

基于改进Bayesian-MCMC的突发水污染事件 预测模型参数率定方法

杨海东, 刘碧玉, 黄建华[†]

(福州大学 经济与管理学院, 福州 350116)

摘要: 预测模型是科学制定应急处置措施的基础. 为快速准确地构建突发水污染事件预测模型, 将预测模型参数的率定问题视为贝叶斯估计问题, 并根据有限差分方法和贝叶斯推理得到参数的后验概率密度函数, 再通过改进的Metropolis-Hastings抽样方法得到较为合理的参数值. 以发生在某明渠段的突发水污染事件为例, 分析讨论等容量控制非均匀流和非等容量控制非均匀流两种情景下不同观测噪声对参数率定值的影响, 并与由贝叶斯-马尔科夫链蒙特卡罗方法得到的参数值和真实值进行对比. 结果表明: 改进Bayesian-MCMC方法在计算精度、适用性和抗噪声等方面优于贝叶斯-马尔科夫链蒙特卡罗方法, 能较好地率定模型参数, 并为构建突发水污染事件预测模型提供了新思路.

关键词: 参数率定; 预测模型; 贝叶斯; 马尔科夫链蒙特卡罗; 突发水污染事件

中图分类号: X522 **文献标志码:** A

Forecast model parameters calibration method for sudden water pollution accidents based on improved Bayesian-Markov chain Monte Carlo

YANG Hai-dong, LIU Bi-yu, HUANG Jian-hua[†]

(School of Economics and Management, Fuzhou University, Fuzhou 350116, China)

Abstract: The prediction models are the basis for scientific formulation of emergency disposal and rescue measures. In order to quickly and accurately construct the forecasting model of sudden water pollution accidents, this paper regards the problem as the Bayesian estimation problem and obtains the posterior probability density function of the model parameters according to the finite difference method and Bayesian inference. Then, by using the improved Metropolis-Hastings sampling method, more reasonable prediction model parameters are obtained. Taking the sudden water pollution event in a certain open channel as an example, the effects of different observation noises on the parameters calibration results are discussed under the two scenarios of non-uniform flow control with non-uniform flow and non-equal capacity control, and compared with the parameter and true value of the Bayesian-Markov chain Monte Carlo method. The results show that, the improved Bayesian-Markov chain Monte Carlo method can give a better parameter value, more application and has a good anti-noise, which can provide a new approach to build the forecast model of sudden water pollution accidents.

Keywords: parameters calibration; prediction model; Bayesian; Markov chain Monte Carlo; sudden water pollution accidents

0 引言

随着“一带一路”战略的实施和新型城镇化建设速度的加快,我国危化品的生产、贮存、运输和使用也随之增加,极易发生突发性泄漏事故. 这类事故不仅破坏当地生态环境,还对人们身体健康构成威胁,

甚至影响社会稳定^[1]. 如2015年甘肃省兰州市自来水苯超标、2014年江苏省靖江市饮用水异常和2013年山西省长治市苯胺泄漏等突发水污染事件. 为有效减轻和控制此类事件带来的影响,决策者应在事发第一时间掌握水体中污染物迁移扩散规律,并据此

收稿日期: 2017-03-04; 修回日期: 2017-06-26.

基金项目: 国家社科基金一般项目(17BGL179).

责任编辑: 刘民.

作者简介: 杨海东(1981—),男,副教授,博士,从事复杂系统理论与方法,应急管理,应急物流等研究; 黄建华(1972—),男,副教授,博士,从事复杂系统与复杂网络、物流与供应链优化等研究.

[†]通讯作者. E-mail: everam@fzu.edu.cn

发布准确的预警级别和采取针对性的补救措施^[2]. 其中, 预测模型参数的合理取值是准确反映水体中污染物迁移扩散规律的关键^[3]. 因此, 能否依据应急信息迅速准确地率定突发水污染事件预测模型中的参数, 对于应急处置与救援策略的制定十分重要.

目前, 参数率定方法主要有试错法、理论公式法、经验公式法和示踪试验法等^[4-6]. 其中, 试错法具有经验性强、效率低和精度差等特点, 已远远不能满足工程实践的需要; 理论公式法是基于 Fischer 张量形式理论模型的一类方法, 但在介质速度规律与公式等方面的研究没有形成一个统一的结论; 经验公式法虽然具有方便快捷、工作量小等特点, 但其适用性和准确性仍有待商榷; 示踪试验法是指在人财物许可条件下通过开展示踪试验率定相关参数的一类方法, 该方法能比较准确地表征污染物的输移规律^[6-7].

本文所提出的方法属于示踪试验法. 利用示踪试验法研究率定参数问题属于反问题研究范畴, 已先后发展了矩量法、拟合法、演算法、优化法等^[7-8]. 矩量法是指基于污染物时间浓度分布的方差率定参数的一类方法, 但基于该方法得到的参数具有较大的误差^[6]. 为此, 众多学者通过分析污染物迁移扩散模型的解析解, 提出了诸如相关系数法^[9]、抛物线方程^[10]和非线性最小二乘法^[11]等拟合逼近方法, 但这些方法的理论假设不够严谨, 且易受初始值的影响.

