

# Analysis on the Discharge of Main Pollutants from Wastewater in China

Lin Zhang, Xiang Gao\*

School of Mathematical Sciences, Ocean University of China, Qingdao Shandong  
Email: linzhang312@163.com, gaoxiangshuli@126.com

Received: Mar. 29<sup>th</sup>, 2018; accepted: Apr. 16<sup>th</sup>, 2018; published: Apr. 23<sup>rd</sup>, 2018

---

## Abstract

Life was originally born in the water. Water is one of the basic life-sustaining substances that are indispensable to life and are indispensable to our life. Our country is experiencing the problems of water shortage and serious water pollution. We must recognize the importance of protecting water resources and preventing and controlling water pollution. We must also recognize that protecting the water environment is sacrosanct and obligatory responsibility for mankind. This paper collects the water pollution discharges in 31 major areas in China in 2016 and analyzes them using SPSS software and R language. The main pollutants include 12 indicators of chemical oxygen demand, ammonia nitrogen, total nitrogen, total phosphorus, petroleum, Fermented fermentation, lead, mercury, cadmium, hexavalent chromium, chromium, arsenic. Firstly, the correlation analysis was made on 12 indicators, and it was found that there was a strong correlation between variables. Considering the method of dimension reduction to process the data, 12 principal components were used to extract the 3 principal components according to relevance, then cluster analysis of 31 provinces, autonomous regions and municipalities. Based on the principle of systematic clustering, 31 regions are divided into 4 categories. According to the characteristics of each type of region, the sources of pollutants are analyzed. Finally, the corresponding control measures are put forward for the current situation of water pollution in China.

## Keywords

Water Pollution, Principal Component Analysis, Cluster Analysis

---

# 我国废水中主要污染物排放情况分析

张琳, 高翔\*

中国海洋大学数学科学学院, 山东 青岛  
Email: linzhang312@163.com, gaoxiangshuli@126.com

---

\*通讯作者。

收稿日期: 2018年3月29日; 录用日期: 2018年4月16日; 发布日期: 2018年4月23日

## 摘要

生命最初诞生在水中, 水是生命的基础物质之一, 是维持生命所不可缺少的, 也是我们生活中所不可缺少的物质。我国正面临水资源短缺, 水体污染严重等问题, 我们必须认识到保护水资源、防治水污染的重要性, 更要认识到保护水环境是全人类神圣和义不容辞的责任。本文搜集了2016年全国31个省、自治区、直辖市的水污染排放情况并运用SPSS软件和R语言进行了分析, 主要污染物有12个指标, 化学需氧量, 氨氮, 总氮, 总磷, 石油类, 挥发酚, 铅, 汞, 镉, 六价铬, 铬, 砷。首先对12个指标做相关性分析, 发现变量之间存在很强的相关性, 考虑采取降维的方法对数据进行处理, 将12个指标用主成分法根据相关性提取出3个主成分, 然后将31个省、自治区、直辖市进行聚类分析, 根据系统聚类法原则将31个地区分为4类, 根据每一类的地区特点分析污染物来源, 最后对我国水污染现状提出了相应的治理措施。

## 关键词

水污染, 主成分分析, 聚类分析

Copyright © 2018 by authors and Hans Publishers Inc.

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

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



Open Access

## 1. 引言

近年来, 随着国民经济的迅猛发展, 大量生活污水和生产废水对生态环境造成破坏, 严重的水污染危机又摆在了中国人面前: 根据世界卫生组织(WHO)调查统计, 现阶段已经发现的水污染物种类为 2221 种, 有机化合物和金属为主要污染物。每年因水污染而引发的疾病导致全世界几千万人的死亡。党的十九大报告指出, 必须树立和践行绿水青山就是金山银山的理念, 形成绿色发展方式和生活方式。可见在未来很长一段时间内保护生态环境治理水污染仍是我国实现经济社会可持续发展的重点。

国内外学者用不同的方法从不同的角度对我国水污染情况进行了分析。王慧敏, 傅涛对 1992~2010 年苏、锡、常地区经济增长与水环境污染情况运用 ARDL 模型与 ECM 模型进行短息动态调整分析[1]。孙世明等(2012)为了判断城市水污染的主要来源, 利用 GIS 空间聚类方法对温州市水体污染情况进行了分析[2]。张文艺等(2010)对温州市水环境质量, 污水排放情况, 污染因子及其贡献等方面进行了综合分析[3]。因子分析法、主成分分析法、聚类分析法[4] [5]也为水环境的分析提供了相关依据。基于以上探究, 本文将相关性分析、主成分分析和聚类分析相结合运用于废水中主要污染物情况的研究中, 对我国水环境发展提出相关建议和措施。

