

专论与综述

系统识别与参数估计方法在河流水质模型中的应用

侯 然 杰

(北京师范大学地理系)

一、引 言

在建立河流水质数学模型的过程中,模型结构一旦决定之后,剩下的最重要工作就是确定模型中的未知参数(或参变函数)。一般情况下,这些参数既不能通过实验得到,也没有现成的公式可以计算,只能根据实际河流水质监测数据进行估计。

在控制理论中,系统模型的参数估计是系统识别的重要内容。而系统识别则是随着系统可观测性、可控制性研究迅速发展起来的一个重要研究领域。

1971年, A. J. Koivo 和 G. R. Phillips 第一次把参数估计方法引入河流水质模型^[1]。此后,该方法在河流水质模型中获得越来越广泛的应用^[2-5],近年来已经形成了一个独特的研究领域。

本文试图结合图们江的实际工作谈谈这种方法在建立河流水质模型中的应用。

二、参数估计方法的数学原理

我们所说的河流水质模型通常是指描述河流中 DO-BOD 动态平衡关系的微分(或差分)方程^[6]。假定通过模型结构识别我们确定了如下的 DO-BOD 模型结构:

$$\frac{\partial X(L, t)}{\partial t} + v \frac{\partial X(L, t)}{\partial L} = KX(L, t) + S \quad (1)$$

其中,

$$X(L, t) = \begin{bmatrix} x_1(L, t) \\ x_2(L, t) \end{bmatrix}$$

$$K = \begin{bmatrix} -(K_1 + K_3) & 0 \\ -K_1 & -K_2 \end{bmatrix}$$

$$S = \begin{bmatrix} u_1 + S_b \\ u_2 + K_2 C_s + D_B \end{bmatrix}$$

$x_1(L, t)$: BOD 浓度, (毫克/升);

$x_2(L, t)$: DO 浓度, (毫克/升);

K_1 : 耗氧系数, (1/天);

K_2 : 复氧系数, (1/天);

K_3 : BOD 沉淀——再悬浮系数, (1/天);

u_1 : BOD 的排放量, (毫克/升·天);

u_2 : DO 的输入量, (毫克/升·天);

C_s : DO 的饱和浓度, (毫克/升);

S_b : BOD 的分散输入量, (毫克/升·米);

D_B : 光合作用与生物呼吸引起的 DO 净增加量, (毫克/升·天);

v : 河流的流速, (公里/天);

L : 距离, (公里);

t : 时间, (天);

设该河流水质系统的观测方程为:

$$Y(L_i, t) = X(L_i, t) + \eta_i(t) \quad (2)$$

其中,

$$Y(L_i, t) = \begin{bmatrix} y_1(L_i, t) \\ y_2(L_i, t) \end{bmatrix}$$

$$X(L_i, t) = \begin{bmatrix} x_1(L_i, t) \\ x_2(L_i, t) \end{bmatrix}$$

$$\eta_i(t) = \begin{bmatrix} \eta_{i,1}(t) \\ \eta_{i,2}(t) \end{bmatrix}$$

$x_1(L_i, t)$, $x_2(L_i, t)$ 分别为 t 时刻在空间点 L_i 处 BOD、DO 浓度的实际值; $y_1(L_i, t)$, $y_2(L_i, t)$ 分别为 t 时刻在空间点 L_i 处 BOD、DO 浓度的观测值; $\eta_{i,1}(t)$ 、 $\eta_{i,2}(t)$ 分别为 BOD、DO 的观测噪声, 假定它是平均值为零、协方差矩阵为已知的白噪声。

如果把根据模型(1)计算的浓度值记为:

$$X_p(L_i, t, K) = \begin{bmatrix} x_{p,1}(L_i, t, K) \\ x_{p,2}(L_i, t, K) \end{bmatrix} \quad (3)$$

那么参数估计的目的就是选择参数矩阵 K , 使得 X_p 与系统观测值 Y 最佳拟合。为了评价拟合的好坏, 通常需要定义一个适当的测度。在水质模型参数估计中, 最常使用的测度是形如下式:

$$J = \sum_{i=1}^n \varepsilon_i W [\varepsilon_i]^T, i = 1, 2 \cdots n \quad (4)$$

的平方误差准则。式中 W 为一矩阵, 选择不同形式的 W 可以得到不同的参数估计算法。而式中,

$$\varepsilon_i = \begin{bmatrix} \varepsilon_{i,1} \\ \varepsilon_{i,2} \end{bmatrix} = \begin{bmatrix} y_1(L_i, t) - x_{p,1}(L_i, t) \\ y_2(L_i, t) - x_{p,2}(L_i, t) \end{bmatrix} \quad (5)$$

