

因子分析法在黄柏河下游水质评价中的应用

张彩香¹, 王焰新¹, 张兆年²

(1. 中国地质大学环境学院, 湖北 武汉 430074 2. 宜昌市环境保护监测站, 湖北 宜昌 443000)

摘要 运用因子分析法对黄柏河下游平湖水域进行综合评价, 通过分析得出, 黄柏河污染主要因子有五大类, 其中富营养化污染最为严重, 其次是油及有毒物质污染。分析结果系统整体地认识和评价了黄柏河下游的污染现状和污染性质, 根据此结果提出改善黄柏河下游水污染现状的生态环境修复方案。

关键词 水质评价; 因子分析; 黄柏河

中图分类号: X824 文献标识码: A 文章编号: 1004-693X(2005)04-0011-04

Application of factor analysis method to the water quality evaluation of lower Huangbaihe River

ZHANG Cai-xiang¹, WANG Yan-xin¹, ZHANG Zhao-nian²

(1. School of Environment, China University of Geosciences, Wuhan 430074, China; 2. Yichang Environmental Monitoring Station, Yichang 443000, China)

Abstract Factor analysis method was applied to the evaluation of water quality of Pinghu area on the lower reach of Huangbaihe River. Five main pollution factors were pointed out. Eutrophication is the most serious one of them, and the next is contaminant of oil and toxic substances. The research provides a scientific basis not only for assessing the state of pollution and pollution characteristics of lower Huangbaihe River, but also for planning the plans for its ecological remediation.

Key words water quality evaluation; factor analysis; Huangbaihe River

因子分析法(FA)是研究一组变量(或样品)之间相关关系的一种多元统计方法,是化学计量学的重要内容。它运用合适的统计方法,对已知的数据进行计算,分析和识别出对高维度变量有公共影响的公因子,并对其加以评析,得出科学合理的评价,通过分析所得信息为决策提供科学的依据^[1]。本文利用因子分析法,通过分析黄柏河下游水质监测数据,建立各因子之间的相关矩阵,从不同环境质量级别的标度^[2]反映群体之间及因子之间的关联性。通过各主因子的方差贡献及因子得分,得出各水质污染程度的综合评价、分析、排序,从而对各断面的污染特性及污染程度做出合理的评价,提出相应的整顿建议。

1 因子分析的原理及基本思路

因子分析的实质就是找出某个问题中可直接测量的、具有一定相关性的诸指标如何受少数几个在专业上有意义且不可直接测得、相对独立的因子支配的规律,从而可以用诸指标的测定值直接确定诸因子的状态,其目的是:用有限个不可观测的潜在变量来解释原变量间的相关性或协方差关系。

因子模型^[3]:

$$\begin{cases} X_1 = a_{11}F_1 + \dots + a_{1n}F_n + \epsilon_1 \\ X_2 = a_{21}F_1 + \dots + a_{2n}F_n + \epsilon_2 \\ \dots \\ X_m = a_{m1}F_1 + \dots + a_{mn}F_n + \epsilon_m \end{cases} \quad (1)$$

式中: X_1, X_2, \dots, X_m 为实测变量; a_{ij} 为因子荷载; F_i ($i = 1, 2, \dots, n$) 为公共因子; ϵ_i ($i = 1, 2, \dots, m$) 为残差。

因子分析的任务就是求出公因子负荷系数 A 和残差 E 。在分析主因子时,为了避免不同量纲、不同数据类型及绝对值大小等对分析结果的影响,必须先对原始数据进行标准化,得到标准化数据矩阵,再计算其相关矩阵,根据相关矩阵解特征值,由方差累计贡献率(一般取值在 85% 以上)确定因子个数,在此基础上建立因子载荷矩阵。当因子意义不明显时,应旋转初始因子。经比较一般采用方差极大(Varimax)旋转,从而构造主因子得分和综合因子得分,最后给出正确评价。整个分析过程可用统计分析软件 SPSS(Statistical Package for the Social Science)执行。

2 因子分析法评价黄柏河下游水质

黄柏河是长江一级支流,下游平湖区域地处葛洲坝左坝头上首。葛洲坝工程建成后,库区水位大幅度抬高,尾水深入黄柏河。由于黄柏河下游区域远离长江主流,枝蔓出来的港汊状水库的水体与外界交换一次的时间变得非常长,基本上成为一泓静水,形成一人工湖泊。水库建成后河水流速减缓,支流泥沙沉降增加,从而壅高河口河床,乃至形成“拦门沙”,使内水与外水更难交换。随着经济的发展,黄柏河上游已发展成为磷矿产业聚集地,大小磷矿十几个,该水域已成为矿石、砂石、原煤、货物、客运等运转的中转码头以及船舶修理服务中心,加上夷陵区域城镇工业的发展和居民生活污水的排放,河口水质逐步恶化,严重影响人们的健康,因此应加强对黄柏河下游水污染趋势及性质的研究。

