

DOI:10.3876/j.issn.1000-1980.2022.02.001

基于 EnKF 法的径流数据同化对 SWAT 模型 参数优化效果评估

刘永伟¹,王 文²,刘元波¹,刘 庆³,凌 哲⁴

(1. 中国科学院南京地理与湖泊研究所流域地理学重点实验室,江苏 南京 210008;

2. 河海大学水文水资源与水利工程科学国家重点实验室,江苏 南京 210098;

3. 江西省水利规划设计研究院,江西 南昌 330029; 4. 江苏省水利工程规划办公室,江苏 南京 210029)

摘要: 基于集合卡尔曼滤波(EnKF)法,在合理量化模型模拟和径流观测误差、有效处理模型参数演变及过拟合问题的基础上,以淮河上游淮滨水文站以上流域为研究区,构建了基于径流数据同化的 SWAT 分布式水文模型参数优化方案,就站点实测径流数据同化对模型参数的优化效果进行了评估。结果显示,数据同化过程中,被更新的模型参数集合逐渐收敛并趋于稳定,基于稳定后的参数获得的模拟径流与实测径流过程接近,流域出口径流模拟的纳什效率系数可达 0.88,说明基于 EnKF 法的径流数据同化对 SWAT 模型参数具有一定的优化估计能力,且采用数据同化方式进行模型参数率定具有一定的可行性。

关键词: EnKF 法; SWAT 模型; 径流; 数据同化; 参数优化; 淮河

中图分类号: TU122 **文献标志码:** A **文章编号:** 1000-1980(2022)02-0001-10

Evaluation of the optimization effect of streamflow data assimilation on SWAT model parameters based on the EnKF approach

LIU Yongwei¹, WANG Wen², LIU Yuanbo¹, LIU Qing³, LING Zhe⁴

(1. Key Laboratory of Watershed Geography Sciences, Nanjing Institute of Geography and Limnology,
Chinese Academy of Sciences, Nanjing 210008, China;

2. State Key Laboratory of Hydrology-Water Resources and Hydraulic Engineering, Hohai University, Nanjing 210098, China;

3. Jiangxi Provincial Water Conservancy Planning Design and Research Institute, Nanchang 330029, China;

4. Jiangsu Water Conservancy Engineering Planning Office, Nanjing 210029, China)

Abstract: In this study, a hydrological model parameter optimization scheme was constructed in SWAT based on streamflow assimilation using the ensemble Kalman filter(EnKF) approach, in which the model simulation and streamflow observation errors were reasonably quantified and the model parameter evolution and over-fitting issues were effectively treated. Based on this scheme, the capacity of streamflow assimilation on SWAT model parameter optimization was evaluated in the upper Huai River above Huaibin hydrological station. The results showed that the updated parameter ensemble gradually converged and stabilized during the data assimilation process. The runoff process based on the stabilized parameter ensemble was close to the measured runoff. The deterministic coefficient at the outlet of the basin reached 0.88. The results indicated that the EnKF based streamflow assimilation had the ability to optimize the parameters in SWAT. It was feasible to calibrate model parameters using data assimilation method.

Key words: EnKF method; SWAT model; streamflow; data assimilation; parameter optimization; Huai River

数据同化是将多源观测信息应用于模型模拟预报的一种有效手段,被公认是提高模型模拟预报精度、量化不确定性方面最具前途的方法^[1-2]。径流是流域产汇流模拟中最为重要的输出变量,可以反映整个流域或

基金项目: 国家重点研发计划(2018YFE0105900);江苏省科技计划青年项目(BK20191097);国家自然科学基金(41901049)

作者简介: 刘永伟(1987—),女,助理研究员,博士,主要从事水文遥感数据同化与气象水文干旱研究。E-mail: ywliu@niglas.ac.cn

通信作者: 刘元波(1969—),男,研究员,主要从事水文遥感研究。E-mail: ybliu@niglas.ac

引用本文: 刘永伟,王文,刘元波,等. 基于 EnKF 法的径流数据同化对 SWAT 模型参数优化效果评估[J]. 河海大学学报(自然科学版), 2022, 50(2): 1-10.

