

SCIENTIFIC REPORTS



OPEN

Spatiotemporal Pattern of PM_{2.5} Concentrations in Mainland China and Analysis of Its Influencing Factors using Geographically Weighted Regression

Jieqiong Luo^{1,2,3}, Peijun Du^{1,2,3}, Alim Samat⁴, Junshi Xia⁵, Meiqin Che^{1,2,3} & Zhaohui Xue⁶

Based on annual average PM_{2.5} gridded dataset, this study first analyzed the spatiotemporal pattern of PM_{2.5} across Mainland China during 1998–2012. Then facilitated with meteorological site data, land cover data, population and Gross Domestic Product (GDP) data, etc., the contributions of latent geographic factors, including socioeconomic factors (e.g., road, agriculture, population, industry) and natural geographical factors (e.g., topography, climate, vegetation) to PM_{2.5} were explored through Geographically Weighted Regression (GWR) model. The results revealed that PM_{2.5} concentrations increased while the spatial pattern remained stable, and the proportion of areas with PM_{2.5} concentrations greater than 35 µg/m³ significantly increased from 23.08% to 29.89%. Moreover, road, agriculture, population and vegetation showed the most significant impacts on PM_{2.5}. Additionally, the Moran's I for the residuals of GWR was 0.025 (not significant at a 0.01 level), indicating that the GWR model was properly specified. The local coefficient estimates of GDP in some cities were negative, suggesting the existence of the inverted-U shaped Environmental Kuznets Curve (EKC) for PM_{2.5} in Mainland China. The effects of each latent factor on PM_{2.5} in various regions were different. Therefore, regional measures and strategies for controlling PM_{2.5} should be formulated in terms of the local impacts of specific factors.

Clean air is regarded as a basic requirement of human comfort, health and well-being. As we all know, air pollution has many serious adverse effects, such as negative impacts on the climate, ecosystem services, biodiversity, and food security¹ and continuously posing severe threats to human health worldwide. According to a study of World Health Organization (WHO), more than two million people prematurely died each year attribute to the effects of air pollution since the 21st century. It is worth noting that more than half of these deaths happened in developing countries, particularly in China and India². Particulates, especially PM_{2.5} (aerosol particles with aerodynamic diameter less than 2.5 µm), are the deadliest form of air pollutants. Epidemiologic studies have shown that exposures to PM_{2.5} are associated with increased cardiovascular and respiratory morbidity and mortality³. With the rapid development of urbanization and industrialization, PM_{2.5} pollutions in China, especially in the North part, is becoming serious, which can be clearly observed by satellite measurements^{4–6}. Thus, the analysis of the spatiotemporal pattern of PM_{2.5} across China is very imperative for carrying out environmental epidemiologic studies.

¹Key Laboratory for Satellite Mapping Technology and Applications of State Administration of Surveying, Mapping and Geoinformation of China, Nanjing University, Nanjing 210023, China. ²Jiangsu Provincial Key Laboratory of Geographic Information Science and Technology, Nanjing University, Nanjing 210023, China. ³Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China. ⁴State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi 830011, China. ⁵Research Center for Advanced Science and Technology, The University of Tokyo 4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan. ⁶School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China. Correspondence and requests for materials should be addressed to P.D. (email: dupjrs@126.com)

Received: 03 May 2016
Accepted: 08 December 2016
Published: 12 January 2017

For the past couple of years, most of the studies have investigated the spatiotemporal pattern of PM_{2.5} in China on a regional scale^{7–10}. For instance, Liu *et al.*⁷ found that there were no significant differences of PM_{2.5} concentrations between urbanized and suburban areas of Shanghai. Meanwhile, from 2006 to 2008, the seasonal variation of PM_{2.5} in Hangzhou was noticeably characterized by higher concentration in winter and lower in summer⁸, which is consistent with the results of Xi'an⁹ and Beijing cities¹⁰. In recent years, the exploration of spatial and temporal variability of PM_{2.5} pollutions in these areas has expanded to more broad regional^{11,12} and national scales^{13,14}. As aforementioned, most studies of the spatiotemporal pattern of PM_{2.5} at a regional scale are conducted with a short time scale and high temporal resolution, whereas those on a large regional or national scales usually use inter-annual and long time series data sets with a lower spatial resolution (e.g., 50 km). There has been a lack of research on the spatiotemporal pattern of PM_{2.5} concentrations in China with high spatial resolution using long time series data. Therefore, one of the primary objectives of our study is to analyze the spatiotemporal pattern of PM_{2.5} concentrations in China using more than a decade of data.

From the microcosmic perspective, a considerable number of existing studies have been dedicated to investigating the main sources of PM_{2.5} in China. They mainly focused on one large city at a specific time spot. For example, Liu *et al.*¹⁵ investigated the source apportionment of PM_{2.5} using observations in Beijing, a highly polluted city in Northern China. Wang *et al.*¹⁶ found that industrial process and vehicle emission are the dominant local contributors to total PM_{2.5} mass in the whole city. Also, analysis of the source apportionment of PM_{2.5} in the Yangtze River Delta of China indicated that PM_{2.5} is primarily from secondary pollutants and primary emissions from vehicles and biomass burning¹⁷.

From the macroscopic perspective, many previous studies have evaluated the impacts of natural geographical factors on PM_{2.5} concentrations at a regional scale, including vegetation^{18–22} and meteorological parameters^{23–25}. However, limited attention was paid to explore the relationships between PM_{2.5} and socioeconomic factors. For instance, Lin *et al.*¹³ found that there was a significant positive correlation between population growth, economic development, urban expansion and PM_{2.5} concentrations. Other studies also identified that urbanization indicators (i.e., urban built-up area, population and industry fraction) have great impacts on urban PM_{2.5} concentrations^{26,27}. In addition, Han *et al.*²⁸ found the evidence of the increasing effects of human activities on PM_{2.5} pollutions through analyzing the relationship between artificial surface, cropland and PM_{2.5} concentrations.

Unfortunately, to the best of our knowledge, there is no research to systematically and comprehensively analyze the impacts of natural geographical and socioeconomic factors on PM_{2.5} at a national scale. Given the fact that China is facing severe air pollution and PM_{2.5} is its main component, to investigate the relationships between PM_{2.5} concentrations and its influencing factors is very important and valuable for drafting appropriate air pollution control policies. Thus, considering the fact that PM_{2.5} concentrations vary over space and there probably exists spatial autocorrelation within PM_{2.5} concentrations of surrounding regions, another primary objective of our study was to explore the contributions of the influencing factors, including natural geographical and socioeconomic factors on PM_{2.5} in 343 cities of Mainland China using geographically weighted regression (GWR) model. It is necessary for China in achieving the goals of sustainable development and the National New-type Urbanization Plan^{28,29}.

In brief, this study has three main contributions. First, using more than a decade of data, we analyzed the spatiotemporal patterns of PM_{2.5} concentrations in China at a high spatial resolution. Second, to our knowledge, this research was the first attempt to investigate the relationships between PM_{2.5} concentrations and the natural geographical and socioeconomic factors across Mainland China on a city level from the macroscopic perspective. Last but not least, GWR, a local form of linear regression was applied in this study to fully consider the spatial heterogeneity of PM_{2.5}. And it is of paramount importance for formulating and refining local pollution control strategies.

Results and Discussion

Spatiotemporal patterns of PM_{2.5} concentrations in Mainland China. As Supplementary Fig. S1 shown, in general, PM_{2.5} concentrations increased over most of Mainland China and its spatial pattern remained stable during the study period, and there were two main clusters of regions with low PM_{2.5} concentrations. One is located at the northern Inner Mongolian Plateau, and northeastern plains. The other is lied in the southwestern Tibetan Plateau. On the contrary, there are three main clusters of regions with high PM_{2.5} concentrations, including the North China plains, Yangtze plains and Central China, followed by Tarim Basin, and the Sichuan Basin. Many studies have partly attributed the higher PM_{2.5} concentrations in these regions to the coal-based industries such as coal-fired power plants, iron and steel manufacturing^{17,30}. Comparatively, partly thanks to the developed tertiary industry which produces little pollution, the PM_{2.5} concentrations in the Pearl River Delta are the lowest in the three main economic zones²⁷. The spatial distribution areas of PM_{2.5} concentrations greater than 100 µg/m³ during the two periods of 2005–2007 and 2006–2008 were larger than in other periods. During the period of 1998–2012, areas that exceeded Interim Target-1 35 µg/m³ expanded from the central eastern region to south western region, and eventually connected in the Sichuan Basin. In the southwestern Tibetan Plateau, northern Inner Mongolian Plateau, and northeastern plains, PM_{2.5} concentrations are lower than the WHO air quality guideline (AQG) of 10 µg/m³. Besides, PM_{2.5} concentrations of the southwestern Tibetan Plateau changed very little, while those of the northern Inner Mongolian Plateau and northeastern plains significantly increased. We can also observe that the tendency of high PM_{2.5} concentrations was still remaining high with a little decreasing trend.

Furthermore, in order to quantitate the annual variations of all grades of PM_{2.5} concentrations in the whole Mainland China and seven geographical subareas from 1998 to 2012, the cumulative proportion of area of each concentration range was calculated as shown in Fig. 1.

