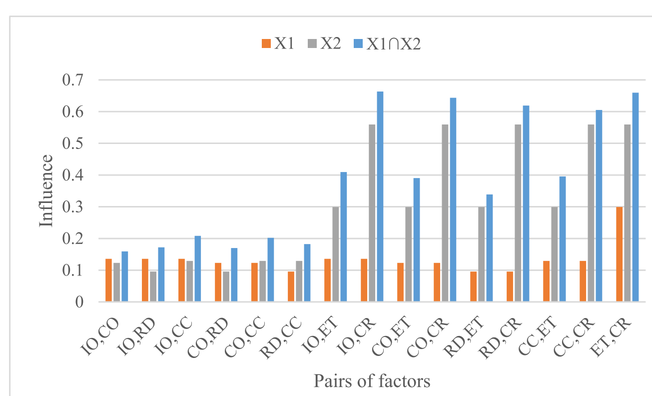

Fig. 4. The comparison of the interactive q value and the original q value of each pair of factors.
Note that X1 donates the first factor, X2 donates the second factor, and X1∩X2 is the interaction of the two factors. For example, in the pair of (IO, CO), X1 donates IO, X2 donates CO, and X1∩X2 is the interaction of IO and CO.

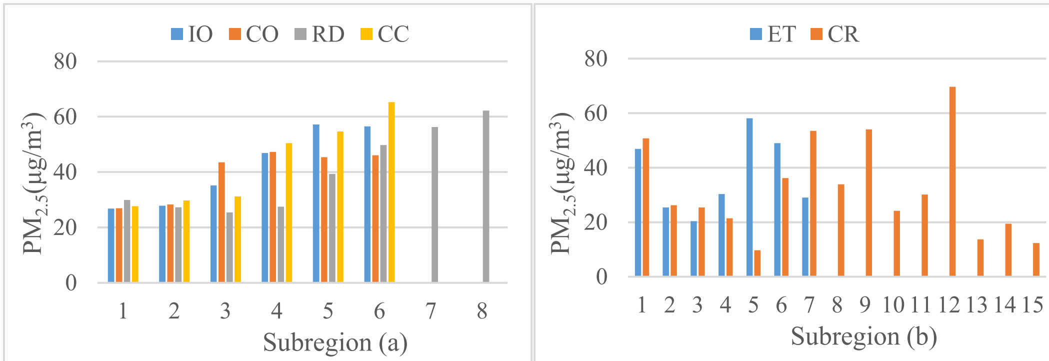

Fig. 5. The comparison of the interactive *q* value and the original q value of each pair of factors.
Note: In Figure a, numbers 1–6 (or 8) in the abscissa axis denote the number of subregions (subclasses) of the socioeconomic factors, which were divided based on their values from small to large. In Figure b, numbers 1–7 (or 15) in the abscissa axis represent the number of subregions (subclasses) of the natural factors; the corresponding relations are presented in Table A. 1.

factors were associated with higher $PM_{2.5}$ concentrations. The average $PM_{2.5}$ concentration was highest (58.15 μg/m^3) in the subregion of human settlement in ET, followed by desert (49.04 μg/m^3) and farmland (46.89 μg/m^3). Among subregions of CR, the average $PM_{2.5}$ concentration was the highest (69.67 μg/m^3) in the subregion with a warm temperate arid climate, followed by warm temperate semi-humid climate (54.02 μg/m^3), warm temperate humid climate (53.52 μg/m^3), and north subtropics (50.74 μg/m^3).

The area where the $PM_{2.5}$ concentrations were high can be identified as a high-risk area for $PM_{2.5}$ pollution. The first several subregions of each factor where the average $PM_{2.5}$ concentrations were relatively high (more than 40 μg/m^3 generally) were mapped to inform the areas with high risk of $PM_{2.5}$ pollution (Fig. 6). The eastern plain area, especially the North China Plain, was affected by multiple factors and was a high risk area for $PM_{2.5}$ pollution, while the northwest region, especially the Tarim Basin region which is mainly affected by deserts and a warm temperate arid climate, was also a high-risk area for $PM_{2.5}$ pollution.