LIU Yongwei, WANG Wen, LIU Yuanbo, et al. Evaluation of the optimization effect of streamflow data assimilation on SWAT model parameters based on the EnKF approach[J]. Journal of Hohai University(Natural Sciences), 2022, 50(2): 1-10.

汇水区域综合的水文信息,且地面实测径流数据具有很高的可靠性。因此,将站点径流实测信息同化到水文模型模拟预报过程中,对于改善流域状态变量及模型参数的估计精度,进而改进流域的降雨径流过程模拟及预测具有重要意义。

径流的同化始于20世纪80年代,主要是在流域/区域的水文模拟预报中基于站点径流观测数据进行模型输出结果的修正,也称为误差校正^[3]。当时所使用的同化方法主要为卡尔曼滤波(KF)法和扩展KF法。随着数据同化技术与模型模拟技术的发展,径流的同化不仅体现在线性系统或弱非线性系统对系统状态变量以及输出结果的校正上,更表现为在非线性系统中通过对系统状态变量及参数的更新或优化来改进整个系统的过程模拟及预报精度^[4-9]。用于同化的水文模型由简单的集总式概念性模型(如新安江模型^[10]、格林-安普特模型^[11])逐渐发展为基于物理机制的复杂的分布式模型(如SWAT模型^[12]、TopNet模型^[13]),同化方法也由线性KF法逐渐发展为各种集合KF法(如EnKF法)、粒子滤波和变分法等。基于简单的集总式概念性模型的径流数据同化方面,大量研究表明,通过径流数据同化能够获得较好的模型参数与状态估计结果,从而有效地改进流域水文模拟的预报精度和可靠性^[8,10,14]。但是,基于物理机制的分布式水文模型,由于其复杂的模型结构、大量的模型变量与参数,导致其径流数据同化过程中存在高维度现象,大大增加了数据同化效果的不确定性及其同化实施的难度^[12-13,15-16]。分布式水文模型中,径流数据的同化效果更容易受到模型的物理机制、模型结构、模型对流域空间异质性的概化能力以及同化系统规则等方面的影响。目前,分布式水文模型中径流数据的同化研究,基于径流观测数据进行模型状态更新与参数估计方面都相对滞后,主要表现在对分布式水文模型模拟误差与径流观测误差的合理量化^[7,13,17]、对模型参数的同化处理(包括参数在同化过程中的演变以及参数过拟合问题的处理等)^[5,15]等方面。

因此,本文在分布式水文模型SWAT中开展基于集合卡尔曼滤波(EnKF)法的径流数据同化研究,在合理量化模型模拟误差和径流观测误差、有效处理参数演变及过拟合问题的基础上,构建基于数据同化的水文模型参数优化方案,综合评估径流数据同化对SWAT模型参数的优化效果。

1 基于EnKF法和SWAT模型的数据同化方案

1.1 SWAT模型

SWAT模型是由美国农业部农业研究中心(ARG-USA)开发的分布式流域尺度水文模型^[18],主要包含水文过程子模型、土壤侵蚀子模型和污染负荷子模型。SWAT模型水文过程子模型涉及的关键模块有地表径流模块、蒸散发模块、土壤水分运动模块、地下径流模块和河道汇流模块。地表径流的模拟一般采用SCS曲线法;蒸散发的模拟是在潜在蒸散发量的基础上计算实际蒸散发量,包括植被冠层截留蒸发、植被蒸腾和土壤蒸发;土壤水分运动采用垂直分层水量平衡模拟,考虑下渗与壤中流过程;地下径流的模拟分为浅层和深层地下径流,涉及潜水再蒸发、灌溉等过程,基于水量平衡方程实现;河道汇流基于变动蓄量法或马斯经跟法对进入河道中的水流进行汇流演算。

