

文章编号:1000-2375( 2016) 06-0567-05

# 主成分分析在城市大气环境质量评价中的应用

李成 李海波 高丹丹 杨小露

( 湖北大学资源环境学院 湖北 武汉 430062)

**摘要:**以武汉市 2015 年的 5 种大气主要污染物的监测数据为依据,运用主成分分析法( PCA),对武汉市的大气环境质量进行综合评价。结果表明,武汉市 2015 年 12 个月的大气质量状况为:6 月、7 月、8 月和 9 月的环境空气质量要优于国家空气质量一级标准;4 月、5 月和 11 月的空气质量要优于国家空气质量二级标准;1 月、2 月、3 月、10 月和 12 月这 5 个月均差于国家空气质量二级标准,且整体上表现为夏季空气质量状况明显优于冬季空气质量状况。

**关键词:**主成分分析法;大气环境质量评价;综合评价;武汉

中图分类号:X823 文献标志码:A DOI: 103969/j.issn.1000-2375.2016.06.017

## Application of principal component analysis in the evaluation of urban atmospheric environment quality

LI Cheng, LI Haibo, GAO Dandan, YANG Xiaolu

( Faculty of Resources and Environmental Science, Hubei University, Wuhan 430062, China)

**Abstract:** Based on the monitoring data of five kinds of main atmospheric pollutants in Wuhan city in 2015, using principal component analysis ( PCA), on the atmospheric environmental quality comprehensive evaluation of Wuhan city. The results showed that the condition of atmospheric quality about twelve months of 2015 in Wuhan city were as follows: ambient air quality in June, July, August and September was superior to the national air quality standards; air quality in April, May and November was better than national secondary air quality standards; in January, February, March and October and December it was inferior to secondary air quality standards of the state, and overall performance for the summer air quality was better than winter air quality condition.

**Key words:** principal component analysis ( PCA); atmospheric environmental quality assessment; comprehensive evaluation; Wuhan

## 0 引言

随着人口的快速增长和经济的迅猛发展,导致大气环境质量不断恶化,因此,客观、合理地评价大气环境质量就显得尤为重要。这一问题,20 世纪 90 年代以来,就是环境研究方面的热点和难点。目前应用于大气环境质量评价的方法主要有指数评价法、灰色聚类法、神经网络、改进密切值法等方法。由于这些方法的变量较多,具有一定的缺陷性,因此,很难体现主要指标的作用。

近年来,随着多元统计方法的普及和应用,主成分分析方法( PCA) 作为一种新兴的方法得到广泛的应用。考虑到多个变量之间的相关关系,主成分分析法<sup>[1]</sup>能够最大限度地保留原有数据的信息,通过对

收稿日期:2016-01-09

作者简介:李成( 1991-) 男,硕士生;李海波,通信作者,教授, E-mail: 1278289302@qq.com

多维数据进行最佳的综合降维处理 将许多变量转化为几个不相关的变量 ,更加合理、客观地确定了各个指标的权重数 ,避免了主观随意性带来的误差.本文中利用该方法对武汉市 2015 年 12 个月的大气环境质量进行综合评价 ,以期该方法能为环保部门对环境质量的监测、污染的控制以及污染源的解析提供一定的科学参考依据.

1 基础数据来源

依据《中华人民共和国环境空气质量标准》( GB 3095-2012) ,选择了武汉市 9 个国控环境监测点 ( 东湖梨园、汉阳月湖、汉口花桥、武昌紫阳、青山钢花、沌口新区、汉口江滩、东湖高新和吴家山) 2015 年 12 个月的二氧化硫(  $\text{SO}_2$ ) 、二氧化氮(  $\text{NO}_2$ ) 、可吸入颗粒物(  $\text{PM}_{10}$ ) 、臭氧(  $\text{O}_3$ ) 和细颗粒物(  $\text{PM}_{2.5}$ ) 5 个空气污染物的实测资料 ,作为综合评价的数据.

2 主成分分析的原理和方法

2.1 基本原理 主成分分析法主要是将原来多个具有一定相关性的指标 ,进行线性组合 ,重新组合成少量的几组无相关的新的综合指标.我们将它们称之为主成分.当前  $n$  个主成分的方差累积贡献率达到 80% 以上时 ,我们就认为这  $n$  个主成分能够代表原始数据的绝大部分信息.

