

DOI:10.3876/j.issn.1000-1980.2024.05.003

基于 SWMM-贝叶斯耦合方法的排水管网污染溯源

杨立园¹, 黄 标¹, 刘甲春¹, 钱尚拓², 冯建刚²

(1. 宁波大学土木工程与地理环境学院, 浙江 宁波 315211; 2. 河海大学农业科学与工程学院, 江苏 南京 210098)

摘要:为解决雨水管道中未经处理的污水直接排放造成的一系列水污染问题和排水管网环境复杂性导致的反问题不确定性问题,基于暴雨洪水管理模型(SWMM)模拟的排水管网水质变化,采用 MatSWMM 工具箱构建了基于贝叶斯-马尔可夫链蒙特卡罗法的 SWMM-贝叶斯溯源模型,估算每个潜在污染源的位置、排放量和排放时间的概率分布。实例验证结果表明:在推理单污染源单个未知参数的情况下,SWMM-贝叶斯溯源模型可准确判断未知参数的真实值;当单污染源 3 个参数均未知时,由于存在多种组合性,溯源准确度将显著降低,只能判断出近似范围,但可以通过增加适当的水质监测点,提高 SWMM-贝叶斯溯源模型的准确性和效率;对于最大污染源为 2 个的情况,SWMM-贝叶斯溯源模型容易陷于局部最优解,通过利用似然函数对迭代过程的初始值进行择优处理的方法改进溯源模型,可以有效解决局部最优解的问题。

关键词:排水管网;污染溯源;贝叶斯-MCMC 法;SWMM-贝叶斯溯源模型;非法排放

中图分类号:TU992.0

文献标志码:A

文章编号:1000-1980(2024)05-0020-10

Pollution source tracing in sewer networks based on a Bayesian-SWMM coupled approach

YANG Liyuan¹, HUANG Biao¹, LIU Jiachun¹, QIAN Shangtuo², FENG Jiangang²

(1. School of Civil and Environmental Engineering, Ningbo University, Ningbo 315211, China;

2. College of Agricultural Science and Engineering, Hohai University, Nanjing 210098, China)

Abstract: In order to address the problem of directly discharging of untreated sewage from storm pipes causing a range of water pollution problems and the uncertainty problem of source inference due to the complexity of drainage system conditions, the storm water management model (SWMM) was employed to simulate the variation of water quality within the drainage system. A SWMM-Bayesian traceability model based on the Bayesian-MCMC (Monte Carlo Markov Chain) method was constructed using the MatSWMM toolbox to predict the probability distributions of the location of each potential source, the extent of discharge, and the time of discharge. Example validation results show that in the case of reasoning about a single unknown parameter of a single pollution source, the SWMM-Bayesian traceability model can accurately determine the true value of the unknown parameter. When all three parameters of a single pollution source are unknown, due to the existence of a variety of combinatorial properties, the traceability accuracy will be reduced significantly, and can only be judged to be in the approximate range, but it can be increased by adding the appropriate water quality monitoring points, to improve the accuracy and efficiency of SWMM-Bayesian traceability model. For the case that the maximum number of pollution sources is 2, the SWMM-Bayesian traceability model is easy to be trapped in the local optimal solution, and the improvement of the traceability model by using the likelihood function to optimize the initial value of the iterative process can effectively solve the problem of local optimal solution.

Key words: sewer network; source identification; Bayesian-MCMC method; SWMM-Bayesian traceability model; illicit discharge

城市分流制雨水管网中的雨污混接是一个普遍存在的问题。雨水和污水管道在设计时是分开的,但在实际运行中,生活污水和工业污水等会进入雨水管道直接排入河道,造成水质污染、水体返黑返臭,甚至导致病毒传播,严重威胁人民的健康安全^[1-2]。国外关于雨污混接的研究起步较早,美国国家环境保护局于 1993 年颁布了针对雨污混接的检测和清理用户手册^[3],并于 2004 年针对雨污混接的检测推出了以特征因子法为重

基金项目:国家重点研发计划项目(2022YFC3203200)

作者简介:杨立园(1997—),女,硕士研究生,主要从事污染溯源研究。E-mail:1692072154@qq.com

通信作者:黄标(1988—),男,副教授,博士,主要从事城市排水管网研究。E-mail:huangbiao@nbu.edu.cn

引用本文:杨立园,黄标,刘甲春,等.基于 SWMM-贝叶斯耦合方法的排水管网污染溯源[J].河海大学学报(自然科学版),2024,52(5):20-29.

YANG Liyuan, HUANG Biao, LIU Jiachun, et al. Pollution source tracing in sewer networks based on a Bayesian-SWMM coupled approach[J]. Journal of Hohai University(Natural Sciences), 2024, 52(5): 20-29.

点的导则^[4]。Irvine 等^[5]利用特征因子法,以大肠杆菌作为指标研究了美国某些城市的雨污混接问题。Karri^[6]利用蒙特卡罗算法结合化学质量平衡法建立溯源模型,判别了污染源成分。我国从 2005 年开始开展雨污混接的相关研究,主要集中在上海市。李田等^[7]根据水量平衡原理分析了上海市某雨水管网的水位和流量变化规律,证明该雨水管网中存在雨污混接并确定了非雨水源的混接量。水质特征因子法定量判断雨水管网旱天水量来源的核心是化学质量平衡模型(chemical mass balance model, CMBM),将某一混接类型的水质特征因子质量浓度数据基于确定性算法代入 CMBM,但该方法忽略了水质特征因子数据的不确定性,会导致计算结果存在偏差、解析结果不闭合等问题^[8-9],同时由于水化学参数通常不稳定,因此其适用范围有限,仅适用于稳定性的特征因子识别。有学者通过结合其他溯源方法提高水质特征因子法的精度:徐祖信等^[10]将水质特征因子法与蒙特卡罗算法结合,对分流制排水管网雨污混接问题进行定量分析,证明了考虑混接来源中多个混接点空间质量浓度差异性以及测量误差等不确定因素的蒙特卡罗算法与确定性算法相比,具有更高的分析结果可靠性;Zhao 等^[11]通过结合水质特征因子法和蒙特卡罗算法,确定了地下水进入城市排水管网的高风险子集水区,并利用微生物遗传算法判断地下水来源及其时空流动。

