

DOI :10.3969/j.issn.1004-6933.2011.06.002

辽河流域水环境时空差异性评价

付保荣¹, 苏 杰¹, 吴 丹¹, 李法云¹, 徐成斌¹, 谢 轶², 张 峥²

(1. 辽宁大学环境学院, 辽宁 沈阳 110036 2. 辽宁省环境监测实验中心, 辽宁 沈阳 110036)

摘要 采用聚类分析和主成分分析等多元统计方法对辽河流域的水环境时空差异性进行分析评价。依据流域内 4 条主要河流(辽河、浑河、太子河和大辽河)的 25 个监测点 2000—2009 年资料,监测点分为重度污染区、中度污染区和水质清洁区,反映不同采样点的水质水平和污染程度。通过重度污染区和中度污染区的各个站点主成分综合得分的计算,得到辽河流域水质总体优劣状况的排序为:太子河、大辽河、浑河、辽河,2009 年这 4 条河的主成分综合得分均为最小值。

关键词 辽河流域;聚类分析;主成分分析;时空差异性

中图分类号 X824 **文献标识码** A **文章编号** 1004-6933(2011)06-0005-04

Evaluation of temporal and spatial differences of water environment in Liaohe Basin

FU Bao-rong¹, SU Jie¹, WU Dan¹, LI Fa-yun¹, XU Cheng-bin¹, XIE Yi², ZHANG Zheng²

(1. College of Environmental Sciences, Liaoning University, Shenyang 110036, China; 2. Liaoning Environmental Monitoring and Experiment Center, Shenyang 110036, China)

Abstract : The multivariate statistical methods, including cluster analysis and principal component analysis, were used to evaluate the temporal and spatial differences of the water environment in the Liaohe Basin in Liaoning Province. The monitoring data from 25 monitoring sites from 2000 to 2009 for the major rivers, including the Liaohe River, Hunhe River, Taizihe River, and Large Liaohe River, were selected. The monitoring sites were classified as seriously polluted areas, moderately polluted areas, and non-polluted areas, which represent different levels of water quality and water pollution. Through the calculation of the composite scores of the principal components of the sites in the seriously polluted areas and the moderately polluted areas, the descending order of the water quality of the rivers in the Liaohe Basin was determined as follows: the Taizihe River, Large Liaohe River, Hunhe River, and Liaohe River. All the composite scores of the principal components for the four rivers in 2009 were the lowest.

Key words : Liaohe Basin; cluster analysis; principal component analysis; temporal and spatial differences

辽河流域位于我国重工业发达的东北地区,水质污染严重,水环境形势严峻。为了有效地反映流域水环境的质量状况,准确评价河流水质就显得极为重要。由于河流水环境总是处于不断变化的时空之中,所以水质评价必须同时考虑时间和空间两个方面。对水环境动态进行时空一体化研究可以在时

间和空间上把握水环境要素的分布与变化情况,发现水环境的变化规律^[1]。目前,多元统计方法是研究水质时空变化的一个有效工具,在实践中已得到广泛应用^[2-3]。笔者用多元统计分析中的聚类分析和主成分分析方法^[4-5]对辽河流域水环境的时空差异性进行分析评价。

基金项目 国家水体污染控制与治理科技重大专项(2008ZX07526-006);辽宁省科技攻关基金(2005229003);辽宁省教育厅重点实验室项目(2009S044)

作者简介 付保荣(1965—),女,辽宁沈阳人,教授,主要从事环境生物学研究。E-mail: fubaorong@yahoo.com.cn

通讯作者 吴丹,副教授。E-mail: xlanwu001@yahoo.cn

1 研究方法与研究区概况

1.1 研究方法

1.1.1 聚类分析

聚类分析是根据“物以类聚”的道理,其实质是根据变量或样品之间的亲疏程度,以逐次聚合的方法,将特性最接近的对象结合在一起,直到聚成一类。采用常用的欧氏距离和离差平方和法,进行采样点的空间差异性分析,使水质特征相似的采样点聚类在一起。

1.1.2 主成分分析

主成分分析是将多个变量通过线性变换以选出较少个数重要变量的一种多元统计分析方法。笔者通过主成分分析中主成分的确定和综合得分的计算对辽河流域水质时间和空间差异性分析评价。