## 2. 相关性分析

本文选取了《中国统计年鉴 2017》[6]中分地区废水中主要污染物排放情况 31 个省、自治区、直辖市的 12 个指标进行分析。首先进行变量相关性分析, 相关性分析是指对两个或多个具备相关性的变量间的相关密切程度进行度量的统计方法。各变量间的相关性如表 1 所示:

**Table 1.** Covariance matrix**表 1.** 协方差矩阵

	化学需氧量	氨氮	总氮	总磷	石油	挥发酚	铅	汞	镉	六价铬	总铬	砷化
化学需氧量	1	0.97	0.95	0.96	0.48	0.03	0	0.41	0.3	0.62	0.64	0.21
氨氮	0.97	1	0.99	0.97	0.52	0.04	0	0.33	0.2	0.68	0.72	0.09
总氮	0.95	0.99	1	0.97	0.52	0.03	0	0.3	0.2	0.7	0.77	0.05
总磷	0.96	0.97	0.97	1	0.49	0.05	0	0.29	0.3	0.7	0.77	0.1
石油	0.48	0.52	0.52	0.49	1	0.28	0	0.22	0.1	0.4	0.38	0.12
挥发酚	0.03	0.04	0.03	0.05	0.28	1	-0	-0	0	-0.07	-0	-0.02
铅	0.38	0.32	0.3	0.32	0.1	-0.02	1	0.7	1	0.1	0.06	0.65
汞	0.41	0.33	0.3	0.29	0.22	-0.04	1	1	0.7	0.11	0.01	0.65
镉	0.28	0.24	0.23	0.25	0.08	0.01	1	0.66	1	0.06	0.02	0.55
六价铬	0.62	0.68	0.7	0.7	0.4	-0.07	0	0.11	0.1	1	0.86	-0.04
总铬	0.64	0.72	0.77	0.77	0.38	-0.02	0	0.01	0	0.86	1	-0.13
砷化	0.21	0.09	0.05	0.1	0.12	-0.02	1	0.65	0.6	-0.04	-0.1	1

从变量之间的相关图中可以看出化学需氧量, 氨氮, 总氮, 总磷三者之间的相关性非常强, 都达到了 95% 以上, 而挥发酚与铅, 汞, 六价铬, 总铬, 砷之间, 砷与六价铬, 总铬之间存在负相关性, 其余变量之间存在正相关性。因为变量较多, 而变量之间的相关性较强, 接下来我们用主成分分析对变量进行降维。

### 3. 主成分分析

主成分分析是将多指标化为少数几个综合指标的一种统计方法, 当变量较多时, 在高维空间中研究样本的分布规律比较复杂, 人们希望用较少的综合变量来代替原来较多的变量, 而这几个综合变量又能够尽可能多的反映原来变量的信息, 并且彼此之间互不相关[7]。图 1 是主成分分析结果。

通过碎石图可以看出前三个主成分值比较大, 均大于 1, 而且前三个主成分以后图形趋于平缓, 所以选取前三个主成分作为 12 个变量的概括。接下来对三个主成分进行提取。

图 2 是用最大方差法进行正交旋转, 对 3 个主成分进行选取, 从 12 个变量与 3 个主成分之间的相关性可以看出, 第一主成分与化学需氧量、氨氮、总氮、总磷、六价铬、总铬之间的相关性较大, 第二主成分与铅、汞、镉、砷之间的相关性较大, 第三主成分与石油、挥发酚之间的相关性较大, 3 个主成分共解释了 12 个变量 82% 的信息, 各主成分的方差贡献率及累积方差贡献率如表 2 所示。

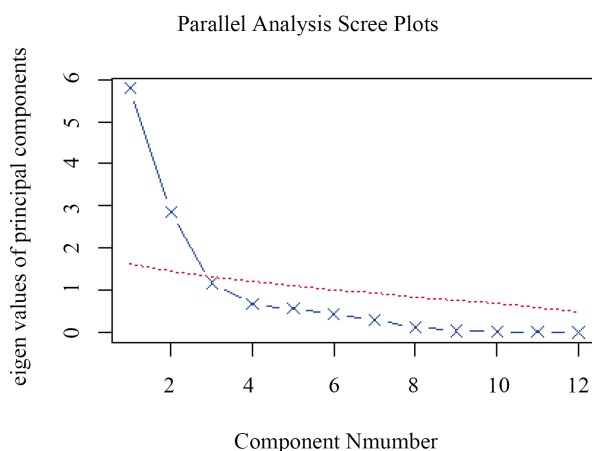
### 4. 聚类分析

聚类分析是研究对样品或指标进行分类的一种多元统计方法, 所谓“类”, 通俗地说就是相似元素的集合[7]。如表 3、图 3, 以下用系统聚类法对 31 个省、自治区、直辖市进行分类。

