

ASSESSING SPATIAL DISTRIBUTION AND IMPACT FACTORS OF FINE PARTICULATE MATTER (PM_{2.5}) AND NITROGEN DIOXIDE (NO₂) BY COMBINING HIGH-RESOLUTION GEOGRAPHICAL CENSUS DATA, AND METEOROLOGICAL DATA IN HEBEI PROVINCE, CHINA

YAO, S. Y.¹ – QIAO, B. S.¹ – YU, X. T.¹ – LIU, C.^{2*} – SUN, W.³ – ZHANG, F.⁴ – CAO, Y. M.⁵ – LI, Z. Y.⁶ – PENG, Z. R.^{7,8} – WANG, Y.³

¹*School of Traffic and Transportation, Shijiazhuang Tiedao University, Shijiazhuang 050043, China*

²*College of Architecture and Urban Planning, Tongji University, Siping Rd. 1239, Shanghai 200082, China*

³*The Third Institute of Surveying and Mapping of Hebei Province, Shijiazhuang 050032, China*

⁴*Environmental Monitoring Center of Hebei Province, Shijiazhuang 050030, China*

⁵*Hebei Zhengrun Environmental Technology Co., Ltd, Shijiazhuang 050051, China*

⁶*College of Resources and Environmental Sciences, Shijiazhuang College, Shijiazhuang 050035, China*

⁷*School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiaotong University, Shanghai 200240, China*

⁸*Department of Urban and Regional Planning, University of Florida, P. O. Box 115706, Gainesville, FL 32611-5706, USA*

**Corresponding author*

e-mail: liuchao1020@gmail.com

(Received 29th Jan 2019; accepted 10th Apr 2019)

Abstract. This paper was the first to employ land use regression (LUR) with high-resolution geographical census data for Hebei, one of the most severely polluted regions in China, to evaluate its spatial distribution characteristics of PM_{2.5} and NO₂ concentrations and identify influencing factors. To develop the LUR model, PM_{2.5} and NO₂ concentrations recorded at 53 sites in Hebei were selected as dependent variables. Independent variables include buffer-related and location-based factors. At first, 169 independent variables were chosen in total. Then pre-processing of bivariate correlation was performed to prevent multicollinearity. Lastly, step-wise regression was processed to identify the impacting factors. Different to other cities which have been studied like Shanghai or Beijing, we find that the results showed that PM_{2.5} and NO₂ concentrations were positively correlated with the industrial pollution sources in a buffer area. NO₂ concentrations displayed significant negative correlations with forestland within the distance of 1 km and from the coastline. This study showed that the introduction of high-resolution geographical data into the LUR model significantly improved the fitting. More importantly, our study identified industries within a 9 km-buffer as important influencing factors in Hebei and was also consistent with empirical observations. It provided data on effective buffers to support future policy-making and designations of residential areas.

Keywords: *land use regression, fine particulate matter, air pollution*

Introduction

Fine particulate matter with a diameter of less than 2.5 µm (PM_{2.5}) and nitrogen dioxide (NO₂) have been recognized as causes of adverse health effects (Duki et al., 2003; Roswall et

al., 2017; Chen et al., 2018; Yin et al., 2017). Empirical studies have shown a significant increase in the proportion of related diseases in regions affected by air pollution (Shekarrizfard et al., 2016; Su et al., 2008). Research by the World Health Organization (WHO) has shown that outdoor pollution caused 3.0 million extra deaths in 2012 (WHO, 2016). After three decades of rapid economic development, China has become the world's second largest economy (Morrison et al., 2015). At the same time, China is suffering from serious air pollution (Chambers et al., 2015; Chan et al., 2008; He et al., 2016; Hu et al., 2015; Wang et al., 2013; Fengwen et al., 2015; Song et al., 2017). There are emerging studies from various perspectives under pressures of urban and social development (Liu et al., 2017), but interdisciplinary research linking urban land use pattern, air pollution and meteorology at the regional level in China is lacking.

Many studies in Europe and North America have shown that land use patterns directly and indirectly influence air pollutant concentrations, and the optimization of land use patterns is feasible to alleviate air pollution problems (Briggs et al., 2000). In order to optimize land use, it is necessary to understand the impacts of its various influencing factors. There are numerous approaches to the quantitative assessment of the influence of land use on air pollution, such as diffusion models (Ciocănea et al., 2013) (for example, land use regression [LUR]), numerical simulations of the atmosphere (Chen et al., 2009), inversion of MODIS remote sensing images (Wu et al., 2016; Yao et al., 2017), and artificial neural network models (Jerret et al., 2005; Lu et al., 2003), etc. Jedynska et al. (2015) and Liang et al. (2016) respectively applied the LUR model on relationship of air pollutants, land use variables in Oslo, The Netherlands, Munich/Augsburg, and Catalonia of Europe and Houston of USA.

Among these tools, LUR has been proved to be feasible and appropriate of regions of China and other areas worldwide (Hoek et al., 2008; Ross et al., 2007). Therefore, LUR was chosen as the method to study spatial influencing factors in air pollution in this paper. As of the study areas, Hebei Province was selected because it is one of the most polluted area in China and worldwide, and in-depth research is needed to understand its pollution characteristics (He et al., 2016; Zhang et al., 2016).

LUR models are important methods for studying the relation between the spatial distribution of air pollutants and elements of geographical space. Usually, an LUR model is built by constructing a regression equation to analyze factors affecting air quality, in which data for air pollutant concentrations serve as the dependent variable while data for socioeconomic, land use, and meteorological factors, etc., serve as the independent variables.

After the Chinese government established national networks for monitoring air pollutants in 2013, relevant studies using LUR models gradually increased in number. Some researchers employed an LUR model to study land use and air pollution levels in Shanghai, one of the largest megacities in China (Liu et al., 2016; Meng et al., 2015). Although, Meng et al. (2015) was the first to combine aerosol optical depth (AOD), meteorological information and the land use regression (LUR) model to predict ground PM₁₀ levels on high spatiotemporal resolution. However, in their studies, the resolution of land use data is not high.

Previous studies showed that the Beijing–Tianjin–Hebei region is one of the most seriously areas affected by the pollutant haze in China (Wang et al., 2015b; Gang et al., 2014; He et al., 2018). Beijing and Tianjin are two of the most important cities in China. Beijing is the capital of China, and Tianjin is an important municipality directly under the jurisdiction of the central government. Both cities have a high degree of urbanization and high levels of financial industry. Although the level of urbanization in Hebei Province is not high, its industry is dominated by steel and coal. The sources of pollution are therefore different, and the spatial patterns and influencing factors of pollution are also potentially different.

Some researchers have pointed out that the impact of motor vehicle exhaust emissions on the atmosphere is the most significant pollution source in Beijing and Tianjin, whereas in Hebei Province industry is the most notable contributor (Yang et al., 2018; Wang et al., 2017).

By the end of 2012, the total coal consumption of Hebei Province reached 271 million tons, which was much higher than the combined consumption of Beijing and Tianjin. Among 74 major cities in China, pollution by PM_{2.5} in seven Hebei cities is the most serious (Wang et al., 2013).

An LUR model has been used to analyze the spatial distribution and influencing factors of air pollutants in Tianjin and Beijing and good results have been achieved (Chen et al., 2010; Wu et al., 2014). However, an LUR model has not been employed in Hebei Province alone in any recent researches. Because of the heavy industry-based industrial structure and complex geographical and meteorological environment, we focused exclusively on Hebei and used high-resolution land use data to employ LUR for the identification of influencing factors.

