

不确定水质模型在城市河流水质模拟中的应用

田一梅¹, 刘扬¹, 王彬蔚^{1,2}

(1. 天津大学环境科学与工程学院, 天津 300072; 2. 陕西省环境监测中心站, 西安, 710054)

摘要: 鉴于城市景观河流受沿河排水污染, 水质波动较大, 建立了内嵌神经网络的一维不确定性水质模型, 利用改进适应度函数的遗传算法, 优化水质模型的参数解。经实例验证, 不确定性水质模型拟合的精度更高, 对排入污染物的波动更敏感, 其对景观河流水质预测的平均准确度基本在80%以上, 普遍高于确定性水质模型, 尤其是在靠近污染源的监测断面, 其不确定性水质模型预测优势更加明显, 更能适应变化的景观河流水体环境。

关键词: 不确定性水质模型; 人工神经网络; 遗传算法; 模拟

中图分类号: X522 **文献标志码:** A **文章编号:** 1674-4764(2011)03-0119-05

Application of Uncertain Model in Urban River Quality Simulation

TIAN Yi-mei¹, LIU Yang¹, WANG Bin-wei^{1,2}

(1. School of Environmental Science and Engineering, Tianjin University, Tianjin 300072, P. R. China;
2. Shanghai Environmental Monitoring Center of Shanxi Province, Xian 710054, P. R. China)

Abstract: Owing to the fluctuation of water quality in urban river which polluted by drainage along river, one-dimension uncertain water quality model embeded neural network is established. Genetic algorithms and a modified fitness function are used to optimize parameters of the uncertain model. Examples illustrate that the uncertain model has higher prediction accuracy with the average accuracy over 80% than the certain model, and is more sensitive to the fluctuation of pollutants discharged into the river. The uncertain model has a significant advantage of prediction and could better adapt to the changing urban water environment, especially at points close to the pollution sources.

Key words: uncertain water quality model; artificial neural network; genetic algorithms; simulation

水质模型是用来描述水体中污染物与时间、空间的定量关系, 描述物质在水环境中的混合、迁移过程的数学方程。根据模型中的变量是否为随机变量, 水质模型可分为确定性水质模型和不确定性水质模型。确定性水质模型, 如 SIMCAT, TOMCAT, QUAL2E, QUASAR, MIKE-11 和 ISIS^[1-2] 等已广泛应用于水环境的模拟预测, 但由于水环境的不确定性、复杂性, 使水质变化呈现出基于一定变化规律的不确定性变化特征, 同时由于监测数据本身均带有随机性, 故建立不确定性水质模型更能反映水质的实际变化规律^[3]。

早期的不确定性水质模型较多是利用随机理论建立的, 如 1966 年 Loucks 和 Lgmn 基于随机过程理论首先提出了预测最小溶解氧的概率模型, 1969 年 Custer 与 Krutcheff^[4] 提出了随机游动模型, 1984 年, Deway 提出了 BOD-NOD-DO 随机模型。而近年来, 随着神经网络理论的成熟, 将其应用于各类不确定性水质模拟问题已逐渐成为重要的研究方向之一, 2002 年 Huang 和 Foo^[5] 利用人工神经网络模拟了受支流流入、潮汐以及风力等不确定性因素影响的河流中盐度的变化; Maier 等人^[6] 利用人工神经网络对混凝剂的投加量进行了优化; Mozejko 等

收稿日期: 2010-03-10

基金项目: 国家水体污染控制与治理重大专项(2008ZX07314-003); 天津市科技创新专项资金资助项目(O6FZZDSH0090)

作者简介: 田一梅(1959-), 女, 教授, 博士, 主要从事环境系统优化研究, (E-mail) ymtian_2000@yahoo.com.cn

人^[7]和 Chen 等人^[8]模拟了河流中氮磷浓度的变化;利用人工神经网络模拟可以增强处理非线性问题的能力,使得水质预测精度得以提高。一般来说,由于人工神经网络算法局部精确搜索的特性,容易陷入局部最优,故对于波动性较大的数据,易导致过拟合而影响预测精度,利用全局搜索的遗传算法与之结合,能够有效的摆脱局部极点的困扰^[9]。以北方某工业园区 J 河为研究对象,通过监测、分析该河主要污染物指标的变化,研究沿河排水对河流水质的影响,利用内嵌神经网络建立了一维不确定水质模型,利用遗传算法求解模型参数,提高了对城市河流水质模拟预测的精度。

