scientific reports

OPEN

Heterogeneity and the determinants of PM$_{2.5}$ in the Yangtze River Economic Belt

Siyou Xia[1,2], Xiaojie Liu[3], Qing Liu[4], Yannan Zhou[1,2] & Yu Yang[1,2]✉

Haze has reached epidemic levels in many Chinese cities in recent years. Few studies have explored the determinants and heterogeneity of PM$_{2.5}$. This paper investigates the spatiotemporal characteristics of PM$_{2.5}$ through spatial analytical methods based on aerosol optical depth data from the Yangtze River Economic Belt (YREB) between 2000 and 2017. Geographically weighted regression and geodetector models were applied to assess the heterogeneity of key factors influencing PM$_{2.5}$. The results indicate that the annual concentrations of PM$_{2.5}$ in the YREB were 23.49–37.37 µg/m³, with an initial increase and a later decrease. PM$_{2.5}$ pollution showed a diagonal high spatial distribution pattern in the northeast and a low spatial distribution in the southwest, as well as a noticeable spatial convergence. The spatial variability of PM$_{2.5}$ was enlarged, and its main fractal dimension was in the northeast-southwest direction. There were clear spatiotemporal variations in the impacts of natural and anthropogenic factors on PM$_{2.5}$. Our findings contribute to a better understanding of the impact mechanisms of PM$_{2.5}$ and the geographic factors that form persistent and highly polluted areas and imply that more specific coping strategies need to be implemented in various areas toward successful particulate pollution prevention and control.

Haze is present in most urban regions of the world and is a global environmental problem. The rapid increase in the urbanization and industrialization of many cities in China has caused haze pollution, which is dominated by the enhanced concentration of PM$_{2.5}$ (particulate matter with an aerodynamic diameter ≤ 2.5 µm). PM$_{2.5}$ has significant adverse impacts on public health, the environment, and climate because of its small diameter, high activity, and ability to transport noxious substances in the air with long residence times[1–3]. Epidemiological studies have confirmed that PM$_{2.5}$ can induce various respiratory and cardiovascular diseases, impair the body's immune system, and increase the mortality of people exposed to it[4,5]. More than 1.3 million people in China die prematurely every year due to prolonged exposure to polluted air, which is approximately 40% of the global total[6]. The International Agency for Research on Cancer (IARC) classified PM$_{2.5}$ as a human carcinogen in 2013[7]. Moreover, continuous haze damages the overall image of a city and weakens the attraction of tourism, talent, and investments, which restricts the sustainable development of the economy[8–10]. The issues related to cleaning the air have aroused attention both domestically and globally.

PM$_{2.5}$ has attracted significant attention from the atmospheric and climate science community, and many studies have explored environmental issues related to it, including spatiotemporal distribution and drivers. Research on the spatial patterns of PM$_{2.5}$ mainly focuses on environmental coping policies by identifying its distribution patterns and spatial effects. These studies reveal that PM$_{2.5}$ has regional, cumulative, and compound effects and possesses significant spatiotemporal variability[11–13]. In addition, PM$_{2.5}$ is not restricted to the local environment and can be diffused or transferred to neighboring areas largely through external forces, such as atmospheric circulation causing spillover effects[14]. The identification and estimation of PM$_{2.5}$ with spatial characteristics should be studied by means of spatial analytical methods rather than by traditional statistical theory based on independent observations[15]. Spatial statistics started in the 1970s with the goal of understanding spatial dependence, spatial association, and other relations among data related to geographical locations with wide applications in many fields. Accurate identification of PM$_{2.5}$ determinants can provide a strong theoretical basis for the prevention and control of air pollution. Therefore, many scholars are committed to understanding source analyses, emission inventories, chemical conversion, and regional transport of PM$_{2.5}$ and have achieved significant results. In general, PM$_{2.5}$ is considered to be emitted from local sources, including primary sources,

[1]Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China. [2]College of Resources and Environment, University of Chinese Academy of Sciences, Beijing 100049, China. [3]School of Geographic and Oceanographic Sciences, Nanjing University, Nanjing 210023, China. [4]School of Geography and Planning, Sun Yat-Sen University, Guangzhou 510275, China. ✉email: yangyu@igsnrr.ac.cn

such as traffic fumes, industrial activities, soil dust, biomass, and coal combustion, and secondary sources of gaseous pollutants (such as SO_2, NOx, and NH_3) formed by complex chemical reactions[10,16–18]. Thus, a local $PM_{2.5}$ concentration change is the result of the combined action of natural and human factors[19,20]. Natural conditions, such as meteorology, topography, and vegetation, play important roles in the generation, accumulation, transfer, diffusion, and settlement of $PM_{2.5}$ and profoundly affect its local concentration[20,21]. With respect to the influence of socioeconomic factors related to anthropogenic activities on the distribution of $PM_{2.5}$, studies suggest that energy-intensive economic growth and nonecological urbanization increase $PM_{2.5}$ concentrations, implying that economic development, urbanization, industrialization, land use, and energy use mixtures and efficiency can affect urban air quality[13,22–24]. These studies provide many insights into urban particle pollution from the perspectives of artificial and natural conditions.

The Yangtze River Economic Belt (YREB) is one of China's most crucial economic and ecological corridors that connects three national urban agglomerations of the Sichuan–Chongqing, the middle reaches of the Yangtze River, and the Yangtze River Delta (YRD). The YREB accounts for more than 40% of China's population and GDP, and its unique geographical advantages and vast economic hinterland make it the region with the greatest economic growth potential in the next 30 years[25]. However, urban air quality along the Yangtze has been deteriorating and has suffered from varying degrees of haze pollution due to continuous and intense industrial and human activities[26]. Recently, some areas in the middle-lower reaches of the Yangtze River have had more than 100 haze d year^{-1}, with some cities even exceeding 200 d year^{-1}[27]. In the upper reaches of the Yangtze River, the ecological environment is relatively vulnerable as economic development is scaled up[28]. The alleviation of haze pollution involves the integrity of the ecological and environmental system and the quality of local people's lives[9]. In 2014, the development of the YREB was given national strategy status, leading to the fact that potential conflicts between economic growth and environmental conservation may be more prominent than ever before[29]. In 2018, President Xi emphasized the protection of the eco-environmental status of the YREB. Hence, the eco-environmental conservation of the Yangtze River is a priority for the government of China[25].

