Study on the Influence of Socio-economic Activities on PM$_{2.5}$ in Henan Province Based on Principal Component Regression Analysis

Wei Yan$^{1,2,*}$, Shuangshuang Hou$^3$, Feng Zhou$^4$, Minghua Liu$^{1,2}$ and Haihua Sun$^{1,2}$

$^1$School of Geographic Sciences, Xinyang Normal University, 46400 Xinyang, China
$^2$Key Laboratory for Synergistic Prevention of Water and Soil Environmental Pollution, Xinyang Normal University, 46400, Xinyang, China
$^3$College of Life Sciences, Guizhou University, 550001 Guiyang, China
$^4$College of Computer and Information Technology, Xinyang Normal University, 464000, Xinyang, China
$^*$ Corresponding author: yanwei.jw@xynu.edu.cn

Abstract. In the context of increasingly serious haze pollution in Henan Province, which is harmful to human health, it is important to clarify the impact of socio-economic activities on PM$_{2.5}$, the key indicator of haze. Here, based on the analysis of the spatio-temporal variation characteristics of PM$_{2.5}$ concentration in Henan Province during 2015-2018, this paper combined the data of ten socio-economic indicators and adopted the principal component regression (PCR) analysis method to explore the influence of socio-economic activities on PM$_{2.5}$ concentration. The results showed that the annual average concentration of PM$_{2.5}$ during 2015-2018 ranged from 59.0 μg/m$^3$ to 83.8 μg/m$^3$, with a decreasing rate of -5.955 μg/(m$^3$·a). Spatial distribution of PM$_{2.5}$ showed "high in the north and low in the south", and the PM$_{2.5}$ concentration in Anyang, Zhengzhou, Jiaozuo and Xinxiang were more than 75μg/m$^3$, indicating a slightly polluted state. Among all the socio-economic indicators, gross domestic product (GDP), population density (PD), industrial value-added (IVA) and actual utilization of foreign capital (AUFC) played main positive roles in PM$_{2.5}$ from 2015 to 2018 but with a diminishing trend, and afforestation area (AA) played a negative role in PM$_{2.5}$. In addition, the effect of the sown area of major crops (SAMC) on PM$_{2.5}$ was negative with enhanced before 2018 and reversed in 2018. Our findings can provide a basis for the government to formulate PM$_{2.5}$ control policies.

1. Introduction

Haze is an aerosol consisting of sulfur dioxide, nitrogen dioxide, fine particulate matter, inhalable particulate matter and other atmospheric pollutants, which obscures the clarity of the sky and limits horizontal visibility to less than 10 km [1]. Haze has gained wide concern due to its huge harm to human society, including traffic accidents rising [2], harming human mental and physical health [3]. It is well known that socio-economic activities, such as economic growth, heavy industries, rapid increase in private car ownership and coal consumption for winter heating, have contributed to haze issues [4, 5]. The PM$_{2.5}$ (particle matter less than 2.5 μm in aerodynamic diameter), harming greatly human health [6, 7], is one of the important haze components and is a crucial index in studying haze. There are many
studies about PM$_{2.5}$, including variations and influence factors of PM$_{2.5}$. For variations of PM$_{2.5}$, spatio-temporal dynamics are the necessary contents that suggest the changing trend of PM$_{2.5}$ using linear regression and other statistic methods [8, 9]. For the influence factors of PM$_{2.5}$, it is suggested that the natural environment and socio-economic factors drive the generation and development of PM$_{2.5}$ [10, 11].

Henan Province is one of the most serious regions where haze issues occur frequently attributed to climate, terrain and social development [12]. For healthy social-economic development, haze pollution control is necessary. At present, there are a mass of studies on haze by PM$_{2.5}$ as the main index, and the variations and influence factors of PM$_{2.5}$ have been researched in Henan Province. The researches show that haze forms mainly in winter [13] because of the steady atmosphere layer [14], and social-economic activities influence PM$_{2.5}$ variations greatly [15, 16]. Discovering the influence of social-economic activities on PM$_{2.5}$ is of great significance for effective control of haze pollution. Although the influence of various possible socio-economic factors on PM$_{2.5}$ is discussed in other regions [11, 17, 18], how much these socio-economic factors affect PM$_{2.5}$ in Henan Province is uncertain, and that is the study objective of this work which can guide the control of haze pollution in Henan province.

