

稀疏数据下复杂流域的水质模拟 :以赣江为例

邓义祥,陈吉宁,程声通 (清华大学环境科学与工程系,北京 100084)

摘要:以赣江流域为例,对稀疏数据条件下复杂流域水质模型的建立和参数识别进行了探讨.采用了结构相对比较简单的 CSTR 模型,将其参数划分为水文参数和水质参数,并分别进行识别.水文参数采用回归方法进行识别,对水文参数的扰动实验表明,水质对于水文参数在可能的取值范围内的变化不敏感,因此在随后的水质参数的识别中,将其固定在回归方法求出的最优值上.水质参数的确定则采用了模型方法和资料方法相结合的手段,综合考虑了赣江流域可获得的数据信息、文献资料中对于参数的经验取值和模拟者在大量数学建模中的经验,最终根据河流的不同类型确定了赣江流域的水质参数,并对模型进行了验证.

关键词:复杂流域;水质模型;赣江流域;CSTR 模型

Water quality simulation in a complicated large watershed: A case study of the Ganjiang watershed

DENG Yixiang, CHEN Jining, CHENG Shengting (Department of Environmental Science and Engineering Tsinghua University, Beijing 100084, China)

Abstract: Due to lack of data, it is difficult to develop a mathematical water quality model for a large and complicated watershed. It is thus crucial to fully utilize both quantitative and qualitative information from the studied river system as well as elsewhere. Taking the Ganjiang watershed as a case study, a practical way for identifying a watershed water quality model was illustrated. The results suggested that the coefficients, describing the relationships between flow and average depth and between flow and section area, across the river, presented little change along time in most observed hydrological stations. This was consistent with the observation that the hydrological conditions of the Ganjiang River were stable over years which indirectly provided evidence for the validity of the hydrological model. For the water quality components, optimization could easily lead to the estimated parameters highly based. Combination of optimization and the *a priori* information, in particular from similar watersheds, could give a reliable and better identified model parameters. Validation results indicated that the proposed identification approach was successful in simulating the water quality of the complicated large watershed.

Key words: complicated large watershed; water quality model; Ganjiang watershed; CSTR

数学模型是流域水质管理的重要手段,为水质预测、环境容量计算等提供了关键的技术支持.然而对于大型复杂流域,由于获得全面的水量和水质信息的困难,相应数学模型的建立和应用存在着严重的技术障碍^[1],如果仍然仅仅依赖传统的优化方法进行参数识别,得到的参数与真实情况可能偏离很大.而人们在大量的数学模型应用中所积累起来的经验无疑提供了数学建模的另外一种信息.因此,综合考虑根据实际数据所体现的信息和人们在建模经验中所积累的信息,大大地提高了数学模型的可靠性.本文以赣江流域为例,对复杂流域的数学建模进行了探讨.

1 赣江流域概况

赣江是江西省境内第一大河,自南向北纵贯全省,从源头到河口全长 766 km,水系发育,支

收稿日期:2002-12-24;修订日期:2003-04-04

基金项目:国家环保局赣江流域环境遥感与数字化管理研究,“985”项目

作者简介:邓义祥(1974—),男,博士研究生,E-mail:dxy99@mails.tsinghua.edu.cn

流众多,就水量而言,为长江第二大支流.赣江沿岸有 42 条支流汇入,其中流域面积在 1000 km² 以上的主要支流有 13 条.通常以赣州市和新干县城为界,将赣江分上、中、下游.上游多山,中游丘陵、盆地相间,下游以冲积平原为主.赣江流域范围包括 46 个县(市),流域面积 83 500 km²,约占江西省总面积的 51%;流域内人口数量约占江西全省的 40% 左右;工农业总产值占江西省 60% 以上;很多重要指标均占江西全省总量的一半以上.

2 模型简介

赣江流域水环境模型的选择主要考虑了以下几个因素:(1)赣江流域的河流特点.赣江流域范围广,水系复杂.(2)模型用途.本研究面向赣江流域的污染控制规划,预测赣江在一定排污情景下的水质,并选择最优控制方案.(3)可获得的数据.数据采样时间频率小(特别是水质和污染源数据)、水质和水文断面少且空间分布对数学模型的适宜性低、数据的精度较低以及数据之间的配套性差等.考虑到上述特点,决定采用 CSTR 模型构建赣江流域水质模型.该模型实际上是结合了传统的水力学模型和化学工程(或更普遍的说是模拟混合或工业处理过程)模型.其最基本的思想是把河道划分成若干连续的段,段内划分箱子,在每一段内模型参数认为保持不变,且每个箱内水质完全均匀混合.该模型是由零维模型串联而成的一维模型,其完全均匀混合的概念具有高度的概括性,适于处理大流域水环境问题,因此 CSTR 模型曾被广泛地应用于国内外的河流水质模拟中.

