

城市雨水管理

DOI: 10. 19853/j. zgjsps. 1000-4602. 2024. 05. 020

基于机器学习理论的SWMM参数自动率定方法

康得军^{1,2}, 赖李保壹³, 邱福杰⁴, 温儒杰⁵, 吴端炜⁶, 赵颖⁷

(1. 福州大学 土木工程学院, 福建 福州 350108; 2. 福州大学晋江科教园, 福建 晋江 362251; 3. 福建省龙岩市长汀县濯田镇政府, 福建 龙岩 366300; 4. 华东勘测设计院 <福建>有限公司, 福建 福州 350001; 5. 中冶京诚工程技术有限公司, 北京 100176; 6. 中国市政工程中南设计研究总院有限公司, 湖北 武汉 430010; 7. 中国环境科学研究院, 北京 100012)

摘要: SWMM作为一种模拟降雨情况的软件,由于其模型参数在经验范围内选取的不确定性,随着模型不断地向前运行演绎,误差会持续累积,最终影响模型的模拟结果,而传统的参数率定方法则具有效率低、准确度差等缺点。为此,利用MATLAB软件调用SWMM的水力演算程序来完成两者的数据交互与集成,结合多元逐步回归法筛选出敏感性高的参数作为率定对象,进而通过数值实验的方式以敏感性分析结果为依据使用Bayes-MH机器学习算法实现SWMM参数的自动率定。对于不同的实测降雨场景,自动率定的结果均能够顺利输出,参数匹配度达95%以上,表明了该方法具备优秀的稳定性、自适应性与全范围寻参能力。

关键词: 机器学习理论; Metropolis-Hastings算法; 雨洪管理模型(SWMM); MATLAB; 参数自动率定

中图分类号: TU992 **文献标识码:** A **文章编号:** 1000-4602(2024)05-0122-08

Automatic Calibration Method of SWMM Parameters Based on Machine Learning Theory

KANG De-jun^{1,2}, LAI Li-bao-yi³, QIU Fu-jie⁴, WEN Ru-jie⁵, WU Duan-wei⁶, ZHAO Ying⁷

(1. College of Civil Engineering, Fuzhou University, Fuzhou 350108, China; 2. Fuzhou University Jinjiang Science and Education Park, Jinjiang 362251, China; 3. Zhuotian Town Government of Changting County in Longyan City of Fujian Province, Longyan 366300, China; 4. Huadong Engineering <Fujian> Corporation, Fuzhou 350001, China; 5. Capital Engineering & Research Incorporation Limited, Beijing 100176, China; 6. Central & Southern China Municipal Engineering Design and Research Institute Co. Ltd., Wuhan 430010, China; 7. Chinese Research Academy of Environmental Sciences, Beijing 100012, China)

Abstract: As a kind of software for simulating rainfall situation, errors of SWMM will accumulate continuously with the operation of the model, and ultimately affect the simulation results of the model due to the uncertainty in the selection of model parameters within the empirical range. The traditional

基金项目: 北京建筑大学未来城市设计高精尖创新中心开放课题(UDC2017032512)

通信作者: 赵颖 E-mail: zhaoying@craes.org.cn

parameter calibration methods have the disadvantages such as low efficiency and poor accuracy. Therefore, this paper completed the data interaction and integration between MATLAB and SWMM by using MATLAB to call the hydraulic calculus program of SWMM, selected the parameters with high sensitivity as the calibration object by using multiple stepwise regression method, and eventually realized the automatic calibration of SWMM parameters by using Bayes-MH machine learning algorithm and numerical experiment based on the sensitivity analysis. In different measured rainfall scenarios, the results of automatic calibration were output smoothly, and the parameter matching degree was more than 95%, indicating that the method had excellent stability, self-adaptability and full-range parameter searching ability.

Key words: machine learning theory; Metropolis-Hastings algorithm; storm water management model (SWMM); MATLAB; automatic parameter calibration

利用软件模型模拟城市降雨径流相关的水文水力过程,从而对城市水环境进行评估和管理是重要的研究手段,雨洪管理模型(SWMM)是当前众多主流动态降雨-径流模型之一^[1]。在完成SWMM模型的初步概化后,应进一步对其进行水文水力参数的敏感性分析,选取出对模型输出结果影响较大的若干个敏感性参数,然后对其进行参数率定,从而保证模型的准确性与稳定性^[2]。

