

湖泊水质预测模型的建立及其应用*

万幼川 李植生 梁小民 刘良明

(武汉测绘科技大学信息工程学院, 武汉市珞喻路 39 号, 430070)

摘 要 根据东湖多年监测数据分析,建立了多元相关预测模型、基于灰色理论灰色系统的预测模型和质量平衡模型。为了提高灰色预测精度,利用残差对模型进行了相应的修正,并利用 3 种模型对影响东湖水质主要因子——经济、人口、BOD、COD、TN、TP 进行了 2000 年的预测,提出了东湖污染防治对策。

关键词 预测;多元回归;灰色模型;质量平衡

分类号 TP391;X824

武汉东湖是我国和世界闻名的城市湖泊,由于大量生产和生活污水排入湖中,加之开发不当、保护不力,其富营养化在城市湖泊中具有代表性。1992~1994 年水质监测表明,东湖的污染已相当严重,全湖 COD 和 DCOD 年平均值分别达 21.5 mg/L 和 13.01 mg/L。由于湖泊本身的自净能力下降,其水质在国家地面水标准 IV~V 之间。目前全湖平均总氮为 1.45 mg/L;总磷为 0.121 mg/L;藻数量达 87.4×10^6 个/L,已属富营养湖泊类型,因此东湖的治理已迫在眉睫。本文针对东湖的污染状况,在多年监测资料(武汉市环境保护科学研究所提供)分析的基础上,提出和建立了多元相关预测模型、灰色预测模型及质量平衡模型;对东湖地区的人口、经济及主要水质污染参数 BOD、COD、TN、TP 等进行了预测,为东湖区域规划、综合治理及社会经济可持续发展提供了依据。

1 多元相关模型建立

设预测对象为 y , 其相关的因子为 x_1, x_2, \dots, x_m , 则多元相关模型可定义为:

$$y_i = \bar{b}_0 + \bar{b}_1 x_{i1} + \bar{b}_2 x_{i2} + \dots + \bar{b}_m x_{im} + \Delta_i \quad (1)$$

矩阵形式为: $Y = \begin{matrix} X \\ n \end{matrix} \begin{matrix} B \\ m+1 \end{matrix} + \Delta$, $E(\Delta) = 0$, $D(Y) = \epsilon_0^2 I$ (2)

式中 $R(X) = m+1$, 以及

$$X_{n, m+1} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1m} \\ 1 & x_{21} & \dots & x_{2m} \\ \dots & \dots & \dots & \dots \\ 1 & x_{n1} & \dots & x_{nm} \end{bmatrix} \quad B = \begin{bmatrix} \bar{b}_0 \\ \bar{b}_1 \\ \dots \\ \bar{b}_m \end{bmatrix}$$

其误差方程为: $V = XB - Y$ (3)

根据最小二乘解: $\bar{B} = (X^T X)^{-1} X^T Y$, $Q_{BB} = (X^T X)^{-1}$, $\epsilon_0^2 = V^T V / (n - m - 1) = K / (n - m - 1)$,

检验原假设 $H_0: \bar{b}_1 = \bar{b}_2 = \dots = \bar{b}_m = 0$, 则有:

收稿日期: 1996-06-03. 万幼川, 男, 36 岁, 副教授, 博士生, 现从事 RS 与 GIS 研究。

* 测绘遥感信息工程国家重点实验室开放研究基金资助项目, 编号 WKL(95)101

$$H\bar{B} = 0, H_1 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} = \begin{bmatrix} 0 & \\ m_1 & m_m \end{bmatrix} \quad (4)$$

相应将 X 和 B 分块:

$$X = \begin{bmatrix} e & X_1 \\ n_1 & nm \end{bmatrix}, \bar{B} = \begin{bmatrix} b_0 \\ B_1 \\ n_1 \end{bmatrix}$$