The significance of varying influence among different subregions of IO showed that there were no significant differences between subregions 5 and 6, and there were significant differences between other pairs of subregions (Table A. 2). Similarly, there were no significant differences between subregions 5 and 6, and there were significant differences between other pairs of subregions of CO (Table A. 3). For factor RD, there were significant differences between all pairs of subregions except the pair of subregions 2 and 4 (Table A. 4). The results of significant differences were all true between all pairs of subregions of both CC and ET (Table A. 5 and A. 6). The significance of varying influences of CR was true except the pair of warm temperate semi-humid climate region and warm temperate humid climate region (Table A. 7).

3.4. Statistical significance of differences among influencing factors

The significance of varying influence among the six factors was examined using the ecological detector. The results (Table 3) showed that more than two-thirds were not statistically significant, while there were statistically significant differences between IO and other socioeconomic factors (CC, RD, and CO) and between CO and RD. By combining the findings from the factor detector, it could be concluded that IO has a significantly stronger effect on $PM_{2.5}$ than CO, RD, and CC; and CC has a significantly stronger effect on $PM_{2.5}$ than RD.

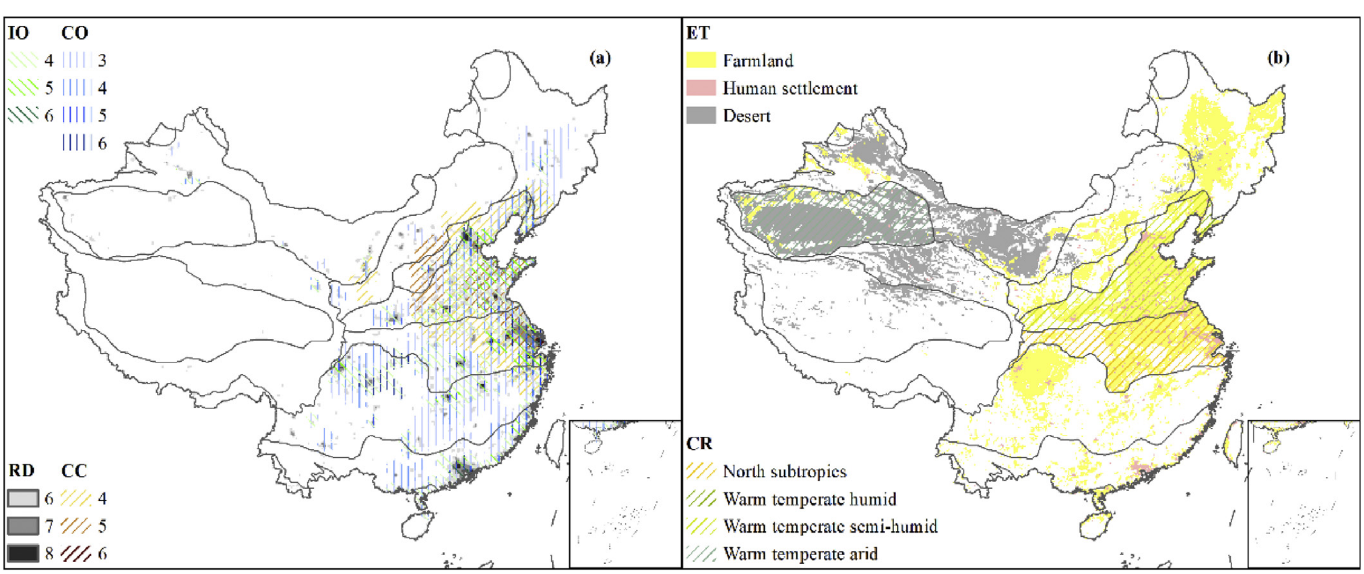

Fig. 6. The spatial distribution of the leading impact areas of factors.

Table 3
Statistically significant differences in the factors' influence on the spatial pattern of $PM_{2.5}$ concentration.

Difference	CR	ET	CC	RD	CO	IO
CR						
ET	N					
CC	N	N				
RD	N	N	N			
CO	N	N	N	Y		
IO	N	N	Y	Y	Y	

Note: Y means the difference of the influence of the two factors is significant with the confidence of 95%, while N means no significant difference.

4. Discussion

Integrated natural and socioeconomic factors and their interactions drove the discrepant geographical pattern of $PM_{2.5}$ concentrations in China. This paper quantified the influences of the socioeconomic and natural factors, as well as their interactive impact on $PM_{2.5}$ concentrations using the GeogDetector model. The results showed that natural factors, including ecosystem and climate, were more influential than socioeconomic factors in driving $PM_{2.5}$ pollution. Among the influencing factors, climate was the dominant factor, and industry was more influential than the other socioeconomic factors. Among all the interactions of the six influencing factors, the interaction between industry and climate presented the largest influence, which was larger than the interaction of ecosystem and climate. The Tarim Basin was a severely contaminated region located mainly in the warm temperate arid climate and desert ecosystem. Another severely contaminated region distributed in the plain areas in the eastern region was primarily in the warm temperate semi-humid climate and influenced by multiple socioeconomic factors.

