

# Spatial Heterogeneity Analysis of PM<sub>2.5</sub> Concentration in Central Plains Economic Region

Kaiguang Zhang, Hongling Meng, Mingting Ba, Yanmin Sun

School of Geography and Tourism, Zhengzhou Normal University, Zhengzhou, China

Email: zsgis@zznu.edu.cn

**How to cite this paper:** Zhang, K. G., Meng, H. L., Ba, M. T., & Sun, Y. M. (2020). Spatial Heterogeneity Analysis of PM<sub>2.5</sub> Concentration in Central Plains Economic Region. *Journal of Geoscience and Environment Protection*, 8, 244-254.

<https://doi.org/10.4236/gep.2020.812015>

**Received:** November 7, 2020

**Accepted:** December 25, 2020

**Published:** December 28, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

The research of the spatial heterogeneity of PM<sub>2.5</sub> concentration in an area, is of great significance for understanding its regional spatial distribution structure, exploring the transmission relationship between regions, in order to formulate joint prevention and control measures within the entire area. Based on the daily monitoring data of PM<sub>2.5</sub> concentration in the Central Plains Economic Region in 2019, this paper utilizes cluster analysis to divide the regional PM<sub>2.5</sub> concentration into 5 classes, builds their spatial semi-variogram model, and then utilizes interpolation analysis method to study the regional overall distribution characteristics and transmission law. The results show that the PM<sub>2.5</sub> concentration in the Central Plains Economic Region has a medium or higher spatial autocorrelation. The critical value of the overall PM<sub>2.5</sub> concentration in the area is 150 µg/m<sup>3</sup>, as the overall PM<sub>2.5</sub> concentration less than the value, the PM<sub>2.5</sub> in a region mainly comes from local emissions, as the overall PM<sub>2.5</sub> concentration higher than the value, the influence of spatial structure on the distribution of PM<sub>2.5</sub> concentration is gradually obvious. PM<sub>2.5</sub> has a certain degree of spatial transmission, which mainly includes two routes as Puyang-Xingtai and Puyang-Zhengzhou, and the transmission intensity of the former is greater than the latter.

## Keywords

Cluster Analysis, Semi-Variogram, Spatial Heterogeneity, Spatial Structure, Central Plains Economic Region

## 1. Introduction

With the continuous development of China's economy, the rapid urbanization and industrialization, the increase of energy and natural resources consumption, a large amount of harmful substances have been discharging into the atmos-

phere. Air pollution is becoming more and more seriously, the problem of air pollution control has been becoming one of the main problems to be solved in the sustainable development of China's economy and satisfying people's yearning for a better life (Liu et al., 2017; Ma et al., 2016; Lin & Wang, 2016).

The components of atmospheric pollutants in China mainly include PM<sub>2.5</sub>, PM<sub>10</sub>, SO<sub>2</sub>, NO<sub>2</sub>, CO and O<sub>3</sub>. The statistical results show that PM<sub>2.5</sub> is the primary pollutant in most regions in pollution weather, which causes more than 50% pollution days of the total pollution weather. PM<sub>2.5</sub> is atmospheric particulate matter with aerodynamic equivalent diameter less than 2.5 microns, mainly composed of water-soluble ions, carbon components and chemical elements, all of which have the characteristics of small diameter, large surface area, strong activity, strong adsorb ability for toxic and harmful substances (for example, heavy metals, microorganisms, etc.), long residence time in the atmosphere and long transport distance (Yuan et al., 2018; Bie et al., 2018; Zhang et al., 2019b; Zhang et al., 2019a).

In recent years, Many researchers have focused on the components analysis and diffusion model analysis, and the spatial distribution characteristics analysis in multi-scale of geoscience vision, as well as the influencing factors analysis, have achieved a lot of useful results (Yuan et al., 2018; Bie et al., 2018; Zhang et al., 2019b; Zhang et al., 2019a; Chen et al., 2017; Wen et al., 2018; Zhao et al., 2017; Zhang et al., 2019c; Chen et al., 2019; Li et al., 2019; Liao et al., 2016).