考虑约束条件 (4), 残差平方和为:

$$K_H = K + R \quad (5)$$

根据 R 定义得:

$$R = (H\bar{B})^T (H(X^T X)^{-1} H^T)^{-1} H\bar{B}$$

$$= ([0 \ I] \begin{bmatrix} b_0 \\ B_1 \end{bmatrix})^T ([0 \ I] \begin{bmatrix} e^T \\ X_1^T \end{bmatrix} [e \ X_1])^{-1} \begin{bmatrix} 0 \\ 1 \end{bmatrix})^{-1} ([0 \ I] \begin{bmatrix} b_0 \\ B_1 \end{bmatrix}) = \bar{B}_1^T Q_{B_1, B_1}^{-1} \bar{B}_1 \quad (6)$$

其中,

$$\begin{bmatrix} e^T e & e^T X_1 \\ X_1^T e & X_1^T X_1 \end{bmatrix}^{-1} = \begin{bmatrix} Q_{e, e_0} & Q_{e, B_1} \\ Q_{B_1, e_0} & Q_{B_1, B_1} \end{bmatrix}$$

如果 H_0 成立则统计量为:

$$F = \frac{B_1^T Q_{B_1, B_1}^{-1} B_1 / m}{K / (n - m - 1)} \quad (7)$$

在作回归分析时, 须将数据作中心化处理, 即

$$x'_{ij} = x_{ij} - \bar{x}_j, \bar{x} = 1/n \sum_{i=1}^n x_{ij}$$

参数 b_0 亦作相应变换, 取

$$b'_0 = b_0 + b_1 \bar{x}_1 + b_2 \bar{x}_2 + \dots + b_m \bar{x}_m \quad (8)$$

则相应的函数模型为:

$$y_i = b'_0 + b_1 x'_{i1} + b_2 x'_{i2} + \dots + b_m x'_{im} + \Delta_i \quad (9)$$

此时相应的法方程为:

$$\begin{bmatrix} n & 0 \\ 0 & X'^T X' \end{bmatrix} \begin{bmatrix} b'_0 \\ B_1 \end{bmatrix} = \begin{bmatrix} n\bar{y} \\ x'^T y \end{bmatrix} \quad (10)$$

式中,

$$X' = \begin{bmatrix} x'_{11} & x'_{12} & \dots & x'_{1m} \\ x'_{21} & x'_{22} & \dots & x'_{2m} \\ \dots & \dots & \dots & \dots \\ x'_{n1} & x'_{n2} & \dots & x'_{nm} \end{bmatrix}$$

解得: $b'_0 = \bar{y}, B_1 = (X'^T X')^{-1} X'^T Y, Q_{B_1, B_1} = (X'^T X')^{-1}, R = B_1^T Q_{B_1, B_1}^{-1} B_1 = B_1^T X'^T Y = (\sum_i x_{i1} y_i) b_1 + (\sum_i x_{i2} y_i) b_2 + \dots + (\sum_i x_{im} y_i) b_m$

2 灰色预测模型建立

设有变量 $x^{(0)}$ 为:

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\} \quad (11)$$

其相应的微分模型为:

$$dx^{(1)}/dt + ax^{(1)} = u \tag{12}$$

(12)式中 u 为内生变量, dx/dt 与背景量 \mathcal{H} 为线性组合, 即有:

$$a^{(1)}(x^{(1)}(k+1)) + a\mathcal{H}^{(1)}(k+1) = u \tag{13}$$

考虑

$$a^{(1)}(x^{(1)}(k+1)) = x^{(0)}(k+1) \tag{14}$$

$$\mathcal{H}^{(1)}(k+1) = (x^{(1)}(k) + x^{(1)}(k+1)) \tag{15}$$

便有:

$$k = 1, \quad x^{(0)}(2) = a(-1/2(x^{(1)}(1) + x^{(1)}(2))) + u$$

$$k = 2, x^{(0)}(3) = a(-1/2(x^{(1)}(2) + x^{(1)}(3))) + u$$

.....

$$k = n - 1, x^{(0)}(n) = a(-1/2(x^{(1)}(n-1) + x^{(1)}(n))) + u$$

引入下述符号:

$$y^N = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \dots \\ x^{(0)}(n) \end{bmatrix}, \quad \mathcal{H} = \begin{bmatrix} -1/2(x^{(1)}(1) + x^{(1)}(2)) \\ -1/2(x^{(1)}(2) + x^{(1)}(3)) \\ \dots \\ -1/2(x^{(1)}(n-1) + x^{(1)}(n)) \end{bmatrix}$$