现在, 我们只要使做为 K 的函数的 J 为最小, 就能在平方意义下最佳地确定参数矩阵 K 的估计值 K^* , 这个过程可以表达为:

$$\lim_{r \rightarrow \infty} \{E[(K^r - K^*)]\} = 0 \quad (6)$$

式中 K^r 代表第 r 次迭代时的 K 值, r 是迭代次数。当 r 趋近于无穷时, K^r 将收敛于它的真值 K^* 。

三、参数估计的算法

许多算法都可以用来进行参数估计。这里我们给出一种建立在统计近似基础上的梯度法。为了简化, 我们取(4)式中的矩阵 W 为对角线矩阵, 即令:

$$W = \begin{bmatrix} W & 0 \\ 0 & (1 - W) \end{bmatrix} \quad (7)$$

这时我们得到:

$$J = \sum_{i=1}^n [W[y_1(L_i, t) - x_{p,1}(L_i, t)]^2 + (1 - W)[y_2(L_i, t) - x_{p,2}(L_i, t)]^2] \quad (8)$$

对于(8)式, 在三维参数空间里, 它的梯度矢量为:

$$\nabla J = \left[\frac{\partial J}{\partial K_1} \quad \frac{\partial J}{\partial K_2} \quad \frac{\partial J}{\partial K_3} \right]^T \quad (9)$$

梯度的负方向 $-\nabla J$ 是函数 J 下降最快的方向, 因此用它做为探查 J 最小值的方向能得到最陡的下降。至于(9)式中的偏导数, 可以从下式

$$\frac{\partial J(K'_\theta)}{\partial K'_\theta} = \frac{J(K'_\theta) - J(K'^{-1}_\theta)}{K'_\theta - K'^{-1}_\theta}, \quad \theta = 1, 2, 3; \quad (10)$$

求出。式中 K'_θ 代表 K_θ 的第 r 次近似值。从(10)式我们能得到下面的参数估计算法:

$$K'^{r+1}_\theta = K'_\theta - \rho_r \frac{\frac{\partial J}{\partial K'_\theta}}{\sqrt{\sum_{\theta=1}^3 \left[\frac{\partial J}{\partial K'_\theta} \right]^2}} \quad (11)$$

式中 ρ_r 是一个由探查程序调节的正数, 它的选取应满足下面三个条件:

$$\lim \rho_r = 0 \quad (12)$$

$$\sum_{r=1}^{\infty} \rho_r = \infty. \quad (13)$$

$$\sum_{r=1}^{\infty} \rho_r^2 < \infty \quad (14)$$

条件(12)要求 ρ_r 是一个以零为极限的无穷递减序列, 条件(13)保证探查的每一步参数值都有充分的变化, 条件(14)保证探查程序在平方意义下的收敛性。

在确定性(即 $\eta_i(t) = 0$) 情况下, 当 $K_\theta = K^*_\theta$ ($\theta = 1, 2, 3$) 时, 梯度矢量(10)变为零, 这时函数 J 有最小值。因此, 我们能利用(11)式得到模型参数的最佳估计值 K^*_θ 。

四、参数估计方法的应用

在建立图们江水质模型的过程中, 我们

使用了上述非线性参数估计方法.

在图们江,经过模型结构识别,我们得到了下面的 Dobbins/Camp 型模型结构:

$$\left\{ \begin{array}{l} \text{BOD: } \frac{dx_1}{d\tau} = -(k_1 + k_3)x_1 \end{array} \right. \quad (15a)$$

$$\text{DO: } \frac{dx_2}{dx} = -k_1 x_1 + k_2 (c_s - x_2) \quad (15b)$$

式中 τ 为从参考点计算的流动时间, 其他符号的意义同前。这个方程可以解析求解, 其解为:

$$\text{BOD: } x_1 = x_{1,0} \cdot e^{-(k_1+k_3)\tau} \quad (16a)$$

$$\text{DO: } x_2 = c_s - (c_s - x_{2,0})e^{-k_2\tau} + \frac{k_1 x_{1,0}}{k_1 - k_2 + k_3} (e^{-(k_1+k_3)\tau} - e^{-k_2\tau}) \quad (16b)$$

式中, $x_{1,0}$ 和 $x_{2,0}$ 分别为 BOD、DO 浓度的初始值,

把根据方程(16a)和(16b)计算得到的 x_1, x_2 值分别做为(5)式中的 $x_{p,1}$ 和 $x_{p,2}$ 。根据(8)式能够得到下面的目标函数:

$$J^i = \sum_{i=1}^n [W[y_{1,i} - x_{1,i}(\tau_i, k_1^i, k_3^i)]^2 + (1 - W)[y_{2,i} - x_{2,i}(\tau_i, k_1^i, k_2^i, k_3^i, c_s^i)]^2] \quad (17)$$