1.2 基于EnKF法的状态-参数优化更新

1.2.1 基于EnKF法的状态更新

EnKF法是Evensen^[19]于1994年基于随机动态预报理论,将集合预报思想引入KF法,提出的一种顺序数据同化算法。EnKF法主要分为两步:预报(状态转移过程)和分析(状态更新过程)。基于Monte Carlo方法,将模型状态变量加入初始状态误差(通常为高斯白噪声),生成一组(数目 N)随机变量集合 $X_{k,i}(k=0, i=1, 2, \dots, N)$,每组随机变量 $X_{k,i}$ 在 k 时刻的一步预报方程为

$$\begin{cases} X_{k+1,i}^f = f_k(X_{k,i}^a, \theta, u_k) + \omega_k \\ \omega_k \sim N(0, Q_k) \end{cases} \quad (1)$$

式中: k 为第 k 个同化时刻; i 为第 i 个集合成员; $X_{k+1,i}^f$ 为第 i 个随机变量在 $k+1$ 时刻的模型状态预报值; $X_{k,i}^a$ 为第 i 个随机变量在 k 时刻的模型状态分析值; $f_k(\cdot)$ 为非线性模型算子,本文即SWAT模型; θ 为模型参数,通常不随时间变化; u_k 为 k 时刻的模型驱动项,如降雨、气温等; Q_k 为模型误差协方差矩阵; ω_k 为模型误差扰动项,其服从均值为0、方差为 Q_k 的正态分布。

当 $k+1$ 时刻有观测存在时,模型状态变量的预报值基于观测信息更新如下:

$$\begin{cases} \mathbf{X}_{k+1,i}^a = \mathbf{X}_{k+1,i}^f + \mathbf{K}_{k+1} [(\mathbf{Y}_{k+1} + v_{k+1}) - h(\mathbf{X}_{k+1,i}^f)] \\ v_{k+1} \sim N(0, \mathbf{R}_{k+1}) \end{cases} \quad (2)$$

其中 $\mathbf{K}_{k+1} = \mathbf{P}_{k+1}^f \mathbf{H}^T (\mathbf{H} \mathbf{P}_{k+1}^f \mathbf{H}^T + \mathbf{R}_{k+1})^{-1}$

式中: $\mathbf{Y}_{k+1}$ 为 $k+1$ 时刻的观测变量值集合, 本文为站点实测径流; $\mathbf{R}_{k+1}$ 为 $k+1$ 时刻的观测误差协方差矩阵; v_{k+1} 为观测误差扰动, 服从均值为 0、方差为 $\mathbf{R}_{k+1}$ 的高斯分布; $h(\cdot)$ 为观测算子, 表征模型状态变量 $\mathbf{X}_{k+1}^f$ 到观测变量 $\mathbf{Y}_{k+1}$ 所对应的模型输出变量的映射过程; $\mathbf{K}_{k+1}$ 为 $k+1$ 时刻的卡尔曼增益矩阵, 代表模型模拟误差与观测误差的权重; $\mathbf{H}$ 为观测算子 $h(\cdot)$ 的线性化表达式; $\mathbf{P}_{k+1}^f$ 为 $k+1$ 时刻模型状态变量 $\mathbf{X}_{k+1}^f$ 的预报误差协方差; $\mathbf{P}_{k+1}^f \mathbf{H}^T$ 为 $k+1$ 时刻模型状态变量 $\mathbf{X}_{k+1}^f$ 与对应于观测变量的模型输出变量 $h(\mathbf{X}_{k+1}^f)$ 的预报误差协方差; $\mathbf{H} \mathbf{P}_{k+1}^f \mathbf{H}^T$ 为 $k+1$ 时刻对应于观测变量的模型输出变量 $h(\mathbf{X}_{k+1}^f)$ 的预报误差协方差。

