

主分量分析法在环境空气监测数据分析中的应用

刘得守¹, 周康², 张夏夏², 左瑞亭^{2*}

(1. 西宁市生态环境局, 青海 西宁 810001; 2. 南京雨后地软环境技术有限公司, 江苏 南京 210013)

摘要:根据西宁市13个环境空气监测站点2013—2017年大气污染物细颗粒物($PM_{2.5}$)、可吸入颗粒物(PM_{10})、二氧化硫(SO_2)、二氧化氮(NO_2)、臭氧最大8 h平均(O_3-8h)和一氧化碳(CO)的监测数据,采用主分量分析法对西宁市环境空气质量进行了综合评估。结果表明,2013—2017年西宁市大部分环境空气监测站点周边环境空气质量逐渐提升,4个国控站综合得分(F)趋势变化幅度较大,其周边环境空气质量状况改善较为明显;城南新区、湟源县气象局和西钢监测站点周边环境空气质量呈逐年下降趋势,与其附近工业生产有关。

关键词:主分量分析法;环境空气质量;综合评估;西宁

中图分类号:X823

文献标志码:B

文章编号:1674-6732(2019)04-0016-04

Application of Principal Component Analysis in Evaluating of Ambient Air Automatic Monitoring Station

LIU De-shou¹, ZHOU Kang², ZHANG Xia-xia², ZUO Rui-ting^{2*}

(1. Xining Ecological and Environment Bureau, Xining, Qinghai 810001, China; 2. Nanjing Star-jelly Environmental Consultant Company Limited, Nanjing, Jiangsu 210013, China)

Abstract: According to the monitoring data of fine particulate matter ($PM_{2.5}$), respirable particulate matter (PM_{10}), sulfur dioxide (SO_2), nitrogen dioxide (NO_2), ozone up to 8 hours average (O_3-8h) and carbon monoxide (CO) in Xining City from 2013 to 2017, the principal component analysis method was used to comprehensively evaluate the ambient air quality in Xining City. The results showed that the ambient air quality of most environmental air monitoring stations in Xining City has gradually increased in recent years, and the integrated score (F) trend of the 4 state controlled stations varied mostly, the ambient air quality surrounding them improved obviously. While the air quality observed in the sites of Chengnan new area, Meteorological Bureau of Huangyuan county and Xigang showed declining year by year, which was related to the industrial production nearby.

Key words: Principal component analysis; Ambience air quality; Comprehensive evaluation; Xining

目前评价环境空气质量的方法有灰色聚类法^[1-3]、综合指数评价法^[4-5]、模糊综合评判法^[6-8]、改进密切值法^[9-11]和神经网络^[12-14]等。以上方法在评价时均存在一定的不足,如模糊综合评价评判法需对每个指标人为的给定一个权数,另外由于指标数量多,难以体现主要指标的作用,同时增加了评价工作量^[15]。而主分量分析法在保证原始数据信息损失最小的前提下,以少数的综合变量取代原有的多维变量,可达到评价环境空气质量的目的^[16-19]。

现根据西宁市13个环境空气监测站点2013—2017年大气污染物细颗粒物($PM_{2.5}$)、可吸入颗粒物(PM_{10})、二氧化硫(SO_2)、二氧化氮

(NO_2)、臭氧(O_3-8h)和一氧化碳(CO)的监测数据,采用主分量分析法对西宁市环境空气质量进行了综合评估。

1 研究区概况

西宁市位于青海省东北部,青藏高原东部。全市总面积7 649 km²,其中市辖区面积510 km²,平均海拔2 261 m,地理坐标东经101°77′、北纬36°62′。下辖城东区、城中区、城西区、城北区以及