根据聚类分析结果, 最终将 31 个地区分为四类, 如表 4 所示。

第二类河北、江苏、浙江、福建、山东、广东废水排放物中的氨氮、总氮、六价铬、总铬的含量较高, 在第二主成分中得分较高。主要来源是工业废水。

第三类内蒙古、云南、甘肃废水排放物中铅、镉、砷的含量较高, 在第二主成分的得分较高。污染物主要来源是含重金属、石油、农药及采矿等工业的废水。



**Figure 1.** Gravel map

**图 1.** 碎石图

	RC1	RC2	RC3	h2	u2	com
化学需氧量	0.89	0.29	0.12	0.90	0.103	1.3
氨氮	0.94	0.20	0.13	0.94	0.059	1.1
总氮	0.96	0.16	0.11	0.95	0.046	1.1
总磷	0.95	0.19	0.11	0.94	0.056	1.1
石油	0.50	0.09	0.60	0.62	0.382	2.0
挥发酚	-0.07	-0.03	0.91	0.83	0.165	1.0
铅	0.16	0.93	-0.03	0.90	0.100	1.1
汞	0.17	0.84	0.04	0.74	0.263	1.1
镉	0.09	0.90	-0.02	0.82	0.183	1.0
六价铬	0.84	-0.06	-0.08	0.72	0.279	1.0
总铬	0.88	-0.14	-0.04	0.90	0.198	1.1
砷	-0.05	0.81	0.05	0.66	0.338	1.0
			RC1	RC2	RC3	
SS loadings			5.31	3.26	1.25	
Proportion Var			0.44	0.27	0.10	
Cumulative Var			0.44	0.71	0.82	
Proportion Explained			0.54	0.33	0.13	
Cumulative Proportion			0.54	0.87	1.00	

**Figure 2.** The main component extraction

**图 2.** 主成分提取

**Table 2.** Eigenvalues of each principal component and variance contribution rate

**表 2.** 各主成分特征值及方差贡献率

主成分	特征值	方差贡献率	累积贡献率
1	5.79	44%	44%
2	2.86	27%	71%
3	1.17	10%	82%

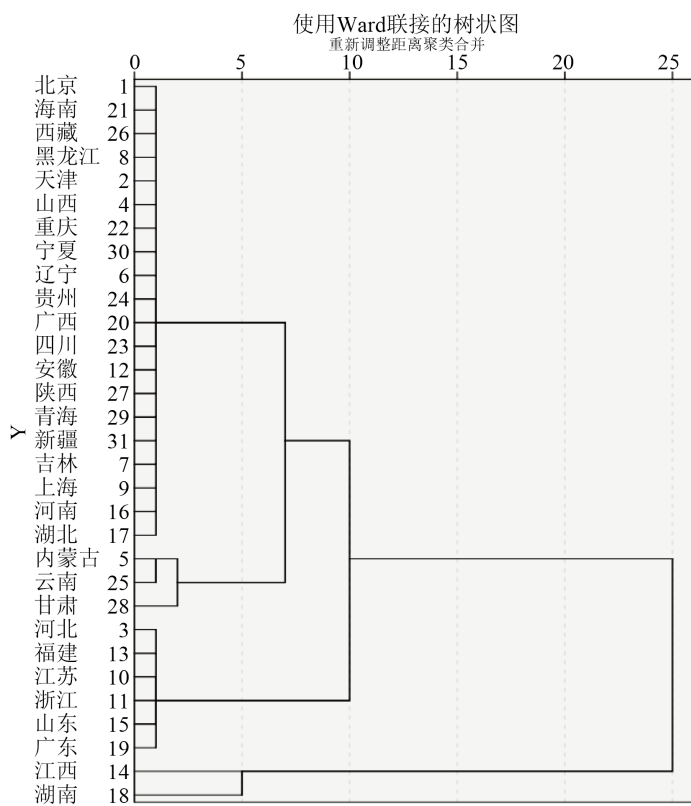
第四类江西、湖南废水排放物中的化学需氧量，铅，汞，镉，砷的含量几乎都是最高的，在第二主成分的得分均较高。主要来源是各种富含有机物的废水，比如大便，含油脂的厨房废水和含重金属化合物的工业废水。

剩下的城市废水污染物排放情况比较相似归为一类。

通过以上分析可知，废水中的污染物主要来源是各种工厂排放的工业废水，生活污水占比较小。

**Table 3.** Cluster analysis  
**表 3.** 聚类分析

案例	4 群集	5 群集	案例	4 群集	5 群集
1: 北京	1	1	17: 湖北	1	1
2: 天津	1	1	18: 湖南	4	5
3: 河北	2	2	19: 广东	2	2
4: 山西	1	1	20: 广西	1	1
5: 内蒙古	3	3	21: 海南	1	1
6: 辽宁	1	1	22: 重庆	1	1
7: 吉林	1	1	23: 四川	1	1
8: 黑龙江	1	1	24: 贵州	1	1
9: 上海	1	1	25: 云南	3	3
10: 江苏	2	2	26: 西藏	1	1
11: 浙江	2	2	27: 陕西	1	1
12: 安徽	1	1	28: 甘肃	3	3
13: 福建	2	2	29: 青海	1	1
14: 江西	4	4	30: 宁夏	1	1
15: 山东	2	2	31: 新疆	1	1
16: 河南	1	1			