式中的下标 i 代表第 i 个测量点, 上标 j 代表

第 i 次监测, 其余符号的意义同前。利用这个目标函数, 根据从图们江上得到的数据 (表 1) 进行参数估计能够得到表 2 所示的结果。从表 2 我们可以看到, 对于每个以一组水温、流量、流速数据为特征的稳态, 我们都得到了一组参数估计值。因此, 我们有可能通过预先设定参数 k_1, k_2, k_3 的函数表达式结构, 然后再次使用参数估计方法确定表达式中的常参数从而把 k_1, k_2, k_3 表达为水温、流量、流速的函数。在图们江, 我们假定 k_1, k_2, k_3 具有下面的函数形式:

$$k_1 = a_1 \cdot a_2^{(T-20)} \quad (18)$$

$$k_2 = a_3 \cdot Q^{a_4} \cdot a_5^{(T-20)} \quad (19)$$

$$k_3 = a_6 \cdot e^{-a_7 V} - a_8 \quad (20)$$

式中 Q 、 V 、 T 分别代表流量、流速、水温, $a_1 \cdots a_8$ 是待定的常参数。把表 2 中的 k_1^i 、 k_2^i 、 k_3^i 做为测量值,把由式 (18)、(19)、(20) 给出的 k_1 、 k_2 、 k_3 值做为模型计算值。根据 (8) 式我们能得到类似 (17) 式的目标函数:

$$J_1 = \sum_{i=1}^m [k_1^i - a_1 \cdot a_2^{(T_i - 20)}]^2 \quad (21)$$

$$J_2 = \sum_{i=1}^m [k_2^i - a_3 \cdot Q^{a_4} \cdot a_5^{(T_i - 20)}]^2 \quad (22)$$

表 1 建立图们江水质模型所用的部分数据

		-1Km	0	6.33	10.4	23.65	36.35	流量 m ³ /s	流速 Km/day	水温 ℃
		Ref	0	1	2	3	4	Q	V	T
1	BOD	8.92		83.50	73.50	36.15	36.75	8.20	30.79	0.2
	DO	5.47		1.80	3.15	0.00	0.00			
2	BOD	2.09		72.44	72.44	70.44	58.44	9.85	38.59	2.5
	DO	10.90		5.90	5.90	5.10	2.30			
3	BOD	1.90		21.83	14.15	11.53	13.73	25.00	45.96	13.6
	DO	8.95		4.28	4.97	4.65	5.37			
4	BOD	4.80		8.90	5.93	5.00	4.86	36.60	60.48	20.6
5	BOD	6.70		10.33	5.33	5.23	5.21	96.50	106.3	21.6
	DO	6.30		4.56	4.30	4.88	3.91			

表2 图们江水质模型的参数估计结果

i	Q_i (m^3/s)	v_i (Km/day)	T_i ($^{\circ}\text{C}$)	k_1^i ($1/\text{day}$)	k_2^i ($1/\text{day}$)	k_3^i ($1/\text{day}$)
1	8.20	30.79	0.2	0.270	1.03	0.280
2	9.85	38.59	2.5	0.380	1.010	0.091
3	25.00	45.96	13.6	0.940	1.820	-0.170
4	36.60	60.48	20.6	3.440	2.360	-0.250
5	96.50	106.30	21.6	3.460	4.290	-0.280

表3 常参数 a_1 — a_8 的估计结果

a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
2.635	1.117	0.443	0.484	1.014	2.756	0.051	0.306

$$J_3 = \sum_{i=1}^m [k_3^i - (a_6 \cdot e^{-a_7 v} - a_8)]^2 \quad (23)$$

再次利用(11)式给出的算法,通过分别使(21)、(22)、(23)式取最小值,我们就能得到常参数 $a_1 \cdots a_8$ 的估计值(见表3)。

这样,利用两步参数估计法我们就确定了图们江水质模型(15)中的参数 k_1 、 k_2 、 k_3 的函数表达式:

$$k_1 = 2.635 \times 1.117^{(T-20)} \quad (24)$$

$$k_2 = 0.443 Q^{0.484} \cdot 1.014^{(T-20)} \quad (25)$$

$$k_3 = 2.756 \cdot e^{-0.051v} - 0.31 \quad (26)$$

通过参数估计是否得到了适当的参数值尚需由模型检验来证明。由于篇幅所限,这里省略了模型的检验。对图们江水质模型参数估计结果的检验证明,我们上面得到的结果是符合实际情况的。

