Comprehensive Evaluation and Analysis of Spatial Dynamic Transition of Air Pollution

Li Fangfeng1, Cai Jieping2, Liao Zicen3
1-3Xinhua College of Sun Yat-sen University,Guanzhou,Guangdong, 510520, China
*Corresponding author’s e-mail: etlf@xhsysu.edu.cn

Abstract. The prevention and control of air pollution has a long way to go. This paper makes a comprehensive evaluation of air pollution and explores its spatial distribution characteristics in order to provide suggestions for environmental control. In this paper, a comprehensive evaluation index is constructed by factor analysis method, and Moran's I index is calculated to analyze the spatial distribution and dynamic transition of air pollution. The results show that there is a positive spatial correlation of air pollution in China, and it has a high degree of spatial stability. It is possible to alleviate the pressure on the general atmospheric environment by improving the environment in highly polluted areas.

1. Introduction
With the rapid development of China's economy, environmental pollution has become more and more prominent. In order to improve air pollution, the departments of ecological environment have taken a series of prevention and control measures, and have made certain achievements. Judging from the long-term effect of air pollution prevention and control in China, the work of air pollution prevention and control has a long way to go.

In the face of increasingly serious ecological and environmental problems, the report of the 19th National Congress of the Communist Party of China clearly points out that the construction of ecological civilization should be vigorously promoted, and the concept of urban green development should be fully implemented. Based on this, this paper is to make a comprehensive assessment of the air pollution in different areas and to explore its spatial distribution characteristics, so as to provide some references for the related departments.

2. Materials and Methods
There are many indexes of air pollution, and it is one-sided to evaluate only by single index of air pollution. The commonly used evaluation index is air quality index (AQI). But with the development of science and technology, it’s easy to get the result with the help of statistical software. Factor analysis method is used in the paper, and the factor comprehensive score is taken as the comprehensive index of air pollution assessment. Considering the vast territory of our country, the different geographical conditions, development level and industrial structure of different regions, the environmental pollution situation must have different characteristics, so Moran’s I Index is used to test the spatial distribution pattern of environmental pollution.

2.1. Factor analysis
Factor analysis is a statistical method dealing with dimensionality reduction. Its mathematical model is as follows:
Where $x_i$ is the $i$th variable, $F_i$ is the $i$th principal factor and $\xi_k$ is the special factor. The main steps of factor analysis are as follows:

1. Data check. The Bartlett’s test and KMO test were used to determine whether the data were suitable for factor analysis. The Bartlett’s test is based on the Correlation Coefficient Matrix of variables, assuming that the correlation Coefficient Matrix is a unit matrix. If the Correlation Matrix is a unit Matrix, the variables are independent and can't do a factor analysis. The KMO test is used to test the partial correlation among variables. The closer the statistic value is to 1, the stronger the partial correlation is, and the better the effect of factor analysis is. When KMO is above 0.7, the effect of factor analysis is good; when KMO statistic is below 0.5, the effect of factor analysis is bad.

2. Factor extraction. In this paper, the cumulative contribution rate of model variance is over 85% as the standard to extract common factor.

3. Rotation.

4. Calculate the factor score

The calculation of comprehensive score with $p$ principal factors was defined as:

$$F = \sum_{i=1}^{p} \frac{\lambda_i}{\sum_{i=1}^{n} \lambda_i} F_i$$

Where $F$ represents the Composite Index, $F_i$ represents the $i$th principal factor and $\lambda_i$ represents the eigenvalues corresponding to the $i$th principal factor.

### 2.2. Moran’s I index

#### 2.2.1. Global Moran Index

Spatial autocorrelation analysis is a common analysis method in ecology and soil science, which is mainly used to test whether spatial variables have spatial dependence. The Global Moran Index is the most commonly used to test space Autocorrelation\[1\][2][4]. The formula is as follows:

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

Where $x_i$ and $x_j$ are the value of x on the adjacent pairing space points $i$ and $j$, $n$ is the total number of space units, $W_{ij}$ is a neighboring weight. $W_{ij}$ is a weight Matrix based on distance:

$$W_{ij} = \begin{cases} 1, & \text{when region } I \text{ is adjacent to region } J, \\ 0, & \text{when region } I \text{ is not adjacent to region } J, \\ 0, & \text{where } i = j. \end{cases}$$

The Moran Index is only an important index in the measurement of spatial correlation. It needs to be verified and synthetically considered:

$$z = \frac{I - E(I)}{\sqrt{V(I)}}$$
Where $E(I)$, $V(I)$ are the expectation and variance of I.

