

文章编号:1005-3085(2005)07-0035-06

# 水质的评价和预测模型

张 震, 张 超, 张 昊

指导教师: 指导组

(解放军信息工程大学信息工程学院, 郑州 450002)

**编者按:** 本文构造了“S”型的变权函数, 对属于不同水质类别的同种污染指标进行了动态加权; 根据7个观测站的位置将干流分为8段, 计算中间6段的排污量, 将本段内所有污染源等效为一个段中央的连续稳定源, 计算出其对该段段末观测站浓度的影响值。以上两点具有独到想法。全文思路正确, 表述清晰, 假设可靠。

**摘 要:** 本文首先考虑到水质类别的差异和相同类别水质在数量上的差异对综合评价的影响, 构造“S”形的变权函数, 对属于不同水质类别的同种污染指标进行“动态加权”, 建立基于逼近理想点排序法的评价模型和利用灰色关联度的分析方法, 对长江水质状况做出了综合评价; 其次, 根据7个观测站的位置将干流分成8段, 把每段河道内所有污染源都等效为一个段中央的连续稳定源, 分别利用稳态条件下的一维水质模型及质量守恒定律, 得出中间6段每个月的排污量, 综合比较各河段一年多来的总排污量得到主要污染源的分布区域; 然后, 用每年不可饮用类水的百分比之和刻画水质状况, 综合利用灰色 GM(1,1) 模型和时间序列分析方法, 对变化趋势进行了预测; 最后, 建立不可饮用类水的百分比与长江水总流量和废水排放量的线性回归模型, 计算在满足约束条件下排污量的极限值, 用排污量的预测值减去极限值, 得到未来10年的污水处理量。

**关键词:** 逼近理想点排序法; 一维水质模型; GM(1,1) 模型; 时间序列分析; 线性回归

**分类号:** AMS(2000) 76Z10

**中图分类号:** O212

**文献标识码:** A

## 1 问题分析

问题一要求对长江两年多的水质情况做出定量的综合评价, 应先对一个月内各观测站的水质情况做出评价, 然后再综合考虑每个月的情况给出评价。附表所给出的水质评价仅仅体现了污染最严重的一项指标所达到的水质类别, 即突出主因素, 为进一步细化评判, 还要综合兼顾其他各因素的作用, 并且对于此种多属性问题, 可以借助“空间距离”概念的角度来解决, 因此采用“逼近理想解的排序法”(TOPSIS 法)。在综合评价中, 各指标水质类别区间的差异可以通过数据标准化消除, 在确定各指标权重时, 为体现出各项指标水质类别的差异以及相同类别水质在数量上的差异, 想到构造一个“S”形曲线作为变权函数, 对属于不同水质类别的同种污染指标进行动态赋权。

问题二要求污染物主要的分布区域, 可用七个观测站把干流分成八段, 将每段作为一个区域, 考虑中间的六个区域。由于段尾观测站的观测值受上游以及本段污染源的影响, 所以, 考虑用段首观测站的观测值代替上游的污染源, 将本段内所有污染源等效为一个段中央的连续稳定源, 利用一维水质模型求出两者对段尾观测站的影响值, 将两个影响值相加就等于段尾观测站的观测值, 从而求解每段排污源的排污量。

问题三要求对长江未来水质污染的发展趋势做出预测。首先, 要确定预测水质污染的物理量。因为IV、V、劣V类的水质属于不可饮用类水, 所以用这三类水质每个时段的每个评价范围所对应百分比之和刻画水质的污染。分析可知, 此数据具有确定的增长趋势和平稳的周期波动特性。则用以下方法预测: ①利用灰色 GM(1,1) 模型对此序列的确定性增长趋势进行预测, ②利用时间序列分析法对此序列的平稳周期变化趋势进行预测。两种预测值之和在一定的

程度上体现出水质污染的发展趋势。

问题四是要在满足一定约束条件下, 求出未来10年内每年需要处理的废水量。首先, 要对未来10年的废水排放量做出预测; 然后, 在满足约束条件时, 求出所允许的废水排放量的极限值, 这两个量之差就是每年要处理的废水量。每年长江水的总流量及废水排放总量会对不可饮用类水的百分比含量产生影响, 可以采用“线性回归”的方法来刻画此关系, 由此能够得到所允许废水排放量的极限值。

## 2 模型假设与符号说明

### 2.1 模型假设

- 1) 假设问题中所给出的数据能客观地反映现实情况, 值得相信;
- 2) 假设河道的长度远大于其宽度与深度;
- 3) 假设相邻观测点间河道中的污染源可等效成稳定连续点源, 且位于该段河道的中央;
- 4) 假设在短时期内, 河道中各观测点间的水流速度保持稳定。