Although the amount of PM<sub>2.5</sub> emissions is the main factor determining regional concentration, the regional topographic conditions, wind direction, wind speed, precipitation and temperature have some impacts on the its spread, regional transmission and transport is one of its important features, the concentration in a region is affected by the concentrations of its adjacent regions, and there is some certain correlations between regions (Zhao et al., 2017; Zhang, Meng, 2019b).

However, for the limited number of monitoring stations with an uneven distribution and the short of observation time, there are certain limitations in studying its temporal and spatial distribution characteristics and changing trends in a large-scale range. Therefore, based on the daily monitoring data of PM<sub>2.5</sub> concentration in the Central Plains Economic region in 2019, this article uses the semi-variogram model to analyze the spatial heterogeneity of PM<sub>2.5</sub> concentration, study the characteristics of its spatial variation and explore its transmission relationship in the area, in order to provide a scientific reference for the formulation of its joint prevention and control measures in the Central Plains Economic Region.

## 2. Data and Research Methods

### 2.1. Research Area Overview and Data

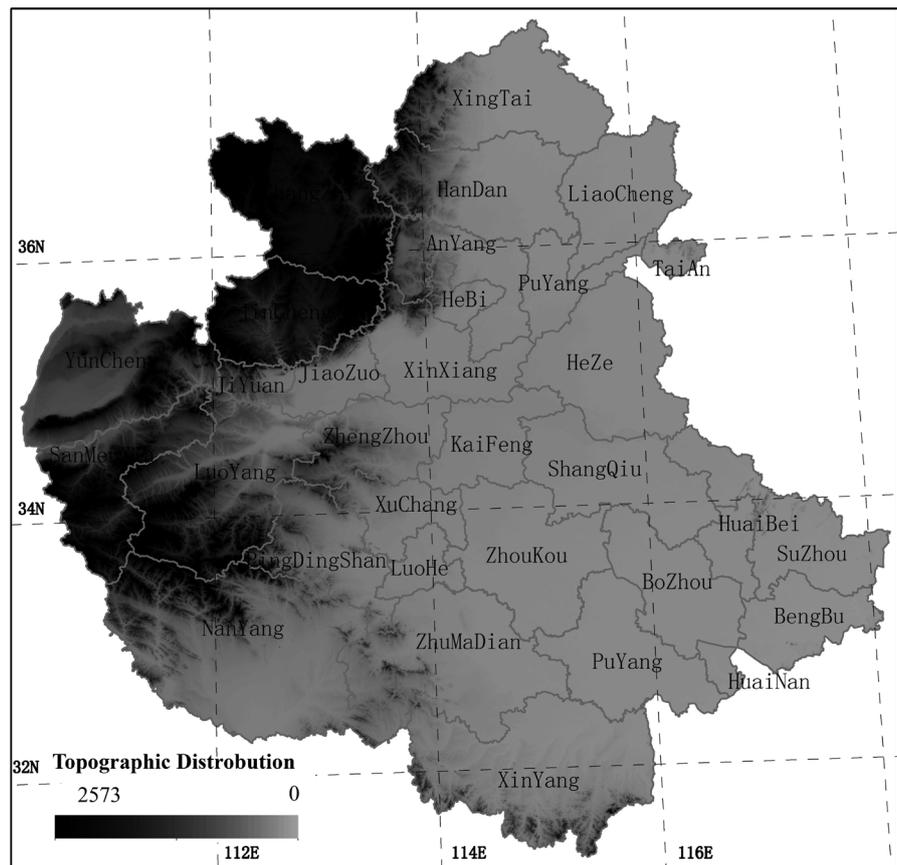
The Central Plains Economic region (31°23'N - 37°49'N, 110°21'E - 118°11'E) with

a total area of 280,000 km<sup>2</sup>, located in the central part of China, is a key development area in the planning of China's main functional zone, covering the whole province of Henan, the northwest part of Anhui Province (including Huaibei, Suzhou, Fuyang, Bozhou, Bengbu, and a part of Huainan), the southwest part of Shandong Province (including Heze, Liaocheng, and a part of Tai'an), the south part of Hebei Province (including Handan and Xingtai), the southeast part of Shanxi Province (including Jincheng, Changzhi, and Yuncheng). The west and south sides are distributed by Funiu Mountain, Tongbai Mountain and Dabie Mountain along the boundary in a semi ring shape, the middle and east parts are the Huanghuaihai alluvial plain. Topographically, the west of the area is higher and lower in the east, higher in the south and lower in the north (see **Figure 1**).