1.2.2 状态扩展、参数演变与过拟合问题处理

顺序数据同化中参数估计可以通过状态扩展方式^[20]实现, 即将待估计的参数在其物理区间内均匀采样后, 获得待估计的参数集合, 将其纳入状态向量 $\mathbf{X}$ 中, 并基于观测信息与更新状态变量一样对估计参数进行更新^[21]。假设待优化估计的参数向量为 $\boldsymbol{\delta}$, 则扩展后的状态向量可以表示为 $[\mathbf{X}^T \boldsymbol{\delta}^T]^T$, 基于 EnKF 法, k 时刻状态与参数的更新方程如下:

$$\begin{bmatrix} \mathbf{X} \\ \boldsymbol{\delta} \end{bmatrix}_k^u = \begin{bmatrix} \mathbf{X} \\ \boldsymbol{\delta} \end{bmatrix}_k^f + \mathbf{K}_k^* \left(\mathbf{Y}_k - \mathbf{H}^* \begin{bmatrix} \mathbf{X} \\ \boldsymbol{\delta} \end{bmatrix}_k^f \right) \quad (3)$$

式中: $\begin{bmatrix} \mathbf{X} \\ \boldsymbol{\delta} \end{bmatrix}_k^u$ 为 k 时刻更新后的扩展状态向量; $\begin{bmatrix} \mathbf{X} \\ \boldsymbol{\delta} \end{bmatrix}_k^f$ 为 k 时刻预报的扩展状态向量; $\mathbf{Y}_k$ 为 k 时刻的观测向量; $\mathbf{H}^*$ 为扩展状态向量对应的观测算子; $\mathbf{K}_k^*$ 为 k 时刻扩展状态向量对应的卡尔曼增益矩阵。由于 $\boldsymbol{\delta}$ 不属于观测变量, 则 $\mathbf{H}^*$ 为

$$\mathbf{H}^* = [\mathbf{H} \quad \mathbf{0}] \quad (4)$$

式中 $\mathbf{0}$ 为零矩阵, 其维数为 $n_{\text{ob}} \times n_{\text{par}}$, n_{ob} 为观测向量 $\mathbf{Y}_k$ 中观测变量的个数, n_{par} 为 $\boldsymbol{\delta}$ 中参数的个数。

在以上基于状态扩展方式的 EnKF 法参数估计框架下, 随着 EnKF 法更新过程的继续, 参数集合将快速收敛, 使得分析场很快趋近于背景场, 最终导致观测信息对参数更新不起作用, 即出现滤波发散。为避免滤波发散问题, 模型参数在同化过程中需要与状态变量一样具有一个随时间的演变过程 (状态变量随时间演变过程通过模型本身的物理机制实现)。但是, 这与一般意义上对模型参数的理解 (随机动态模型中的参数通常被假定为常数) 有冲突, 因此, 通常期望参数的变化相比于状态变量的变化要小得多^[1]。采用条件协方差法, 模拟参数演变:

$$\begin{cases} \delta_k^f = \delta_{k-1}^a & \text{var}(\delta_{k-1}^a) \geq v \\ \delta_k^f = \delta_{k-1}^a + \xi_k & \text{var}(\delta_{k-1}^a) < v \quad \xi_k \sim N(0, v - \text{var}(\delta_{k-1}^a)) \end{cases} \quad (5)$$

式中: δ_k^f 为 k 时刻的预报参数集合; δ_{k-1}^a 为 $k-1$ 时刻的分析参数集合; v 为给定的阈值, 通常设置为初始参数集合的协方差乘以一个给定的系数 γ 。当分析参数集合的协方差大于 v 时, 不对 δ_{k-1}^a 进行扰动; 当分析参数集合的协方差小于 v 时, 对 δ_{k-1}^a 加入均值为 0、方差为 $v - \text{var}(\delta_{k-1}^a)$ 的高斯扰动 ξ_k , 以增大参数的分布范围, 防止滤波发散。这样, 基于条件协方差法的参数演变过程不易出现滤波发散, 但当参数集合方差较大时, 参数将不存在演变过程。本文设 γ 为 0.0025。