1.2 研究区概况

辽河流域位于我国东北地区西南部,流域总人口为3029万人。全流域面积21.9万km²,南北长约7.6km,东西宽约490km,属温带、暖温带半湿润大陆性季风气候,季风显著,年平均气温4~9℃,年降水量350~1000mm。辽河流域由辽河、浑河、太子河以及浑河与太子河汇聚而成的大辽河组成,流经11个市36个县(市区),流域面积7万km²,占辽宁省陆域面积的47%。

1.3 资料选取

根据辽宁省环境质量报告书的资料,数据集合2000—2009年辽河流域4个水系(辽河、浑河、太子河和大辽河)25个监测地点(图1)的监测资料^[6],每个监测点选取7个水质指标,包括COD、BOD₅、COD_{Mn}、石油类、挥发酚、DO和NH₃-N。采用的数据分析工具为Excel 2003和SPSS 13.0^[7]。

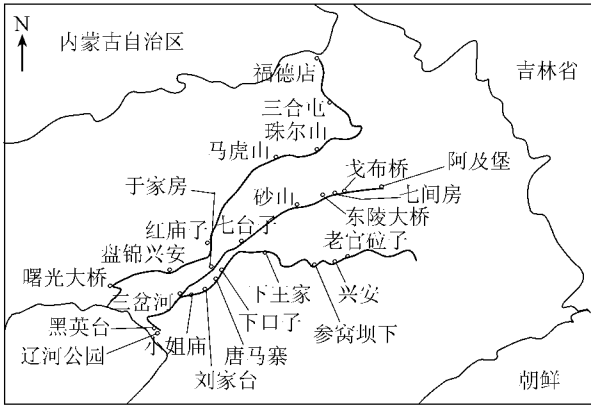


图1 辽河流域水系

2 分析结果

2.1 数据预处理

多元统计方法要求污染指标呈正态或接近正态

分布^[8-10]采用Kolmogorov-Smirnov和Shapiro-Wilk检验判别水质指标是否符合正态分布,检验结果为:概率P值大于0.05时,大部分参数通过了正态性检验,符合正态分布,而氨氮和挥发酚两个参数偏离正态分布,进行对数转化后,所有污染指标符合正态分布。同时,考虑到不同水质指标数量级和计量单位的差异,在进行聚类分析时,对数据进一步进行标准化,即均值为0,方差为1。

2.2 聚类分析水环境因子空间差异性

利用SPSS软件对辽河流域25个监测地点的水质参数进行聚类分析,空间差异性聚类结果如图2所示。在 $(D_{link}/D_{max}) \times 100 < 60$ 时,结合水质污染程度并参照监测地点空间分布图,如图3所示,25个监测地点分为3类,其中第1类用方形表示,包括A1、A2、A3、A4、A5、A6、A7、B5、B6、B7、D1(11个地点,依次为福德店、三合屯、珠尔山、马虎山、红庙子、盘锦兴安、曙光大桥、砂山、七台子、于家房、三岔河);第2类用圆点表示,包括B2、B3、B4、C2、C3、C4、C5、C6、C7、C8、D2、D3(12个地点,依次为戈布桥、七间房、东陵大桥、兴安、参窝坝下、下王家、下口子、唐马寨、刘家台、小姐庙、黑英台、辽河公园);第3类用三角形表示,包括B1、C1(2个地点,依次为阿及堡、老官砬子)。