收稿日期:2018-12-05;修订日期:2019-03-26

作者简介:刘得守(1981—),男,工程师,本科,主要从事环境监测分析工作。

* 通讯作者:左瑞亭 E-mail:zrt@njyhdr.com

大通县、湟中县、湟源县。地处中纬内陆,属于大陆性高原半干旱气候,具有气压低、日照长、雨水少、太阳辐射强、日夜温差大等特点。从行业分类看,西宁市主要有 5 大行业:有色金属冶炼和压延加工业,电子、电力、热力生产和供应业,黑色金属冶炼和压延加工业,化学原料和化学制品制造业,医药制造业。以金属冶炼及制造业为主的工业结构,加上重工业为主的特点导致能源消耗量大,给西宁市空气质量进一步改善带来较大压力。近年来全市首要污染物主要是颗粒物,受冬季供暖影响, SO_2 、 NO_2 、 $\text{PM}_{2.5}$ 月均值较高,在春季受沙尘天气影响 $\rho(\text{PM}_{10})$ 较高,西宁市空气质量整体呈现“冬春季污染重,夏秋季污染轻”,且春季沙尘污染较严重的总体特征。

2 研究方法

2.1 分析方法

2.1.1 数据标准化

为了排除数量级和量纲不同带来的影响,需要对原始数据进行标准化处理,公式^[20]如下。

$$Y_{ij} = \frac{X_{ij} - X_j}{R_j} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n) \quad (1)$$

式中: X_{ij} ——第 i 个指标类型的第 j 个指标的值;

X_j 和 R_j ——第 j 个指标样本的均值和标准差;

Y_{ij} ——标准化后数据。

2.1.2 相关系数矩阵

利用标准化后的数据,采用 SPSS19.0 软件^[21]进行计算得到数据相关系数矩阵。

2.1.3 数据特征值及贡献率

根据相关矩阵 R 的特征方程 $|\lambda E_p - R| = 0$, 解得特征值 $\lambda_1, \lambda_2, \dots, \lambda_p$, 依据特征值 $\lambda \geq 1$ 确定主成分个数 m , E_p 为标准矩阵。

2.1.4 主成分得分及综合评价

主成分载荷矩阵表示主成分 F_i 与原始变量 X_j 之间的相关关系,即原来变量 $X_j (j = 1, 2, \dots, n)$ 在各主成分 $F_i (i = 1, 2, \dots, m)$ 上的载荷 $l_{ij} (i = 1, 2, \dots, m; j = 1, 2, \dots, n)$ 。 l_{ij} 为第 i 个指标类型的第 j 个指标的载荷。

$$l(Z_i, X_j) = \sqrt{\lambda_i} a_{ij} \quad (2)$$

式中: l ——主成分载荷矩阵;

a_{ij} ——主成分 F_i 表达式的系数;

λ_i ——相关系数矩阵的特征值;

Z_i ——标准化后的矩阵;

X_j ——第 j 个指标样本的均值。

根据各主成分在各变量上的载荷和各变量标准化后的值可以求出主成分 F_i 的得分,再以第 n 主成分的贡献率为权重,计算出综合得分。

$$F = \sum_{n=1}^n F'_n \times \omega_n \quad (3)$$

式中: F ——主成分综合得分;

F'_n ——第 n 主成分得分;

ω_n ——第 n 主成分贡献率。

2.2 数据时间

2013—2017 年。

2.3 数据来源

根据《西宁市 2013—2017 年环境质量状况公报》等资料,选取 $\text{PM}_{2.5}$ 、 PM_{10} 、 SO_2 、 NO_2 、 O_3 - 8h 和 CO 共 6 种污染物的监测数据作为评估依据。西宁市目前有 4 个国控站点和 13 个市控站点,见表 1。

表 1 17 个环境空气监测站点信息^①

序号	站点名称	所在区/县
1	市环境监测站	城西区
2	城北区政府	城北区
3	四陆医院	城东区
4	第五水厂	湟中县
5	城东区政府	城东区
6	西宁宾馆	城西区
7	海湖新区	城西区
8	城南新区	城西区
9	西钢	城西区
10	西宁市植物园	城西区
11	大通县朔山中学	大通县
12	大通县北川工业园区	大通县
13	湟中县园林所	湟中县
14	湟中县甘河西区	湟中县
15	甘河工业园管委会	湟中县
16	湟源县气象局	湟源县
17	湟源县大华工业园区	湟源县

①西宁市植物园、大通县北川工业园区、甘河工业园管委会和湟源县大华工业园区共 4 个监测站点由于设备尚处在初期调试中,实时监测数据未上报;1~4 为国控站,5~17 为市控站。