Until now, studies on the environmental problems of the YREB have mainly focused on the macro-level, including environmental quality and risk assessment, a low-carbon economy, and a sustainable development strategy[9,28]. Studies on environmental problems at the micro-level focus more on water pollution, while studies on air pollution are still insufficient[27]. As the China National Environmental Monitoring Centre (CNEMC) did not include $PM_{2.5}$ in its monitoring index system until 2012, relatively few studies focused on $PM_{2.5}$ in the YREB prior to 2012[30]. The CNEMC has begun widespread real-time monitoring of $PM_{2.5}$ since 2013. Furthermore, the use of remote sensing data for long-term time series research was limited to local areas and specific cities, and studies on factors of influence were mostly in terms of the overall perspective[31]. Obvious spatial differences in the natural conditions and socioeconomic development of various regions exist, and hence, a global analysis cannot reveal the spatial heterogeneity of the effects of various factors[32]. On the other hand, the majority of research has not conducted spatial analysis from a geographical viewpoint. Although some research has considered the spatial dependence of $PM_{2.5}$ pollution, it has still neglected the spatial differentiation of the $PM_{2.5}$ level at the urban agglomeration scale and lacked analysis of the interactive mechanism between factors, which may have led to incomplete and biased conclusions[33].

In this study, we attempt to fill the aforementioned knowledge gaps. We consider 130 cities within the YREB as the research area and employ geostatistical and spatial analytical methods to summarize the spatiotemporal evolutionary features of $PM_{2.5}$. Furthermore, we adopt a geographically weighted regression model and geographical detectors to quantitatively analyze the spatial differentiation and interactive effects of socioeconomic and natural conditions on $PM_{2.5}$ during various periods. Thus, our study is significant in terms of meeting the demand for air quality improvements in the YREB and sheds light on implementing effective urban $PM_{2.5}$ pollution abatement policies.

Data and methods

Study area. The YREB spans the eastern, central, and western regions of China, including nine provinces (Zhejiang, Jiangsu, Jiangxi, Anhui, Hubei, Hunan, Sichuan, Yunnan, and Guizhou) and two municipalities (Shanghai and Chongqing) (Fig. 1). The YREB comprises an area of 2.05×10^6 km², accounting for 21.3% of the country's land area. Moreover, its economy and population density are 6.2 and 4.5 times the national average, respectively, establishing its incomparable role in China's development strategies[8]. However, the YREB is an ecologically vulnerable zone, and hence, assessing $PM_{2.5}$ pollution in the region is of great strategic significance for national ecological security and regional sustainable development[28].

Data acquisition and processing. Our data sources have four components: (1) $PM_{2.5}$ concentration data from satellite retrievals. The annual concentrations of $PM_{2.5}$ were obtained from the global surface raster with a resolution of $0.01° \times 0.01°$, which was released by the Atmospheric Composition Analysis Group (ACAG) (https://sites.wustl.edu/acag/). This $PM_{2.5}$ remote sensing inversion dataset has the largest global coverage, the longest time span, and the highest accuracy, and it has been widely verified and effectively applied in China[34]. To verify the accuracy of the dataset in the YREB, we calculated the annual concentration of each station through real-time monitoring data of the YREB from 2015 to 2017 obtained from the CNEMC (http://www.cnemc.cn) and correlated it with the $PM_{2.5}$ concentration from the remote sensing inversion data at the corresponding position. The correlation coefficient was 0.82 and significant, indicating a strong correlation and consistency between these datasets. (2) Data on natural parameters include wind, precipitation, vegetation, and topography[21,35]. The normalized differential vegetation index (NDVI) is the best indicator of vegetation coverage and growth status, and the original data were derived from the National Aeronautics and Space Administration (NASA) (https://www.nasa.gov/), while other supporting data were obtained from the Resource and Environment Data Cloud

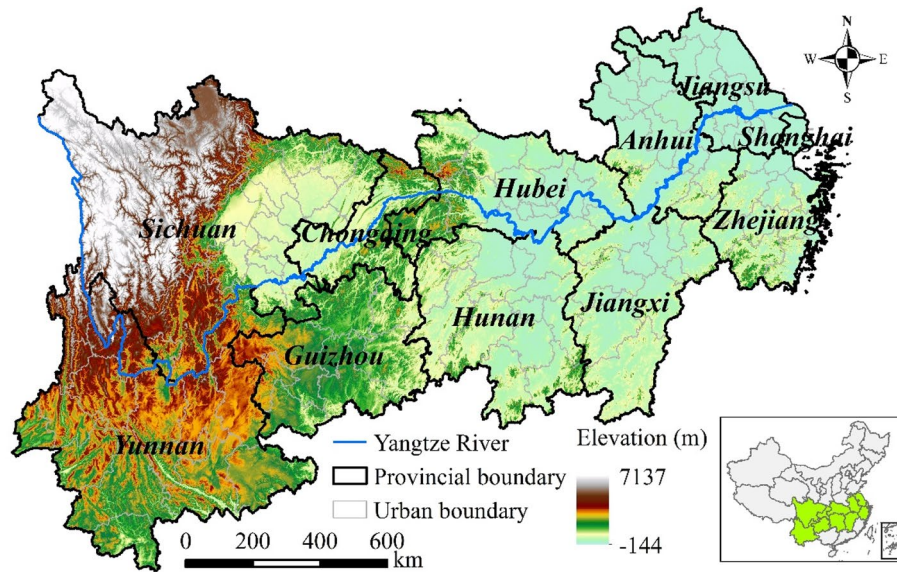**Figure 1.** General overview of the YREB. Standard map services are provided by the Ministry of Natural Resources of China (http://bzdt.ch.mnr.gov.cn/), GS (2020)4619.

Platform (http://www.resdc.cn). (3) Socioeconomic data includes per capita GDP, population density, economic density, urbanization rate, industrial structure, and energy consumption[33,36]. Studies have shown that there is a significant linear correlation between nighttime light data from the Defense Meteorological Satellite Program/ Operational Line Scanner (DMSP/OLS) and energy consumption[37]. We converted the light intensity to gray pixel values and used the sum of all DMSP/OLS raster gray values in each region as an indicator of energy consumption in the region. The DMSP/OLS dataset was obtained from the National Oceanic and Atmospheric Administration (NOAA) (https://www.ngdc.noaa.gov), and the remaining data were collected from the China Urban Statistics Yearbook published by the National Bureau of Statistics (http://www.stats.gov.cn/). (4) Geographic information data includes the spatial vector map, which was derived from the National Catalog Service for Geographic Information (http://www.webmap.cn).

As multicollinearity in the selected variables may cause information redundancy, we adopted the variance inflation factor (VIF) for multicollinearity diagnosis with a threshold of 5. We excluded the indicators of economic density and energy consumption because their VIFs were greater than 5. Thus, 8 explanatory variables were used in the model, including per capita GDP (*pgdp*), population density (*popd*), urbanization rate (*urba*), secondary industry share (*indu*), annual wind speed (*wind*), annual precipitation (*prec*), NDVI (*ndvi*), and topographic relief (*topo*). To reduce data heteroscedasticity, all the indicators were treated with Z–standardization.