污水进入雨水管网后会沿管道向下游扩散造成水污染,因此需要获取污染源的相关信息,并实施混接改造工程,减少水环境的外部污染,提高水体质量^[12-13]。混接点的溯源属于水动力反问题,而反问题具有数据不确定性以及观测数据噪声大等问题,使得反问题的求解非常困难^[14]。Tikhonov 正则化法是求解反问题的方法之一,Skaggs 等^[15]利用该方法重现了污染源释放过程,并表明当数据足够时,该方法重现污染物浓度变化过程是有效的。优化方法也可以求解反问题,将监测数据与模拟数据之差设为目标函数,应用最优化法寻找目标函数的最小值,将反问题作为寻求正问题参数的最优解来处理。Banik 等^[16]基于多目标优化方法,使用 NSGA-II 进行求解,研究了污水系统中监测点的最佳放置位置。Xu 等^[17]结合了 CMBM 与蒙特卡罗算法,计算进入雨水管网的污水量,并利用 PySWMM 和微生物遗传算法建立逆优化模型,查明了非法连接和管道损坏的高风险聚集区域。李敏慧等^[18]将权重系数法引入传统遗传算法中,并通过控制种群进化方向提高溯源效率,发现该方法比传统遗传算法能更快地溯源出污染源的排放位置和排放时间。虽然 Tikhonov 正则化法和优化法可以求解反问题,但没有考虑观测数据的不确定性问题,所以得到的溯源结果均受不确定性扰动的影响较大。

Neupauer 等^[19-20]考虑到确定性算法没有充分利用先验信息,也没有考虑反演参数的不确定性对反演结果的影响,将能够充分利用先验信息并考虑不确定性的贝叶斯推理法引入反问题求解中,并将发生污染事故的大概位置及大致排放时间序列写成概率密度函数的形式,通过已知的污染源信息得到最优估计值,即可得到反问题的解。Ghane 等^[21]利用倒向概率法(BPM)溯源了地表水中污染源位置和污染时间。Snodgrass 等^[22]最先将贝叶斯推理法用于地下水污染源识别问题,在只考虑一个已知位置污染源情况下反演了污染物的排放过程。郑卓乐^[23]采用贝叶斯-马尔可夫链蒙特卡罗(Monte Carlo Markov Chain, MCMC)法将先验信息和测量误差融合到溯源过程中,然后结合暴雨洪水管理模型(storm water management model, SWMM)反演污水管网中的瞬态排放过程,对污水管网中偷排漏排超标废水的污染源进行溯源,并利用增加水质监测点的方法提高溯源算法的精度。我国多位学者利用贝叶斯-MCMC 法实现了在顺直河道中反演多个污染源的排污过程^[24-25]。顾文龙等^[26]针对贝叶斯-MCMC 法采用的 Metropolis 抽样算法提出了新的改进方案,仅使用似然度达到 0.9 以上的抽样值作为初始值,且在接受新的抽样值时似然度更要高达 0.96 以上,以解决迭代过程中陷入局部最优或难以收敛的问题,保证每次抽取的样本值始终在一定精度范围内,显著提高了计算效率。

为考虑数据不确定性的问题,本文采用 SWMM 重现了以混接点为污染源的雨水管网的水力动态和水质属性,采用 MatSWMM 工具箱构建了基于贝叶斯-MCMC 法的 SWMM-贝叶斯溯源模型,并确定了与污染源相关的 3 个参数(排放点、排放时间和排放量),针对在溯源多个污染源的情况下结果易于收敛到局部最优解的问题,提出了改进的 SWMM-贝叶斯溯源模型,以期溯源雨水管网中的非法排放提供参考。

1 溯源方法

本文方法以 Shao 等^[27]建立的 SWMM-贝叶斯溯源框架为基础,主要包括两个核心计算部分:①基于 SWMM 模拟整个雨水管网的流量和污染物扩散过程;②贝叶斯-MCMC 法:利用 MCMC 抽样进行贝叶斯推理计算。利用 MatSWMM 将两个核心计算部分耦合成 SWMM-贝叶斯溯源模型,针对单个未知参数验证 SWMM-贝叶斯溯源模型的有效性,再针对污染源 3 个参数(排放点、排放时间和排放量)均缺失的情况下,利

用不同的水质监测点提高溯源精度,最后针对同时存在两个污染源的复杂参数情况,利用改进的 SWMM-贝叶斯溯源模型估算未确定污染源参数的概率。

1.1 SWMM

SWMM 是一种分布式离散时间模拟模型,通常用于模拟城市排水管网的水力变化和污染物迁移扩散。在 SWMM 传输系统中,假设污染物在断面上混合均匀,污染物质量浓度只随流程的方向变化,即污染物的模拟被假定为连续搅动水箱式反应器,其他参数始终不变。在确定每个计算时间步长的水力状态后,计算节点和管段污染物质量浓度变化过程^[28],并通过求解一个完整反应器的质量守恒方程解决污染物的传输问题^[29]。

1.2 贝叶斯-MCMC 法

为了处理环境水力学反演问题中固有的不确定性问题,所有形式的不确定性都表示为概率分布,采用贝叶斯推理法进行统计反演,通过结合观测数据,将先验知识转化为后验知识^[30]。先验分布本身具有主观性,可以根据经验选择。本文先验信息是指在获取质量浓度数据 Y 之前,污染源参数 X 的概率分布,即先验概率 $p(X=(J_x,M,T_d))$;排放点的先验概率密度函数为离散均匀分布(式(1)),排放量、排放时间的先验概率密度函数为连续均匀分布(式(2)(3))。

$$p(J_x)=1/N \tag{1}$$

$$p(M)=\begin{cases} \frac{1}{M_{\max}-M_{\min}} & M_{\min}\leq M\leq M_{\max} \\ 0 & \text{其他} \end{cases} \tag{2}$$

$$p(T_d)=\begin{cases} \frac{1}{T_{\max}-T_{\min}} & T_{\min}\leq T_d\leq T_{\max} \\ 0 & \text{其他} \end{cases} \tag{3}$$

式中: J_x 为潜在的排放点; N 为雨水管网节点总数; M 为排放量; T_d 为污染物的初始排放时间; $M_{\min}$ 、 $T_{\min}$ 为参数值的下限; $M_{\max}$ 、 $T_{\max}$ 为参数值的上限。