The proportion of area experiencing low PM_{2.5} concentrations of Grade 1 significantly decreased from 32.92% to 25.67% and this decrease mostly occurred in Northeast, North and South China, slightly in Southwest and Northwest China. However, the proportion of area under high PM_{2.5} concentrations increased annually

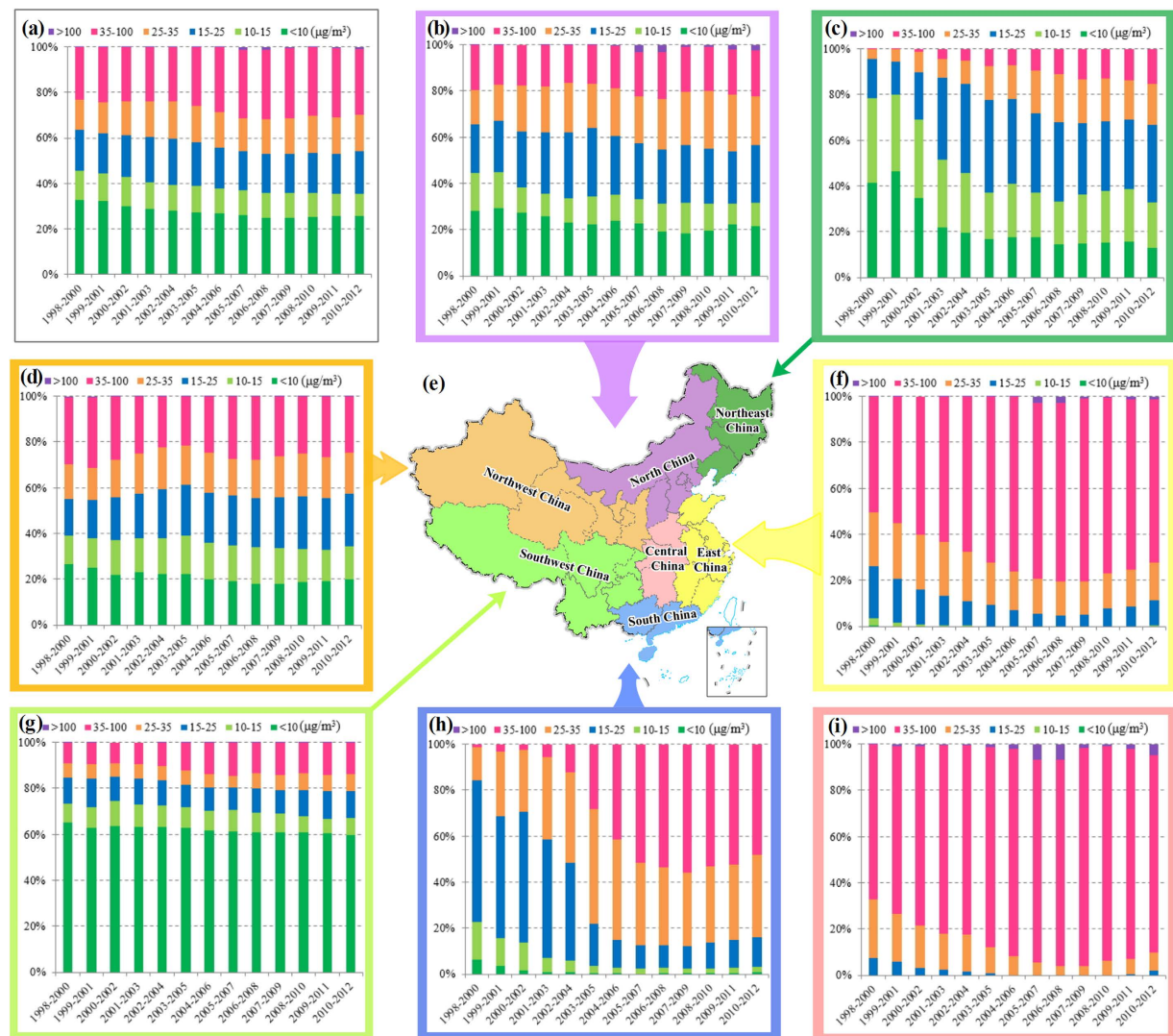


Figure 1. Annual variations of different grades of annual average $PM_{2.5}$ concentrations in whole Mainland China and seven geographical subareas from 1998 to 2012. (a) Mainland China, (b) North China, (c) Northeast China, (d) Northwest China, (e) Subarea Map, (f) East China, (g) Southwest China, (h) South China, (i) Central China. The subarea map was generated in ArcGIS10.2, URL: <http://www.esrichina-bj.cn/softwareproduct/ArcGIS/>.

during the study period. For example, the proportion of area in Grade 6 increased from 0.27% to 0.91% and this increase primarily happened in East, North and Central China, while the proportion of area in Grade 5 significantly increased from 22.81% to 28.98%. Similarly, the proportion of area under mid-range concentrations also increased, with Grades 3 and 4 respectively increasing from 17.93% to 18.76% and from 13.32% to 15.88%, respectively, while Grade 2 decreased from 12.75% to 9.79%, mostly locating in South, Northeast and North China. The increase of Grade 3 was mainly located in Northeast, North and Northwest China, whereas declined in South, East and Central China. Meanwhile, the increase of Grade 4 came from South, North, Northeast and Northwest China, while declined in Central and East China.

The proportion of area whose $PM_{2.5}$ concentrations is greater than the secondary standard of $35 \mu g/m^3$ increased from 23.08% to 29.89%. Likewise, the areal extent of $PM_{2.5}$ concentrations between the primary and secondary standard increased from 31.25% to 34.64%, which is slightly more than one-third of the study area. Among the concentration grades, Grade 5 ($35-100 \mu g/m^3$) shows the biggest area of increasing distribution, and this increase occurred primarily in East, South and Central China, slightly in Northeast China. The inter-annual undulation in area proportion was the largest for comprising Grades 3 and Grade 4, with both of these areas first rapidly decreasing and then slightly increasing. The largest area proportion of Grade 6 with $PM_{2.5}$ concentrations greater than $100 \mu g/m^3$ happened at the periods 2005–2007 and 2006–2008. The largest area proportion under Grade 3 with $PM_{2.5}$ concentrations of $15-25 \mu g/m^3$ occurred during 2002–2004, with an area ratio of 20.35%, then declined to 17.00% during 2006–2008, and finally increased to 18.76% during 2010–2012.

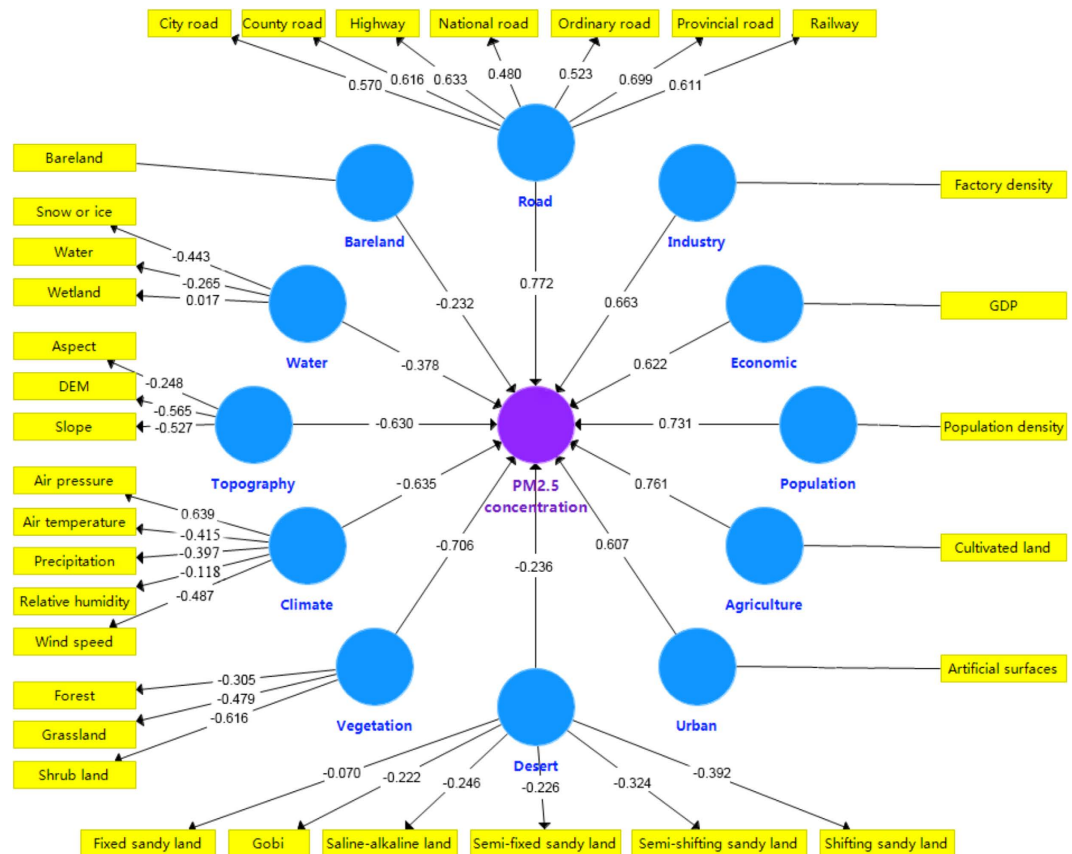


Figure 2. The correlation coefficients between latent geographic factors and $PM_{2.5}$ concentrations.