由于反问题可转化为系统优化问题, 基于优化思想的率定方法随着计算机技术的发展而得到了广泛的应用, 如单纯形法^[12]、蚁群算法^[13]、神经网络^[14-15]、遗传算法^[16-17]和改进微分进化算法^[18]等. 但基于优化思想的率定方法受制于一些非常严格的限制条件, 如限于研究矩形顺直且水流近似为均匀流的渠道或地貌平坦且空气流均匀的区域, 当未知参数越多时, 其响应曲面的非线性程度越高, 参数无法率定的程度越大^[19].

此外, 观测误差或模型计算误差还可能会使预测结果出现严重偏差. 鉴于此, 概率方法被引入研究参数率定问题. 其中, 贝叶斯方法是概率方法中运用最为广泛、理论最为成熟的一种方法. 与优化算法不同的是, 贝叶斯方法将反问题视为一个扩展随机空间上的适定问题, 并认为后验概率密度能表示问题的“完全”解^[20]. 换言之, 基于贝叶斯的参数率定方法是通过计算待求参数的后验概率分布及其统计特征值获取待求参数点估计的一种方法, 该方法能很好解决由观测数据噪声带来的非唯一解的问题^[21-23]. 如朱嵩等^[24]利用贝叶斯推理建立了二维含源对流扩散方程

参数率定的数学模型; Zhang 等^[25]为处理观测噪声和模型的不确定性, 设计了一种耦合贝叶斯和区间分析的反演方法. 然而, 当前基于贝叶斯方法的参数率定问题研究大多围绕预测模型解析解或数值解等方面, 很难满足应急处置与救援的需求.

本文以发生在某一段长距离明渠输水工程的突发污染事件为研究对象, 在一维污染物迁移扩散修正模型的基础上采用有限差分方法 (FDM) 求出突发污染事件正向数值预测模型; 然后基于贝叶斯推理采用改进的马尔科夫链蒙特卡罗 (MCMC) 方法率定相关参数; 之后通过对不同情景不同观测噪声下该明渠段预测模型参数率定值进行分析; 最后与贝叶斯-马尔科夫链蒙特卡罗 (Bayesian-MCMC) 方法得到的结果进行对比分析.

1 预测模型

水体中污染物的迁移规律通常遵循流体运动的基本物理规律, 并满足根据质量守恒推导出来的迁移转换基本方程. 对主要考虑纵向污染物浓度变化的河渠而言, 可以用下式描述事件演化进程^[26]:

$$\frac{\partial(AC)}{\partial t} + \frac{\partial(QC)}{\partial x} = \frac{\partial}{\partial x} \left(AE_x \frac{\partial C}{\partial x} \right) + W_c + S_1 - AKC. \quad (1)$$

其中: A 为断面面积, C 为污染物质量浓度, Q 为断面 A 纵向流量, t 为时间, x 为纵向空间坐标, E_x 为纵向离散系数, W_c 为外源输入项, S_1 为与污染物质量浓度 C 无关的内部源汇项, K 为污染物自身的衰减速度常数.

根据输入输出信息准则, 系统预测模型的一般形式^[27]为

$$y = f(x) + e. \quad (2)$$

其中: x 为系统输入; y 为系统输出; e 为噪声; f 为一个抽象函数, 代表从模型空间到参数空间的映射, 可以用解析解或数值解表示. 在已知污染源特性 (如污染源的位置、排放强度与时间, 以及初、边界条件) 时, 基于式 (2), 式 (1) 的解可表示为

$$C(x, t) = C(x, t|\theta) + e. \quad (3)$$

其中: x 为纵向空间坐标; t 为时间; θ 为控制浓度分布的参数集, 如污染物扩散系数 E_x 、降解系数 K 等.

由于河渠的流速、水位和水质参数均是沿程变化的, 根据式 (1) 无法得到河渠中污染物质量浓度的解析解. 为此, 本文对河渠研究区域进行空间离散, 同时将求解变量 (物质浓度) 定义在各断面上^[28], 采用隐式迎风差分格式将式 (1) 中的每一项进行离散, 得

到

$$\begin{cases} \frac{\partial(AC)}{\partial t} = \frac{(AC)_i^{k+1} - (AC)_i^k}{\Delta t}, \\ \frac{\partial(QC)}{\partial x} = \frac{(QC)_i^{k+1} - (QC)_{i-1}^{k+1}}{\Delta x_{i-1}}, \\ \frac{\partial}{\partial x} \left(AE_x \frac{\partial C}{\partial x} \right) = \\ \left[(AE_x)_i^{k+1} \frac{C_{i+1}^{k+1} - C_i^{k+1}}{(\Delta x_i + \Delta x_{i-1})/2} - \right. \\ \left. (AE_x)_{i-1}^{k+1} \frac{C_i^{k+1} - C_{i-1}^{k+1}}{(\Delta x_{i-1} - \Delta x_{i-2})/2} \right] \frac{1}{\Delta x_{i-1}}, \\ S_c + W_c = (\overline{S}_1 + \overline{W}_c)_{i-1}^{k+1} - (A\overline{K}C)_i^{k+1}. \end{cases} \quad (4)$$

其中: C_i^{k+1} 为第 i 段下游断面第 k 个时段末污染物质量浓度值; C_i^k 为第 i 段下游断面第 k 个时段初污染物质量浓度值; $(\overline{S}_1)_i$ 、 $(\overline{W}_c)_i$ 、 $(\overline{K})_i$ 分别为第 i 段的内部源汇、外源输入项和反应速率; $(AE_x)_i$ 式中时间项采用前差分。

