

# 我国污染排放情况的评价研究及实证分析

## ——基于中国31个省(市、自治区)的主要城市数据

赵慧盈, 梁天维

广东财经大学统计与数学学院, 广东 广州

收稿日期: 2024年6月29日; 录用日期: 2024年8月1日; 发布日期: 2024年8月26日

### 摘要

本研究收集了2022年中国各省31个主要城市的污染排放与环境的相关数据, 变量包括降水量、平均气温、平均相对湿度、日照时数、工业废水COD排放量(吨)、工业氨氮排放量等16类数据对中国各省主要城市进行污染排放情况研究。首先对数据进行简单分析, 可以知道重庆、成都、哈尔滨三个城市在污染排放方面应该是存在一定的问题。再对数据进行简单的系统聚类 and K-Means 聚类, 当将31个城市分为四类的时候, 两种聚类方法得到的结果都是将重庆和哈尔滨单独出了一类。为了进一步研究, 对31个城市数据进行污染排放处理情况的评价分析, 分别使用主成分分析法和因子分析法进行分析, 对比两种方法的好坏, 最终考虑到对主成分和公因子的解释原因, 认为使用因子分析法得到的结果更优, 结论为: 在中国31个主要城市中, 污染排放处理情况存在较大问题的城市为重庆、哈尔滨、上海、武汉(依次的综合得分排位从一到四)。最后给出合理的改善建议。

### 关键词

污染排放处理, 系统聚类, K-Means 聚类, 主成分分析法, 因子分析法

# Evaluation Study and Empirical Analysis of Pollution Emissions in China

## —Based on Data from Major Cities in 31 Provinces (Municipalities and Autonomous Regions) in China

Huiying Zhao, Tianwei Liang

School of Statistics and Mathematics, Guangdong University of Finance & Economics,  
Guangzhou Guangdong

Received: Jun. 29<sup>th</sup>, 2024; accepted: Aug. 1<sup>st</sup>, 2024; published: Aug. 26<sup>th</sup>, 2024

文章引用: 赵慧盈, 梁天维. 我国污染排放情况的评价研究及实证分析[J]. 可持续发展, 2024, 14(8): 2070-2082.  
DOI: 10.12677/sd.2024.148238

## Abstract

This study collects data related to pollution emission and environment of 31 major cities in each province of China in 2022, and the variables include 16 types of data, such as precipitation, average temperature, average relative humidity, sunshine hours, industrial wastewater COD emission (tons), industrial ammonia nitrogen emission, and so on. First of all, the data are briefly analyzed, and it can be known that the three cities of Chongqing, Chengdu and Harbin should have certain problems in pollution emission. Then the data were subjected to simple systematic clustering and K-Means clustering, and when the 31 cities were divided into four categories, the results obtained from both clustering methods were that Chongqing and Harbin were singled out in one category. In order to further study the evaluation of pollution emission treatment, data from 31 cities were analyzed using both principal component analysis and factor analysis. By comparing the strengths and weaknesses of these two methods, and considering the interpretation of the principal components and common factors, it is concluded that the results obtained using the factor analysis method are more favorable. Specifically, among the 31 major cities in China, Chongqing, Harbin, Shanghai, and Wuhan (ranked from one to four based on their composite scores in descending order) are found to have the biggest problems with pollution emission treatment. Finally, reasonable suggestions for improvement are provided.

## Keywords

Pollution Discharge Treatment, Systematic Clustering, K-Means Clustering, Principal Component Analysis Method, Factor Analysis Method

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

当前, 中国经济已经进入高质量发展阶段, 但大气污染、水污染等环境问题依然严峻复杂。党的二十大报告提出我们坚持绿水青山就是金山银山的理念, 坚持山水林田湖草沙一体化保护和系统治理, 生态文明制度体系更加健全, 生态环境保护发生历史性、转折性、全局性变化, 我们的祖国天更蓝、山更绿、水更清。因此对近些年来中国主要城市的相关环境指标进行研究评价, 对以后的环境发展可以起到研究参考的作用。

### 1.1. 文献综述

近年来, 不少学者利用主成分分析法开展空气环境质量评价并取得了一系列成果。对于重庆地区, 有学者利用层级阶梯评价法、气象要素相关分析和统计计算方法分析各类污染物的年月变化特点[1]-[3]。对于哈尔滨地区, 有学者利用主成分分析法和通过 HYSPLIT 模型分析空气质量[4] [5]。卜兴兵, 方自力等人[6]利用 2017 年四川省 21 个城市 6 种大气污染物浓度监测数据, 结合 Pearson 相关性分析、主成分分析和复合污染特征分析方法, 对 21 个城市空气质量状况和影响大气质量的关键指标进行了综合评价与分析。李如玥[7]针对 2015~2020 年的湖北省宜昌市大气监测数据, 以主要大气污染物 SO<sub>2</sub>、NO<sub>2</sub>、PM<sub>10</sub>、PM<sub>2.5</sub>、O<sub>3</sub>、CO 为研究对象, 利用 Pearson 相关性分析和主成分分析分析了各污染物变化趋势及特征。刘

兰芳[8]运用主成分分析法,对湖南省 14 个地级市雾霾风险进行定量评价,获得雾霾风险评估综合指标评价公式。胡荣明[9]基于 2017 年西安市空气质量监测数据和土地利用数据,利用主成分分析法探讨了大气污染物质量浓度与土地利用类型的关系。朱常琳[10]利用主成分分析法研究了西安市主要大气污染物的相关性分析及时空分布特征。Tao Ma 等[11]分析了 2014~2016 年长三角地区空气污染特征及其与排放和气象的关系,并发现长三角北部地区的  $PM_{2.5}$ 、 $PM_{10}$ 、 $SO_2$  和  $CO$  浓度较高,长三角中部地区的  $NO_2$  和  $O_3$  浓度较高。Tao Y [12]等利用每日死亡率和空气污染数据,对位于中国珠江三角洲劳动力和资源密集型城市群的四个城市进行了时间序列分析,评估一氧化碳对死亡率的影响。2000 年 Neha Khanna [13]提出利用多种污染物的综合评判给出空气污染指数(API)体系,1988 年日本对氮氧化物进行每日预报,Patricio Perez 等[14]发展了综合人工神经网络模型,用来预测智利圣地亚哥  $PM_{10}$  的日均浓度最大值,Elias Kalapanidas 等[15]探讨了希腊雅典的短期  $NO_2$  化最大浓度预测模式 NEMO 样板。