2.2 方法步骤

- 1) 数据标准化<sup>[2]</sup>.为了将数量级和量纲的影响排除在外 ,使原始数据能在同一尺度上比较 ,需要对数据进行标准化处理.
- 2) 利用标准化后的矩阵求解相关系数矩阵<sup>[3]</sup>.
- 3) 计算特征值和特征向量<sup>[4]</sup>.
- 4) 计算解释方差 ,确定主成分.一般要求主成分的累积方差贡献率达到 80% 以上即可.
- 5) 对  $m$  个主成分进行综合评价<sup>[5]</sup>.通过对  $m$  个主成分进行加权求和 ,得到最终的评价值 ,权数为每一个主成分的方差贡献率.根据各主成分与其相应的贡献率之积的和 ,计算各环境要素的污染综合得分 ,得分越大表示污染的程度越严重.

3 主成分分析评价结果

本文中以武汉市 2015 年 12 个月的大气环境质量状况( 见表 1) 为例 ,参照 GB 3095-2012《环境空气质量标准》( 见表 2) ,通过主成分分析法( 基于 Spss18.0) 对大气环境质量进行综合评价.

| 表 1 武汉市 2015 年 12 个月的环境质量状况 $\mu\text{g}/\text{m}^3$ |               |               |                  |              |                   |
|--|---------------|---------------|------------------|--------------|-------------------|
| 月份   | $\text{SO}_2$ | $\text{NO}_2$ | $\text{PM}_{10}$ | $\text{O}_3$ | $\text{PM}_{2.5}$ |
| 1  | 35            | 80            | 149              | 68           | 131               |
| 2  | 24            | 59            | 119              | 98           | 110               |
| 3  | 21            | 53            | 100              | 101          | 71                |
| 4  | 18            | 56            | 113              | 152          | 60                |
| 5  | 15            | 49            | 113              | 193          | 66                |
| 6  | 13            | 42            | 79               | 162          | 41                |
| 7  | 10            | 35            | 69               | 174          | 36                |
| 8  | 13            | 37            | 83               | 197          | 41                |
| 9  | 16            | 41            | 86               | 188          | 48                |
| 10   | 19            | 65            | 128              | 168          | 71                |
| 11   | 15            | 46            | 73               | 70           | 57                |
| 12   | 18            | 58            | 137              | 61           | 108               |

| 表 2 《环境空气质量标准》( GB 3095-2012) $\mu\text{g}/\text{m}^3$ |               |               |                  |              |                   |
|--|---------------|---------------|------------------|--------------|-------------------|
| 大气质量级别   | $\text{SO}_2$ | $\text{NO}_2$ | $\text{PM}_{10}$ | $\text{O}_3$ | $\text{PM}_{2.5}$ |
| 一级   | 20            | 40            | 40               | 100          | 15                |
| 二级   | 60            | 40            | 70               | 160          | 35                |

| 表 3 相关系数矩阵        |               |               |                  |              |                   |
|-------------------|---------------|---------------|------------------|--------------|-------------------|
| 指标名称              | $\text{SO}_2$ | $\text{NO}_2$ | $\text{PM}_{10}$ | $\text{O}_3$ | $\text{PM}_{2.5}$ |
| $\text{SO}_2$     | 1.000         |               |                  |              |                   |
| $\text{NO}_2$     | 0.193         | 1.000         |                  |              |                   |
| $\text{PM}_{10}$  | 0.046         | 0.876         | 1.000            |              |                   |
| $\text{O}_3$      | -0.158        | -0.517        | -0.266           | 1.000        |                   |
| $\text{PM}_{2.5}$ | 0.119         | 0.877         | 0.907            | -0.562       | 1.000             |

数据来自湖北省环境保护厅

3.1 数据标准化处理 为了消除 5 个指标的量纲所带来的影响,对原始数据进行标准化处理,使处理后的数据具有可比性.通过 Spss18.0 可以快速地将数据进行标准化处理.

3.2 计算相关系数矩阵 利用 Spss 软件可以得到 5 个评价指标的相关系数矩阵(见表 3)以及每个变量的提取度(见表 4).

由表 3 可知,PM<sub>2.5</sub>、NO<sub>2</sub>、PM<sub>10</sub>、O<sub>3</sub> 四者之间具有较强的相关性,SO<sub>2</sub>与其他 4 个指标间的相关性较弱.由表 4 可知,这 5 个变量的共性方差,除了 O<sub>3</sub>接近 0.5 以外,其他的变量的共性方差都大于或接近 0.9,故表示提取的公共因子能够较好地反映原始变量的主要信息.

