Comparison And Analysis Of Air Quality In Major Cities Of Environmental Protection

Yunpeng Chen¹, Ling Jiang ² and Hui Shi ¹

¹Standard Intelligence Institute  China National Institute of Standardization  Beijing, China
²Department of Data Science and Intelligent Media Communication University of China Beijing, China

chenyp@cnis.gov.cn; cuc_jl@cuc.edu.cn; shih@cnis.gov.cn;

Abstract. As one of the most important environmental problems, air quality has always been highly valued by the government and society. The analysis of air quality in key cities can provide a basis for how to evaluate air quality and how to classify air quality levels and from what aspects it is helpful to improve air quality. The air quality of 51 key cities of environmental protection in China is analyzed and evaluated by selecting 7 indicators of air quality as the original variables. Cluster analysis was used to classify the key environmental protection cities preliminarily, and it was found that the air quality of the cities with fast development level was relatively poor, and the air quality of coastal cities was similar. It is found that the air quality of northern cities is generally worse than that of southern cities, and that of inland cities is generally worse than that of coastal cities. At last, according to the research results, some reasonable conclusions and suggestions are given.

1 Introduction

1.1 The research background

Since the reform and opening up, the economy continues to develop, the pace of industrialization is getting faster and faster, China's air pollution problem is becoming more and more serious. Especially in recent years, many cities have seen serious sandstorms, haze and other bad weather. It not only seriously hinders residents' travel, but also causes great harm to residents' health, which is not conducive to the sustainable development of society. Air quality reflects the extent of air pollution. In order to control air pollution sources and prevent air pollution, we need to study the air quality in key cities of environmental protection.

1.2 Data source and index interpretation

We selected 7 indicators to reflect the air quality, namely: x₁ -- annual average concentration of sulfur dioxide (µg/m³); x₂ -- average annual concentration of nitrogen dioxide (µg/m³); x₃ -- annual average concentration of inhalable particles (PM10) (µg/m³); x₄ -- carbon monoxide concentration at the 95th percentile per day (µg/m³); x₅ -- 90th percentile concentration of ozone (O₃) at maximum 8 hours per day (µg/m³); x₆ -- annual average concentration of fine particles (PM2.5) (µg/m³); x₇ -- the number of days (days) where air quality is at or above level 2. The selection of indicators refers to the setting of indicators in China statistical yearbook. Data are from China statistical yearbook in 2013.

2 Analysis

2.1 Cluster analysis

The first step is to standardize the data to eliminate the effects of dimensional differences. After that, standardized scores were stored as variables, and cluster analysis was conducted on these variables. The aggregation coefficient in the table of clustering process was used to make the curve of aggregation coefficient changing with classification number by Excel, as shown in Figure 1.

Figure 1  The curve of aggregation coefficient with classification number
As can be seen from the Figure 1, when the classification number is 6, the curve starts to flatten, so we should divide the 51 cities into 6 categories. At this time, the clustering number is set as 6, from which specific classification conditions can be obtained:

**Category1:** Beijing, Shanghai, Hangzhou, Wenzhou, Jiaxing, Huzhou, Jinhua, Wuhu, Guangzhou, Foshan, Dongguan, Zhongshan, Chongqing, Chengdu (14 cities)

**Category2:** Tianjin, Zhengzhou, xi’an, Urumqi (4 cities)

**Category3:** Shijiazhuang (1 city)

**Category4:** Taiyuan, Hohhot, Shenyang, Changchun, Harbin, Nanjing, Shaoxing, Hefei, Nanchang, Qingdao, Changsha, Lanzhou, Xining, Yinchuan (14 cities)

**Category5:** Ningbo, Quzhou, Zhoushan, Taizhou, Lishui, Fuzhou, Xiamen, Shenzhen, Zuhai, Jiangmen, Zhaoqing, Huizhou, Nanning, Haikou, Guiyang, Kunming, Lhasa (17 cities)

**Category6:** Jinan (1 city)

According to the results, cities with high development level cluster together, coastal cities cluster together, and cities with moderate development level cluster together.

### 2.2 Factor analysis

Before factor analysis, the seven variables to be analyzed should be determined whether they are suitable for factor analysis on the basis of standardized data (starting from the covariance matrix). Factor analysis requires strong correlation between variables, otherwise common factors that can reflect common characteristics between variables cannot be obtained. Common correlation tests are: correlation matrix, KMO and Bartlett sphericity tests, etc. The results of correlation tests are shown in Table 1 and Table 2.