2.1 水质监测及数据预处理

2000 年,宜昌市对黄柏河风景区及城市绿化景观规划设计方案,公开进行了国际招标,一个世界水准的风景区建设蓝图已初具雏形。因此为了加强治理整顿黄柏河流域的污染问题,2001 年宜昌市环境保护监测站对黄柏河下游平湖区域选择 70 多个监测断

面 18 个污染指标进行监测,具体监测点位见图 1。



图 1 黄柏河下游平湖区域水环境监测点位

由于 Cr^{6+} 、Cd、AS、Pb 等污染因子监测值较低,本次水质评价选择 COD_{Mn} 、 NO_2-N 、 NO_3-N 、 NH_3-N 、TN、非离子氨、TP、氟化物、SS、DO、 BOD_5 、油 12 项指标,其中 11 项指标属于越大污染越严重的污染指标,仅 DO 属于越大越好质量评价指标,在数据处理中选择观察值的倒数,且对同一横断面不同断面深度和断面宽度的监测数据取平均值。将原始数据 x_{ij} ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$, 其中 m 为变量个数, n 为样本点)标准化为均值为 0。方差为 1 的变量^[4]有:

$$x'_{ij} = \frac{x_{ij} - \bar{x}_j}{\sigma_j} \quad (2)$$

其中

$$\bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (3)$$

$$\sigma_j = \left[\frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \right]^{1/2} \quad (4)$$

2.2 水质评价

2.2.1 建立相关矩阵

对标准化数据矩阵计算 $R = \frac{1}{n} x'x$, 得相关矩阵(表 1)。

表 1 水质数据相关矩阵

污染指标	COD_{Mn}	NO_2-N	TN	NO_3-N	NH_3-N	非离子氨	氟化物	SS	DO	BOD_5	油	TP
COD_{Mn}	1.000											
NO_2-N	0.030	1.000										
TN	0.829	0.282	1.000									
NO_3-N	0.303	0.634	-0.177	1.000								
NH_3-N	0.923	-0.129	0.735	-0.342	1.000							
非离子氨	0.700	0.349	0.678	0.247	0.720	1.000						
氟化物	0.168	-0.239	0.097	-0.421	-0.020	-0.229	1.000					
SS	-0.076	-0.162	-0.181	0.390	-0.037	0.198	-0.189	1.000				
DO	0.351	-0.409	0.151	-0.431	0.354	0.010	0.191	-0.069	1.000			
BOD_5	0.807	0.010	0.737	-0.318	0.724	0.517	0.510	-0.167	0.261	1.000		
油	0.098	-0.044	0.333	0.241	0.053	0.237	-0.072	0.342	-0.053	0.098	1.000	
TP	0.855	-0.124	0.504	-0.191	0.827	0.649	0.113	0.298	0.350	0.597	-0.009	1.000

表 2 特征值与贡献率和累积贡献率

因子序号	特征值	贡献率/%	累计贡献率/%	对综合因子的贡献率/%	因子序号	特征值	贡献率/%	累计贡献率/%
1	4.854	40.449	40.449	44.82	7	0.219	1.822	97.002
2	2.519	20.989	61.438	23.26	8	0.157	1.307	98.309
3	1.534	12.782	74.220	14.16	9	0.128	1.064	99.373
4	1.092	9.103	83.323	10.09	10	0.049	0.410	99.783
5	0.831	6.927	90.251	7.68	11	0.018	0.146	99.929
6	0.592	4.930	95.180		12	0.008	0.070	100

表 3 12 项水质污染指标定量在 5 个主因子上的载荷矩阵 P

污染指标	第一主因子	第二主因子	第三主因子	第四主因子	第五主因子	共同度
COD _{Mn}	0.960	-0.157	-0.069	0.121	0.017	0.966
NO ₂ -N	0.170	0.889	-0.234	-0.130	-0.066	0.894
TN	0.835	0.086	-0.318	0.088	0.345	0.933
NO ₃ -N	-0.129	0.777	0.376	-0.285	0.135	0.861
NH ₃ -N	0.931	-0.281	-0.031	-0.078	-0.018	0.953
非离子氨	0.851	0.291	0.165	-0.209	0.116	0.894
氟化物	0.021	-0.181	-0.074	0.963	-0.034	0.967
SS	0.010	0.035	0.947	-0.096	0.190	0.943
DO	0.266	-0.711	-0.042	-0.019	-0.048	0.580
BOD ₅	0.777	-0.070	-0.150	0.523	0.082	0.911
油	0.096	0.043	0.188	-0.026	0.967	0.982
TP	0.845	-0.220	0.380	0.074	-0.181	0.946