3 各监测站点数据的评估分析

3.1 市环境监测站点

2013—2017 年市环境监测站点 $\text{PM}_{2.5}$ 、 PM_{10} 、 SO_2 、 NO_2 、 O_3 - 8h 和 CO 的日均值监测数据见表 2。计算标准化后数据的主成分特征值、贡献率和累计贡献率见表 3,因子载荷矩阵见表 4。

表 2 市环境监测站点监测数据^① $\mu\text{g}/\text{m}^3$

年份	PM _{2.5}	PM ₁₀	SO ₂	NO ₂	O ₃ -8h	CO
2013	65	160	57	48	67	2.1
2014	58	120	41	40	75	1.3
2015	45	95	26	42	76	1.4
2016	40	90	25	45	77	1.5
2017	38	102	24	40	103	1.4

①CO 单位: mg/m^3 。

表 3 主成分特征值及贡献率

主成分	λ	$w/\%$	累计 $w/\%$
F_1	4.594	76.6	76.6
F_2	0.758	12.6	89.2
F_3	0.636	10.6	99.8
F_4	0.011	0.19	100
F_5	4.957×10^{-16}	8.3×10^{-15}	100
F_6	-4.264×10^{-18}	-7.1×10^{-17}	100

表 4 因子载荷矩阵

污染物	F_1
PM _{2.5}	0.911
PM ₁₀	0.928
SO ₂	0.962
NO ₂	0.808
O ₃ -8h	-0.751
CO	0.872

由表 3 可见, F_1 的 $\lambda > 1$, 因此提取 F_1 , 除 O₃-8h 以外, F_1 与 PM_{2.5}、PM₁₀、SO₂、NO₂ 和 CO 年日均值成正相关, F_1 主要代表了该 6 种污染物日均值总的变化程度。各年份 F_1' 和 F 见表 5。

表 5 各年份 F_1' 和 F 及排序

年份	F_1		F	排序
	F_1'	排序		
2013	6.281	5	4.809	5
2014	-0.212	4	-0.163	4
2015	-2.374	1	-1.818	1
2016	-1.967	2	-1.506	2
2017	-1.728	3	-1.323	3

由表 5 排序可见, F_1' 在 2013—2014 年呈下降趋势, 在 2015—2017 年呈上升趋势, 2015 年得分最低, 说明该监测站点周边区域的污染程度在 2013—2014 年呈减轻趋势, 在 2015—2017 年呈加重趋势, 2015 年污染程度最轻。 F 的大小表示大气环境空气污染程度的轻重, 即 F 越大, 大气环境

污染程度越严重。从 F 来看, 2013、2014 年 F 呈下降趋势, 在 2015—2017 年呈上升趋势, 说明该监测站点周边近 5 年整体大气污染程度呈减轻趋势, 但近 3 年来大气环境质量有所下降。

3.2 其他监测站

各监测站点 F 趋势变化见图 1。由图 1 可见, 城南新区、湟源县气象局和西钢 3 个监测站点周边空气污染呈逐年上升趋势, 这是由于城南新区监测站点处于工业企业众多的南川工业园, 湟源县气象局监测站点北侧有工厂, 西钢监测站点位于西宁特种钢铁厂区内, 与其监测站点周边工业生产活动有关, 后续应加强相应企业的排污管理; 其余监测站点虽有小幅波动, 但总体呈下降趋势, 说明近些年西宁市整体大气污染状况得到改善, 环境空气质量有所提升; 西宁市 2013—2017 年各污染物年均值汇总见表 6。由表 6 可见, 各污染物年均值呈逐年下降趋势。