将式(4)代入(1)得

$$\begin{aligned} \frac{(AC)_i^{k+1} - (AC)_i^k}{\Delta t} + \frac{(QC)_i^{k+1} - (QC)_{i-1}^{k+1}}{\Delta x_{i-1}} = \\ \left[(AE_x)_i^{k+1} \frac{C_{i+1}^{k+1} - C_i^{k+1}}{(\Delta x_i + \Delta x_{i-1})/2} - \right. \\ \left. (AE_x)_{i-1}^{k+1} \frac{C_i^{k+1} - C_{i-1}^{k+1}}{(\Delta x_{i-1} - \Delta x_{i-2})/2} \right] \frac{1}{\Delta x_{i-1}} + \\ (\overline{S}_1 + \overline{W}_c)_{i-1}^{k+1} - (A\overline{K}C)_i^{k+1}. \end{aligned} \quad (5)$$

令

$$\begin{aligned} a_i &= 2 \left[\frac{(AE_x)_{i-1}^{k+1}}{\Delta x_{i-1} + \Delta x_{i-2}} + Q_{i-1}^{k+1} \right], \\ C_{i-1} &= C_{i-1}^{k+1}, C_i = \begin{bmatrix} C_i^{k+1} \\ C_i^k \end{bmatrix}, \\ c_i &= -2 \frac{(AE_x)_i^{k+1}}{\Delta x_i + \Delta x_{i-1}}, \\ b_i &= \frac{A_i^{k+1} \Delta x_{i-1}}{\Delta t} + Q_i^{k+1} + \\ & 2 \left[\frac{(AE_x)_i^{k+1}}{\Delta x_i + \Delta x_{i-1}} + \frac{(AE_x)_{i-1}^{k+1}}{\Delta x_{i-1} + \Delta x_{i-2}} \right] + \\ & \Delta x_{i-1} (A\overline{K})_i^{k+1} - \frac{A_i^{k+1} \Delta x_{i-1}}{\Delta t}, \\ C_{i+1} &= C_{i+1}^{k+1}, d_i = \Delta x_i (\overline{S}_1 + \overline{W}_c)_{i-1}^{k+1}. \end{aligned}$$

经整理得

$$a_i C_{i-1} + b_i C_i + c_i C_{i+1} = d_i, \quad i = 1, 2, \dots, N. \quad (6)$$

其中: a_i 、 b_i 、 c_i 为系数项; d_i 为常数项; C_i 为第 i 个渠道段末的物质浓度值。式(6)是由 N 个方程组成的线性隐式差分方程组, 结合上下游边界条件可进行数值

求解。

2 参数率定方法

基于贝叶斯推理的参数率定方法是利用未知参数的不确定性分布信息, 在一定程度上避免因“最优”参数失真而带来的决策风险的一种方法^[29]。根据贝叶斯理论^[30-31], 贝叶斯推理可以表述如下:

$$p(\theta|y) = \frac{p(\theta)p(y|\theta)}{p(y)} \propto p(\theta)p(y|\theta). \quad (7)$$

其中: θ 为待率定参数, y 为观测数据, $p(y|\theta)$ 为似然函数, $p(\theta|y)$ 为待率定参数的后验概率密度函数, $p(\theta)$ 为待率定参数的联合先验概率密度函数。

实际中, $p(y)$ 不便于用分析的方法计算, 且不依赖于 θ , 仅起到一个正则化因子的作用。若假设 $\theta = (\theta_1, \theta_2, \dots, \theta_m)$ 服从均匀分布且相互独立, 则未知参数的联合先验概率密度函数可以写成

$$p(\theta) = \prod_{i=1}^m p(\theta_i) = \begin{cases} \prod_{i=1}^m \frac{1}{b_i - a_i}, & \theta_i \in [a_i, b_i]; \\ 0, & \text{else.} \end{cases} \quad (8)$$

其中 m 为模型参数的个数。

将观测误差和预测误差分别表示为

$$\varepsilon_i = f_i - y_i, \quad e_i = h_i(\theta) - f_i(\theta),$$

其中 $y = (y_1, y_2, \dots, y_M)$ 为 M 个观测点的观测值, $f = (f_1, f_2, \dots, f_M)$ 为 M 个观测点的预测值, $h = (h_1, h_2, \dots, h_M)$ 为 M 个理论值。假定观测误差和预测误差均服从正态分布, 似然函数分别为

$$\begin{aligned} p(y_i|h_i, \theta) &= \frac{1}{(2\sigma_{d,i})^{1/2}} \exp \left[-\frac{(y_i - h_i)^2}{2\sigma_{d,i}^2} \right], \\ p(h_i|f_i, \theta) &= \frac{1}{(2\sigma_{f,i})^{1/2}} \exp \left[-\frac{(f_i - h_i)^2}{2\sigma_{f,i}^2} \right]. \end{aligned} \quad (9)$$

若观测误差和预测误差不相关且每个观测点相互独立, 则

$$\begin{aligned} p(y|\theta) &= \prod_{i=1}^M p(y_i|\theta) \propto \\ & \frac{1}{2^M \prod_{i=1}^M \sqrt{\sigma_{f,i} \sigma_{d,i}}} \exp \left[-\sum_{i=1}^M \frac{(C_i(x, t) - C_i(x, t|\theta))^2}{2(\sigma_{f,i}^2 + \sigma_{d,i}^2)} \right]. \end{aligned} \quad (10)$$