The data used in the study mainly includes a 30 m × 30 m digital elevation model, which comes from Geospatial Data Cloud (<http://www.gscloud.cn>), the regional PM<sub>2.5</sub> concentration data comes from China Air Quality Online Monitoring and Analysis Platform (<https://www.aqistudy.cn/historydata/>).

## 2.2. The Spatial Heterogeneity of PM<sub>2.5</sub> Concentration

PM<sub>2.5</sub> concentration  $Z(P) = Z(x, y)$  is a regionalized variable, which is different from an ordinary random variable. The value of an ordinary random



**Figure 1.** Topographic distribution characteristics in central plains economic region.

variable conforms to a certain probability distribution, while a regionalized random variable takes a value according to its location in a region. In other words, the regionalized random variable is the specific value of the ordinary random variable at a certain time for a specific location. It is a random function based on the position, an ordinary random variable in one place (Chen et al., 2019; Li et al., 2019; Liao et al., 2016).

PM2.5 concentration has two significant characteristics of randomness and structure. Firstly, it is a random function, has locality, randomness and abnormality. Secondly, it has average structural properties, to some extent reflects the spatial correlation characteristics of the PM2.5 concentration  $Z(P_0)$  and the PM2.5 concentration  $Z(P_1)$ , where  $P_1$  is  $h$  away from  $P_0$ , these spatial correlations depends on the distance  $h$  between the two points and the characteristics of the PM2.5 concentration itself.

The correlation coefficient between two points is

$$c(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} [Z'_i(P_1)Z'_i(P_2)] \quad (1)$$

where  $N(h)$  is the number of observation value pairs,  $Z'$  is the normalized observation value of  $Z$ .

The research PM2.5 results with the confidence level of 0.95%, were shown. The PM2.5 concentration between adjacent regions in the Central Plains Economic Region presents a significant correlation characteristic, with the correlation coefficient of  $r \geq 0.90$  ( $R^2 \geq 0.96$ ) (Zhang et al., 2019a). There was a significant spatial autocorrelation characteristic showing that Xinxiang and Zhengzhou are the center of HH aggregation region, and the aggregation degree tends to be significant with the increase of the average PM2.5 concentration (Zhang et al., 2019b).

In order to study the overall diffusion and transport characteristics of PM2.5 in an area, explore the spatial variation of PM2.5 concentration at different distances, the distribution of the spatial variation at a specific distance, as well as the relationship between PM2.5 concentration distribution and distance, defines the semi-variogram as a half of the sum of covariance of the observation point pairs  $(P_i, P'_i)$ , where  $P'_i$  is  $h$  apart from  $P_i$ .

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(P_i) - Z(P'_i)]^2 \quad (2)$$

$\gamma(h)$  reflects the PM2.5 concentration variation of all point pairs at the distance of  $h$  in the discrete point set, as well as the concentration variation and its spatial distribution characteristics. The semi-variogram has four important parameters, namely the nugget or localized discontinuity,  $C_0 = \lim_{h \rightarrow 0} \gamma(h)$ , which describes the spatial variation caused by random factors; with the interval distance increase,  $\gamma(h)$  from  $C_0$  gradually tends to a constant  $C_0 + C = \lim_{h \rightarrow \infty} \gamma(h)$ , called the sill value, which reflects the largest variation of PM2.5 concentration; when  $\gamma(h)$  reaches near the sill value and tends to stabilize, the corresponding

value  $h$  is called the range ( $a$ ), which reflects when  $h \geq a$ , the spatial correlation of PM2.5 concentration disappears.

### 2.3. The Spatial Distribution Characteristics and Interpolation Model of PM2.5 Concentration

The regional PM2.5 concentration has a significant spatial correlation. In order to study the spatial variation and spatial heterogeneity at different distances, fit the discrete variation value  $\gamma(h)$  with a mathematical function, by using the least square method to ensure the maximum regression  $R^2$  and the minimum residual  $RSS$ , obtain the mathematical model of the semi-variogram. Commonly the fitting functions include Exponential model, Spherical model, Gaussian model, and Linear model.