The latent geographic factors under the blue circle, including “Road”, “Industry”, “Economic”, “Population”, “Agriculture”, “Urban”, “Desert”, “Vegetation”, “Climate”, “Topography”, “Water”, “Bareland”, were extracted from the variables in the yellow rectangle. For instance, “Climate” was based on “Air pressure”, “Air temperature”, “Precipitation”, “Relative humidity”, “wind speed”. For factors that contain multiple variables, the coefficients between blue circle and yellow rectangle directly are the contribution of the variables to $PM_{2.5}$. Similarly, the coefficients between blue and purple circle are the contribution of the factors to $PM_{2.5}$. For univariate factors, the coefficients between blue and purple circle are also the contribution of the variables to $PM_{2.5}$.

From Fig. S1 and Fig. 1, it should be stressed that three-fourths of Mainland China had $PM_{2.5}$ concentrations that exceeded the WHO AQG of $10 \mu\text{g}/\text{m}^3$. The concentrations were particularly high in Central, North, Southwest and East China, and the most severe areas were the northern Henan, southern Hebei and Shanxi, central Shaanxi, eastern Sichuan and western Shandong. It is worth noticed that there was a rapid increase of highest $PM_{2.5}$ concentrations ($>100 \mu\text{g}/\text{m}^3$) over southern Hebei, northern Henan and western Shandong during 1998–2008, with peaks during 2007–2009, followed by a significant reduction during 2008–2010. It is believed that this reduction was associated with the national energy conservation policies that were carried out during the “Eleventh Five-Year” period¹⁰.

Relationship between $PM_{2.5}$ concentrations and latent geographic factors. With the help of ordinary least squares (OLS), multivariate correlation analysis was carried out between latent geographic factors and $PM_{2.5}$ concentrations to identify the decisive factors of $PM_{2.5}$ concentrations before GWR. Figure 2 showed all the correlation coefficients between latent geographic factors and $PM_{2.5}$ concentrations. Overall, most of the correlations are significant in 343 cities except for the factors of bareland and desert.

From Fig. 2, there is a significantly positive correlation between road, agriculture, population, industry, economic, urban and $PM_{2.5}$ concentrations. It implies that $PM_{2.5}$ originates primarily from motor vehicle emissions (diesel and gasoline exhaust), dust (road dust, surface dust, building construction dust), industrial and combustion sources including coal and biomass combustion (e.g. straw, bark residuals, sawdust and shavings)^{15,16,28,31}. In particular, traffic (47.9%) and combustion (29.7%) aerosol were two conclusive factors of $PM_{2.5}$ concentrations²². This is consistent with the results of our study that road and agriculture have the strongest relationships with $PM_{2.5}$ concentrations.

A significantly negative correlation was found between vegetation and $PM_{2.5}$ concentrations. Previous studies have shown that vegetation can mitigate particulate air pollution through a number of mechanisms, such as intercepting and accumulating atmospheric particles through leaf pubescence and stomata^{18–20}. The maximum effect on $PM_{2.5}$ concentrations is the shrub land, followed by the grass land and forest. This is due to the fact that different species have different properties, such as leaf size, stomata, vegetation structure and leaf microstructure

which will affect the capture efficiency²¹, whilst forest may influence the microclimate, e.g. blocking effective ventilation and thus lead to higher local PM_{2.5} concentrations²².

The results also indicated that there is a significantly negative correlation among all variables of topography and PM_{2.5} concentrations. The decreasing order of ranking is that DEM (altitude), slope and aspect. This may be the reason why PM_{2.5} concentrations of the Sichuan Basin are much higher than that of its circumjacent regions.

Meanwhile, the correlations between all variables of climate and PM_{2.5} concentrations were negative except air pressure. That is to say, an increase (decrease) of each meteorological factor except air pressure will result in the decrease (increase) of PM_{2.5} concentrations. Specifically, there was a moderate negative relation between air temperature and PM_{2.5} concentrations, with a correlation coefficient of -0.415 . This is because when temperature rises, air convection becomes quick and frequent, which leads to the diffusion and dilution of PM_{2.5}, decreasing PM_{2.5} concentrations, and vice versa. There was a moderate negative correlation between wind speed and PM_{2.5} concentrations, with a correlation coefficient of -0.487 and higher wind speed is conducive to the diffusion of PM_{2.5}, which results in lower concentrations of PM_{2.5}. A weak negative correlation was found between relative humidity and PM_{2.5} concentrations because when air humidity increases, particles will adsorb moisture, occur condensation, and finally fall in the form of precipitation that results in lower concentrations of PM_{2.5}²³. The correlation coefficients between precipitation and PM_{2.5} concentrations was -0.397 and there are two main impacts of precipitation on PM_{2.5}. One is that the adsorption and collision of raindrops upon PM_{2.5} result in the wet sedimentation of PM_{2.5}. The other one is that after rainy weather PM_{2.5} concentrations significantly decrease due to the notable reducing of dust and fugitive dust which previously suspended in the atmosphere. PM_{2.5} concentrations display a strong positive correlation with air pressure, with significant correlation coefficients of 0.639 . This is due to the fact that when there was high pressure, the down draft hinders the upward movement of PM_{2.5}, causing an accumulation of particles²⁴.

However, there are weak negative correlations among all variables of desert and bare land. The Variance Inflation Factor (VIF) was used to detect whether collinearity problems existed among the variables by OLS. It turned out that the VIF values of most variables are less than 7.5 except water. In conclusion, socioeconomic factors of road, agriculture, population, industry, economic, urban and natural geographic factors of topography, climate and vegetation were chosen as independent variables of the GWR model.

A higher R² value means that the explanatory variables explain more variance in PM_{2.5} concentrations. From the report of the GWR model, the R² value of GWR was 0.9633, whereas it was 0.6788 of OLS, indicating that the global OLS model only can explain 67.88% of the variance in PM_{2.5} concentrations, but the GWR method had a significant improvement. In terms of the analysis of variance (ANOVA), the model fit at a significant level ($F = 13.97$, $p < 0.01$), indicating that the GWR model outperformed the OLS model. Moreover, the corrected Akaike information criterion (AICc)³² value of GWR (720.18) was much lower than that of the OLS (1182.77). In other words, the performance of GWR model was much better than that of the global OLS model. Through golden-section search^{33,34}, a bandwidth size of 60 was selected as an appropriate value for the GWR model, and it means that 60 samples were provided for each local estimation within the adaptive bi-square kernel.

Figure 3 shows the spatial heterogeneity of standardized residuals for PM_{2.5} concentrations derived from the GWR model and spatial autocorrelation analysis. As it shown, the regions with studentized residual (StdResid) value between -2 and 2 account for 93.58% of the whole Mainland China, which indicates that the relations between each of the nine factors and PM_{2.5} are stable. However, few unusually high (red areas) or low (blue areas) residuals can be observed. Red areas are under predictions in which the actual PM_{2.5} concentrations are higher than the model fitted value. Blue areas are over predictions in which the actual PM_{2.5} concentrations are lower than fitted value. Regions that have a notable under-prediction of PM_{2.5} concentrations need further examination to detect the possible explanations. For instance, those regions of the Tarim Basin in some desert areas, e.g. Hotan prefecture and Bayingolin Mongol Autonomous prefecture have highest residuals (StdResid > 2). This is because that the higher concentrations of PM_{2.5} in the desert regions are mainly connected with sand and dust weather phenomena¹¹. The Yuncheng city of southern Shanxi province and Tianjin city have much higher residuals because they are rich in marine salt and coal mine. The eastern Sichuan basin, e.g. Deyang, Zigong and Yibin city has high residuals because of its unique geographical climate conditions³⁵. Shenzhen city of southern Guangdong province also has much higher residuals because of the influence of coal used as fuel in this area for industrial plants^{27,36}. Spatial autocorrelation among residuals of GWR was detected by Moran's I test. Moran's I for residuals is 0.025 and is not significant at a 0.01 level (Z -score = 1.550, $p = 0.1212$), indicating that the pattern of the residuals does not appear to be significantly different than random. It also suggests that the GWR model does not miss any key explanatory variables and it is properly specified.

The GWR model also indicates that the effects of independent variables on PM_{2.5} concentrations vary across space. Figure 4 shows the classification maps for the local estimates and p values of nine factors' regression coefficients. These results clearly illustrate the existence of an unstable local spatial dependence between PM_{2.5} concentrations and its nine latent geographical factors. In general, it is evident that all coefficients resulting from global models are significant. Such inference is also strong for local coefficients in terms of sign and magnitude.