1 一维不确定性水质模型的建立

J 河部分为天然河道、部分经人工开挖,所处区域地势平坦,地面坡度小于 1/1 000,河水流速缓慢。河平均宽度 30 m,平均水深 2.2 m,河长约 13 km。其水源主要为附近河网来水及该区域径流雨水。经监测上游来水及沿河排水中部分指标经常不满足地表水 V 类标准,且波动较大,并影响了 J 河水质。

由于 J 河的宽度、深度与其长度相比较小,河流的竖向与横向均匀混合所需的距离远远小于 J 河的河长,因此可假设污染物在河流的横断面上瞬时达到均匀混合,故采用一维水质模型即可取得很好的近似。

利用一维水质模型预测 COD_{Cr} 和氨氮的变化,以 COD_{Cr} 为例,当考虑纵向弥散条件时,一维水质模型的方程为:

$$u_x \frac{\partial C}{\partial t} = D_x \frac{\partial^2 C}{\partial x^2} - kC \quad (1)$$

式中: C 为河段中某污染物质的浓度, mg/L; u_x 为河流的流速, m/s; x 为距污染源排放断面的距离, m; k 为污染物降解系数, d⁻¹; D_x 为河流的纵向弥散系数, m²/s。

J 河的纵向弥散系数可利用矩阵法^[10]求得,结果如表 1 所示, k 值可利用 MATLAB 中的 Fminsearch 函数拟合得到,为 0.095 2d⁻¹。

表 1 J 河各断面处纵向弥散系数

断面编号	2	3	4	5
断面距离/km	2	7	11	13
$D_x / (m \cdot s^{-1})$	0.6	7.04	20.1	28.2
$\bar{D}_x / (m \cdot s^{-1})$	14			

当不考虑弥散条件时,一维水质模型方程如(2)式所示,拟合得到的 k 值为 0.095 1 d⁻¹:

$$u \frac{dC}{dx} = -kC \quad (2)$$

图 1 所示的上述两种条件下拟合出的不同时刻各断面处 COD_{Cr} 浓度结果十分相近,因此为了简化一维模型,可忽略 J 河的纵向弥散作用。图 1 显示确定性一维水质模型对 J 河中污染物降解的拟合效果不理想,估计与沿河排水的水质波动有关,因此在忽略河流弥散作用的一维模型中增加一项对排水水质的波动的观测变量,建立了一维不确定水质模型,见式(3)。

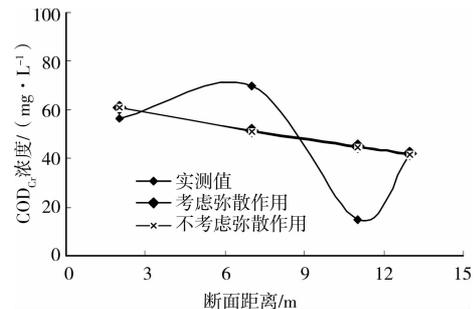


图 1 各断面的 COD_{Cr} 监测值与一维确定性水质模型预测值的比较

$$u \frac{dC}{dx} = -kC + U \quad (3)$$

式中: U 为水质变化不确定因素影响的外部观测量,可以是一个变量,也可以是一个函数。

式(3)建立的关键在于 U 与模型参数的合理确定。

1.1 不确定因素表征

水环境是一个充满不确定因素的复杂系统,其不确定性主要来源于污染物排放量与河流水文条件的不确定性、由于对水环境中复杂的物理生化反应机理认知不足造成的水质模型结构的不确定性、水质模型参数确定所需的河流及水质资料的不确定性、未观察到的模型输入过程中各种扰动所造成的不确定性^[11]等等。其中有些不确定因素是可以通过一定时间内的水质监测数据反映其主要变化信息。而神经网络是以数据为基础,通过对系统历史数据的学习训练来掌握系统内部的变化规律,无需构建模型,同时,神经网络应用简便,具有较强的自适应性,非常适合于处理具有不确定性和非线性的水环境问题。