2. Study area, Data and Methods

2.1. Study area

Sits in the central part of China, Henan Province is located from 31°23'-36°22'N, 110°21'-116°39'E (Fig. 1), with an altitude of -65~2392 m and a total area of 167,000 km$^2$. It is surrounded by Taihang Mountains, Funiu Mountains, Tongbai Mountains and Dabie Mountains in the north, west and south side, respectively, forming a semi-circular distribution along the provincial boundary. And Huang-Huai-Hai alluvial plain is located in the middle and east part of Henan Province. In terms of administrative divisions, Henan Province is surrounded by Anhui, Shandong, Hebei, Shanxi, Shaanxi and Hubei.
province, so Henan Province is an important comprehensive transportation hub and information flow centre of human flow and logistics in China. It belongs to continental monsoon climate with an annual average air temperature of 10.5~16.7 °C, annual precipitation between 407.7~1295.8 mm, annual sunshine hours of 1285.7~2292.9 hours, and an annual frost-free period of 201~285 days.

2.2. Data sources
In this paper, there are mainly two datasets, including PM$_{2.5}$ concentration and socio-economic data. PM$_{2.5}$ concentration data is obtained from the website of the Ministry of Ecology and Environment of the People's Republic of China (https://datacenter.mee.gov.cn/). The socio-economic dataset is obtained from the Statistical Yearbook of Henan Province (2016-2019).

Here, we constructed a set of indexes to quantify the socio-economic activities, which were confirmed to have an impact on PM$_{2.5}$ concentration [4, 18]. The index system consists of nine socio-economic indexes corresponding to ten indicators. Among them, gross domestic product (GDP) reflects the local economic level, and population density (PD) reflects the city scale. Industrial pollutants, energy consumption, dust pollutants, straw pollutants and tail gas pollution can be reflected by the indicators corresponding to industrial value-added (IVA), total social electricity consumption (TSEC), area of building under construction (ABUC), smoke and dust emission (SDE), the sown area of major crops (SAMC), and highway cargo throughput (HCT), respectively. The level of invite investment, with actual utilization of foreign capital (AUFC) as the indicator, can also generate PM$_{2.5}$ pollution. In order to promote economic development, a large number of foreign investments may bring some heavily polluting industries from abroad, which plays a significant role in PM$_{2.5}$ pollution in Henan Province. Green vegetation plays an important role in reducing PM$_{2.5}$, and greening level, with afforestation area (AA) as the indicator, can reflect the effect of government in environmental protection.

2.3. Methods
In this study, principal component regression (PCR), which can effectively avoid multicollinearity among explanatory variables [19], was used to estimate the model parameters, reflecting the individual impact of each influencing factor on PM$_{2.5}$. The process of PCR mainly consists of three steps, as follows: Firstly, the dimensionality of explanatory variables is reduced by using principal component analysis (PCA) to extract principal components, and the principal component scores are calculated. Here, if the cumulative variance contribution exceeds a certain threshold, the principal components will be considered to represent most of the indicators and selected to be used in the next step. Secondly, PM$_{2.5}$ is taken as the explained variable and the selected principal components as the explanatory variables to estimate parameters of the regression model, and its significance is tested in the meanwhile. Finally, the principal component function is put into the estimation regression equation to obtain the regression equation of the original variables. The final regression formula is as follows:

$$y = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \ldots + \beta_mx_m$$

(1)

Where $y$ represents the explained variable, $\beta_1, \beta_2, \beta_3, \ldots, \beta_m$ represent partial regression coefficients estimated by PCR for explanatory variables $x_1, x_2, x_3, \ldots, x_m$, respectively.

3. Results

3.1. Spatio-temporal dynamics of PM$_{2.5}$
Fig. 2 displays the spatial distribution (Fig. 2a) and the temporal trend (Fig. 2b) of the annual mean PM$_{2.5}$ concentration in Henan Province from 2015 to 2018. The temporal trends of the annual mean PM$_{2.5}$ in 17 cities are shown in Fig. 3.