Based on the semi-variogram model, using interpolation analysis obtains the PM2.5 concentration at unknown points, could present the overall spatial distribution and variation characteristics of PM2.5 concentration in the whole area.

The Kriging interpolation method is based on the data of known points in the finite neighborhood of one unknown point, taking into their spatial position relationship and the structural information provided by the best fitting function model, using unbiased optimal estimation to calculate the value of the unknown point (Li et al., 2019).

$$Z^*(P) = \sum_{i=1}^n \lambda_i Z(P_i), \quad \lambda = K_{n \times n}^{-1} D_{n \times 1} \quad (3)$$

where  $\lambda_i$  is the contribution of the PM2.5 concentration value of the known point  $P_i$  to the unknown point  $P$ .  $n$  is the number of points within the specified distance neighborhood of  $P$ .  $K$  is the augmented matrix of the semi-variogram values on the known points,  $D$  is an augmented column vector composed of the semi-variogram values from the known points to the unknown point.

## 3. Results and Analysis

### 3.1. Statistical Characteristics of PM2.5 Concentration

The daily monitoring data of PM2.5 concentration presents the average level of PM2.5 emissions around the monitoring stations and the diffusion from surrounding regions, which is affected by many factors such as daily emissions, meteorological conditions and etc., with a certain degree of randomness. for accurate research its spatial distribution and variation characteristics in the area, carries K-mean cluster analysis on the data, using the monitoring stations as the variables, the Euclidean distance between days as the criterion function, and the average concentration of PM2.5 corresponding to the first 5 levels of air quality, defined in the “Ambient Air Quality Index (AQI) Technical Regulations” (HJ 633-2012), as the starting clustering center vector, to classify the days into 5 classes, respectively corresponding to PM2.5 concentration of 5 levels, as Good, Moderate, lightly polluted, moderately polluted and heavily polluted. Calculate

their corresponding class centers as the basic data of variation characteristics research, and their spatial distribution is shown in the **Figure 2**. The spatial distributions of 5 classes show some non-uniformity and great difference, the high concentration regions show a zonal distribution, mainly located in the northern and central areas. The high concentration regions also have relatively high population densities, and relatively developed economy and relatively large proportion of total industrial output value, this is one of the main factors that lead to significantly higher concentration of PM2.5 in these areas than that in other areas.

### 3.2. Analysis of the Spatial Heterogeneity of PM2.5 Concentration

In order to quantitatively analyze the spatial heterogeneity characteristics of PM2.5 concentration at different levels, the semi-variogram model is used to calculate the digital characteristics of the spatial heterogeneity function of each level, the results are showed in **Table 1**.