Methods

Spatial autocorrelation analysis. Tobler's first law of geography states that everything is related to everything else, but nearby things are more related than distant things[15]. Therefore, we adopted the classical spatial autocorrelation method to quantitatively measure the spatial dependence of $PM_{2.5}$ in neighboring regions. Global Moran's I is written as[38]:

$$I \frac{n \sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} W_{ij} \sum_{i=1}^{n} (x_i - \overline{x})^2} \tag{1}$$

where x_i and x_j are the observations of spatial units i and j, respectively; $\overline{x}$ is the mean of n locations; and W_{ij} is a spatial weight matrix. $I \in [-1, 1]$, where $I > 0$ indicates positive correlation, $I < 0$ indicates negative correlation, and $I = 0$ indicates mutual independence.

Spatial variogram analysis. The spatial variogram is a geostatistical method to describe the structure and randomness of regionalized variables and is often used to effectively measure the spatial structure characteristics and variation pattern of geographical variables[39]. The variogram is expressed as:

$$\gamma(\lambda) = \frac{1}{2N(\lambda)} \sum_{i=1}^{N(\lambda)} [\xi(x_i) - \xi(x_i + \lambda)]^2 \tag{2}$$
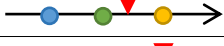
Diagram	Criterion	Interaction
	$q(X_1 \cap X_2) < \text{Min}(q(X_1), q(X_2))$	Nonlinear weakening
	$\text{Min}(q(X_1), q(X_2)) < q(X_1 \cap X_2) < \text{Max}(q(X_1), q(X_2))$	Univariate nonlinear weakening
	$q(X_1 \cap X_2) > \text{Max}(q(X_1), q(X_2))$	Bivariate enhancement
	$q(X_1 \cap X_2) = q(X_1) + q(X_2)$	Independent
	$q(X_1 \cap X_2) > q(X_1) + q(X_2)$	Nonlinear enhancement

Table 1. Types of interaction between two covariates. ● $\text{Min}(q(X_1), q(X_2))$ is the minimum value between $q(X_1)$ and $q(X_2)$; ● $\text{Max}(q(X_1), q(X_2))$ is the maximum value between $q(X_1)$ and $q(X_2)$; ● $q(X_1) + q(X_2)$ is the sum of $q(X_1)$ and $q(X_2)$; ▼$q(X_1 \cap X_2)$ is the interaction between $q(X_1)$ and $q(X_2)$.

where $\xi(x_i)$ and $\xi(x_i + \lambda)$ are the values of the regionalized variables at points x_i and $x_i + \lambda$, respectively, and $N(\lambda)$ is the sample size with separation distance λ.

There are three major parameters derived from the variogram model[39–41]. The nugget parameter (C_0) is a random spatial variance; the partial sill parameter (C) is a structural spatial variance; the sill parameter ($C_0 + C$) represents the total degree of spatial variation. The nugget effect ($C_0/(C_0 + C)$) indicates whether regional or local factors are more important for $PM_{2.5}$. If the nugget effect is less than 0.25, it indicates strong spatial correlation; if it is between 0.25 and 0.75, it shows moderate spatial correlation; and if it is greater than 0.75, it indicates weak spatial correlation[41]. The range parameter (A_0) represents the maximum spatial distance of the correlation.

The fractal dimension is another important parameter that characterizes the variogram, and its value is determined by the relationship between the variogram $\gamma(\lambda)$ and distance λ:

$$2\gamma(\lambda) = \lambda^{(4-2D)} \tag{3}$$

where D is the slope of the double logarithm linear regression equation; the higher the value is, the higher the heterogeneity caused by the spatial autocorrelation. The closer the value is to 2, the more balanced the spatial distribution.

Geographically weighted regression modeling (GWR).   Geographically weighted regression (GWR) is a spatial regression model based on the idea of local smoothness[42]. This technique constructs an independent equation for each unit in the study area and incorporates the spatial attributes of data into the regression model so that the relationship between variables can change with the change in spatial location, thus reflecting the spatial nonstationarity of parameters in different regions. Its formula is as follows[16]:

$$y_i = \varphi_0(\omega_i, \alpha_i) + \sum_{i=1}^{p} \varphi_p(\omega_i, \alpha_i) x_{ip} + \varepsilon_i \tag{4}$$

where x_{ip} is a dimensional interpretation variable matrix; (ω_i, α_i) represents the longitude and latitude coordinates at the ith observation point; $\varphi_p(\omega_i, \alpha_i)$ is the regression coefficient of the ith observation point; and ε_i is a random error term.

Geographical detector technique.   The geographical detector technique is a set of statistical methods to detect spatially stratified heterogeneity and reveal the driving forces behind it[43]. It can detect the possible causal relationship between variables by verifying the consistency of the spatial distribution of two variables[44]. The explanatory power of the factor is measured by the q value, and its expression is as follows:

$$q = 1 - \frac{\sum_{h=1}^{L} N_h \sigma_h^2}{N \sigma^2} \tag{5}$$

where L refers to the strata of variable Y ($PM_{2.5}$) or factor X; N_h and σ_h^2 are the number of units and variance of strata h, respectively; N and σ^2 are the total number of units and variance, respectively; $q \in [0, 1]$, and the greater the value is, the stronger the explanatory power of this factor is.

The purpose of interactive detection is to assess whether the factors X_1 and X_2 work together to increase or decrease the explanatory power on Y or whether the impact of these factors on Y is independent. The evaluation method first calculates the q value of the two factors X_1 and X_2 acting on Y, namely, $q(X_1)$ and $q(X_2)$, and then calculates the q value for their interaction, namely, $q(X_1 \cap X_2)$; finally, $q(X_1)$ and $q(X_2)$ are compared with $q(X_1 \cap X_2)$. The discriminant criteria can be divided into 5 categories (Table 1).
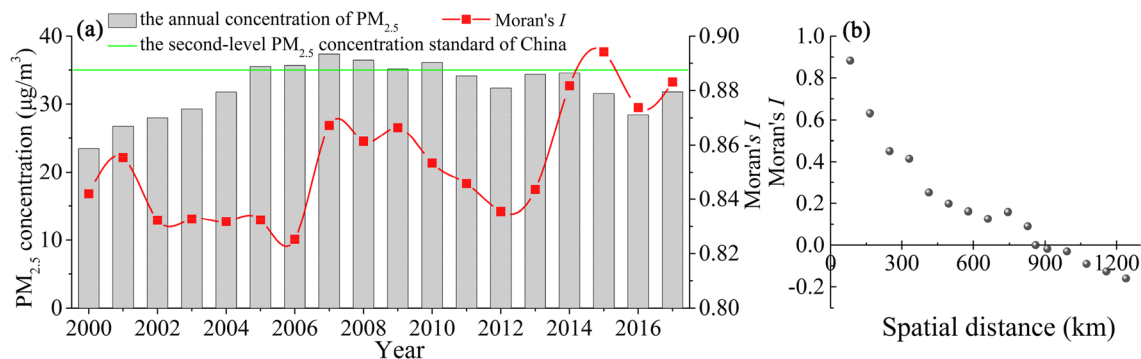
Figure 2. Variation in the PM$_{2.5}$ concentration and Moran's I value in the YREB.
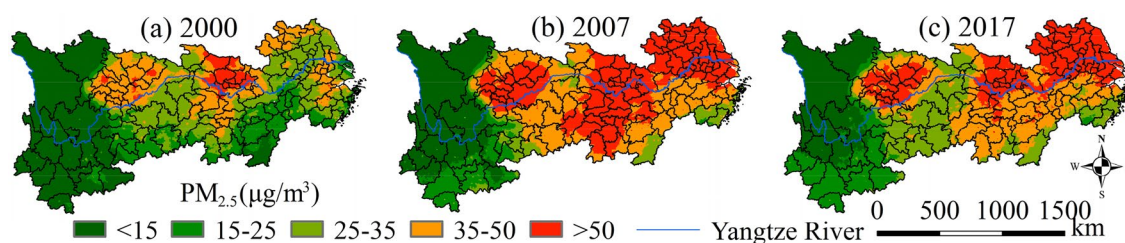