### 2.2 符号说明 (略)

## 3 模型的建立及求解

### 3.1 问题一

#### 3.1.1 某个月内河流沿线的17个观测点水质情况的量化评价

##### 1) 构造变权函数

四项污染物对水质的污染程度分为 I、II、III、IV、V、劣 V 类 6 个等级, 不妨设对应的数值分别为 1, 2, 3, 4, 5, 6 (见图1)。分析可知, 当水质的类别从 I 变化到 III 时, 其权值变化应该较缓慢; 从 III 类水变到 IV 水时, 其权值变化应该非常大, 这体现出水将发生质的变化 (从可饮用到不可饮用); 而且水质越差, 相应的权值也要越大, 这样才能突出首要污染物。依据以上情况, 构造增长的“S”形曲线作为变权函数:

$$f(x) = \begin{cases} \alpha \sqrt[3]{x - \beta} + \gamma, & x \geq 0 \\ 0, & x < 0, \end{cases}$$

其中  $\alpha, \beta, \gamma$  为待定的常数。

当  $x = 1$ , 即水质最好时, 令相应的量化值  $f(1) = 0.05$ ; 为了说明 I、II、III 的之间的相对变化较小, 令  $f(3) = 0.25$ ; 当  $x = 6$  时, 其值为 1, 此时污染最严重。对应以上三个点, 求得  $\alpha = 0.35, \beta = 3.48, \gamma = 0.52$ 。于是得到  $f(x)$  的具体表达式为:

$$f(x) = \begin{cases} 0.35 \sqrt[3]{x - 3.48} + 0.52, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

最后, 经计算可得水质类型 I 至劣 V 所对应的量化值分别为 (0.05, 0.12, 0.25, 0.8, 0.92, 1)。

##### 2) 求权重矩阵

以其中的某一个月为例, 将 17 个观测点的 4 项监测指标的浓度提取出来, 构造出原始评判矩阵  $R = (r_{ij})_{17 \times 4}$ ; 再将元素  $r_{ij}$  的值变换成污染类别, 把类别对应的数值代入变权函数, 得到量化矩阵; 对量化矩阵中每一行的指标做归一化, 得到权重矩阵  $W$ 。对于两年来的每个月, 都能得到一个权重矩阵。

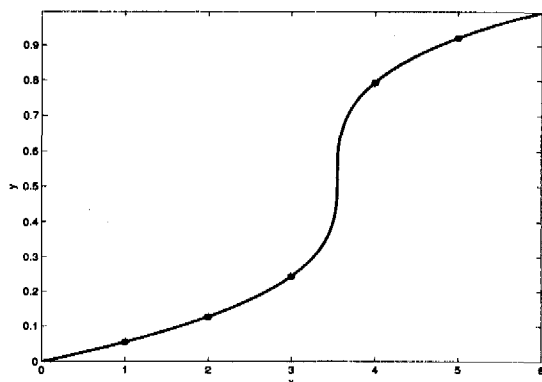


图 1: 变权函数示意图

### 3) 基于 TOPSIS 法的综合评价模型

step1: 评价指标的极性和极值处理, 得到规范化矩阵  $X$

溶解氧的指标是极大值, PH 值的指标是居中型的, 其它两种是极小型的。对评价指标进行极小型处理, 得到评价指标的极型一致化矩阵  $R^* = (r_{i1}^*, r_{i2}^*, r_{i3}^*, r_{i4}^*)$ , 然后进行极值处理, 得到规范化矩阵  $X = (x_{ij})_{17 \times 4}$ , 其中

$$x_{ij} = \frac{r_{ij}^* - \min_{1 \leq i \leq 17} r_{ij}^*}{\max_{1 \leq i \leq 17} r_{ij}^* - \min_{1 \leq i \leq 17} r_{ij}^*} \quad (i = 1, 2, 3, \dots, 17; \quad j = 1, 2, 3, 4).$$

step2: 构造加权规范决策矩阵  $Z$

依据得到权重矩阵  $W$ , 令:  $z_{ij} = x_{ij} \times w_{ij} (i = 1, 2, 3, \dots, 17; j = 1, 2, 3, 4)$ , 由此就在决策矩阵  $Z$  中体现了水质类别的差异和相同类别水质数量上的差异。

step3: 确定理想解及负理想解, 并计算综合评价值

设理想解为  $z^+$ , 负理想解为  $z^-$ , 则综合评价值为:  $b_i = \frac{d_i^+}{d_i^+ + d_i^-}$

其中

$$d_i^+ = \sqrt{\sum_{j=1}^4 (z_{ij} - z_i^+)^2}, \quad d_i^- = \sqrt{\sum_{j=1}^4 (z_{ij} - z_i^-)^2}.$$