因此, 参数的后验概率函数可以表示为

$$p(\theta|y) = p(y|\theta_1, \theta_2, \dots, \theta_m)p(\theta_1, \theta_2, \dots, \theta_m). \quad (11)$$

由于随机误差的似然函数表示模型预测值与实测数据的拟合程度, 参数率定问题可以转化为求参数

的后验概率密度函数

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{2^M \prod_{i=1}^M \sqrt{\sigma_{f,i} \sigma_{d,i}}} \exp \left[- \sum_{i=1}^M \frac{(C_i(x, t) - C_i(x, t|\theta))^2}{2(\sigma_{f,i}^2 + \sigma_{d,i}^2)} \right]. \quad (12)$$

贝叶斯公式表达形式简单,但在具体确定每一项时十分困难,比如似然函数的准确属性表达,归一化常数积分区间的选择等,通常采用MCMC方法进行抽样计算.其中MCMC是构造合适的马尔科夫链进行抽样而使用蒙特卡罗方法进行积分计算的一种方法;Metropolis-Hastings(M-H)方法是一种生成-拒绝样本形式的抽样方法,它能从任意分布中随机抽取样本值,并以某一标准决定接受还是拒绝该样本值,避免了计算 $p(y)$ 的问题^[32].因此,本文为提高M-H方法的抽样效率,在M-H抽样过程中加入一次人为筛选,其具体求解步骤如下:

1) 将研究区域空间离散为 N 段,每段水流断面形状变化不大.

2) 根据变量个数 N 及其部分先验信息,确定未知参数的样本空间和先验概率密度函数 $p(\theta)$.

3) 在其先验范围内随机生成 N 个初始值 $X^i = \{x_i(1), x_i(2), \dots, x_i(N)\}$,并设定 $i = 1$.

4) 设定Proposal分布 $U(x_i(s) - \text{step}, x_i(s) + \text{step})$,并生成 $x'(s)$,其中 U 表示均匀分布,step为随机游走的步长.

5) 分别计算出 $x_i(s)$ 和 $x'(s)$ 对应的污染物浓度值 Y 和 Y_0 ,即

$$B = \sum |Y - Y_0|. \quad (13)$$

6) 如果 $B > 0.6$,则接受该测试参数并设定为当前模型参数,即 $x_i(s) = x'(s)$;否则不接受该测试参数, $x_i(s) = x_i(s)$.

7) 利用分布 $U(x_i(s) - \text{step}, x_i(s) + \text{step})$ 生成 $X^* = \{x^*(1), x^*(2), \dots, x^*(N)\}$.

8) 计算能够反映模型参数和观测数据之间关系的似然函数 $p(y|\theta)$.

9) 计算未知参数的后验概率密度 $p(\theta|y)$.

10) 计算Markov链从 $X^{(i)}$ 位置移动到 $X^{(*)}$ 的接受概率

$$A(X^{(i)}, X^{(*)}) = \min \left\{ 1, \frac{p(X^{(*)})}{p(X^{(i)})} \right\}. \quad (14)$$

11) 产生一个 $0 \sim 1$ 间均匀分布的随机数 R ,如果 $R < A(X^{(i)}, X^{(*)})$,则接受该测试参数并设定为当前模型参数,即 $X^{(i+1)} = X^{(*)}$;否则,不接受该测试参数, $X^{(i+1)} = X^{(i)}$.

12) 重复步骤1)~步骤11),直至达到预定迭代次数.

3 案例分析

3.1 研究区域描述

为简便计算,本文选择一条长为3 km包含5个调节闸门的某明渠段,在不考虑污染物的任何生物化学条件下,研究率定等容量控制非均匀流和非等容量控制非均匀流情景下各渠段纵向离散系数集合^[13],如图1所示.若该明渠段上游起始端($x = 0$)径流来水污染物浓度为 1.0 mg/L ,下游为自由出流,即浓度梯度为0,渠段流场沿程分布为 $u = 5 + 0.001x$,则根据该明渠段的几何特征,可将其离散成6个微渠段,对应 $(E_x)_i (i = 1, 2, \dots, 6)$ 的真值分别为50, 70, 90, 110, 130和140,单位为 m^2/s .

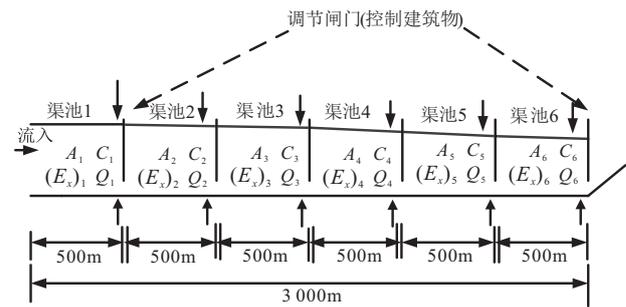


图1 某明渠示意图

3.2 等容量控制非均匀流

若该明渠流量恒为 $10 \text{ m}^3/\text{s}$,则首先根据FDM方法和 $(E_x)_i (i = 1, 2, \dots, 6)$ 的真值得到各渠池断面污染物浓度序列数据、待求参数 E_x 的先验分布 $p(E_x)$ 和似然函数 $p(E_x|y)$;然后按改进Bayesian-MCMC方法的思路迭代20000次,得到在误差水平 $\sigma = 0.01$ 时各渠段参数 E_x 的直方图和迭代曲线,如图2和图3所示.