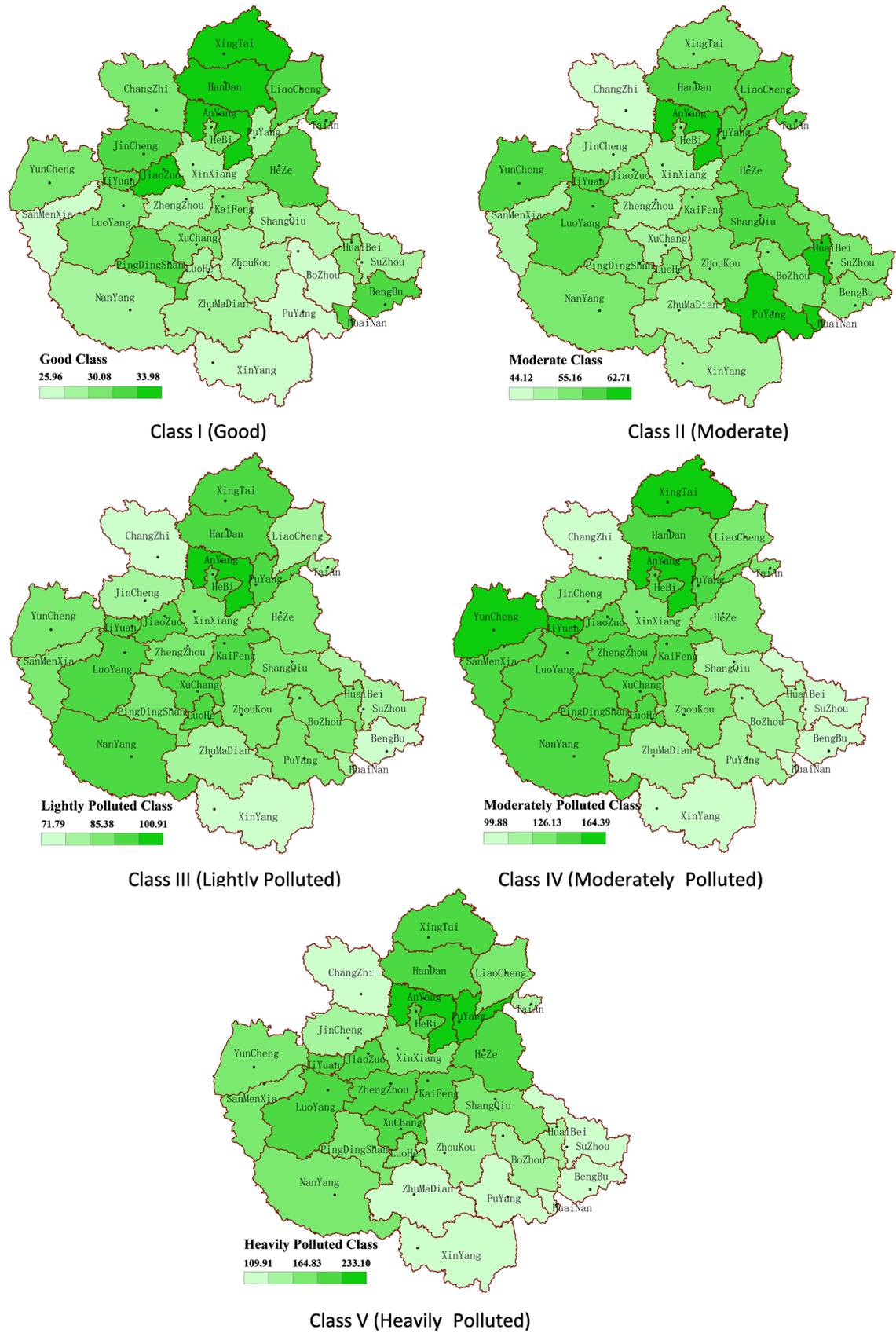
The variation coefficient (Li et al., 2019) increases with the increase of PM2.5 concentration, which means that with the increase of overall regional concentration, the spatial variability gradually strengthens. The nugget value ( $C_0$ ) represents the spatial heterogeneity caused by random factors, its variation range is between 2.00 and 549.00, as the overall concentration increases in the area, it exhibits a large fluctuation, and the value in level III is the minimum.

The partial sill value ( $C$ ) represents the spatial heterogeneity of PM2.5 concentration caused by the spatial structure, and its variation range is 4.19 - 2175, as the overall concentration increases, it shows a rapid growth trend.

The sill value ( $C + C_0$ ), is the sum of the nugget value and the partial sill value, which describes the total spatial heterogeneity, its variation is similar to that of the nugget value, and the variation range has been expanded to 6.85 - 2724.

The nugget-sill ratio ( $C_0/(C + C_0)$ ) indicates the proportion of the spatial heterogeneity caused by random factors to the overall spatial heterogeneity, shows a relatively large fluctuation, ranging from 3% to 39%. This means that the spatial heterogeneity of PM2.5 concentration presents some certain nugget effect. As PM2.5 concentration is in level I, the proportion caused by random factors in the area is the maximum; as PM2.5 concentration is in level III, the proportion is the minimum, in other words, the spatial heterogeneity caused by the regional spatial structure is the maximum. In terms of correlation degree (Chen et al., 2019; Liao et al., 2016), levels II, III, and V have strong spatial autocorrelation, and levels 1 and 4 have moderate spatial correlation. In general, the fluctuation of the nugget-sill ratio is mainly affected by the changes of nuggets. Therefore, when studying the spatial heterogeneity of PM2.5 concentration, the influence of random factors must be considered.