结合赣江流域的实际情况,本研究对模型进行了一定的改进,使 CSTR 模型处理河网的能力得到了增强(图 1),包括:(1)CSTR 模型能够处理单向流情况下任意复杂的河网关系;(2)CSTR 模型具有灵敏性分析的能力;(3)建立了以 Access 为基础的数据库和友好的程序界面,方便了数据录入.CSTR 模型更为详细的介绍见文献[2].

考虑赣江流域的环境污染现状和规划要求,本研究模拟的水质项目只包括 COD_{Mn} 和 NH₄⁺-N.

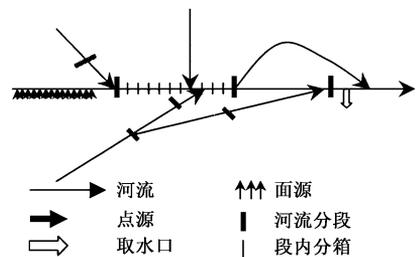


图 1 CSTR 模型河流系统

Fig. 1 Rivers of CSTR model

3 水文参数的识别

3.1 水文参数的识别

在 CSTR 模型中,需要建立流量-平均水深和流量-过水断面面积之间的关系.通常以指数关系来描述,即:

$$Q = k_1 H^{n_1} \quad (1a)$$

$$Q = k_2 S^{n_2} \quad (1b)$$

式中: Q ——流量 (m³/s); H ——平均水深 (m); S ——过水断面面积 (m²); k_1 、 n_1 、 k_2 、 n_2 为相应的参数.

采用 1996—1998 年的水文实测数据确定上述参数.赣江流域在设定的水文站有比较详细的水文数据,但是缺少成对的水文站,因此不能利用上下游断面的输入输出关系进行参数识别.在这种情况下,认为各个水文站的数据能够代表一定距离内河段的水文响应特点.采用回归分析方法来确定 CSTR 模型的水文参数,对公式(1)进行对数变换后,将其转化为线性形式,

表1 赣江流域芦溪水文站的水文参数取值

Table 1 Hydrological parameter estimation of Luxi station,
Ganjiang watershed

参数	Q-H			Q-S		
	均值	下限	上限	均值	下限	上限
k	71.68	68.09	75.26	0.025	0.023	0.026
n	2.122	2.016	2.229	1.884	1.790	1.979

利用最小二乘法求出 k_1 、 n_1 以及 k_2 、 n_2 。为了考虑误差的影响,将 k_1 和 n_1 分别上下扰动 10%,将 k_2 和 n_2 分别上下扰动 5%,分别作为其取值的上限和下限,并将回归求出的参数值作为均值。为演示起见,这里只以芦溪水文站为例,所求得的水文参数见表 1。

图 2(a)是当 k_1 、 n_1 分别取均值和上下限时,预测的流量及其变化范围。图 2(b)是当 k_2 、 n_2 分别取均值和上下限时,预测的流量及其变化范围。从图 2 可以看出,当 k_1 、 n_1 、 k_2 、 n_2 在其取值区间内扰动时,能够比较准确地预测流量的取值及其变化范围。