Specifically, although local parameter estimates for "Road" are positive in most areas of Mainland China, the intensity of the relationships is not constant (Fig. 4(a1)). The positive and strong relationships are found in the northern regions of East China, southern districts of Northwest, North and East China, and a distinct region located in South China (Fig. 4(a2)). In such areas, the effect of "Road" upon PM_{2.5} is relatively higher than in other areas, which indicates very efficient strategies to abate the PM_{2.5} concentrations. The map is produced for "Agriculture" (Fig. 4(b1)) illustrates that there are strong positive trends in the central parts towards the northeast and southwest, while weak negative relationships in the North and South China. Meanwhile, the regression coefficients of "Agriculture" were primarily significant, except in South China, northern areas of the Northeast and East China, and central districts of Southwest China (Fig. 4(b2)). In Southwest China, North China, Northeast

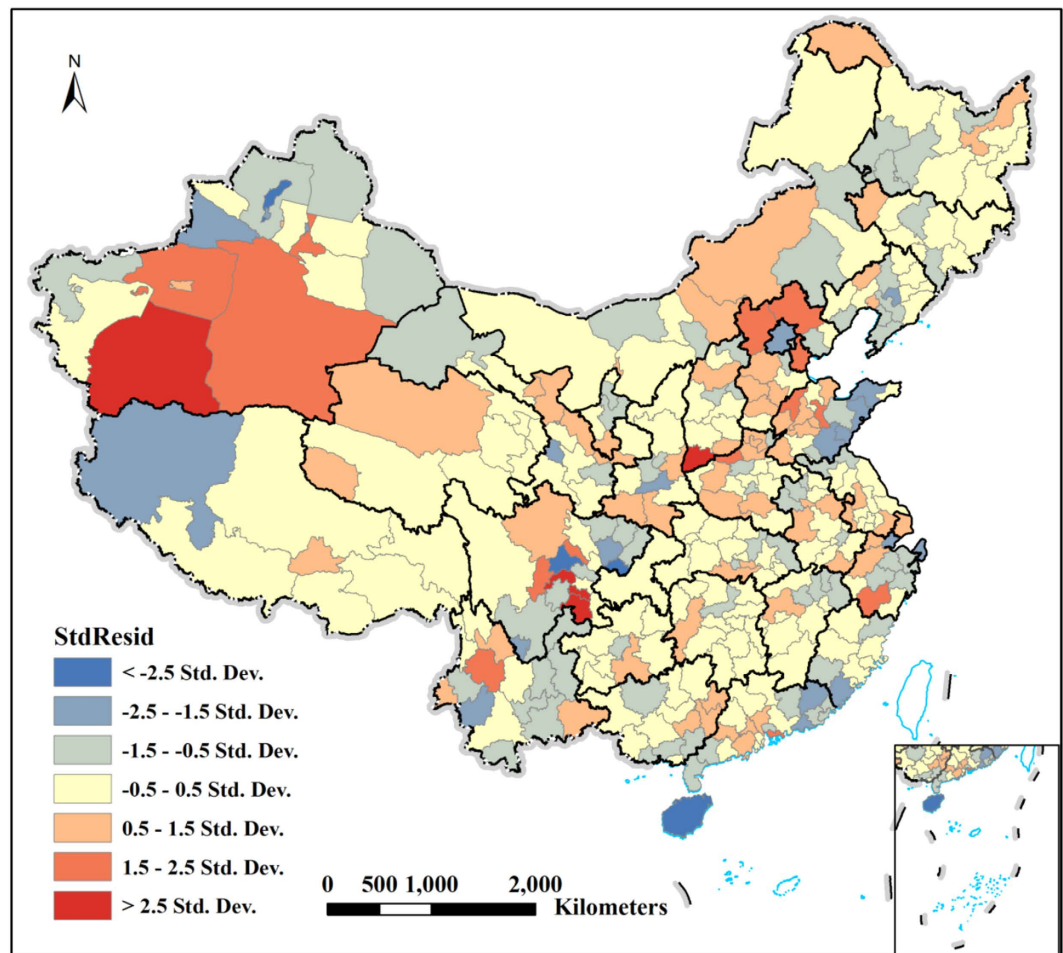


Figure 3. Spatial heterogeneity of standardized residuals for $PM_{2.5}$ concentrations derived from the GWR model. All the maps were generated in ArcGIS10.2, URL: <http://www.esrichina-bj.cn/softwareproduct/ArcGIS/>.

China and southern regions of Northwest China, the regression coefficients of “Population” were notable positive (Fig. 4(c1,2)).

The direct effect of “Industry” has been illustrated with greater magnitude in the central areas of Southwest and Northwest China, whole of South China, most areas of East China, and distinct areas of North and Central China (Fig. 4(d1)). And the regression coefficients were primarily significant in the northern regions of the Northwest, North, South and East China, southern areas of the Northeast and East China (Fig. 4(d2)). In most areas, the local coefficient estimates of “Economic” were positive. Comparatively, it is noteworthy that negative estimates were found in some areas. These results indicated that an inverted-U shaped Environmental Kuznets Curve (EKC)²⁷ relationship between GDP and $PM_{2.5}$ concentrations indeed exists. Most estimates were significant except in Northeast China, South China, central areas of East China and eastern regions of North China (Fig. 4(e1,2)). The regression coefficient estimates of “Urban” were primarily positive in the west areas of Northwest and Southwest China, Northeast and North China, and the estimates were highly significant in Northeast China, northeast and south of North China, south and central regions of Southwest China and most areas of Northwest China (Fig. 4(f1,2)).

“Topography” had negative effects on “ $PM_{2.5}$ Concentrations” except in the southern regions of Southwest and Northeast China, western areas of South China, and eastern districts of East and North China. And their relationships were significant in most areas of Northwest and Southwest China, north and south areas of East China, and the distinct regions of North and Central China (Fig. 4(g1,2)). Estimated local coefficients of “Climate” were for the most part negative over the study area which echoes its direct effect on “ $PM_{2.5}$ Concentrations” (Fig. 4(h1)). The stronger effects were found in the central districts of North China, toward eastern expansion. And in most regions of North China, west of South China, east of East China, and south of the Northeast and Southwest China, the estimates were significant (Fig. 4(h2)). In Fig. 4(i1), the direct effect of “Vegetation” has been illustrated with greater magnitude in the southeast areas of Central China, eastern and western regions of Northwest China, some specific areas of North China, and most areas of East China. And its regression coefficients in Northeast China, western regions of Northwest China and Southwest China and some districts of East China and North China were significant (Fig. 4(i2)).

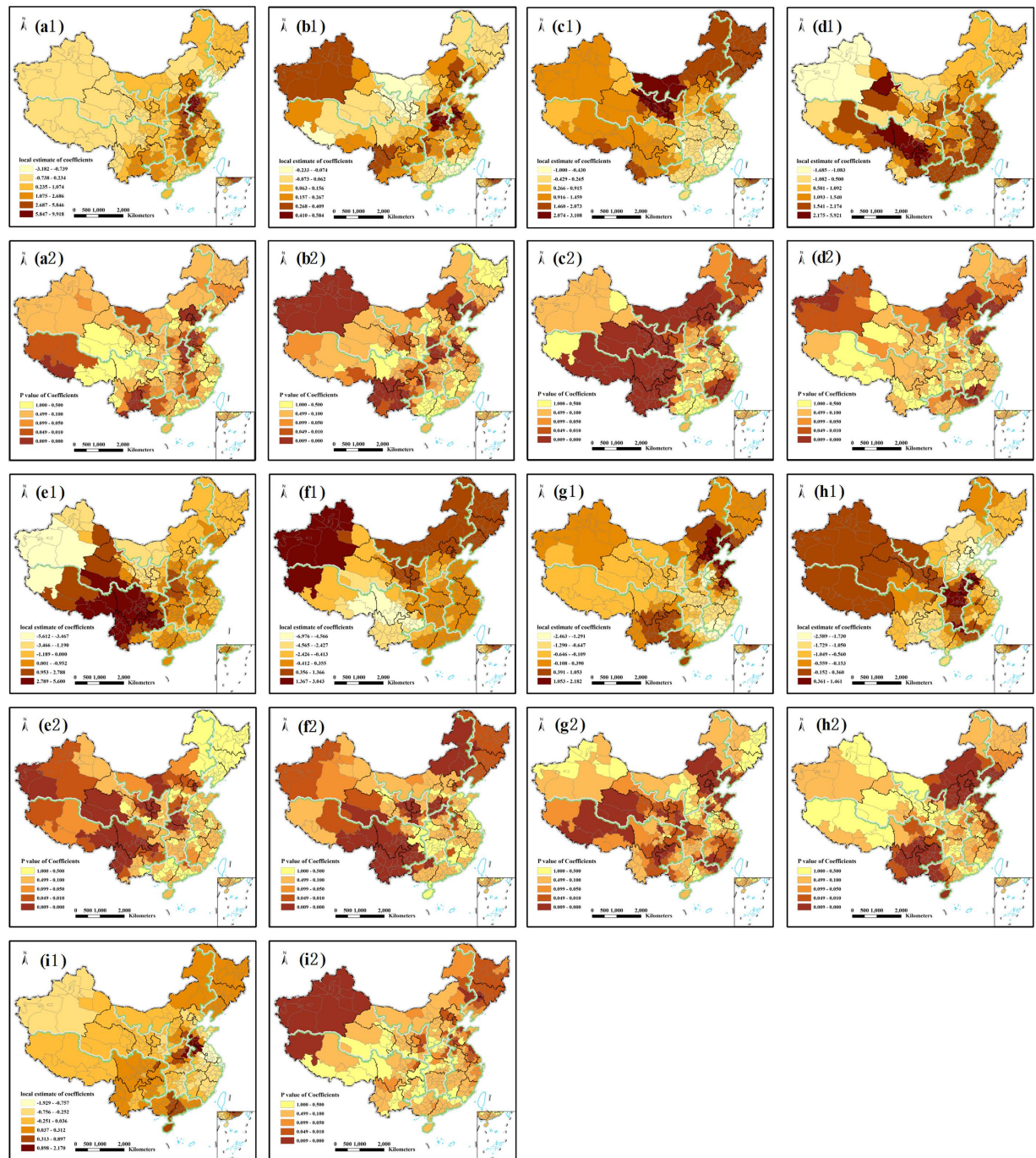


Figure 4. Spatial distribution of local estimates (a1–i1) and p values (a2–i2) of nine factors' coefficients derived from the GWR model. Local estimates of (a1) Road, (b1) Agriculture, (c1) Population, (d1) Industry, (e1) Economic, (f1) Urban, (g1) Topography, (h1) Climate, (i1) Vegetation, p values of (a2) Road, (b2) Agriculture, (c2) Population, (d2) Industry, (e2) Economic, (f2) Urban, (g2) Topography, (h2) Climate, (i2) Vegetation. All the maps were generated in ArcGIS10.2, URL: <http://www.esrichina-bj.cn/softwareproduct/ArcGIS/>.