Figure 3. Spatial patterns of PM$_{2.5}$ in the YREB between 2000 and 2017.

Year	A_0 (km)	C_0	$C_0 + C$	$C_0/(C_0 + C)$	R^2	Residual sum of squares	Optimal fitting model
2000	625.27	0.0029	0.0543	0.0534	0.960	2.061E−04	Gaussian
2007	737.85	0.0036	0.0492	0.0732	0.963	1.444E−04	Gaussian
2017	635.66	0.0011	0.0484	0.0227	0.987	5.271E−05	Gaussian

Table 2. Fitting parameters of the PM$_{2.5}$ concentration variogram.

Results

Spatiotemporal evolutionary characteristics of PM$_{2.5}$. During 2000–2017, the overall PM$_{2.5}$ level in the YREB showed a trend of first increasing and then decreasing (Fig. 2a). In 2005–2010, the annual PM$_{2.5}$ concentration in the YREB was higher than the second-level standard of China of 35 μg/m^3. During 2000–2007, the PM$_{2.5}$ level increased from 23.49 to 37.37 μg/m^3, an increase of 59.07%. Afterward, it decreased to 31.79 μg/m^3 in 2017, a decrease of 14.93%. This improvement may be due to the effects of the national tenth 5-year plan on controlling the total emissions of major pollutants, adjusting the industrial structure, and establishing a monitoring, statistics, and assessment system for energy conservation and pollution emission reduction. It was further reinforced by the implementation of the Action Plan for Air Pollution Prevention and Control in 2013 (http://www.mee.gov.cn).

Figure 2a shows that the sliding interval of the global Moran's I value over the years was [0.825, 0.894], and all the values were significant at the 99% level. This indicates that the PM$_{2.5}$ distribution was not random but was a significant spatial agglomeration. We pursued the association of the evolution rules of spatial correlation with distance and found that the distance threshold to maintain spatial correlation was approximately 870 km (Fig. 2b). Within this spatial range, PM$_{2.5}$ had significant positive interaction effects, which increased with the shortening of distance.

Figure 3 displays the spatial patterns and evolution of PM$_{2.5}$ in the YREB from 2000 to 2017. Its main features are as follows: (1) cities with an annual PM$_{2.5}$ level of less than 15 μg/m^3 were mainly concentrated in ethnic minority areas, where the environmental conditions were relatively good. However, air quality continuously deteriorated, albeit at low levels, which needs attention. (2) The PM$_{2.5}$ level was higher in the lower reaches than in the upper reaches and higher on the north bank than on the south bank, presenting a diagonal spatial distribution pattern with an obvious lowland plain directivity. (3) Urban economic activities and population density were closely related to PM$_{2.5}$ in three centers, namely, the Cheng-Yu area, the Wuhan metropolitan area, and the northern Anhui–Jiangsu region.

Spatiotemporal variation characteristics of PM$_{2.5}$. The value of the variogram increased with increasing separation distance, indicating that the spatial autocorrelation of PM$_{2.5}$ changed from strong to weak with increasing distance (Table 2). During 2000–2017, the variation range was 625–738 km, and it showed an overall

Year	Isotropic		South–North (0°)		Northeast–Southwest (45°)		East–West (90°)		Southeast–Northwest (135°)	
	D	R^2	D	R^2	D	R^2	D	R^2	D	R^2
2000	1.536	0.959	1.517	0.847	1.430	0.985	1.603	0.918	1.527	0.922
2007	1.482	0.977	1.554	0.763	1.409	0.997	1.464	0.914	1.520	0.756
2017	1.453	0.936	1.509	0.774	1.292	0.985	1.258	0.787	1.567	0.611

Table 3. Variable difference dimension of PM$_{2.5}$ concentrations.
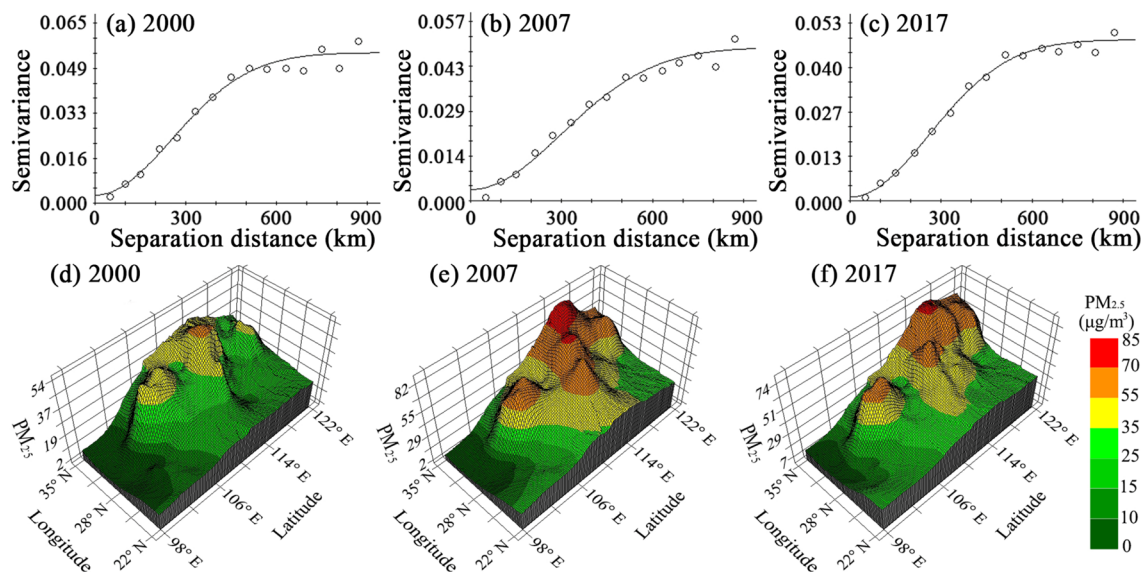

Figure 4. Evolution of PM$_{2.5}$ concentrations in the YREB based on variograms.

upward trend, implying that the spatial correlation of PM$_{2.5}$ was partly expanded in scope. In addition, the nugget effect indicated that regional-scale factors are more important for the distribution of PM$_{2.5}$.