## 1.2. 数据符号象征

文章符号象征意义如下表 1。

**Table 1.** Table of symbolic meaning of article symbols

**表 1.** 文章符号象征意义表

$y$	空气质量优良天数比例(百分比)
$x_1$	降水量
$x_2$	平均气温
$x_3$	平均相对湿度
$x_4$	日照时数
$x_5$	工业废水 COD 排放量(吨)
$x_6$	工业氨氮排放量
$x_7$	生活废水 COD 排放量
$x_8$	生活氨氮排放量
$x_9$	工业二氧化硫排放量
$x_{10}$	工业氮氧化物排放量
$x_{11}$	工业颗粒物排放量
$x_{12}$	生活及其他二氧化硫排放量
$x_{13}$	生活及其他氮氧化物排放量
$x_{14}$	生活及其他颗粒物排放量
$x_{15}$	一般工业固体废物综合利用量

## 2. 数据分析

### 2.1. 数据来源

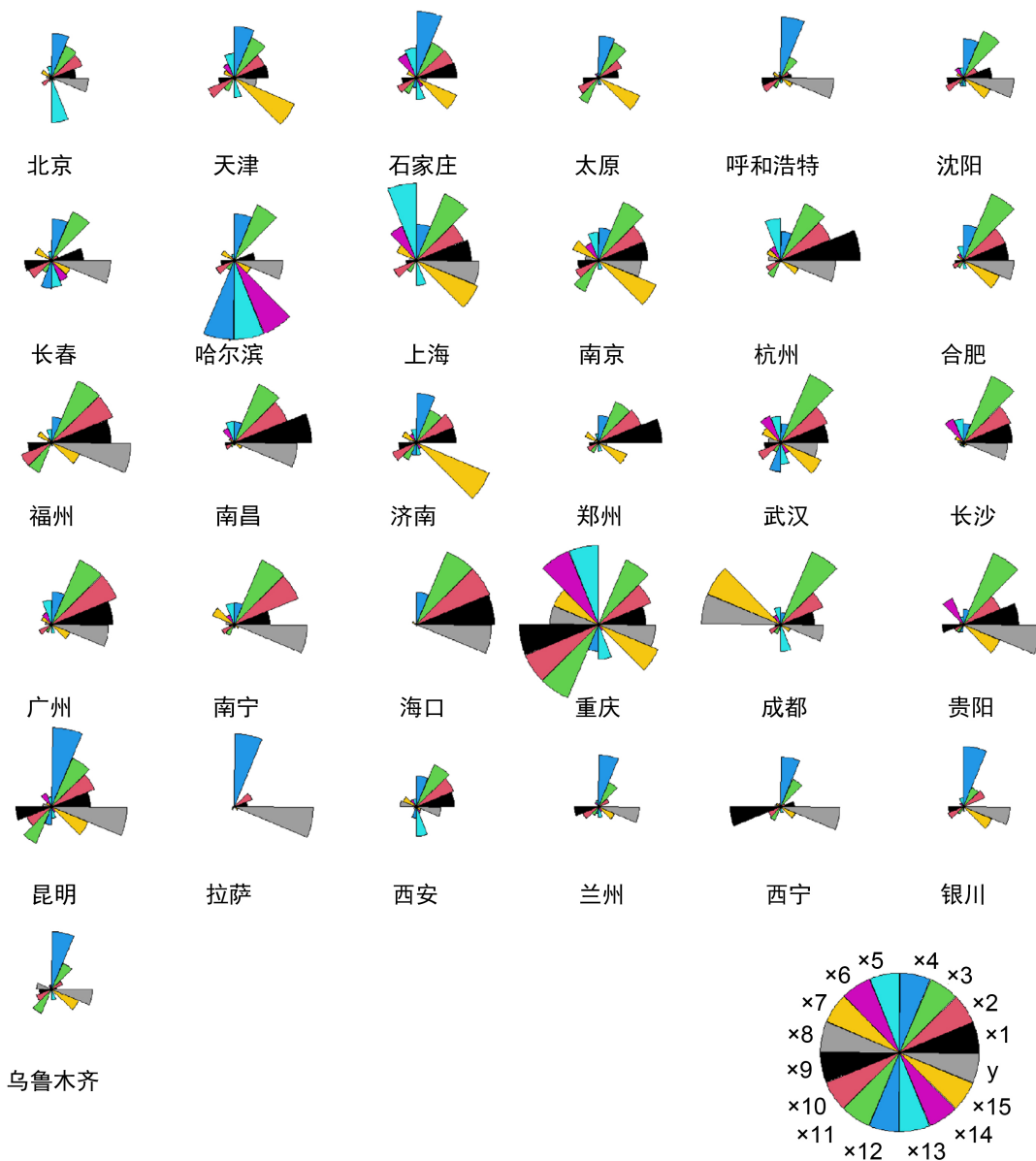
本文收集了 2022 年中国各省主要城市的资源与环境的相关数据进行研究分析,数据来源于中国统计年鉴(中国经济社会大数据研究平台(cnki.net)),数据包括各主要城市北京、天津、石家庄、太原等 31 个城市地区,变量包括降水量、平均气温、平均相对湿度、日照时数、工业废水 COD 排放量(吨)、工业氨