3.3 计算特征值和主成分贡献率 通过协方差矩阵,可以求出每一个主成分所对应的特征值、解释方差以及累积方差贡献率<sup>[6]</sup>,如表 5 所示.

表 4 指标的提取度

| 指标名称              | 初始    | 提取    |
|-------------------|-------|-------|
| SO <sub>2</sub>   | 1.000 | 0.904 |
| NO <sub>2</sub>   | 1.000 | 0.913 |
| PM <sub>10</sub>  | 1.000 | 0.882 |
| O <sub>3</sub>    | 1.000 | 0.464 |
| PM <sub>2.5</sub> | 1.000 | 0.949 |

表 5 特征值及主成分贡献率

| 主成分 | 特征值   | 方差贡献率  | 累积贡献率   |
|-----|-------|--------|---------|
| 1   | 3.098 | 61.959 | 61.959  |
| 2   | 1.014 | 20.276 | 82.235  |
| 3   | 0.738 | 14.769 | 97.005  |
| 4   | 0.123 | 2.463  | 99.467  |
| 5   | 0.027 | 0.533  | 100.000 |

从上表 5 可以看出,第一主成分的方差贡献率已达到 61.959%,说明第一主成分已可以代表大多原始数据的信息.再加上第二主成分,前两个主成分的方差累积贡献率达到 82.235%,已经大于 80%.因此,前两个主成分能够反映原始数据提供的绝大部分信息.根据特征值大于或等于 1 的原则,确定主成分个数为 2 个.利用它们,可以对环境空气质量进行综合评价.

3.4 计算主成分表达式 利用 Spss 软件先求出主成分载荷矩阵<sup>[7]</sup>,然后将主成分载荷矩阵中的数据除以主成分相对应的特征值,再开平方根便可得到两个主成分中每个指标所对应的系数,如表 6 所示.

表 6 两个主成分的特征向量

| 主成分 | SO <sub>2</sub> | NO <sub>2</sub> | PM <sub>10</sub> | O <sub>3</sub> | PM <sub>2.5</sub> |
|-----|-----------------|-----------------|------------------|----------------|-------------------|
| 1   | 0.119           | 0.543           | 0.512            | -0.355         | 0.551             |
| 2   | 0.927           | -0.030          | -0.261           | -0.270         | -0.101            |

由上表 6 可得,这两个主成分与各个变量的线性组合关系<sup>[8]</sup>分别为:

$$Z_1 = 0.119Z_{SO_2} + 0.543Z_{NO_2} + 0.512Z_{PM_{10}} - 0.355Z_{O_3} + 0.551Z_{PM_{2.5}},$$
$$Z_2 = 0.927Z_{SO_2} - 0.030Z_{NO_2} - 0.261Z_{PM_{10}} - 0.270Z_{O_3} - 0.101Z_{PM_{2.5}}.$$

从第一主成分的特征向量构成来看,PM<sub>2.5</sub>、NO<sub>2</sub>、PM<sub>10</sub>的绝对值较大,对空气质量起主导作用,其中 PM<sub>2.5</sub>的绝对值最大,为主要污染因子;在第二主成分中,SO<sub>2</sub>的绝对值最大,说明 SO<sub>2</sub>是影响空气的主要因子.两个独立的主成分代表了武汉市大气的两种污染机制,分别为无机物污染和扬尘污染.说明现阶段武汉市空气污染的三类污染,主要是建设过程中产生的粉尘污染、工业生产产生的 SO<sub>2</sub>以及汽车尾气污染.

3.5 计算主成分得分及综合评价 利用 Spss 软件计算出各主成分得分,然后将各主成分得分与对应的方差贡献率相乘以后的总和,即为综合得分<sup>[9]</sup>.综合得分  $Z = 0.753Z_1 + 0.247Z_2$ .对武汉市 2015 年空气质量状况进行定量化描述,得分越高的,表明其受污染的程度越高,以此来对环境空气质量状况进行排序和分级.结果见表 7.根据表 7 综合评价的结果,按照主成分综合得分排序<sup>[10]</sup>可以判断武汉市 2015 年 12 个月的大气环境质量状况由优到劣依次为:7 月、8 月、6 月、9 月、5 月、11 月、4 月、3 月、10 月、2 月、12 月、

1 月.其中 7 月、8 月、6 月和 9 月达到空气质量一级标准;5 月、11 月和 4 月达到空气质量二级标准;3 月、10 月、2 月、12 月和 1 月的大气质量状况差于国家空气质量二级标准.可以看出,夏季的空气质量明显要优于冬季的空气质量,而且污染最严重的是 12 月、1 月和 2 月.经分析,空气污染主要是与本地污染源排放、外来污染输入和大气扩散条件等多方面因素有关.冬季的大气活动减弱,大气扩散条件不好,容易造成本地污染积累.再加上,冬季的偏北风影响,使北方的污染物输入武汉.