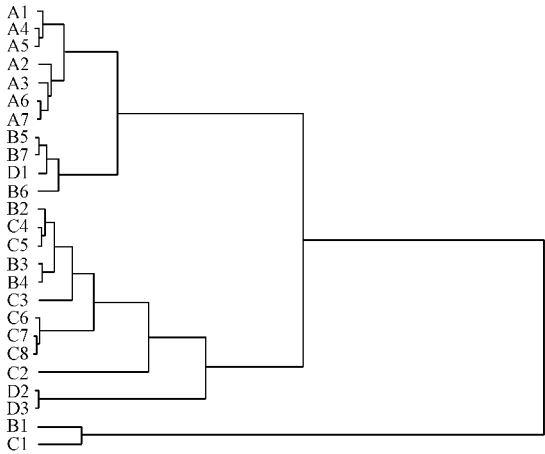


图2 监测地点空间聚类树状图

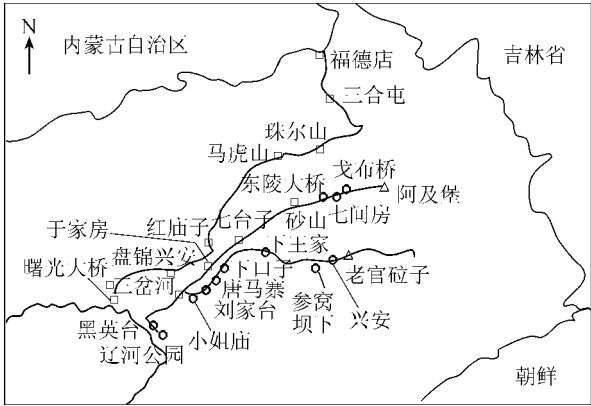


图3 监测地点空间分区

根据辽宁省 2000—2009 年环境质量报告书 第 1 类中的监测地点 A1~A7 位于辽河、B5~B7 主要位于浑河中下游及大辽河的三岔河断面。辽河全程 7 个断面水质均超标,其中 $\rho(\text{COD})$ 年均值 62.0 mg/L,超标 0.6 倍。辽河上游三合屯断面和下游曙光大桥断面 $\rho(\text{COD})$ 在 2006 年达到最高,分别是 119.9 mg/L、102.4 mg/L。浑河水质从上游到下游污染不断加重,首要污染指标是 $\text{NH}_3\text{-N}$,浓度高且污染河段长,其下游 3 个断面 $\rho(\text{NH}_3\text{-N})$ 年均值 10.42 mg/L,为全省之首。浑河、太子河汇合后的三岔河断面受两条河的影响, $\text{NH}_3\text{-N}$ 和 BOD_5 均超标。这些监测点水质均超标,污染严重,属于重度污染区域,故称第 1 类重度污染区;第 2 类中的监测点 B2~B4 主要位于浑河上游、C2~C8 位于太子河、D2 和 D3 位于大辽河下游。浑河上游流经大伙房水库,污染物在水库中得到稀释,污染轻。太子河 $\rho(\text{COD})$ 为辽河流域中最低,年均值 22.4 mg/L,大辽河下游受海水倒灌影响明显,水质污染较轻。这些监测点水质污染轻,属于中度污染区域,称第 2 类中度污染区;第 3 类中 B1 阿及堡位于浑河上游,C1 老官砬子位于太子河源头,均远离主要点源和非点源污染,水质清洁,所以称第 3 类为水质清洁区。

2.3 主成分分析水环境因子时空差异性

2.3.1 Kaiser-Meyer-Olkin 检验

为了验证主成分分析的适用性,首先需进行 Kaiser-Meyer-Olkin(KMO)检验。KMO 值在(0,1)之间变化,一般认为 0.7 以上适合主成分分析,小于 0.5 以下则不合适,介于二者之间的比较适用。本研究中,在 2000—2009 年期间,水质清洁区水质保持良好,因而未对水质清洁区进行主成分分析。在 $P>0.05$ 时,中度污染区和严重污染区的 KMO 值为 0.558,表明主成分分析适用。

2.3.2 主成分的确定

将处理好的数据在 SPSS 软件支持下进行主成分分析,可以得到主成分碎石图(图 4),总方差解释(表 1)及初始因子荷载矩阵(表 2)。根据图 4,从第 1 至第 3 个主成分特征值出现陡峭的拐点,由表 1 可知,前 3 个主成分方差解释原始变量总方差的 93.848%,超过 85%,仅有 6.1%多一点的信息丢失,且特征值大于 1,涵盖了分析数据的大部分信息,所以确定主成分个数为 3 个。由原来的 7 个分析项目