便有:

$$y^N = a\mathcal{H} + uE = [\mathcal{H}; E] \begin{bmatrix} a \\ u \end{bmatrix} = [\mathcal{H}; E] \hat{a} \tag{16}$$

记:

$$B = [\mathcal{H}; E] = \begin{bmatrix} -1/2(x^{(1)}(1) + x^{(1)}(2)) \\ -1/2(x^{(1)}(2) + x^{(1)}(3)) \\ \dots \\ -1/2(x^{(1)}(n-1) + x^{(1)}(n)) \end{bmatrix}$$

$$y^N = B\hat{a} \tag{17}$$

则式 (16) 为:

根据最小二乘得解:

$$\hat{a} = (B^T B)^{-1} B^T y^N, x^{\hat{1}}(k+1) = (x^{(0)}(1) - u/a) e^{-ak} + u/a$$

为了提高模型精度, 利用残差对原模型进行修正. 修正模型既可以是生成模型 $x^{\hat{1}}(k+1)$, 也可以是还原模型 $x^{\hat{0}}(k+1)$, 即

$$x^{\hat{0}}(k+1) = (-a)(x^{(0)}(1) - u/a) e^{-ak} \tag{18}$$

设生成残差为 $X^{(0)}(k) = x^{(1)}(k) - x^{\hat{1}}(k)$, 则生成残差列 $X^{(0)}$ 为: $X^{(0)} = \{X^{(0)}(1'), X^{(0)}(2'), \dots, X^{(0)}(n')\}$, $X^{(0)}$ 所对应的生成数列分别为:

$$x^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\}, x^{\hat{1}} = \{x^{\hat{1}}(1), x^{\hat{1}}(2), \dots, x^{\hat{1}}(n)\}$$

一般有 $n' \leq n$, $X^{(0)}$ 的生成数列为: $X^{(1)} = \{X^{(1)}(1), X^{(1)}(2), \dots, X^{(1)}(n')\}$, 则解为:

$$\hat{X}^{(1)}(k+1) = (X^{(0)}(1) - u^X/l^X) e^{-a^X k} + u^X/l^X \tag{19}$$

$X^{(1)}(k+1)$ 的导数为:

$$\hat{X}^{(0)}(k+1) = (-a^X)(X^{(0)}(1) - u^X/l^X) e^{-a^X k} \tag{20}$$

以 $X^{(0)}(k+1)$ 修正 $x^{\hat{1}}(k+1)$ 得修正后的模型为:

$$x^{\hat{1}}(k+1) = (x^{\hat{0}}(1) - u/l a) e^{-ak} + u/l a + W(k-i)(-ax)(X^{(0)}(1) - u^X/l^X) e^{-a^X k} \tag{21}$$

其中,

$$W(k-i) = \begin{cases} 1, & k \geq i \\ 0, & k < i \end{cases}, i = n - n'$$

对还原模型修正, 则相应数列:

$$x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}, x^{\hat{0}} = \{x^{\hat{0}}(1), x^{\hat{0}}(2), \dots, x^{\hat{0}}(n)\}$$

残差 $q^{(0)}(k)$ 为:

$$q^{(0)} = \{q^{(0)}(1), q^{(0)}(2), \dots, q^{(0)}(n')\}, q^{(0)} = x^{(0)}(k) - \hat{x}^{(0)}(k)$$

通过残差 $q^{(0)}$ 建立的模型为:

$$\hat{q}^{(1)}(k+1) = (q^{(0)}(1) - u_q / a_q) e^{-a_q k} + u_q / a_q \quad (22)$$

相应 $\hat{q}^{(1)}(k+1)$ 的导数为:

$$\dot{q}^{(0)}(k+1) = (-a_q)(q^{(0)}(1) - u_q / a_q) e^{-a_q k}$$

修正后的模型为:

$$\hat{x}^{(0)}(k+1) = (-a)(x^{(0)}(1) - u/a) e^{-ak} + W(k-i)(-a_q)(q^{(0)}(1) - u_q/a_q) e^{-a_q k} \quad (23)$$