### 2.2.2. Local Moran Index

The local Moran Index is used to see if the number of strongly correlated regions increases over time. The Moran Index of region I is used to measure the degree of correlation between region I and its neighborhood:\(^5\):

$$I_i = \frac{n(x_i - \bar{x}) \sum_{j \neq i}^n W_{ij} (x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$  \hspace{1cm} (5)

The symbols in the expression are the same as the Global Moran Index.

When $I_i$ is greater than 0, the high value is surrounded by the high value, and the low value by the low value, marked as "high-high" or "low-low" respectively. When $I_i$ is less than 0, the high value is surrounded by the low value, and the low value is surrounded by the high value, marked "high-low (HL)" or "low-high (LH)" respectively. By plotting Moran’s scatterplot, the local spatial clustering can be visualized.

### 3. Analysis of air pollution in China

In this paper, the air pollution index data were collected, including PM2.5, PM10, SO2, NO2, Co and O3, which data of 168 cities in China from June 2018 to June 2020. Factor analysis is carried out in this paper. Table 1-4 below shows the results of factor analysis with SPSS for June 2020.

The P value of the Bartlett’s test is 0, reject the null hypothesis of independence of variables at the 0.05 significance level. And KMO is 0.837. The results show that this set of data is very suitable for factor analysis. Except for SO2 and Co, the communalities of each variable is above 0.9, which indicates that each factor has strong explanatory power. The cumulative contribution rate of the first three principal factors to the model is 88.958%, which shows that the information of the original data can be interpreted by extracting the three principal factors.

| Table 1. KMO and Bartlett's test |
|----------------------------------|
| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | .837 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 731.173 |
| df | 15 |
| Sig. | .000 |

| Table 2. Communalities |
|------------------------|
| Initial | Extraction |
| PM2.5 | 1.000 | .932 |
| PM10 | 1.000 | .913 |
| so2 | 1.000 | .776 |
| no2 | 1.000 | .963 |
| co | 1.000 | .833 |
| o3 | 1.000 | .921 |
Table 3. Total Variance Explained

| Component | Initial Eigenvalues | Extraction Sums of Squared Loadings | Rotation Sums of Squared Loadings |
|-----------|---------------------|--------------------------------------|----------------------------------|
|           | Total               | % of Variance | Cumulative % | Total               | % of Variance | Cumulative % | Total               | % of Variance | Cumulative % |
| 1         | 3.850               | 64.159       | 64.159       | 3.850               | 64.159       | 64.159       | 2.614               | 43.571       | 43.571       |
| 2         | 0.881               | 14.683       | 78.842       | 0.881               | 14.683       | 78.842       | 1.632               | 27.208       | 70.779       |
| 3         | 0.607               | 10.115       | 88.958       | 0.607               | 10.115       | 88.958       | 1.091               | 18.179       | 88.958       |
| 4         | 0.432               | 7.198        | 96.155       | 0.432               | 7.198        | 96.155       | 0.879               | 14.488       | 100.000      |
| 5         | 0.130               | 2.167        | 98.322       | 0.130               | 2.167        | 98.322       | 0.130               | 2.167        | 98.322       |
| 6         | 0.101               | 1.678        | 100.000      | 0.101               | 1.678        | 100.000      | 0.101               | 1.678        | 100.000      |

Table 4. Component Score Coefficient Matrix

| Component | 1  | 2  | 3  |
|-----------|----|----|----|
| PM2.5     | 0.430 | -0.054 | -0.171 |
| PM10      | 0.362 | -0.023 | -0.065 |
| so2       | -0.245 | 0.582 | 0.177 |
| no2       | -0.205 | -0.182 | 1.120 |
| co        | -0.092 | 0.731 | -0.334 |
| o3        | 0.475 | -0.210 | -0.084 |

According to the analysis, the formula of the comprehensive index of environmental pollution is given in Table 3:

$$F = \frac{3.85F_1 + 0.881F_2 + 0.607F_3}{5.338}$$

The average monthly comprehensive pollution index of cities in 2018-2020 is in table 5, ranking see figure 1, figure 2.