In terms of the fractal dimension (Table 3), the isotropic dimension continuously decreased from 1.536 in 2000 to 1.453 in 2017, indicating that the spatial difference in PM$_{2.5}$ was continuously expanding. The northeast-southwest direction had the greatest goodness of fit and the smallest fractal dimension and showed a downward trend. This result denoted that the spatial variation in PM$_{2.5}$ in this direction was continuously strengthened, making it the main direction in terms of spatial difference. The southeast-northwest fractal dimension was the largest, and its decisive coefficient continued to decrease, showing that the spatial difference in PM$_{2.5}$ in this direction continued to weaken and remained relatively evenly balanced. We conducted 3D-kriging interpolation, which further depicted the spatial distribution and evolution morphology of PM$_{2.5}$ (Fig. 4). It was evident that the spatial pattern steadily transitioned from a gradient differentiation to a relatively balanced structure that formed a trend indicating that the middle-lower reaches of the Yangtze River drove the whole basin PM$_{2.5}$ level to increase.

Analysis models of the factors of influence

GWR model results. The GWR model fitting results are shown in Table 4, in which the adjusted R^2 values are above 0.9, indicating good fitting performance.

The regression coefficients of socioeconomic factors, such as per capita GDP, population density, urbanization rate, and industrial structure, were mainly positive. Among them, population density had the largest impact on PM$_{2.5}$ and the most obvious spatial difference, followed by per capita GDP, while the coefficients of urbanization and industrial structure were relatively small. The regression coefficients of natural factors, such as wind, precipitation, vegetation, and topography, were distributed in positive and negative intervals, and the instability was striking in different years. Among them, the coefficient of the distribution interval for topographic relief was the longest, indicating that spatial heterogeneity was the largest.

Spatial heterogeneity of factors of influence. The coefficient of anthropogenic factors increased, signifying that extensive economic development and intensive human activities aggravated haze pollution. Existing studies have argued that economic growth strongly correlates with regional environmental pollution, but the relationship between GDP and PM$_{2.5}$ is significantly different in various regions[19]. Per capita GDP had positive impacts on PM$_{2.5}$ in the middle-upper reaches of the Yangtze River (Fig. 5a–c). This implied that PM$_{2.5}$ in economically underdeveloped areas was more sensitive to economic development and that the economic growth of these regions came at the cost of the environment. Some cities in the YRD showed negative correlation effects,

Variables and parameters	2000	2007	2017
pgdp	− 0.2792 to 0.4534*	− 0.1764 to 0.6815**	− 0.1982 to 0.4392*
popd	− 0.2327 to 1.2928**	0.0070–0.9061***	− 0.0448 to 1.4807***
urba	− 0.2099 to 0.5177**	− 0.0996 to 0.3314*	− 0.1814 to 0.1884**
indu	− 0.1569 to 0.1629*	− 0.1711 to 0.2928*	− 0.0612 to 0.1158***
wind	− 1.5755 to 0.2788***	− 0.6810 to 0.1470***	− 0.4908 to 0.3147*
prec	− 0.8223 to 0.4898***	− 0.3110 to 0.0518*	− 0.2301 to 0.3349**
ndvi	− 0.6251 to 0.0474***	− 0.7815 to 0.0027***	− 0.5942 to − 0.0026**
topo	− 0.6420 to 0.1005***	− 0.7488 to − 0.1018***	− 1.1785 to − 0.1100***
Bandwidth	3.321	3.627	3.289
AICc	118.931	51.926	16.542
R^2	0.949	0.966	0.977
Adjusted R^2	0.913	0.944	0.960

Table 4. Fitting results of the GWR model. *, ** and *** indicate significance at the 10%, 5% and 1% levels, respectively.

which indicated that the development planning in the above areas was relatively good. With technological progress, industrial upgrading and economic development, these areas were essentially coordinated with the surroundings.

The coefficient of population density showed approximately the same spatial distribution at the three time nodes (Fig. 5d–f), all of which increased from the coastal areas to the inland areas, and among them, the east–west difference was obvious in 2017. This may be because of the increase in urban traffic flow and production, with a higher population density contributing to an increase in the local $PM_{2.5}$ level. The control of air pollution in the middle-upper reaches of the Yangtze Basin was still weak, thereby making the impact of population density more significant. Some studies have held that an increase in population density may have an agglomeration effect in promoting regional technical progress, thus facilitating a reduction in the local $PM_{2.5}$ level[14]. The technical advances brought by the agglomeration effect for the population size in the YREB were not significant. This may be related to the migration of people to large cities in recent years and the disordered nature of population mobility.

The coefficient of the urbanization rate was positive at the three time nodes. The proportion of positive values in 2007 was relatively large, and the positive values in 2017 marginally declined (Fig. 5g–i). During 2000–2007, areas with low urbanization rates in the Yangtze Basin were experiencing rapid urbanization, and the urban infrastructure industry developed rapidly[9]. A large quantity of building dust entered the atmosphere, aggravating urban $PM_{2.5}$ pollution. However, areas with a high urbanization rate, such as the YRD, tended to mature, and a stagnant infrastructure industry was conducive to the reduction in emissions of fine particles.

The industrial structure had a negative impact on $PM_{2.5}$ in the middle-lower Yangtze and Cheng–Yu areas, aggravating local air pollution (Fig. 5j–l). This is consistent with the findings of existing studies confirming that industrial activities were the main drivers of $PM_{2.5}$ in most areas[19]. The impact of the industrial structure in the Chengdu–Chongqing areas was strong, which may be because of heavy industries, such as the energy, chemical, and machinery industries, with relatively high direct energy consumption and pollutant emissions[25]. Optimization of the industrial structure significantly affects the local $PM_{2.5}$ level. The coefficient had a weak impact on the YRD because the local industrial structure was dominated by the service industry and was relatively stable; thus, there was limited scope for further optimization.

The coefficients of wind in the inland areas were mainly positive and decreased from the central region to the west (Fig. 5m–o). The negative impact was dominant in the eastern areas, and as the distance to the coast decreased, the negative impact increased. This may be related to the impacts of topography and monsoons[27]. Coastal areas were mostly alluvial plains with flat terrains, and they were affected by the local circulation caused by monsoon and temperature differences. Clean air from the ocean had important dilution effects on pollutants, and thus, the coefficient was mainly negative. In the Sichuan Basin, the impact of closed topography restricted the diffusion of airflow, and the transport of wind caused pollutants in the region to interact with each other; thus, the coefficient was mainly positive. Similar findings were made in the Fenwei Plain, where basin topography exists[30].