The purpose of this study is to conduct Land use Regression model to analyze the relationship between spatial characteristic and air pollutants in Hebei province. In comparison with previous studies, this paper improved both data fineness and model integrity in cooperation with multiple local and international institutes. Firstly, high-resolution (1 × 1 m) land use data from national geographical census data were used. Secondly, we set 20 buffer distances to refine the influencing factors that significantly influenced the concentrations of air pollutants. Finally, this study introduced meteorological data as a variable for a more comprehensive analysis on air pollution problems.

Data and methodology

Hebei Province was chosen as the study area because it is the most polluted region of China (Lu et al., 2015). In contrast to Beijing and Tianjin, Hebei has a more complex and diverse geographical environment and a much larger area. Furthermore, Hebei is relatively less urbanized and has more heavy industries. Hebei's unique characteristics need customized research designs, and it is necessary to conduct separate studies in Hebei alone. When we developed our LUR model, the PM_{2.5} and NO₂ concentrations at 53 air quality monitoring sites were selected as separate dependent variables, and various influencing indicators (e.g., meteorological factors, land use types, road area, population density, monitoring enterprises, pollution sources, elevation, etc.) were selected as independent variables. The LUR model was analyzed with a statistical package (SPSS) and geographic information system (GIS) software (ArcGIS 10.1). A flow chart of the study is shown in *Figure 1*.

Research area

Hebei Province is a coastal province adjacent to two megacities in northern China, namely, Beijing and Tianjin. The Beijing–Tianjin–Hebei region is also one of China's largest urban agglomerations. Hebei consists of 11 cities, and its land area is approximately 188,500 km². The Yanshan and Taihang Mountains lie to its north and west respectively and the Bohai Sea lies to its northeast Hebei, as illustrated in *Figures 2* and *3*. At the end of 2014, the gross domestic product (GDP) of Hebei reached 2942.12 billion renminbi (RMB), its population was 73.83 million, and its average population density was 392 inhabitants/km² (National Bureau of Statistics, 2015).

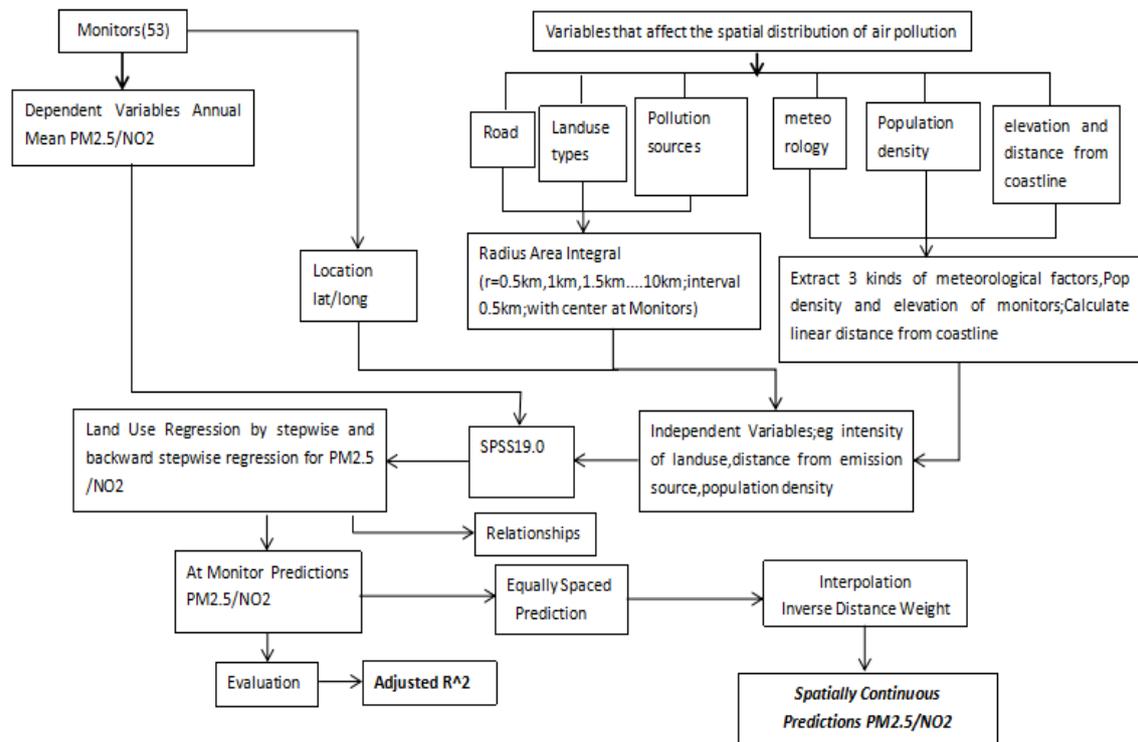


Figure 1. Flow chart of our study

The annual temperature at meteorological monitoring sites in Hebei Province in 2014 ranged is 13°C (Climate Bulletin of 2014 in Hebei Province), and the average wind speed is 2.5 m/s, calculated from data for 20 national meteorological monitoring stations in Hebei in 2014. The average wind speed in southern and western areas is lower than that in northern and eastern areas because the Taihang and Yanshan mountains to the west and north of Hebei block the prevailing northerly winds in winter.

LUR model setting

Dependent variables: PM_{2.5} and NO₂

Hebei Province started to report PM_{2.5} concentrations in 2012. In this study, we extracted hourly monitoring data for PM_{2.5}/NO₂ concentrations from the official website of China Environmental Monitoring Station (<http://www.cnemc.cn>) during the period from March 22, 2014 to March 22, 2015. In total, one-year observations from 53 national air quality monitoring sites in 11 cities of Hebei were collected, and the average annual values of PM_{2.5} and NO₂ concentrations were used as dependent variables.

Independent variables

Independent variables were divided into buffer-related and location-based variables, of which the latter were specific to the 53 air quality monitoring sites. The buffer-related variables are comprised of six land use types, road areas, and the numbers of pollution-emitting industries in different buffer areas around each monitoring site, which were calculated for 20 different buffer distances (from 0.5 km to 10 km at

intervals of 500 m) from each monitoring site. Furthermore, the location-based variables included the longitude, latitude, distance from the monitoring site to the coastline, population density, and elevation, as well as meteorological elements, including air pressure, wind speed, temperature, precipitation, and relative humidity. In the LUR model, 169 independent variables were used, and each independent variable is described as follows.