表 7 环境空气质量状况综合评价结果

| 月份及大气质量级别 | 第一主成分得分 $Z_1$ | 第二主成分得分 $Z_2$ | 综合得分 $Z$  | 主成分得分排序 | 空气质量分类 |
|-----------|---------------|---------------|-----------|---------|--------|
| 1 月       | 2.210 53      | 0.639 37      | 1.822 45  | 12      | 劣于二级   |
| 2 月       | 1.033 05      | 0.052 06      | 0.790 75  | 10      | 劣于二级   |
| 3 月       | 0.307 17      | 0.114 00      | 0.259 46  | 8       | 劣于二级   |
| 4 月       | 0.176 38      | -0.459 70     | 0.019 27  | 7       | 二级     |
| 5 月       | -0.119 69     | -0.898 73     | -0.312 11 | 5       | 二级     |
| 6 月       | -0.735 93     | -0.495 66     | -0.676 58 | 3       | 一级     |
| 7 月       | -1.113 94     | -0.660 31     | -1.001 89 | 1       | 一级     |
| 8 月       | -0.961 95     | -0.706 26     | -0.898 79 | 2       | 一级     |
| 9 月       | -0.716 72     | -0.497 79     | -0.662 64 | 4       | 一级     |
| 10 月      | 0.583 57      | -0.655 19     | 0.277 60  | 9       | 劣于二级   |
| 11 月      | -0.159 44     | 0.135 95      | -0.086 48 | 6       | 二级     |
| 12 月      | 1.278 15      | -0.324 66     | 0.882 26  | 11      | 劣于二级   |
| 一级标准      | -1.113 26     | 0.756 43      | -0.651 45 |         |        |
| 二级标准      | -0.667 90     | 3.000 47      | 0.238 19  |         |        |

根据武汉市环保局监测发布的武汉市 2015 年 12 个月的空气质量状况由优到劣依次为:7 月、6 月、11 月、8 月、9 月、3 月、4 月、5 月、10 月、12 月、2 月、1 月.与本文中评价结果相比,大致趋势都是夏季的空气质量优于冬季,6 月、7 月、8 月、9 月的空气质量较好,12 月、1 月和 2 月的空气质量最差.不同之处在于 3 月、4 月、5 月和 11 月这 4 个月与本文的综合得分排序不一样.原因可能是因为这 4 个月的空气质量优劣程度相近,再加上本文中的指标变量  $O_3$  的提取度稍低,对结果的分析会产生一些误差.本文中使用的的方法不仅能对大气质量状况进行定性和定量的分析,还能有效地反映各环境要素中不同污染物的污染程度以及它们之间的相关性,对推断污染源有一定的参考价值.

## 4 结论与讨论

1) 基于主成分分析法对武汉市 2015 年的大气环境质量进行综合评价,得到的结果与我们实际所处的大气环境大体一致,说明利用主成分分析法在大气环境质量评价方面的可行性.

2) 通过主成分分析法将多个变量指标,转换成少量的综合指标,使数据结构得到大大的简化,为评价大气环境质量提供了一种简单易行的方法.

3) 通过主成分所对应的特征向量,还可以知道各主要污染物之间的相关性<sup>[11]</sup>,为进一步推断污染源提供了有利的信息,为环保部门制定污染防治措施提供一定的参考依据.

## 5 参考文献

- [1] 王丹,倪长健.主成分分析法在大气环境质量评价中的应用[J].平顶山学院报,2011,26(2):112-115.
- [2] 杨仲伟.主成分分析法在农村环境质量评价中的应用[J].甘肃科学学报,2015,27(1):72-75.
- [3] 解素雯,孟昭为.淄博市 2014 年大气环境质量综合评价——基于主成分分析[J].现代商贸工业,2015,6(上):27-28.
- [4] 黄玉平,张庆国,汪水兵,等.主成分分析在大气质量监测优化布点中的应用[J].安徽农业大学学报,2011,38(6):966-969.
- [5] 冯利华.环境质量的主成分分析[J].数学的实践与认识,2003,33(8):32-35.