此时, 各类污染指标越大, 被评价对象的状态越接近负理想解, 得到的综合评价值越大, 河流污染程度就越严重; 反之则接近理想解, 综合评价值小, 污染程度就越轻。

step4: 评价模型的求解

依据附件三中的数据, 利用 MATLAB 编程求出两年多来每个月中 17 个观测点水质的评价值。以 2003 年 6 月为例, 17 个观测站的评价值分别为  $\{0.186, 0.235, 0.448, 0.25, 0.252, 0.244, 0.266, 0.57, 0.197, 0.608, 0.117, 0.471, 0.2, 0.208, 0.475, 0.245, 0.196\}$ 。

#### 3.1.2 基于灰色关联的方法综合分析 17 个观测点水质的污染情况

1) 选定母序列和比较序列

定义评价值矩阵  $H$ , 其中的元素  $h_{ij}$  表示第  $j (j = 1, 2, 3, \dots, 17)$  个观测点在第  $i (i = 1, 2, 3, \dots, 28)$  个月的评价值 (由 3.1.1 求得)。令  $h_{i0} = \min_{1 \leq j \leq 17} h_{ij}$ , 则母序列  $h_{i0} (i = 1, 2, 3, \dots, 28)$  是构造出来的最理想评判值序列。由  $h_{i1}, h_{i2}, h_{i3}, \dots, h_{i17}$  分别构成 17 个子序列 (比较序列)。

## 2) 计算关联度

子序列中每个数值与母序列中相应数值的关联系数为:  $v_j(i) = \frac{a+\rho \times b}{\Delta_j(i)+\rho \times b}$ , 其中

$$\Delta_j(i) = |h_{ij} - h_{i0}|, \quad a = \min_{1 \leq j \leq 17} \min_{1 \leq i \leq 28} \Delta_j(i), \quad b = \max_{1 \leq j \leq 17} \max_{1 \leq i \leq 28} \Delta_j(i), \rho = 0.5.$$

把关联系数取均值得到关联度:

$$u_j = \frac{1}{28} \sum_{i=1}^{28} v_j(i) \quad (j = 1, 2, \dots, 17),$$

关联度表示观测站评判值序列和理想评判值序列相关联的程度, 数值越大, 观测站的水质情况越好。

## 3.1.3 模型的求解

利用 MATLAB 编程得到17个观测点水质的综合评价及排名, 如表1所示:

表 1: 各观测点的综合评价及排名

观测点	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
综合评判值 ( $\times 10^{-3}$ )	797	850	780	668	790	755	741	533	773	667	925	663	632	685	423	682	725
排名	3	2	5	12	4	7	8	16	6	13	1	14	15	10	17	11	9

由综合排名得到: 湖北丹江口胡家岭的水质情况最好, 而江西南昌滁槎的水质情况最差。

## 3.2 问题二

## 3.2.1 一维水质模型

依据文献[2], 我们得到流体中的一维水质模型:  $\frac{\partial C}{\partial t} = D_x \frac{\partial^2 C}{\partial x^2} - u_x \frac{\partial C}{\partial x} - KC$ , 其中  $x$  为河流的长度,  $C$  为流体中污染物的浓度分布,  $u_x$  为水流速度,  $D_x$  为弥散系数,  $K$  为流体的降解系数。稳态条件下的一般河流, 扩散作用可以忽略, 上面的方程变为  $u_x \frac{\partial C}{\partial x} = KC$ , 若给定初始条件为  $C(0) = C_0$  时, 得到  $C = C_0 \exp\{-\frac{Kx}{u_x}\}$ 。

3.2.2 第  $n$  段河道中污染源排放量模型的建立

## 1) 确定段内的水流速度

每一段内的水速取前后两个观测站水速观测值的均值, 即  $\frac{U_n + U_{n+1}}{2}$ , 其中  $U_n$  和  $U_{n+1}$  分别为第  $n$  段河道中段首观测站和段尾观测站的水流速度。

2) 段首观测站上游所有污染源对段尾观测站的影响值  $C_{n+1}^{(1)}$ 

用段首观测站的观测值代替上游污染源的影响, 根据一维水质模型, 得到

$$C_{n+1}^{(1)} = C_n \exp\left\{-\frac{2KL_n}{U_{n+1} + U_n}\right\} \quad (n = 1, 2, 3, \dots, 6) \quad (1)$$