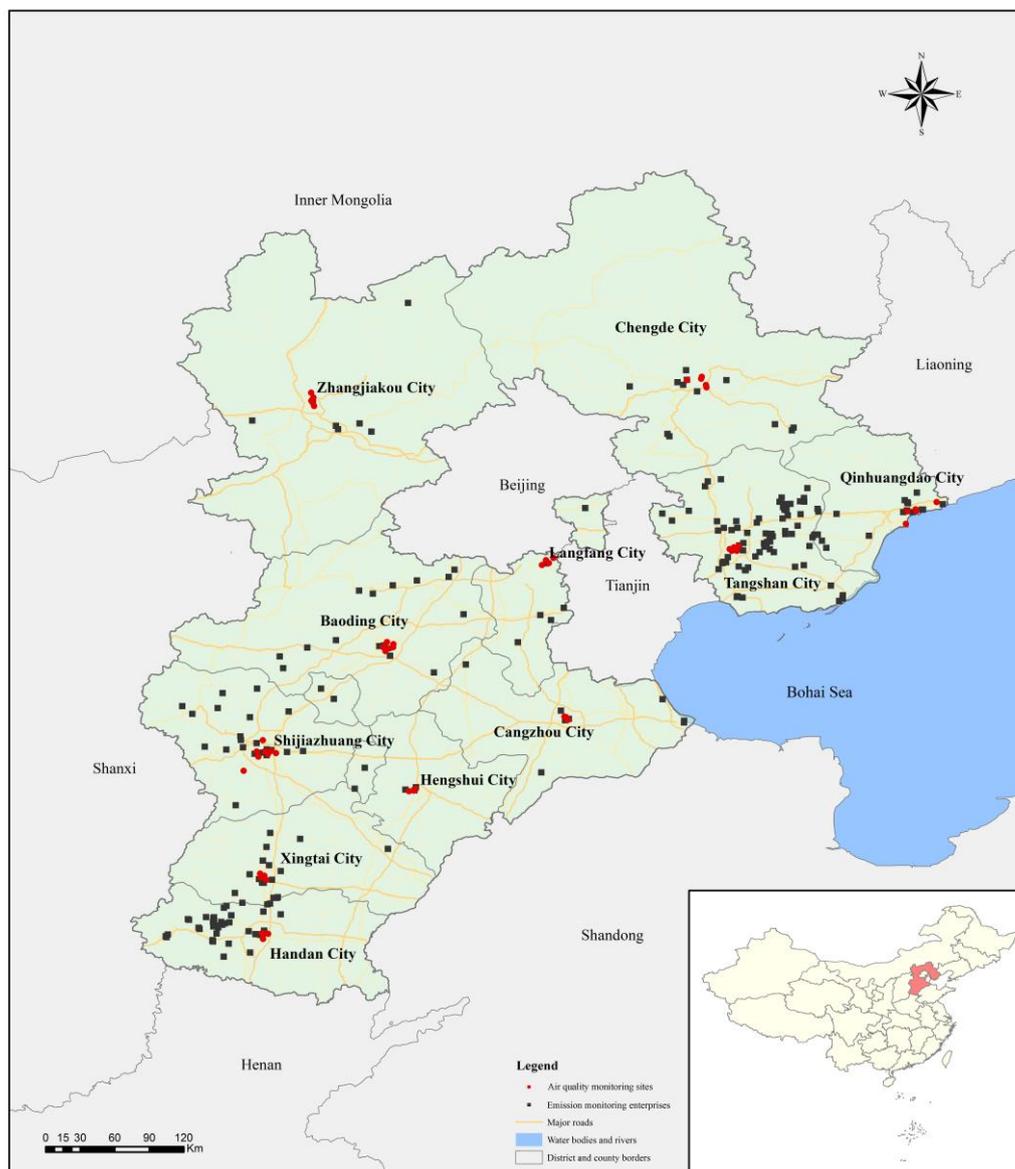


Figure 2. Study area and monitor distributions

Land use

The six land use types adopted are comprised of building land (including workplaces, residences, educational institutions, and buildings used for entertainment), forest land, farmland, bodies of water (rivers and lakes), greenhouses, and garden land. (A greenhouse is defined as a heating facility used to cultivate plants.) The areas occupied

by each land use in each buffer zone were calculated separately. The land use data were derived from national geographical census data for Hebei Province with high resolution (1 × 1 m), which were collected by the Hebei Bureau of Geoinformation (<http://www.hebsm.gov.cn/>) in 2015. The variables were denoted as the land use type plus buffer distance as follows: build_buffer distance, fore_buffer distance, farm_buffer distance, wat_buffer distance, ghl_buffer distance, and gar_buffer distance. The explanation of abbreviation is shown in *Table A3* in the *Appendix*.



Figure 3. Study area and ridgelines distributions

Road areas

The areas of all road types in the different buffer zones in km² were calculated. The road area data were derived from high-resolution national geographical census data for Hebei Province from the Hebei Bureau of Geoinformation (<http://www.hebsm.gov.cn/>). The road types included national roads, provincial roads, county roads in rural areas, expressways, arterial streets, sub-arterial streets, and local streets in urban areas. We did not separate these road types in detail but uniformly calculated the areas of all types of roads in different buffer zones, which were denoted as rdar_buffer in this paper.

Population density

Population density data in inhabitants/km² were extracted from the database of China's population distribution provided by the Data Center for Resources and Environmental Sciences, Chinese Academy of Sciences. We used the GIS to extract the raster values of the 53 air quality monitoring sites in Hebei Province as population density variables from the population distribution dataset for the Chinese population grid (1 × 1 km), which was calculated by Fu et al. (2013). In comparison with the census data, the above-mentioned data have higher resolution and are better for revealing spatial differences in population density, because the relative error is kept under 10% (Fu et al., 2010).

Distance to the coastline

The straight-line distance (m) between each air quality monitoring site and the nearest coastline was measured on a map of Hebei Province with the GIS. The measurement of the distance to the coastline served to determine the effects of the deposition of pollutants in the sea and the dilution of pollutants by sea breezes.

Geographical features of air quality monitoring sites

The geographical coordinates of each air quality monitoring site (including its longitude and latitude) were recorded, which were obtained from the Environmental Monitoring Center of Hebei Province. Elevations were extracted from the data of a digital elevation model of Hebei with a worldwide resolution of 30 m.

Industrial sources of air pollution

This study considered 241 enterprises that carried out national emission monitoring in Hebei Province in 2015, as listed on the website of the Ministry of Environmental Protection of China, as shown in *Figure 2* (<http://www.zhb.gov.cn/>). These included waste gas, coal, and thermal power plants, etc. We counted the number of these enterprises within each buffer zone, which was included as an independent variable in the model.

Meteorological data

We extracted meteorological data for 20 national meteorological monitoring stations in Hebei Province from the official website of the China Meteorological Administration. The mean values of five variables (air pressure, temperature, precipitation, wind speed, and relative humidity) in one year were calculated. By means of the kriging interpolation algorithm in ArcGIS 10.1, the values of the five classes of meteorological data for each air pollution monitoring site were taken as independent variables.

Statistical analysis

Pre-processing was performed to solve problems arising from multicollinearity between independent variables. Firstly, a bivariate (Pearson) correlation test was used to calculate the absolute strength of the correlation between all variables and the measured PM_{2.5}/NO₂ concentrations. Secondly, for land use factors, from two independent variables between which the correlation coefficient was greater than 0.6, only the

variable with a stronger correlation with the dependent variables was employed in the LUR model. In addition, the longitude, latitude, wind speed, population density, precipitation, pressure, temperature, and altitude were added directly to the LUR model.

After the pre-processing step, stepwise backward multiple regression methods were employed to test the remaining variables for their correlations with PM_{2.5} and NO₂ concentrations. The adjusted R² and residuals were used to evaluate the goodness of model fitting. The global Moran's index and Durbin-Watson index were employed to test the assumption that the regression residuals were spatially independent. In the application of GIS model, Moran's (Anselin, 1983) and Durbin-Watson (Mcauliffe, 2015) indexes are generally used for spatial autocorrelation test of data.