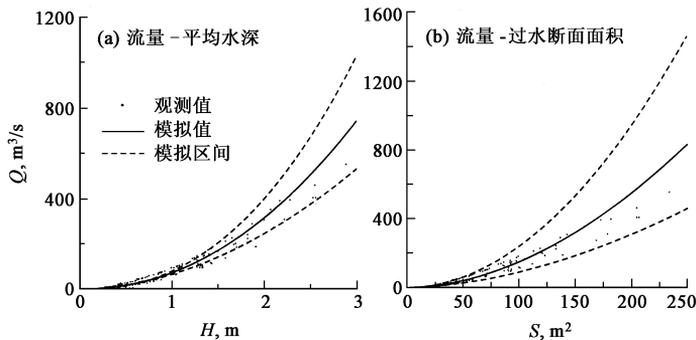


图2 芦溪水文站的水文参数的识别

Fig. 2 Estimation of hydrological parameters of Luxi station

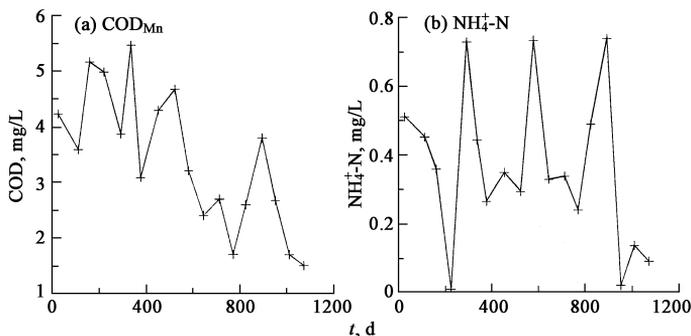


图3 水文参数的扰动对于水质模拟的影响

Fig. 3 Effect of hydrological parameter disturbance on water quality

3.2 水质对于水文参数的敏感性

以宣风-棚下段为例,水文参数取离该段最近的芦溪水文站的参数。不失一般性,随机确定一组水质参数,考虑水文参数在其范围内随机变化时的水质变化情况。图 3 是采用水文参数的均值和极限值的 9 种组合计算宣风-棚下段 1996—1998 年的水质变化情况。从图 3 可以看出,9 种组合的水质曲线几乎完全重合。

更进一步分析,比较各种组合 1996—1998 年模拟水质与 k_1 、 n_1 、 k_2 、 n_2 取均值时模拟水质的偏差。研究发现, COD_{Mn} 最大偏差不超过 1%, NH_4^+-N 不超过 1.5%。这说明,在宣风-棚下段,

水质对于水文参数的取值不敏感. 计算结果显示,其它各段也有类似的情况.

上述分析表明,试图通过水质关系来识别水文参数是不现实的,有必要单独确定河段的水文参数. 由于水文参数在可能的取值范围内对水质模拟的影响不大,因此,在以后水质模型的参数识别中,水文参数始终取均值,从而在一定程度上降低了 CSTR 模型参数识别的复杂性.

4 水质参数识别

由于赣江流域水质数据在时间上监测频率小,上下游断面有相关关系的监测点不多,特别是污染源的信息有可能严重失真,因此参数优化识别的随机性大. 考虑到这种情况,赣江流域的参数识别采用了优化方法与文献资料方法相结合的方法.

4.1 优化方法

采用最小二乘法进行优化方法的参数识别,即目标函数为实测值与模拟值之差的平方和. 采用 1996—1997 两年的数据进行参数识别,采用 1998 年的数据进行检验. 根据赣江流域的河流特点、水文站的位置和水质断面的位置,最终确定 7 个参数识别河段. 优化参数识别的结果见表 2.

表 2 赣江流域 7 河段初步参数识别结果

Table 2 Initial identification of the parameters of the seven river sections

(1/d)

河段	生米 至朝阳水厂	大巷口 至石油库	神岗山 至大巷口	大余市郊 至大余新城	芦溪 至宣风	宣风 至棚下	朝阳水厂 至八一桥北
类型	平原河流	平原河流	平原河流	山区河流	山区河流	山区河流	类似水库
COD _{Mn}	0.88	0.02	1.02	0.06	0.38	0.04	0.26
NH ₄ ⁺ -N	0.8	1.5	0.04	1.3	0.02	0.04	0.66

4.2 文献资料法

由于稀疏数据,优化方法的参数识别结果具有一定的随机性,参数的规律性特点不明显,若仅仅采用优化方法求得的参数进行水质预测,有可能得出不正确的结论. 因此,在现有数据信息有限的情况下,可以参考同类河流的水质参数取值,以最终确定模型的水质参数. 国内相关河流水质参数的取值范围见表 3.

以宣风-棚下段为例,水质参数的初步识别结果和验证见图 4.