其中,

$$W(k-i) = \begin{cases} 1, & k \geq i \\ 0, & k < i \end{cases}, i = n - n'$$

3 物质平衡模型

根据物质平衡原理,进入湖泊的磷量减去从湖泊内流出和沉降的磷量,应等于湖水中磷总量的变化:

$$V \frac{dp}{dt} = I_p - qp - \lambda_p Vp \quad (24)$$

化简后得:

$$dp/dt = I_p/V - (p_w + \lambda_p)p \quad (25)$$

式中, p 为湖水年平均总磷浓度 (mg/L); I_p 为输入湖泊磷的总量 (g/d); p_w 为水力冲刷系数, $p_w = q/V$; q 为出湖河道流量 (m^3/d); V 为湖泊容量 (m^3); λ_p 为磷的沉降速率常数 (1/d); t 为河水入湖时间 (d).

为了求得在均匀混合条件下, V 为稳态时的上述方程的解,本研究采用了求解湖水总磷浓度的合田健公式,即

$$c = L/Z(Q/V + T) \quad (26)$$

式中, c 为湖水总磷的预测浓度 (mg/L); L 为湖泊单位面积年度总磷的负荷量 [$g/cm^2 \cdot a$]; Q 为年入湖水量 (m^3/a); V 为湖泊容量 (m^3); Z 为平均水深; T 为湖水总磷的沉降系数 (1/a)

合田健根据日本 25 个湖泊的调查资料,求得总磷的沉降系数与平均水深之间的关系式为:

$$T = 10/Z \quad (27)$$

4 应用

4.1 相关预测

利用相关模型,设水质浓度与其影响因素人口、工业产值、年渔产量之间存在线性相关,则函数式可表达为:

$$y = b_0 + b_1 X_1 + b_2 X_2 + b_3 X_3 \quad (28)$$

式中, y 为东湖 COD BOD TN TP 的年平均浓度 (mg/L); X_1 为人口数 (万人); X_2 为工业总产值 (亿元); X_3 为东湖年捕鱼量 (t).

根据建立的灰色预测模型对相关因子进行计算和预测,表 1 给出了人口、经济和捕鱼量在

2000年的预测值分别为 763.9万人、567.1亿元和 2123.5t

表 1 相关因素预测

年代	人口 万人	计算值	误差 %	经济 亿元	计算值	误差 %	捕鱼量 /t	计算值	误差 %
1986	608.4	608.4	0	175.4	175.4	0	1 125	1 125	0
1987	620.0	625.4	0.92	187.2	193.9	3.45	1 175	1 119.2	4.98
1988	641.5	635.4	0.95	192.7	210.6	8.48	1 150	1 175.7	2.19
1989	647.4	645.3	0.33	215.6	228.7	5.72	1 200	1 235.1	2.84
1990	653.3	655.2	0.29	238.7	248.4	3.92	1 240	1 297.5	4.43
1991	665.3	665.4	0.01	335.1	269.7	24.21	1 260	1 363.0	7.55
1992	677.0	675.6	0.20	300.2	292.9	2.45	1 450	1 431.8	1.27
1993	684.4	686.1	0.25	311.8	318.2	6.41	1 600	1 504	6.37
2000		763.9			567.1			2123.5	

根据建立起的相关预测模型分别形成了 TN TR BOD和 COD的预测方程,即

$$y_{TN} = - 0.000 798 + 0.009 012X_1 - 0.002 409X_2 - 0.001 638X_3 \quad (29)$$

$$y_{TP} = - 0.000 005 + 0.000 023X_1 + 0.000 280X_2 \quad (30)$$

$$y_{COD} = - 0.000 1 + 0.009 771X_1 + 0.006 335X_2 - 0.002 070X_3 \quad (31)$$

$$y_{BOD} = 0.001 244 - 0.022 023X_1 + 0.001 933X_3 \quad (32)$$

相关系数分别为:

$$R_{TN} = 0.701, \quad R_{TP} = 0.75, \quad R_{COD} = 0.948, \quad R_{BOD} = 0.928$$

其计算值如表 2所示

表 2 东湖水中 4种污染物相关预测浓度 (/mg° L⁻¹)