缩减到 3 个,仅用 3 个主成分表达原来 7 个指标所能表达信息的绝大部分。

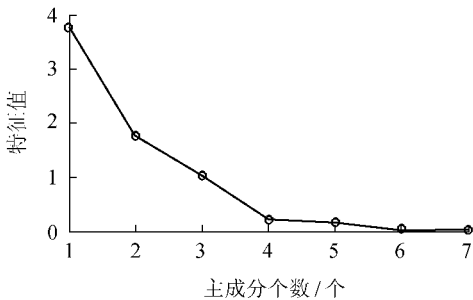


图 4 主成分碎石图

表 1 总方差解释

主成分	初始特征根			提取初始特征根		
	特征根	方差贡献率/%	方差累计贡献率/%	特征根	方差贡献率/%	方差累计贡献率/%
1	3.774	53.911	53.911	3.774	53.911	53.911
2	1.761	25.156	79.076	1.761	25.156	79.076
3	1.035	14.781	93.848	1.035	14.781	93.848
4	0.216	3.087	96.936			
5	0.164	2.339	99.275			
6	0.033	0.468	99.742			
7	0.018	0.258	100.000			

表 2 因子荷载矩阵

初始因子	COD	$\text{NH}_3\text{-N}$	BOD_5	COD_{Mn}	石油类	挥发酚	DO
1	0.939	0.615	0.957	0.927	0.662	-0.132	-0.531
2	-0.021	-0.083	0.151	0.010	0.626	0.914	0.710
3	-0.290	0.751	-0.202	-0.335	0.230	0.223	-0.362

2.3.3 主成分综合得分的计算

主成分综合得分表达式用初始因子荷载矩阵第 i 列向量除以特征根开平方就得到第 i 个主成分的系数向量。根据表 2,利用公式 $A_1 = B_1/\sqrt{3.774}$ (B_1 为第 1 主成分因子荷载矩阵系数,3.774 为第 1 主成分特征值)可以计算出第 1 主成分系数。同理可以确定其他 2 个主成分系数。根据主成分系数和标准化数据相乘再求和,可以计算出各项目 3 个主成分的得分,各主成分得分与各主成分的方差贡献率相乘再求和,即可得到各站点的主成分综合得分。

2.3.4 主成分水环境时空差异性分析

将 23 个水质监测站按照 4 个水系归类,用各水系所有站点每年主成分综合得分的均值作为该分区的得分。具体结果见表 3,根据表中数据制作各水系污染指数时空变化图,结果见图 4。

表 3 各水系主成分综合得分

水系	站点数	2000 年	2001 年	2002 年	2003 年	2004 年	2005 年	2006 年	2007 年	2008 年	2009 年
辽河	7	0.982	1.132	1.280	1.394	1.245	1.055	1.087	1.137	1.023	0.508
浑河	6	1.205	1.151	0.952	0.970	1.037	0.887	0.702	0.551	0.478	0.463
太子河	7	0.821	0.564	0.678	0.802	0.760	0.573	0.488	0.581	0.345	0.356
大辽河	3	1.031	0.531	0.783	0.896	0.737	0.576	0.635	0.564	0.567	0.394

从图 5 可以看出 4 条河的总体状况,水质最好的是太子河,2003 年出现极值的主要原因是挥发酚和硝基苯浓度超标,挥发酚超标是因为本溪钢铁焦化厂部分未经处理的含酚废水超标排放,而硝基苯超标是因为辽阳庆阳化工厂含大量硝基化合物的废水直接排入太子河所造成。水质次好的是大辽河,2000 年出现极值,主要由于 2000 年造纸、印染等企业工业废水集中排放和石油开采污水的排放所引起,导致 COD、COD_{Mn}和 NH₃-N 等严重超标。水质相对较差的是浑河,2000 年出现极值,主要原因是流量变化的影响,由于工农业生产发展需要水资源多功能开发利用,水库拦水闸等水利工程人工调节水量,使浑河流量变化较大,2000 年因浑河流量小,所以水质污染最重。水质最差的是辽河,在此期间水质总体处于较高的污染水平,COD、NH₃-N 4 项指标常年超标。各年度中以 2003 年水质污染最重,是因为 2003 年辽河上游三合屯断面流量很小,仅为 0.02~0.1 m³/s,其中 ρ(COD)高达 317~814 mg/L,致使污染物浓度值高于其他年度。4 条河都在 2009 年出现主成分综合得分极小值,是因为从 2008 年 7 月 1 日起,辽宁境内企业往地表直接排放的水污染物(ρ(COD))最高浓度由 100 mg/L 下降到 50 mg/L,大大减轻了辽河流域的纳污压力。