由图2和图3可看出,所有待求参数 $(E_x)_i (i = 1, 2, \dots, 6)$ 均在真值附近的取值概率最大.其中, $(E_x)_i (i = 1, 2, 3)$ 和 $(E_x)_i (i = 4, 5)$ 大约分别迭代300次和600次后,产生的Markov链值开始接近于真值,迭代次数比采用Bayesian-MCMC方法减少了15%.表1和表2为分别采用改进Bayesian-MCMC方法和Bayesian-MCMC方法得到不同观测误差下 $(E_x)_i (i = 1, 2, \dots, 6)$ 的率定值.

由表1和表2可看出,在 $\sigma = 0.01, 0.05, 0.1$ 时,基于改进 Bayesian-MCMC 得到 $(E_x)_i (i = 1, 2, \dots, 6)$ 的平均相对误差分别为 0.47%, 2.34% 和 3.70%, 分别比采用 Bayesian-MCMC 方法低 0.27%, 1.03% 和 1.11%。

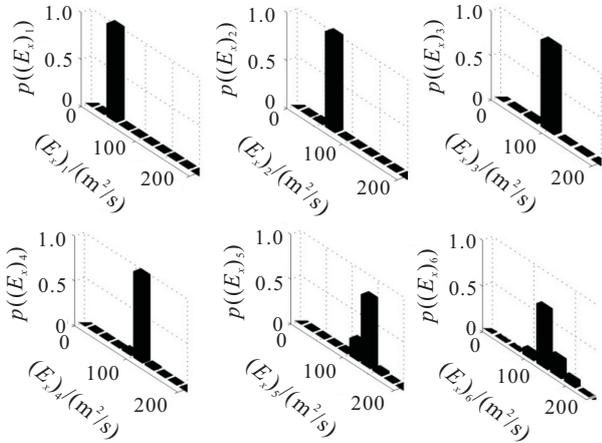


图2 等容量控制非均匀流情景下参数 (E_x) 的直方图

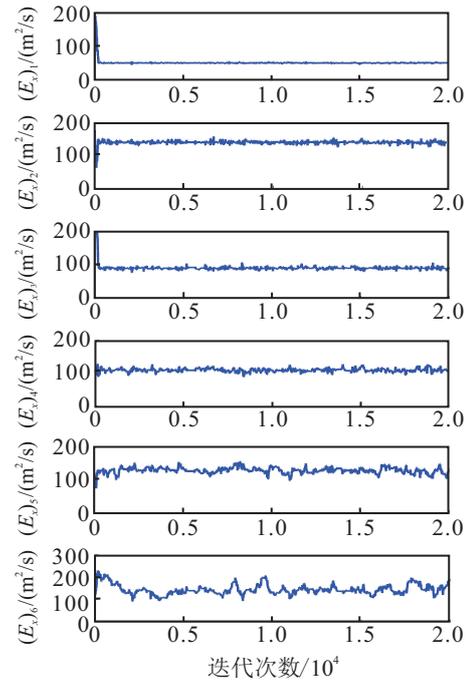


图3 等容量控制非均匀流情景下参数 (E_x) 的迭代曲线

表1 不同观测误差水平下基于改进 Bayesian-MCMC 方法的 $(E_x)_i (i = 1, 2, \dots, 6)$ 率定值

断面编号	$\sigma = 0.01$		$\sigma = 0.05$		$\sigma = 0.1$	
	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %
1	50.0533	0.11	50.1814	0.36	50.2094	0.42
2	70.161	0.23	70.4462	0.64	71.7018	2.43
3	89.7716	0.25	91.4276	1.59	91.4491	1.61
4	109.9792	0.02	111.2201	1.11	114.2726	3.88
5	128.9898	0.78	135.4634	4.20	135.0298	3.87
6	137.9742	1.45	148.6219	6.16	153.9622	9.97

表2 不同观测误差水平下基于 Bayesian-MCMC 方法的 $(E_x)_i (i = 1, 2, \dots, 6)$ 率定值

断面编号	$\sigma = 0.01$		$\sigma = 0.05$		$\sigma = 0.1$	
	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %
1	49.7844	0.43	50.0952	0.19	50.1963	0.39
2	70.2705	0.39	69.397	0.86	69.1707	1.18
3	90.1206	0.13	92.6923	2.99	95.1213	5.69
4	111.3893	1.26	112.4978	2.27	105.295	4.28
5	130.3584	0.28	138.9934	6.92	122.0383	6.12
6	142.7324	1.95	149.8147	7.01	155.6278	11.16

3.3 非等容量控制非均匀流

若该段明渠流量可以通过闸控建筑物控制,且是时间的函数,则 $Q(t) = 10 + 0.001t$. 同理,首先根据 FDM 方法和 $(E_x)_i (i = 1, 2, \dots, 6)$ 的真值得到各渠池断面污染物浓度序列数据、待求参数 E_x 的先验分布 $p(E_x)$ 和似然函数 $p(E_x|y)$; 然后按改进 Bayesian-MCMC 方法的思路迭代 20 000 次,得到在误差水平 $\sigma = 0.01$ 时各渠段参数 $(E_x)_i (i = 1, 2, \dots, 6)$ 的直方图和迭代曲线,如图4和图5所示。