Results of the spatial distribution indicated that the annual mean concentration of PM$_{2.5}$ ranged from 59.0 $\mu$g/m$^3$ to 83.8 $\mu$g/m$^3$, and the mean value of the whole province was 69.4 $\mu$g/m$^3$, with an increasing trend from south to north roughly. The highest concentration zone, with the PM$_{2.5}$ concentration above
75 μg/m³, was located in Anyang, Zhengzhou, Jiaozuo and Xinxiang, attributing to concentrated industry, high population density, high coal consumption and poor diffusion conditions [15, 20]. The lowest concentration zone was mainly distributed in the south part of Henan Province, including Nanyang, Zhumadian and Xinyang, attributing to an underdeveloped economy, high forest coverage and mountainous terrain. Ambient Air Quality Standards (GB3095-2012) defines PM$_{2.5}$ concentrations of 75 μg/m³ as a pollution threshold. Based on this standard, the area of light pollution (the PM$_{2.5}$ concentration ranges 75~115μg/m³) accounted for 6.62% of the total area of Henan Province. What is worse, even in Xinyang, where air quality was best, the PM$_{2.5}$ concentration was close to the pollution threshold of 75 μg/m³. Therefore, it was obvious that PM$_{2.5}$ pollution was serious in Henan Province.

Results of the temporal trend showed that the annual mean PM$_{2.5}$ concentration in Henan Province had a downtrend with a rate of -5.955 μg/(m³·a) from 2015 to 2018, which was similar to that of all the cities of Henan Province (Fig. 3). The trend of annual mean PM$_{2.5}$ concentration declined and was different from different cities. The trend of annual mean PM$_{2.5}$ concentration could be fit into four categories, as follows: (1) Type of sustained downward trend was the most widespread, distributing in Anyang, Jiaozuo, Xinxiang, Zhengzhou, Kaifeng, Xuchang, Sanmenxia, Pingdingshan, Luohe, Zhoukou and Xinyang. The largest declining trend stood in Xinxiang with a rate of -11.51 μg/(m³·a) and smallest in Kaifeng with a rate of -3.09 μg/(m³·a). (2) Type of first increased and then decreased trend distributed in Hebi, Shangqiu and Luoyang with 2016 as the boundary. The annual mean PM$_{2.5}$ concentration in Shangqiu had a significant drop in 2017. (3) Type of first decreased and then increased trend distributed in Zhumadian and Nanyang with 2017 as a boundary. (4) Type of “decreased–increased–decreased” trend only appeared in Puyang City.

Fig.2  The spatial distribution (a) and the temporal trend (b) of the annual mean PM$_{2.5}$ concentration in Henan Province from 2015 to 2018. (Note: the slope was calculated by the method of linear regression)
Fig. 3  The temporal characteristics of the annual mean PM$_{2.5}$ concentration in all cities of Henan Province from 2015 to 2018. (Note: the slope was calculated by the method of linear regression)

To summarize, the trend of annual mean PM$_{2.5}$ concentration showed a decreasing trend roughly, which indicated that the government of Henan Province had been effective in controlling PM$_{2.5}$ from 2015. However, Zhumadian and Nanyang had an increasing trend of annual mean PM$_{2.5}$ concentration after 2017, demonstrating where PM$_{2.5}$ pollution became worse and the government needed to step up efforts to control.
3.2. Influence of socio-economic activities on PM$_{2.5}$

To eliminate possible heteroscedasticity and the effect of different units, PM$_{2.5}$ and ten socio-economic indicators values were taken logarithmic and normalized. Then, Kaiser-Meyer-Olkin (KMO) and Bartlett's test of sphericity method were used to test the selected indicators data. Results showed that KMO values during 2015-2018 were more than 0.5, and the P values of Bartlett's test of sphericity were all less than 0.05, indicating that the selected data can be used for PCR analysis [21].

Fig. 4 displays partial regression coefficients between socio-economic indicators and PM$_{2.5}$ from 2015 and 2018 calculated by PCR analysis, with a positive value indicating positive effect (highlighted in red), a negative value for negative effect (highlighted in blue). Results showed that all the socio-economic indicators had different influence on PM$_{2.5}$ and varied from year to year. Except SDE, SAMC and AA, the effect on PM$_{2.5}$ of other indicators (GDP, PD, IVA, TSEC, ABUC, HCT and AUFC) was positive from 2015 to 2018. Among them, the effect of GDP, IVA, TSEC and AUFC showed the characteristics of “diminishing–strengthening–diminishing”, and PD, ABUC and HCT showed first diminishing and then strengthening trend. The effect of AA on PM$_{2.5}$ was negative but diminished. The effect of SAMC on PM$_{2.5}$ was negative with enhanced before 2018 and reversed in 2018. The effect of SDE on PM$_{2.5}$ was positive and reversed after 2017. Those positive effect (GDP, IVA, TSEC, ABUC, HCT and AUFC) diminished overall during 2015-2018 suggested that policies of optimizing the industrial structure and controlling environment had a certain effect, which was indicated by the negative effect of AA diminished. The effect of SAMC on PM$_{2.5}$ deserves particular concern due to a change from negative effect forward to positive.