It is expected that the variables such as road, agriculture, industry show direct positive effects on $PM_{2.5}$ concentrations, the variables like population, urban with negative signs are also detectable. Moreover, the variables of topography, climate and vegetation are the opposite. One reason for such counterintuitive signs as expressed by Chow *et al.*³⁷ is that the collinearity among some independent variables or the collinearity in estimated local coefficients. It is argued that the local coefficients could be correlated even when there is no collinearity among independent variables³⁸. Another reason is that some variables might be less significant at certain locations than others or are completely insignificant, which is due to the method employed in estimating the standard deviation in GWR models³⁹.

Measures and suggestions for controlling PM_{2.5} concentrations. Given the fact that PM_{2.5} pollutions in Mainland China is more and more severe, it is urgent to take quick measures. For example, actions to strictly forbid straw-burning, install desulphurization and dust removal device for coal-fired power plant, close heavily polluting industry, transfer other large industries elsewhere and replace coal with cleaner renewable energy sources, e.g., solar energy, hydrogen fuel, geothermal. Another effective measure to reduce PM_{2.5} emissions is reducing the use of motor vehicles. Some megacities (e.g., Beijing and Shanghai) have limited the vehicles usage through odd-and-even license plate rule and strict control of high displacement automobiles usage. Such measures could be effective in the short term, but will give considerably rise to the government cost and difficult to control in the near future when the pollution became more intense and integrated with regional economic issues²⁸.

To fundamentally mitigate PM_{2.5} concentrations, the following suggestions are proposed based on our study and the opinions of other investigations. Firstly, the improvement of fuel quality and the implementation of a more stringent vehicle emission standard are the most effective means of reducing PM_{2.5} emissions. Secondly, formulating strict monitoring system, immediately setting out to a nationwide monitoring network, making contingency plans for heavily polluted days and developing mitigation targets are the important means of governing PM_{2.5}. Regional central heating in winter is a substantial measure of the prevention and control of urban PM_{2.5}. Thirdly, the environmental protection policy should be prevention-driven rather than problem-driven. It is also important to emphasize that the impacts of each latent factor on PM_{2.5} in different regions are different. Thus, measures and strategies for controlling PM_{2.5} should be an integration of unified planning with the principle of adaptation to local conditions, e.g. the design of vegetation configuration based on local conditions to achieve the goal of increasing vegetation coverage. Furthermore, because of limited communication and lack of a negotiation process, we did not see significant trust building either among the public or between the public and the government, which could have been a major benefit of public participation⁴⁰. Hence the last and the most important suggestions are raising public awareness of PM_{2.5} perniciousness and environmental conservation and strengthening the involvement of the public and government.

Limitations and future research directions. The models and datasets used in this study reflected an integration of multiple scales, which would inevitably generate uncertainties in spatial statistics. Gridded population and GDP datasets are secondary derived data, and to some extent will introduce new uncertainty. Although there were no collinearity problems existed among the decisive factors, the information redundancy still existed. It affects the performance of the model and even produces counterintuitive signs.

As we all know, the correlations between PM_{2.5} and latent geographic factors that varied at different spatial and temporal scales; therefore, when analyzing their causal relationships, the scale effect should be considered to successfully uncover the spatial and temporal characteristics of PM_{2.5} concentrations. Furthermore, in future studies, it will be necessary to distinguish the influencing ranges of different latent geographic factors and to make certain the inconsistency of their reacting ranges.

Conclusions

In order to arouse the attention of researchers to investigate the causes of severe PM_{2.5} mass frequently in China nowadays from the macroscopic perspective, this paper analyzed the characteristics of spatiotemporal variations of PM_{2.5} concentrations in Mainland China during 1998–2012. This study was an initial attempt to explore the influencing factors, including natural geographical and socioeconomic factors of PM_{2.5} concentrations in 343 cities across Mainland China and dynamically evaluated the potential health risks of PM_{2.5} in 2000, 2005 and 2010. The following conclusions were drawn from this research:

(1) PM_{2.5} concentrations increased over most of Mainland China during the period 1998–2012. The proportion of area with low PM_{2.5} concentrations less than the WHO AQG of 10 µg/m³ declined significantly from 32.92% to 25.67%, while the proportion of area under high range concentrations greater than the WHO Interim Target-1 concentration of 35 µg/m³ increased significantly from 23.08% to 29.89%. The concentrations were particularly high in Central, North, Southwest and East China, and the most severe were in northern Henan, southern Hebei and Shanxi, central Shaanxi, eastern Sichuan and western Shandong.

(2) A significantly strong positive correlation was found between all variables of socioeconomic factors (road, agriculture, population, industry, economic, urban) and PM_{2.5} concentrations, while a significantly strong negative correlation was found between almost all variables of natural geographical factors (vegetation, topography, climate) and PM_{2.5} concentrations. Moreover, Moran's I for its residuals was 0.025 and was not significant at a 0.01 level (Z -score = 1.550, p = 0.1212), indicating that the GWR model didn't miss any key explanatory variables and was properly specified. It is also necessary to emphasize that the effects of each latent factor on PM_{2.5} in various regions are different. Therefore, regional measures and strategies for controlling PM_{2.5} should be integration of unified planning with the principle of adaptation to local conditions.

Methods

The technical flowchart of this study is shown in Fig. 5, mainly includes the following steps:

Step 1: Validate the accuracy of the annual PM_{2.5} concentrations grids dataset in Mainland China based on related literatures and evaluate the spatial and temporal characteristics of PM_{2.5} exposures from 1998 to 2012.

Step 2: Download data and extract latent geographic factors from these datasets.

Step 3: Analyze and compare the impact of each latent geographic factor on PM_{2.5} concentrations through multivariate correlation analysis. Moreover, calculate the VIF by OLS to detect whether collinearity problems existed among the factors. The goal of this step is to determine the decisive factors on PM_{2.5} concentrations.

