

文章编号:1005-3085(2005)07-0059-06

基于回归分析的长江水质预测与控制

路 亮¹, 刘 睿¹, 张 齐²

指导教师: 刘保东²

(1- 山东大学齐鲁软件学院, 济南 250101; 2- 山东大学数学与系统科学学院, 济南 250100)

编者按: 通过数学建模方法, 本文对长江水域受污染的情况作出比较全面和量化的评价, 对污染源进行了比较深入的分析, 得出明确的结论, 同时也对长江未来的水质情况和污水处理形势做出量化的科学预测。特别值得推荐的是, 作者对于污染源的特点和水质的不同性质进行了分类, 对于控制水质与污水处理的策略具有积极的参考作用。作为大学生能够在短时间内, 在一个问题中拓出多处有创意的概念和方法, 实在难能可贵。虽然文章仍有不足, 仍希望引起读者关注, 以期提高中国大学生的创造性能力。

摘 要: 本文采用频度统计和标准化方法对长江近并两年多的水质情况进行了分析和评价, 并根据污染物氨氮和高锰酸盐污染源在时间和空间上的分布特点, 采用一维河道水质模型计算给出了主要污染源的分布。最后讨论了水质的两类特性以及长江水污染的预测与控制。

关键词: 污染源; 污染物; 水质; 统计; 预测

分类号: AMS(2000) 34B05; 91B82

中图分类号: X522

文献标识码: A

1 引言

水资源的保护问题是当今全球范围内讨论的焦点问题之一。本文采用标准化公式对不同成分的污染物统一化处理, 以此对长江水污染进行了定量的综合评价。然后通过分析污染物氨氮和锰酸盐污染源在时间和空间上的分布特点, 得出如下结论: 1) 氨氮属单污染源持续污染型 2) 高锰酸盐属多污染源分时污染型。其中, 赣江(鄱阳湖入口)是长江干流氨氮的主要污染源之一, 而高锰酸盐则有多个污染源。随后我们还引入一维河道水质模型来详细分析污染源。

我们使用指数预测模型对长江后10年废水排放总量预测、应用正态分布模型对长江总流量进行了验证及模拟并且得出了长江水质变化发展的特点: 1) 长江近10年主要特点: II类, III类水具有缓冲性。2) 长江后10年主要特点: I类, 劣V类水具有两极性。然后我们比较了 Logistic 模型、线性回归模型在预测长江水质发展趋势方面的优劣并对两个模型分别进行了综合检验。最后得出了对长江未来10年水质发展趋势估计(2005~2014)不可饮用水占总流量的百分比(%), 通过对回归模型进行控制, 保证了一年中长江劣V类水消失的概率为95%, 并且使得长江干流的IV类和V类水的比例在20%以内。

2 问题与思路

题目通过若干附件, 给出了关于长江水质的统计数据。包括长江沿线17个观测站(地区)近两年多主要水质指标的检测数据, 以及干流上7个观测站近一年多的基本数据(站点距离、水流量和水流速)。通常认为一个河段(地区)的水质污染主要来自于本地区的排污和上游的污水。一般说来, 江河自身对污染物都有一定的自然净化能力, 即污染物在水环境中通过物理

降解、化学降解和生物降解等使水中污染物的浓度降低。反映江河自然净化能力的指标称为降解系数。根据题目的要求，我们认为长江干流的自然净化能力是近似均匀的。

3 基本假设

- 1) 水质评价指标为：溶解氧 (DO)、高锰酸盐指数 (COD_{Mn})、氨氮 (NH₃-N)、pH 值。而忽略水温、渗透压、混浊度（透明度）、色度、悬浮固体、蒸发残渣、含盐量（盐度）、硬度、化学耗氧量等较为复杂的水质情况的评价指标。
- 2) 长江干流的自然净化能力认为是近似均匀的，此处主要污染物高锰酸盐指数和氨氮的降解系数取为0.2（单位：1/天）。
- 3) 不考虑长江支流的自然净化能力（题目中未给出相应数据）。
- 4) 长江的水质均匀变化，污染物浓度不发生剧烈变化。
- 5) 长江在未来的10年自然降雨量均衡分布，冰雪融化速率均衡，忽略发生特大洪水及冰雪大面积融化的情况。
- 6) 假定长江干流河道顺直、平滑。
- 7) 附件中给出的数据都是正确的。