3 个污染源参数的联合先验概率 $p(X=(J_x,M,T_d))$ 可计算如下:

$$p(X=(J_x,M,T_d))=p(J_x)p(M)p(T_d) \tag{4}$$

$p(Y|X)$ 表示在污染源参数为 X 的情况下,质量浓度监测数据为 Y 的条件概率,也称为似然函数,它量化了模型预测值与实际观测值之间的一致性^[31]。假设观测噪声为白噪声,则进一步假设该噪声服从均值为 0、标准差为 σ 的正态分布,因此似然函数可表示为:

$$p(Y|X=(J_x,M,T_d))=\frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\sum_{i=1}^n\frac{[Y_i-N_i(J_x,M,T_d|X)]^2}{2\sigma^2}\right\} \tag{5}$$

式中: Y_i 为实测质量浓度数据集; N_i 为贝叶斯推理法的采样数值在监测点的预测质量浓度数据集; n 为提取质量浓度数据的总数。

后验概率 $p(X|Y)$ 表示获得 Y 后,污染源参数 X 的分布概率,根据贝叶斯推理法,可得:

$$p(X|Y)=\alpha p(X)\frac{1}{(\sqrt{2\pi}\sigma)^n}\exp\left\{-\sum_{i=1}^n\frac{[Y_i-N_i(J_x,M,T_d|X)]^2}{2\sigma^2}\right\} \tag{6}$$

式中 α 为比例常数^[20]。

1.3 梅特罗波利斯-黑斯廷斯 (Metropolis-Hastings, M-H) 算法

高维空间中的后验分布问题是传统解析方法面临的一项艰巨任务,本文采用 MCMC 法简化采样过程^[32]。MCMC 法会生成一系列采样点,这些采样点构成了一条马尔科夫链,马尔科夫链收敛后生成的平稳分布即为所求的后验分布。M-H 算法是应用最为广泛的 MCMC 法,是一种基础算法,其他 MCMC 法可视为其特例或扩展^[33]。为了建立一个后验概率为平稳分布的马尔科夫链,必须选择一个转移概率,表示为 $q(X_t,X_{t+1})$ 以及函数 $a(X_t,X_{t+1})$,其中 $0<a(X_t,X_{t+1})\leq 1$,对任一组合 (X_t,X_{t+1}) ,有 $X_t\neq X_{t+1}$,这些因素共同构成了一个转移核 $p(X_t,X_{t+1})$:

$$p(X_t,X_{t+1})=q(X_t,X_{t+1})a(X_t,X_{t+1}) \tag{7}$$

采样点的接受概率设定如下:在 M-H 算法中,根据参数的先验分布选择转移概率函数分布,选取均匀分布 $q(X_t-\varepsilon,X_t+\varepsilon)$ 或正态分布 $q(X_t,\varepsilon^2)$,其中 ε 为采样步长,步长表示采样过程中未知参数连续值之间的

间隔,是影响采样效率的关键参数。

M-H 算法流程如下:①初始化。根据先验信息生成初始状态或初始点, $X_0=X_t$,利用当前状态 X_t ,从建议分布 $q(X_t, X_{t+1})$ 中抽取候选样本 X^* 。②接受概率计算。计算 X^* 的接受概率 $a(X_t, X^*)$ 。③决定。从均匀分布 $U(0, 1)$ 中生成一个随机数 u ,如果 $u \leq a(X_t, X_{t+1})$,则接受新状态 $X^* = X_{t+1}$;否则拒绝接受,算法保持当前状态 $X_t=X_{t+1}$;然后,算法从步骤①开始,提取新样本。

1.4 SWMM 与贝叶斯推理法的耦合

本文通过 SWMM 建立雨水管网水力水质模型,模拟管道和各节点的水质变化过程^[34],基于下游水质监测点的污染物质量浓度时间序列,采用贝叶斯-MCMC 法对系统内污染源项未知参数(排放点、排放时间和排放量)进行反向求解,给出可能取值区间的概率。采用 Riaño-Briceño 等^[35]开发的 MatSWMM 工具箱,将 SWMM 计算引擎与贝叶斯-MCMC 算法耦合,构建 SWMM-贝叶斯溯源模型,耦合的主要步骤如下:①构建 SWMM。构建 SWMM,设置污染源参数的“真实值”,如排放点、排放量和排放时间,为 SWMM 创建初始输入.inp 文件;运行 SWMM,生成.rpt 文件从而获取水质监测点的质量浓度 Y 。②初始采样。根据先验信息随机生成污染源参数的初始样本 X_0 。③SWMM 模拟和计算似然函数。修改.inp 文件,找到文件中该参数值的位置,并将其替换为初始样本 X_0 ,通过 MatSWMM 工具箱运行 SWMM 进行模拟,从生成的.rpt 报告文件中提取模拟的预测质量浓度值 N ,通过比较预测质量浓度值 N 与步骤①中的“真实值”质量浓度 Y ,计算似然函数 $p(Y|X_0)$ 。④参数采样。使用建议分布从当前参数状态 $X_0=X_t$ 中抽取新的测试参数 X^* ,重复步骤③,计算新样本 X^* 的似然函数 $p(Y|X^*)$ 。⑤验收检查。从区间 $[0, 1]$ 中随机抽取一个均匀数 u ,如果 $u \leq a(X_t, X^*)$,则接受新样本并将其作为下一次迭代的初始样本,否则拒绝接受 X^* ,保留 X_t 用于下一次迭代。⑥迭代。重复步骤③~⑤,直到达到所需的迭代次数,以获得污染源参数的后验样本集合。⑦后验分布。分析后验样本,推断模型参数的后验分布。

1.5 性能评价

与 Shao 等^[27]的方法类似,使用后验分布直方图和标准误差评价本文方法的有效性,其中标准误差包括平均绝对误差(MAE)和中位绝对误差(MedAE)。

2 实例验证

宁波市总面积约76.5hm²,雨水管网由20个管段、20个节点和1个出水口组成,管道直径范围400~1200mm,曼宁粗糙系数设为0.01,平面布置如图1所示。其中,雨水管网数据为shp格式矢量数据,可利用ArcGIS对矢量数据进行处理,最终生成SWMM可识别的输入文件^[36-38]。根据服务区域估算了每个节点的污水流量,得到的流量作为干燥天气下的恒定流量输入到每个节点系统中。在所有数值模拟中,假设:①在任何非法排放之前,雨水管网内的流量达到稳定状态;②假设污染物遵循一阶衰减反应,衰减系数为0.25^[26],在下游节点安装水质监测传感器,获得污染物质量浓度的时间序列数据。