氮排放量等 16 类数据(均为 2022 年数据)。

本文的研究目的为通过对收集到的数据先进行简单的直观研究, 然后使用聚类分析对变量进行分析, 目的在于想看看 31 个城市之间有哪些城市是具有相关的环境特征, 对其中的几类城市可以提出相似的改善意见, 可以对评价研究有一点点的便捷作用。评价方法选用主成分分析法、因子分析法同时进行分析, 通过观察两种不同方法的评价结果, 与现实情况相结合选用相对更好的评价方法, 最终对评价结果提出改善意见。

## 2.2. 可视化分析

这一部分将对收集到的数据进行直观分析展示, 包括扇形星象图、箱线图。



**Figure 1.** Sectoral star map of data for major cities  
**图 1.** 各主要城市数据扇形星象图

图 1 为以每个样本 16 个变量画的扇形星象图, 从图中可以看出, 各主要城市中 16 个变量的值比较高的前三个城市为重庆、成都、哈尔滨。其中重庆在工业废水 COD 排放量(吨)、工业氨氮排放量、工业二氧化硫排放量、工业氮氧化物排放量、工业颗粒物排放量的数值比较大。

再画出数据的箱线图如图 2:

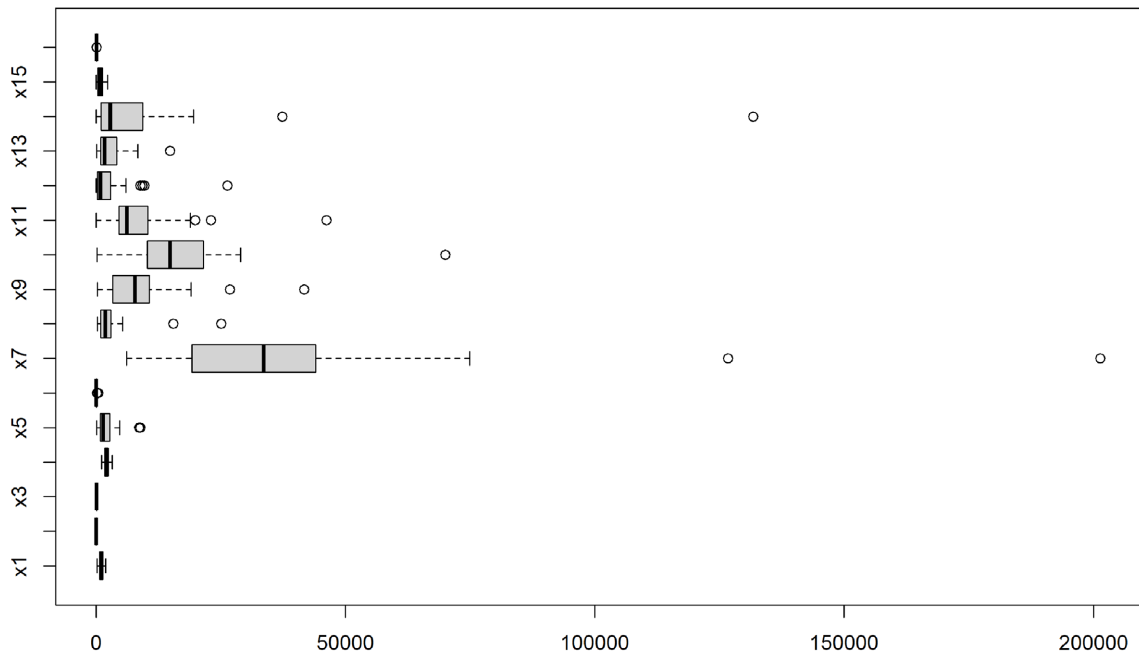


Figure 2. Box line diagram for 16 variables

图 2. 16 个变量的箱线图

从图 2 中可以看出, 在观察到的中国各主要城市中, 生活废水 COD 排放量相对其他的 15 个变量来说明显偏高, 并且在生活废水 COD 排放量中重庆和成都特别突出, 远高于其他地区, 形成离群值。其次是工业氮氧化物排放量也稍高于其他项目, 其中重庆是突出城市, 数值也远高于其他地区。

### 3. 研究方法

#### 系统聚类法:

本次系统聚类法使用离差平方和法(Ward), 该方法的基本思想来自方差分析, 如果分类正确, 同类样品的离差平方和应当较小, 类与类的离差平方和较大。具体的做法是先将  $n$  个样品各自成一类, 然后每次缩小一类, 每缩小一类, 离差平方和就要增大, 选择使方差增加最小的两类合并, 直到所有的样品归为一类为止。

设将  $n$  个样品分成  $k$  类  $G_1, G_2, \dots, G_k$ , 用  $X_{it}$  表示  $G_i$  中的第  $i$  个样品,  $n_i$  为  $G_i$  中样品的个数,  $\bar{X}_i$  是  $G_i$  的重心, 则  $G_i$  的样品离差平方和为:

$$S_i = \sum_{i=1}^{n_i} (X_{it} - \bar{X}_i)' (X_{it} - \bar{X}_i)$$

如果  $G_p$  和  $G_q$  合并为新类  $G_r$ , 类内离差平方和分别为:

$$S_p = \sum_{i=1}^{n_p} (X_{ip} - \bar{X}_p)' (X_{ip} - \bar{X}_p)$$

$$S_q = \sum_{i=1}^{n_q} (X_{iq} - \bar{X}_q)' (X_{iq} - \bar{X}_q)$$

$$S_r = \sum_{i=1}^{n_r} (X_{ir} - \bar{X}_r)' (X_{ir} - \bar{X}_r)$$

它们反映了各自类内样品的分散程度, 如果  $G_p$  和  $G_q$  这两类相距较近, 则合并后所增加的离差平方和  $S_r - S_p - S_q$  应较小; 否则, 应较大。于是定义  $G_p$  和  $G_q$  之间的平方距离为:

$$D_{pq}^2 = S_r - S_p - S_q$$

其中,  $G_r = G_p \cup G_q$ , 可以证明类间距离的递增公式为:

$$D_{kr}^2 = \frac{n_k + n_p}{n_r + n_k} D_{kp}^2 + \frac{n_k + n_q}{n_r + n_k} D_{kq}^2 - \frac{n_k}{n_r + n_k} D_{pq}^2$$

