

长江水质的评价预测模型及其实用性研究

于亮亮, 李 倩, 李良书

指导教师: 数模组

(西安交通大学, 西安 710049)

编者按: 本文思路清晰, 论述疏密有致, 许多细微之处稍显匠心, 构造了模糊评价指数, 可以很好的整合不同水质的影响因素; 在未来10年的预测中, 兼顾了长江流量与污水总量两者的共同影响(文中是对长江流量在不同置信水平的下限预测分析的)。

摘 要: 本文在假设所给数据完全真实的前提下, 对长江流域的水质特征进行了分析。首先, 采用模糊综合分析指数(FCI)对由多种模糊因素影响的水质进行综合评价。将所得的数据制成“FCI 信息表格”。在污染源定位问题上, 我们引入等效污染源的概念, 利用简化的一维水质扩散模型对污染物的扩散进行了简单的计算机仿真。在长江水质的预测问题中, 将已有水分为 A(饮用水)、B(IV、V)和 C(劣 V)三类。鉴于约束条件的存在, 我们定义了相对比的概念, 在进行突变数值的剔除之后, 我们采用灰色 GM(1, 1)带有修正因子的预测模型对两个相对比进行预测。在预测排污量的控制时, 我们通过灰色 GM(1, 1)模型预测出了排污量的变化, 但是由于10年的长江总流量是一个随机数组, 基本没有规律, 我们假设长江每年的水流量服从正态分布, 给出了置信水平为0.95和0.9的长江总流量的单侧置信下限, 然后给出了置信水平分别为0.95和0.9的每年的处理污水状况, 并对两种置信水平分别做了符合实际的分析。

关键词: 模糊综合评价; 置信水平

分类号: AMS(2000) 62P12

中图分类号: TB114

文献标识码: A

1 基本假设

综合评价长江及各地的水质时, 假设水质只受题目中所给的4个主要指标的影响。

假设长江内各排污点属于定常排放, 污染物在水中的扩散是稳定的。

设某点的污染物浓度仅与该点及其上游的排放量有关, 而其下游污染物对其的扩散影响忽略不计。

假设附件中所给的数据都是真实的。

2 符号规定

c : 长江水中污染物的浓度, 单位: mg/L // l : 观察点距排污处的距离, 单位: m // K : 污染物的降解系数, 单位: $1/\text{s}$ // u : 水的流速, 单位: m/s // Q : 水的流量, 单位: m^3/s // A : 长江各处的截面积, 单位: m^2 // m : 某种等级的水占长江水量的百分比// M : 两种等级的水量相对比//

3 模型的建立与求解

问题1 利用模糊综合指数对长江水质进行评价

1.1 模型建立

1.1.1 模型的主体

$U = \{U_1, U_2, \dots, U_n\}$ 表示被考察样本的 n 个指标的实测值;

$v = \{V_1, V_2, \dots, V_n\}$ 为评价集, 集合中 V_j 是与 U_i 中因子相对应的评价标准集合。

$$R = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1m} \\ r_{21} & r_{22} & \cdots & r_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ r_{n1} & r_{n2} & \cdots & r_{nm} \end{pmatrix}$$

为 U 和 V 的模糊关系矩阵。 r_{ij} 表示第 i 种污染物的含量可以被评价为第 j 类水质的可能性。

$W = (w_1, w_2, \dots, w_n)$ 为权重矩阵, 满足关系 $\sum_{i=1}^n w_i = 1$ 。 w_i 越大, 说明样本 U_i 的贡献越大。在本题里, 说明 U_i 是主要污染物。

根据上面给出的各种矩阵, 我们可以得到一个多指标模糊综合评价矩阵

$$B = WR^T,$$

构造水质标准类别向量 $S^T = (1, 2, 3, 4, 5)$ 。模糊综合指数可以表示为

$$FCL = BS,$$

FCI 作为模糊综合指数, 可以对水质进行综合评价。

1.1.2 模型中的参数

对于模型中的权重矩阵 W , 我们有如下求解方法: 对于数值越小指标越好的因素, 如高锰酸盐、氨氮等

$$w_i = \frac{c_i / \bar{s}_i}{\sum_{i=1}^n c_i / \bar{s}}$$

对于数值越大指标越好的因素, 如 DO

$$w_i = \frac{\bar{s}_i / c_i}{\sum_{i=1}^{i-1} c_i / \bar{s} + \sum_{i=1+i}^n c_i / \bar{s} + \bar{s}_i / c_i}.$$