Results and discussion

Descriptive characteristics of PM_{2.5} and NO₂ concentrations

The average PM_{2.5} and NO₂ concentrations at the 53 air quality monitoring sites in Hebei Province were $83.64 \pm 29.45 \mu\text{g}/\text{m}^3$ and $47.18 \pm 12.10 \mu\text{g}/\text{m}^3$, respectively. *Figure 4a* shows an inverse distance weighting (IDW) interpolation map derived from the annual average values of PM_{2.5} concentrations. *Figure 4b* shows an IDW interpolation map derived from the annual average values of NO₂ concentrations. The maps indicate that the concentrations in the southwest were higher than those in other areas. In addition, there were significant differences in concentration between the north and south of Hebei along the Yanshan mountains. Areas under severe air pollution were mainly concentrated in the south of Hebei for both PM_{2.5} and NO₂.

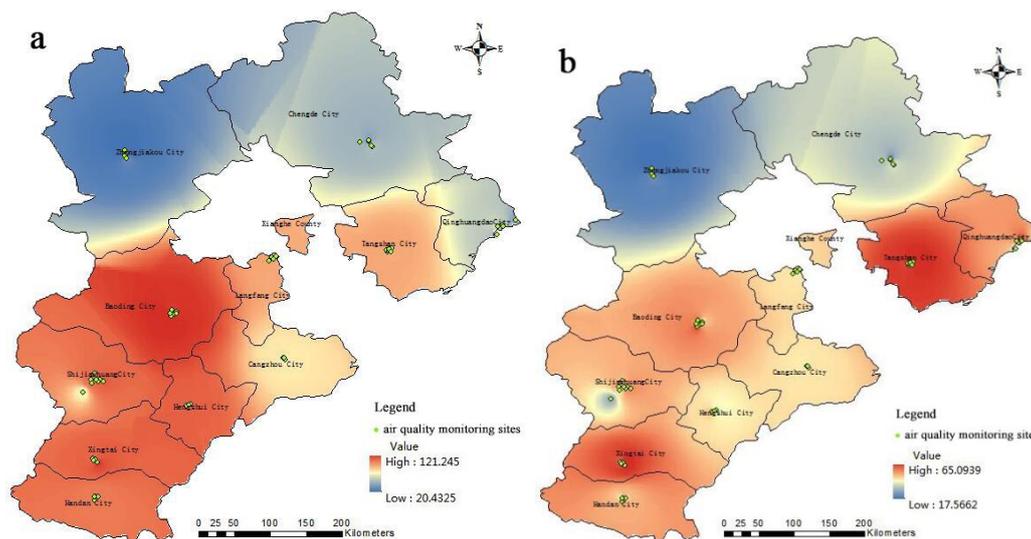


Figure 4. *a* Interpolation map of PM_{2.5} concentrations ($\mu\text{g}/\text{m}^3$) based on observational data. *b* Interpolation map of NO₂ concentrations ($\mu\text{g}/\text{m}^3$) based on observational data

Results of pre-processing for multicollinearity

There were initially a total of 169 indicator variables in this study; after filtering for multicollinearity between the influencing factors with reference to the modelling method proposed by Wu et al. (2014), 20 influencing factors remained for PM_{2.5}

concentrations (*Table 1*) and 19 influencing factors remained for NO₂ concentrations (*Table 1*), which were added to the LUR model at the outset.

Table 1. Screening of optimal correlation factors for bivariate correlation analysis with PM_{2.5}/NO₂ concentration

PM _{2.5}		NO ₂	
Influencing factors	Correlation coefficient	Influencing factors	Correlation coefficient
Longitude	-0.39**	Distance to the coastline	-0.28*
Latitude	-0.75**	Latitude	-0.44**
Population density	0.52**	Population density	0.52**
Elevation	-0.73**	Elevation	-0.76**
Wind speed	-0.51**	Wind speed	-0.46**
Temperature	0.77**	Temperature	0.41**
Air pressure	0.75**	Relative humidity	0.61**
Precipitation	-0.30*	Air pressure	0.72**
Farmland_0.5 km	0.31*	Road area_0.5 km	0.34*
Building land_0.5 km	0.52**	Forest land_1 km	-0.73**
Farmland_3 km	0.28*	Road area_1 km	0.43**
Farmland_3.5 km	0.31*	Emission monitoring enterprises_3.5 km	0.37**
Water_3.5 km	-0.37**	Emission monitoring enterprises_4 km	0.37**
Emission monitoring enterprises_3.5 km	0.34*	Garden land_5 km	-0.42**
Emission monitoring enterprises_4 km	0.33*	Building land_5 km	0.63**
Forest land_6 km	-0.70**	Emission monitoring enterprises_8.5 km	0.76**
Building land_9 km	0.78**	Water_10 km	0.36**
Emission monitoring enterprises_7.5 km	0.48**	Farmland_10 km	0.26*
Farmland_10 km	0.58**	Road area_10 km	0.66**
Road area_10 km	0.62**	-	-

**Significant correlation at level of $p < 0.01$. *Significant correlation at level of $p < 0.05$

LUR model results

PM_{2.5} concentration modeling results

The adjusted R² in the regression model was 0.96 and the F-statistic was 157.71 ($p < 0.01$), which indicates that the LUR model had good fitness (*Table A1* in the *Appendix*). The value of the Durbin–Watson test statistic was 1.954, which indicates that there was no autocorrelation in the residual sequence. Eight independent variables were finally entered into the LUR model, namely, air pressure, precipitation, temperature, wind speed, longitude, elevation, farmland_500 m, and pollution-emitting industries within the 7.5 km buffer zone. The standardized regression coefficients for air pressure and precipitation were 0.94 and 0.23, respectively, which indicates that they

were positively correlated with PM_{2.5} concentrations. In contrast, the standardized regression coefficients for temperature and wind speed were -0.62 and -0.15, respectively, which indicates that they were negatively correlated with PM_{2.5} concentrations. The results also show that the area of farmland within the 500 m buffer zone and emission monitoring enterprises within the 7.5 km buffer zone (Emission monitoring enterprises_7.5 km) were included as significant predictor variables, with standardized regression coefficients of 0.07 and 0.10, respectively. This indicates that the emission source of waste gas enterprises is the main contributor to the concentration of PM_{2.5} pollutants, and the wind speed is the main factor to alleviate PM_{2.5} pollution, which is similar to the conclusions of relevant study (Guan et al., 2017; Liu et al., 2018). Longitude and elevation were also influencing factors negatively correlated with PM_{2.5} concentrations, with standardized regression coefficients of -1.14 and -0.29.

The final model explains 96% of the variation in the measured PM_{2.5} concentrations. There was no apparent bias in the model, as shown in *Figure 5a* and *b*. The results for spatial autocorrelation are shown in *Figure 5c*. The z-score was 0.74, which is less than the critical value of 1.96, which means the probability that the residuals resulted from random factors is greater than 0.95. The value of Moran's index of 0.056 further indicates that there was no autocorrelation and justifies the assumption that errors were spatially independent.