表 3 国内部分河流水质参数的取值范围

Table 3 Range of degradation parameters in some rivers of China (1/d)

序号	类别	下限	上限	位置	类型	文献
1	COD _{Mn}	0.253	0.26	无锡县河网	类似于水库	文献[3]
2	COD _{Mn}	0.3	5.0	图们江	平原河流	文献[4]
3	COD _{Mn}	0.05	0.29	山东小清河	山区河流	文献[5]
4	COD _{Mn}	0.64	1.08	浑江吉林省段	山区河流	文献[6]
5	COD _{Mn}	0.11	0.20	金堤河	平原河流	文献[7]
6	COD _{Mn}	0.06	0.20	上海地区河网	类似于水库	文献[8]
7	COD _{Mn}	0.25	0.25	柳江柳州段	山区河流	文献[9]
8	NH ₄ ⁺ -N	0.16	0.16	柳江柳州段	山区河流	文献[9]
9	NH ₄ ⁺ -N	0.01	0.50	QUAL-	一般河流	文献[10]

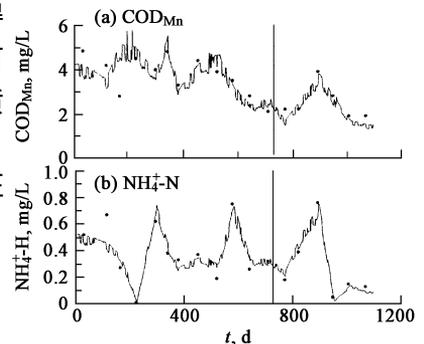


图 4 优化方法的参数识别与检验

(竖线以前(1996—1997年)是标定,以后(1998年)是验证,图中点为观测值,线为模拟值)

Fig. 4 Identification and test of optimization method

(Before the erect line (1996—1997) is identification and after that (1998) is test, the dots are observation values, solid line is the simulation results)

4.3 模型水质参数的确定

综合优化方法和上述文献资料,考虑到赣江流域的河流特征,将赣江流域的河道划分为3种类型,即山区河流、平原河流和水库,分别确定其参数取值。

上述优化方法和文献资料法的数据整理于表4,其中剔除了图们江的上限数据。根据这些数据,计算得到了相应的均值和方差。

确定参数取值时综合考虑模型参数的均值和方差,并倾向于保守估计。另外还考虑了以下两个原则:(1)不同类型河流物质衰减排序为:山区河流 > 平原河流 > 水库;(2) $\text{NH}_4^+ - \text{N}$ 的降解速度略快于 COD_{Mn} 。最终取定不同类型河流水质参数的上限,同时分别各自降低约0.2/d作为取值下限,以避免参数取值过高而带来的风险,最终确定的模型参数上限和下限见表5。

表4 不同类型河流的水质参数

Table 4 Parameters of different types of rivers

(1/d)

序号	山区河流		平原河流		水库型河流	
	COD_{Mn}	$\text{NH}_4^+ - \text{N}$	COD_{Mn}	$\text{NH}_4^+ - \text{N}$	COD_{Mn}	$\text{NH}_4^+ - \text{N}$
1	0.05	0.16	0.02	0.80	0.25	0.01
2	0.29	0.16	0.11	1.50	0.26	0.50
3	0.06	0.01	0.20	0.04	0.26	0.66
4	0.64	0.50	0.88	0.01	0.06	
5	1.08	1.30	0.30	0.50	0.20	
6	0.38	0.02				
7	0.25	0.04				
8	0.25					
9	0.04					
均值	0.34	0.31	0.30	0.57	0.21	0.39
方差	0.11	0.47	0.12	0.62	0.01	0.34

表5 赣江流域水质模型参数的取值范围

Table 5 Range of model parameters of Ganjiang watershed

(1/d)

河流类型	COD_{Mn}		$\text{NH}_4^+ - \text{N}$	
	上限	下限	上限	下限
山区河流	0.4	0.2	0.5	0.3
平原河流	0.3	0.1	0.4	0.2
水库型河流	0.2	0.05	0.3	0.1

5 CSTR 模型验证

考虑规划的目的和项目方的要求,并尽可能减少面源和随机因素的影响,采用每年的9月—12月水质和水量的平均值建立稳态模型进行验证。所采用的流量和水质是1996—1998各年9月—12月共3个时期的平均流量和平均水质。仍然采用上述优化方法时的河段,并分别将取值上限和取值下限代入模型进行检验,模型的检验结果见图5和图6。从图中可以看出:(1)相对误差比较对称地分布在0附近;(2)除