年代	TN	计算值	误差	TP	计算值	误差	COD _{mn}	计算值	误差	BOD	计算值	误差
1986	1.76	1.632	- 0.128	0.067	0.052 8	- 0.014 2	4.35	4.529	0.179	2.33	2.497	0.167
1987	1.49	1.625	0.135	0.055	0.056 4	0.001 4	4.87	4.614	- 0.256	3.13	2.839	- 0.291
1988	1.50	1.846	0.346	0.058	0.058 4	0.000 4	4.87	4.910	0.040	2.30	2.069	- 0.231
1989	1.91	1.761	- 0.149	0.052	0.064 9	0.012 9	4.97	5.009	0.039	2.20	2.537	0.337
1990	1.99	1.693	- 0.297	0.056	0.071 5	0.015 5	5.12	5.131	0.011	2.53	2.885	0.355
1991	1.42	1.535	0.115	0.086	0.098 8	0.012 8	5.80	5.817	0.017	3.67	2.861	- 0.809
1992	1.58	1.412	- 0.168	0.121	0.089 3	- 0.031 7	5.45	4.317	- 0.133	3.80	4.872	1.072
1993	1.06	1.205	0.145	0.09	0.092 7	0.002 7	5.05	5.152	0.102	7.10	6.500	- 0.600
2000	2.031			0.176			6.707			8.531		

4.2 灰色预测

利用建立的灰色预测模型,预测了不同年代的 TN TR COD BOD在湖水中的浓度,结果列表 3

4.3 质量平衡预测

由于相关预测与灰色预测的 TN TP浓度存在差异,而且 TN的差异较大,为了提高预测值精度,又采用质量平衡法进行了预测。根据近 10年东湖的资料分析可以看出,总磷总氮的排放量平均每年增加 3%。在不考虑降雨氮磷输入的情况下,以 1992为基准,求出了 2000年和 2005年的总磷、总氮的量,代入式 (26)并根据东湖数字高程模型推算出湖水的平均深度,最后

分别算出 2000年和 2005年东湖水中总磷、总氮浓度。如表 4所示,与相关预测相比,TP仅差 0.002 mg/L, TN 相差也不过 0.17 mg/L。

表 3 东湖水中 4种污染物灰色预测浓度 ($\text{mg} \cdot \text{L}^{-1}$)

年代	TN	计算值	误差	TP	计算值	误差	COD _{mn}	计算值	误差	BOD	计算值	误差
1986	1.76	1.760	0	0.067	0.067 0	0	4.35	4.35	0	2.33	2.33	0
1987	1.49	1.727	0.237	0.055	0.066 6	0.011 6	4.87	4.711	- 0.159	3.13	2.284	- 0.846
1988	1.50	1.670	0.170	0.058	0.070 4	0.012 4	4.87	4.838	- 0.032	2.30	2.552	0.252
1989	1.91	1.616	- 0.294	0.052	0.074 4	0.022 4	4.97	4.970	0	2.20	2.851	0.651
1990	1.99	1.563	- 0.427	0.056	0.078 7	0.022 7	5.12	5.104	- 0.016	2.53	3.185	0.655
1991	1.42	1.512	0.092	0.086	0.083 2	- 0.002 8	5.80	5.243	- 0.557	3.67	3.558	- 0.112
1992	1.58	1.462	- 0.118	0.121	0.087 9	- 0.033 1	5.45	5.385	- 0.065	3.80	3.975	0.175
1993	1.06	1.414	0.354	0.09	0.092 9	0.002 9	5.05	5.531	0.481	7.10	4.441	- 2.659
2000		1.210			0.137			6.670				9.645

综合上述三种方法预测结果,到 2000年东湖主要污染指标分别达 TN= 0.174 TP= 1.98 COD_{mn}= 6.70 BOD= 9.00 总体趋势表明,东湖的污染将趋向不断恶化,特别是城市污水的不断增长,排入东湖水体的氮、磷营养物质和有机物也随之加大,这是导致东湖水质污染不断加重的主要原因。按 2000年我国国民经济翻两番的发展速度和当前污水治理的经济与技术水平,到本世纪末下世纪初,东湖水质污染程度和范围都将明显加重。也就是说,氮、磷、有机物污染的控制是东湖污染亟待解决的重要环境问题。