表3和表4分别采用改进 Bayesian-MCMC 方法和 Bayesian-MCMC 方法得到不同观测误差下 $(E_x)_i$

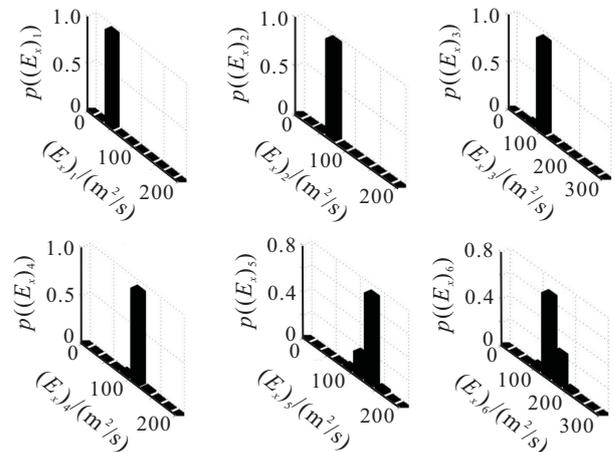


图4 非等容量控制非均匀流下参数 (E_x) 的直方图

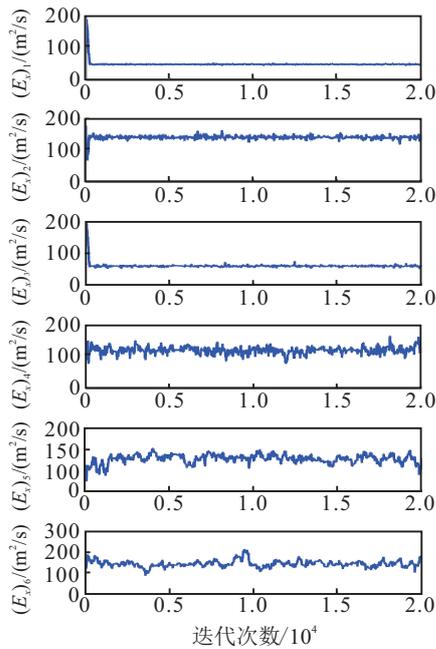


图5 非等容量控制非均匀流情景下参数 $(E_x)_i$ 的迭代曲线

$(i = 1, 2, \dots, 6)$ 的率定值.

由图4和图5可看出,所有待求参数 $(E_x)_i (i = 1, 2, \dots, 6)$ 均在真值附近的取值概率最大. 其中 $P((E_x)_i) \approx 1, i = 1, 2, 3$. $(E_x)_i (i = 1, 2, 3)$ 和 $(E_x)_i (i = 4, 5)$ 大约分别迭代300次和600次后,即迭代次数比采用 Bayesian-MCMC 方法减少了15%,产生的 Markov 链值开始接近于真值.

由表3和表4可看出,在观测误差 $\sigma = 0.01, 0.05, 0.1$ 时,采用改进 Bayesian-MCMC 方法得到 $(E_x)_i (i = 1, 2, \dots, 6)$ 的平均相对误差分别为0.42%、1.54%和4.05%,分别比采用 Bayesian-MCMC 方法低0.45%、1.41%和1.4%.

3.4 案例结果分析

综合图2~图5和表1~表4可得,相对 Bayesian-MCMC 方法而言,基于改进 Bayesian-MCMC 方法具有以下特点:

表3 不同观测误差下基于改进 Bayesian-MCMC 方法的 $(E_x)_i (i = 1, 2, \dots, 6)$ 率定值

断面编号	$\sigma = 0.01$		$\sigma = 0.05$		$\sigma = 0.1$	
	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %
1	50.0323	0.06	50.1627	0.33	50.2830	0.57
2	70.0028	0.00	70.5350	0.76	72.2023	3.15
3	89.9656	0.04	90.0965	0.11	92.2466	2.50
4	110.4561	0.41	113.5114	3.19	117.6315	6.94
5	129.1287	0.67	134.3805	3.37	128.3266	1.29
6	141.8405	1.31	142.0582	1.47	153.7869	9.85

表4 不同观测误差下基于 Bayesian-MCMC 方法的 $(E_x)_i (i = 1, 2, \dots, 6)$ 率定值

断面编号	$\sigma = 0.01$		$\sigma = 0.05$		$\sigma = 0.1$	
	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %	率定值 / (m ² /s)	相对误差 / %
1	50.014	0.03	50.148	0.30	49.967	0.10
2	70.046	0.06	69.269	0.99	69.390	0.67
3	90.142	0.12	93.348	3.77	99.589	10.85
4	109.934	0.09	110.775	0.73	107.939	2.27
5	131.426	1.09	135.627	4.45	125.762	3.21
6	134.696	3.79	150.404	7.43	161.836	15.60

1) 高计算精度. 采用基于改进 Bayesian-MCMC 方法得到 $(E_x)_i (i = 1, 2, \dots, 6)$ 的平均相对误差均小于10%,且总平均相对误差分别比采用 Bayesian-MCMC 方法低0.8%和1.01%,即在相同情景下采用改进 Bayesian-MCMC 方法得到参数值的精度比 Bayesian-MCMC 方法高.

2) 强适用性. 当观测误差 σ 从0.01增加到0.1时,无论等容量还是非等容量控制的非均匀流情景,采用基于改进 Bayesian-MCMC 方法得到 $(E_x)_i (i = 1, 2,$

$\dots, 6)$ 的平均相对误差均小于10%,即改进 Bayesian-MCMC 方法能适应观测误差较大和不同情景.