**Figure 3.** Dendrogram

**图 3.** 树状图

**Table 4.** Cluster analysis results**表 4.** 聚类分析结果

第一类	北京、天津、山西、辽宁、吉林、黑龙江、上海、安徽、河南、湖北 广西、海南、重庆、四川、贵州、西藏、陕西、青海、宁夏、新疆
第二类	河北、江苏、浙江、福建、山东、广东
第三类	内蒙古、云南、甘肃
第四类	江西、湖南

## 5. 结论及建议

第二类河北、江苏、浙江、福建、山东、广东多为沿海城市，沿海城市相对发展较快，城市化水平较高，生活用水工业废水排放量多，而且沿海城市工业废水直接排入海水中，使得近海海域所面临的生态环境压力较大，沿海城市作为对外开放的前哨，在招商引资，吸引外资方面起着重大作用，因此必须把水污染治理放在经济发展的前列，营造一个和谐优美的环境。

第三类内蒙古、云南、甘肃同属于西部地区，自然资源丰富，自西部大开发战略实施以来，经济得到了快速增长，在经济快速增长的同时，随之带来的是能源的消耗和环境的污染，特别是对煤炭的利用，在西部地区一次能源消费结构中，煤炭比重超过 80%，所以在经济发展的同时，必须注重生态环境的保护。

第四类江西、湖南属于内陆地区矿产资源丰富，工厂大量发展冶炼轧钢等高耗能产业，生产废水直接排入河流，给河流造成严重污染，有关部门必须引起重视，经济发展不能以破坏环境为代价。总而言之，对于水环境污染应从以下几个方面进行治理。

### 1) 预防为主，防治结合

水污染是我国目前面临的最重要的环境问题之一，加强水质监管治理应当采取预防为主防治结合的原则，相比先污染后治理可以最大程度的节约治理成本，即通过事先宣传教育，兴建污水处理厂，对高污染排放企业进行整改甚至关闭，对于已经出现的污染，可以采取化学方法进行相关治理。

### 2) 加强水质监管体制

相关部门必须加强对水环境污染的监控体制，强化污染者责任严惩重罚制度，具体来说就是由相关单位对自身造成的污染负责，有关部门必须采取有效措施对涉事单位污水排放带来的损失及费用进行评估，由涉事单位支付相关费用，并对自己造成的损失负责。

### 3) 污水区别处理

强的自净能力，并不是所有的污水排入海都会造成海水污染，一般生活污水不会造成海水污染，所以没必要对所有污水都进行净化处理，为降低污水治理成本，只需对工业废水或污染性强的污水进行处理即可。

综上，对污水的治理必须采取“防、治、管”相结合的原则，从源头进行治理，从根本上解决水污染问题。

## 基金项目

山东省研究生教育创新计划项目“一线贯通式几何方向研究生课程群及信息化平台建设”(SDYY15129)，山东省研究生导师指导能力提升项目“布尔巴基观点下的几何与代数方向研究生课程体系与课程群建设”(SDYY17009)。

## 参考文献

- [1] 王惠敏, 傅涛. 苏, 锡, 常经济增长与水环境污染的关系——基于 ARDL 和 ECM 的实证分析[J]. 软科学, 2014, 28(1): 91-95.
- [2] 孙世明, 李军, 朱元励, 等. 基于 GIS 空间聚类的温州市水环境污染分析[J]. 浙江农业科学, 2012(7): 1023-1025.
- [3] 张文艺, 孙玲玲, 滕加泉, 等. 加工工业城市——常州市水环境污染分析[J]. 中国农村水利水电, 2010(5): 14-17.
- [4] Zhao, S.R. (2016) Factor Analysis on Quantity of Discharged Pollutants in Wastewater in Major Cities. *Agricultural Science & Technology*, **17**, 964-967.
- [5] 明星, 姚建, 程欢, 等. 基于主成分分析与聚类分析的水污染排放分布研究[J]. 绿色科技, 2016(10): 36-38.
- [6] 中国统计局. 中国统计年鉴[M]. 北京: 中国统计出版社, 2017.
- [7] 高惠璇. 应用多元统计分析[M]. 北京: 北京大学出版社, 2005.

### 知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>  
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2325-2251, 即可查询
2. 打开知网首页 <http://cnki.net/>  
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [sa@hanspub.org](mailto:sa@hanspub.org)