SWMM 的工况设置见表1,均为瞬时排放模式。首先,对单点源排污的情形进行水质模拟(A类和B类):A类工况假设某日0:20在上游单个节点13瞬时排入1000g的污染物;工况A1~A6为溯源单个未知参数(M 或 T_d),而工况B1~B2为溯源3个未知参数。其次,对多点源排污的情形进行水质模拟(C类):如工况C1,假设某日0:20在节点5瞬时排入

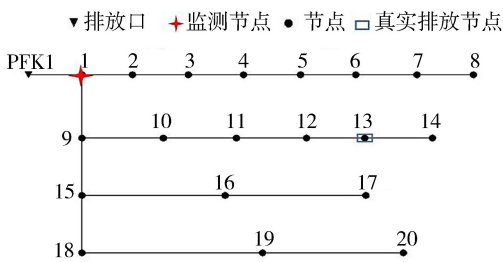


图1 雨水管网平面布置

Fig.1 Plane layout of rainwater pipe network

表1 SWMM 工况设置

Table 1 Working condition setting of SWMM

组别	污染源数量	未知参数个数	J_x	M/g	T_d/min
A1	1	1	13	800	20
A2	1	1	13	1 000	20
A3	1	1	13	1 200	20
A4	1	1	13	1 000	10
A5	1	1	13	1 000	20
A6	1	1	13	1 000	30
B1	1	3	13	1 000	20
B2	1	3	19	1 000	20
C1	2	2	5,13	1 000	20
C2	2	2	5,16	1 000	20
C3	2	2	13,19	1 000	20

1 000 g 的污染物,同时在节点 13 处 1 min 排入质量为 1 000 g 的污染物。

利用下游监测点获取的质量浓度数据反演上游的污染物排放点、排放量及排放时间。MCMC 法迭代的初始点在先验范围内随机产生,利用转移概率函数抽取模拟值,将模拟值作为水质模型边界条件输入 SWMM,计算出此模拟值对应的质量浓度变化过程。将此模拟质量浓度值与实际混入污染物的质量浓度值进行拟合度对比,得到该未知参数的概率分布。

2.1 单一污染源的推理

2.1.1 一个未知参数

在单一未知污染源的情况下,假设 3 个源参数 (J_x , M , T_d) 只有一个 是未知的,其他两个参数已知,工况 A1 ~ A3 仅对 M 进行推理,工况 A4 ~ A6 仅对 T_d 进行推理。质量浓度监测从 0:20 开始读取数据, t 为 0:20、0:23、0:26、0:29、0:32、0:35、0:38、0:41 时在监测点收集质量浓度数据,用于未知参数的反推。为量化 SWMM-贝叶斯溯源模型的性能,计算了识别结果的统计 MAE 和 MedAE(表 2),由表 2 可知,SWMM-贝叶斯溯源模型预测值基本上与真实值相匹配,当只有一个未知参数时,SWMM-贝叶斯溯源模型可有效探索整个参数域,证明了该溯源模型的有效性。

2.1.2 多个未知参数

在 3 个污染源参数均未知的情况下,参数的多组合性以及管网中各管段水力条件不同等因素,使得非真实值的参数组合可能在监测点获得相似甚至相同的质量浓度结果,导致结果的多组合性也随之增加。另外,贝叶斯推理法的有效性取决于似然函数,因为它与未知变量的数量直接相关(式(5)),未知数的增加将对反推结果的准确性产生不利影响。

与一个未知参数的溯源结果相比,识别的结果范围显著增加,排放点、排放量及排放时间的后验概率密度较分散,没有收敛到真实值,不能准确推理出污染源参数信息(图 2 和图 3)。为了定量地评估收敛性,对识别结果的误差进行了分析(表 3),由于排放点是整数离散量,因此没有给出与排放点相关的误差,此误差不具有物理意义,与表 2 对比可以看出,MAE 和 MedAE 均大于仅识别一个未知参数的情况,大部分误差超过了 10%,甚至达到了 25.67%,表明只有一个未知参数的情况具有更清晰的可追溯性。

表 2 一个未知参数情况下 SWMM-贝叶斯溯源模型
预测值与真实值比较

Table 2 Comparison of SWMM-Bayesian predicted and true values in the cases of inferring one unknown source parameter

工况	M			工况	T_d		
	真实值/g	MAE/%	MedAE/%		真实值/min	MAE/%	MedAE/%
A1	800	0.85	0.00	A4	10	1.02	0.02
A2	1 000	0.20	0.13	A5	20	0.60	0.03
A3	1 200	1.60	0.40	A6	30	1.79	0.50

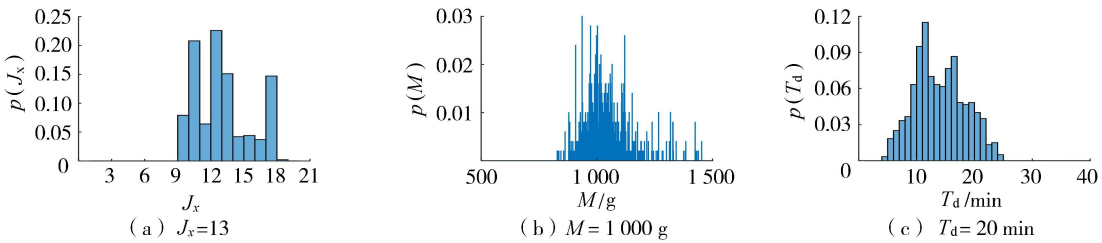


图 2 3 个源参数同时被识别的后验概率直方图(工况 B1)

Fig.2 Posterior probability histogram in which three source parameters are identified simultaneously(Case B1)

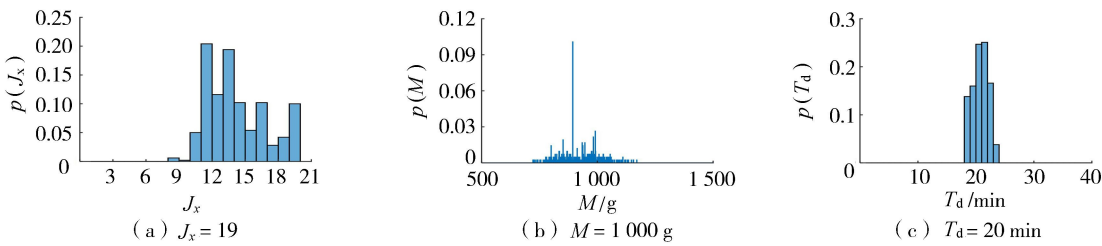


图 3 3 个源参数同时被识别的后验概率直方图(工况 B2)