Although socioeconomic factors such as industry, traffic, construction, and coal consumption were related to the direct sources of $PM_{2.5}$ pollution, the results documented that socioeconomic factors had less influence than natural factors. Meteorological conditions usually have a direct and real-time impact on $PM_{2.5}$ in many different aspects, such as transformation and diffusion caused by wind, purification caused by rainfall, the aggregation of air pollutants, and the formation of secondary particle matter caused by unusual weather conditions. Previous studies have indicated that local meteorology was a relatively strong influencing factor in air pollution (Pearce et al., 2011; Zhao et al., 2013). There are fewer emission sources in a forest system and forests also have certain subduction effects on $PM_{2.5}$, so pollution concentrations in forest regions were relatively low (Nowak et al., 2014, 2006). However, sand-dust from deserts, straw burning and fertilizer use in farmland systems, and anthropogenic emissions from human settlements can enhance $PM_{2.5}$ pollution (Guan et al., 2017; Wang et al., 2016a; Zhang et al., 2015). Thus, the integrated ecosystem also has an important influence on $PM_{2.5}$ pollution (Table 2).

The interaction detector revealed that the interactions between all the factors presented enhanced influence. As the socioeconomic factors were related to anthropogenic emissions, their interactions would increase anthropogenic emissions and reinforce each other in influencing $PM_{2.5}$ pollution. Socioeconomic activities mainly spatially interact with human settlements and farmland systems in an ecological environment, and these interactions also enhance the influence of $PM_{2.5}$ pollution. Gu et al. (2014) have reported similar findings in interpreting the chemical transportation process of urban $PM_{2.5}$ pollution. Nevertheless, the formation of severe haze and the highly polluted areas is usually inseparable from weather and local climatic conditions.

Additionally, the leading impact areas (high-risk areas of $PM_{2.5}$ pollution) were revealed using the risk detector. As the hypothesis that the $PM_{2.5}$ concentrations were generally high in subregions with high values of the socioeconomic variables, and the subclass of ecosystem, mainly human settlements, desert, and farmland and the subclass of climate, including the warm temperate arid climate, warm temperate semi-humid climate, warm temperate humid climate, and north subtropics were also found to have high $PM_{2.5}$ concentrations (Fig. 5).

The risk detector also confirmed that two highly polluted regions, the Tarim Basin (and its surrounding areas) and the eastern plain, were affected by multiple leading impact subregions of either natural or socioeconomic factors, or both. In the Tarim Basin, the high $PM_{2.5}$ concentrations were mainly attributable to the Takla Makan Desert and warm temperate arid climate (Fig. 6). The lack of plant-covered geographical surfaces formed under the influence of dry weather strengthened the effect of the wind and further promoted dust storm outbreaks (Ma et al., 2013; Wang et al., 2009). Thus, the interaction of the desert ecosystem and climate resulted in long-term highly polluted areas of $PM_{2.5}$.

The eastern plain, especially the North China Plain, is the main agricultural area of China and has high concentrations of heavy industries, high road density, and accelerating urban sprawl. This geographical environment contributed a substantial amount to primary emissions and secondary inorganic $PM_{2.5}$ pollutants (Guo et al., 2014). Adverse weather conditions in the planetary boundary layer, such as a continuous strong temperature inversion, downdrafts, and weak surface wind speeds are frequent and continual in the winter, which can further cause pollutants to accumulate in a shallow layer (Zhang et al., 2014; Zhao et al., 2013). Moreover, coal-fired heating and less vegetation coverage (deciduous broad-leaf forests) were related to substantial increases in pollution sources in winter (Wu et al., 2015; Xiao et al., 2015; Yu et al., 2013; Zheng et al., 2005). Hence, the interactions between multiple leading impact subregions of natural and socioeconomic factors led to the formation of severe $PM_{2.5}$ pollution, especially during winter in the North China Plain.