### K-Means 聚类法:

K-Means 算法的处理流程如下: 首先, 随机地选择  $k$  个对象, 每个对象初始地代表了一个簇的平均值或中心。对剩余的每个对象, 根据其与其各个聚类中心的距离将它赋给最近的簇。然后重新计算每个簇的平均值作为聚类中心进行聚类。这个过程不断重复, 直到准则函数收敛, 如下所示, 通常采用平方误差原则, 其定义如下:

$$E = \sum_{i=1}^k \sum_{p \in C_i} (p - m_i)^2$$

其中,  $E$  为数据中所有对象与相应聚类中心的均方差之和,  $p$  代表对象空间中的一个点,  $m_i$  为类  $C_i$  的均值( $p$  和  $m_i$  均是多维的)。该式所示聚类标准旨在使所有获得的聚类有以下特点: 各类本身尽可能紧凑, 而各类之间尽可能分开。

### 主成分分析法:

主成分分析主要目的是希望用较少的变量去解释原来资料中的大部分变异, 亦即期望能将手中许多相关性高的变量转化成彼此互相独立的变量, 能由其中选取较原始变量个数少的, 能解释大部分资料变异的几个新变量, 也就是所谓的主成分, 而这几个主成分也就成为我们用来解释资料的综合性指标。

### 因子分析法:

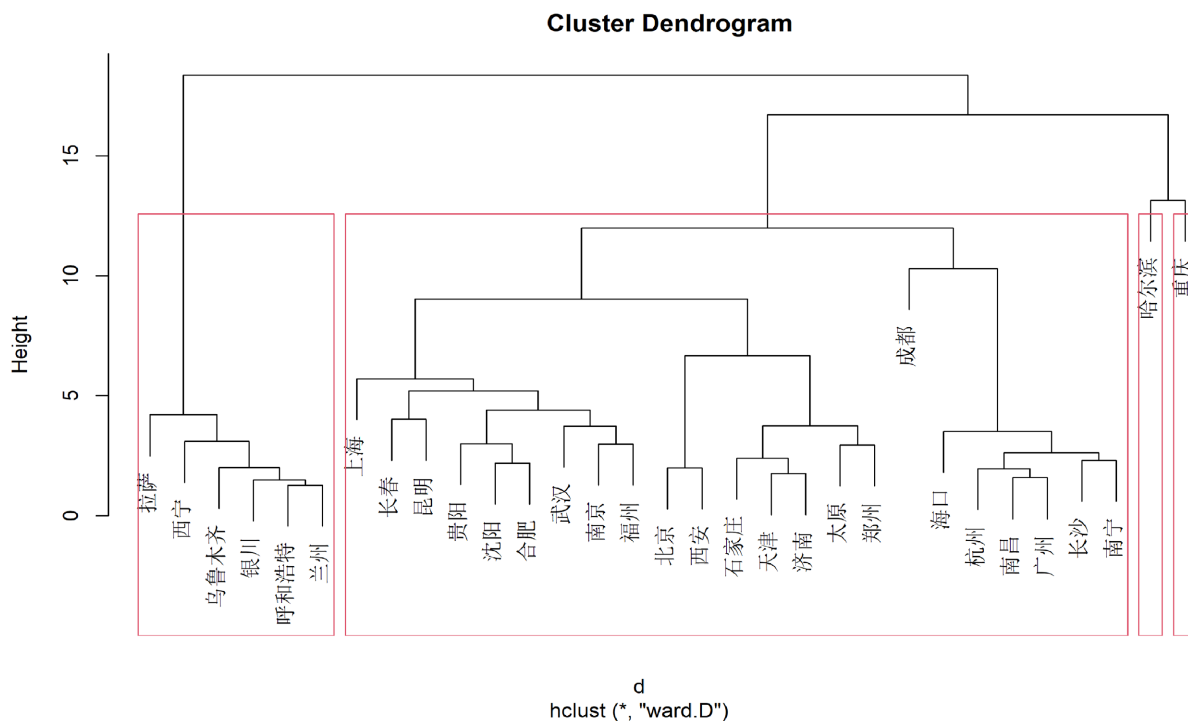
因子分析法是从研究变量内部的依赖关系出发, 把一些具有错综复杂关系的变量归结为少数几个综合因子的一种多变量统计分析方法。它的基本思想是将观测变量进行分类, 将相关性较高, 即联系比较紧密的分在同一类中, 而不同类变量之间的相关性则较低, 那么每一类变量实际上就代表了一个基本结构, 即公共因子。对于所研究的问题就是试图用最少数量的不可测的所谓公共因子的线性函数与特殊因子之和来描述原来观测的每一分量。

可以把因子分析看成是主成分分析的推广, 即可从研究相关系数矩阵内部的依赖关系出发, 把一些具有错综复杂关系的变量归结为少数几个综合因子。因子分析可用于对变量或样品进行分类处理。根据因子的得分值, 在因子轴所构成的空间中把变量或样品点画出来, 达到形象直观的分类目的。

## 4. 实证分析

### 4.1. 聚类分析

在该系统聚类过程中, 采用欧式度量, 选择 Ward.D 的方法, 考虑到各个指标的度量单位不同, 对数据进行了标准化。图 3 为样品被分成四类时的聚类图和分三类时和四类时的结果表:



**Figure 3.** Clustering diagram when the systematic clustering method is divided into four categories  
**图 3.** 系统聚类法分四类时聚类图

从聚类图可以看到, 无论分成三类还是四类, 第一类和第二类的聚类结果是相同的, 第一类结果中的城市都是相对比较繁华、经济环境发展较好, 地理位置偏向内陆的城市, 第二类城市多为中国的西北地区或者少数民族自治区, 这一类地区工业和经济发展相对不发达, 所以污染排放也相对少, 空气质量应相对较好。最后, 分成三类或四类的区别在于重庆和哈尔滨是否分在了同一类, 观察重庆和哈尔滨的变量指标, 可以看到, 重庆的工业废水 COD 排放量、工业氨氮排放量、生活废水 COD 排放量等多个污染排放类指标的数值都可以算得上是处在 31 个主要城市中最高一层的, 甚至有几个指标是其中的最高值, 所以重庆在污染排放方面的情况堪忧, 迫切需要进行改善。而哈尔滨的情况和重庆类似, 生活及其他二氧化硫排放量、生活及其他氮氧化物排放量、生活及其他颗粒物排放量三个生活类污染排放指标在所有城市中是处在最高值的范围内的, 根据这些特征, 所以重庆和哈尔滨相似, 在分成 3 类的时候被分成了同一类。从聚类结果可以看出, 重庆和哈尔滨这两个城市的污染排放有些偏高, 需要对这两个城市采取措施来控制污染排放量的措施。

使用系统聚类法和 K-Means 算法得到的聚类结果如表 2 所示。

从 K-Means 聚类算法得到的结果可以看到, 分成四类的时候与系统聚类算法相似, 哈尔滨和重庆都被单独分成了一类, 说明重庆和哈尔滨这两个城市的数据特征确实与其他城市有所不同, 虽然其余两类的划分有所区别, 但是我们结合数据表可以得到的结论是重庆和哈尔滨两座城市的在污染排放方面存在问题, 所以才会在分类的时候单独被分出一类。下面将对这 31 个主要城市的受污染情况进行评价分析。

## 4.2. 评价分析

### 4.2.1. 主成分分析法

首先计算主成分对象的方差贡献率和确定主成分个数, 要使得方差累积贡献率不低于 80%。经过计算(使用相关阵), 可以得到表 3 的主成分方差信息表。

**Table 2.** Organizing results into four categories using hierarchical clustering and K-Means clustering methods  
**表 2.** 系统聚类法和 K-Means 聚类法分四类整理结果表

分类法	第一类	第二类	第三类	第四类
系统聚类法 (分四类)	北京、天津、石家庄、太原、沈阳、长春、上海、南京、杭州、合肥、福州、南昌、济南、郑州、武汉、长沙、广州、南宁、海口、成都、贵阳、昆明、西安	呼和浩特、拉萨、兰州、西宁、银川、乌鲁木齐	哈尔滨	重庆
K-Means 聚类算法 (分四类)	北京、天津、石家庄、太原、呼和浩特、沈阳、长春、济南、昆明、拉萨、兰州、西宁、银川、乌鲁木齐	上海、南京、杭州、合肥、福州、南昌、郑州、武汉、长沙、广州、南宁、海口、成都、贵阳、西安	哈尔滨	重庆

**Table 3.** Variance information table for each principal component  
**表 3.** 各主成分方差信息表

主成分	方差	贡献率	累计贡献率
主成分 1	5.343518	0.33396989	0.3339699
主成分 2	3.471963	0.21699767	0.5509676
主成分 3	2.424188	0.15151173	0.7024793
主成分 4	1.434860	0.08967874	0.7921580
主成分 5	1.231511	0.07696943	0.8691275

从表 3 的主成分信息表可以看出当保留 5 个主成分的时候, 主成分累计贡献率才达到 80% 以上, 累计贡献率为 86.9128%, 因此可以确定主成分的个数为 5 个, 接下来, 求得主成分的载荷如下表 4:

**Table 4.** Loadings table for the 5 principal components  
**表 4.** 五个主成分的载荷表

变量	主成分 1	主成分 2	主成分 3	主成分 4	主成分 5
$x_1$	0.18349	-0.36445	0.15326	0.31445	-0.13009
$x_2$	0.19735	-0.41088	0.05745	0.1424	-0.08107
$x_3$	0.26485	-0.24967	0.31795	0.16415	0.01115
$x_4$	-0.27232	0.24005	-0.23456	0.07974	0.00923
$x_5$	0.35294	-0.04724	-0.07854	0.11022	-0.14226
$x_6$	0.37959	0.01539	-0.09532	0.12224	-0.00758
$x_7$	0.30181	-0.01393	0.18485	-0.50545	-0.01472
$x_8$	0.26461	-0.00592	0.10831	-0.61201	0.09289
$x_9$	0.24952	0.24802	-0.26545	0.05234	0.34484
$x_{10}$	0.34271	0.22717	-0.2159	0.07126	0.06195
$x_{11}$	0.31918	0.16544	-0.27325	0.03937	0.15478
$x_{12}$	0.09883	0.36982	0.38387	0.21368	0.051
$x_{13}$	0.11126	0.35425	0.37513	-0.01539	-0.16533
$x_{14}$	0.00669	0.36465	0.42348	0.22585	-0.00255
$x_{15}$	0.22062	0.13361	-0.29813	0.20192	-0.42388
$y$	0.02939	-0.14876	0.08891	0.20824	0.76865



从表 4 的主成分的载荷表中可以看到:

主成分一主要由工业废水 COD 排放量、工业氨氮排放量、工业二氧化硫排放量、工业氮氧化物排放量、工业颗粒物排放量决定, 这五个指标是总量指标, 代表城市的工业排放物污染水平。