模型中模糊关系矩阵的求法由于太过繁琐, 我们这里不再写出, 详细求解方法见附录1。

1.2 模型求解

在附件三所给的数据中我们把每个地区每一个月的水质情况作为一个样本, (共约 28×17 个样本) 对其进行综合模糊评价。

我们在求解模型前将劣 V 类的水从样本集中去掉, 只分析剩下的样本, 将所有的结果汇总到一张表上, 我们称之为“FCI 信息表格”, 这张表上包含了全部的从2003年6月到2005年9月长江沿岸17个观测站每月的水质情况, 表格的横向是时间方向, 表格的纵向代表空间(一维)方向。

1.3 基于 FCI 表格的数值分析

由于表格太大, 我们这里不再列出(完整表格请见光盘)。表中, 每一月份纵列的左面是模糊综和评价指数 FCI , 右侧是主要污染源(1代表 DO, 2代表 CODMn, 3代表 NH₃-N)。

1.3.1 长江各地水质的评价

支流上的 FCI 值随着不同的支流各自有不同的情况, 然而总体上比干流上的值要大, 水质不如干流。特别是四川乐山岷江大桥和江西九江蛤蟆石, 水质很差。总的来说, 干流沿岸地区的水质较好。

问题2 污染源的位置判定略

问题3 长江未来水质的预测

3.1 研究对象的确定

我们对附件中给出的长江水按照其对人类生活影响的不同分为三类。A类为可饮用水，包括 I、II 和 III 类的水质；B类为污染水质，包括 IV 和 V 类；C类为严重污染水质，由劣V类水质组成。

我们引入一个量，叫作相对比。当我们选定一类水作为标准量，在同一时间内，其他类别的水占的百分比与标准量的百分比的比值，定义为相对比。即

$$W_{ij} = \frac{w_i}{w_j}.$$

3.2 原始数据的处理

在所给的数据中，会有个别数据发生突变。如果我们将突变的数据带入进行预测，必定会对总趋势产生影响，大大减小预测的准确性。因此，我们应将这些数据剔除，并插上一个符合实际的值。我们将1998年的数据剔除，用1997年和1999年的平均值来代替。

处理后的数据基本上符合优质水不断减少，污染水不断增多的趋势，没有太大的扰动。因此，我们用这些数据计算相对比。我们知道，以较大的数做分母，可以从一定的程度上减少水的波动，因此我们主要预测以下两个相对比： W_{BA} 和 W_{CA} 。

3.3 灰色 GM(1,1) 模型对相对比的预测

以 W_{BA} 为例，进行灰色 GM(1, 1) 模型的建立及求解，并且引入修正因子来调节误差：通过灰色 GM(1, 1) 模型我们可以得到

$$W_{BA}^{(1)}(t) = 4.0162 \times \exp(0.0432 \times (t-1)) - 3.9421,$$

$\widehat{W_{BA}^{(0)}}(t) = \widehat{W_{BA}^{(1)}}(t) - \widehat{W_{BA}^{(1)}}(t-1)$ 做累减变换，变换为原序列得预测值 $\widehat{W_{BA}^{(0)}}(t) (t \leq n)$,

$$\widehat{W_{BA}^{(0)}}(t) = [0.3179 \ 0.3319 \ 0.3466 \ 0.3619 \ 0.3778 \ 0.3945 \ 0.4120 \ 0.4302 \ 0.4492 \ 0.4690]$$

通过修正因子来调节误差项：应用上面的模型的预测，得到下一年的预测值为 $\widehat{W_{BA}^{(0)}}(t+1)$ ，当回验值 $\widehat{W_{BA}^{(0)}}(t)$ 比实际值 $W_{BA}^{(0)}(t)$ 高时，往往预测值 $\widehat{W_{BA}^{(0)}}(t+1)$ 偏高，这时 $\frac{\widehat{W_{BA}^{(0)}}(t)}{W_{BA}^{(0)}(t)} < 1$ ；当回验值 $\widehat{W_{BA}^{(0)}}(t)$ 比实际值 $W_{BA}^{(0)}(t)$ 低时，往往预测值 $\widehat{W_{BA}^{(0)}}(t+1)$ 偏低，这时 $\frac{\widehat{W_{BA}^{(0)}}(t)}{W_{BA}^{(0)}(t)} > 1$ 。