If most of the correlation coefficients in the correlation coefficient matrix are less than 0.3 and fail the statistical test, then these variables are not suitable for factor analysis. However, most of the correlation coefficients in Table 1 are greater than 0.3, so these variables are suitable for factor analysis, and there is a large overlap of information reflected by indicators. In the KMO test, when the sum of squares of simple correlation coefficients among all variables is much larger than the sum of squares of partial correlation coefficients, the value of KMO tends to 1. The closer the value of KMO is to 1, the stronger the correlation between variables and the more suitable for factor analysis. The KMO test value in Table 2 is 0.837, which is suitable for factor analysis. In Bartlett's sphericity test, if the measurement value in the correlation coefficient matrix is large and its associated probability is less than the significance level, it is suitable for factor analysis. Table 2 shows that the test value of sphericity is 353.670, and the p value is 0, less than 0.05, so the null hypothesis is rejected and suitable for factor analysis.

### Table 1 Correlation coefficient matrix

|   | Zx1 | Zx2 | Zx3 | Zx4 | Zx5 | Zx6 | Zx7 |
|---|-----|-----|-----|-----|-----|-----|-----|

### Table 2 KMO and Bartlett's sphericity test

| KMO | Bartlett's sphericity test | chi-square | DOF | Significance |
|-----|-----------------------------|------------|-----|--------------|
| .837| 353.670                     | 21         | .000|              |

Table 3 shows that the variance contribution rate of the first three factors is 90.508%, more than 85%, indicating that extracting the three common factors can explain most information of the original variables, so the three factors should be selected and retained in general.

### Table 3 Eigenvalues and variance contribution rates

| Extract the sum of squares of load | Initial Eigenvalue | Cumulative percentage of variance | Cumulative percentage of variance |
|------------------------------------|--------------------|----------------------------------|----------------------------------|
| 1                                  | 4.742              | 67.739                           | 67.739                           |
| 2                                  | 1.104              | 15.773                           | 83.512                           |
| 3                                  | .490               | 90.508                           | 90.508                           |
| 4                                  | .310               | 4.422                            | 94.929                           |
| 5                                  | .247               | 3.530                            | 98.459                           |
| 6                                  | .069               | .987                             | 99.446                           |
| 7                                  | .039               | .554                             | 100.00                           |

Finally, factor scores were calculated, and the ratio of variance contribution rate of each factor to the total variance contribution rate of the three factors was taken as the weight for weighted summary, and the comprehensive score F of each city was obtained. According to Table 4, namely:

\[ F = \left(43.969 \times F_1 + 30.794 \times F_2 + 15.745 \times F_3\right) / 90.508 \]

The operation result is obtained and the comprehensive score is calculated. If the comprehensive score is positive, the air quality is poor; if the comprehensive score is negative, the air quality is good. It can be seen from the operation results that Shijiazhuang is the key city with the highest comprehensive score for environmental protection, while Haikou is the key city with the lowest comprehensive score, indicating that among the 51 cities, Shijiazhuang has the worst air quality and Haikou has the best air quality. According to the comprehensive score, the
Conclusion is not quite consistent with the previous stage. For example, the air quality of cities with high economic development level, such as Shenzhen and Zhuhai, is not very bad.

Table 4  Total variance explained

| Component | Initial Eigenvalues | Extract the sum of squares of loads | Sum of squares of rotating loads |
|-----------|---------------------|-------------------------------------|----------------------------------|
|           | Percentage of variance | cumulative % | Percentage of variance | cumulative % | Percentage of variance | cumulative % |
| 1         | 4.7                  | 67.7                  | 4.7                  | 67.7                  | 3.0                  | 43.9                  |
|           | 39                   | 67.739                | 39                   | 67.739                | 78                   | 69                   |
| 2         | 1.1                  | 15.7                  | 1.1                  | 15.7                  | 2.1                  | 30.7                  |
|           | 04                   | 73                    | 04                   | 73                    | 56                   | 94                   |
| 3         | 0.4                  | 6.9                   | 0.4                  | 6.9                   | 1.1                  | 15.7                  |
|           | 90                   | 90.508                | 90                   | 90.508                | 02                   | 45                   |
| 4         | 0.3                  | 4.4                   | 0.3                  | 4.4                   | 22                   | 94.929                |
|           | 10                   |                       |                      |                       |                      |                      |
| 5         | 0.2                  | 3.5                   | 0.2                  | 3.5                   | 30                   | 98.459                |
|           | 47                   |                       |                      |                       |                      |                      |
| 6         | 0.0                  | 0.9                   | 0.0                  | 0.9                   | 87                   | 99.446                |
|           | 69                   |                       |                      |                       |                      |                      |
| 7         | 0.0                  | 0.5                   | 0.0                  | 0.5                   | 54                   | 100.00                |
|           | 39                   |                       |                      |                       |                      |                      |