Table 5. Pollution index value and ranking

| Province | City      | comprehensive pollution index | Ranking |
|----------|-----------|------------------------------|---------|
|          |           | 2018 | 2019 | 2020 | 2018 | 2019 | 2020 |
| Xizang   | Lasa      | -1.10004 | -1.50634 | -1.44358 | 5 | 2 | 1 |
| Hainan   | Haikou    | -1.42555 | -1.52351 | -1.37104 | 1 | 1 | 2 |
| Guangdong | Shenzhen | -0.9345 | -1.01463 | -1.2165 | 8 | 7 | 3 |
| Guangdong | Zhuhai    | -0.76911 | -0.99184 | -1.20279 | 17 | 9 | 4 |
| Guangdong | Huizhou   | -0.82008 | -0.89348 | -1.12292 | 13 | 12 | 5 |
| Zhejiang | Lishui    | -1.0138 | -0.95025 | -1.03595 | 6 | 10 | 6 |
| Anhui    | Huangshan | -1.33208 | -1.06552 | -1.02143 | 2 | 5 | 7 |
| Zhejiang | Zhouran   | -1.21876 | -1.23667 | -0.99333 | 3 | 3 | 8 |
| Henan    | Jiaozuo   | 0.951117 | 0.935991 | 0.936717 | 158 | 162 | 160 |
| Shanxi   | Taiyuan   | 1.120854 | 0.83207 | 0.937953 | 161 | 155 | 161 |
According to the relative ranking changes over three years (the bigger the ranking, the worse the pollution is), 38 cities, including Dongguan, Foshan, Guangzhou, Huizhou, Jiangmen, Shenzhen, Zhaoqing, Zhongshan and Zhuhai, have seen their environmental pollution situation improve year by year, especially Hefei, Foshan and Zhaoqing (see Chart 1). In 26 cities, including Huangshan City, Fuzhou, Hengshui, Hebi, Zhengzhou, Harbin, Huanggang, Huangshi and Xiaogan, air pollution has worsened to a certain extent, especially Harbin, Changchun, Deyang, Luzhou and Meishan (see figure 2).

The monthly average comprehensive index data of each year are used for cluster analysis, and the pollution situation of each region is divided into 5 categories. The results of the lowest pollution level (Group 1) and the highest pollution level (Group 5) are shown in Table 6.

| case       | group | case | group | case | group |
|------------|-------|------|-------|------|-------|
| 1:Anhui    | 1     | 56:Hainan | 1     | 74:Henan | 5     |
| 2:Fujian   | 1     | 156:Xizang | 1     | 75:Henan | 5     |
| 24:Sichun  | 1     | 162:Zhejiang | 1     | 135:Shanxi | 5     |
| 43:Fujian  | 1     | 167:Zhejiang | 1     | 136:Shanxi | 5     |
| 48:Guangdong | 1     | 61:Hebei | 5     | 137:Shanxi | 5     |
| 50:Guangdong | 1     | 63:Hebei | 5     | 137:Shanxi | 5     |
| 53:Guangdong | 1     | 64:Hebei | 5     |       |       |
| 55:Guizhou | 1     | 65:Hebei | 5     |       |       |
Fujian, Hainan, Sichuan, Tibet, Guizhou and Guangdong have good air quality, while Hebei and Shanxi, regions with large national economy and developed resource-based industries, are consistently ranked high in overall environmental pollution.

Using the formula (3), the monthly Environmental Pollution Moran Index of 168 cities in China is calculated. As can be seen from table 7 and figure 3, although the Moran Index declined slightly from June 2018 to June 2020, the Moran Index for each month was above 0 and passed the test at the 1% significance level. The results show that there is a positive correlation in the spatial distribution of environmental pollution in our cities, that is, there is a spatial agglomeration of air pollution in China.

In order to find out where the spatial concentration of air pollution exists, the local Molain index was calculated and the scatter diagram of Moran\(^5\) was drawed. Quadrants 1-4 of the Moran scatterplot are HH, LH, LL, HL, where HH indicates that the area is heavily polluted and surrounded by highly polluted areas; LH indicates that the area is low polluted and surrounded by highly polluted areas, and so on.