鉴于上述原因并充分考虑目前水环境管理日常工作需求及水质监测的工作量等多方面条件,采用神经网络函数作为不确定因素影响的外部观测量,即:

$$U = \sum_{i=1}^m a_i h \left(\sum_{j=1}^n \omega_{i,j} C_{t-j} - \theta_j \right) \quad (4)$$

式中: $h(\cdot)$ 为激发函数,选择激发函数 $h(x) = \frac{1}{1 + e^{-x}}$; $a_i, \omega_{i,j}$ 为神经网络权重系数; θ_j 为阈值; $C_{t-1}, C_{t-2}, \dots, C_{t-n}$ 为污染物输入断面的 n 次监测结果。 m : 隐层神经元节点数。

1.2 不确定性水质模型建立

式(4)所构建的神经网络函数是以连续若干天污染物输入断面的水质监测数据作为网络的输入元,通过对一段时间内网络数据的学习、训练,感知污染物随机排放量的变化,从而建立一个具有学习功能的不确定性水质模型,见式(5):

$$u \frac{dC_t}{dx} = -kC_t + \sum_{i=1}^m a_i h \left(\sum_{j=1}^n \omega_{i,j} C_{t-j} - \theta_j \right) \quad (5)$$

式中: C_t 为 t 时刻河段中某污染物质的浓度, mg/L;

根据景观河监测数据分析及网络建模试算,选取 COD_{Cr} 、 $\text{NH}_3\text{-N}$ 作为建模指标,取污染物输入断面(1-1断面)4 d的水质监测数据作为模型训练的样本,则网络输入层节点数为4,其相应隐层节点数采用 Hecht-Nielsen 的“2N+1”法确定为9,经解式(5)可得不确定性水质模型为:

$$C_t^{\text{COD}} = C_0^{\text{COD}} e^{-\frac{k_{\text{COD}}}{u}x} + \frac{1}{k_{\text{COD}}} \sum_{i=1}^9 a_i^{\text{COD}} \frac{1}{1+e^{-x}} \left(\sum_{j=1}^4 \omega_{i,j}^{\text{COD}} C_{t-j}^{\text{COD}} - \theta_j^{\text{COD}} \right) (1 - e^{-\frac{k_{\text{COD}}}{u}x}) \quad (6)$$

$$C_t^{\text{NH}_3\text{-N}} = C_0^{\text{NH}_3\text{-N}} e^{-\frac{k_{\text{N}}}{u}x} + \frac{1}{k_{\text{NH}_3\text{-N}}} \sum_{i=1}^9 a_i^{\text{NH}_3\text{-N}} \frac{1}{1+e^{-x}} \left(\sum_{j=1}^4 \omega_{i,j}^{\text{NH}_3\text{-N}} C_{t-j}^{\text{NH}_3\text{-N}} - \theta_j^{\text{NH}_3\text{-N}} \right) (1 - e^{-\frac{k_{\text{N}}}{u}x}) \quad (7)$$

2 不确定水质模型参数估计

式(5)显示,不确定水质模型中待估参数除污染物衰减系数 k 之外,还增加了网络参数 a_i 、 $\omega_{i,j}$ 、 θ_j ,并构成一个多参数寻优问题,其目标函数可取为污染物模型计算值与实测值的绝对百分比误差之和:

$$F(k, a_i, \omega_{i,j}, \theta_j) = \sum_i \sum_j \frac{|C_{i,j} - C_{i,j}^m|}{C_{i,j}} \quad (8)$$

式中: $C_{i,j}$ 、 $C_{i,j}^m$ 分别为污染物在第 j 断面、第 i 次监测的监测值和模型计算值。

由于模型参数的合理确定,将影响到内嵌的神

经网络函数与传统的水质模型能否形成一个有机整体,可否提高对各种不确定因素带来的水环境时空变化的预测水平,因此,选择适当的解法至关重要。根据遗传算法有利于在不确定的环境中寻找最优控制解^[12],选择改进适应度函数的遗传算法^[13-16],利用景观河水质监测数据,对不确定水质模型参数进行求解。

