Advances in Engineering Research, volume 170
7th International Conference on Energy and Environmental Protection (ICEEP 2018)

Research on Real-time Monitoring of Urban Air Quality Based on Optimal Scale and Factor Analysis

Chun Chen[1],[a] Xiang Gao[2,b] Qian Mou [3,c]

([1]Professor, Business School, Sichuan Normal University, Chengdu 610101, China)

([2]Postgraduate student, Information School, the University of Sheffield, Sheffield S10 2TN, the United Kingdom)

([3]Postgraduate student, Business School, Sichuan Normal University, Chengdu 610101, China)

[a]chenchun@sina.com [b]2846857801@qq.com [c]1369053026@qq.com

Keywords: Air quality; Optimal scale; Multivariate statistical analysis

Abstract. In this paper, based on the 190 monitored city data released by the China National Environmental Monitoring Center in 2018, the optimal scale and factor analysis methods were selected to conduct in-depth research and analysis of urban air quality. The study found that: PM2.5 has a serious impact on air quality, and PM10, ozone, SO2, and NO2 are important factors affecting air quality. At present, cities with good air quality include Sanya, Haikou, Zhanjiang, Maoming, Beihai, Zhuhai, Lhasa, Shenzhen, Heyuan, etc.; cities with severe air pollution include Texas, Tianjin, Zaozhuang, Linyi, Handan, Tangshan, and Hengshui. Finally, based on the conclusions of the study, we propose strategies for effective improvement and control of air pollution.

Introduction

With the rapid development of the industry and transportation industry in China, a large amount of toxic dust is discharged into the air, causing air pollution, affecting the ecological environment. How to effectively control air pollution and improve air quality has become a matter of high concern for people now. To improve air quality, it is necessary to first monitor urban air quality in real time and find causes of air pollution from initial sources so as to achieve targeted environmental management. At present, 190 cities have established thousands of air quality monitoring stations. They continuously monitor air quality in 24 hours and publish PM2.5, PM10, CO, NO2, SO2, and ozone in real time. This study is based on the domestic monitoring of air quality in real-time monitoring of 190 cities, using factor analysis and optimal standards for air quality evaluation and research, hoping to reduce air pollution, improve and promote the development of ecological cities in China.

Factor Analysis and Optimal Scale Analysis

Data Selection and Processing

The data is taken from the 190 monitored city released by the China Environmental Monitoring Center of the Ministry of Environmental Protection. The data is from the PM2.5 data network. The 190 cities' air quality ranking data and the number of days of compliance were selected and released in March 2018, and the number of days of compliance was split into excellent and good days. Table 1 shows the actual data of 190 monitoring urban air quality variables, among which the first seven data are urban air quality monitoring data, which are PM2.5, fine particles, PM10 suction particles, carbon monoxide, nitrogen dioxide, sulfur dioxide, Ozone 1 hour's average and ozone 8 hours' average. AQI is the air quality index, which is also commonly used by developed countries in the air quality evaluation standards. AQI uses the piecewise linear function to convert the monitoring data into calculations. Through the AQI size, international or domestic city rankings can be obtained, and AQIs can also be divided into several types of grades: AQI $\leq$ 50 when the air quality is excellent, 50 <AQI $\leq$ 100 is good, and then you can calculate the monthly or annual air quality days and good days.

Table 1 190 variables from monitored urban air quality

Name	PM2.5	PM10	CO
Content	fine particles	suction particles	carbon monoxide
Name	O3-8h	AQI	EGDAYS
Content	Ozone 8h mean	quality index	excellent days
Name	NO2	SO2	O3-1h
Content	nitrogen dioxide	sulfur dioxide	ozone 1h mean
Name	GDAYS	AQIR	CITY
Content	good days	ranking	cities

In order to facilitate the study, IBM SPSS Statistics 21 was used to test and standardize the air quality monitoring data of 7 cities. It was found that the data skewness was less than 1.5 and the absolute value of kurtosis was less than 4.5. Then the reliability of the standardized data is tested. The reliability statistic Cronbach's Alpha coefficient is 0.661, which means that the corresponding observation variable is highly reliable, and it can be used for factor analysis or optimal scale analysis.