3) 本段污染源对段尾观测站的影响值  $C_{n+1}^{(2)}$ 

首先, 段中央等效污染点的浓度值等于单位时间排污量与水流量的比值, 即  $2\frac{B_n}{V_n + V_{n+1}}$ 。其次, 根据一维水质模型, 得到本段污染源对段尾观测站的影响值为

$$C_{n+1}^{(2)} = \frac{2B_n}{V_{n+1} + V_n} \exp\left\{-\frac{KL_n}{U_{n+1} + U_n}\right\} \quad (n = 1, 2, 3, \dots, 6) \quad (2)$$

其中  $B_n$  为第  $n$  段河道中污染源每月的排污量,  $V_n$  为第  $n$  段河道中段首观测站的水流量,  $L_n$  为第  $n$  段河流的长度。

#### 4) 列方程求解污染源排放量 $B_n$

根据质量守恒定律, 观测点单位时间内流入的污染物的量等于单位时间内流出的量, 于是  $C_{n+1}V_{n+1} = C_{n+1}^{(1)}V_{n+1} + C_{n+1}^{(2)}V_{n+1}$ , 即

$$C_{n+1} = C_{n+1}^{(1)} + C_{n+1}^{(2)} \quad (3)$$

将(1)、(2)、(3)式联立得到

$$B_n = \frac{V_{n+1} + V_n}{2} \left( C_{n+1} - C_n \exp \left\{ -\frac{2KL_n}{U_{n+1} + U_n} \right\} \right) \exp \left\{ \frac{KL_n}{U_{n+1} + U_n} \right\} \quad (4)$$

### 3.2.3 问题二的求解

将附件3中所给出的数据代入(4)式, 求出每个月中各段河道污染源的排污量。最后, 得到 6 段河道在 13 个月中高锰酸盐的排污总量为 {113203.8, 1874725, 9935280, 9763920, 14898750, 12192310} (吨); 氨氮排污总量为 {22425.3, 161422.1, 1266054, 714477, 345246.6, 1014997} (吨)。

进行比较后, 得到河道中高锰酸盐污染源主要分布在第 5 河段 (江西九江—安徽安庆), 氨氮的污染源主要分布在第 3 河段 (湖北宜昌—湖南岳阳)。

### 3.3 问题三

根据分析, 可以将河道中 IV 类、V 类和劣 V 类水的总含量来刻画长江水质情况, 对附件 4 中数据统计后, 得到过去 10 年中各种情况下不可饮用类水的含量百分比序列。将序列  $X_0$  分解为  $Y_0$  和  $Z_0$ , 其中  $Y_0$  反映  $X_0$  的确定性增长趋势,  $Z_0$  反映  $X_0$  的平稳周期变化趋势。

#### 3.3.1 利用灰色 GM(1,1) 模型对 $X_0$ 序列的确定性增长趋势进行预测

利用 MATLAB 编程进行求解, 得到  $X_0$  序列的确定性增长趋势  $Y_0$ 。

#### 3.3.2 利用时间序列分析法对序列 $X_0$ 的平稳周期变化趋势进行预测

从原始序列  $X_0$  中, 消除确定性增长的趋势  $Y_0$ , 就能得到平稳周期变化趋势, 即

$$\{Z_0(1), \dots, Z_0(10)\} = \{X_0(1), \dots, X_0(10)\} - \{Y_0(1), \dots, Y_0(10)\}$$

对于此平稳变化序列, 可利用时间序列分析法中的 ARMA 模型进行预测。根据  $\rho_k$  和  $\varphi_{kk}$  的拖尾性和截尾性, 利用 MATLAB 编程进行预报, 得到平稳变化趋势的预测值  $Z_0$ 。最后, 综合确定性增长趋势  $Y_0$  和平稳随机变化趋势  $Z_0$  的预测值, 得到  $X_0$  序列的预测值。这样, 就可以得到未来 10 年中各种情况下废水含量百分比的预测值。

以全流域在枯水期中不可饮用类水含量百分比为例

$$\{X_0(1), \dots, X_0(10)\} = \{6.9, 17.2, 32.7, 14.3, 23, 26.8, 28.7, 32.4, 27.5, 32.2\}$$

得到未来 10 年废水含量百分比的预测值为

$$\{X_0(11), \dots, X_0(20)\} = \{34.00, 35.93, 37.96, 40.11, 42.37, 44.77, 47.30, 49.98, 52.81, 55.80\}$$