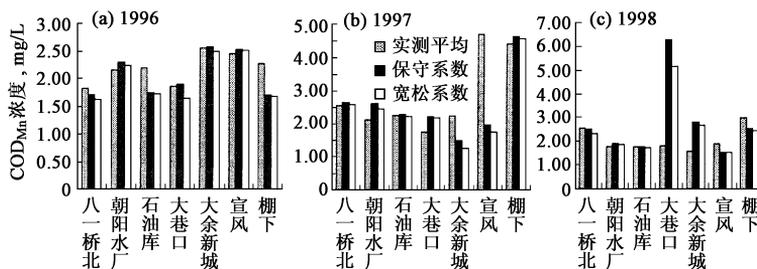


图5 各年9—12月 COD_{Mn} 实测平均值与计算值

Fig. 5 COD_{Mn} average observations and simulation results between Sep. —Dec. in different years

个别异常点以外,相对误差大都在 50% 的误差范围以内.在 30% 误差范围以内 COD_{Mn} 占 81%, NH_4^+-N 占 52%. (3) 模型对参数取值不敏感.考虑到在赣江流域具有复杂水系、面积广大并且可获得数据十分有限,认为所建立的模型是可以接受的.

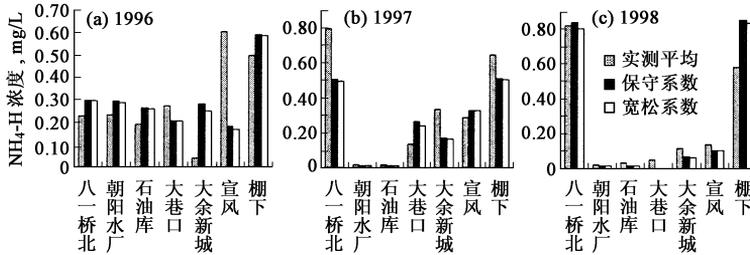


图 6 各年 9—12 月 NH_4^+-N 实测平均值与计算值

Fig. 6 NH_4^+-N average observations and simulation results between Sep. —Dec. in different years

6 小结

(1) 在数据稀疏条件下,仅仅依赖现有的实测数据采用优化方法进行参数识别是危险的.综合考虑从实际数据中获得的信息以及在长期模型应用中所积累的经验,确定数学模型的参数值,将降低由于实测数据的信息不足而带来的风险.

(2) 为了减少数学模型参数识别的复杂性,对某些敏感性不大的参数可单独进行识别,并将其固定为常数,再进行其它参数的识别.

(3) 本文采用上述方法建立了赣江流域的水质模型,并进行了参数识别.验证结果表明,水质对于参数的变化不敏感.在赣江流域稀疏数据的条件下,模型是可以接受的.

参考文献:

- [1] Jining Chen. Modeling and control of the activated sludge process: towards a systematic framework, PhD. thesis. The University of London, 1993. 130—165
- [2] 邓义祥,陈吉宁,杜鹏飞. HSY 算法在水质模型参数识别中的应用[J]. 上海环境科学, 2002, 21(8): 497—500
- [3] 孙勤芳,贺赜和,李锦秀,等. 河网水质数学模型的参数估算[J]. 农村生态环境, 1994, 10(4): 46—50
- [4] 田卫,俞穆清,刘桂琴. 图们江地区水环境容量及其对区域开发的影响研究[J]. 地理科学, 1998, 18(2): 169—175
- [5] 董梅,夏学梅,慕金波. 小清河干流水环境容量的计算[J]. 山东环境, 1998, 85(4): 8—10
- [6] 田卫,俞穆清,朱显梅等. 浑江吉林省段水环境容量及其在总量控制中的利用研究[J]. 东北师大学报(自然科学版), 2000, 32(3): 85—88
- [7] 程志臣,杨志杰,崔书臣,等. 金堤河水质污染模型的建立[J]. 仪器仪表与分析监测, 2000, 4: 58—60
- [8] 徐贵泉,褚君达,吴祖扬,等. 感潮河网水环境容量影响因素研究[J]. 水科学进展, 2000, 11(4): 375—380
- [9] 龚若愚,周源岗. 柳州柳州段水环境容量研究[J]. 水资源保护, 2001, 63(1): 32—48
- [10] 李炜. 环境水力学[M]. 武汉: 武汉水利电力大学出版社, 1999. 500—503