3) 强抗噪声能力. 当观测误差 σ 从0.01增加到0.1时,采用改进 Bayesian-MCMC 方法所得 $(E_x)_i (i = 1, 2, \dots, 6)$ 的平均相对误差变化量比 Bayesian-MCMC 低0.12%,即改进 Bayesian-MCMC 方法具有更强的抗噪能力.

综上所述,基于改进 Bayesian-MCMC 方法能得到更为合理的参数值,它不仅提升了参数率定的精

度,而且扩大了适用范围,增强了抗噪声能力。

4 结 论

突发水污染事件由于其发生概率的不确定性、污染物种类的多样性及事发区域的复杂性等特点,使得率定此类事件预测模型参数问题具有很强的不适应性。为此,本文将突发水污染事件预测模型中的参数视为随机变量,并充分考虑参数先验信息、观测噪声以及模型误差的影响,在 FDM 和 Bayesian-MCMC 的基础上,设计了一种能较好地解决参数率定过程中的不适应性方法。同时,将该方法同 Bayesian-MCMC 方法进行了比较,结果表明所设计的方法具有更高的计算精度、更强的适用性和抗噪声能力。然而,突发污染事件中预测模型往往无明确的数学表达式,并极有可能存在局部极值点,导致该方法的计算效率低,所以提高参数率定方法的计算速度是进一步的研究方向。

参考文献(References)

- [1] 余乐安,李玲,武佳倩,等. 基于系统动力学的危化品水污染突发事件中网络舆情危机应急策略研究[J]. 系统工程理论与实践, 2015, 35(10): 2687-2697.
(Yu L A, Li L, Wu J Q, et al. Emergency policy exploration for network public opinion crisis in water pollution accident by hazardous chemicals leakage based on systematic dynamics[J]. Systems Engineering — Theory & Practice, 2015, 35(10): 2687-2697.)
- [2] 杨海东,肖宜,王卓民,等. 突发性水污染事件溯源方法[J]. 水科学进展, 2014, 25(1): 122-129.
(Yang H D, Xiao Y, Wang Z M, et al. On source identification method for sudden water pollution accident [J]. Advance in Water Science, 2014, 25(1): 122-129.)
- [3] 张质明,王晓燕,于洋,等. 基于 GLUE 法的多指标水质模型参数率定方法[J]. 环境科学学报, 2014, 34(7): 1853-1861.
(Zhang Z M, Wang X Y, Yu Y, et al. A new calibration approach for multi-output water quality model based on GLUE method [J]. Acta Scientiae Circumstantiae, 2014, 34(7): 1853-1861.)
- [4] Albers C, Steffler P. Estimating transverse mixing in open channels due to secondary current-induced shear dispersion[J]. J of Hydraulic Engineering, 2007, 133(2): 186-196.
- [5] Baek K O, Seo I W. Transverse dispersion caused by secondary flow in curved channels[J]. J of Hydraulic Engineering, 2011, 137(10): 1126-1134.
- [6] 刘晓东,姚琪,薛红琴,等. 环境水力学反问题研究进展[J]. 水科学进展, 2009, 20(6): 885-893.
(Liu X D, Yao Q, Xue H Q, et al. Advance in inverse problems of environmental hydraulics[J]. Advance in Wwater Science, 2009, 20(6): 885-893.)
- [7] 顾莉,惠慧,华祖林,等. 河流横向混合系数的研究进展[J]. 水利学报, 2014, 45(4): 450-466.
(Gu L, Hui H, Hua Z L, et al. Advance on river transverse mixing coefficient[J]. J of Hydraulic Engineering, 2014, 45(4): 450-466.)
- [8] 刘毅,陈吉宁,杜鹏飞. 环境模型参数优化方法的比较[J]. 环境科学, 2002, 23(2): 1-6.
(Liu Y, Chen J N, Du P F. Comparison of parameter optimization algorithms for environmental model[J]. Environmental Science, 2002, 23(2): 1-6.)
- [9] 郭建青,王洪胜,李云峰. 确定河流纵向离散系数的相关系数极值法[J]. 水科学进展, 2000, 11(4): 3872-3911.
(Guo J Q, Wang H S, Li Y F. The correlation coefficient extreme value method to determine the dispersion parameters of river[J]. Advance in Water Science, 2000, 11(4): 3872-3911.)
- [10] 郭建青,李彦,王洪胜,等. 确定河流水质参数的抛物方程近似拟和法[J]. 水利水电科技进展, 2005, 25(2): 11-13.
(Guo J Q, Li Y, Wang H S, et al. Approximate fitting method for parabolic equation for determination of water quality parameters for rivers[J]. Advances in Science and Technology of Water Resources, 2005, 25 (2): 11-13.)
- [11] Singh S K, Beck M B. Dispersion coefficient of streams from tracer experiment data[J]. J of Environmental Engineering, 2003, 129(6): 539-546.
- [12] 薛红琴,赵尘,刘晓东,等. 确定天然河流纵向离散系数的有限差分-单纯形法[J]. 解放军理工大学学报: 自然科学版, 2012, 13(2): 214-218.
(Xue H Q, Zhao C, Liu X D, et al. Finite Difference method simplex method for determination of longitudinal dispersion coefficient in natural river[J]. J of PLA Universtiy of Science and Technology: Natural Science Edition, 2012, 13(2): 214-218.)
- [13] 侯景伟,孔云峰,孙九林. 蚁群算法在需水预测模型参数优化中的应用[J]. 计算机应用, 2012, 32(10): 2952-2955.
(Hou J W, Kong Y F, Sun J L. Application of ant colony algorithm for parameter optimization of water demand prediction model[J]. J of Computer Applications, 2012, 32(10): 2952-2955.)
- [14] Khanmirza E, Khaji N, Khanmirza E. Identification of linear and non-linear physical parameters of multistory shear buildings using artificial neural network[J]. Inverse Problems in Science and Engineering, 2015, 23(4): 670-687.
- [15] Benvenuti L, Kloss C, Pirker S. Identification of DEM