参数率定的方法主要分为两类,一类是传统的人工手动率定,另一类是基于程序算法的自动率定。传统的人工手动率定方法,通常情况下需要操作者对参数进行多次的手动调整^[3],无法对不同研究区域产生完全的自适应性,与此同时,调参过程的顺利与否还受限于调参人员对相关参数的认识与理解,因此增加了率定结果产生误差的可能性,随之不断地累积,最终影响模型的输出结果。为解决传统率定方法所具有的模糊性、盲目性与不稳定性等问题,提高参数率定过程的效率与准确度,诸多研究人员提出了基于各种理论参数自动率定方法。目前主要的自动率定算法包括遗传算法、HS算法、人工神经网络算法等^[4],存在着率定结果准确度较低、环境自适应性较差等缺点,同时由于SWMM模型计算的非线性特征,在参数的经验取值范围内可能存在多个局部解。为此,笔者提出了一种基于Bayes机器学习理论参数自动率定方法,可以将先验资料与后验资料进行有效地综合,既可以避免复杂的计算与仅利用先验资料可能带来的主观误差,亦可避免使用后验资料可能带来的数据噪音影响,能够有效地提高参数率定的效率、适用

性以及准确度,解决传统自动率定方法无法遍历所有解集空间和易陷入局部解循环的问题^[5-6]。最后以泉州某滨海城区为研究对象,根据实际降雨数据,以数值实验的形式对水文水力参数的自动率定方法加以验证。

1 研究方法

1.1 机器学习理论

机器学习是指计算机通过编制的算法模拟人脑学习的行为,其核心就是通过新信息与旧信息两者不断地综合,使自身对外部事物的认知产生持续性的改善,即“认识-实践-再认识”的意识过程,而Bayes学习是机器学习理论中的一个重要分支^[7]。Bayes机器学习理论利用先验资料与样本资料对未知的参数总体进行求解,而概率密度分布则是两者在学习过程中具体的表达方式。Bayes定理记作:

$$P(\theta|D) = \frac{P(\theta)P(D|\theta)}{P(D)} \quad (1)$$

式中: θ 为需率定的未知参数向量; D 为给定的观测向量; $P(\theta)$ 为未知参数向量的先验概率密度分布函数; $P(D|\theta)$ 为似然函数; $P(D)$ 为归一化常数; $P(\theta|D)$ 为未知参数向量的后验概率密度分布函数。

Bayes机器学习理论主要包括3个基本变量:先验资料对应的先验概率密度分布函数、样本资料对应的似然函数、后验资料对应的后验概率密度分布函数。

① 先验概率密度分布函数

先验概率密度分布通常来自于研究者的主观认识和专家经验,与给定的观测向量(样本资料)无关,用于描述参数在指定区间内的位置信息。根据

Bayes假设,函数的表达式在参数的取值范围内是均匀分布的,则先验概率密度函数具体形式如下:

$$P(\theta) = \prod_{i=1}^n P(\theta_i) \quad (2)$$

$$P(\theta_i) = \begin{cases} \frac{1}{b_i - a_i} & \theta_i \in [a_i, b_i] \\ 0 & \text{else} \end{cases} \quad (3)$$

式中: i 为第 i 个需要率定的参数; n 为需要率定的参数总个数; a_i 为第 i 个待率定参数的取值范围上限; b_i 为第 i 个待率定参数的取值范围下限; $P(\theta_i)$ 为第 i 个未知参数在其取值范围内的均匀分布概率密度函数。

② 似然函数

似然函数主要用于参数的估计,其具体形式会对参数自动率定结果的准确性与稳定性产生深远影响,因此构造一个合理的似然函数形式是至关重要的环节。

此外, Bayes理论能够将实际观测过程中可能产生的数据噪音等因素纳入考虑,根据误差理论,观测误差以及模型误差通常服从均值为0、标准差为 σ 的正态分布,因此构建的似然函数具体形式如下:

$$P(D|\theta) = \prod_{j=1}^k \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{[T_j(\theta) - D_j]^2}{2\sigma^2}\right\} \quad (4)$$

式中: j 为模拟或观测时的第 j 个时刻; k 为观测的总时刻数; σ 为似然函数的标准差; $T_j(\theta)$ 为未知参数组 $(\theta_1, \theta_2, \dots, \theta_i, \dots, \theta_n)$ 在第 j 个时刻所对应的SWMM水文水力模拟值; D_j 为第 j 个时刻所对应的观测值。

③ 后验概率密度分布函数

将式(2)、式(3)与式(4)代入式(1)得到后验概率密度分布函数,如式(5)所示,其函数值表示根据先验资料与观测向量所求解出的参数初始取值范围内各取值的概率大小情况,即参数率定结果在区间范围内取不同值时的可能性。

$$P(\theta|D) = \prod_{i=1}^n P(\theta_i) \prod_{j=1}^k \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{[T_j(\theta) - D_j]^2}{2\sigma^2}\right\} \quad (5)$$