Fig.3 Posterior probability histogram in which three source parameters are identified simultaneously (Case B2)

根据 SWMM 中污染物输送过程的控制方程,任意节点的监测质量浓度为初始质量浓度 ρ_0 、一阶降解系数 k 、管段长度 x 和水流流速 v 的组合,表示为 $\rho_t = \rho_0 e^{-k \frac{x}{v}}$ 。对于具有树状结构的管网,这 3 个参数有许多可能的组合,可以产生相同的监测质量浓度数据。

2.1.3 水质监测点的布置

由第 2.1.2 节可知,虽然同时识别 3 个源参数的结果不理想,但是可以看出 SWMM-贝叶斯溯源模型的优势,与优化方法相比,它能够考虑所有可能的值,并为所有潜在的候选值提供概率。对于污染溯源,质量浓度监测数据是所有溯源方法的必要依据,对污染源的可识别性和识别精确度具有重要作用^[39]。当监测数据不能充分反演实际的污染源特征时,可以利用增加水质监测点的方法,并结合雨水管网的拓扑结构,根据潜在节点的频率统计,在上游高概率的区域增加水质监测点,缩小先验信息范围,提高 SWMM-贝叶斯溯源模型的准确性和效率。

一般情况下,实际雨污混接的 3 个污染源参数信息均未知,掌握排放点位置信息对后续实施混接改造工程十分重要,于是对排放点落在先验信息各区域的频数进行统计。工况 B1 真实排放点在节点 9~14 之间的概率最大,工况 B2 真实排放点在节点 9~20 之间的概率最大。为了更精确地获取污染源位置,在上游高概率区域设置新的水质监测点。分支相遇的节点具有相对较高的熵,在这个节点上放置一个监测点,与其他节点相比,能够获取更多的信息,减少更多的不确定性^[40]。因此选定节点 9 作为新的水质监测点,再次采用 SWMM-贝叶斯溯源模型对高概率区域进行推理,进一步优化污染源 3 个参数同时溯源的结果。优化后的结果如图 4 和图 5 所示,与图 2 和图 3 相比,所有参数的不确定性都显著降低,表明增加监测点可显著提高贝叶斯推理性能,特别是在确定源位置方面,更加接近真实值。由表 4 可知,在增加一个监测点后,中值误差和平均误差均显著降低,验证了增加监测点数量的方法能够提高识别排放点的能力,这在实际应用中十分关键。

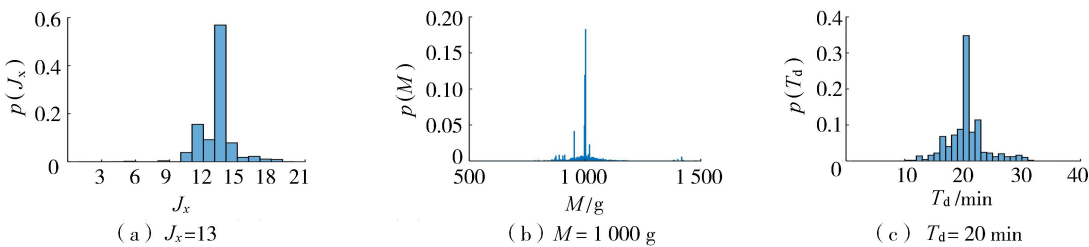


图 4 增加监测点后 3 个源参数同时被识别的后验概率直方图(工况 B1)

Fig. 4 Posterior probability histogram of case B1 in which three source parameters are identified simultaneously after adding monitoring sites

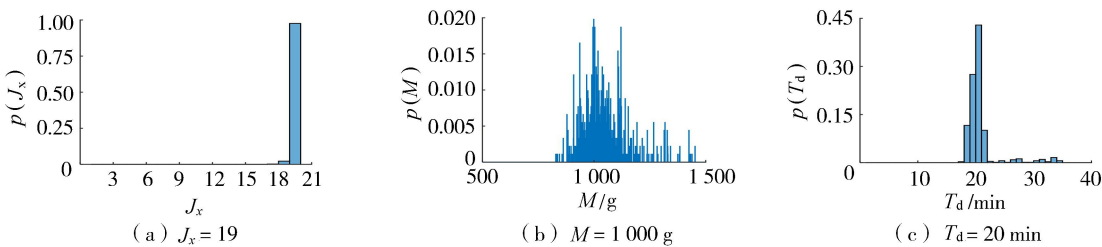


图 5 增加监测点后 3 个源参数同时被识别的后验概率直方图(工况 B2)

Fig. 5 Posterior probability histogram of case B2 in which three source parameters are identified simultaneously after adding monitoring sites

2.2 多个污染源的推理

在真实的排水管网中通常存在多个污染源,当管网多个分支存在多个污染源时,SWMM-贝叶斯溯源模

表 3 3 个未知参数情况下 SWMM-贝叶斯溯源模型
预测值与真实值比较

Table 3 Comparison of SWMM-Bayesian predicted and true values in the cases of inferring all three unknown source parameters

工况	M			Td		
	真实值/g	MAE/%	MedAE/%	真实值/min	MAE/%	MedAE/%
B1	1 000	8.10	11.40	20	25.67	23.02
B2	1 000	14.20	10.09	20	16.04	9.10

型只能查找潜在的污染源位置,因为不同分支内的多个污染源在监测点可能产生类似的质量浓度过程。将构建的SWMM-贝叶斯溯源模型应用在排水管网多个污染源溯源的场景中,加大迭代次数后发现,预测结果容易陷入局部最优解,导致溯源结果的稳定性与精度降低,原因为传统的M-H算法针对较为复杂参数空间的随机游走抽样,通常无法在全局范围内进行搜索。因此,基于具有潜在非唯一解的反问题,根据陷入局部最优解的随机游走产生的样本来求解反问题,是不全面的甚至可能是误导的^[41]。