表 4 东湖水中 TP TN 质量平衡预测浓度 ($\text{mg} \cdot \text{L}^{-1}$)

年份	TP	TN
1991~ 1993	0.138	1.51
2000	0.174	1.86
2005	0.202	2.16

5 东湖污染防治对策

5.1 东湖入湖污染物削减量

根据湖泊中污染物质平衡原理,在稳态情况下反推得计算公式:

$$w = k(c - \alpha)V + q(c - \alpha) \quad (33)$$

总氮、总磷削减量按合田健公式反推可得计算公式:

$$w = k(c - \alpha) \cdot s \cdot \bar{Z}(Q/V + 10/\bar{Z}) \quad (34)$$

东湖削减量计算结果列于表 5 由表 5 显示,随着国民经济的发展,到 2000年和 2005年东湖的 COD 浓度将超过 V 类标准,比较干净的湖区也将超过 III 类标准,达到重污染级别。按湖泊营养级评价,东湖在 2000年和 2005年将处于极富营养水平。

表 5 东湖污染物削减总量 ($\text{t} \cdot \text{d}^{-1}$)

年份	COD	TP	TN
1992	1 152.1	35.2	222.8
2000	3 515.4	61.4	425.8
2005	8 130.0	75.3	574.4

5.2 防治对策

1) 尽量减少入湖营养物质和有机物。据东湖污染总负荷统计,污染主要来自点源,而其中

生活污水中氮、磷分别为污染负荷总量的 83.7% 和 82.3%。由此可见,截污和污水处理工程是东湖污染治理的关键

2) 尽可能进行水体置换。根据西湖的经验,调水搞活水体,稀释营养物质和有机物是可取的。建议调水前放水,在出水处改溢流为底部放水,这一方面在调水前可适当降低湖面,为调水空出容积,达到换水目的;另一方面,底部放水可以带出湖底污泥

3) 有计划清理底泥,消除二次污染源。底泥中的氮、磷物质足以造成湖泊的富营养化,加上湖底裂层中的富磷,严重影响和干扰其它治理措施的效果,亦是东湖富营养化难以根除的根源。为此,必须集中力量,短期内疏浚,疏一段,铺盖一段适度的砂砾,以防底泥上泛。从湖体氮、磷平衡来看,现在东湖内负荷量大大超过外来负荷,而且还在不断增加。据 1992 年湖北省环保所统计资料,东湖底泥中积累氮、磷分别为 21 823.22 和 1 469.71t,可见疏浚底泥尤为重要

总之,东湖的污染治理各项措施必须高效而且各项技术措施又应协调,以发挥综合效能的作用

参 考 文 献

- 1 顾丁锡,舒多华.湖泊水污染及其规划方法.北京:中国环境科学出版社,1988.
- 2 刘鸿亮,韩国刚,严济民,等.中国水环境预测与对策概论.北京:中国环境科学出版社,1988.
- 3 傅国伟,程声通.水污染控制系统规划.北京:清华大学出版社,1985.
- 4 邓聚龙.灰色预测与决策.武汉:华中理工大学出版社,1988.
- 5 肖庭延.实用预测技术及其应用.武汉:华中理工大学出版社,1988.
- 6 陶本藻.测量数据统计分析.北京:测绘出版社,1992.
- 7 金相灿.中国湖泊水库环境调查研究.北京:中国环境科学出版社,1989.

Establishment and Application of Water Quality Forecasting Model of Lake

Wan Youchuan Li Zhisheng Liang Xiaomin Liu Liangming
(School of Information Engineering, W TU SM, 39 Luoyu Road, Wuhan, China, 430070)

Abstract In this paper, forecasting models, such as multivariate regression model, grey model and mass balance model have been established on the basis of analysis of monitoring data for East Lake, Wuhan City. In order to improve precision, the grey model has been revised by using residual error equation. The population, economy, BOD, COD, TN, TP about 2000 have been forecasted with these models. The proposals of administering for East Lake have been advanced.

Key words forecasting; multivariate regression; grey model; mass balance