The impact of precipitation on regional $PM_{2.5}$ presented a negative correlation at the three time nodes (Fig. 5p–r). A positive influence was mainly distributed in the Sichuan Basin and some areas of the Yunnan–Guizhou Plateau in western China, while the negative effect decreased from the coastal to inland areas. Concretely, the regions with a high regression coefficient were mainly located in the upper and middle reaches of the Yangtze Basin, while the eastern coastal areas with abundant rainfall had a small regression coefficient, indicating that abundant precipitation had a positive impact on $PM_{2.5}$ in most cities, and this effect was more distinct in areas with relatively deficient precipitation.

The impact of the NDVI on $PM_{2.5}$ had a negative correlation that was mainly distributed in the middle-lower Yangtze Plains (Fig. 5s–u). Research has shown that vegetation growth is correlated with climate and is affected by topography and human activities, resulting in a complex correlation between $PM_{2.5}$ and vegetation[21]. The high-value areas were mainly in Yunnan and Hunan Provinces, while the low-value areas were concentrated in the middle-lower Yangtze Plain. In 2017, the NDVI showed a negative impact, indicating that vegetation's
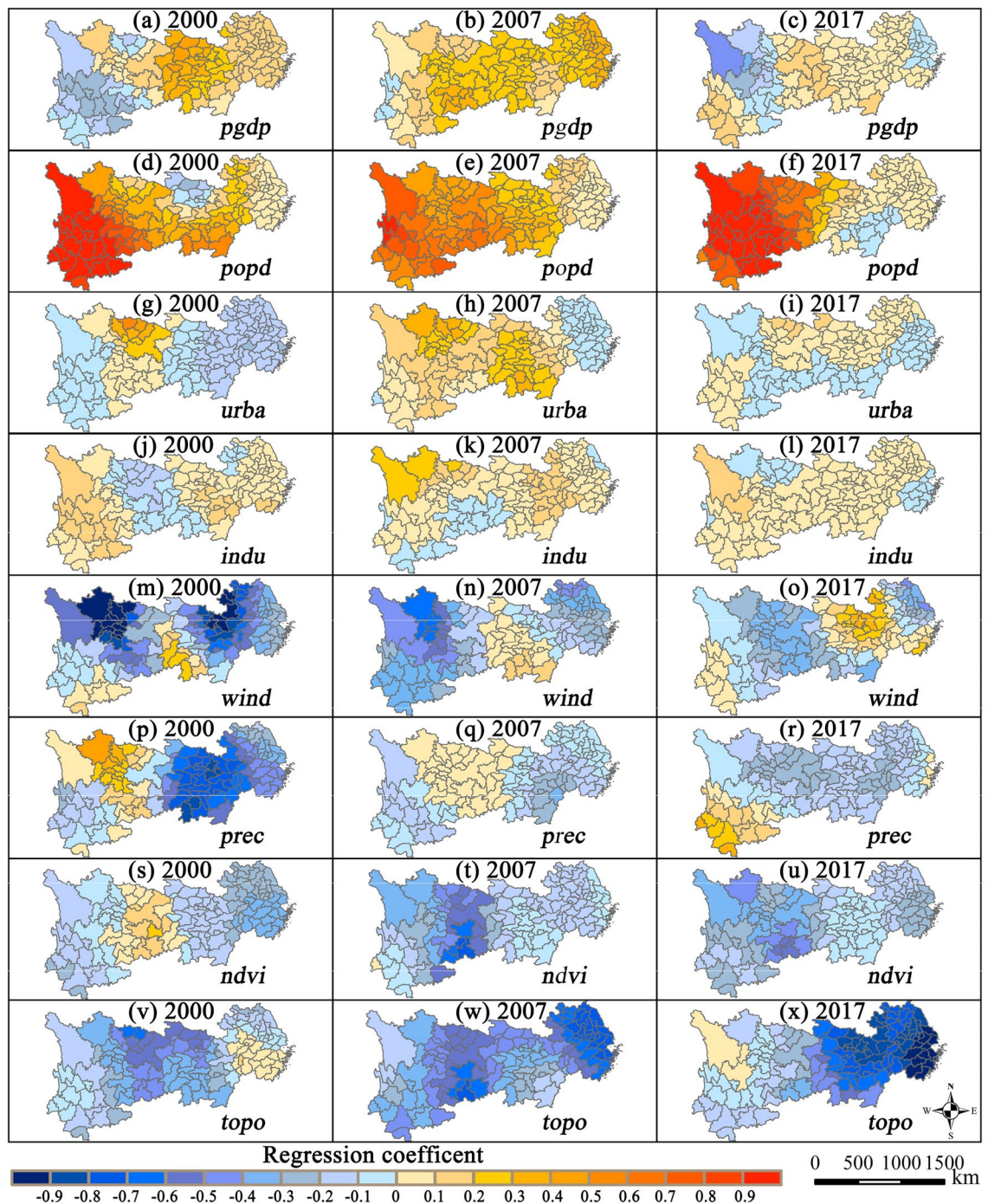
Figure 5. Spatial distribution of regression coefficients of the GWR model.

inhibitory effect on PM$_{2.5}$ was distinctly enhanced. This may be related to the rapid restoration and development of vegetation in the Yangtze Basin, whose forest coverage rate rose from 24% in 2000 to 39% in 2017[9].

Topographic relief was negatively correlated with regional PM$_{2.5}$ in the Yangtze Basin, which was conducive to the improvement in air quality (Fig. 5v–x). Specifically, the high-value regions were located in Sichuan, Chongqing, and Zhejiang. Lower values were concentrated in the middle-lower reaches of the Yangtze River. Local topography affects the diffusion and dilution of atmospheric pollutants in an area by influencing meteorological conditions[20].