为了避免出现局部最优解,在其陷入局部最优时设计一种方案使其逃逸局部最优,使用不同的初始值多次进行迭代,增加找到全局最优解的机会。每次迭代可能会收敛到不同的局部最优解,但通过多次尝试可以提高找到全局最优解的概率,具体改进方法如下:记录每一个样本的模拟质量浓度值与观测值的似然度,在迭代过程中,如果当前的候选采样值没有被接受,则将似然度为最大的样本作为下一次迭代的初始值,使用不同的初始值多次进行迭代,使计算能够直接采用一个较为准确的初始值,从而增加找到全局最优的概率,也避免了马尔科夫链前期大幅波动的情况。

假定污染源个数为2,在上游两个节点同时排放污染物,在节点1进行下游质量浓度变化过程的监测,对工况C1~C3中的未知参数排放点及排放量同时利用改进SWMM-贝叶斯溯源模型进行推理,结果表明两个排放点均收敛到设置的真实节点(图6)。

由图7可知,经历2000次的迭代后均能达到收敛,在真实值附近的振荡逐渐变小;排放点离水质监测点

表4 3个未知参数的情况下增加一个监测点后
SWMM-贝叶斯溯源模型预测值与真实值比较

Table 4 Comparison of SWMM-Bayesian predicted and true values in the cases of inferring all three unknown source parameters with one additional monitoring site

工况	M			T _d		
	真实值/g	MAE/%	MedAE/%	真实值/min	MAE/%	MedAE/%
B1	1 000	4.31	2.70	20	2.20	0.20
B2	1 000	7.50	0.90	20	2.70	0.00

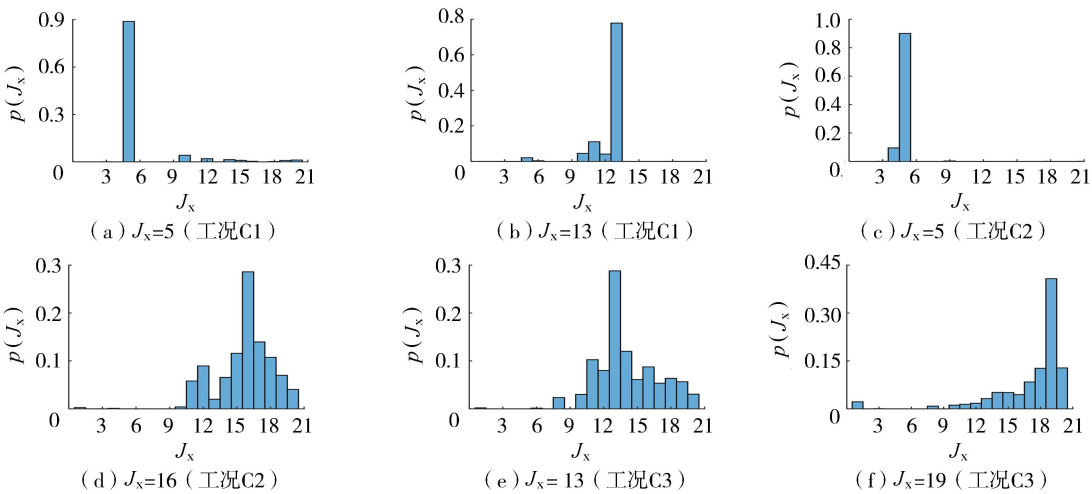


图6 参数J_x的后验概率直方图

Fig. 6 Posterior probability histogram of parameter J_x

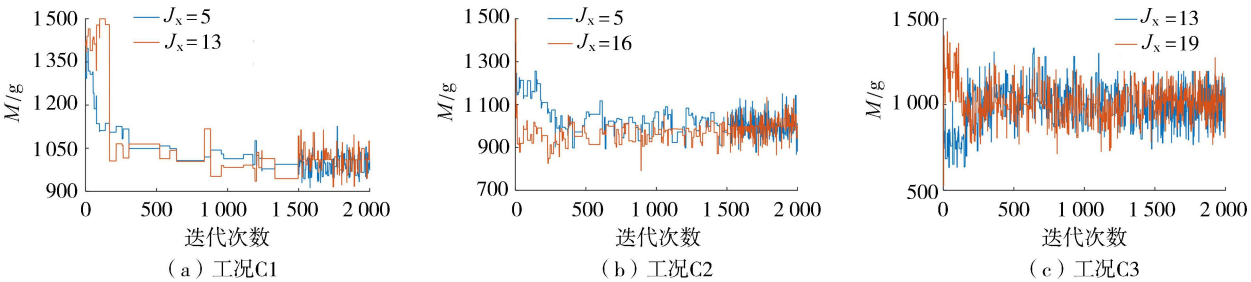


图7 工况C1~C3的迭代过程

Fig. 7 Iteration processes for Case C1 to C3

越远,迭代过程越震荡,收敛的速度越慢;验证了提出的 SWMM-贝叶斯溯源模型在同时识别两个污染源位置和排放量时的收敛性及有效性。

为了计算 M 的统计误差,取 2 000 步的马尔科夫链序列计算各个未知参数抽样值的均值和中值,得到未知参数的反演误差。由表 5 可知,反推污染物排放量的结果误差均没有超过 10%,但排放量的识别结果有略微偏差,原因为上游两个点源的排污过程在下游的水质监测点产生了叠加作用,如果其中一个节点排放污染物的量的反推结果要大于 1 000 g,那么另一个节点排放污染物的量小于 1 000 g 是合理的。为了提高两个污染源的 4 个参数的反演精度,可以增加马尔科夫链的长度,但能够证明提出改进 SWMM-贝叶斯溯源模型具有全局收敛能力,能够很好地收敛在真实值附近。在实际应用中,对于排放点能否精准溯源才是至关重要的,此改进方法可以有效提供排放点的信息。

3 结 论

- a. 在推理单一未知参数的情况下,SWMM-贝叶斯溯源模型可准确判断未知参数的真实值。当污染源 3 个参数均未知时,由于存在多种组合性,溯源准确度将显著降低,只能判断出近似范围。
- b. 利用多个水质监测点的质量浓度数据筛选概率较大的潜在污染源,可以提升溯源准确性及可靠性,证明了 SWMM-贝叶斯溯源模型结合优化设置水质监测点的方法,能够有效推断复杂情况下的污染源信息。
- c. 将管网系统中的污染情况扩展到多个污染源的场景,假定最大污染源个数为 2 时,SWMM-贝叶斯溯源模型的有效性不仅受限于参数的多重组合,还受限于多点排放质量浓度过程线的叠加性,导致溯源结果容易陷于局部最优解。
- d. 改进 SWMM-贝叶斯溯源模型充分利用似然函数对迭代过程的初始值进行择优处理,避免了因局部最优解导致对溯源结果判断的干扰,证明提出的改进方法能够提高 SWMM-贝叶斯溯源模型的全局搜索能力。