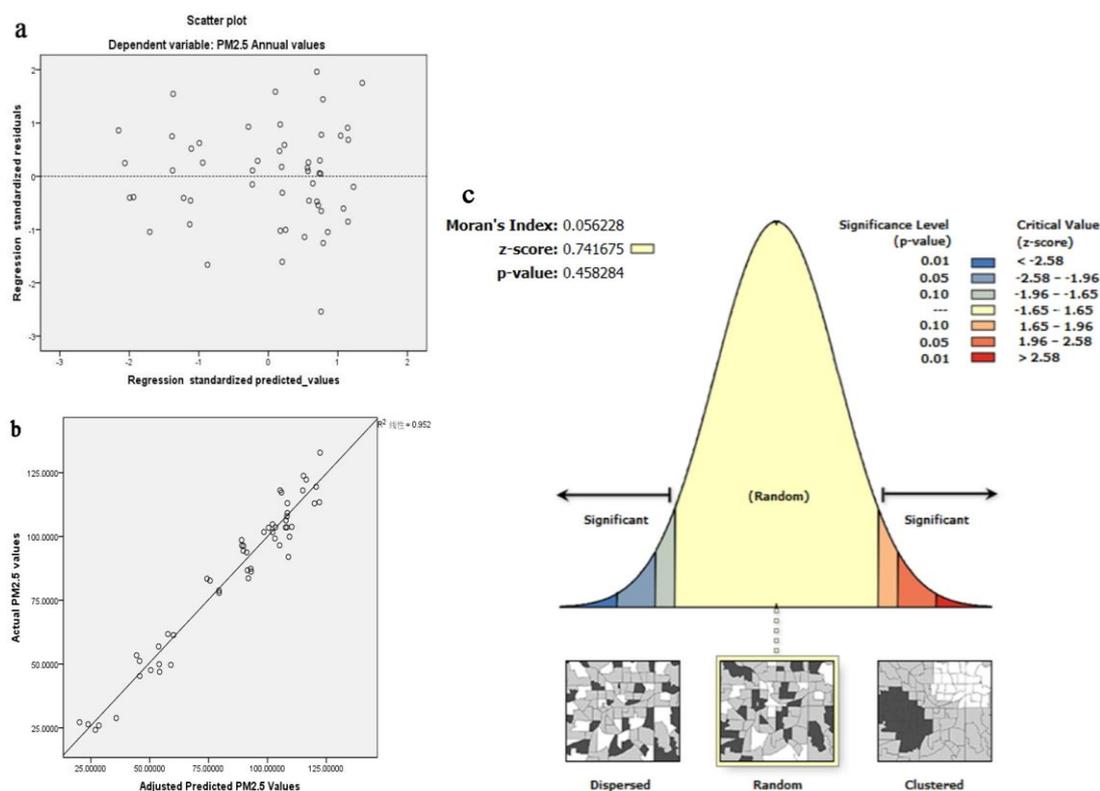


Figure 5. *a* Plot of residuals versus predicted PM_{2.5} values. *b* Plot of monitored versus predicted PM_{2.5} values (μg/m³). *c* Results of spatial autocorrelation test for LUR residuals

An interpolation map of measured values of PM_{2.5} concentrations is shown in *Figure 4a*, and an interpolation map of LUR predictions of PM_{2.5} concentrations is shown in *Figure 6a*. From these two figures, we can conclude that there existed

significant heterogeneity in the spatial distribution of PM_{2.5} concentrations in Hebei Province. The variables of longitude and elevation were negatively correlated with PM_{2.5} concentrations. Correspondingly, the LUR model results show that western areas were more heavily polluted than eastern areas and that southern areas were more heavily polluted than northern areas. The regression coefficients for the number of emission monitoring enterprises within the 7.5 km buffer zone and farmland within the 500 m buffer zone show that these variables were positively correlated with PM_{2.5} concentrations.

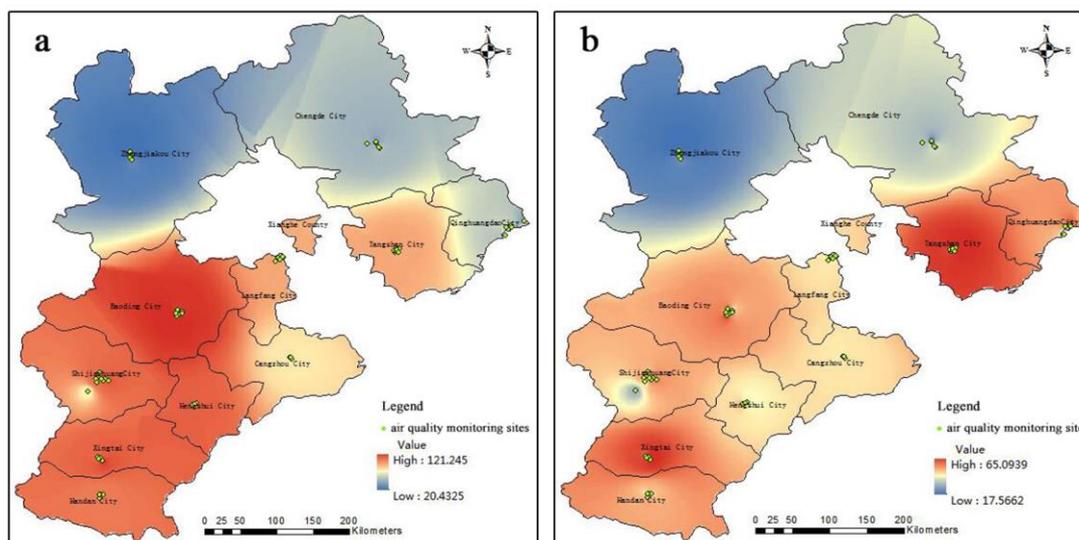


Figure 6. *a* LUR predictions of PM_{2.5} concentrations ($\mu\text{g}/\text{m}^3$). *b* LUR predictions of NO₂ concentrations ($\mu\text{g}/\text{m}^3$)

This is due to the following reasons. Firstly, industrial emissions of air pollutants make a great contribution to PM_{2.5} concentrations (Rohde et al., 2015). Secondly, the destruction of the soil structure and surface vegetation cover of farmland by human farming processes (Bogman et al., 2005; Goossens, 2004; Pattey et al., 2012) and the burning of biomass lead to the increases in PM_{2.5} concentrations (Cheng et al., 2014). Therefore, we conclude that PM_{2.5} concentrations are influenced by meteorological factors, certain land uses, and heavy industries.

Spatial distribution of NO₂ concentrations and influencing factors

NO₂ is the main air pollutant indicator that originates from transportation. In this study, we also developed an LUR model for NO₂ for comparison with the features of PM_{2.5} pollution. The adjusted R² in the regression model was 0.81 and the F-statistic was 25.31 ($p < 0.01$), which indicates that the LUR model was also highly predictive for NO₂ concentrations (*Table A2* in the *Appendix*). The value of the Durbin–Watson test statistic was 2.08, which is very close to 2, indicating that there was no autocorrelation in the residual sequence. Finally, nine independent variables were entered into the model, namely, distance to the coastline, latitude, forest land within the 1 km buffer zone, emission monitoring enterprises within the 8.5 km buffer zone, area of water within the 10 km buffer zone, wind speed, temperature, air pressure, and relative humidity.

An interpolation map of measured values of NO₂ concentrations is shown in *Figure 4b*, and an interpolation map of LUR predictions of NO₂ concentrations is shown in *Figure 5b*.