Previous studies had similar findings and determined that the air pollution over eastern China was associated with both anthropogenic emissions and meteorological conditions (Yang et al., 2017b; Zhang and Cao, 2015). Although there were also massive cities, industries, and transportation in other regions, such as in the Greater Changsha Metropolitan Region, the Chengdu-Chongqing Urban Agglomeration, and the Pearl River Delta, the emissions of industrial waste gas, such as SO_2 and soot (He et al., 2014), and coal combustion were less than the North China Plain, and natural conditions (climate and ecology) were not conducive to forming severe haze, and their $PM_{2.5}$ concentrations were much lower than in the North China Plain. No other region had as many anthropogenic emission sources and such favorable natural conditions for developing severe $PM_{2.5}$ pollution as the North China Plain (Fig. 6).

The ecological detector's results indicated that industry has a significantly stronger effect on $PM_{2.5}$ than coal consumption, traffic, and the building industry, and coal consumption has a significantly stronger effect on $PM_{2.5}$ than traffic on the national scale. Socioeconomic factors differ widely across regions, and empirical studies have indicated varying contributions of these factors to $PM_{2.5}$ pollution in different cities or regions (Wang et al., 2013; Wei et al., 1999; Yu et al., 2013). Some previous studies on the national scale have also explored the effected factors of $PM_{2.5}$. Huang et al. (2014) indicated that China's emitted $PM_{2.5}$ was coarser mainly due to strong industry emissions; however, nonindustrial sources had a larger contribution in developed countries. Timmermans et al. (2017) indicated that industry and traffic were the largest contributors during summer months, and residential combustion was the largest during winter months in China.

Spatial stratified heterogeneity is a spatial representation of natural and socioeconomic phenomena, including $PM_{2.5}$ pollution. Theoretically, the GeogDetector in this study is based on a spatial variance analysis of the spatial consistency between $PM_{2.5}$ concentrations and suspect geographical strata (subregions or subclasses). The key assumption is that if the spatial distribution of two variables tends to be consistent, then there is a statistical association between the two variables. Thus, the statistical association can imply potential mechanisms within distinct spatial strata and suggest potential factors involved in the observed process (Wang et al., 2010; Wang and Hu, 2012). This assumption reflects a general understanding of natural phenomena.

In this paper, discrepant influences and interactions among possible factors involved in $PM_{2.5}$ pollution were examined. Although anthropogenic emissions related to socioeconomic factors are the primary cause for increases in $PM_{2.5}$ concentrations, this study verified that natural factors had a stronger influence in driving regional severe and persistent $PM_{2.5}$ pollution, especially in the Tarim Basin and the North China Plain. The existence of spatial stratified heterogeneity within $PM_{2.5}$ pollution determined that the association between $PM_{2.5}$ concentrations and the possible factors was usually weak on a large scale; relationships derived from the regression model were explicable on a limited local scale, but lack explanatory power on a large scale (Lin et al., 2014; Lu et al., 2017).

Some limitations of the study should be clarified. The first limitation is that the results are statistical and do not prove causality, the specific factors for $PM_{2.5}$ pollution are quite complex and diverse in different regions. However, the results can screen out highly suspicious factors for confirming the causality with further analysis. The second limitation is from the raw data of $PM_{2.5}$. The raw data of $PM_{2.5}$ are estimated based on aerosol optical depth (AOD). The accuracy is considerably high, but it also shows uncertainty in different regions (Van Donkelaar et al., 2016). The error in the raw $PM_{2.5}$ data would affect the results inevitably. In addition, different industrial structure may contribute to different amount of polluting emissions and produce different effects on $PM_{2.5}$ pollution. The potential impacts were not fully considered, which would also introduce some uncertainties. Thus, data for emission inventory with high spatial resolution can be used in further studies to improve the accuracy and efficiency of the quantification of anthropogenic influences on $PM_{2.5}$ pollution.

5. Conclusions

The study involved a quantitative analysis of natural and social-economic factors' influence on $PM_{2.5}$ pollution in China. The discrepant influences among natural and social-economic factors in $PM_{2.5}$ pollution were revealed; the spatial differences of $PM_{2.5}$ pollution was related to the spatial disparity of the factors and their interactions. The findings can contribute to better understanding the influencing mechanisms of $PM_{2.5}$ pollution and the geographical factors for forming persistent and highly polluted areas, and imply that more specific control strategies need to be targeted in different regions toward successful $PM_{2.5}$ pollution control and abatement.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 41601419, 41571138, 41761021) and the Collaborative Project of Henan Key Laboratory of Integrative Air Pollution Prevention and Ecological Security. The authors are grateful to the Resource and Environmental Science Data Center (http://www.resdc.cn) of the Chinese Academy of Sciences and the China Meteorological Data Sharing Service System (http://cdc.cma.gov.cn/) for providing data.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.envpol.2018.05.043.