参考文献:

[1] 张厚强,尹海龙,金伟,等. 分流制雨水系统混接问题的调研技术体系[J]. 中国给水排水,2008,24(14):95-98. (ZHANG Houqiang,YIN Hailong,JIN Wei,et al. Investigation technology study on non-stormwater entries into separate storm drainage system[J]. China Water & Wastewater,2008,24(14):95-98. (in Chinese))

[2] 方国华,王鸿祯,于凤存,等. 城市地表集中式水源地管理风险评价[J]. 水利经济,2022,40(6):62-67. (FANG Guohua,WANG Hongzhen,YU Fengcun,et al. Research on risk assessment of urban centralized surface water source management[J]. Journal of Economics of Water Resources,2022,40(6):62-67. (in Chinese))

[3] FIELD R,PITT R,LALOR M,et al. Investigation of dry-weather pollutant entries into storm-drainage systems[J]. Journal of Environmental Engineering,1994,120(5):1044-1066.

[4] PITT R,CHATURVEDULA S,KARRI V,et al. Source verification of inappropriate discharges to storm drainage systems[J]. Proceedings of the Water Environment Federation,2004,2004(16):1192-1218.

[5] IRVINE K,ROSSI M C,VERMETTE S,et al. Illicit discharge detection and elimination:low cost options for source identification and trackdown in stormwater systems[J]. Urban Water Journal,2011,8(6):379-395.

[6] KARRI V R. Source quantification of inappropriate discharges to storm drainage systems[D]. Tuscaloosa:University of Alabama,2003.

[7] 李田,周永潮,冯仓,等. 分流制雨水系统雨污混接水量的模型分析[J]. 同济大学学报(自然科学版),2008,36(9):1226-1231. (LI Tian,ZHOU Yongchao,FENG Cang,et al. Model analysis of illicit connections flow of storm sewer system[J]. Journal of Tongji University (Natural Science),2008,36(9):1226-1231. (in Chinese))

[8] 孟莹莹,冯沧,李田,等. 不同混接程度分流制雨水系统旱流水量及污染负荷来源研究[J]. 环境科学,2009,30(12):3527-3533. (MENG Yingying,FENG Cang,LI Tian,et al. Identifying dry-weather flow and pollution load sources of separate storm

表 5 识别 2 个未知参数情况下 SWMM-贝叶斯溯源模型 M 预测值与真实值比较

Table 5 Comparison of SWMM-Bayesian predicted and true values of M in the cases of inferring two unknown source parameters

工况	真实值/g	MAE/%		MedAE/%	
		M_1	M_2	M_1	M_2
C1	1 000	0.00	5.40	0.65	4.40
C2	1 000	1.30	8.28	1.30	3.13
C3	1 000	1.53	1.10	0.90	1.40

注:工况 C1、C2 的 M_1 为节点 5 的排放量,C3 的 M_1 为节点 13 的排放量;工况 C1、C2、C3 的 M_2 分别为节点 13、16、19 的排放量。

- sewer systems with different degrees of illicit discharge[J]. *Environmental Science*, 2009, 30(12): 3527-3533. (in Chinese))
- [9] 尹海龙, 郭龙天, 解铭, 等. 基于特征因子的食品工业废水排放追踪方法[J]. *同济大学学报(自然科学版)*, 2019, 47(8): 1168-1174. (YIN Hailong, GUO Longtian, XIE Ming, et al. Source tracking of food industry wastewater discharge into sewers using marker species[J]. *Journal of Tongji University (Natural Science)*, 2019, 47(8): 1168-1174. (in Chinese))
- [10] 徐祖信, 汪玲玲, 尹海龙. 基于水质特征因子和 Monte Carlo 理论的雨水管网混接诊断方法[J]. *同济大学学报(自然科学版)*, 2015, 43(11): 1715-1721. (XU Zuxin, WANG Lingling, YIN Hailong. Quantification of non-storm water flow entries into storm drains using monte carlo based marker species approach[J]. *Journal of Tongji University (Natural Science)*, 2015, 43(11): 1715-1721. (in Chinese))
- [11] ZHAO Zhichao, YIN Hailong, XU Zuxin, et al. Pin-pointing groundwater infiltration into urban sewers using chemical tracer in conjunction with physically based optimization model[J]. *Water Research*, 2020, 175: 115689.
- [12] 李一平, 郑可, 周玉璇, 等. 南方城市污水处理系统效能评估与提质增效策略制定[J]. *水资源保护*, 2022, 38(3): 50-57. (LI Yiping, ZHENG Ke, ZHOU Yuxuan, et al. Efficiency evaluation of a sewage system and strategy formulation of quality and efficiency improvement in city of Southern China[J]. *Water Resources Protection*, 2022, 38(3): 50-57. (in Chinese))
- [13] QIAN Shangtuo, QIAO Xuyang, ZHANG Wenming, et al. Machine learning-based prediction for settling velocity of microplastics with various shapes[J]. *Water Research*, 2024, 249: 121001.
- [14] MOGHADDAM M B, MAZAHARI M, SAMANI J M V. Inverse modeling of contaminant transport for pollution source identification in surface and groundwaters: a review[J]. *Groundwater for Sustainable Development*, 2021, 15: 100651.
- [15] SKAGGS T H, KABALA Z J. Recovering the release history of a groundwater contaminant[J]. *Water Resources Research*, 1994, 30(1): 71-79.
- [16] BANIK B K, ALFONSO L, TORRES A S, et al. Optimal placement of water quality monitoring stations in sewer systems: an information theory approach[J]. *Procedia Engineering*, 2015, 119: 1308-1317.
- [17] XU Zuxin, QU Yang, WANG Siyu, et al. Diagnosis of pipe illicit connections and damaged points in urban stormwater system using an inversed optimization model[J]. *Journal of Cleaner Production*, 2021, 292: 126011.
- [18] 李敏慧, 荆立, 朱彦君, 等. 基于改进型遗传算法的水质溯源问题[J]. *河海大学学报(自然科学版)*, 2018, 46(5): 402-407. (LI Minhui, JING Li, ZHU Yanjun, et al. Research on the water source identification based on an improved genetic algorithm [J]. *Journal of Hohai University (Natural Sciences)*, 2018, 46(5): 402-407. (in Chinese))
- [19] NEUPAUER R M, WILSON J L. Adjoint method for obtaining backward-in-time location and travel time probabilities of a conservative groundwater contaminant[J]. *Water Resources Research*, 1999, 35(11): 3389-3398.
- [20] KEATS A, YEE E, LIEN F S. Bayesian inference for source determination with applications to a complex urban environment[J]. *Atmospheric Environment*, 2007, 41(3): 465-479.
- [21] GHANE A, MAZAHARI M, SAMANI J M V. Location and release time identification of pollution point source in river networks based on the backward probability method[J]. *Journal of Environmental Management*, 2016, 180: 164-171.
- [22] SNODGRASS M F, KITANIDIS P K. A geostatistical approach to contaminant source identification [J]. *Water Resources Research*, 1997, 33(4): 537-546.
- [23] 郑卓乐. 排水管道排放污染物贝叶斯统计算法-SWMM 耦合溯源模型[D]. 重庆: 重庆大学, 2018.
- [24] WANG Jiabiao, ZHAO Jianshi, LEI Xiaohui, et al. New approach for point pollution source identification in rivers based on the backward probability method[J]. *Environmental Pollution*, 2018, 241: 759-774.
- [25] YANG Haidong, SHAO Dongguo, LIU Biyu, 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.
- [26] 顾文龙, 卢文喜, 张宇, 等. 基于贝叶斯推理与改进的 MCMC 方法反演地下水污染源释放历史[J]. *水利学报*, 2016, 47(6): 772-779. (GU Wenlong, LU Wenxi, ZHANG Yu, et al. Reconstructing the release history of groundwater contamination sources based on the bayesian inference and improved mcmc method[J]. *Journal of Hydraulic Engineering*, 2016, 47(6): 772-779. (in Chinese))
- [27] SHAO Zhiyu, XU Lei, CHAI Hongxiang, et al. A Bayesian-SWMM coupled stochastic model developed to reconstruct the complete profile of an unknown discharging incidence in sewer networks[J]. *Journal of Environmental Management*, 2021, 297: 113211.
- [28] ZOPPOU C. Review of urban storm water models[J]. *Environmental Modelling & Software*, 2001, 16(3): 195-231.
- [29] ROSSMAN L, HUBER W. Storm water management model reference manual volume III-water quality[R]. Washington: U. S. EPA Office of Research and Development, 2016.