2.2.2 特征值和特征向量的计算

根据各变量的相关系数矩阵, 求出特征值 λ_j , 即第 j 个主因子的“贡献”, 并计算主因子方差贡献率, 确定主因子个数为 5 个 (90.251% > 85%) (表 2), 最后确定主因子上的载荷矩阵 (表 3)。

由表 2 可知, 用 5 个主因子代替原来的 12 个因子累计贡献率达 90.25%, 前 5 个因子作为综合因子损失只有 9.75%, 并且在 SPSS 分析中, 各个观测变量的共同度几乎都在 90% 以上。表 3 中 COD_{Mn}、BOD₅、TN、NH₃-N、非离子氨、TP 对第一主因子的贡献明显, 这 6 个指标主要是有机污染物。对第二主因子作主要贡献的是 NO₂-N、NO₃-N 和 DO, 这 3 个指标主要是营养元素, 第一、二主因子是引起水体富营养化的主要指标, 对综合因子的贡献率为 68.08%, 表明黄柏河下游平湖区域的污染现状主要表现为富营养化污染。结合表 2 和表 3 可知, 第三、四、五主因子对综合因子的贡献率分别为 14.16%、10.09%、7.68%。第三个主因子主要是 SS, 第四个主因子主要是无机有毒物质氟化物, 第五个主因子是油污染。

2.2.3 各主因子的得分

旋转主因子载荷矩阵并转置, 求出各主因子的得分系数矩阵 (表 4)。 $F_{ij} = \sum_{k=1}^{12} x'_{ik} \cdot p_{jk}$ (x'_{ik} 为标准化合后第 i 个监测断面第 k 个变量的数值; $i = 1, 2, \dots, 20$; $j = 1, 2, 3, 4, 5$) 求出各监测断面上各主因子的得分, 由 $F_i = \left(\sum_{j=1}^5 F_{ij} \cdot \lambda_j \right) / \sum_{j=1}^5 \lambda_j$ ($i = 1, 2, \dots$,

20) 求出各断面的综合得分 (表 5)。

表 4 主因子得分系数矩阵

污染指标	第一主因子	第二主因子	第三主因子	第四主因子	第五主因子
COD _{Mn}	0.206	-0.021	-0.025	0.005	-0.037
NO ₂ -N	0.087	0.459	-0.169	0.029	-0.135
TN	0.159	0.053	-0.275	-0.021	0.297
NO ₃ -N	0.012	0.345	0.232	0.002	-0.009
NH ₃ -N	0.208	-0.128	-0.034	-0.189	-0.056
非离子氨	0.213	0.138	0.084	-0.136	-0.019
氟化物	-0.054	0.121	0.139	0.799	-0.023
SS	0.004	0.003	0.684	0.110	0.014
DO	0.035	-0.372	-0.046	-0.198	0.022
BOD ₅	0.139	0.099	-0.017	0.374	0.034
油	-0.053	-0.084	-0.022	-0.003	0.864
TP	0.201	-0.034	0.344	0.052	-0.281

2.3 结果与分析

表 5 中综合因子的排序反映了监测断面的污染状况, 主因子的排序反映了各断面污染性质, 由此可得出如下结论:

a. 从各断面的综合因子排序可知, 罗家小河河段综合评价得分最高, 表明该段污染状况最为严重; 其次是庙嘴、3 号船闸、丁家坝冲和冯家湾; 其他断面 (如上游断面黄柏河东西支、两河口和石碑滩) 水质状况较好, 表明上游断面区域污染源少及源强较小, 少量的污染物在水体自净能力的调节下不至于恶化水体。此外, 黄柏河东西支的污染状况差于下游两河口, 可能是由于两支流汇入后增加了水环境容量, 从而减轻了对水体的污染。