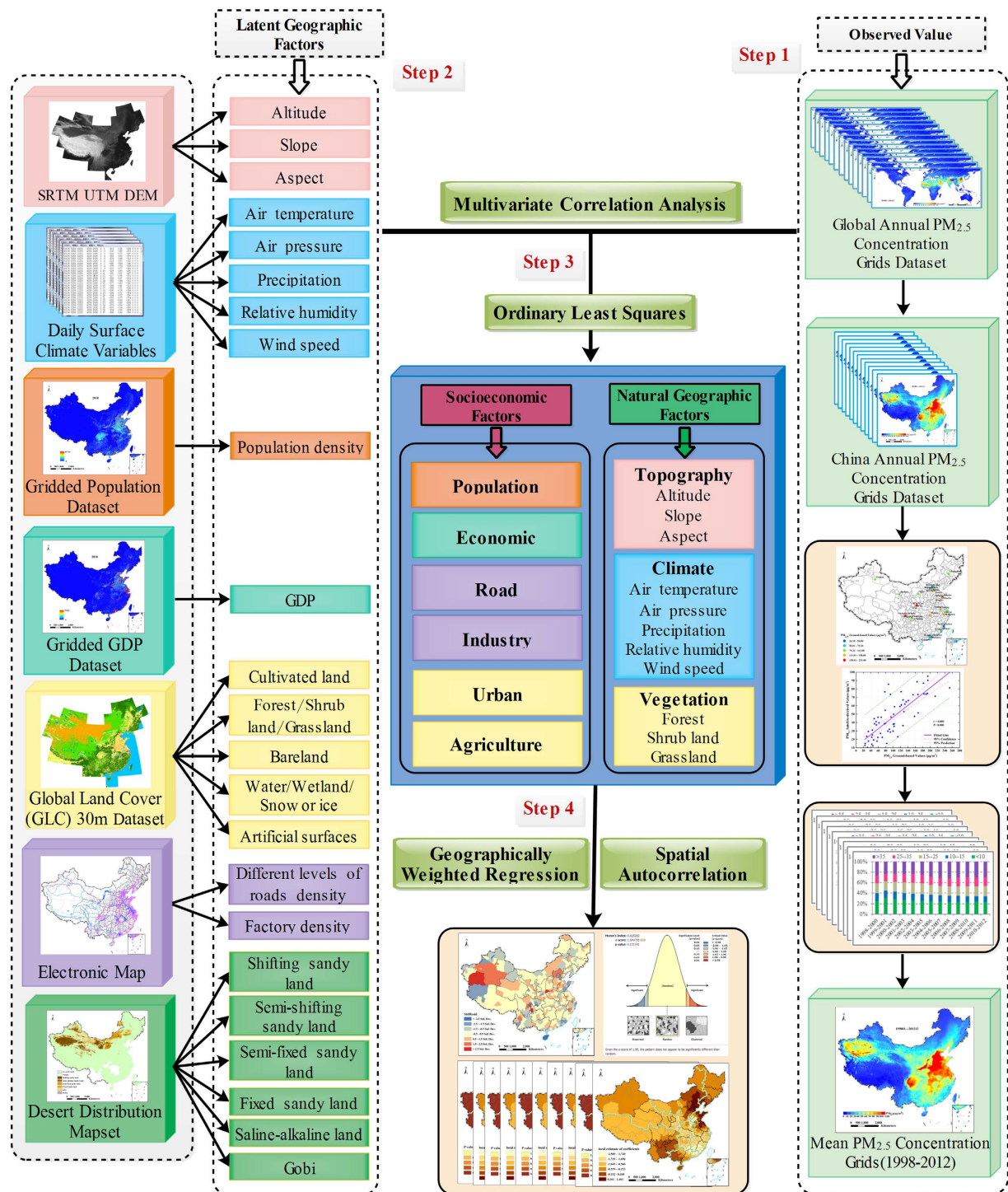


Figure 5. The technical flowchart of this study. All the maps were generated in ArcGIS10.2, URL: <http://www.esrichina-bj.cn/softwareproduct/ArcGIS/>.

Step 4: Apply the GWR method to explore spatial non-stationarity and varying relationships between $PM_{2.5}$ concentrations and decisive factors. Moreover, spatial autocorrelation analysis was used to detect the performance of the GWR.

In brief, conventional statistical analysis and GWR model were adopted in this study.

Conventional statistical analysis. In order to make the analysis straightforward, annual average $PM_{2.5}$ concentrations were categorized sequentially into six grades (Grade 1: <10 , Grade 2: 10–15, Grade 3: 15–25, Grade 4: 25–35, Grade 5: 35–100, Grade 6: $>100 \mu\text{g}/\text{m}^3$) according to WHO's air quality guidelines² and the latest

Geographical subarea	Provinces and municipalities
Northeast China	Liaoning, Jilin, and Heilongjiang
North China	Beijing, Tianjin, Hebei, Shanxi and Inner Mongolia
East China	Jiangsu, Zhejiang, Shanghai, Anhui, Fujian, Shandong and Jiangxi
Central China	Hunan, Hubei and Henan
South China	Guangxi, Hainan and Guangdong
Southwest China	Sichuan, Yunnan, Guizhou, Chongqing and Tibet
Northwest China	Gansu, Qinghai, Ningxia, Shaanxi and Xinjiang

Table 1. Seven geographical subareas of Mainland China.

version of China's ambient air quality standard (GB 3095–2012)⁴¹. WHO's air quality guidelines have four standards, including one air quality guideline (AQG:10 $\mu\text{g}/\text{m}^3$) and three interim targets. As for the interim targets, Interim Target-1 (IT-1:35 $\mu\text{g}/\text{m}^3$) was in line with the annual average secondary standard of $\text{PM}_{2.5}$ concentrations of China's ambient air quality standard; Interim Target-2 (IT-2:25 $\mu\text{g}/\text{m}^3$) and Interim Target-3 (IT-3:15 $\mu\text{g}/\text{m}^3$) was consistent with the annual average primary standard of $\text{PM}_{2.5}$ concentrations of China's ambient air quality standard. The characteristics of the variations in the seven regions were analyzed to quantify the degree of spatial variation across Mainland China. To objectively analyze the influencing factors of $\text{PM}_{2.5}$ concentrations in Mainland China as well as to reduce the influence of data from different years, mean $\text{PM}_{2.5}$ concentrations were calculated at city level using the $\text{PM}_{2.5}$ concentrations dataset and city boundary layer derived from electronic map. Similarly, values of each latent geographic factors (e.g. population density, road density, slope, wind speed, forest cover ratio, etc.) of 343 cities were also summarized.

Geographically weighted regression. Geographically weighted regression (GWR) was adopted to explore the local spatial heterogeneity of the causal relationships between $\text{PM}_{2.5}$ concentrations and geographic factors. It is a powerful technique to examine geographically non-stationarity and varying relationships between dependent/response variable Y and a set of independent/explanatory variables X_j ($j = 1, 2, \dots, m$) at regional scale⁴². Through adding the geographical location information into the conventional regression process, the GWR attempts to show how the relationship between the dependent variable and the independent variables varies over the entire space. Also, it is described by the equation:

$$Y_i = \beta_0(U_i, V_i) + \beta_1(U_i, V_i)X_{i1} + \beta_2(U_i, V_i)X_{i2} + L + \beta_m(U_i, V_i)X_{mi} + \varepsilon_i \quad (1)$$

where (U_i, V_i) and ε_i are respectively space coordinate and regression residual of the i th location. And unlike conventional global regression, the coefficients $\beta_j(U_i, V_i)$ ($j = 1, 2, \dots, m$) are varying conditionals on the location.

In GWR model, the regression coefficients show the local spatial variation, and the standard errors of the coefficients illustrate the reliability of the estimated coefficients⁴³. Considering the samples are not regularly spaced in our study, GWR v4.0^{33,34} with the adaptive bandwidth and bi-square kernel was implemented to build the model. Meanwhile, golden-section search which can efficiently identify the optimal bandwidth size in most cases was used in this research. Furthermore, the AICc was extensively adopted to compare the global OLS model with a local GWR model.

Data and Study Area

Study area. Cities, including county-level, prefectural-level cities and municipalities are the basic administrative units which can be used to reveal Chinese Mainland's natural geographic features and socioeconomic condition, as well as its air pollution. Additionally, natural geographic factors and socioeconomic factors match well with the $\text{PM}_{2.5}$ concentrations distribution at the city level. According to the above mentioned analysis, 343 cities were used as the basic study unit to explore the influencing factors of $\text{PM}_{2.5}$ in Mainland China. In order to better and more easily elaborate the results, Mainland China was divided into seven geographical subareas: Northeast China (NEC), North China (NC), East China (EC), Central China (CC), South China (SC), Southwest China (SWC), and Northwest China (NWC) (Table 1).

Multisource data used in this study is listed in Table 2, which is classified into two broad categories: 1) $\text{PM}_{2.5}$ data used as dependent/response variable Y in GWR model; 2) natural geographical and socioeconomic data, such as DEM, population, GDP and so on used as independent/explanatory variables X_j ($j = 1, 2, \dots, m$).

$\text{PM}_{2.5}$ data. There is a lack of publically available global remote sensing data related to $\text{PM}_{2.5}$, until van Donkelaar *et al.*^{4–6} used the GEOS-Chem global chemical transport model (<http://geos-chem.org/>) to successfully map global ground-level $\text{PM}_{2.5}$ concentrations. They are based on total column aerosol optical depth (AOD) from a combination of MODIS (Moderate Resolution Imaging SpectroRadiometer), MISR (Multi-angle Imaging SpectroRadiometer) and SeaWiFS (Sea-Viewing Wide Field-of-View Sensor) AOD satellite instruments and coincident aerosol vertical profiles. The global annual $\text{PM}_{2.5}$ concentrations grids dataset represents a series of three-year running mean grids (1998–2012) of fine particulate matter (solid particles and liquid droplets) that provides the highest accuracy, largest coverage (from 70°N to 55°S), longest temporal range and highest resolution (6 arc-minutes, 0.1 degree or approximately 10 km at the equator). It has been effectively applied on a national and regional scale^{26,28,44}. These data is derived from Socioeconomic Data and Applications Center (sedac)—Hosted by the Center for International Earth Science Information Network (CIESIN) at Columbia University (<http://sedac.ciesin.columbia.edu/>). A subset of the global $\text{PM}_{2.5}$ concentrations grids dataset (1998–2012) covering the

Data	Sources	Year	Resolution or scale
Global Annual PM _{2.5} Grids Datasets	http://sedac.ciesin.columbia.edu/	1998–2012	10000 m
SRTM UTM DEM	http://www.gscloud.cn	2003	90 m
Daily Surface Climate Variables of China	http://cdc.cma.gov.cn/home.do	2010	1000 m
Gridded GDP Dataset	http://www2.geodata.cn/index.html	2010	1000 m
Gridded Population Dataset	http://www2.geodata.cn/index.html	2010	1000 m
Global Land Cover 30 m Dataset	http://www.globallandcover.com/GLC30Download/index.aspx	2004–2010	30 m
Desert Distribution Mapset	http://www2.geodata.cn/index.html	2000	1:100000
Electronic Map		2010	1:10000

Table 2. Basic information of the eight datasets used in this study.