The main similarity was that for both PM_{2.5} and NO₂ the number of emission monitoring enterprises and air pressure had a positive influence on pollution levels. There were also differences between the two models. In general, land uses were more influential in the NO₂ model than in the PM_{2.5} model, whereas meteorological factors were more significant in the PM_{2.5} model than in the NO₂ model. A possible reason is that PM_{2.5} remains suspended for a longer time and could be transported over longer distances with the help of meteorological conditions in comparison with NO₂ (Liu et al., 2016). NO₂ is considered to be a local pollutant and is thus more influenced by local land uses (Liu et al., 2016). The results of the NO₂ model differed from those of the PM_{2.5} model. Unlike the LUR model results for PM_{2.5}, the distance to the coastline and latitude affected NO₂ pollution but were not entered into the PM_{2.5} model. Secondly, wind speed, temperature, and relative humidity were positively correlated with NO₂ concentrations but negatively correlated with PM_{2.5} concentrations. Thirdly, forest land was helpful for NO₂ pollution reduction but did not exhibit significant effects in decreasing PM_{2.5} concentrations. Besides, the adjusted R² for the NO₂ regression was lower than that for the PM_{2.5} model.

There is a negative correlation between wind speed and NO₂ concentration in *Table 1*. This suggests that without being influenced by anything else, the higher the wind speed, the more favorable is the diffusion of NO₂ concentration. But in the LUR model, the influence factor coefficient of wind speed is negative, this may be because the influence of wind speed on NO₂ concentration may be interfered by other factors.

Discussion

A comparison was carried out with LUR results for similar regions. Owing to the lack of specialized research in Hebei Province using LUR, we chose previous studies of the Beijing–Tianjin–Hebei region as a whole to compare our results. Each study had a specific combination of influencing variables for the prediction of PM_{2.5} concentrations, as shown in *Table 2*.

The results of the comparison showed that this study displayed the best fitness and the highest adjusted R². The primary reason was that Hebei Province and the cities of Beijing and Tianjin have distinct economic structures. Beijing and Tianjin are two of the most highly developed cities of China, with high levels of urbanization and much fewer polluting industries. Therefore, the fitness of the LUR model was lower when Beijing–Tianjin–Hebei as a whole was considered as the study region and the spatial characteristics were assumed to be the same in the model.

In all three sets of results, elevation was a common variable and was negatively correlated with PM_{2.5} concentrations. Emission monitoring enterprises played an important role both in this study and in the model used by Gang et al. (2016), which had buffer zones of 9 km and 10 km, respectively. This consistency indicates that the influence of pollution-emitting industries on PM_{2.5} concentrations was significant.

Our study was the first to consider the impacts of meteorological factors on pollutant levels. Our model showed that meteorological factors were even more significant for predicting PM_{2.5} concentrations, and hence a comprehensive LUR including meteorological factors is valuable and more convincing.

Table 2. Comparison of LUR results for PM_{2.5} concentrations for the Beijing–Tianjin–Hebei region

Authors	Study area	Buffers	Adjusted R ²	Influencing variables in the model
Our study	Hebei Province	0–10 km total, 20 different buffers, interval 500 m	0.96	Pressure, longitude, precipitation, temperature, elevation, wind speed, emission monitoring enterprises_9 km, farmland_0.5 km
Xu et al. (2016)	Beijing–Tianjin–Hebei region	0–10 km total, 22 different buffers	0.81	Temperature, elevation, emission monitoring enterprises_10 km, water_10 km
Jian-Sheng et al. (2017)	Beijing–Tianjin–Hebei region	100, 200, 300, 500, 1000, 2000, 3000, 5000 m	0.81	Elevation, water_2 km, relative humidity

Conclusions

Our LUR models indicate that the PM_{2.5} and NO₂ concentrations in Hebei Province had significant spatial heterogeneity, and that the PM_{2.5} and NO₂ concentrations in southern cities were higher than in northern cities.

In contrast to relevant research studies in Europe, America, and other areas in China, the distribution of PM_{2.5} and NO₂ concentrations in Hebei Province is affected by pollution-emitting industries, whereas the road network is not the most significant variable. This is because Hebei is a developing region, whereas other areas of the Beijing–Tianjin–Hebei region are already developed. Specifically, Beijing’s per capita GDP reached 99,995 RMB and the per capita GDP of Tianjin reached 105,231 RMB, whereas the per capita GDP of Hebei was only 39,984 RMB (National Bureau of Statistics, 2015).

In general, the model results for PM_{2.5} concentrations demonstrate a gradually decreasing tendency from west to east, and the pollutant concentration is lower near the coastline than inland, which can be attributed to the influence of the onshore monsoon. The effect of dilution by sea breezes is consistent with the conclusions of a study performed in Shanghai (Liu et al., 2016).

The results show that the rational planning of land uses, such as the designation of land as industrial land or as green space, plays an important role in mitigating air pollution and can effectively reduce PM_{2.5} and NO₂ concentrations.

Our study was the first to use high-resolution geographical data to study Hebei Province as the sole study area. Furthermore, our models made significant improvements with the use of meteorological factors, which were usually not incorporated into traditional LUR models. The results indicate the importance of meteorological factors. In addition, we identified the most distinctive characteristic of the spatial patterns of pollution in Hebei in comparison with developed cities, namely, the effective buffers around polluting industry. Our findings can be utilized for the planning of residential land and the design of buffers around heavily polluting industry and can help to reveal the land use factors with the greatest influences on PM_{2.5} and NO₂ concentrations.

Some limitations of the LUR models used in this study should be recognized. Firstly, the total number of national monitoring stations in Hebei Province used in this study was 53. Although this conforms to the number of monitoring sites (40–80)

recommended by Hoek et al. (2008), most monitoring sites were located in urban areas and there was a lack of rural monitoring sites. Secondly, our study was limited by the availability of refined variables, including land use types, pollutant emission sources, real-time meteorological data, and transportation emission data. This is a common constraint that also existed in previous studies.

Further studies should be performed with more monitoring site data and more dynamic variables. In addition, empirical observations are needed to support our findings, such as the impact of buffers around industry, meteorological impacts, etc. As the most severely polluted region in China, Hebei Province needs more comprehensive study from an interdisciplinary perspective to identify the relevant problems, reasons, and solutions.

Acknowledgements. This study was supported by the Hebei Province Natural Science Fund (Grant No. D2016210008). This study was supported by the National Natural Science Foundation of China (51308358), the Hebei Province Natural Science Fund (D2016210008), the Science and Technology Department of Hebei Province (18390324D), the Science and Technology Bureau of Shijiazhuang City (181130034A).