4. Conclusions and discussion

In order to deeply understand the relationship between PM$_{2.5}$ and socio-economic factors, this paper analyzed the spatio-temporal variation characteristics of PM$_{2.5}$ of Henan Province during 2015-2018, and clarified the impact of socio-economic activities on PM$_{2.5}$ by the method of PCR. Our conclusions are as follows:

(1) The annual mean PM$_{2.5}$ distributed inhomogeneous in space. The annual mean PM$_{2.5}$ concentration ranged 59.0–83.8 μg/m$^3$, with an increasing trend from south to north roughly during 2015-2018. There was a decreasing temporal trend from 2015 to 2018 roughly. However, the temporal trend of different cities showed differences including four categories, that was, type of sustained downward trend, first increased and then decreased trend, first decreased and then increased trend and “decreased–increased–decreased” trend. The decreasing of PM$_{2.5}$ indicated that the government of Henan Province has been effective in controlling PM$_{2.5}$ since 2015.

(2) Among all the socio-economic factors, GDP, PD, IVA and AUFC were the main factors that played a positive role in promoting the development of PM$_{2.5}$ during 2015-2018. Therefore, in the future economic development layout of Henan Province, it is important to further optimize the industrial structure, including developing industries of science and technology, increasing the proportion of new-type industrialization and reducing the proportion of high-polluting industries in the national economy [4, 13]. In addition, it is necessary to consider the optimization of population distribution by population
migration. People are possibly attracted by the flourishing economy, perfect public facilities and services, which is finished through some policies in the areas where the population density is relatively low. The optimization of population distribution can alleviate the problem of air haze pollution caused by excessive population density [22].

(3) It is obvious that AA can effectively restrain the PM$_{2.5}$ pollution. Therefore, it is important to create city parks and green belts being carried out in our country [4]. It is worth noting that the negative effect of SAMC on PM$_{2.5}$ was strengthening due to enforcement of the government policy on straw burning from 2015 to 2017. However, the positive effect of SAMC on PM$_{2.5}$ was aggravated in 2018. Agriculture plays an important role in Henan Province, even in China, producing a huge number of crop straw. Therefore, it needs to further strengthen the research and development of crop straw processing technology, and expand crop straw using channels [23].

Acknowledgements
This work was supported by the Key Scientific and Technological Project of Henan Province (Grant No. 202102310338), the Natural Science Foundation of Jiangsu Province (Grant No. BK20181059) and the Young Backbone Teachers Assistance Scheme of Xinyang Normal University (Grant No. 2019GGJS-07).