1.2 Metropolis-Hastings(MH)算法

Bayes机器学习理论的后验概率密度分布函数

就是参数自动率定结果的概率密度分布函数,为获取未知参数的分布函数则需要通过MH算法进行求解,其核心思想是构造一条随机且能够收敛到指定概率分布的马尔科夫链,最后根据细致平稳条件进行采样求解,该算法对SWMM的多未知参数高维空间尤为有效^[8],过程如下:

① 旧状态值 θ_i 根据游走步长在待率定参数的先验取值范围内进行随机抽样并产生新状态值 θ^* 。为此,引入与后验概率密度分布函数具有相同形态的提议密度分布函数 $G(\theta, w^2)$,并将其标准差 w 作为衡量游走步长的指标对率定参数在初始取值范围内进行随机抽样,其原因是标准差 w 的大小影响当前参数值 θ_i 所对应的提议密度分布函数的离散程度,标准差越大,函数形态越离散,则在函数分布范围内所能随机抽取到的新状态采样值 θ^* 的范围就越大;标准差越小,函数形态越集中,则在函数分布范围内所能随机抽取到的新状态采样值 θ^* 的范围就越小。此外,若游走步长过大,会导致 θ^* 相较于 θ_i 的变化幅度过大,使得收敛难度加大;游走步长过小,会导致 θ^* 相较于 θ_i 的变化幅度过小,从而易使抽样值陷入局部解中,导致率定结果的准确度降低。多个参数率定时,游走步长的大小依据各参数自身初始取值范围所对应的精度并结合抽样误差 e 进行实时动态调整,最终通过MATLAB代码编程实现。

② 计算新状态的采样参数值 θ^* 被接受的概率,由细致平稳条件可得:

$$A(\theta_i, \theta^*) \triangleq \min \left\{ 1, \frac{P(\theta^*)G(\theta^*|\theta_i)}{P(\theta_i)G(\theta_i|\theta^*)} \right\} = \min \left\{ 1, \frac{P(\theta^*)}{P(\theta_i)} \right\} = \alpha \quad (6)$$

式中: A 为接受概率; P 为似然函数; θ_i 为旧状态 t 时刻对应的参数值; θ^* 为新状态对应的参数值; G 为提议密度分布函数,函数形态为正态分布, $G(\theta^*|\theta_i)$ 是指均值为 θ_i 时抽取 θ^* 的概率, $G(\theta_i|\theta^*)$ 是指均值为 θ^* 时抽取 θ_i 的概率,与 $G(\theta_i|\theta^*)$ 两者关于 $(\theta_i + \theta^*)/2$ 对称分布, $G(\theta^*|\theta_i)$ 与 $G(\theta_i|\theta^*)$ 的函数值相等,两者相除抵消,从而使接受概率 A 的解得到简化。

③ 从0~1的均匀分布 $\gamma \sim \text{Uniform}[0, 1]$ 进行采样,若 $\gamma < \alpha$,则 $t+1$ 时刻的参数值接受新状态的采样参数值 θ^* ,否则拒绝新状态的采样参数值 θ^* ,保

持旧状态的参数值 θ_i ,即 $\theta_{i+1} = \theta_i$ 。

④ 重复步骤②和③,直至马尔科夫链完成收敛。

MH算法满足以下任意一个条件时即可达到收敛状态,完成机器学习,得到参数的自动率定结果:收敛条件一,SWMM模拟值与实际观测值之间的抽样误差 e 为零,表明此时SWMM模拟值与实际观测值相等,没有抽样误差,则参数候选值与参数率定结果相等;收敛条件二,机器学习次数达到150次以上,抽样误差 $e < 10$,同时对于当前学习次数而言,本次参数接受值与之前50次学习产生的50组参数接受值的平均值之间的变化率,即收敛偏差,不大于0.5%,且所有待率定参数均达到收敛状态时认为完成学习,率定结束。

$$e = \sqrt{\frac{\sum_{j=1}^k [T_j(\theta) - D_j]^2}{k}} \quad (7)$$

1.3 SWMM模型与MATLAB软件的集成技术

SWMM软件对输入和输出的文件的类型与格式有着严苛的要求。SWMM软件的输入文件类型*.inp包含了SWMM运行需要的所有参数数据与空间资料,以*.txt的格式储存。*.rpt与*.out文件则是SWMM输入*.inp文件执行后生成的两种输出文件类型,*.rpt文件是对模拟过程的报告总结,以*.txt格式储存;*.out文件包含了SWMM模拟结果的时间序列,以二进制格式储存,数据结构复杂。