4 水质现状综合评价与分析

4.1 近两年多长江水质现状综合评价

对于问题1，通过对题目附件中给出近两年多（2003.06~2005.9）长江水质的检测数据的分析，我们发现长江干流的水质在近两年内能达到可饮用水的标准（即至少能达到国家规定的Ⅲ类水质级别），满足：溶解氧 (DO)≥5mg/L；高锰酸盐指数 (COD_{Mn})≤6mg/L；氨氮 (NH₃-N)≤1.0mg/L。但有部分支流的污染比较严重，部分支流的高锰酸盐指数 (COD_{Mn})及氨氮 (NH₃-N) 严重超标，这部分污水不经处理直接排入长江，造成了长江水质的下降。当然，通过对目前的数据的分析，我们没有发现长江干流有达不到国家饮用水标准的情况。但这并不能说明长江干流没有污染源，这很可能是由于长江干流流量较大，流速较快，对污水造成了极大的“稀释”和“自然净化”作用。

由题目中的已知数据，利用频度统计得出近两年长江水质情况的定量综合统计如表1所示。

表 1: 近两年长江水质情况的定量综合统计表(按水文年统计，单位：%)

年份	I 类	II 类	III 类	IV 类	V 类	劣 V 类	I、II、III 类合计	IV、V、劣 V 类合计
2003	4.7	41.5	31.3	6.4	5.8	10.3	77.5	22.5
2004	1.2	26.9	39.9	14.8	5.9	11.3	68	32

4.2 各地区水质污染状况分析

为比较和分析各地区水质的污染状况，对附件中给出的17个检测点位4类污染物质近两年多共28次检测数据，采用如下标准化方法：

- 1) 各点位单一污染物的标准指数公式

$$r_{ij} = \frac{\bar{c}_{ij} - \bar{c}_j}{\sqrt{L_j}}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, 3, 4. \tag{1}$$

其中， n 位检测点位个数， $n = 17$ ； m 为检测时段数， $m = 28$ ； c_{ijk} 为第 i 检测点位第 j 类污染物第 k 个月份的检测结果； r_{ij} 为第 i 检测点位第 j 类污染物的标准指数值； \bar{c}_{ij} 为第 i 检测点位第 j 类污染物的检测均值，

$$\bar{c}_{ij} = \sum_{k=1}^m c_{ijk}/m, \quad i = 1, 2, \cdots, n, \quad j = 1, 2, 3, 4,$$

\bar{c}_j 为所有检测点位第 j 类污染物的检测均值

$$\bar{c}_j = \sum_{i=1}^n \sum_{k=1}^m c_{ijk}/(n \times m), \quad j = 1, 2, 3, 4;$$