References
[1] Lu, Y., Wang, Y., Li, B., Li, J., and Jiang, H. (2020) Temporal and spatial variations in haze research: a bibliometric analysis. Environmental Reviews, 28(1): 12-20.
[2] Wan, Y., Li, Y., Liu, C., and Li, Z. (2020) Is traffic accident related to air pollution? A case report from an island of Taihu Lake, China. Atmospheric Pollution Research, 11(5): 1028-1033.
[3] Matus, K., Nam, K.-M., Selin, N.E., Lamsal, L.N., Reilly, J.M., and Paltsev, S. (2012) Health damages from air pollution in China. Global Environmental Change, 22(1): 55-66.
[4] Wu, J., Zhang, P., Yi, H., and Qin, Z. (2016) What Causes Haze Pollution? An Empirical Study of PM$_{2.5}$ Concentrations in Chinese Cities. Sustainability, 8(2): 132.
[5] Li, M. and Zhang, L. (2014) Haze in China: Current and future challenges. Environmental Pollution, 189: 85-86.
[6] Pope, C.A., III, Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K., and Thurston, G.D. (2002) Lung Cancer, Cardiopulmonary Mortality, and Long-term Exposure to Fine Particulate Air Pollution. JAMA: Journal of the American Medical Association, 287(9): 1132-1141.
[7] Franklin, M., Zeka, A., and Schwartz, J. (2006) Association between PM2.5 and all-cause and specific-cause mortality in 27 US communities. Journal of Exposure Science & Environmental Epidemiology, 17(3): 279-287.
[8] Lu, D., Xu, J., Yang, D., and Zhao, J. (2017) Spatio-temporal variation and influence factors of PM$_{2.5}$ concentrations in China from 1998 to 2014. Atmospheric Pollution Research, 8(6): 1151-1159.
[9] Gang, L., Fu, J., Dong, J., Hu, W., Dong, D., Huang, Y., and Zhao, M. (2014) Spatio-Temporal Variation of PM$_{2.5}$ Concentrations and Their Relationship with Geographic and Socioeconomic Factors in China. International Journal of Environmental Research and Public Health, 11:173-186.
[10] Wang, Y., Liu, C., Wang, Q., Qin, Q., Ren, H., and Cao, J. (2021) Impacts of natural and socioeconomic factors on PM$_{2.5}$ from 2014 to 2017. J Environ Manage, 284(1): 112071.
[11] Yang, D., Wang, X., Xu, J., Xu, C., Lu, D., Ye, C., Wang, Z., and Bai, L. (2018) Quantifying the influence of natural and socioeconomic factors and their interactive impact on PM$_{2.5}$ pollution in China. Environ Pollut, 241:475-483.
[12] Tian, S. and Zhao, P. (2016) An Empirical Study on the Dynamic Comprehensive Evaluation Model of Six Central Provinces' Haze Pressure in China. Statistics & Information Forum, 31(9): 56-61. (in Chinese with English abstract)
[13] Li, M. (2017) Temporal-Spatial Characteristics and Factors of The Haze Pollution in Henan Province. Chongqing: Chongqing University. (in Chinese with English abstract)

[14] Wang, C. (2016) Analysis on Differences and Similarities of Haze Formation and Continuous Meteorological Elements in Henan Province. Journal of Agricultural Catastrophology, 6(2): 36-38. (in Chinese with English abstract)

[15] Dong, K. (2017) Study on the Influencing Factor of Haze Pollution in Zhengzhou. Kaifeng: Henan University. (in Chinese with English abstract)

[16] Du, X., Chu, M., and Yang, L. (2016) Comparative Analysis and Influencing Factors of PM$_{2.5}$ Pollution in Five Districts in Zhengzhou during Haze Seasons. Journal of Green science and technology, 7(14): 101-103,108. (in Chinese with English abstract)

[17] Li, X., Zheng, W., Yin, L., Yin, Z., Song, L., and Tian, X. (2017) Influence of Social-economic Activities on Air Pollutants in Beijing, China. Nephron Clinical Practice, 9(1): 314-321.

[18] Zhang, Y., Shuai, C., Bian, J., Chen, X., Wu, Y., and Shen, L. (2019) Socioeconomic factors of PM$_{2.5}$ concentrations in 152 Chinese cities: Decomposition analysis using LMDI. Journal of Cleaner Production, 218: 96-107.

[19] Song, Z., Deng, Q., and Ren, Z. (2020) Correlation and principal component regression analysis for studying air quality and meteorological elements in Wuhan, China. Environmental Progress & Sustainable Energy, 39(1): e13278.

[20] Lin, Y., Li, M., and Song, D. (2013) Causes and countermeasures of haze weather in Jiaozuo city. Journal of Henan Institute of Science and Technology, 41(5): 27-31. (in Chinese with English abstract)

[21] Dziuban, C.D. and Shirkey, E.C. (1974) When is a correlation matrix appropriate for factor analysis? Some decision rules. Psychological Bulletin, 81(6): 358-361.

[22] Li, Y., Zhu, K., and Wang, S. (2020) Polycentric and dispersed population distribution increases PM$_{2.5}$ concentrations: Evidence from 286 Chinese cities, 2001–2016. Journal of Cleaner Production, 248:119202.

[23] Zhang, L., Liu, Y., and Hao, L. (2016) Contributions of open crop straw burning emissions to PM$_{2.5}$ concentrations in China. Environmental Research Letters, 11(1): 014014.