为了精确起见，可将预测值 $\widehat{W_{BA}^{(0)}}(t+1)$ 再乘以修正因子 $\delta = \frac{W_{BA}^{(0)}(t)}{\widehat{W_{BA}^{(0)}}(t)}$ ，将它修正为

$$\widehat{y}^{(0)}(t+1) = \delta \times \widehat{W_{BA}^{(0)}}(t+1) = \widehat{W_{BA}^{(0)}}(t+1) \left[\frac{\widehat{W_{BA}^{(0)}}(t)}{W_{BA}^{(0)}(t)} \right].$$

通过 $\widehat{y}^{(0)}(t+1)$ 便可以比较精确的对模型进行预测，由以上的模型，我们可以预测出未来10年各类水的百分含量，其随时间走势如图1所示。

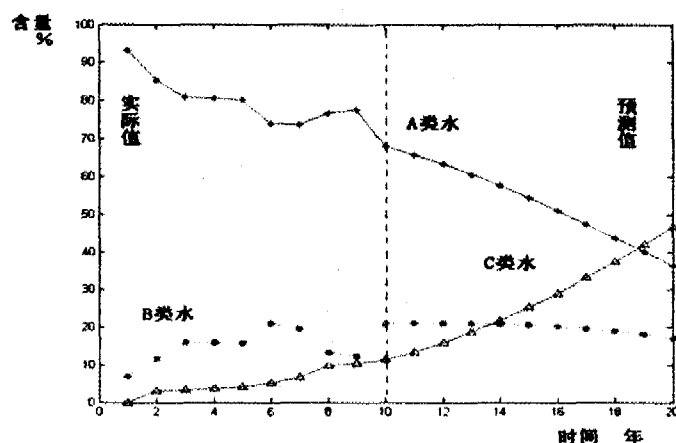


图1: 各类水百分含量预测图

发展系数:

$$-a_{BA} = 0.0432$$

$$-a_{CA} = 0.2046$$

据[6], 当 $-a \leq 0.3$ 时, GM(1,1) 可适用于中长期预测, 所以我们对于10年后长江的预测应该还是比较准确的。

3.4 模型检验

进行模型检验时, 我们选取已有的7个进行预测, 然后用3个作为检验, 通过计算我们可以得到预测值与实际值的关系如下表所示

原始数据	预测数据	残差	标准差
0.130378	0.113	0.017378	0.133289
0.132903	0.1384	-0.0055	0.041361
0.166176	0.1694	-0.00322	0.019401

通过上面的分析我们可以看出该模型对于第一年的预测误差比较大, 对于第二、第三年的预测误差则比较小, 即对于中长期的预测比较准确, 这也符合前面对于发展系数的计算。

问题4 带有置信水平的污水处理量的模型

4.1 研究对象的确定和数据的处理

首先, 我们根据已有的10年的数据, 确定与 A、B、C 类水质的比例有关的变量, 因为 $w_A + w_B + w_C = 1$, 所以 w_A, w_B, w_C 是线性相关的, 我们只要考虑与 w_B, w_C 有关的变量即可。

根据题中附件4给出的数据, 我们考察各水文年干流的 w_B, w_C , 发现与 w_B, w_C 有关的变量为每年的长江的总流量、污水直排量。

由统计结果我们可以看出, w_B, w_C 与废水直排量/长江的总流量的变化趋势几乎十分吻合, 但是只有2003年数据出现了异常, 在以后的计算中, 我们将2003年的数据剔除, 并通过插值法给2003年的 w_B, w_C 赋值分别为: 22.4%、8.8%。

4.2 最大排污量的确定:

要使10年后每年的 B 类水控制在20%以内, C 类水为0, 我们必须先通过前10年的数据确定出允许的最大直排量。

通过前10年中废水直排量/长江的总流量与 w_B , w_C 插值我们可以知道当 B 类水到达20%之前, C 类仍然处在0的状态, 所以我们只要求出 B 类水20%时对应的 μ , 然后通过长江的总流量, 就可以确定出最大排污量。

在区间内预测, 通过插值的方式得到 $\mu = 0.0273$ 。

4.3 每年处理污水量的确定

4.3.1 2005~2014年排污总量的确定

利用第三问的灰色模型进行求解, 并且进行了发展系数 $-\alpha$ 的检验, 得到 $-\alpha = 0.0624 < 0.3$, 知道灰色模型对于排污总量的中长期求解仍然是十分准确的。