[30] 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):30-35. (ZHU Song,LIU Guohua,WANG Lizhong,et al. A Bayesian approach for the identification of pollution source in water quality model coupled with hydrodynamics[J]. Journal of Sichuan University (Engineering Science Edition),2009,41(5):30-35. (in Chinese))

[31] VRUGT J A. Markov chain Monte Carlo simulation using the DREAM software package: theory, concepts, and MATLAB implementation[J]. Environmental Modelling & Software,2016,75:273-316.

[32] KASS R E,CARLIN B P,GELMAN A,et al. Markov chain Monte Carlo in practice;a roundtable discussion[J]. The American Statistician,1998,52(2):93-100.

[33] HASTINGS W K. Monte Carlo sampling methods using Markov chains and their applications[J]. Biometrika,1970,57(1):97-109.

[34] 王鹏,董国庆,阴祖荣,等. 面向复杂情景的 SWMM 水质参数敏感性分析[J]. 河海大学学报(自然科学版),2023,51(5):31-37. (WANG Peng,DONG Guoqing,YIN Zurong,et al. Sensitivity analysis of water quality parameters in SWMM for complex scenarios[J]. Journal of Hohai University (Natural Sciences),2023,51(5):31-37. (in Chinese))

[35] RIAÑO-BRICEÑO G,BARREIRO-GOMEZ J,RAMIREZ-JAIME A,et al. MatSWMM-an open-source toolbox for designing real-time control of urban drainage systems[J]. Environmental Modelling & Software,2016,83:143-154.

[36] 栾广学,侯精明,王添,等. 基于复杂高密度管网结构优化技术的银川城区排水过程高效模拟方法[J]. 水资源保护,2023,39(3):205-212. (LUAN Guangxue,HOU Jingming,WANG Tian,et al. Efficient simulation method of Yinchuan urban drainage process based on structure optimization technology of complex high-density pipe network[J]. Water Resources Protection,2023,39(3):205-212. (in Chinese))

[37] 廖如婷,徐宗学,叶陈雷,等. 基于 SWMM 和 InfoWorks ICM 模型的大红门排水区暴雨内涝模拟[J]. 水资源保护,2023,39(3):109-117. (LIAO Ruting,XU Zongxue,YE Chenlei,et al. Simulation of rainstorm waterlogging in Dahongmen drainage area based on SWMM and InfoWorks ICM models[J]. Water Resources Protection,2023,39(3):109-117. (in Chinese))

[38] 宋耘,李琼芳,牛铭媛,等. 基于 SWMM 模型的南京典型易涝区暴雨内涝模拟[J]. 水利水电科技进展,2019,39(6):56-61. (SONG Yun, LI Qiongfang, NIU Mingyuan, et al. Rainstorm and waterlogging simulation in typical inundated districts of Nanjing based on SWMM[J]. Advances in Science and Technology of Water Resources,2019,39(6):56-61. (in Chinese))

[39] YAZDI J. Water quality monitoring network design for urban drainage systems, an entropy method[J]. Urban Water Journal,2018,15(3):227-233.

[40] BANIK B K,ALFONSO L,DI CRISTO C,et al. Evaluation of different formulations to optimally locate sensors in sewer systems[J]. Journal of Water Resources Planning and Management,2017,143(7):04017026.

[41] 朱嵩,毛根海,刘国华,等. 改进的 MCMC 方法及其应用[J]. 水利学报,2009,40(8):1019-1023. (ZHU Song,MAO Genhai,LIU Guohua,et al. Improved MCMC method and its application[J]. Journal of Hydraulic Engineering,2009,40(8):1019-1023. (in Chinese))

(收稿日期:2024 - 04 - 15 编辑:刘晓艳)