SWMM是一个开源软件,用户能很方便地利用MATLAB对其进行二次开发,集成路线见图1。

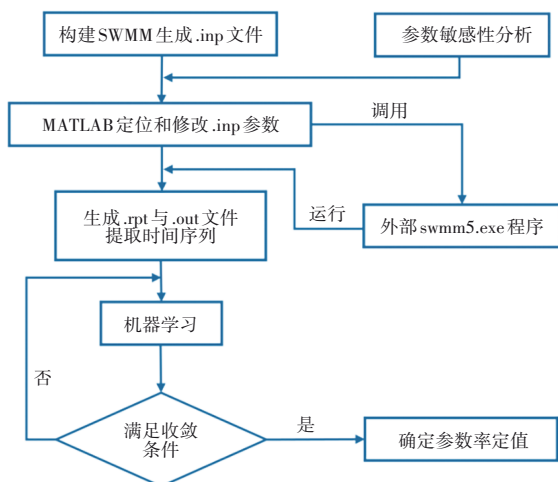


图1 SWMM与MATLAB的集成路线

Fig.1 SWMM and MATLAB integration roadmap

通过MATLAB的程序代码编写,将生成的参数候选值替换至前一个*.inp文件的相应位置,并利用system函数与SWMM核心代码直接生成*.rpt文件与*.out文件,逐层读取以二进制形式保存在*.out文件的时间序列数据,从而有效解决*.rpt与*.out文件模拟数据读取困难、工作量大等问题。

2 模型构建

2.1 研究区域概况

泉州市位于福建省东南部沿海,季风气候显著,降水充沛,但时空分布不均匀,全市年降水量为1 000~1 800 mm,自东南部向西北部递增,2021年泉州市的常住人口城镇化率达到了69.7%,城镇化水平较高。选取泉州市丰泽区南部为研究区域,该区域位于北滨江公园汀洲园以北,区域内有一条自东北向西南贯通的交通干道,干道两侧建有工业园、商业广场、法院、中学以及住宅小区,占地面积为83.9 hm²,其中不透水面积为63.7 hm²,不透水率为75.92%。

2.2 SWMM模型概化

基于ArcGIS平台完成对研究区域的子汇水区域划分、下垫面等资料的提取,并根据当地管网数据资料以及前期水文分析结果,通过自开发软件GIS to Excel to SWMM (GES)完成GIS数据结果向SWMM程序的导入。最终,研究区域概化为由126个子汇水区、132个管段、128个节点、12个雨水排放口构建成的SWMM模型,如图2所示。选取2019年5月17日的实测降雨作为降雨时间序列输入SWMM软件执行模拟,其降雨量为73.5 mm,降雨历时为2 h,如图3所示。

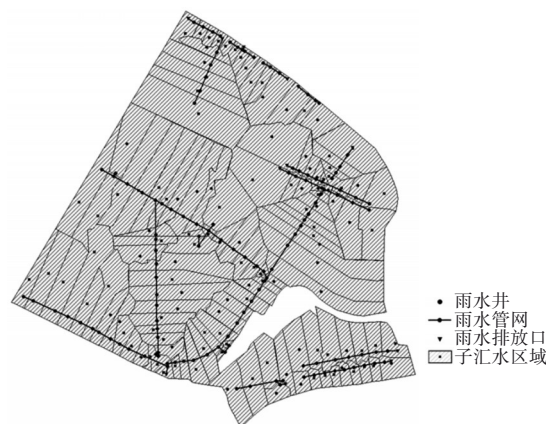


图2 研究区域概化

Fig.2 Generalized map of study area

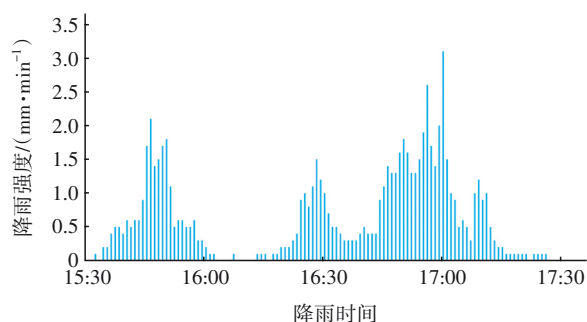


图3 2019年5月17日实测降雨

Fig.3 Measured rainfall on May 17, 2019

2.3 率定参数的选取

在自动率定参数之前应进行参数的敏感性分析,筛选出对输出结果影响较大的若干个敏感性参数作为优化前提,以提高参数率定的整体速度。通过多元逐步回归法对目标函数管网排放总量和地表径流总量进行分析,得出5个具有显著敏感性的参数^[9],其初始取值范围参考相关文献经验值与模型手册推荐值加以确定,如表1所示。