References

Aslam, M.Y., Krishna, K.R., Beig, G., Tinmaker, M.I.R., Chate, D.M., 2017. Diurnal evolution of urban heat island and its impact on air quality by using ground observations (SAFAR) over New Delhi. Open J. Air Pollut. 6, 52–64. https://doi.org/10.4236/ojap.2017.62005.

Bloomer, B.J., Stehr, J.W., Piety, C.A., Salawitch, R.J., Dickerson, R.R., 2009. Observed relationships of ozone air pollution with temperature and emissions. Geophys. Res. Lett. 36, 1–5. https://doi.org/10.1029/2009GL037308.

Butt, E.W., Turnock, S.T., Rigby, R., Reddington, C.L., Yoshioka, M., 2017. Global and regional trends in particulate air quality and attributable health burden over the past 50 years, 19, 3519.

Cao, C., Lee, X., Liu, S., Schultz, N., Xiao, W., Zhang, M., Zhao, L., 2016. Urban heat islands in China enhanced by haze pollution. Nat. Commun. 7, 1–7. https://doi.org/10.1038/ncomms12509.

Chowdhury, Z., Zheng, M., Schauer, J.J., Sheesley, R.J., Salmon, L.G., Cass, G.R., Russell, A.G., 2007. Speciation of ambient fine organic carbon particles and source apportionment of $PM_{2.5}$ in Indian cities. J. Geophys. Res. 112, D15303. https://doi.org/10.1029/2007JD008386.

de Miranda, R.M., de Fatima Andrade, M., Fornaro, A., Astolfo, R., de Andre, P.A., Saldiva, P., 2012. Urban air pollution: a representative survey of $PM_{2.5}$ mass concentrations in six Brazilian cities. Air Qual. Atmos. Heal 5, 63–77. https://doi.org/10.1007/s11869-010-0124-1.

Donkelaar, A., Van Martin, R.V., Brauer, M., Boys, B.L., 2015. Use of satellite observations for long-term exposure assessment of global concentrations of fine particulate matter, 123, 135–143.

Green, M.C., Chow, J.C., Watson, J.G., Dick, K., Inouye, D., 2015. Effects of snow cover and atmospheric stability on winter $PM_{2.5}$ concentrations in western U.S. Valleys. J. Appl. Meteorol. Climatol 54, 1191–1201. https://doi.org/10.1175/JAMC-D-14-0191.1.

Gu, B., Sutton, M.A., Chang, S.X., Ge, Y., Chang, J., 2014. Agricultural ammonia emissions contribute to China's urban air pollution. Front. Ecol. Environ. 12, 265–266. https://doi.org/10.1890/14.WB.007.

Guan, D., Su, X., Zhang, Q., Peters, G.P., Liu, Z., Lei, Y., He, K., 2014. The socioeconomic drivers of China's primary $PM_{2.5}$ emissions. Environ. Res. Lett. 9, 24010. https://doi.org/10.1088/1748-9326/9/2/024010.

Guan, Q., Cai, A., Wang, F., Yang, L., Xu, C., Liu, Z., 2017. Spatio-temporal variability of particulate matter in the key part of Gansu Province, Western China. Environ. Pollut. 230, 189–198. https://doi.org/10.1016/j.envpol.2017.06.045.

Guo, S., Hu, M., Zamora, M.L., Peng, J., Shang, D., Zheng, J., Du, Z., Wu, Z., Shao, M., Zeng, L., Molina, M.J., Zhang, R., 2014. Elucidating severe urban haze formation in China. Proc. Natl. Acad. Sci. Unit. States Am. 111, 17373–17378. https://doi.org/10.1073/pnas.1419604111.

Guo, J., Xia, F., Zhang, Y., Liu, H., Li, J., Lou, M., He, J., Yan, Y., Wang, F., Min, M., Zhai, P., 2017. Impact of diurnal variability and meteorological factors on the PM2.5-AOD relationship: implications for $PM_{2.5}$ remote sensing. Environ. Pollut. 221, 94–104. https://doi.org/10.1016/j.envpol.2016.11.043.