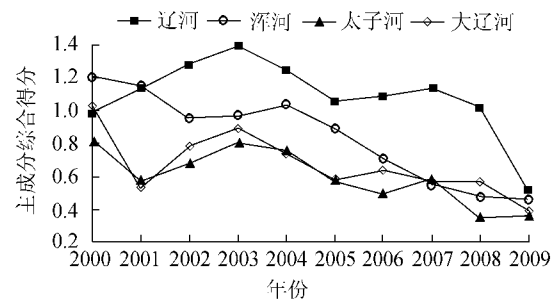


图 5 各水系主成分综合得分时空变化

3 结 论

- a. 空间聚类分析将采样监测点分为 3 类,分别是重度污染区、中度污染区和水质清洁区,反映不同采样点的水质水平和污染程度。
- b. 通过重度污染区和中度污染区的各个站点主成分综合得分的计算,得到辽河流域水质总体优劣状况的排序为:太子河>大辽河>浑河>辽河,2009 年这 4 条河的主成分综合得分均为最小值。从整体上较好地客观地反映辽河流域 4 条主要河流水质情况。

参考文献:

[1]袁中智.水环境时空分析方法研究[J].生态经济,2009 (5):116-119.

[2]李一平,严莹,韩广毅.太湖水质时空相关性分析[J].河海大学学报:自然科学版,2005,33(5):505-508.

[3]周丰,郭怀成,刘永,等.基于多元统计分析和 RBFNNs 的水质评价方法[J].环境科学学报,2007,28(5):846-853.

[4]李大鹏,段沛霞,倪长健.基于投影寻踪动态聚类法的湖泊水质评价模型[J].水资源保护,2007,22(S2):85-87.

[5]李哲强,侯美英,白云鹏.基于 SPSS 的主成分分析在水环境质量评价中的应用[J].海河水利,2008(3):49-52.

[6]辽宁省环境监测实验中心.辽宁省环境质量报告书[M].沈阳:辽宁省环境监测实验中心,2008:49-52.

[7]张红兵.SPSS 宝典[M].北京:电子工业出版社,2007.

[8]PAPATHEODOROU G,DEMOPOULOU G,LAMBRAKIS N. A long-term study of temporal hydrochemical data in a shallow lake using multivariate statistical techniques[J]. Ecological Modelling, 2006, 193: 759-776.

[9]JOHNSON R A, WICHERN D W. Applied multivariate statistical analysis[M]. 5th ed. New Jersey:Prentice-Hall, 1992.

[10]LATTIN J,CARROLL D,GREEN P. Analyzing multivariate data[M]. New York:Duxbury Press, 2003.

(收稿日期:2010-09-14 编辑:高渭文)

(上接第 4 页)

[3]李如忠,汪家权,钱家忠.区域水污染负荷分配 Delphi-AHP 法[J].哈尔滨工业大学学报,2005,37(1):84-88.

[4]陈丁江,吕军,金树权,等.非点源污染河流的水环境容量估算和分配[J].环境科学,2007,28(7):1416-1423.

[5]陈燕华,李彦武,谋海省,等.长江九江段水环境容量研究[J].环境科学研究,1994,7(1):24-29.

[6]李永军,陈余道,孙涛.地理信息模型方法初探河流环境容量:以漓江桂林市区段为例[J].水科学进展,2005,16(2):280-283.

[7]王勤耕,李宗恺,陈志鹏,等.总量控制区域排污权的初始分配方法[J].中国环境科学,2000,20(1):68-72.

[8]包存宽,张敏,尚金城.流域水污染物排放总量控制研究:以吉林省松花江流域为例[J].地理科学,2000,20(1):61-64.

[9]重庆市环境保护局.丹麦王国环境保护署.梁滩河试点工程项目[R].重庆:重庆市环境保护局,2009.

[10]重庆市环境保护局.重庆市水环境功能区划报告[R].重庆:重庆市环境保护局,2008.

(收稿日期:2010-11-01 编辑:高渭文)