表1 待率定参数及初始取值范围

Tab.1 Uncalibrated parameters and initial range

参数	物理含义	初始取值范围	目标函数
N-Imperv	不透水区域曼宁系数	0.011~0.024	管网排放总量
PctZero/%	无洼地不透水面积百分比	5~85	管网排放总量
Roughness	管道曼宁系数	0.009~0.015	管网排放总量
S-Imperv/mm	不透水区域洼蓄深度	1.27~2.54	地表径流总量
MinRate/(mm·h ⁻¹)	最小入渗率	1~10	地表径流总量

3 参数自动率定与结果分析

3.1 水文水力参数的率定步骤

① 根据研究区域现铺设管网资料、数字地形高程模型(DEM),基于ArcGIS平台完成对其SWMM模型的概化与确定性参数的提取,并通过GES软件生成初始*.inp文件。

② 使用MATLAB软件打开需要率定的初始*.inp文件,完成率定参数的定位修改与保存。

③ 通过MATLAB调用外部的swmm5.exe应用程序,批量运行更新后的*.inp文件以完成对多个未知参数的率定,并将*.inp批量运行生成的*.rpt与*.out文件储存,以供后续调用。

④ 通过fseek、fread等命令对*.rpt和*.out文

件中的相关水文水力模拟资料进行读取,结合输入的观测向量数据进行Bayes机器学习。

⑤ 根据设定好的蒙特卡洛马尔科夫链规则对参数进行抽样,并通过MH算法判断是否接受新状态的采样值,若接受则替换旧*.inp文件并批量运行获得相应的*.rpt和*.out文件,否则继续抽样。

⑥ 通过若干次对观测向量数据的机器学习,循环运行至满足收敛条件,最终完成参数的自动率定并输出结果。

3.2 水文水力参数的自动率定与验证分析

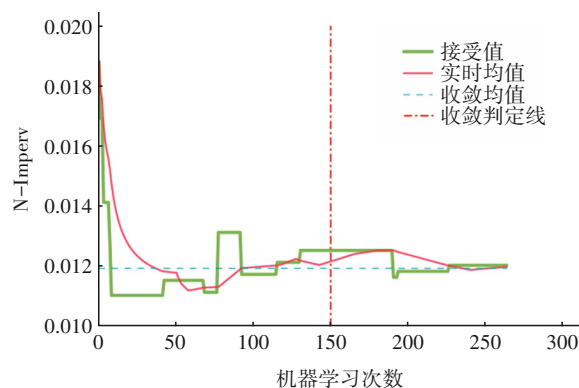
数值实验可以有效地对不同工况进行反复的计算机模拟,并对参数加以精确控制。因此,本研究在遵循数值实验数据产生流程的基础上,根据2.3节的参数敏感性分析结果,水文水力参数在先验取值范围内,N-Imperv取0.012、PctZero取50%、Roughness取0.014、S-Imperv取1.43 mm、MinRate取3 mm/h。输入以上参数值对SWMM模型执行正向模拟,并将正向模拟结果从*.rpt与*.out文件中提取出来作为目标函数的观测时间序列数据,对基于机器学习理论的自动率定方法进行验证。