2.1 适应度函数

基于遗传算法在运行早期时,初始群体中可能存在特殊个体的适应度值超常现象,为了防止其统治整个群体并误导群体的发展方向而使算法收敛于局部最优解,即遗传算法的早熟现象,本文采用 Goldberg^[17]提出的线性变化适应度拉伸法构造适应度函数,使种群个体适应度成等差数列均匀分布,从而避免早熟和停滞问题,并加快函数的收敛速度,减少起始收敛代数。线性适应度函数如式(9)所示:

$$F^* = \frac{(c-1)\bar{F}}{F_{\text{max}} - \bar{F}} F + \frac{F_{\text{max}} - c\bar{F}}{F_{\text{max}} - \bar{F}} \bar{F} \quad (9)$$

式中: F^* 为目标函数式(8)通过线性变换得到的适应度函数; \bar{F} 为当前群体个体平均适应度; F_{max} 为当前群体中最佳个体的适应度; c 为常数,一般取 $1 \sim 2$ 。

2.2 模型参数求解

利用 MATLAB 编写遗传算法程序确定水质模型参数,模型计算所需参数设置:交叉概率 $p_m = 0.2$,变异概率 $p_c = 0.2$,各待估参数的取值范围:根据景观河监测数据试算,其 COD_{Cr} 、 $\text{NH}_3\text{-N}$ 衰减系数均小于 0.5,故令 $k_{\text{COD}}, k_{\text{NH}_3\text{-N}} \in (0, 0.5)$,同时,选择 $a_i, \omega_{ij}, \theta_j \in (-100, 100)$,先在此范围内搜索最佳适应度,若最佳适应度未能小于 0.5,需进一步放大 $a_i, \omega_{i,j}, \theta_j$ 取值范围,直至最佳适应度小于 0.5 为止。表 2、表 3 为 COD_{Cr} 、 $\text{NH}_3\text{-N}$ 不确定性水质模型式(6)、式(7)的参数估计结果。

表 2 COD_{Cr} 不确定性水质模型参数估计

编号	a_i	$\omega_{i,j}$				θ_j	$k / (\text{d}^{-1})$	
		$j = 1$	$j = 2$	$j = 3$	$j = 4$			
$i = 1$	5.125 0	2.462 9	5.248 0	-3.712 2	-10.327 3	$j = 1$	-72.623 8	0.189 8
$i = 2$	6.625 0	0.836 8	-3.165 4	1.258 5	1.993 2	$j = 2$	-18.483 4	
$i = 3$	1.625 0	-3.456 1	9.706 5	-0.305 4	-4.492 0	$j = 3$	36.577 8	
$i = 4$	47.500 0	7.927 3	9.661 7	-8.939 2	1.416 6	$j = 4$	6.302 9	
$i = 5$	0.500 0	9.666 4	9.759 4	2.057 5	-35.090 8			
$i = 6$	23.000 0	-4.783 4	1.986 3	1.243 3	6.403 4			
$i = 7$	21.000 0	-6.018 5	1.033 3	3.474 7	10.541 4			
$i = 8$	4.875 0	2.167 5	8.762 8	2.875 5	-18.912 6			
$i = 9$	29.875 0	-5.560 9	-9.002 4	9.406 1	-18.243 1			

表 3 NH₃-N 不确定性水质模型参数估计

编号	a_i	$\omega_{i,j}$				θ_j	$k / (d^{-1})$	
		$j = 1$	$j = 2$	$j = 3$	$j = 4$			
$i = 1$	46.875 0	-0.091 3	-7.740 8	-3.008 3	37.096 2	$j = 1$	-84.032 9	0.214 5
$i = 2$	18.375 0	2.007 9	-9.445 7	0.109 1	32.630 9	$j = 2$	85.636 7	
$i = 3$	5.500 0	-8.077 4	-9.677 7	8.696 8	34.765 5	$j = 3$	-5.351 5	
$i = 4$	4.500 0	-0.219 7	-5.973 0	-3.273 0	32.774 2	$j = 4$	79.389 6	
$i = 5$	38.250 0	2.127 1	5.556 8	-9.085 2	29.908 7			
$i = 6$	0.625 0	4.803 6	2.269 3	4.656 9	23.135 8			
$i = 7$	30.875 0	6.406 6	-1.917 2	-2.402 4	26.221 6			
$i = 8$	62.000 0	-0.712 0	-4.527 6	-3.273 8	29.645 5			
$i = 9$	6.000 0	-9.610 0	-0.651 0	-8.046 1	27.446 3			