REFERENCES

- [1] Anselin, L. (1983): Spatial processes, models and applications. – *Economic Geography* 59(3): 322-325.
- [2] Bogman, P., Cornelis, W., Rollé, H., Gabriels, D. (2005): Prediction of TSP and PM10 emissions from agricultural operations in Flanders, Belgium. – Proceedings of the 14th International Conference “Transport and Air Pollution”, Graz, Austria, June.
- [3] Briggs, D. J., Collins, S., Elliott, P., Fischer, P., Kingham, S., Lebet, E., Pryn, K., Van Reeuwijk, H., Smallbone, K., Van Der Veen, A. (1997): Mapping urban air pollution using GIS: a regression-based approach. – *International Journal of Geographical Information Science* 11(7): 699-718.
- [4] Chambers, S. D., Wang, F., Williams, A. G., Deng, X., Zhang, H., Lonati, G., Crawford, J., Griffiths, A. D., Ianniello, A., Allegrini, I. (2015): Quantifying the influences of atmospheric stability on air pollution in Lanzhou, China, using a radon-based stability monitor. – *Atmospheric Environment* 107: 233-243.
- [5] Chan, C. K., Yao, X. (2008): Air pollution in mega cities in China. – *Atmospheric Environment* 42(1): 1-42.
- [6] Chen, H., Bai, S., Eisinger, D. S., Niemeier, D., Claggett, M. (2009): Predicting near-road PM2.5 concentrations: comparative assessment of CALINE4, CAL3QHC, and AERMOD. – *Transportation Research Record: Journal of the Transportation Research Board* 2123(1): 26-37.
- [7] Chen, J., Zhou, C., Wang, S., Hu, J. (2018): Identifying the socioeconomic determinants of population exposure to particulate matter (PM 2.5) in China using geographically weighted regression modeling. – *Environmental Pollution* 241: 494-503.
- [8] Chen, L., Bai, Z., Kong, S., Han, B., You, Y., Ding, X., Du, S., Liu, A. (2010): A land use regression for predicting NO2 and PM10 concentrations in different seasons in Tianjin region, China. – *Journal of Environmental Sciences* 22(9): 1364-1373.
- [9] Cheng, Z., Wang, S. X., Fu, X., Watson, J. G., Jiang, J. K., Fu, Q. Y., Chen, C. H., Xu, B. Y., Yu, J. S., Chow, J. C. (2014): Impact of biomass burning on haze pollution in the Yangtze River Delta, China: A case study of summer in 2011. – *Atmospheric Chemistry & Physics Discussions* 14(9): 4573-4585.

- [10] Ciocănea, A., Dragomirescu, A. (2013): Modular ventilation with twin air curtains for reducing dispersed pollution. – *Tunnelling & Underground Space Technology Incorporating Trenchless Technology Research* 37(6): 180-198.
- [11] Duki, M. I., Sudarmadi, S. S., Kawada, T., Tri, T. A. (2003): Effect of air pollution on respiratory health in Indonesia and its economic cost. – *Archives of Environmental Health an International Journal* 58(3): 135-143.
- [12] Fengwen, W., Tian, L., Jialiang, F., Huaiyu, F., Zhigang, G. (2015): Source apportionment of polycyclic aromatic hydrocarbons in PM_{2.5} using positive matrix factorization modeling in Shanghai, China. – *Environ Sci Process Impacts* 17(1): 197-205.
- [13] Gang, L., Jingying, F., Dong, J., Wensheng, H., Donglin, D., Yaohuan, H., Mingdong, Z. (2014): Spatio-temporal variation of PM_{2.5} concentrations and their relationship with geographic and socioeconomic factors in China. – *International Journal of Environmental Research & Public Health* 11(1): 173-186.
- [14] Gang, X. U., Jiao, L., Xiao, F., Zhao, S., Zhang, X. (2016): Applying land use regression model to estimate spatial distribution of PM_(2.5) concentration in Beijing-Tianjin-Hebei region. – *Journal of Arid Land Resources & Environment* 2016(10): 116-120.
- [15] Gilbert, N. L., Goldberg, M. S., Bernardo, B., Brook, J. R., Michael, J. (2005): Assessing spatial variability of ambient nitrogen dioxide in Montréal, Canada, with a land-use regression model. – *Journal of the Air & Waste Management Association* 55(8): 1059-1063.
- [16] Goossens, D., Riksen, M. (2004): Wind Erosion and Dust Dynamics at the Commencement of the 21st Century. – In: Goossens, D., Riksen, M. (eds.) *Wind Erosion and Dust Dynamics: Observations, Simulations, Modelling*. ESW Publications, Wageningen, pp. 7-13.
- [17] Guan, Q., Cai, A., Wang, F., Yang, L., Xu, C., Liu, Z. (2017): Spatio-temporal variability of particulate matter in the key part of Gansu Province, Western China. – *Environmental Pollution* 230: 189-198.
- [18] He, G., Fan, M., Zhou, M. (2016): *The Effect of Air Pollution on Mortality in China: Evidence from the 2008 Beijing Olympic Games*. – Social Science Electronic Publishing, New York.
- [19] He, Q., Bo, H. (2018): Satellite-based high-resolution PM_{2.5} estimation over the Beijing-Tianjin-Hebei region of China using an improved geographically and temporally weighted regression model. – *Environmental Pollution* 236:1027-1037..
- [20] Hoek, G., Beelen, R., Hoogh, K. D., Vienneau, D., Gulliver, J., Fischer, P., Briggs, D. (2008): A review of land-use regression models to assess spatial variation of outdoor air pollution. – *Atmospheric Environment* 42(33): 7561-7578.
- [21] Hu, J., Ying, Q., Wang, Y., Zhang, H. (2015): Characterizing multi-pollutant air pollution in China: Comparison of three air quality indices. – *Environment International* 84: 17-25.
- [22] Jedynska, A., Hoek, G., Wang, M., Eeftens, M., Cyrus, J., Beelen, R., Cirach, M., Nazelle, A. D., Keuken, M., Visschedijk, A. (2015): Spatial variations of levoglucosan in four European study areas. – *Science of the Total Environment* 505: 1072-1081.
- [23] Jerrett, M., Arain, A. P., Beckerman, B., Potoglou, D., Sahuvaroglu, T., Morrison, J., Giovis, C. (2005): A review and evaluation of intraurban air pollution exposure models. – *J Expo Anal Environ Epidemiol* 15(2): 185-204.
- [24] Jian-Sheng, W. U., Wang, X., Jia-Cheng, L. I., Yuan-Jie, T. U. (2017): Comparison of models on spatial variation of PM_(2.5) concentration: a case of Beijing-Tianjin-Hebei region. – *Environmental Science*. DOI: 10.13227/j.hjxk.201611114.
- [25] Liang, Z., Zou, B., Xin, F., Luo, Y., Wan, N., Shuang, L. (2016): Land use regression modeling of PM_{2.5} concentrations at optimized spatial scales. – *Atmosphere* 8(1): 1.
- [26] Liu, C., Henderson, B. H., Wang, D., Yang, X., Peng, Z. R. (2016): 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. – *Science of the Total Environment* 565: 607-615.
- [27] Liu, M., Huang, Y., Jin, Z., Ma, Z., Liu, X., Zhang, B., Liu, Y., Yu, Y., Wang, J., Bi, J. (2017): The nexus between urbanization and PM_{2.5} related mortality in China. – *Environmental Pollution* 227: 15.
- [28] Liu, W., Liu, Q., Liu, W. X., Yu, S. Y., Wang, X., Xu, Y. S., Liu, Y., Tao, S. (2018): Oxidative potential of ambient PM_{2.5} in the coastal cities of the Bohai Sea, Northern China: seasonal variation and source apportionment. – *Environmental Pollution* 236: 514-528.
- [29] Lu, W. Z., Fan, H. Y., Lo, S. M. (2003): Application of evolutionary neural network method in predicting pollutant levels in downtown area of Hong Kong. – *Neurocomputing* 51(2): 387-400.
- [30] McAuliffe, R. E. (2015): Durbin–Watson Statistic. – In: McAuliffe, R. E. (ed.) *The Blackwell Encyclopedia of Management*. Vol. 8: Managerial Economics. Blackwell, Hoboken, NJ.
- [31] Meng, X., Fu, Q., Ma, Z., Chen, L., Zou, B., Zhang, Y., Xue, W., Wang, J., Wang, D., Kan, H. (2015): Estimating ground-level PM₁₀ in a Chinese city by combining satellite data, meteorological information and a land use regression model. – *Environmental Pollution* 208(Pt A): 177.
- [32] Morrison, W. M. (2015): China’s Economic Rise: History, Trends, Challenges, and Implications for the United States. – *Congressional Research Service Reports* 7-5700, www.crs.gov, RL33534.
- [33] National Bureau of Statistics (2015): *Hebei Economic Yearbook*. – China Statistics Press, Beijing.
- [34] Pattey, E., Qiu, G. (2012): Trends in primary particulate matter emissions from Canadian agriculture. – *Journal of the Air & Waste Management Association* 62(7): 737-747.
- [35] Rohde, R. A., Muller, R. A. (2015): Air pollution in China: mapping of concentrations and sources. – *Plos One* 10(8): e0135749.
- [36] Ross, Z., Jerrett, M., Ito, K., Tempalski, B., Thurston, G. D. (2007): A land use regression for predicting fine particulate matter concentrations in the New York City region. – *Atmospheric Environment* 41(11): 2255-2269.
- [37] Roswall, N., Raaschou-Nielsen, O., Ketzel, M., Gammelmark, A., Overvad, K., Olsen, A., Sørensen, M. (2017): Long-term residential road traffic noise and NO₂ exposure in relation to risk of incident myocardial infarction - a Danish cohort study. – *Environmental Research* 156: 80-86.
- [38] Shekarrizfard, M., Faghieh-Imani, A., Dan, L. C., Goldberg, M., Ross, N., Eluru, N., Hatzopoulou, M. (2016): Individual exposure to traffic related air pollution across land-use clusters. – *Transportation Research Part D Transport & Environment* 46: 339-350.
- [39] Song, C., Lin, W., Xie, Y., He, J., Xi, C., Wang, T., Lin, Y., Jin, T., Wang, A., Yan, L. (2017): Air pollution in China: Status and spatiotemporal variations. – *Environmental Pollution* 227: 334-347.
- [40] Su, J. G., Brauer, M., Ainslie, B., Steyn, D., Larson, T., Buzzelli, M. (2008): An innovative land use regression model incorporating meteorology for exposure analysis. – *Science of the Total Environment* 390(2): 520-529.
- [41] Wang, J., Zimei, H. U., Chen, Y., Chen, Z., Shiyuan, X. U. (2013): Contamination characteristics and possible sources of PM₁₀ and PM_{2.5} in different functional areas of Shanghai, China. – *Atmospheric Environment* 68(2): 221-229.
- [42] Wang, L., Zhang, N., Liu, Z., Sun, Y., Ji, D., Wang, Y. (2015): The Influence of Climate Factors, Meteorological Conditions, and Boundary-Layer Structure on Severe Haze Pollution in the Beijing-Tianjin-Hebei Region during January 2013. – *Advances in Meteorology* 2014(7): 1-14.