① 根据初始取值范围以及观测时间序列(见表2)对目标函数为管网排放总量的参数进行自动率定,过程如图4所示。

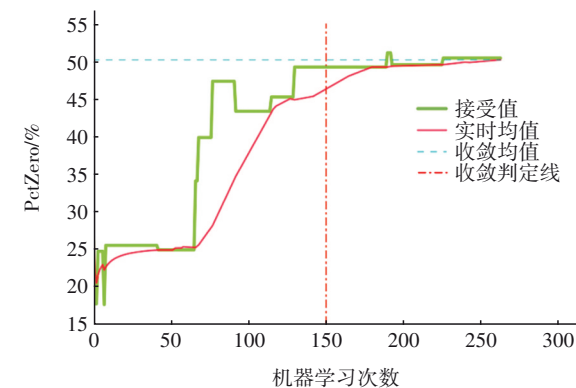
表2 目标函数管网排放总量的观测时间序列数据

Tab.2 Time series of total pipe network emission

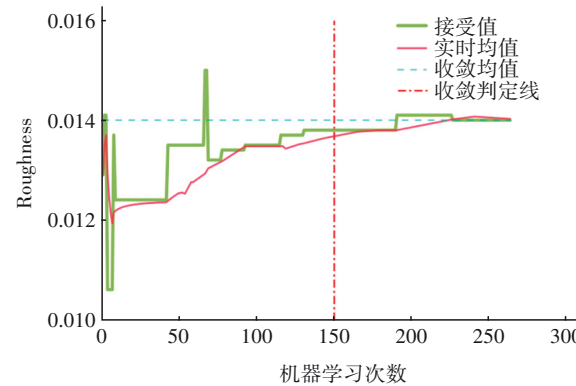
观测时刻数	观测时间	出流量/(m ³ ·s ⁻¹)	观测时刻数	观测时间	出流量/(m ³ ·s ⁻¹)
1	00:45	12.33	6	02:00	22.60
2	01:00	8.03	7	02:15	10.36
3	01:15	8.77	8	02:30	5.47
4	01:30	13.86	9	02:45	3.34
5	01:45	27.15	10	03:00	2.13



a. N-Imperv的自动率定过程



b. PctZero的自动率定过程



c. Roughness的自动率定过程

图4 管网排放总量的参数率定

Fig.4 Parameter calibration of total pipe network emission

在图4中,绿色实线是参数当前接受值的率定过程线,即经机器学习后所接受的参数值构成的折线;红色实线是参数均值曲线,代表第1次机器学习至最后1次机器学习过程中参数接受值的平均值变化情况;蓝色水平虚线是参数均值线,即相对当前学习次数而言,前50次机器学习产生的50组参数接受值的平均值,若机器学习次数未50次,则以当前机器学习总次数的参数接受值的平均值作为取值;红色竖直点划线是参数收敛与否的判断起始线,即开始判断参数是否收敛的学习次数位置,本研究设定机器学习次数达到150次时计算收敛偏差,并判断是否满足收敛条件。

目标函数为管网排放总量的待率定参数经过264次机器学习后成功收敛,参数的抽样误差 e 为0.91972,收敛偏差均小于0.5%,符合收敛条件二,输出结果如表3所示,说明参数的自动率定值与假定的参数真实值基本吻合,参数率定值所对应的目标函数的时间序列与观测时间序列基本一致。

表3 目标函数管网排放总量的参数率定值

Tab.3 Parameter calibration values of total pipe network emission

项目	N-Imperv	PctZero/%	Roughness
观测值	0.012	50	0.014
率定结果	0.01195	50.399	0.01403
收敛均值	0.0119	50.166	0.014

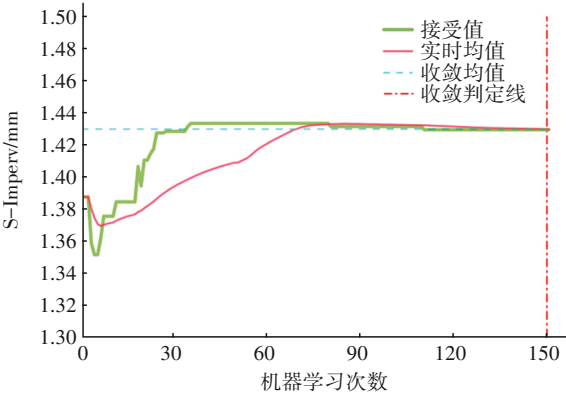
注: N-Imperv、PctZero、Roughness的收敛偏差分别为0.42%、0.47%、0.21%,率定误差分别为0.42%、0.80%、0.21%。

根据初始取值范围以及观测时间序列(见表4)对目标函数为地表径流总量的参数进行自动率定,过程如图5所示。

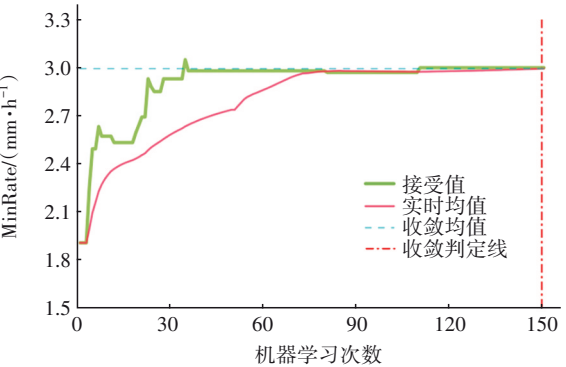
表4 目标函数地表径流总量的观测时间序列数据

Tab.4 Time series of total runoff

观测时刻数	观测时间	出流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)	观测时刻数	观测时间	出流量/ ($\text{m}^3 \cdot \text{s}^{-1}$)
1	00:45	19.26	6	02:00	30.87
2	01:00	7.25	7	02:15	10.58
3	01:15	12.80	8	02:30	4.24
4	01:30	17.82	9	02:45	2.15
5	01:45	44.40	10	03:00	1.28



a. S-Imperv的自动率定过程



b. MinRate的自动率定过程

图5 地表径流总量的参数率定

Fig.5 Parameter calibration of total runoff

目标函数为地表径流总量的待率定参数经过151次机器学习后成功收敛,参数的抽样误差 e 为0.086 274,收敛偏差均小于0.5%,符合收敛条件二,输出结果如表5所示,说明参数的自动率定值与假定的参数真实值基本吻合,参数率定值所对应的目标函数的时间序列与观测时间序列基本一致。