Interactions between factors.   Table 5 shows the results of the interactions between factors. The impact of the interactions between factors was greater than that for individuals, and the interaction effects included nonlinear enhancement and bifactor enhancement. When wind speed, precipitation, and vegetation interacted with PM$_{2.5}$ in pairs, the nonlinear enhancement effect was generated at all three time nodes, and the explana-

Factors	2000	2007	2017	Factors	2000	2007	2017
$pgdp \cap popd$	NE (0.682)	NE (0.745)	BE (0.709)	$wind \cap prec$	NE (0.624)	NE (0.488)	NE (0.749)
$pgdp \cap urba$	NE (0.281)	NE (0.420)	BE (0.421)	$wind \cap ndvi$	NE (0.570)	NE (0.486)	NE (0.814)
$pgdp \cap indu$	NE (0.370)	NE (0.225)	NE (0.472)	$wind \cap topo$	NE (0.666)	NE (0.794)	BE (0.843)
$popd \cap urba$	NE (0.654)	BE (0.757)	BE (0.748)	$prec \cap ndvi$	NE (0.442)	NE (0.427)	NE (0.373)
$popd \cap indu$	BE (0.657)	NE (0.747)	BE (0.762)	$prec \cap topo$	NE (0.602)	BE (0.757)	NE (0.841)
$urba \cap indu$	NE (0.397)	NE (0.432)	NE (0.516)	$ndvi \cap topo$	NE (0.649)	NE (0.842)	BE (0.847)

Table 5. Results of interaction detecting. *NE* nonlinear enhancement, *BE* bifactor enhancement.

tory power was notably varied in different periods. When the industrial structure interacted with the per capita GDP and urbanization level on $PM_{2.5}$, a nonlinear enhancement effect was exerted at all three time nodes, and the explanatory power was continuously improved. Nonlinear enhancement means that the interactive impact of two factors is greater than the sum of the impacts when they act alone. The interactive types of $pgdp \cap popd$, $pgdp \cap urba$, $wind \cap topo$, $prec \cap topo$, and $ndvi \cap topo$ were dominated by nonlinear enhancement, although they varied at different times. The types of $popd \cap urba$ and $popd \cap indu$ exerted a bifactor enhancement effect, which was not as significant as that of the nonlinear enhancement.

Conclusions and policy implications

Conclusions. Haze pollution in Chinese cities has escalated to hazardous levels in recent years. This environmental problem has become a great challenge for public health and urban sustainable development. Our exploration of $PM_{2.5}$ in the YREB provides useful results for haze prediction, which is an important step toward protecting people from health damage caused by poor air quality. The results indicate that the annual $PM_{2.5}$ level of the YREB displayed an upward trend before 2007 and a fluctuating downward trend thereafter. $PM_{2.5}$ had significant spatial heterogeneity and convergence characteristics. There were clear spatiotemporal differences in the impact of various factors on the $PM_{2.5}$ pollution pattern. In the socioeconomic layer, population has the greatest impact, followed by the economy and industry, while urbanization was a relatively stable factor causing the rise in the $PM_{2.5}$ level. Among the natural factors, topography and vegetation mainly exerted a negative impact on $PM_{2.5}$, while the potency and direction of others changed with spatiotemporal changes.

Policy implications. As the Yangtze River plays a vital role in China's eco-environmental systems, stricter measures should be implemented to meet the goal of sustainable development. First, special attention should be given to natural factors when distributing industries and residences; for instance, topography and wind, which notably affect $PM_{2.5}$, should be considered, while urban air ducts and green belt designs must be optimized. Second, region-targeted policies should be considered based on spatial differentiation. Downstream areas should play a leading role in promoting pollution prevention and control technologies, while upstream areas should actively drive ecological protection. More emphasis should be placed on transregional linkage governance when formulating mitigation measures. Third, heavily and highly polluted industrial sectors must be closed or upgraded, as they are major sources of particulates, while sectors with environmental protection, new energy, and a low-carbon economy and technology should be highly encouraged, such as building hi-tech eco-industrial parks. Last, it seems inevitable that air pollution will be aggravated because of burgeoning urbanization, and therefore, reducing the impact of anthropogenic activity on the atmosphere through environmental policy and education will reduce $PM_{2.5}$ pollution.

Limitations and recommendation. Our study mainly focuses on the exploration of the heterogeneity and determinants of $PM_{2.5}$ at the macro-level. However, haze reduction ultimately needs to be implemented at micro-level enterprises, so investigations of micro-enterprises should be strengthened. In addition, $PM_{2.5}$ is closely related to other air pollutants, which leads to compound air pollution with multiple pollutants. Therefore, it is necessary to explore the mechanism of compound pollution in the future to better formulate control measures.

References
1. Lelieveld, J., Evans, J. S. & Fnais, M. The contribution of outdoor air pollution sources to premature mortality on a global scale. *Nature* **525**(7569), 367–371 (2015).
2. Kan, H. D., Chen, R. J. & Tong, S. L. Ambient air pollution, climate change, and population health in China. *Environ Int.* **42**(7), 10–19 (2012).
3. Figueres, C. & Landrigan, P. J. Tackling air pollution, climate change, and NCDs: Time to pull together. *Lancet* **392**, 1502–1503 (2018).
4. Jin, Y., Andersson, H. & Zhang, S. Air pollution control policies in China: A retrospective and prospects. *Int. J. Environ. Res. Public Health* **13**(12), 1219 (2016).
5. Samet, J. M., Dominici, F. & Curriero, F. C. Fine particulate air pollution and mortality in 20 U.S. cities, 1987–1994. *N. Engl. J. Med.* **343**(24), 1742–1749 (2000).

6. Yin, P., Michael, B., Aaron, J. C. & Wang, H. D. The effect of air pollution on deaths, disease burden, and life expectancy across China and its provinces, 1990–2017: An analysis for the Global Burden of Disease Study 2017. *Lancet Planet Health* https://doi.org/10.1016/S2542-5196(20)30161-3 (2020).

7. Xu, P., Chen, Y. F. & Ye, X. J. Haze, air pollution, and health in China. *Lancet* **382**(9910), 2067–2067 (2013).

8. Tian, X., Dai, H., Geng, Y., Wilson, J. & Wu, R. Economic impacts from $PM_{2.5}$ pollution-related health effects in China's road transport sector: A provincial-level analysis. *Environ. Int.* **115**, 220–229 (2018).

9. Chen, Y. S., Zhang, S. H. & Huang, D. S. The development of China's Yangtze River Economic Belt: How to make it in a green way?. *Sci. Bull.* **62**, 1–20 (2017).

10. Li, M. N. & Zhang, L. L. Haze in China: Current and future challenges. *Environ. Pollut.* **189**, 85–86 (2014).

11. Ye, W. F., Ma, Z. Y. & Ha, X. Z. Spatial–temporal patterns of $PM_{2.5}$ concentrations for 338 Chinese cities. *Sci. Total Environ.* **631–632**, 524–533 (2018).

12. Ma, Z. *et al.* Satellite-based spatiotemporal trends in $PM_{2.5}$ concentrations: China, 2004–2013. *Environ. Health Perspect.* **124**, 184–192 (2016).

13. Wang, S. J., Zhou, C. S. & Wang, Z. B. The characteristics and drivers of fine particulate matter ($PM_{2.5}$) distribution in China. *J. Clean. Prod.* **142**, 1800–1809 (2017).

14. Zhou, C. J., Chen, J. & Wang, S. J. Examining the effects of socioeconomic development on fine particulate matter ($PM_{2.5}$) in China's cities using spatial regression and the geographical GIS detector technique. *Sci. Total Environ.* **619–620**, 436–445 (2018).