Factor Analysis

Multivariate statistical analysis includes factor analysis and optimal scale analysis. Factor analysis can regroup original variables into new unrelated several comprehensive main variables. These new main variables keep the original variable information as much as possible. Optimal Scaling provides a multivariate correspondence analysis method that display multiple variable information in a low-dimensional plan. In multivariate statistical analysis, they can be grouped according to the size of the original variable's relevance, so that the correlation within the group variable is higher, and the correlation between different groups of variables is lower. Each group is represented by an unobservable comprehensive variable. Factor analysis method can obtain correlation coefficient matrix by SPSS, obtain common factor, total factor variance, factor load matrix through variable correlation, then use factor rotation to obtain factor score, and finally calculate comprehensive score by variance contribution rate, and obtain city air quality rankings.

Table2 shows the eigenvalue, variance contribution rate and cumulative contribution rate. According to the first two factors, the initial eigenvalue is greater than 1, the cumulative variance contribution rate is 72.3%, and the rotation square and cumulative variance contribution rate is 66.6%. Therefore, the first two are selected. These factors have already evaluated the overall level of urban air quality.

Table 2 Eigenvalues, variance contribution rates and cumulative contribution rates

Factor	Initial feature value			Rotation squared and loaded		
	total	variance %	accumulation %	total	variance %	accumulation %
1	3.108	44.404	44.404	2.601	37.155	37.155
2	1.953	27.905	72.308	2.061	29.446	66.601
3	.815	11.647	83.956			

Table 3 Rotation factor score coefficient matrix

Variable	The first factor	The second factor	Variable	The first factor	The second factor
PM2.5	.891	-.019	CO	.353	-.231
PM10	.859	.031	O_3-1h	-.074	.997
SO_2	.742	-.059	O_3-8h	-.071	.996
NO_2	.618	-.131			

Table 3 shows the twiddle factor score matrix. The first factor's scores from high to low are PM2.5, PM10, SO2, NO2, and CO. The second are O3-1h and O3-8h, respectively. The variables with greater influence on air quality are PM2.5, PM10, SO2, NO2, and CO. In addition, the variable that have a greater impact on air quality is ozone. The analysis results are also basically consistent with the actual results. The scores FAC1_1 and FAC2_1 can be automatically calculated in the database.

Calculating the city air quality total score function based on the weight variance contribution factor corresponding to the rotation factor is:

$$S1=(2.601FAC1\_1+2.061FAC2\_1)/(2.601+2.061) \qquad (1)$$

Table 4 lists the 190 city air quality index AQI rankings and comprehensive score rankings. We can see that the factor analysis comprehensive score ranking is basically the same as the national air quality index, indicating that the factor analysis comprehensive score is credible.