表5 目标函数地表径流总量的参数率定值

Tab.5 Parameter calibration values of total runoff

参数名称	S-Imperv/mm	MinRate/(mm·h ⁻¹)
观测值	1.43	3
率定结果	1.429	2.999
收敛均值	1.429 4	2.994
注: S-Imperv、MinRate的收敛偏差分别为0.03%、0.17%,率定误差分别为0.07%、0.03%。		

② 为避免单场降雨以及降雨强度所带来的各种偶然性因素影响,本研究采用另外两场实测降雨作为降雨时间序列输入SWMM软件执行模拟,对所提出的自动率定方法做进一步检验。

对于2020年6月8日的实测降雨数据(降雨量为32.5 mm),目标函数为管网排放总量的待率定参数经过191次机器学习后成功收敛,参数抽样误差 e 为6.112 7,收敛偏差均小于0.5%,符合收敛条件二;目标函数为地表径流总量的待率定参数经过90次机器学习后成功收敛,参数的抽样误差 e 为0,符合收敛条件一,机器学习次数小于150次,收敛偏差不做计算。输出结果如表6所示。

表6 2020年6月8日的率定值

Tab.6 Calibration values on June 8, 2020

目标函数	参数	观测值	率定结果	收敛均值
管网排放 总量	N-Imperv	0.012	0.012 17	0.012 2
	PctZero/%	50	50	50.108
	Roughness	0.014	0.013 84	0.013 9
地表径流 总量	S-Imperv/mm	1.43	1.43	1.487
	MinRate/(mm·h ⁻¹)	3	2.92	2.67
注: N-Imperv、PctZero、Roughness的收敛偏差分别为0.25%、0.22%、0.43%,率定误差分别为1.42%、0.00%、1.14%;S-Imperv、MinRate的率定误差分别为0.00%、2.67%。				

对于2021年2月10日的实测降雨数据(降雨量为45.0 mm),目标函数为管网排放总量的待率定参数经过151次机器学习后成功收敛,参数抽样误差 e 为0.815 31,收敛偏差均小于0.5%,符合收敛条件二;目标函数为地表径流总量的待率定参数经过

121次机器学习后成功收敛,参数抽样误差 e 为0,符合收敛条件一,机器学习次数小于150次,收敛偏差不做计算。输出结果如表7所示。

表7 2021年2月10日的率定值

Tab.7 Parameter calibration values on February 10, 2021

目标函数	参数名称	观测值	率定结果	收敛均值
管网排放 总量	N-Imperv	0.012	0.012 1	0.012 16
	PctZero/%	50	51.5	51.245
	Roughness	0.014	0.014	0.013 94
地表径流 总量	S-Imperv/mm	1.43	1.43	1.632
	MinRate/(mm·h ⁻¹)	3	2.85	2.84
注: N-Imperv、PctZero、Roughness的收敛偏差分别为0.49%、0.50%、0.43%,率定误差分别为0.83%、3.00%、0.00%;S-Imperv、MinRate的率定误差分别为0.00%、5.00%。				

由表6和表7可知,对于不同的实测降雨情况,参数的机器学习率定算法亦能够满足收敛条件并完成参数率定,体现了该方法的自适应性与稳定性。相比于文献[3]和[10]中的传统人工手动率定与自动率定方法,本研究中参数的率定值误差更低,率定结果匹配度达95%以上,对于所需率定参数的连续性取值范围而言具有相当高的准确度,同时由于学习过程中采用了实时调整的动态游走步长,参数率定结果不易陷入局部解,大大减少了机器学习的次数,具备更加优秀的参数全范围搜寻能力与率定效率。

4 结论

① 根据机器学习理论确定先验概率密度分布函数与似然函数,并得出后验概率密度分布函数,利用MATLAB软件从外部调用SWMM程序,并通过代码编写读取*.out文件中的时间序列数据作为输入的观测向量,进而通过MH算法对其进行求解,经若干次的抽样学习后得到满足收敛条件的后验概率密度分布函数,最终基于此密度函数得出SWMM参数的自动率定结果,为SWMM模型的参数自动率定提供了新思路。