The range indicates the variation range and uniformity of spatial heterogeneity. In general, the range shows a certain fluctuation with the increase of the overall concentration, ranging from 121.40 to 589.070 km. The PM2.5 concentrations in



**Figure 2.** PM2.5 spatial distribution of 5 classes in central plains economic region.

**Table 1.** Characteristic parameters of the spatial heterogeneity.

Level	Variation Coefficient	Nugget	Partial Sill	Sill	nugget-sill ratio	Range	$R^2$	Fitting Coefficient
1	8.61	2.66	4.191	6.851	0.39	441.30	0.64	Exponential
2	8.38	5.53	18.31	23.84	0.23	125.80	0.38	Spherical
3	9.54	1.90	65.53	67.43	0.03	121.40	0.56	Spherical
4	17.82	227.00	604.90	831.9	0.27	589.07	0.91	Gaussian
5	25.43	549.00	2175.00	2724	0.20	564.20	0.91	Spherical

level IV and V have a wide range of spatial heterogeneity, and the spatial distribution is relatively uniform; the PM<sub>2.5</sub> concentrations in level II and III have a small spatial heterogeneity range, and the spatial distribution presents non-uniform characteristics.

### 3.3. Spatial Distribution Characteristics and Interpolation Analysis of PM<sub>2.5</sub> Concentration

To study the overall spatial distribution of PM<sub>2.5</sub> concentration, using the four semi-variogram models, mentioned before, to fit on the discrete variance values. According to the principle of optimal fitting criterion, obtain 5 optimal heterogeneity functions as

$$\gamma_1(h) = 2.66 + 4.19(1 - e^{-2.67 \times 10^{-6} h}) \quad (4)$$

$$\gamma_2(h) = \begin{cases} 5.53 + 2.18 \times 10^{-4} - 4.60 \times 10^{-14} h^3 & (0 \leq h < 1.25 \times 10^5) \\ 23.84 & (h \geq 1.25 \times 10^5) \end{cases} \quad (5)$$

$$\gamma_3(h) = \begin{cases} 1.90 + 7.84 \times 10^{-4} h - 1.83 \times 10^{-14} h^3 & (0 \leq h < 1.21 \times 10^5) \\ 67.63 & (h \geq 1.21 \times 10^5) \end{cases} \quad (6)$$

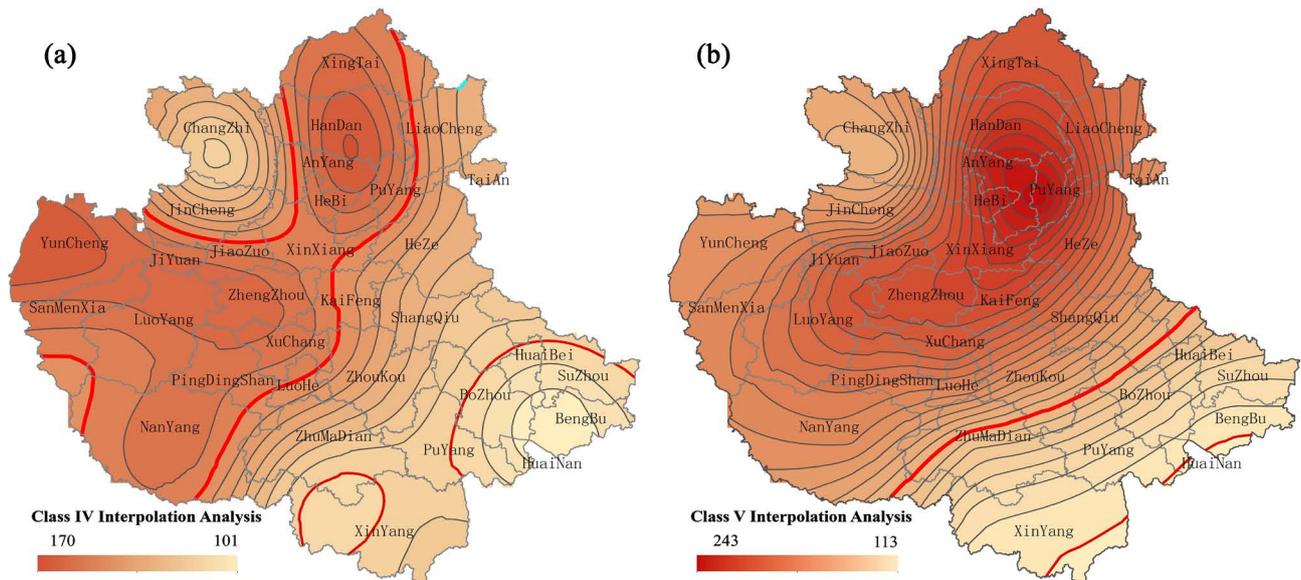
$$\gamma_4(h) = 227.00 + 604.30(1 - e^{-2.88 \times 10^{-12} h^2}) \quad (7)$$