- [43] Wang, S., Zhou, C., Wang, Z., Feng, K., Hubacek, K. (2017): The characteristics and drivers of fine particulate matter (PM 2.5) distribution in China. – *Journal of Cleaner Production* 142(4): 1800-1809.
- [44] Wheeler, A. J., Smith-Doiron, M., Xu, X., Gilbert, N. L., Brook, J. R. (2008): Intra-urban variability of air pollution in Windsor, Ontario—Measurement and modeling for human exposure assessment. – *Environmental Research* 106(1): 7-16.
- [45] WHO (2016): *World Health Statistics 2016. Monitoring Health for the SDGs.* – WHO, Geneva.
- [46] Wu, J., Li, J., Jian, P., Li, W., Xu, G., Dong, C. (2014): Applying land use regression model to estimate spatial variation of PM 2.5 in Beijing, China. – *Environmental Science & Pollution Research* 22(9): 7045-7061.
- [47] Wu, J., Yao, F., Li, W., Si, M. (2016): VIIRS-based remote sensing estimation of ground-level PM2.5 concentrations in Beijing–Tianjin–Hebei: a spatiotemporal statistical model. – *Remote Sensing of Environment* 184: 316-328.
- [48] Yang, G., Huang, J., Li, X. (2018): Mining sequential patterns of PM2.5 pollution in three zones in China. – *Journal of Cleaner Production* 170: 388-398.
- [49] Yao, F., Si, M., Li, W., Wu, J. (2017): A multidimensional comparison between MODIS and VIIRS AOD in estimating ground-level PM2.5 concentrations over a heavily polluted region in China. – *Science of the Total Environment* 618: 819-828.
- [50] Yin, H., Pizzol, M., Xu, L. (2017): External costs of PM2.5 pollution in Beijing, China: Uncertainty analysis of multiple health impacts and costs. – *Environmental Pollution* 226: 356.
- [51] Zhang, Y., Lin, Y., Cai, J., Liu, Y., Hong, L., Qin, M., Zhao, Y., Ma, J., Wang, X., Zhu, T. (2016): Atmospheric PAHs in North China: spatial distribution and sources. – *Science of the Total Environment* 565: 994-1000.

APPENDIX

Table A1. Regression results of Hebei PM_{2.5} LUR model

Model parameter	Unit	Unstandardized coefficient β	Standardized coefficient β	p-values
Constant	–	1472.79	–	< 0.01
Air pressure	hPa	0.95	0.94	< 0.01
Longitude	°	-19.34	-1.14	< 0.01
Precipitation	mm	0.09	0.23	< 0.01
Temperature	°C	-8.39	-0.62	< 0.01
Farmland_500 m	km ²	33.1	0.07	0.02
Elevation	m	-0.03	-0.29	< 0.01
Wind speed	m/s	-10.05	-0.15	< 0.01
Emission monitoring enterprises_7.5 km	N	1.7	0.1	< 0.01
Adjusted R2		0.96		
Durbin–Watson statistic		1.95		
F-values		157.71		p < 0.01

Table A2. Regression results of Hebei NO₂ LUR model

Model parameter	Unit	Unstandardized coefficient β	Standardized coefficient β	p-values
Constant	–	-1500.62	–	< 0.01
Distance to coastline	km	0.20	1.88	< 0.01
Latitude	°	19.11	2.14	< 0.01
Forest land_1 km	km ²	-3.28	-0.20	0.073
Water_10 km	km ²	0.88	0.20	0.082
Emission monitoring enterprises_8.5 km	N	2.93	0.48	< 0.01
Wind speed	m/s	14.17	0.51	0.04
Temperature	°C	6.09	1.09	0.04
Air pressure	hPa	0.40	0.96	0.02
Relative humidity	%	4.56	1.35	0.02
Adjusted R ²	0.81			
Durbin–Watson statistic	2.08			
F-values	25.31			p < 0.01

Table A3. Explanation of abbreviations of land use

Abbreviation for land use data	Description
build_buffer	The area of building land in a circular area centered on a monitoring station
fore_buffer	The area of forest in a circular area centered on a monitoring station
farm_buffer	The area of farmland in a circular area centered on a monitoring station
wat_buffer	The area of water in a circular area centered on a monitoring station
ghl_buffer	Greenhouse area in a circular area centered on a monitoring station
gar_buffer	Garden area in a circular area centered on a monitoring station