最后,我们还需要指出:第一,在进行模型参数估计时,参数初值的选择对收敛速度和估计结果都有很大的影响,因此在实际应用时,必须根据河流的具体情况和计算经验谨慎地选择参数的初值。第二,如果能够同时得到 BOD、DO 数据,在参数估计中应尽可能同时使用两种数据。这时式(17)中的加权系数 W 代表 BOD 数据对 DO 数据的相对可信度。根据我们的研究^[7],选择 $W = 0.5$

一般能得到合理的参数估计结果。

五、结 束 语

在控制理论中发展起来的系统识别与参数估计方法已经成功地应用于河流水质系统的模型化,并且正在形成一个引人注目的研究领域。但是,在河流水质模型的参数估计中仍然存在一些不确定性,这些不确定性主要来自下述三个方面:

(1) 在建立河流水质模型过程中,分析者常常无法确定所研究系统的因果关系,有时即使了解了那些关系,也无法确定其中一个关系是基本的。这就给模型结构识别造成了许多困难。参数估计建立在模型结构识别基础之上,结构识别的不确定性必然引起参数估计的不确定性。

(2) 各种随机干扰的影响

处于自然界中的河流水质系统经常受一些随机事件(如降雨,地表径流等等)的影响。这就使得水质监测数据具有很高的噪声水平。要想通过参数估计最佳地拟合这样的数据几乎是不可能的。通常只能在“白噪声”的假设下进行拟合。由于实际的噪声往往不是“白”的,因此也可能使水质模型的参数估计产生一些不确定性。

(3) 水质监测数据的可用性

在目前的河流水质监测中, 取样次数少且在时间上不连续是很常见的现象. 这种监测方式不可避免地会在水质数据中引入一些偶然性. 因此以这些数据为基础的参数估计也必然会带有某种程度的不确定性.

目前, 河流水质模型参数估计方法的主要研究课题就是设法减少这些不确定性. 这固然可以通过改善监测方式, 使用更先进的监测仪器来部分解决, 但更主要的是设法发展一些适合水质监测数据特点的新参数估计方法.

参 考 文 献

[1] Koivo, A. J. and Phillips, G. R., *Water Re-*

sour. Res., 7(4), 853—962 (1972).

[2] Koivo, H. N. and Koivo, A. J., *Proc., 3rd IFAC Symposium on Identification and System Parameter Estimation*, The Hague, pp. 12—15, June 1973.

[3] Rinaldi, S. et al., *Proc., 4th IFAC Symposium on Identification and System Parameter Estimation*, Tbilisi, USSR, 1976.

[4] Shastry, J. S., Fan, L. T. and Erickson, L. E., *Jour. Env. Eng. Div., A. S. C. E.*, 99(EE3), 315—331 (1973).

[5] Rinaldi, S. et al., *Parameter Estimation of Streeter-Phelps Models*, Centro Teoria dei Sistemi, C. N. R., Milano, Italy, 1979.

[6] Dobbins, W. E., *Jour. San. Eng. Div. A. S. C. E.*, 90(SA3), Proc. Paper 3949, 53—78 (1964).

[7] 侯然杰等, 环境科学学报, 1 (2), 166—167 (1981).

热污染的影响及其预测预报

万 肇 忠

(广东省环境保护研究所)

一、前 言

把温度上升了的冷却循环水直接排放到环境水圈中时, 就称为温排水. 当有高于环境水体温度 4°C 以上的热废水持续流入时, 就可以认为该水体受到了热污染.

当工矿企业把热废水(或温排水, 下同)直接排放到水体时, 其能量就以热的形式加到水体中去, 使得该水体的水温比周围水体的温度要高些. 如果仅考虑热废水的热, 而暂不考虑其中是否含有其它有害物质对环境的影响, 那么, 水温的升高也许是污染的, 也许是有益的. 例如, 水温的升高(在此假定是缓慢升高而不是突然升高), 会影响鱼类的迴游、迁徙、干扰其繁殖、发育、生长, 甚至也会使之死亡. 但是, 寒冷地带的热废水, 往往可

用来调节鱼类越冬和催产. 据说黑龙江省有此成功之例. 这说明对热废水存在一个如何合理利用的问题. 因此, 笔者认为: 所谓热污染, 应由受热废水影响区域的生物适温范围(见表 1, 2)和有益生物的减少程度或速率来决定.

随着工农业的发展和人民生活水平的提高, 对电力的需求量也势必增加. 从日本对电力需求量的预测来看(表略), 无论是火力、水力, 还是原子能, 对电力的绝对需求量都在增加, 尤以原子能增加得最快, 即无论是电力的绝对增量还是每年度所占的百分比, 均是递增的. 于是, 佐佐木忠义根据近藤的电力使用流程图, 对日本 2000 年所需 5 亿千瓦电力预测之分析后提出, 应把热污染作为全球性问题来考虑, 并呼吁“在引起气候变化之