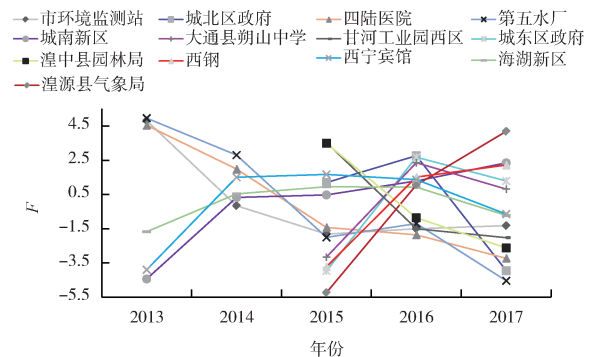


图 1 各监测站点 F 趋势变化

表 6 2013—2017 年各污染物年均值汇总^① $\mu\text{g}/\text{m}^3$

年份	SO ₂	NO ₂	CO	O ₃ -8h	PM _{2.5}	PM ₁₀
2013	48	41	1.8	84	70	163
2014	41	38	1.3	60	63	120
2015	31	38	1.4	81	49	106
2016	31	42	1.4	80	49	113
2017	24	40	1.4	92	39	100

①CO 单位: mg/m^3 。

4 结论

(1) 2013—2017 年西宁市整体环境空气质量有所提升;

(2) 4 个国控站 F 趋势变化幅度较大, 其周边环境空气质量状况改善较为明显;

(3) 城南新区、湟源县气象局和西钢监测站点

周边环境空气质量呈逐年下降趋势,与其附近工业生产有关。

[参考文献]

- [1] 胡军,许丽忠,张江山. 基于改进灰色聚类法的大气环境质量综合评价[J]. 福建师大学报(自然科学版),2012,28(1):49-54.
- [2] 麻冰涓,贺添. 基于改进的灰色聚类法的大气环境质量评价方法的研究[J]. 环境研究与监测,2008(3):51-54.
- [3] 朱庆峰,张淑娟. 灰色聚类法在大气环境质量评价中的应用[J]. 广州环境科学,2004(2):41-44.
- [4] 吉祝美. 新《环境空气质量标准》下的环境空气质量评价方法探讨[J]. 污染防治技术,2014(3):47-50.
- [5] 普映娟,王琳邦. 环境空气质量综合指数评价方法探讨[J]. 环境科学导刊,2010,29(2):93-94.
- [6] 肖继先. 应用模糊综合评价级数评价大气环境质量[J]. 环境工程,1994(6):46-49.
- [7] 马媛媛,孙世群. 模糊综合评价在合肥市大气环境评价中的应用[J]. 环境科学与管理,2012,37(5):188-191.
- [8] 吴新国,王杰,彭书时,等. 武汉市2001~2005年大气环境质量模糊综合评价[J]. 环境科学与技术,2007,30(10):58-59.
- [9] 许丽忠,张江山. 改进密切值法优化大气环境监测布点[J]. 环境工程,2000,18(2):50-53.
- [10] 许剑海,张江山,陈庆华,等. 改进密切值法及其在大气环境质量评价中的应用[J]. 环境科学与管理,2008,33(10):188-191.
- [11] 杨红玉,张江山. 改进密切值法在环境综合评价中的应用[J]. 环境科学与技术,2005,28(s1):112-113.
- [12] 杨文东,程洪斌. 神经网络在大气环境质量评价中的应用研究[J]. 工业安全与环保,2001,27(9):31-33.
- [13] 王红芳,康慕宁,邓正宏. BP神经网络在大气环境质量评价中的应用研究[J]. 科学技术与工程,2009,9(7):1997-2000.
- [14] 张虹冕,孙世群. 基于BP神经网络的大气环境质量评价模型[J]. 安徽农业科学,2010,38(31):17656-17657.
- [15] 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.
- [16] 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]. 平顶山学院学报,2011,26(2):112-115.
- [17] 姜新华,薛河儒,张存厚,等. 基于主成分分析的呼和浩特市空气质量影响因素研究[J]. 安全与环境工程,2016,23(1):75-79.
- [18] 解素雯,孟昭为. 淄博市2014年大气环境质量综合评价——基于主成分分析法[J]. 现代商贸工业,2015,36(11):27-28.
- [19] 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.
- [20] 薛燕芝,刘选卫,易文利,等. 主成分分析法在大气环境质量评价中的应用——以陕西省铜川市为例[J]. 河南科学,2016,34(8):1307-1311.
- [21] 时立文. SPSS 19.0 统计分析从入门到精通[M]. 北京:清华大学出版社,2012.

栏目编辑 李文峻