L_j 为离差平方和， $L_j = \sum_{i=1}^n (\bar{c}_{ij} - \bar{c}_j)^2$ 。

2) 各点位多污染物综合指数公式

$$R_i = \sum_{j=1}^4 r_{ij}. \tag{2}$$

由于各城市 pH 的平均值都满足7-9的范围内，因此本文只考虑污染物 NH_3-N ， DO ， COD_{Mn} ，计算结果见表2所示。

表 2: 各地区各类污染物及综合水质排序指数计算结果

点位名称	NH ₃ -N	DO	COD _{Mn}	综合排序指数
湖北丹江口胡家岭	-0.36	-0.31	-0.13	-0.79
重庆朱沱	-0.27	-0.27	-0.07	-0.61
四川攀枝花龙洞	-0.33	-0.16	-0.1	-0.6
四川宜宾凉姜沟	-0.28	-0.07	-0.05	-0.4
江苏南京林山	0.09	-0.27	-0.12	-0.3
湖北宜昌南津关	-0.17	-0.03	-0.09	-0.28
江西九江河西水厂	0.02	-0.17	-0.11	-0.25
江苏扬州三江营	-0.07	0.01	-0.08	-0.14
安徽安庆皖河口	0.1	-0.12	-0.09	-0.12
湖南岳阳城陵矶	-0.21	0.24	-0.07	-0.04
湖南长沙新港	0.18	-0.15	0.07	0.1
湖北武汉宗关	0.1	0.1	-0.1	0.11
江西九江蛤蟆石	-0.02	0.23	-0.08	0.13
湖南岳阳岳阳楼	-0.12	0.36	-0.06	0.19
四川泸州沱江二桥	0.24	0.11	0.04	0.4
江西南昌滁槎	0.53	-0.2	0.95	1.29
四川乐山岷江大桥	0.57	0.68	0.07	1.32

5 污染物特性分析与污染源定位

对于问题2, 针对主要污染物高锰酸盐和氨氮, 本文通过统计高锰酸盐和氨氮污染物在两年多的28个月检测信息, 发现长江干流的污染主要源自注入长江的几条支流。其中以高锰酸盐为主要污染的区域, 在28个月里总共出现了14次, 主要分布在湖南岳阳岳阳楼, 江西九江蛤蟆石, 四川乐山岷江大桥处, 其污染程度相对较为平均。氨氮的主要污染区域则大不相同, 在28个月里总共出现了48次, 其中鄱阳湖入口的江西南昌滁槎出现了20次占总数的41.7%。

为简便起见, 本文仅给出几个主要污染检测点位的分析结果。

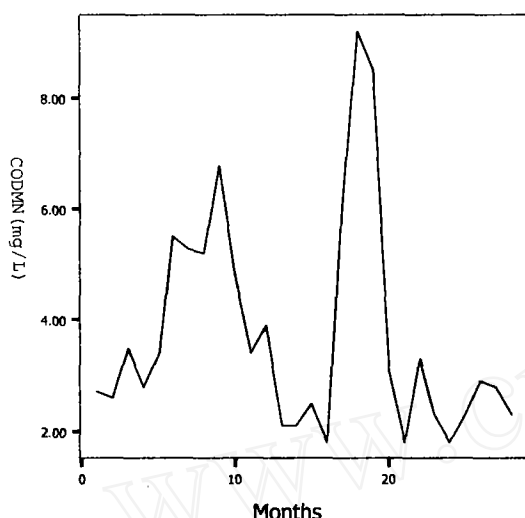


图 1: 江西九江蛤蟆石

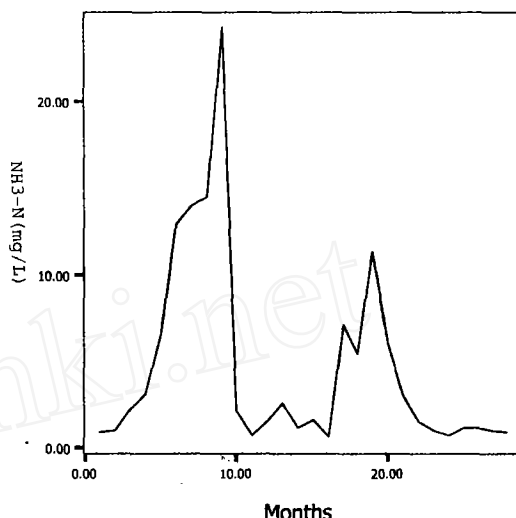


图 2: 江西南昌滁槎

图1是江西南昌滁槎江西九江蛤蟆石(鄱阳湖出口)处2003.06至2005.09的高锰酸盐指数(COD_{Mn})的统计图。图2是赣江(鄱阳湖入口)处2003.06至2005.09的氨氮(NH_3-N)含量的统计图。通过观察、分析, 注意到高锰酸盐的污染与氨氮的污染有不同的特点: 由于以氨氮污染为主的III类与IV类水质的分界点位于 1.0mg/L , 江西南昌滁槎常年处于高污染的状态, 而以氨氮污染为主的江西九江蛤蟆石处波动较大, 多数时间还是以可饮用水为主。同时还注意到, 以上两图在污染的高峰上存在一定的相似, 可能是相同的原因导致污染物的增多。因此综合上述污染物在时间与空间上的分布特点, 我们得出如下结论:

- 1) 氨氮属单污染源持续污染型。
- 2) 高锰酸盐指数属多污染源分时污染型。

为分析长江干流上各河段主要污染源分布情况, 假定污染物浓度的衰减符合一阶化学反应动力方程, 同时对足够长的大中型河流的较长河段, 河宽与河长相比可忽略不计, 即横向和竖向的污染物浓度梯度可以忽略, 而采用只考虑纵向(即水流方向)浓度的推移和衰减变化的一维水质模型来模拟河水的水质: 污染物浓度方程:

$$\frac{\partial c}{\partial t} + u \frac{\partial c}{\partial x} = \frac{\partial}{\partial x} \left(D_L \frac{\partial c}{\partial x} \right) - kc + S \quad (3)$$

式中, u 为流速, c 为污染物浓度, k 为降解系数, D_L 为浓度扩散系数, S 为污染源项, t 为时间, x 为沿河流方向位置。当流速较大时, 污染物的推移作用远大于扩散, 故可忽略 D_L 不计, 同时若假定个污染源排放为连续稳定, 且 $S = 0$, 即污染物浓度变化与时间无关, 此时 $dc/dt = 0$, 则一维模型可简化为:

$$u \frac{\partial c}{\partial x} = -kc. \quad (4)$$

若给定边界条件 $c|_{x=x_i} = c_i$, 则可求得下游断面处的污染物浓度 $c|_{x=x_{i+1}} = c_i e^{-k\Delta x/u}$ 。

记流经站点 i 的流量为 q_i 、流速为 v_i , 进一步若假定流速连续变化, 为计算简便起见, 简化为由站点 i 流至站点 $i+1$ 的流速为平均速度 $\bar{v}_i = (v_i + v_{i+1})/2$, 则流经这一水段需时为 $\Delta t_i = \Delta x_i / \bar{v} = (x_{i+1} - x_i) / \bar{v}$, Δx_i 为该河段长度。由质量守恒定律, 河岸两侧单位时间内排入该水段内的污染质量为

$$m_i = q_{i+1}c_{i+1} - q_i c_i - q_i c_i e^{-k\Delta t_i}, \quad (5)$$

其中 k 为污染物降解速率系数。则整个干流的单位时间内纳污量为 $M = \sum_{i=1}^7 m_i$, 评价河段 i 及支流的污染情况均为

$$\text{deg}_i = \frac{m_i}{M}. \quad (6)$$

经代入近一年多长江干流个断面的水文观测资料, 和附件一中相应的水质资料, 最后求出重庆-湖北段, 湖北-湖南段, 湖南-九江段高锰酸盐污染严重, 湖南-九江段, 湖北-湖南段, 重庆-湖北段氨氮污染严重, 以上河段为干流上的主要污染源。

6 水质特性分析与回归预测控制

对问题3和问题4, 因为题目中未给出相应的水文参数和水质参数资料, 故无法利用一维或高维模型来进行预测计算。为了在宏观上正确地把握长江未来水质污染的发展趋势, 本文依照过去10年的主要统计数据首先对长江的水质特性作了一些简单的探索。图3和图4给出了长江干流过去10年可饮用水与污染水(不可饮用水, 即IV类、V类和劣V类水)百分比统计结果。