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, C., Huang, Z., Ye, X., 2014. Spatial heterogeneity of economic development and industrial pollution in urban China. Stoch. Environ. Res. Risk Assess. 28, 767–781. https://doi.org/10.1007/s00477-013-0736-8.

Huang, Y., Shen, H., Chen, H., Wang, R., Zhang, Y., Su, S., Chen, Y., Lin, N., Zhuo, S., Zhong, Q., Wang, X., Liu, J., Li, B., Liu, W., Tao, S., 2014. Quantification of global primary emissions of $PM_{2.5}$, PM_{10},and TSP from combustion and industrial process sources. Environ. Sci. Technol. 48, 13834–13843. https://doi.org/10.1021/es503696k.

Jin, Q., Fang, X., Wen, B., Shan, A., 2017. Spatio-temporal variations of $PM_2.5$ emission in China from 2005 to 2014. Chemosphere 183, 429–436. https://doi.org/10.1016/j.chemosphere.2017.05.133.

Li, G., Fang, C., Wang, S., Sun, S., 2016. The effect of economic growth, urbanization, and industrialization on fine particulate matter ($PM_{2.5}$) concentrations in China. Environ. Sci. Technol. 50, 11452–11459. https://doi.org/10.1021/acs.est.6b02562.

Lin, G., Fu, J., Jiang, D., Hu, W., Dong, D., Huang, Y., 2014. Spatio-Temporal variation of $PM_{2.5}$ concentrations and their relationship with geographic and socioeconomic factors in China, 11, 173–186. https://doi.org/10.3390/ijerph110100173.

Liu, Y., Cao, G., Zhao, N., Mulligan, K., Ye, X., 2018. Improve ground-level PM2.5 concentration mapping using a random forests-based geostatistical approach. Environ. Pollut. 235, 272–282. https://doi.org/10.1016/j.envpol.2017.12.070.

Lu, F., Xu, D., Cheng, Y., Dong, S., Guo, C., Jiang, X., Zheng, X., 2015. Systematic review and meta-analysis of the adverse health effects of ambient $PM_{2.5}$ and PM_{10} pollution in the Chinese population. Environ. Res. 136, 196–204. https://doi.org/10.1016/j.envres.2014.06.029.