表5 黄柏河监测断面主因子得分及污染排序

监测断面	第一主因子得分 F_1	第二主因子得分 F_2	第三主因子得分 F_3	第四主因子得分 F_4	第五主因子得分 F_5	断面综合因子得分 F	F 值排序
黄柏河东支	-0.7823	-0.2433	0.4897	-0.3585	-0.5117	-0.41335	15
黄柏河西支	-0.6281	-0.2050	-0.2406	-0.2770	-0.4682	-0.42717	16
两河口	-0.5557	-1.7957	-0.1525	-1.0270	-0.2944	-0.81457	20
下坪	-0.4299	-0.4741	-0.6884	-0.5641	-0.3222	-0.48210	17
石碑滩上	-0.7373	-0.4471	-0.5619	-0.0369	-0.0495	-0.52154	18
石碑滩下	-0.8568	-0.4303	-0.7155	-0.1854	-0.1170	-0.61311	19
鄢家河	0.2931	0.0503	-0.6552	-0.0827	-0.1501	0.03042	10
丁家坝冲	0.0050	-0.3085	-0.4624	4.1104	-0.2843	0.25791	4
冯家湾	0.2848	0.3599	-0.2482	0.1278	0.1316	0.19922	5
潘家湾	0.1101	0.2367	-0.5398	-0.0776	-0.1187	0.01102	12
罗家小河	3.8952	-1.1078	-0.3846	-0.3639	-0.1351	1.38660	1
对颈窝中	-0.1749	0.3435	-0.3985	-0.3728	-0.4400	-0.12633	14
对颈窝出	-0.1105	0.4825	-0.1265	-0.1212	-0.1521	0.02088	11
黄柏河桥	-0.2467	0.5120	0.1219	-0.0489	-0.3605	-0.00684	13
虾子沟	0.0629	0.6868	0.0631	-0.3315	-0.0628	0.15861	6
夜明珠	0.0477	0.5096	-0.0424	-0.4191	0.0468	0.09522	7
长江溪桥	-0.0364	0.4582	-0.0624	-0.1191	0.1103	0.07788	9
3号船闸	-0.0449	0.3343	0.5476	-0.0271	4.1180	0.44870	3
胡家岗	-0.2531	0.5832	0.5880	0.0791	-0.3929	0.08328	8
庙嘴	0.3432	0.4254	3.7841	0.1417	-0.8087	0.74079	2

b. 从各断面的主因子值可知,罗家小河河段的 F_1 最大,表明该断面有机污染最为严重,河段呈现富营养化状态,这与现场采样观察该段水体已出现“水华”等富营养化现象相符,该河段边人口密度大,农业比较发达,雨量充沛,生活污水及工农业废水对水体富营养化产生严重影响,3号船闸的 F_5 最大,反映该断面油污染最为严重。该水域已成为矿石、砂石、原煤、货物、客运等运转的中转码头以及船舶修理服务中心,两岸沿线的船舶修造厂大量的含油废水、船上无组织管理的漏油、废机油、生活垃圾及粪便等直接排入水体,对水质和沿河环境造成极大污染;丁家坝冲的 F_4 最大,表明该断面的氟化物无机有毒污染最为严重。该河段边工业比较发达,大量的工业无机有毒物质直接进入水体,加重水体的污染。

与目前国内外常用的环境质量评价方法(如综合指数法、灰色关联度、模糊综合评价等)相比,因子分析法尽管不能明确水体污染的类别,但避免了确定权重的主观随意性,直接从水体实际监测数据出发,探索水体的性质及时、空变化规律。

3 黄柏河下游治理整顿建议

根据因子分析结果,提出相应的黄柏河下游治理整顿建议:

a. 为了改善水体富营养化的现状,必须提高当地居民的环保意识,采取强有力的措施减少面源污

染,加强区域城镇生活污水的综合治理,强制推广应用国家环保总局和建设部颁布最佳实用技术“地理式无动力生活污水净化装置”,并进一步深度处理,加快宜昌市集中污水处理厂的建设。

b. 为了从根本上减轻油污染,必须搬迁临时码头,整顿船舶修造业,防止修理过程中直接将船上的废油、杂物、垃圾倾泄水中。

c. 为了减缓工业有毒物的污染,必须采取有效的措施,整顿两岸工矿企业,控制矿产开发,搞好矿山开发的治理与保护,加强排污管理,加大执法力度,真正对上游及流域内所有厂矿企业实现达标排放,对污染物实行总量控制,防止污水和污染物直接进入水体。

d. 适当调度上游水库的排泄水量,三峡和葛洲坝电厂的运行方式适当调整,以利于三峡库区下游及黄柏河的减污。

参考文献:

- [1] 于秀林,任雪松.多元统计分析[M].北京:中国统计出版社,1999.154~162.
- [2] 曾江宁,黄韦良,曾淦宁,等.近岸海水质量量化方法及分类评价[J].海洋学报,2004,26(4):139~146.
- [3] 卢纹岱.SPSS for Windows 统计分析[M].北京:电子工业出版社,2000.444.
- [4] 陈东景,马安青,徐中民.因子分析在水质评价中的应用[J].水文,2002,22(3):29~31.

(收稿日期:2004-12-08 编辑:傅伟群)