#### 3.3.3 分析长江未来水质污染的发展趋势

根据预测值分析可知, 若不采取任何治理手段, 在 10 年后, 长江的非饮用水 (IV 类、V 类和劣 V 类) 含量在一年中的大多数情况下都将超过 50%, 而且其比例还在逐年增加。正如专家给出的预测 (附件 1): 若不及时拯救, 长江生态 10 年内将濒临崩溃。

### 3.4 问题四

#### 3.4.1 污水百分比与水的总流量和废水排放总量的函数关系

根据分析可知, 一年内污水百分比不仅和废水排放总量有关, 而且也与水的总流量有关系, 并且可以看成是线性关系。所以建立二元线性回归模型, 在此问题中, 以一年中所有检测

数据的平均值来刻画出整体的情况(即水文年中的数据),利用 MATLAB 编程,可分别求得一年内干流中IV、V类水百分比之和  $p_1$  与总流量  $v$  和排放量  $m$  的函数关系。以及IV、V、劣V类水百分比之和  $p_2$  与  $v$  和  $m$  的函数关系。实际情况中,总流量的变化很小,所以取前10年总流量的平均值来代替,即  $v = 9894.1$  亿立方米,得到简化的模型

$$p_1 = 0.15497m - 20.1926$$

$$p_2 = 0.2059m - 28.9729$$

### 3.4.2 满足函数关系约束下废水排放量的极限值

一年内干流中IV、V类水百分比之和要小于20%,即  $p_1 = 0.15497m - 20.1926 \leq 0.2$ ; 一年内干流中劣V类水百分比要不大于0,那么IV、V、劣V类水百分比之和应该小于等于IV、V类水百分比之和,即  $p_2 = 0.2059m - 28.9729 \leq 0.2$ ; 取两种约束范围的交集,得到废水排放量的极限为  $m^{max} = 131.59$  亿吨。

### 3.4.3 未来10年内废水排放量的预测值

利用灰色 GM(1,1) 模型进行预测,得到未来10年污水排放量的预测值:

$$\{303.01, 322.52, 343.29, 365.39, 388.92, 413.96, 440.61, 468.98, 499.18, 531.32\} \quad (\text{亿吨})$$

### 3.4.4 未来10年内的污水处理量的预测值

用排污量的预测值减去极限值,得到未来十年的污水处理量:

$$\{171.42, 190.93, 211.7, 233.8, 257.33, 282.37, 309.02, 337.39, 367.59, 399.73\} \quad (\text{亿吨})$$

## 4 模型的评价和改进 (略)

### 参考文献:

- [1] 岳超源. 决策理论与方法[M]. 北京: 科学出版社, 2003
- [2] 郑彤等. 环境系统数学模型[M]. 北京: 化学工业出版社, 2003
- [3] 韩中庚. 数学建模方法及其应用[M]. 北京: 高等教育出版社, 2005
- [4] 张树京等. 时间序列分析简明教程[M]. 北京: 清华大学出版社, 2003

## An Evaluation and Prediction Model of Water Quality

ZHANG Zhen, ZHANG Chao, ZHANG Hao

Instructor: Instructor Group

(Institute of Information Engineering, Information Engineering University, PLA, Zhengzhou 450002)

**Abstract:** The difference of water quality and quantity have an effect on the comprehensive evaluation of water quality. At first an "S"-like variable-value function dynamically weights the pollution evaluation standards of identical kinds for different water quality categories. Then, by making an evaluation model based on TOPSIS and using the analysis method of Grey Correlation, the water quality of Changjiang River is evaluated. Secondly, the main stream is divided into eight parts according to the positions of the seven observation points. All the polluting sources in each part are equal to a continuously stable source in the center of the part. By using the one-dimensional water quality model in the stable state and Quality Conservation Law to get the monthly pollutant amount in the middle six parts, and by comprehensively comparing the pollutant amount of different parts over one year, the distribution area of the main polluting sources can be got. Thirdly, the description of the state of water quality by the yearly undrinkable water percentage is made, and by using of the Grey GM(1,1) model and time sequence analysis method, the variation tendency is estimated. Last, in light of developing a linear regression model of the undrinkable water percentage in relation to the overall flow and waste water flow in Changjiang River, the limit of pollutant amount under restrictions is resolved. By subtracting the limit from the estimation value of pollutant amount, the waste water disposal amount can be predicted in ten years.

**Keywords:** TOPSIS; grey correlation; one-dimensional water quality model; GM(1,1) model; time sequence analysis; linear regression