4.3.2 2005~2014年长江总流量的确定

通过数据, 我们可以看到长江总流量波动很不规律, 而且考虑到实际情况, 长江虽然受到污染, 但10年后长江的总流量不会受到太大的影响, 每年的总流量都是一个随机的数据。

我们假设长江每年的总流量符合正态分布 (由于数据较少, 无法实现正态分布的 χ^2 检验), 则我们可以根据概率论的知识得到置信水平为 α 的单侧置信下限为

$$\mu = \bar{x} - \frac{s}{\sqrt{n}} t_{\alpha}(n-1),$$

其中 \bar{x} 为期望的无偏估计量

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i,$$

s 为方差的无偏估计量

$$s = \sqrt{\frac{1}{n-1} \left(\sum_{i=1}^n x_i^2 - n\bar{x}^2 \right)}.$$

带入1995~2004年的数据进行求解,

当 $\alpha = 0.05$ 时, 我们可以得到置信水平为0.95的单侧置信下限为: 9103.2

$\alpha = 0.1$ 时, 可以得到置信水平为0.9的单侧置信下限为: 9295.1

4.3.3 带有置信水平的每年污水处理量的计算

当置信水平为0.95时, 每年的允许最大直排量为: 248.12, 所以我们可以得到2005~2014年的置信水平为0.95的污水处理量 P 为:

置信水平: 0.95										
年份	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
P	57.02	78.59	101.69	126.41	152.89	181.24	211.59	244.09	278.88	316.13

当置信水平为0.9时, 每年的允许最大直排量为: 253.35

同理我们可以得到2005~2014年置信水平为0.9的污水处理量 P 为:

置信水平: 0.95										
年份	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
P	49.99	69.53	90.32	112.45	136	161.07	187.75	216.15	246.38	278.55

按照题目的要求的, 我们应该选取置信水平为0.95的污水处理方式, 但是考虑到实际情况, 置信水平越高, 对于企业治理污水花费的成本也就越高, 所以在兼顾环保跟企业积极性的情况下, 我们最好选取置信区间为0.9的污水处理方式。

对于长江,可以在前期采取0.95的污水处理方式,使污染控制在一定的水平并使水质能够得到缓解,到后期采取0.9的污水处理方式,提高企业的积极性。

问题5: 关于未来改善长江水质的建议 略

参考文献:

- [1] 沈时兴, 王国明, 张辉, 顾丽华. 模糊综合指数法评价巢湖原水水质及其应用研究[EB/OL]. <http://www.xa.cnki.net>, 2005-9-17
- [2] 刘培桐. 环境学概论[M]. 北京: 高等教育出版社, 1995
- [3] 刘圣勇. 一维水质模型对河流污染物扩散的简单模拟[EB/OL]. <http://www.xa.cnki.net>, 2005-9-17
- [4] 何秉宇, 朱建新. 中小河流水质模型应用研究[EB/OL]. <http://www.xa.cnki.net>, 2005-9-17
- [5] 盛骤, 谢世千, 潘承毅. 概率论与数理统计[M]. 北京: 高等教育出版社, 2002
- [6] 刘思峰, 党耀国, 方志耕等. 灰色系统理论及其应用[M]. 北京: 科学出版社, 2004

Evaluation and Forecasting Model of Yangtse River Water Quality and Practical Research

YU Liang-liang, LI Qian, LI Liang-shu

(Xi'an Jiaotong University, Xi'an 710049)

Abstract: Supposing the data provided is absolutely authentic, the thesis analyzes the water quality of the Yangtse River. Firstly, it makes a comprehensive evaluation of the water quality affected by manifold factor by using the FCI. FCI information table is formed with these data. We introduce the concept of equivalent contamination when pitching the contamination. Simple computer simulation on the contamination diffusion is done based upon the simplified mode of one-dimensional water diffusion. On the matter of Yangtse River quality prediction, the water is divided into three categories, that is, A (drinking water), B (IV,V) and C (inferior V). As the restriction exists, we has defined the concept of comparative ratio. After eliminating the mutation variable, we predict the comparative ratio according to the GM(1, 1) predicting-mode with revised factors. When forecasting the control on the contamination quantity, we get it by the GM(1, 1) mode. However, total flux of yangtse River in the 10 years is stochastic array almost without rule. Thus we assume it applies to the normal school with low confidence level of 0.90-0.95. The confidence level of annual sewage disposal is 0.95 and 0.90. Practical analysis is done to testify the result.

Keywords: blur comprehensive evaluation; confidence level