② 利用建立的机器学习模型,输入3场实测降雨,依据输入的目标函数观测时间序列数据进行参数率定,各参数均能够顺利收敛并输出结果,体现出机器学习方法的稳定性与自适应性。此外,采用了动态游走步长的参数率定方法,各参数的率定值匹配度达95%以上,率定结果不易陷入局部解而

无法收敛,表明该方法具有出色的参数校准能力与全范围寻参能力。

参考文献:

- [1] 陈韬,夏蒙蒙,刘云鹏,等. 基于SWMM的海绵改建小区雨水径流调控研究[J]. 中国给水排水, 2020, 36(11): 103-111.
CHEN Tao, XIA Mengmeng, LIU Yunpeng, *et al.* Regulation of stormwater runoff in sponge reconstructed community based on SWMM [J]. China Water & Wastewater, 2020, 36(11): 103-111 (in Chinese).
- [2] 杨森雄,卿晓霞,朱韵西. 一种耦合SWMM计算的参数自动率定算法及实现[J]. 给水排水, 2021, 47(1): 148-154.
YANG Senxiong, QING Xiaoxia, ZHU Yunxi. An auto-calibration method by coupling SWMM with optimization algorithm and its implementation [J]. Water & Wastewater Engineering, 2021, 47(1): 148-154 (in Chinese).
- [3] 刘兴坡. 基于径流系数的城市降雨径流模型参数校准方法[J]. 给水排水, 2009, 35(11): 213-217.
LIU Xingpo. Parameter calibration method for urban rainfall-runoff model based on runoff coefficient [J]. Water & Wastewater Engineering, 2009, 35(11): 213-217 (in Chinese).
- [4] 袁绍春,李迪,陈焱,等. 基于BP神经网络算法的SWMM参数自动率定方法[J]. 中国给水排水, 2021, 37(21): 125-130.
YUAN Shaochun, LI Di, CHEN Yao, *et al.* Automatic calibration procedure of storm water management model parameters based on back propagation neural network algorithm [J]. China Water & Wastewater, 2021, 37(21): 125-130 (in Chinese).
- [5] 朱军,胡文波. 贝叶斯机器学习前沿进展综述[J]. 计算机研究与发展, 2015, 52(1): 16-26.
ZHU Jun, HU Wenbo. Recent advances in Bayesian machine learning[J]. Journal of Computer Research and Development, 2015, 52(1): 16-26 (in Chinese).
- [6] 杨海东,刘碧玉,黄建华. 基于改进 Bayesian-MCMC 的突发水污染事件预测模型参数率定方法[J]. 控制与决策, 2018, 33(4): 679-686.
YANG Haidong, LIU Biyu, HUANG Jianhua. Forecast model parameters calibration method for sudden water pollution accidents based on improved Bayesian-Markov chain Monte Carlo [J]. Control and Decision, 2018, 33(4): 679-686 (in Chinese).
- [7] 何立杰,何洪林,任小丽,等. 基于贝叶斯机器学习的生态模型参数优化方法研究[J]. 地球信息科学学报, 2017, 19(10): 1270-1278.
HE Lijie, HE Honglin, REN Xiaoli, *et al.* Parameters optimization method of ecosystem model based on Bayesian machine learning [J]. Journal of Geo-information Science, 2017, 19(10): 1270-1278 (in Chinese).
- [8] 陈平,徐若曦. Metropolis-Hastings 自适应算法及其应用[J]. 系统工程理论与实践, 2008, 28(1): 100-108.
CHEN Ping, XU Ruoxi. Metropolis-Hastings adaptive algorithm and its application [J]. Systems Engineering-Theory & Practice, 2008, 28(1): 100-108 (in Chinese).
- [9] 李伟雨. 基于SWMM模型的悦来会展城雨洪模拟研究[D]. 重庆: 重庆大学, 2018: 64-68.
LI Weiyu. Study on Stormwater Simulation in Convention and Exhibition City of Yuelai Based on SWMM [D]. Chongqing: Chongqing University, 2018: 64-68 (in Chinese).
- [10] GEEM Z W, KIM J H, LOGANATHAN G V. A new heuristic optimization algorithm: harmony search [J]. Simulation, 2001, 76(2): 60-68.

作者简介:康得军(1981-),男,湖北十堰人,博士,副教授,从事海绵城市建设理论与技术研究。

E-mail: dj kang@fzu.edu.cn

收稿日期:2022-06-01

修回日期:2022-08-20

(编辑:刘贵春)