Table 4 City Air Quality Index AQI Ranking and Overall Score Ranking

CITY	AQIR ranking	Synthesis ranking	CITY	AQIR ranking	Synthesis ranking	CITY	AQIR ranking	Synthesis ranking	CITY	AQIR ranking	Synthesis ranking
Sanya	1	1	Chaozhou	49	52	Hohhot	97	167	Jinchang	145	104
Haikou	2	4	Zunyi	50	54	Rushan	98	41	Sanmenxia	146	146
Zhanjiang	3	6	Wendeng	51	33	Wuxi	99	116	Xianyang	147	153
Maoming	4	11	Zhang Qiu	52	120	Qingdao	100	110	Jintan	148	78
North Sea	5	5	Hangzhou	53	60	Zhangjiagang	101	100	Qinhuangdao	149	155
Zhuhai	6	8	Fushun	54	72	Changzhou	102	125	Suqian	150	135
Yunfu	7	29	Wuhu	55	50	Jieyang	103	57	Huai'an	151	83
Lhasa	8	2	Shanghai	56	40	Nanchong	104	95	Tongchuan	152	75
Shenzhen	9	17	Shaoxing	57	81	Chifeng	105	69	Zhengzhou	153	170
Heyuan	10	22	Haimen	58	19	Yantai	106	108	Pingdu	154	111
Jiujiang	11	32	Wenzhou	59	124	Wafangdian	107	53	Qingyuan	155	99
Zhoushan	12	3	Jiaxing	60	55	Suzhou	108	98	Jiayuguan	156	51
Huizhou	13	14	Yiwu	61	73	Lacey	109	71	Jingzhou	157	151
Meizhou	14	18	Zhuzhou	62	96	Yan'an	110	148	Weinan	158	145
Qiqihar	15	23	Jin Hua	63	47	Yibin	111	80	Yangquan	159	175
Shanwei	16	10	Kunshan	64	77	Taiyuan	112	150	Jiaozuo	160	174
Nanchang	17	31	Zhangzhou	65	36	Yinchuan	113	136	Urumqi	161	179
Guiyang	18	24	Panjin	66	123	Shizuishan	114	117	Kaifeng	162	130
Yangjiang	19	25	Dongguan	67	63	Jinzhou	115	142	Beijing	163	157
Daqing	20	9	Xiangtan	68	92	Hefei	116	74	Weifang	164	152
Qujing	21	26	Taicang	69	82	Fuyang	117	79	Jining	165	164
Zhongshan	22	48	Mianyang	70	66	Yue Yang	118	119	Heze	166	138
Quanzhou	23	16	Harbin	71	139	Yancheng	119	46	Xi'an	167	173
Fuzhou	24	37	Changsha	72	102	Benxi	120	162	Jinan	168	165
Shantou	25	27	Guilin	73	45	Xining	121	121	Laiwu	169	181
Kunming	26	42	Luzhou	74	88	Jiangyin	122	126	Cangzhou	170	147
Yuxi	27	61	Changde	75	76	Luoyang	123	156	Pingdingshan	171	154
Nanning	28	59	Linyi	76	141	Deyang	124	112	Liaocheng	172	177
Xiamen	29	20	Liuzhou	77	67	Maanshan	125	115	Binzhou	173	182
Jiangmen	30	62	Rongcheng	78	12	Nanjing	126	128	Yichang	174	169
Lishui	31	21	Chengde	79	101	Lanzhou	127	39	Dongying	175	180
Shaoguan	32	90	Huludao	80	134	Zhenjiang	128	118	Zibo	176	187
Dandong	33	35	Zhuji	81	70	Yangzhou	129	140	Shouguang	177	166
Panzhihua	34	93	Yixing	82	103	Shenyang	130	160	Anyang	178	178
Mudanjiang	35	43	Wu Jiang	83	84	Jiaonan	131	86	Langfang	179	176
Karamay	36	15	Zhangjiajie	84	30	Jurong	132	131	Dezhou	180	171
Datong	37	129	Huzhou	85	91	Taizhou	133	113	Tianjin	181	184
Zhangjiakou	38	97	Penglai	86	68	Lianyungang	134	106	Zaozhuang	182	168
Foshan	39	107	Zigong	87	58	Rizhao	135	127	Linyi	183	185
Chongqing	40	89	Changshu	88	87	Taian	136	163	Handan	184	183
Ordos	41	7	Lin'an	89	38	Anshan	137	159	Tangshan	185	189
Fuyang	42	56	Jiaozhou	90	109	Zhaoyuan	138	114	Hengshui	186	172
Taizhou	43	34	Changzhi	91	137	Baotou	139	143	Baoding	187	188
Jimo	44	49	Weihai	92	13	Xuzhou	140	132	Shijiazhuang	188	186
Jilin	45	44	Changchun	93	122	Wuhan	141	144	Xingtai	189	190
Ningbo	46	28	Dalian	94	105	Chengdu	142	149	Korla	190	158
Yingkou	47	65	Zhaoqing	95	133	Laizhou	143	94			
Guangzhou	48	85	Nantong	96	64	Baoji	144	161			

From Table 4, a few cities have different AQIR rankings and composite scores, and cities with higher comprehensive scores indicate that the comprehensive evaluation index of factor analysis is better than AQIR. The cities with high comprehensive scores are Erdos, Rongcheng, Weihai, Karamay, Zhangjiajie, etc. The cities with low overall scores include Jinchang, Yuxi, Jiangmen and Hohhot.