3 模型预测结果分析比较

3.1 模型预测结果

利用上述建立的不确定性水质模型,对景观河 7 月下半月 4 个断面的 COD_{Cr}、NH₃-N 值进行预测。预测时网络的输入数据随预测日的变化而向前推移,即总是选取预测日前 4 d 的污染物输入断面的监测数据。此外,为比较不确定性水质模型预测效果,采用相同监测数据,建立了确定性水质模型(公式略),并分别对各断面的 COD_{Cr}、NH₃-N 进行预测,图 2、图 3 为河流 2-2 断面的预测结果。

表 4 为确定性水质模型与不确定性水质模型的部分计算参数和预测平均误差的对比。

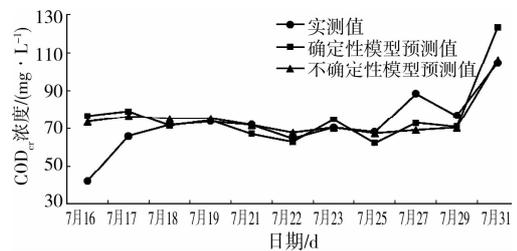


图 2 2-2 断面的 COD_{Cr} 监测值与预测值

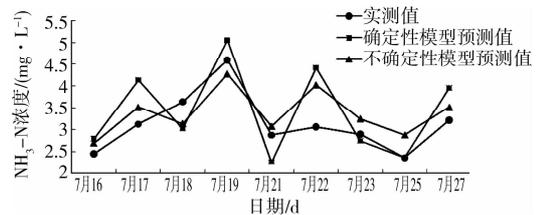


图 3 2-2 断面的 NH₃-N 监测值与预测值

表 4 两种水质模型部分计算参数和预测平均误差对比

水质指标	模型类型	目标值 (适应度)	衰减系数 /d ⁻¹	预测平均误差/%			
				2-2 断面	3-3 断面	4-4 断面	5-5 断面
COD _{Cr}	确定性	2.542 6	0.194 3	15.64	21.65	18.99	15.56
	不确定性	0.485 8	0.189 8	12.48	18.98	16.49	15.32
NH ₃ -N	确定性	3.447 4	0.233 4	18.55	20.94	23.23	37.00
	不确定性	0.246 7	0.214 5	13.81	18.22	17.14	21.10

分析表 4 可知:

1) 不确定性模型给出的污染物衰减系数均小于确定性模型的相应系数,说明增加了神经网络观测项,对污染物衰减规律有所调整。

2) 两类模型的目标值(适应度函数值)不同,说明不确定性水质模型拟合的精度更高,更能适应变化的景观河水体环境。

3) 从预测效果分析,不确定性水质模型的预测平均水平高于确定性水质模型,平均预测误差减

小幅度达 0.24%~15.90%,预测精度基本在 80% 以上,尤其是在靠近污染物输入断面的 2-2 断面,预测精度明显高于其它下游断面,说明增加的神经网络函数对排水水质的波动是比较敏感的,而其他下游各断面虽然预测精度也有提高,但因距污染物输入断面距离逐渐加大,其水质变化将会受到更多的其他随机因素的扰动。

4) 当然,不确定性水质模型的预测结果中仍有部分预测误差大于确定性水质模型的预测,其中 COD_{Cr}

占41%, $\text{NH}_3\text{-N}$ 占33.33%,但其误差增幅不大;而且所有预测结果中对COD_{Cr}的预测误差超过30%的仅为11.36%,对 $\text{NH}_3\text{-N}$ 的预测误差超过30%的也只有16.66%,基本满足日常水质管理的需要。

4 结 语

针对排入景观河水体的排水水质的不确定性,本文建立内嵌神经网络的一维不确定性水质模型,利用改进适应度函数的遗传算法,获得模型参数的最优解。经模型计算参数及预测结果的比较分析,不确定性水质模型拟合的精度更高,更能适应变化的景观河水体环境,其预测平均水平普遍高于确定性水质模型,且靠近污染源的监测断面,其不确定性水质模型预测优势更加明显。

参考文献:

- [1] COX B A. A review of currently available in-stream water-quality models and their applicability for simulating dissolved oxygen in lowland rivers[J]. *Science of the Total Environment*, 2003, 314: 335-377.
- [2] FANG X B, ZHANG J X, CHEN Y X, et al. QUAL2K model used in the water quality assessment of qiantang river, China[J]. *Water Environment Research*, 2008, 80(11): 2125-2133.
- [3] 许劲, 龙腾锐. 不确定性河流水质模型的应用及进展[J]. *中国给水排水*, 2007, 23(16): 4-8.
XU JIN, LONG TENG-RUI. Application and Progress of uncertainty river water quality model [J]. *China Water & Wastewater*, 2007, 23(16): 4-8.
- [4] KARMAKARA S, MUJUMDAR P P. An inexact optimization approach for river water-quality management [J]. *Journal of Environmental Management*, 2006, 81(3): 233-248.
- [5] HUANG W R, FOO S. Neural network modeling of salinity variation in Apalachicola river [J]. *Water Research*, 2002, 36(1): 356-362.
- [6] MAIER H R, CHOW C W. Use of artificial neural networks for predicting optimal alum doses and treated water quality parameters[J]. *Environmental Modelling & Software*, 2004, 19(5): 485-494.
- [7] MOZEJKO J, GNIOT R. Application of neural networks for the prediction of total phosphorus concentrations in surface waters [J]. *Polish Journal of Environmental Studies*, 2008, 17(3): 363-368.
- [8] CHEN D J, LU J, SHEN Y N. Artificial neural network modelling of concentrations of nitrogen, phosphorus and dissolved oxygen in a non-point source polluted river in Zhejiang province, southeast China[J]. *Hydrological Processes*, 2009, 24: 290-299.
- [9] 翟宜峰, 李鸿雁, 刘寒冰, 等. 用遗传算法优化神经网络初始权重的方法[J]. *吉林大学学报: 工学版*, 2003, 33(2): 45-50.
ZHAI YI-FENG, LI HONG-YAN, LIU HAN-BING, et al. Method for optimizing initial weights of ANNs by GAs[J]. *Journal of Jinlin University: Engineering and Technology Edition*, 2003, 33(2): 45-50.
- [10] 程声通, 陈毓龄. 环境系统分析[M]. 北京: 高等教育出版社, 1990.
- [11] 李如忠. 河流水环境系统不确定性问题研究[D]. 南京: 河海大学环境科学与工程学院, 2004.
- [12] CHEN H W, CHANG N B. Water pollution control in the river basin by fuzzy genetic algorithm-based multiobjective programming modeling[J]. *Water Science and Technology*, 1998, 37(8): 55-68.
- [13] 刘臣辉, 刘超, 徐香, 等. 基于改进遗传算法率定水质参数的研究[J]. *环境保护科学*, 2008, 34(3): 47-49.
LIU CHEN-HUI, LIU CHAO, XU XIANG, et al. Parameter estimation of water quality model based on improved genetic algorithm [J]. *Environmental Protection Science*, 2008, 34(3): 47-49.
- [14] 闫欣荣, 史忠科. 反演-遗传算法在河流水质 BOD-DO 耦合模型参数识别中的应用[J]. *水资源与水工程学报*, 2007, 18(4): 41-43.
YAN XIN-RONG, SHI ZHONG-KE. Application of Inversion-Genetic algorithms to parameters identification for the BOD-DO water quality coupling model[J]. *Journal of Water Resources and Water Engineering*, 2007, 18(4): 41-43.
- [15] 梁婕, 曾光明, 郭生练, 等. 变尺度混沌-遗传算法在复杂河流水质模型参数优化中的应用[J]. *环境科学学报*, 2007, 27(2): 342-347.
LIANG JIE, ZENG GUANG-MING, GUO SHENG-LIAN, et al. Application of mutative scale chaos genetic algorithm (MSCGA) to parameters estimation for river water quality model [J]. *Acta Scientiae Circumstantiae*, 2007, 27(2): 342-347.
- [16] 王建平, 程声通, 贾海峰. 基于 MCMC 法的水质模型参数不确定性研究[J]. *环境科学*, 2006, 27(1): 24-30.
WANG JIAN-PING, CHEN SHENG-TONG, JIA HAI-FENG. Markov Chain monte carlo scheme for parameter uncertainty analysis in water quality model [J]. *Environmental Science*, 2006, 27(1): 24-30.
- [17] GOLDBERG D E. Genetic algorithms in search, optimization and machine learning[M]. United States: Addison Wesley Publishing, 1989.