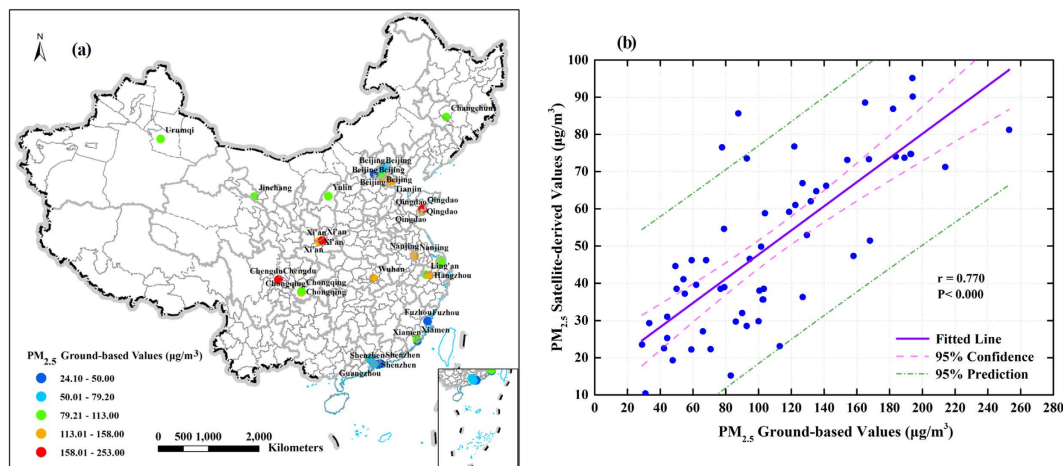


Figure 6. (a) The ground-based PM_{2.5} measurements collected from the literature for locations in Mainland China and (b) linear correlation of the two datasets (PM_{2.5} ground-based values and satellite-derived values). The left map was generated in ArcGIS10.2, URL: <http://www.esrichina-bj.cn/softwareproduct/ArcGIS/>.

Mainland China was used in this research. Furthermore, there exist data voids in the original PM_{2.5} dataset due to snow-covered mountains, perennial cloud and sensor malfunctions, in particular for the Tibetan Plateau. With the purpose of alleviating this problem, the spline interpolation method⁴⁵ was utilized to make up data voids.

Although van Donkelaar *et al.* has validated the accuracy of the global annual PM_{2.5} concentrations grids dataset based on the agreement between satellite-derived estimates and ground-based measurements^{5,6}. However, considering that the relationship between AOD-PM_{2.5} can differ by space and countries, it was still necessary to evaluate the reliability of the dataset for the specific area of Mainland China. Because China has not built national PM_{2.5} network monitoring sites until the end of 2012, continuous observation data with which to validate satellite-derived air quality data was not available. Therefore, we collected Chinese ground-based PM_{2.5} measurements of different sites at different times from relevant published literature. Consequently, as Supplementary Table S1 shown, 61 sample points with site location, geocoordinates, ground-based PM_{2.5} values, and sampling periods were extracted from 34 relevant studies.

The ground measured PM_{2.5} concentrations and spatial distribution of these sample points are shown in Fig. 6(a). Furthermore, according to the location and period of the sample points, the corresponding satellite-derived values of PM_{2.5} concentration data were calculated. Also, the linear correlation between “PM_{2.5} Ground-based Values” and “PM_{2.5} Satellite-derived Values” is shown in Fig. 6(b). A significant overall agreement is found ($r = 0.770$), which indicated that the satellite-derived PM_{2.5} concentration data were reliable. The agreement was higher than that for Europe ($r = 0.730$) and North America ($r = 0.760$)⁶. However, the satellite-derived values tend to be lower than the ground-based values, and this was consistent with results of Europe and North America⁶. Some of this underestimate may arise from the different PM_{2.5} measurement methods which were used in remote sensing and site monitoring. While others might be due to that the time spans of the sample points and remote sensing were inconsistent, e.g., some sample points were monitored only in a few days or months of one year, while remote sensing monitored yearly.

Natural geographical and socioeconomic data.

- (1) DEM data: The SRTM UTM DEM was obtained from the processing of a void-filled version of the SRTM3 dataset (SRTM3 V4.1). The SRTM3 V4.1 was offered by the International Center for Tropical Agriculture

- (CIAT) and its data gaps were filled with different interpolation algorithms⁴⁶ and SRTM DEM. The SRTM DEM was acquired by radar interferometry (InSAR) during about ten days in February 2000, when a Space Shuttle mapped the Earth surface between 60°N and 56°S with C-band radar. Based on these data and other products, a covering more than 80% of earth's land surface DEM with three arc-seconds resolution (~90 m, SRTM3) was derived and distributed for free in 2003. This data has been validated through comparison with ground control points. SRTM UTM DEM is downloaded from International Scientific & Technical Data Mirror Site, Computer Network Information Center, and Chinese Academy of Sciences (<http://www.gscloud.cn>).
- (2) Meteorological site data: Meteorological site data are derived from the “Daily Surface Climate Variables of China” catalog (SURF_CLI_CHN_MUL_DAY_V3.0 on June-25-2014), which is released by the Climatic Data Center, National Meteorological Information Center, China Meteorological Administration and China Meteorological Data Sharing Service System (<http://cdc.cma.gov.cn/home.do>). The dataset starts on 1 January, 1951, and keeps running till now, with a total of 824 fundamental stations (The number of stations does vary in a different year, but has remained stable since 1980) throughout Mainland China. It consists of over 22 climate variables, although we focus on daily average air temperature (TEM), average air pressure (PRS), average relative humidity (RHU), average wind speed (WIN) and total precipitation (PRE) in this study. The raster/grid maps with 1 km resolution of the annual mean air temperature, air pressure, wind speed, rainfall and relative humidity are generated using thin plate spline spatial interpolation method⁴⁷. The spline surfaces are fitted as functions of latitude, longitude, and elevation. The same elevation data used is SRTM UTM DEM 90 m data. It is noted that the elevation is a co-predictor, and thus a topographic correction for the gridded data is calculated during the interpolation.
 - (3) GDP and population data: GDP and population are common indicators of socioeconomic development. The gridded GDP (2010) and population datasets (2010) in Mainland China were all provided by the National Data Sharing Infrastructure of System Science (<http://www2.geodata.cn/index.html>). Respectively according to the relationship between demographic data, GDP data and land use types, these gridded datasets with a spatial resolution of 1 km were transformed from statistical yearbook⁴⁸.
 - (4) Land cover data: Global Land Cover 30 m dataset (GLC30) was downloaded from Global Land Cover Information Service System (<http://www.globallandcover.com/GLC30Download/index.aspx>). This kind of land cover data was derived by using multisource data, including Landsat TM, ETM+, Environmental Disaster Alleviation Satellite (HJ-1) multispectral images. The dataset has a grid cell resolution of 30 m, covers the global land surface from 80°N to 80°S, and consists of ten land cover types: cultivated land, forest, grassland, shrubland, wetland, water bodies, tundra, artificial surfaces, bareland, permanent snow and ice.
 - (5) Desert distribution mapset and electronic map: As with the GDP and population datasets, Desert distribution Mapset was also derived from the National Data Sharing Infrastructure of System Science (<http://www2.geodata.cn/index.html>). This mapset covers most area of Mainland China excluding Hunan, Yunnan province, Shanghai and Chongqing municipality. Desert was divided into six types: shifting sandy land, semi-shifting sandy land, fixed sandy land, semi-fixed sandy land, saline-alkaline land and gobi.

Administrative boundary, factory and road density of Mainland China used in this work were extracted from an electronic map in 2010. All datasets were rectified to the Universal Transverse Mercator (UTM) projection system (datum WGS 1984, zone 48), and integrated into a geodatabase.

References

1. Martini, F. M. S., Hasenkopf, C. A. & Roberts, D. C. Statistical analysis of PM_{2.5} observations from diplomatic facilities in China. *Atmos. Environ.* **110**, 174–185 (2015).
2. World Health Organization. Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide. Global Update 2005 Available at: http://apps.who.int/iris/bitstream/10665/69477/1/WHO_SDE_PHE_OEH_06.02_eng.pdf. (Date of access: 19/04/2014) (2006).
3. Dominici, F. *et al.* Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases. *J. Am. Med. Assoc.* **295**, 1127–1134 (2006).
4. van Donkelaar, A. *et al.* Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environ. Health Perspect.* **118**, 847–855 (2010).
5. van Donkelaar, A. *et al.* Optimal estimation for global ground-level fine particulate matter concentrations. *J. Geophys. Res. Atmos.* **118**, 5621–5636 (2013).
6. van Donkelaar, A., Martin, R., Brauer, M. & Boys, B. Use of satellite observation for long-term exposure assessment of global concentration of fine particulate matter. *Environ. Health Perspect.* **123**, 135–143 (2015).
7. Liu, C. *et al.* A land use regression application into assessing spatial variation of intra-urban fine particulate matter (PM_{2.5}) and nitrogen dioxide (NO₂) concentrations in City of Shanghai, China. *Sci. Total Environ.* **565**, 607–615 (2016).
8. Hong, S., Jiao, L. & Ma, W. Variation of PM_{2.5} concentration in Hangzhou, China. *Particulology* **11**, 55–62 (2013).
9. Wang, P. *et al.* Spatial and seasonal variations of PM_{2.5} mass and species during 2010 in Xi'an, China. *Sci. Total Environ.* **508**, 477–487 (2015).
10. Chen, Y. *et al.* Long-term variation of black carbon and PM_{2.5} in Beijing, China with respect to meteorological conditions and governmental measures. *Environ. Pollut.* **212**, 269–278 (2016).
11. Xie, Y., Zhao, B., Zhang, L. & Luo, R. Spatiotemporal variations of PM_{2.5} and PM₁₀ concentrations between 31 Chinese cities and their relationships with SO₂, NO₂, CO and O₃. *Particulology* **20**, 141–149 (2015).
12. Ma, Z. *et al.* Satellite-derived high resolution PM_{2.5} concentrations in Yangtze River Delta Region of China using improved linear mixed effects model. *Atmos. Environ.* **133**, 156–164 (2016).
13. Lin, G. *et al.* Spatio-temporal variation of PM_{2.5} concentrations and their relationship with geographic and socioeconomic factors in China. *Int. J. Environ. Res. Public Health.* **11**, 173–186 (2014).
14. Fang, X., Zou, B., Liu, X., Sternberg, T. & Zhai, L. Satellite-based ground PM_{2.5} estimation using timely structure adaptive modeling. *Remote Sens. Environ.* **186**, 152–163 (2016).
15. Liu, Z., Hu, B., Liu, Q., Sun, Y. & Wang, Y. Source apportionment of urban fine particle number concentration during summertime in Beijing. *Atmos. Environ.* **96**, 359–369 (2014).