Optimal Scale Analysis

The study selects the optimal scale analysis in IBM SPSS Statistics 21, which provides homogeneity analysis, nonlinear correlation analysis and principal component analysis. It can do multiple correspondence analysis and will analyze multiple nonlinear variables in the low dimension. In order to facilitate the study with urban air quality, it is divided into three dimensions. Table 5 shows the scores of the principal components of variables. Analyzing the data in Table 5, PM2.5 and PM10 are the main factors affect air quality, and they are also important culprits that damage people's health.

Table 5 Principal Component Factor Scores for Variables

Variable	Factor 1	Factor 2	Factor 3	Mean
PM2.5	.848	.849	.952	.883
PM10	.924	.944	.674	.847
CO	.556	.139	.106	.267
NO2	.798	.785	.866	.817
SO2	.886	.856	.802	.848
O3-1h	.738	.810	.865	.804
O3-8h	.622	.823	.857	.767
Total amount	5.372	5.206	5.122	5.233
Total percentage	76.743	74.370	73.168	74.761

Figure 1 shows the three-dimensional difference map of urban air quality variables. Each box shows a two-dimensional difference map between a pair of main factors. It shows differences between the original value and the difference value, and differences between urban air quality variables. There is a big gap between CO and other variables, and CO has less influence on urban air quality; the other six variables are close to each other and have a stronger impact on urban air quality; O3-1h is close to O3-8h, and impacts of two variables on urban air quality are similar.

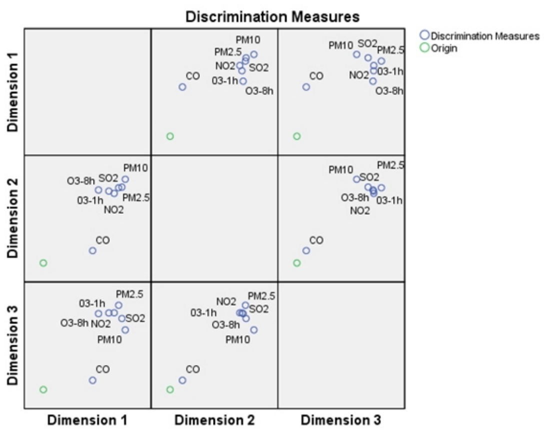

Fig.1 Three-dimensional difference of urban air quality variables

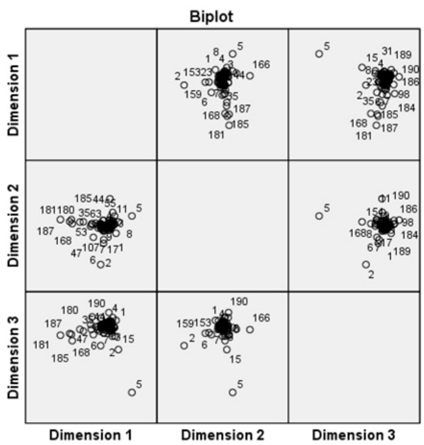

Figure 2 Three-dimensional difference of urban air quality sample points

Figure 2 shows the three-dimensional difference map of urban air quality sample points. Each box shows a two-dimensional difference map of urban samples between two main factors. Each sample point is the city air quality AIRR city ranking point, and the AQI ranking corresponds to cities can be found in Table 4. Cities with AQI ranks 1 to 10 are Sanya, Haikou, Zhanjiang, Maoming, Beihai, Zhuhai, Yunfu, Lhasa, Shenzhen, and Heyuan. From this map, the AIRR city rankings overlap more in the middle of the AIRR rankings, some gaps in the city sample points are the main AIRR ranking points, such as Sanya, Haikou, Zhanjiang, Maoming, Beihai, Zhuhai, etc., or AIRR ranking points, such as Zaozhuang, Linyi, Handan, Tangshan, Hengshui, Baoding, Shijiazhuang, Xingtai, Korla, etc.