Lu, D., Xu, J., Yang, D., Zhao, J., 2017. Spatio-temporal variation and in fluence factors of $PM_{2.5}$ concentrations in China from 1998 to 2014. Atmos. Pollut. Res. 8, 1151–1159. https://doi.org/10.1016/j.apr.2017.05.005.
Ma, Z., Xue, B., Geng, Y., Ren, W., Fujita, T., Zhang, Z., de Oliverira, J.P., Jacques, D.A., Xi, F., 2013. Co-benefits analysis on climate change and environmental effects of wind-power: a case study from Xinjiang, China. Renew. Energy 57, 35–42. https://doi.org/10.1016/j.renene.2013.01.018.
Ma, Z., Hu, X., Sayer, A.M., Levy, R., Zhang, Q., Xue, Y., Tong, S., Bi, J., Huang, L., Liu, Y., 2016. Satellite-based spatiotemporal trends in $PM_{2.5}$ concentrations: China, 2004-2013. Environ. Health Perspect. 124, 184–192. https://doi.org/10.1289/ehp.1409481.
Marcazzan, G.M., Ceriani, M., Valli, G., Vecchi, R., 2003. Source apportionment of PM_{10} and $PM_{2.5}$ in Milan (Italy) using receptor modelling. Sci. Total Environ. 317, 137–147. https://doi.org/10.1016/S0048-9697(03)00368-1.
Mcguinn, L.A., Ward-caviness, C., Neas, L.M., Schneider, A., Di, Q., Chudnovsky, A., Schwartz, J., Koutrakis, P., Russell, A.G., Garcia, V., Kraus, W.E., Hauser, E.R., Cascio, W., Diaz-sanchez, D., Devlin, R.B., 2017. Fine particulate matter and cardiovascular disease : comparison of assessment methods for long-term exposure. Environ. Res. 159, 16–23. https://doi.org/10.1016/j.envres.2017.07.041.
Meng, J., Liu, J., Xu, Y., Tao, S., 2015. Tracing Primary $PM_{2.5}$ emissions via Chinese supply chains. Environ. Res. Lett. 10 https://doi.org/10.1088/1748-9326/10/5/054005.
Nowak, D.J., Crane, D.E., Stevens, J.C., 2006. Air pollution removal by urban trees and shrubs in the United States. Urban for. Urban Green 4, 115–123. https://doi.org/10.1016/j.ufug.2006.01.007.
Nowak, D.J., Hirabayashi, S., Bodine, A., Greenfield, E., 2014. Tree and forest effects on air quality and human health in the United States. Environ. Pollut. 193, 119–129. https://doi.org/10.1016/j.envpol.2014.05.028.
Pant, P., Shukla, A., Kohl, S.D., Chow, J.C., Watson, J.G., Harrison, R.M., 2015. Characterization of ambient PM2.5 at a pollution hotspot in New Delhi, India and inference of sources. Atmos. Environ. 109, 178–189. https://doi.org/10.1016/j.atmosenv.2015.02.074.
Pearce, J.L., Beringer, J., Nicholls, N., Hyndman, R.J., Tapper, N.J., 2011. Quantifying the influence of local meteorology on air quality using generalized additive models. Atmos. Environ. 45, 1328–1336. https://doi.org/10.1016/j.atmosenv.2010.11.051.
Tao, M., Chen, L., Xiong, X., Zhang, M., Ma, P., Tao, J., Wang, Z., 2014. Formation process of the widespread extreme haze pollution over northern China in January 2013: implications for regional air quality and climate. Atmos. Environ. 98, 417–425. https://doi.org/10.1016/j.atmosenv.2014.09.026.
Timmermans, R., Kranenburg, R., Manders, A., Hendriks, C., Segers, A., Dammers, E., Denier van der Gon, H., Schaap, M., Dammers, E., Zeng, L., Wang, L., Liu, Z., 2017. Source apportionment of PM2.5 across China using LOTOS-EUROS. Atmos. Environ. 164, 370–386. https://doi.org/10.1016/j.atmosenv.2017.06.003.
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, models, and monitors. Environ. Sci. Technol. 50, 3762–3772. https://doi.org/10.1021/acs.est.5b05833.
Wang, J., Hu, Y., 2012. Environmental health risk detection with GeogDetector. Environ. Model. Software 33, 114–115. https://doi.org/10.1016/j.envsoft.2012.01.015.
Wang, S., Feng, X., Zeng, X., Ma, Y., Shang, K., 2009. A study on variations of concentrations of particulate matter with different sizes in Lanzhou, China. Atmos. Environ. 43, 2823–2828. https://doi.org/10.1016/j.atmosenv.2009.02.021.
Wang, J., Li, X., Christakos, G., Liao, Y., Zhang, T., Gu, X., Zheng, X., 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., Hu, Z., Chen, Y., Chen, Z., Xu, S., 2013. Contamination characteristics and possible sources of PM_{10} and $PM_{2.5}$ in different functional areas of Shanghai, China. Atmos. Environ. 68, 221–229. https://doi.org/10.1016/j.atmosenv.2012.10.070.