| time  | Moran's I  | z     | p  | time  | Moran's I  | z     | p  |
|-------|------------|-------|----|-------|------------|-------|----|
| Jun-18| 0.29667    | 25.1  | 0  | Jul-19| 0.30115    | 25.493| 0  |
| Jul-18| 0.2755     | 23.389| 0  | Aug-19| 0.14188    | 12.314| 0  |
| Aug-18| 0.2388     | 20.352| 0  | Sep-19| 0.19951    | 17.059| 0  |
| Sep-18| 0.21445    | 18.308| 0  | Oct-19| 0.23953    | 20.399| 0  |
| Oct-18| 0.23891    | 20.326| 0  | Nov-19| 0.24352    | 20.697| 0  |
| Nov-18| 0.27822    | 23.584| 0  | Dec-19| 0.24585    | 20.928| 0  |
| Dec-18| 0.23723    | 20.196| 0  | Jan-20| 0.29287    | 24.81 | 0  |
| Jan-19| 0.21098    | 18.02 | 0  | Feb-20| 0.26743    | 22.748| 0  |
| Feb-19| 0.23332    | 19.909| 0  | Mar-20| 0.15477    | 13.372| 0  |
| Mar-19| 0.26146    | 22.221| 0  | Apr-20| 0.12031    | 10.7  | 0  |
| Apr-19| 0.28321    | 24.019| 0  | May-20| 0.20594    | 17.647| 0  |
| May-19| 0.3049     | 25.794| 0  | Jun-20| 0.35274    | 29.747| 0  |

According to Moran Scatterplot, there are about 140 regions in "HH" and "LL" regions, accounting for more than 80% of the total number of regions. It shows that there is a strong positive spatial correlation of air pollution in different regions of China from 2018 to 2020. Among them, 63 cities in Henan Province, Inner Mongolia Province, Hebei Province, Beijing City, Shandong Province, Anhui Province, Shanxi Province, Gansu Province, Tianjin City, Shaanxi Province, Xinjiang province, Hubei Province and Jiangsu Province have been in the HH region for three years.
The space-time transition of air pollution can be divided into four types, as shown in Table 8.

| Type                        | Performance |
|-----------------------------|-------------|
| Adjacent space adjacent transition | HH→HL, HL→HH, LH→LL, LL→LH |
| Relative shift transition   | HH→LH, HL→LH, LH→HH, LL→HL |
| Space Integral Transition   | HH→LL, HL→LH, LH→HL, LL→HH |
| Keep the space level        | HH→HH, HL→HL, LH→LH, LL→LL |

According to the local Moran value, there are four cities in “Adjacent space adjacent transition”, which are Chuzhou, Lhasa, Xiaogan and Yichang; 19 cities in “Relative shift transition”, including Banbu, Chengdu, Deyang, Harbin, Taizhou, Wuhan, Xinyang, Suqian, Yinchun, Changchun, Zigong; 1 city in “Space Integral Transition”, which is Jinzhou, and its air pollution has improved. 144 cities in “Keep the space level”, accounting for 85.7% of the total, which shows that there is a high spatial stability of air pollution in China.
4. Conclusion
From the analysis, it can be seen that the provinces with a lower overall level of air pollution in China are Fujian, Hainan, Sichuan, Tibet, Guizhou and Guangdong, while Hebei and Shanxi, where the proportion of the national economy is relatively large and the resource-based industries are relatively developed, the overall ranking of environmental pollution remains high. The air pollution in China is highly stable and positively correlated in space. The high pollution area is surrounded by the high pollution area, while the low pollution area is surrounded by the low pollution area. In order to effectively solve the problem of air pollution, we can improve the pollution situation in highly polluted areas to achieve the impact on surrounding cities.

Acknowledgments
This paper is supported by the project of young creative talents of Guangdong Education Department (Natural Science)—Application of energy distance in statistical inference (2018KQNCX366) and the project of "public management" of Key discipline projects with characteristics of Guangdong Province in 2016.

References
[1] Liu,H.,Du,G.(2016) Spatial Pattern and Distributional Dynamics of Urban Air Pollution in China—An Empirical Study Based on Aqi and Six Sub-Pollutants of 161 Cities. ECONOMIC GEOGRAPHY, 10:33-38.
[2] Ma,L. Zhang,X. Spatial Effects of Regional Air Pollution and the Impact of Industrial Structure. CHINA POPULATION,RESOURCES AND ENVIRONMENT, 7:157-163.
[3] Xiang,K., Song,D. (2015) Spatial Analysis of China’s PM2.5 Pollution at the Provincial Level. CHINA POPULATION,RESOURCES AND ENVIRONMENT, 9:153 - 159.
[4]Yang, H., Liu,J. (2020) Study on the Spatial Agglomeration Effect of Regional Air Pollution. Technology and Economy of Changjiang, 4(2):7-8.
[5] Zhang,L. (2017) Temporal and Spatial Variation Characteristics of Air Quality and its Influencing Factors in Shandong Peninsula Urban Agglomeration. Computer Knowledge and Technology,17: 207-210.