主成分二主要由生活及其他二氧化硫排放量、生活及其他氮氧化物排放量、生活及其他颗粒物排放量、降水量、平均气温决定, 这五个指标可以表示为自然天气影响情况和生活空气类污染排放水平。

主成分三主要由工业二氧化硫排放量、工业氮氧化物排放量、工业颗粒物排放量、生活及其他二氧化硫排放量、生活及其他氮氧化物排放量、生活及其他颗粒物排放量、一般工业固体废物综合利用率七个指标决定, 这七个指标可以表示为空气类污染物工业和生活总污染水平。

主成分四主要由生活废水 COD 排放量、生活氨氮排放量、降水量三个指标决定, 这三个指标可以表示为生活类水污染水平和自然降水情况。

主成分五主要由一般工业固体废物综合利用率、空气质量优良天数比例(百分比)、工业二氧化硫排放量三个指标决定, 这三个指标可以表示为空气污染改善措施生效情况。

16 个指标最终可以用 5 个指标来代替, 而综合指标的信息没有损失多少。最终利用线性加权的方法, 以各主成分的贡献率为权数, 计算各综合指标的综合得分并据此排名。

#### 4.2.2. 因子分析法

##### 1) 检验因子分析的可行性

首先, 在进行因子分析之前, 需验证因子分析方法的可行性与科学性, 本文使用 Bartlett's 球体检验。

Bartlett's 球体检验的目的是检验相关矩阵是否是单位矩阵(Identity Matrix), 单位矩阵是指主对角线为 1, 其余元素都是 0 的  $n$  阶方阵, 显然单位矩阵变量间不相关。如果是单位矩阵, 则认为因子模型不合适。检验的虚无假设为相关矩阵是单位阵, 如果不能拒绝该假设的话, 就表明数据不适合用于因子分析。一般说来, 显著水平值越小( $<0.05$ ), 表明原始变量之间越可能存在有意义的关系; 如果显著性水平很大(如 0.10 以上), 可能表明数据不适宜做因子分析。

对数据进行 Bartlett 球体检验, 得到的检验  $p$  值为小于  $10^{-6}$ , 该值明显小于 0.05 的显著水平, 表明变量之间有很大可能存在有意义的关系, 适合用来做因子分析。如下图 4 为根据数据的 16 个变量的相关矩阵画出来的热力图, 其中线越靠近单元格四侧, 相关系数的值越大。

##### 2) 确定因子个数

本文使用了基于主成分估计的因子分析法, 并且采用方差最大正交旋转法(Varimax)进行正交旋转。

下表 5 为因子个数的方差信息表。

从表 5 中得知, 当因子个数达到 5 个时候, 方差累计贡献率就已经达到了 86.913%, 基本能全面地反映 16 个变量指标的信息, 所以我们选择前五个因子作为公因子。

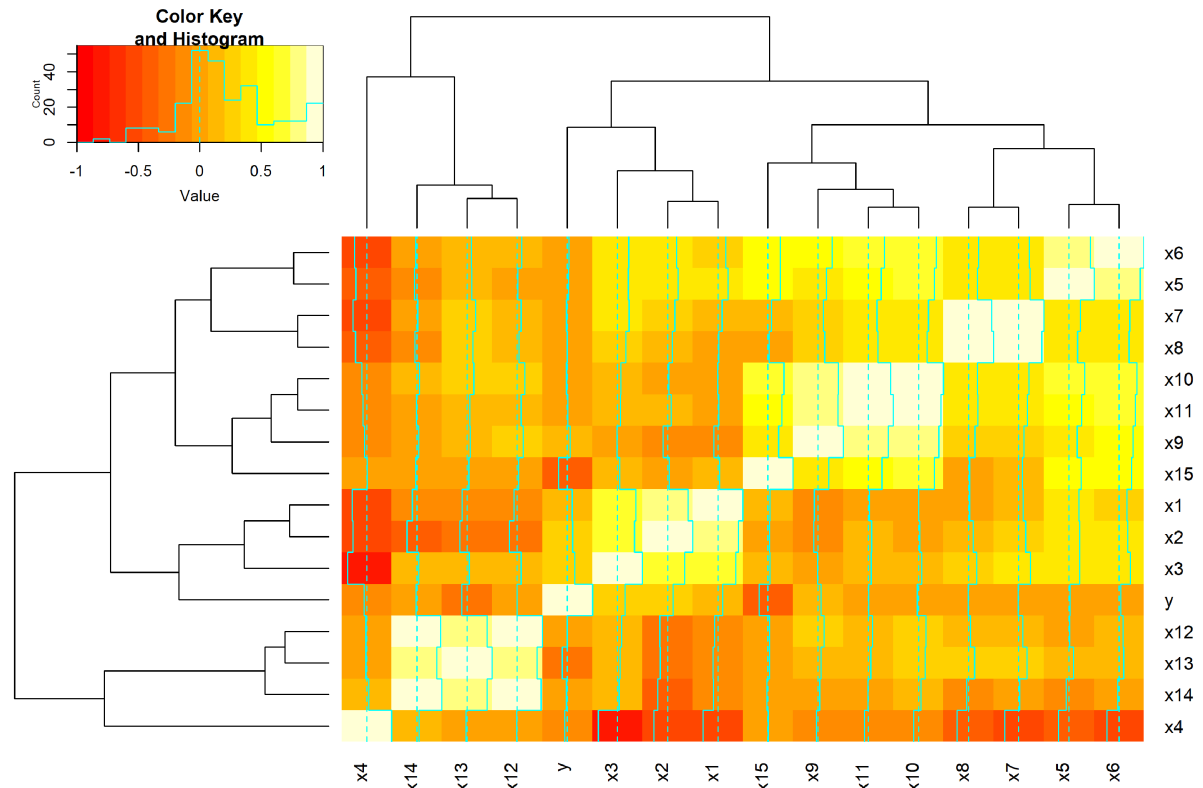
##### 3) 构造因子变量

根据上述确定的公因子个数构造因子变量, 计算出 5 个公因子的载荷如下表 6 所示。

从表 6 中可以看到, 因子 Factor1 在工业氨氮排放量、工业二氧化硫排放量、工业氮氧化物排放量、工业颗粒物排放量、一般工业固体废物综合利用率上的载荷都很大, 可视为反映工业污染处理情况的公因子。