Wang, Y.S., Yao, L., Wang, L.L., Liu, Z.R., Ji, D.S., Tang, G.Q., Zhang, J.K., Sun, Y., Hu, B., Xin, J.Y., 2014. Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China. Sci. China Earth Sci. 57, 14–25. https://doi.org/10.1007/s11430-013-4773-4.
Wang, G., Zhang, R., Gomez, M.E., Yang, L., Levy Zamora, M., Hu, M., Lin, Y., Peng, J., Guo, S., Meng, J., Li, J., Cheng, C., Hu, T., Ren, Y., Wang, Y., Gao, J., Cao, J., An, Z., Zhou, W., Li, G., Wang, J., Tian, P., Marrero-Ortiz, W., Secrest, J., Du, Z., Zheng, J., Shang, D., Zeng, L., Shao, M., Wang, W., Huang, Y., Wang, Y., Zhu, Y., Li, Y., Hu, J., Pan, B., Cai, L., Cheng, Y., Ji, Y., Zhang, F., Rosenfeld, D., Liss, P.S., Duce, R.A., Kolb, C.E., Molina, M.J., 2016a. Persistent sulfate formation from London Fog to Chinese haze. Proc. Natl. Acad. Sci. Unit. States Am. 113, 13630–13635. https://doi.org/10.1073/pnas.1616540113.
Wang, J.F., Zhang, T.L., Fu, B.J., 2016b. A measure of spatial stratified heterogeneity. Ecol. Indicat. 67, 250–256. https://doi.org/10.1016/j.ecolind.2016.02.052.
Wei, F., Teng, E., Wu, G., Hu, W., Wilson, W.E., Chapman, R.S., Pau, J.C., Zhang, J., 1999. Ambient concentrations and elemental compositions of PM_{10} and $PM_{2.5}$ in four Chinese cities. Environ. Sci. Technol. 33, 4188–4193. https://doi.org/10.1021/es9904944.
Wu, J., Li, J., Peng, J., Li, W., Xu, G., Dong, C., 2015. Applying land use regression model to estimate spatial variation of $PM_{2.5}$ in Beijing, China. Environ. Sci. Pollut. Res. 22, 7045–7061. https://doi.org/10.1007/s11356-014-3893-5.
Wu, C., Ye, X., Du, Q., Luo, P., 2017. Spatial effects of accessibility to parks on housing prices in Shenzhen, China. Habitat Int. 63, 45–54. https://doi.org/10.1016/j.habitatint.2017.03.010.
Xiao, Q., Ma, Z., Li, S., Liu, Y., 2015. The impact of winter heating on air pollution in China. PLoS One 10, 1–11. https://doi.org/10.1371/journal.pone.0117311.
Xie, R., Sabel, C.E., Lu, X., Zhu, W., Kan, H., Nielsen, C.P., Wang, H., 2016. Long-term trend and spatial pattern of $PM_{2.5}$ induced premature mortality in China, 97, 180–186. https://doi.org/10.1016/j.envint.2016.09.003.
Yang, D., Lu, D., Xu, J., Ye, C., Zhao, J., Tian, G., Wang, X., Zhu, N., 2017a. Predicting spatio-temporal concentrations of $PM_{2.5}$ using land use and meteorological data in Yangtze River Delta, China. Stoch. Environ. Res. Risk Assess. https://doi.org/10.1007/s00477-017-1497-6.
Yang, Y., Russell, L.M., Lou, S., Liao, H., Guo, J., Liu, Y., Singh, B., Ghan, S.J., 2017b. Dust-wind interactions can intensify aerosol pollution over eastern China. Nat. Commun. 8, 15333. https://doi.org/10.1038/ncomms15333.
Yu, L., Wang, G., Zhang, R., Zhang, L., Song, Y., Wu, B., Li, X., An, K., Chu, J., 2013. Characterization and source apportionment of $PM_{2.5}$ in an urban environment in Beijing. Aerosol Air Qual. Res. 13, 574–583. https://doi.org/10.4209/aaqr.2012.07.0192.
Zhang, Y.-L., Cao, F., 2015. Fine particulate matter ($PM_{2.5}$) in China at a city level. Sci. Rep. 5, 14884. https://doi.org/10.1038/srep14884.
Zhang, D.L., Shou, Y.X., Dickerson, R.R., 2009. Upstream urbanization exacerbates urban heat island effects. Geophys. Res. Lett. 36, 1–5. https://doi.org/10.1029/2009GL041082.
Zhang, R.H., Li, Q., Zhang, R.N., 2014. Meteorological conditions for the persistent severe fog and haze event over eastern China in January 2013. Sci. China Earth Sci. 57, 26–35. https://doi.org/10.1007/s11430-013-4774-3.
Zhang, R., Wang, G., Guo, S., Zamora, M.L., Ying, Q., Lin, Y., Wang, W., Hu, M., Wang, Y., 2015. Formation of urban fine particulate matter. Chem. Rev. 115, 3803–3855. https://doi.org/10.1021/acs.chemrev.5b00067.
Zhao, X.J., Zhao, P.S., Xu, J., Meng, W., Pu, W.W., Dong, F., He, D., Shi, Q.F., 2013. Analysis of a winter regional haze event and its formation mechanism in the North China Plain. Atmos. Chem. Phys. 13, 5685–5696. https://doi.org/10.5194/acp-13-5685-2013.
Zheng, M., Salmon, L.G., Schauer, J.J., Zeng, L., Kiang, C.S., Zhang, Y., Cass, G.R., 2005. Seasonal trends in $PM_{2.5}$ source contributions in Beijing, China. Atmos. Environ. 39, 3967–3976. https://doi.org/10.1016/j.atmosenv.2005.03.036.
Zheng, J., Yin, Y., Li, B., 2010. A new scheme for climate regionalization in China. ACTA Geogr. Sin 65, 3–12.