16. Wang, Y. *et al.* Source apportionment of fine particulate matter during autumn haze episodes in Shanghai, China. *J. Geophys. Res. Atmos.* **119**, 1903–1914 (2014).
17. Hua, Y. *et al.* Characteristics and source apportionment of PM_{2.5} during a fall heavy haze episode in the Yangtze River Delta of China. *Atmos. Environ.* **123**, 380–391 (2015).
18. Chen, L., Liu, C., Zou, R., Yang, M. & Zhang, Z. Experimental examination of effectiveness of vegetation as bio-filter of particulate matters in the urban environment. *Environ. Pollut.* **208**, 198–208 (2016).
19. Beckett, K. P., Freer-Smith, P. H. & Taylor, G. Particulate pollution capture by urban trees: effect of species and windspeed. *Glob. Change Biol.* **6**, 995–1003 (2000).
20. Irga, P. J., Burchett, M. D. & Torpy, F. R. Does urban forestry have a quantitative effect on ambient air quality in an urban environment? *Atmos. Environ.* **120**, 173–181 (2015).
21. Räsänen, J. V. *et al.* Effects of species-specific leaf characteristics and reduced water availability on fine particle capture efficiency of trees. *Environ. Pollut.* **183**, 64–70 (2013).
22. Gromke, C. & Ruck, B. Influence of trees on the dispersion of pollutants in an urban street canyon— Experimental investigation of the flow and concentration field. *Atmos. Environ.* **41**, 3287–3302 (2007).
23. Zhang, C., Ni, Z. & Ni, L. Multifractal detrended cross-correlation analysis between PM_{2.5} and meteorological factors. *Physica A Statistical Mechanics & Its Applications*, **438**, 114–123 (2015).
24. Li, Y., Chen, Q., Zhao, H., Wang, L. & Tao, R. Variations in PM₁₀, PM_{2.5} and PM_{1.0} in an urban area of the Sichuan basin and their relation to meteorological factors. *Atmos.* **6**, 150–163 (2015).
25. Zhang, H., Wang, Y., Hu, J., Ying, Q. & Hu, X. Relationships between meteorological parameters and criteria air pollutants in three megacities in China. *Environ. Res.* **140**, 242–254 (2015).
26. Han, L., Zhou, W., Li, W. & Li, L. Impact of urbanization level on urban air quality: a case of fine particles (PM_{2.5}) in Chinese cities. *Environ. Pollut.* **194**, 163–170 (2014).
27. Hao, Y. & Liu, Y. The influential factors of urban PM_{2.5} concentrations in China: a spatial econometric analysis. *J. Clean. Prod.* **112**, 1443–1453 (2016).
28. Han, L., Zhou, W. & Li, W. City as a major source area of fine particulate (PM_{2.5}) in China. *Environ. Pollut.* **206**, 183–187 (2015).
29. Bai, X., Shi, P. & Liu, Y. Realizing China's urban dream. *Nature* **509**, 158–160 (2014).
30. Zhang, Q. *et al.* Asian emissions in 2006 for the NASA INTEX-B mission. *Atmos. Chem. Phys.* **9**, 5131–5153 (2009).
31. Pui, D. Y. H., Chen, S. & Zuo, Z. PM_{2.5} in China: measurements, sources, visibility and health effects, and mitigation. *Particuology* **13**, 1–26 (2014).
32. Hurvich, C. M., Simonoff, J. S. & Tsai, C. L. Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *J. R. Statist. Soc. B* **60**, 271–293 (1998).
33. Nakaya, T. GWR4 User Manual: Windows application for geographically weighted regression modelling. Available at: <https://gwrtools.github.io/>. (Date of access: 18/06/2014) (2014).
34. Fotheringham, A. S., Brunsdon, C. & Charlton, M. *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships* Ch. 9, 212–213 (John Wiley and Sons, 2002).
35. Mo, H. *et al.* Characterization of summer PM_{2.5} aerosols from four forest areas in Sichuan, SW China. *Particuology* **20**, 94–103 (2015).
36. Zheng, Y., Zhang, Q., Liu, Y., Geng, G. & He, K. Estimating ground-level PM_{2.5} concentrations over three megalopolises in China using satellite-derived aerosol optical depth measurements. *Atmos. Environ.* **124**, 232–242 (2016).
37. Chow, L. F. *et al.* Transit Ridership Model Based on Geographically Weighted Regression. *Trans. Res. Record J. Trans. Res. Board.* **1972**, 105–114 (2006).
38. Shariat-Mohaymany A., Shahri M., Mirbagheri B. & Matkan A. A. Exploring Spatial Non-Stationarity and Varying Relationships between Crash Data and Related Factors Using Geographically Weighted Poisson Regression. *Trans. Gis.* **19**, 321–337 (2014).
39. Congdon, P. *Applied Bayesian Modeling* Ch. 7 308–313 (John Wiley and Sons, 2003).
40. Huang, G. PM_{2.5} opened a door to public participation addressing environmental challenges in China. *Environ. Pollut.* **197**, 313–315 (2015).
41. Ministry of Environmental Protection of the People's Republic of China. China National Ambient Air Quality Standards. Available at: <http://kjs.mep.gov.cn/hjbhzbz/bzwb/dqjhbd/dqjzlbz/201203/W020120410330232398521.pdf>. (Date of access: 29/03/2015) (2012).
42. Xu, S. *et al.* A new satellite-based monthly precipitation downscaling algorithm with non-stationary relationship between precipitation and land surface characteristics. *Remote Sens. Environ.* **162**, 119–140 (2015).
43. Gao, Y., Huang, J., Li, S. & Li, S. Spatial pattern of non-stationarity and scale-dependent relationships between NDVI and climatic factors—A case study in Qinghai–Tibet Plateau, China. *Ecol. Indic.* **20**, 170–176 (2012).
44. Peng, J., Chen, S., Lü, H., Liu, Y. & Wu, J. Spatiotemporal patterns of remotely sensed PM_{2.5} concentration in China from 1999 to 2011. *Remote Sens. Environ.* **174**, 109–121 (2016).
45. Franke, R. Smooth interpolation of scattered data by local thin plate splines. *Comp. and Math. with Appl.* **8**, 273–281 (1982).
46. Reuter, H. I., Nelson, A. & Jarvis, A. An evaluation of void-filling interpolation methods for SRTM data. *Int. J. Geogr. Inf. Sci.* **21**, 983–1008 (2007).
47. Hutchinson, M. F. & Xu, T. ANUSPLIN Version 4.4 User Guide. Available at: <http://fennerschool.anu.edu.au/files/anusplin44.pdf> (Date of access: 22/03/2014) (2013).
48. Yang, X., Jiang, D., Wang, N. & Liu, H. Method of pixelizing population data. *J. Geogr. Sci.* **57**, 70–75 (2002).

Acknowledgements

The authors would like to thank Climatic Data Center, International Scientific & Technical Data Mirror Site, National Data Sharing Infrastructure of System Science, Socioeconomic Data and Applications Center, and Global Land Cover Information Service System for sharing data. This work was supported by the Key Program of Chinese National Natural Science Foundation (No. 41631176), the Jiangsu Provincial Natural Science Foundation (No. BK2012018), and China Geological Survey project (No. 12120115050901).

Author Contributions

J.L. and P.D. designed the study; J.L. and M.C. collected and processed the data; A.S. and Z.X. analyzed the results; and J.L. wrote the main manuscript text. J.X. provided constructive comments on the whole manuscript. All authors reviewed the manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Luo, J. *et al.* Spatiotemporal Pattern of PM_{2.5} Concentrations in Mainland China and Analysis of Its Influencing Factors using Geographically Weighted Regression. *Sci. Rep.* 7, 40607; doi: 10.1038/srep40607 (2017).

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2017