因子 Factor2 在降水量、平均气温、平均相对湿度、日照时数上的载荷都很大, 可视为不可控的自然环境因子。

因子 Factor3 在生活及其他二氧化硫排放量、生活及其他氮氧化物排放量、生活及其他颗粒物排放量上的载荷都很大, 可以认为 Factor3 为生活类空气污染物情况因子。



**Figure 4.** Heat map constructed from the matrix of variable correlation coefficients

**图 4.** 变量相关系数矩阵构造的热力图

**Table 5.** Variance information table for the number of factors

**表 5.** 因子个数的方差信息表

因子个数	方差	贡献率(%)	累计贡献率(%)
Factor1	4.058	25.365	25.365
Factor2	3.478	21.740	47.105
Factor3	2.854	17.835	64.940
Factor4	2.204	13.778	78.718
Factor5	1.311	8.195	86.913

因子 Factor4 在生活废水 COD 排放量、生活氨氮排放量上的载荷很大, 因此可以认为 Factor4 为生活类水污染物排放情况因子。

因子 Factor5 在一般工业固体废物综合利用量、空气质量优良天数比例上的载荷很大, 可以认为因子 Factor5 为污染物改善处理情况因子。

有了对五个公因子的合理解释, 并结合各个主要城市在五个公因子上的得分和综合得分, 就可以对中国各主要城市的污染物处理情况进行评价了。

### 4.3. 综合得分和排名

最后, 31 个主要城市主成分和因子分析法综合得分排名如表 7 所示。

**Table 6.** Loadings of the common factor after orthogonal rotation  
**表 6.** 经过正交旋转后公因子的载荷

变量	Factor1	Factor2	Factor3	Factor4	Factor5
$x_1$	0.003	0.9161	-0.1201	-0.0797	0.0258
$x_2$	0.0381	0.8488	-0.3306	0.0812	0.0402
$x_3$	0.0652	0.8741	0.1728	0.2287	0.1574
$x_4$	-0.0885	-0.7207	-0.0243	-0.4533	-0.075
$x_5$	0.637	0.5097	5.00E-04	0.1905	-0.1698
$x_6$	0.7601	0.433	0.0627	0.2074	-0.0455
$x_7$	0.2218	0.2664	0.1276	0.892	-0.0697
$x_8$	0.2159	0.0996	0.0055	0.9455	0.0123
$x_9$	0.8676	-0.2071	0.081	0.1074	0.2318
$x_{10}$	0.9343	0.0241	0.1575	0.1729	-0.0596
$x_{11}$	0.9061	0.0053	-0.0013	0.172	0.0396
$x_{12}$	0.1595	-0.0286	0.9596	0.0141	0.0749
$x_{13}$	0.0884	-0.0556	0.8784	0.2402	-0.192
$x_{14}$	-0.0371	-0.0929	0.9759	-0.0783	0.0372
$x_{15}$	0.6944	0.1012	-0.0195	-0.1566	-0.5477
$y$	0.0429	0.2023	-0.0514	-0.0712	0.9165

**Table 7.** Ranking of 31 major cities in terms of composite scores of principal components and factor analysis methods  
**表 7.** 31 个主要城市主成分和因子分析法综合得分排名

地区	主成分分析		因子分析	
	得分	综合排名	得分	综合排名
北京	-0.5529	23	-0.3348	25
天津	-0.1124	13	-0.1897	20
石家庄	0.0139	11	-0.1554	18
太原	-0.6521	25	-0.4635	29
呼和浩特	-0.7323	28	-0.4397	27
沈阳	-0.3215	17	-0.1765	19
长春	0.6784	5	0.1834	10
哈尔滨	2.4758	2	0.7009	2
上海	0.8061	4	0.4463	3
南京	0.3884	8	0.2022	8
杭州	-0.1302	14	0.1408	11
合肥	-0.4178	19	-0.0986	17
福州	0.2474	9	0.3092	6
南昌	-0.4567	21	0.0473	14

续表

济南	-0.0754	12	-0.2291	21
郑州	-0.7144	27	-0.3135	24
武汉	0.9344	3	0.422	4
长沙	-0.412	18	-0.008	15
广州	-0.3036	16	0.061	13
南宁	-0.2391	15	0.1178	12
海口	-0.9311	29	-0.0728	16
重庆	4.1245	1	1.8027	1
成都	0.667	6	0.3634	5
贵阳	0.037	10	0.2135	7
昆明	0.4671	7	0.1889	9
拉萨	-1.5756	31	-0.7352	31
西安	-0.4593	22	-0.2371	22
兰州	-0.6351	24	-0.4372	26
西宁	-0.4553	20	-0.2878	23
银川	-0.9924	30	-0.5653	30
乌鲁木齐	-0.671	26	-0.4551	28

从表 7 中可以看到, 受污染情况排名第一的城市是重庆, 其次是哈尔滨, 第三名是武汉。综合得分排名结果和数据特征基本吻合, 同时和聚类结果也有可解释之处, 重庆和哈尔滨的聚类被单独分出来, 和污染情况得分排名第一、第二相契合。接下来, 再使用因子分析法对污染情况进行评价分析。

从最终综合因子得分和排名中, 排名第一和第二的分别是重庆和哈尔滨, 这个排名情况和用主成分分析法得到的排名情况是相同的, 但是用因子分析法得到的第三名是上海, 第四名才是武汉, 而主成分分析法的第三名是武汉, 第四名是上海。