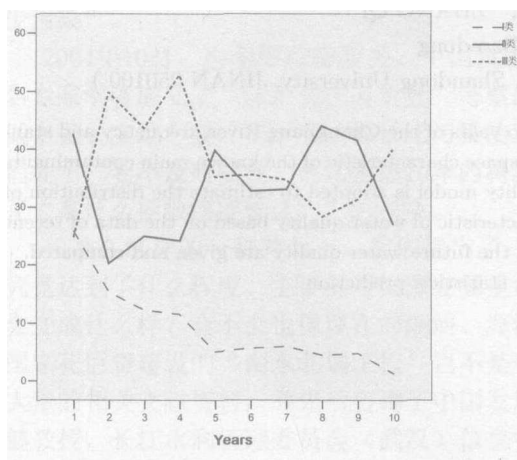


图 3: 长江干流可饮用水百分比统计图

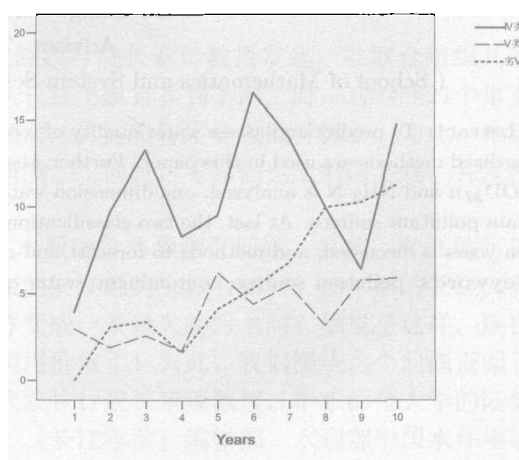


图 4: 长江干流污染水百分比统计图

通过观察统计图3和图4,可以发现污染类 II 与 III 的线图呈现出此消彼长的走向并且很像一只蝴蝶结,具有极强的共线性,如果把两者相合并就会显示出非常平稳的走势,我们称之为“缓冲区”。进一步地,将 II 与 III 类合并,IV 与 V 类合并后作图,则可以看出,I 类水质与劣 V 类水质,合并后的 II、III 类与合并后的 IV、V 类水质同样显示出了比较明显的共线性。其中,缓冲区的作用在过去10年中至关重要。随着时间的推移,废水排放量的增多,在近几年中 I 类水锐减,到2004年几乎消失,而缓冲区中的水(即 II 类和 III 类水)却保持相对较平缓的变化趋势,与此同时,IV 类水及劣 V 类水百分比会急剧升高,V 类水相对比较平缓。即优质水消失,污染严重的水剧增,这就是我们说的“两极性”。由此,我们得到如下结论:

- 1) 长江10年内主要特点:II 类,III 类水具有缓冲性。
- 2) 长江后10年主要特点:I 类,劣 V 类水具有两极性。

然后我们比较了 Logistic 模型、线性回归模型在预测长江水质发展趋势方面的优劣并对两个模型分别进行了综合检验。最后得出对长江未来10年水质发展趋势估计(2005~2014)不可饮用水占总流量的百分比(%),通过对回归模型进行控制,保证了一年中长江劣 V 类水消失的概率为95%,并且使得长江干流的 IV 类和 V 类水的比例在20%以内。限于篇幅,不再详细说明。

7 结束语

水资源是人类赖以生存的重要自然资源,从检测数据评价结果及预测结果来看,长江水资源的保护问题已迫在眉睫,刻不容缓。

参考文献:

- [1] 何晓群. 应用回归分析[M]. 北京: 中国人民大学出版社, 2001
- [2] 刘兰芬, 张祥伟, 夏军. 河流水环境容量预测方法研究[J]. 水利学报, 1998,7:16-20
- [3] 姜启源. 数学模型(第三版)[M]. 北京: 高等教育出版社, 2003

Prediction and Control of Water Quality of the Changjiang River Based on the Regressive Method

LU Liang, LIU Rui, ZHANG Qi

Advisor: LIU Bao-dong

(School of Mathematics and System Sciences, Shandong University, JINAN 250100)

Abstract: To predict and assess water quality of recent two years of the Changjiang River, frequency and standardized methods are used in this paper. Further, the time-space characteristic of the known main contaminants GOD_{Mn} and NH_3-N is analyzed, one dimension water quality model is adopted to estimate the distribution of main pollutant sources. At last, the two classification characteristic of water quality based on the data of recent ten years is discussed, and methods to forecast and control the future water quality are given and compared.

Keywords: pollutant sources; contaminant; water quality; statistics; prediction