15. Tobler, W. R. A computer movie simulation urban growth in the Detroit region. *Econ. Geogr.* **46**, 234–240 (1970).

16. Wang, Z. & Fang, C. Spatial–temporal characteristics and determinants of $PM_{2.5}$ in the Bohai Rim urban agglomeration. *Chemosphere* **148**, 148–162 (2016).

17. Zhang, Y. & Cao, F. Fine particulate matter ($PM_{2.5}$) in China at a city level. *Sci. Rep.* **5**, 14884 (2015).

18. Huang, R. J. *et al.* High secondary aerosol contribution to particulate pollution during haze events in China. *Nature* **514**, 218–222 (2014).

19. Liu, H. *et al.* The effect of natural and anthropogenic factors on haze pollution in Chinese cities: A spatial econometrics approach. *J. Clean. Prod.* **165**, 323–333 (2017).

20. Alvarez, H. B., Echeverria, R. S., Alvarez, P. S. & Krupa, S. Air quality standards for particulate matter (pm) at high altitude cities. *Environ. Pollut.* **173**, 255–256 (2013).

21. Wang, S. J., Liu, X. P. & Yang, X. Spatial variations of $PM_{2.5}$ in Chinese cities for the joint impacts of human activities and natural conditions: A global and local regression perspective. *J. Clean. Prod.* **203**, 143–152 (2018).

22. Guan, D. *et al.* The socioeconomic drivers of China's primary $PM_{2.5}$ emissions. *Environ. Res. Lett.* **9**, 024010 (2014).

23. Xu, X., Gonzalez, J. E., Shen, S., Miao, S. & Dou, J. Impacts of urbanization and air pollution on building energy demands–Beijing case study. *Appl. Energy* **225**, 98–109 (2018).

24. Xu, S. C., Miao, Y. M., Gao, C. & Long, R. Y. Regional differences in impacts of economic growth and urbanization on air pollutants in China based on provincial panel estimation. *J. Clean. Prod.* **208**, 340–352 (2019).

25. Lu, D. D. Conservation of the Yangtza River and sustainable development of the Yangtze River Economic Belt: An understanding of General Secretary Xi Jinping's important instructions and suggestions for their implementation. *Acta Geogr. Sin.* **73**(10), 1829–1836 (2018).

26. Zhong, Y., Lin, A., He, L. J., Zhou, Z. G. & Yuan, M. X. Spatiotemporal dynamics and driving forces of urban land–use expansion: A case study of the Yangtze River Economic Belt, China. *Rem. Sens.* **12**(2), 287 (2020).

27. Yang, M. & Wang, Y. Spatial–temporal characteristics of $PM_{2.5}$ and its influencing factors in the Yangtze River economic Belt. *China Popul. Resour. Environ.* **27**, 91–100 (2017).

28. Zhu, W. W., Wang, M. C. & Zhang, B. B. The effects of urbanization on $PM_{2.5}$ concentrations in China's Yangtze River Economic Belt: New evidence from spatial econometric analysis. *J. Clean. Prod.* **239**, 1–10 (2019).

29. Feng, Y. Y. *et al.* Defending blue sky in China: Effectiveness of the "Air Pollution Prevention and Control Action Plan" on air quality improvements from 2013 to 2017. *J. Environ. Manag.* **252**, 1963 (2019).

30. Zhou, L. *et al.* Spatio–temporal evolution and the influencing factors of $PM_{2.5}$ in China between 2000 and 2015. *J. Geogr. Sci.* **29**(2), 253–270 (2019).

31. Guo, J., Zhu, D., Wu, X. & Yan, Y. Study on environment performance evaluation and regional differences of strictly–environmental–monitored cities in China. *Sustainability* **9**, 2094 (2017).

32. Cheng, Z., Li, L. & Liu, J. Identifying the spatial effects and driving factors of urban $PM_{2.5}$ pollution in China. *Ecol. Indicat.* **82**, 61–75 (2017).

33. Wu, W. Q., Zhang, M. & Ding, Y. T. Exploring the effect of economic and environment factors on $PM_{2.5}$ concentration: A case study of the Beijing–Tianjin–Hebei region. *J. Environ. Manag.* **268**, 1103 (2020).

34. Van, D. A., Martin, R. V. & Brauer, M. Global estimates of fine particulate matter using a combined geophysical statistical method with information from satellites, models, and monitors. *Environ. Sci. Technol.* **50**(7), 3762–3772 (2016).

35. Ma, J. H., Cao, Y., Xu, J. M., Qu, Y. H. & Yu, Z. Q. $PM_{2.5}$ concentration distribution patterns and influencing meteorological factors in the central and eastern China during 1980–2018. *J. Clean. Prod.* **311**, 127565 (2021).

36. Lim, C. H., Ryu, J., Choi, Y. Y., Jeon, S. W. & Lee, W. K. Understanding global $PM_{2.5}$ concentrations and their drivers in recent decades (1998–2016). *Environ. Pollut.* **144**, 106011 (2020).

37. Xie, Y. & Weng, Q. World energy consumption pattern as revealed by DMSP–OLS nighttime light imagery. *GISCI Remote Sens.* **53**(2), 265–282 (2016).

38. Moran, P. Notes on continuous stochastic phenomena. *Biometrika* **37**, 17–23 (1950).

39. Song, W. Z., Jia, H. F., Li, Z. L. & Tang, D. L. Using geographical semi–variogram method to quantify the difference between NO_2 and $PM_{2.5}$ spatial distribution characteristics in urban areas. *Sci. Total Environ.* **631–632**, 688–694 (2018).

40. Ye, L., Tan, W., Fang, L., Ji, L. & Deng, H. Spatial analysis of soil aggregate stability in a small catchment of the loess plateau, China: I. Spatial variability. *Soil Tillage Res.* **179**, 71–81 (2018).

41. Liu, X. M. *et al.* Application of geostatistics and GIS technique to characterize spatial variabilities of bioavailable micronutrients in paddy soils. *Environ. Geol.* **46**, 189–194 (2004).

42. Fotheringham, A. S. & Brunsdon, C. Local forms of spatial analysis. *Geogr. Anal.* **31**(4), 340–358 (1999).

43. Wang, J. F., Li, X. H. & Christakos, G. 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**(1), 107–127 (2010).

44. Wang, J. F. & Xu, C. D. Geodetector: principle and prospective. *Acta Geographica Sin.* **1**, 116–134 (2017).

Acknowledgements

Author contributions

All authors contributed to the study conception and design. S.X.: methodology, writing–original draft preparation; Y.Y.: conceptualization, reviewing, editing and project administration; X.L.: writing–original draft preparation, methodology, formal analysis; Q.L.: data analysis, investigation; Y.Z.: visualization, data analysis.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Y.Y.

Reprints and permissions information is available at www.nature.com/reprints.

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