#### 4.4. 研究结论

使用主成分分析法和因子分析法最终得到的结果基本一致, 但是考虑到根据因子分析法构造出来的 5 个公因子更能够清晰的解释和表达出来, 对现实情况来说更具说服力和方便性, 所以本文认为使用因子分析法得到的结果更优, 采用因子分析法得到的 31 个中国主要城市综合得分和排名作为结果。在污染物处理情况评价分析中, 排名第一和第二的城市分别是重庆和哈尔滨, 排名第三和第四的城市分别是上海和武汉。可以认为重庆和哈尔滨两个城市在中国 31 个主要城市中的污染处理情况存在较大的问题, 急需采取措施进行改善治理, 接下来将结合现实情况进行分析和提出改善意见。

##### 针对重庆市治理建议:

- 1) 加强工业企业排放管理: 严格监管重点行业企业, 确保其大气污染物排放符合标准。
- 2) 推进清洁能源使用: 支持和鼓励企业和居民使用清洁能源, 减少对高污染燃料的依赖, 特别是在冬季取暖方面。
- 3) 实施大气污染治理方案: 持续执行冬季大气污染治理方案, 特别是关注散煤治理等措施, 以提高冬季空气质量。

4) 应急预案完善: 不断优化应急预案, 确保在重污染天气情况下, 及时采取有效措施, 减少对空气质量的不利影响。

#### 针对哈尔滨治理建议:

1) 加强工业排放监管: 强化对工业企业排放的监管, 确保其排放符合国家标准, 降低对空气质量的不利影响。

2) 推进清洁能源使用: 支持和鼓励清洁能源的使用, 特别是在工业和供暖领域, 减少对高污染能源的依赖。

3) 提高城市绿化覆盖率: 加大城市绿化力度, 增加植被覆盖, 有助于吸收污染物, 改善空气质量。

4) 完善大气污染治理方案: 持续改进治理方案, 注重细颗粒物(PM<sub>2.5</sub>)的控制, 采取有力措施提高空气质量。

### 参考文献

- [1] 杨洵, 曹聪, 谢亚巍, 等. 重庆主城都市区地下水质量与污染评价——基于层级阶梯评价法[J]. 三峡生态环境监测, 2023, 8(3): 36-44.
- [2] 刘煜椿, 胡春梅, 陈道劲. 重庆春季一次臭氧污染过程分析[C]//中国环境科学学会(Chinese Society for Environmental Sciences). 中国环境科学学会 2022 年科学技术年会论文集(一). 2022: 546-551.
- [3] 韩余, 周国兵, 陈道劲, 等. 重庆地区大气污染分布特征及气象因子相关分析[J]. 四川环境, 2022, 41(3): 79-87.
- [4] 冯雅杰, 綦峥, 韩馥蕊, 等. 哈尔滨市大气 PM<sub>2.5</sub> 中多环芳香烃污染特征及来源解析[J]. 环境化学, 2023, 42(5): 1593-1603.
- [5] 艾昕. 哈尔滨市两次污染事件下降雪中 PAHs 的大气污染指示意义研究[D]: [硕士学位论文]. 哈尔滨: 哈尔滨师范大学, 2023.
- [6] 卜兴兵, 方自力, 俸强, 等. 基于主成分分析的空气质量综合评价研究——以四川省 21 个城市为例[J]. 四川环境, 2023, 42(3): 51-56.
- [7] 李如玥, 刘琳, 陶世鑫, 等. 宜昌市大气污染物变化趋势及污染特征分析[J]. 工业安全与环保, 2021, 47(z1): 58-62.
- [8] 刘兰芳, 谭秉霖, 张可, 等. 基于主成分分析的湖南省城市雾霾灾害风险评估[J]. 灾害学, 2021, 36(1): 76-81.
- [9] 胡荣明, 王睿哲, 李朋飞, 等. 西安市大气污染物时空特征及其与土地利用因素的关系[J]. 中国科技论文, 2021, 16(9): 925-934.
- [10] 朱常琳, 孟双双, 张荣国. 西安市主要大气污染物的相关性分析及时空分布特征[J]. 环境工程, 2017, 35(12): 86-91.
- [11] Ma, T., Duan, F., He, K., Qin, Y., Tong, D., Geng, G., *et al.* (2019) Air Pollution Characteristics and Their Relationship with Emissions and Meteorology in the Yangtze River Delta Region during 2014-2016. *Journal of Environmental Sciences*, **33**, 8-20. <https://doi.org/10.1016/j.jes.2019.02.031>
- [12] Tao, Y., Zhong, L., Huang, X., Lu, S., Li, Y., Dai, L., *et al.* (2011) Acute Mortality Effects of Carbon Monoxide in the Pearl River Delta of China. *Science of the Total Environment*, **410**, 34-40. <https://doi.org/10.1016/j.scitotenv.2011.09.004>
- [13] Khanna, N. (2000) Measuring Environmental Quality: An Index of Pollution. *Ecological Economics*, **35**, 191-202. [https://doi.org/10.1016/s0921-8009\(00\)00197-x](https://doi.org/10.1016/s0921-8009(00)00197-x)
- [14] Perez, P. (2012) Combined Model for PM<sub>10</sub> Forecasting in a Large City. *Atmospheric Environment*, **60**, 271-276. <https://doi.org/10.1016/j.atmosenv.2012.06.024>
- [15] Kalapanidas, E. and Avouris, N. (2001) Short-Term Air Quality Prediction Using a Case-Based Classifier. *Environmental Modelling & Software*, **16**, 263-272. [https://doi.org/10.1016/s1364-8152\(00\)00072-4](https://doi.org/10.1016/s1364-8152(00)00072-4)