$$\gamma_5(h) = \begin{cases} 549 + 2.78 \times 10^{-3} h - 6.06 \times 10^{-14} h^3 & (0 \leq h < 5.89 \times 10^5) \\ 2724 & (h \geq 5.89 \times 10^5) \end{cases} \quad (8)$$

The fitting coefficient  $R^2$  is the ratio of the sum of regression squares to the total deviation in the fitting model. Among the five models, only IV and V, corresponding to levels IV and V,  $R^2$  reach more than 90%, the fitting is significant, while the rest the three models is not significant.

Carrying Kriging interpolation analysis on the PM<sub>2.5</sub> concentrations of levels IV and V, the results are shown in **Figure 3**, the solid line is the 5  $\mu\text{g}/\text{m}^3$  contour line.

As the PM<sub>2.5</sub> concentration is in level IV, see **Figure 3(a)**, the regional PM<sub>2.5</sub> concentration presents a large difference. The area with a concentration value within the corresponding concentration range of level 4 air quality ((115 - 150)  $\mu\text{g}/\text{m}^3$ ),



**Figure 3.** Interpolation analysis of class IV and V in central plains economic region.

accounts for 45.03% of the total area, distributes in the southwest-northeast distribution along Zhumadian-Heze. The proportion of 11.45% in the level belongs to level 3 air quality ( $(75 - 115] \mu\text{g}/\text{m}^3$ ), they mainly distribute in the south of the Xinyang-HuaiBei. The proportion of 43.52% in the level belongs to level 5 air quality ( $150 - 250] \mu\text{g}/\text{m}^3$ ), they show two pollution systems, the first is distributed in the north of Hebi, centers as Handan with north-south distribution; the second is located to the west of Zhengzhou, with YunCheng as the center, shows north by east to west distribution. Xinxiang links two systems together forming a level 5 air quality area.

As the PM<sub>2.5</sub> concentration is in level V, see **Figure 3(b)** the regional PM<sub>2.5</sub> concentration presents a central characteristic, spreading from to Puyang the surroundings. The area with a concentration value within the corresponding concentration range of level 5 ( $150 - 250] \mu\text{g}/\text{m}^3$ ) accounts for 76.39% of the total area, mainly distributes in the north of Zhumadian-Shangqiu. The area with the air quality of level 4 ( $115 - 150] \mu\text{g}/\text{m}^3$ ) in the level is mainly distributed in the south of the Xinyang-HuaiBei, accounts for 22.44% of the total area. There are also 1.17% area with the concentration of level 3 air quality ( $\leq 115 \mu\text{g}/\text{m}^3$ ), they distribute in the southern part of Xinyang, Huainan and Bengbu.

In the process of evolving from level IV to level V, the relative intensity of the first pollution system gradually weakens, and its center gradually moves to east. The relative intensity of the second pollution system gradually strengthens, and the center gradually moves southward, then they integrate into one, and the center is located near Puyang.

#### 4. Conclusion

The research results show that the PM<sub>2.5</sub> concentration in the study area has obvious spatial heterogeneity characteristics. With the increase of overall PM<sub>2.5</sub>

concentration, the spatial heterogeneity of concentration gradually increases. The critical value of the overall PM<sub>2.5</sub> concentration in the area is 150 µg/m<sup>3</sup>. As the overall PM<sub>2.5</sub> concentration is less than it, the PM<sub>2.5</sub> in the area mainly comes from local emissions. When the overall PM<sub>2.5</sub> concentration is higher than this value, the influence of geospatial structure on the distribution of PM<sub>2.5</sub> concentration is gradually obvious, and PM<sub>2.5</sub> has a certain degree of spatial transmission. The transmission mainly includes two routes as Puyang-Xingtai and Puyang-Zhengzhou, and the transmission intensity of the former is greater than the latter.

### Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

### References

- Bie, T., Han, L. J., Tian, S. F. et al. (2018). Method for Quantifying the Contribution of Urbanization on Population Exposure to Air Pollution. *Acta Ecologica Sinica*, *38*, 4570-4583. <https://doi.org/10.5846/stxb201706141086>
- Chen, Y. H., Wang, Y. J., & Zhang, M. Y. (2017). Temporal and Spatial Distribution of PM<sub>2.5</sub> in Shanghai Based on Clustering Analysis. *Chinese Journal of Environmental Engineering*, *11*, 3671-3677.
- Chen, Z., Luo, X. B., Gao, Y. H. et al. (2019). Modeling and Multi-Scale Analysis of the Spatial Heterogeneity of Land Surface Temperature in Chongqing Based on Semi-Variogram. *Journal of Geo-Information Science*, *21*, 1051-1060.
- Li, L., Zhang, S. K., & Yang, S. Q. (2019). Spatial-Temporal Variation of Mass Fraction of Soil Total phosphorus in Henan Province. *Science of Soil and Water Conservation*, *12*, 75-81.
- Liao, C. L., Fu, L. Y., Zhang, P. B., et al. (2016). Assessment and Spatial Variation of Soil Fertility Quality in Tobacco Field after Land Consolidation from Jinchenshi County, Hunan Province. *Chinese Journal of Soil Science*, *47*, 1077-1083.
- Lin, X. Q., & Wang, D. (2016). Spatio-Temporal Variations and Socio-Economic Driving Forces of Air Quality in Chinese Cities. *Acta Geographica Sinica*, *71*, 1357-1371.
- Liu, H. M., Fang, C. L., Zhang, X. L. et al. (2017). The Effect of Natural and Anthropogenic Factors on Haze Pollution in Chinese Cities: A Spatial Econometrics Approach. *Journal of Cleaner Production*, *165*, 323-333. <https://doi.org/10.1016/j.jclepro.2017.07.127>
- Ma, Z., Hu, X., Sayer, A. et al. (2016). Satellite-Based Spatiotemporal Trends in PM<sub>2.5</sub> Concentrations: China, 2004-2013. *Environmental Health Perspectives*, *124*, 184-192. <https://doi.org/10.1289/ehp.1409481>
- Wen, J., Shi, X. R., Tian, Y. Z. et al. (2018). Analysis of Chemical Composition of the Fine Particulate Matter in Summer in Tianjin City via a Single Particle Aerosol Mass Spectrometer (SPAMS). *Environmental Science*, *39*, 3492-3501.
- Yuan, C., Zhou, J. B., Xiong, Y. et al. (2018). Chemical Compositions and Long-Range Transport of PM<sub>2.5</sub> in Downtown Area of Wuhan. *Environmental Science & Technology*, *41*, 79-86.
- Zhang, H. T., Tian, Y. Z., Liu, B. S. et al. (2019a). Spatial Temporal Characteristics and Clus-

- ter Analysis of Chemical Components for Ambient PM<sub>2.5</sub> in Wuhan. *Environmental Science*, 40, 4764-4773. <https://doi.org/10.13227/j.hjcx.201904069>
- Zhang, K. G., Ba, M. T., Meng, H. L. et al. (2019b). An Air Quality Composite Index Based on Pollutant Concentration Factor Analysis. *African Journal of Mathematics and Computer Science Research*, 12, 10-16. <https://doi.org/10.5897/AJMCSR2019.0788>
- Zhang, K. G., Ba, M. T., Meng, H. L. et al. (2019c). Spatial Distribution Characteristics and Evolution Pattern of Air Quality in Henan Province. *Journal of Progressive Research in Mathematics*, 15, 2464-2468.
- Zhao, X. Y., Gu, C., Yang, H. M. et al. (2017). Chemical Composition and Source Apportionment of PM<sub>2.5</sub> during a Winter Air Pollution Episode in the Kui DuWu Area of Xinjiang Uygur Autonomous Region. *Research of Environmental Sciences*, 30, 1515-1523.