These 51 cities are graded according to their comprehensive scores, as shown in Table 5.

Table 5  Urban air quality rating

| air quality is very bad | air quality is poor | average air quality | air quality is very good |
|-------------------------|---------------------|---------------------|--------------------------|
| Shijiazhuang            | Wuhan               | Shaoxing            | Guiyang                  |
| Jinan                   | Huzhou              | Qingdao             | Kunming                  |
| Zhenzhou                | Nanjing             | Nanchang            | Lishui                   |
| Xi’an                   | Hohhot              | Foshan              | Huizhou                  |
| Tianjin                 | Xining              | Lanzhou             | Xiamen                   |
| Taiyuan                 | Jiaxing             | Hefei               | Zhuhai                   |
| Chengdu                 | Jinhua              | Harbin              | Shenzhen                 |
| Shenyang                | Yinchuan            | Zhaqing             | Lhasa                    |
| Beijing                 | Changchun           | Shanghai            | Zhoushan                 |
| Urumqi                  | Hangzhou            | Wenzhou             | Fuzhou                   |
| Changsha                | Quzhou              | Haikou              |                          |
| Chongqing               | Jiangmen            | Guangzhou           |                          |
|                          | Dongguan            | Taizhou             | Ningbo                   |
|                          | Zhongshan           | Nanning             |                          |

3 Conclusion

This paper evaluates the air quality in 51 key environmental protection cities. According to the research, the air quality in key cities of environmental protection varies greatly. Generally speaking, air quality in northern cities is generally worse than that in southern cities. The reason may be that most cities in the north have more developed industries than those in the south, less vegetation coverage, more dust and sand, serious soil erosion, dry climate and less rainfall. The vegetation coverage rate in the south is higher than that in the north, more rainfall, and the air quality is naturally better than that in the north. Despite the difference between north and south, air quality is generally better in coastal cities than inland ones. The reason may be that economic conditions are good and the government has the ability to better manage the environment. What's more, the coastal climate is good, the precipitation is abundant, and the sea with strong self-purification capacity and large sewage discharge capacity can be used as a clean source.
In key cities with poor air quality and poor environmental protection, priority should be given to improving air quality, reducing emissions of harmful substances in the air, and improving residents' lives and environment. And cities with average air quality and better air quality cannot relax and continue to adhere to measures conducive to improving air quality so as to better realize the harmonious coexistence of human and nature.

Acknowledgement

This research was financially supported by Research on Measurement Method Of Standard Adoption Impact Based on Five-Dimensional Model (292019Y-6690)

References

1. Qiu-Hong Wang, Qiao Zhao. Research on the comprehensive development level of 35 central cities in China[J]. *Productivity Research*, 2017(6): 97-101.
2. Xiao-qiu He. *Multivariate statistical analysis (4th edition)*. China Renmin University Press: Beijing, China, 2015.
3. Si-ying Li. Analysis and research on air quality of major cities in China: In 2013, 51 key environmental protection cities were cited as examples[J]. *Technology Innovation and Application*, 2016(13).
4. Wen-chang Zhao. Study on the health risk and economic loss of urban residents caused by air pollution[D]. Shanghai Jiaotong University, 2012.
5. Chun-ying Zhou, Tao-zhen Huang. Comprehensive evaluation of social and economic development level of central cities[J]. *COMMERCIAL RESEARCH*, 2005(19): 37-40.
6. Qian-qian Yang, Zhong-li Li. Performance evaluation of new energy listed companies based on factor analysis[J]. *TECHNOLOGY AND INNOVATION MANAGEMENT*, 2013(2).
7. Xuan Yang, Guo Tang, Xiang He. Cluster analysis for charity development of cities in China: Taking 15 coastal cities as the example[J]. *Science-Technology and Management*, 2013(3): 1-4.