- [6] 朱蕾.基于主成分分析法的扬州市生态环境质量评价[D].扬州:扬州大学,2013.
- [7] 卢文岱.SPSS for Windows 统计分析[M].北京:电子工业出版社,2003.
- [8] 郑成德,尹德松.主成分分析法在城市大气环境质量评价中的应用[J].大连铁道学院学报,2002,23(3):39-41.
- [9] 孙成宝,李云丽.环境质量评价中的主成分分析与全局主成分分析方法[J].甘肃联合大学学报(自然科学版),2006,20(2):21-25.
- [10] 卜鹏,吕惠进,黄秋香,等.基于主成分分析法的南通市市区大气环境质量评价[J].科技通报,2016,32(2):214-217.
- [11] 张曙红,武鹏程.武汉市大气环境质量的综合评价[J].环境科学与技术,2008,31(3):110-113.

(责任编辑 游俊)

(上接第560页)

的.紫外线诱变可以产生适宜低温生长的 *A. ferrooxidans* 菌种,但诱变时间 30 s、60 s 的效果无显著差异.20 ℃、15 ℃ 与 10 ℃ 时诱变效果相比较,15 ℃ 的诱变效果最佳,诱变组的  $\text{Fe}^{2+}$  氧化率比未诱变的空白组高了 14%.

### 3 结论与讨论

本研究可得出如下结论:1)  $\text{Fe}^{2+}$  浓度随时间变化的曲线呈反“S”型, $\text{Fe}^{2+}$  浓度又与 *A. ferrooxidans* 菌的活性成反比,因而 *A. ferrooxidans* 菌的活性变化是符合有限环境容量下微生物的增长曲线——“S”型增长曲线的;2) 各组亚铁氧化率在第二、三培养阶段的变化大致相同,这表明此时 *A. ferrooxidans* 菌的生长已稳定,故取第三培养阶段的数据作为有效数据;3) 紫外线诱变可以产生适宜低温生长的 *A. ferrooxidans* 菌种,但诱变时间 30 s、60 s 的效果无显著差异;4) *A. ferrooxidans* 菌的活性是随温度的降低而减小的,诱变产生的菌种并不能逆转这一生长规律;5) 3 个诱变温度相比较,15 ℃ 的诱变效果最佳,诱变组的  $\text{Fe}^{2+}$  氧化率比未诱变的空白组高了 14%.

研究可能存在的问题及对后续研究的展望:a) 未直接测定 *A. ferrooxidans* 菌的活性,而是通过  $\text{Fe}^{2+}$  的含量变化间接表示;b) 就不同诱变时间、次数、温度对 *A. ferrooxidans* 菌的活性进行更详细、深入的研究,可得到更全面的数据支持,有利于建立 *A. ferrooxidans* 菌活性与诱变时间和温度的关系的数学模型.

### 4 参考文献

- [1] Kelly D P, Wood A P. Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halothiobacillus* gen. nov. and *Thermithiobacillus* gen. nov. [J]. International Journal of Systematic and Evolutionary Microbiology, 2000, 50: 489-500.
- [2] Temple K L, Colmer A R. The autotrophic oxidation of iron by a new bacterium *Thiobacillus ferrooxidans* [J]. Journal of Bacteriology, 1951, 61: 605-611.
- [3] Abhilash K D, Mehta Bansi D, Pandey. Efficacy of bacterial adaptation on copper biodissolution from a low grade chalcopyrite ore by *A. ferrooxidans* [J]. International Journal of Nonferrous Metallurgy, 2012, 1(1): 1-7.
- [4] 徐晓军,孟运生,宫磊,等.氧化亚铁硫杆菌紫外线诱变及对低品位黄铜矿的浸出[J].矿冶工程,2005,25(1):34-36.
- [5] 周顺桂,周立祥,黄焕忠.黄钾铁矾的生物合成与鉴定[J].光谱学与光谱分析,2004,24(9):1140-1143.
- [6] 张德远,李雅芹,吴依陶.生物加工纯铜的动力学与热力学研究[M].中国科学(C辑),1999,29:132-137.
- [7] 徐晓军,孟运生,宫磊,等.氧化亚铁硫杆菌紫外线诱变及对低品位黄铜矿的浸出[J].矿冶工程,2005,25(1):34-36.
- [8] 鲍士旦.土壤农化分析[M].3版.北京:中国农业出版社,1981:225-226.

(责任编辑 游俊)