- simulation parameters by artificial neural networks and bulk experiments[J]. Powder Technology, 2016, 291: 456-465.
- [16] Pencheva T, Angelova M, Atanassova V, et al. Inter criteria analysis of genetic algorithm parameters in parameter identification[J]. Notes Intuitionistic Fuzzy Sets, 2015, 21(2): 99-110.
- [17] Liu X, Zhou Y, Hua Z, et al. Parameter identification of river water quality models using a genetic algorithm[J]. Water Science and Technology, 2014, 69(4): 687-693.
- [18] 付翠, 刘元会, 郭建青, 等. 识别河流水质模型参数的单纯形-差分进化混合算法[J]. 水力发电学报, 2015, 34(1): 125-130.
(Fu C, Liu Y H, Guo J Q, et al. Simplex-differential evolution hybrid algorithm for parameter identification of river water quality model[J]. J of Hydroelectric Engineering, 2015, 34(1): 125-130.)
- [19] Amirov A, Ustaoglu Z, Heydarov B. Solvability of a two dimensional coefficient inverse problem for transport equation and a numerical method[J]. Transport Theory and Statistical Physics, 2011, 40(1): 1-22.
- [20] Agapiou S, Larsson S, Stuart A M. Posterior contraction rates for the Bayesian approach to linear ill-posed inverse problems[J]. Stochastic Processes and their Applications, 2013, 123(10): 3828-3860.
- [21] Drovandi C C, Pettitt A N, Lee A. Bayesian indirect inference using a parametric auxiliary model[J]. Statistical Science, 2015, 30(1): 72-95.
- [22] Aslett L J M, Coolen F, Wilson S P. Bayesian inference for reliability of systems and networks using the survival signature[J]. Risk Analysis, 2015, 35(9): 1640-1651.
- [23] Liao M Y. Markov chain Monte Carlo in Bayesian models for testing gamma and lognormal S-type process qualities[J]. Int J of Production Research, 2016, 54(24): 7491-7503.
- [24] 朱嵩, 刘国华, 王立忠, 等. 水动力-水质耦合模型污染源识别的贝叶斯方法[J]. 四川大学学报: 工程科学版, 2009, 41(5): 31-35.
(Zhu S, Liu G H, Wang L Z, et al. A Bayesian Approach for the identification of pollution source in water quality model coupled with hydrodynamics[J]. J of Sichuan University: Engineering Science Edition, 2009, 41(5): 31-35.)
- [25] Zhang W, Liu J, Cho C, et al. A hybrid parameter identification method based on Bayesian approach and interval analysis for uncertain structures[J]. Mechanical Systems and Signal Processing, 2015, 60: 853-865.
- [26] Mojtabi A, Deville M O. One-dimensional linear advection-diffusion equation: Analytical and finite element solutions[J]. Computers & Fluids, 2015, 107: 189-195.
- [27] 张铁柱, 宋仁学, 韩志刚. 离散时间非线性系统线性化的泛模型方法[J]. 控制与决策, 2002, 17(2): 249-251.
(Zhang T Z, Song R X, Han Z G. Universal model method of linearization of discrete time nonlinear system[J]. Control and Decision, 2002, 17(2): 249-251.)
- [28] 岳良毅, 汪学锋, 尤云祥. 扩展型 Boussinesq 水波方程的一种改进有限体积/有限差分混合算法[J]. 水动力学研究与进展: A 辑, 2015, 30(5): 477-484.
(Yue L Y, Wang X F, You Y X. A modified hybrid finite-volume/finite-difference scheme for extended Boussinesq-type model[J]. J of Hydrodynamics: Ser A, 2015, 30(5): 477-484.)
- [29] Toubia O, Johnson E, Evgeniou T, et al. Dynamic experiments for estimating preferences: An adaptive method of eliciting time and risk parameters[J]. Management Science, 2013, 59(3): 613-640.
- [30] Efron B. Bayes' theorem in the 21st century[J]. Science, 2013, 340(6137): 1177-1178.
- [31] Yang H, Shao D, Liu B, et al. Multi-point source identification of sudden water pollution accidents in surface waters based on differential evolution and Metropolis-Hastings-Markov chain Monte Carlo[J]. Stochastic Environmental Research and Risk Assessment, 2016, 30(2): 507-522.
- [32] Kennedy D A, Dukic V, Dwyer G. Combining principal component analysis with parameter line-searches to improve the efficacy of Metropolis-Hastings MCMC[J]. Environmental and Ecological Statistics, 2015, 22(2): 247-274.

(责任编辑: 孙艺红)