Conclusions and Prospects

(A)The variable with the greatest impact on air quality is PM2.5, which is also a culprit for people's health. PM10, SO2, and NO2 are also important factors affecting air quality. CO has limited impact on air quality, and ozone is also an important factor affecting air quality. The state needs to establish the concept of environmental governance for formulate systems, rules and laws, and protect the environment according to the law. It requires to limit the growth rate of urban vehicles and strictly control the emission of pollutants from enterprises.

(B) The comprehensive scores are basically the same as those published by the national air quality index. The cities ranked highest among AQI rankings are Sanya, Haikou, Zhanjiang, Maoming, Beihai, Zhuhai, Lhasa, Shenzhen, Heyuan, Zhoushan and Huizhou. They need supervision to prevent air pollution rebound. The cities ranked high in both AQI rankings and comprehensive scores are Tianjin, Zaozhuang, Linyi, Handan, Tangshan, Hengshui, Baoding, Shijiazhuang, Xingtai, and Korla. Further strengthening the governance of cities and ensuring the effectiveness of governance is needed.

(C) At present, the urban air quality supervision is being strengthened, which is conducive to the discovery of urban environmental pollution sources. Therefore, it is necessary to make full use of monitoring results and take effective measures to govern the environment in accordance with the law. In recent years, the rapid development of urbanization in China's industrialization has caused air pollution, especially smog weather, which is mainly based on PM2.5, and seriously endangers people's health. Industrial development will cause a large amount of energy consumption. At present, energy consumption mainly uses coal. The burning of large amounts of coal will increase carbon emissions, causing dust, CO2, and SO2 pollution. Therefore, while reducing energy consumption, we also attach importance to the research, development, and use of green energy, replace green coal with high pollution, reduce carbon emissions, and gradually eliminate or reduce the causes of environmental pollution. The development of urban transportation, the natural evaporation of vehicle oil products and the emission of exhaust gas have also become important killers of urban air pollution, restricting urban automobiles and strict automobile emission standards are effective measures.

Acknowledgements

This work was a Fund Project: The Ministry of Education Planning Fund Project (17XJA630005); Sichuan Provincial Department of Education Key Project (17SA0135).

References

[1] Chen Renjie, Chen Bingheng, Yan Haidong. China's air quality health index preliminary study [J]. Chinese Environmental Science, 2013, 33(11): 2081~2086.

[2] Wang Ju, Hou Jie, Yang Mengxi, Xu Zhiyu, Fang Chunsheng, Chen Kehua, Xiao Zhongshen. Correspondence analysis in the application of atmospheric particle source analysis [J]. China Environmental Monitoring, 2012.6: 41-46.

[3] Zhu Xiuhua, Li Xiaoxu. Application of Correspondence Analysis Method in Urban Atmospheric Environmental Quality Assessment [J]. Journal of Dalian Jiaotong University, 2010.2: 89-93.

[4] Chen Mingyu. Urban Air Quality Evaluation System Based on MATLAB[J]. Ecological Environment Protection, 2013.10: 202-203.

[5] Huang Wenfei, Lu Yiying, Wang Hongxiao, Wen Changling. U.S. Air Quality Management Means Based on Pollution Permits and Its Reference[J]. Environmental Protection, 2014.3: 63-64.

[6] Chen Chun, Research and Exploration of Air Quality in China: Analysis Based on Structural Equation Model [J]. Ecological Economy, 2016.12: 142-145.

[7] (German) Wolfgang Hardle, (by) Leopld Simar, Translated by Chen Shiyi. Applied Multivariate Statistical Analysis [M]. Peking University Publishers, 2011.

[8] Greenacre, M. Correspondence analysis of raw data [J]. Ecology 2010: 91(4): 958-963.

[9] Leibold, C.Signal-to-noise-optimal scaling of heterogeneous population codes[J].Network 2013: 24(3): 114-128.