另外, 参数更新过程中, 一个重要的问题是参数的过拟合 (过度更新) 问题, 其有可能导致模型模拟/预报过程中断或模型无法正常运行。参数的过度更新, 可能是由于数据同化过程中异常信号的加入 (如强降雨) 或对模型模拟与观测误差的量化不合理, 造成数据同化过程中对模型输出结果与实测数据的过度拟合导致。为抑制参数同化过程中的过度更新, 采用参数平滑法, 引入平滑因子 $\beta \in (0, 1.0)$ 进行处理^[22]:

$$\delta_k^f = (1 - \beta) \delta_{k-1}^a + \beta \delta_{k-1}^f \quad (6)$$

当 $\beta = 0$ 时, 表示参数更新过程中没有平滑作用, β 越大, 表示参数更新过程中的平滑作用越强。参考已有研究^[22], 取 β 为 0.8。

1.2.3 模型误差与径流观测误差的概化

SWAT 模型产汇流模拟中, 模型输入的不确定性主要考虑降雨和最高/最低气温输入误差。假定站点降

雨输入误差服从对数正态分布^[23],且该降雨输入误差在时间和空间上独立:

$$P_e = P\eta_p \quad \eta_p \sim \ln N(1, \sigma_p^2)$$

(7)

式中: P 为站点实测降水量; P_e 为降雨经误差扰动后的降雨集合; η_p 为降雨误差扰动项,服从均值为1(为保证降雨扰动后无系统性偏差)、方差为 σ_p^2 的对数正态分布; $\ln N(1, \sigma_p^2)$ 表示均值为1、方差为 σ_p^2 的对数正态分布。

最高/最低气温输入误差 η_T ,假定其服从均值为0、方差为 σ_T^2 的高斯分布:

$$T_e = T + \eta_T \quad \eta_T \sim N(0, \sigma_T^2)$$

(8)

式中: T 为实测最高/最低气温; T_e 为误差扰动后最高/最低气温集合。

模型结构的不确定性对模型模拟预报的影响主要通过扰动模型状态变量来刻画,即对模型状态变量的预报集合加上一定的高斯误差扰动^[15]:

$$X_e^f = X^f \eta_s \quad \eta_s \sim N(1, \sigma_s^2)$$

(9)

式中: X^f 为模型预测状态变量集合; X_e^f 为经状态变量误差扰动后的模型预测状态变量集合; η_s 为均值为1、方差为 σ_s^2 的高斯误差扰动项。

模型参数的不确定性通过对敏感参数加以高斯误差扰动:

$$\delta_e = \delta \eta_{par} \quad \eta_{par} \sim N(1, \sigma_{par}^2)$$

(10)

式中: δ_e 为扰动后的模型参数集合; η_{par} 为均值为1、方差为 σ_{par}^2 的高斯误差扰动项。

径流观测误差主要源于水位观测误差和水位流量转换误差。假定径流观测误差服从正态分布:

$$Q_e = Q(1 + \eta_{obs}) \quad \eta_{obs} \sim N(0, \sigma_{obs}^2)$$

(11)

式中: Q 为站点实测径流; Q_e 为经观测误差扰动后的径流观测集合; η_{obs} 为均值为0、方差为 σ_{obs}^2 的高斯误差扰动项。

1.3 基于 EnKF 法的径流数据同化方案

数据同化方案主要包括模型算子、观测算子和同化算法。模型算子为 SWAT 模型,同化算法为 EnKF 法。观测算子用于连接需要优化的模型状态变量/参数并辅助同化的观测数据。辅助同化的观测数据即为水文站点的实测径流量,而需要优化的模型状态变量与参数的选择根据其对降雨径流过程模拟的重要性和敏感性,选择地表径流滞留量 Q_{stor} 、土壤含水量 SW_{ly} 和河道蓄水量 V_{stored} 三类状态变量和控制地表径流、壤中流、地下径流以及河道汇流过程的7类参数,见表1。当状态变量与观测变量为同一类型的变量(如都为径流)时,观测算子为一个权重矩阵,其权重值可以通过状态变量到观测变量的插值过程获得。然而,由于 SWAT 模型的高度非线性,无法用一个矩阵显式地表示出 SWAT 模型中状态变量 $\mathbf{X}$ 到与观测变量 $\mathbf{Y}$ 所对应的模型输出变量的转化过程,为此,需要把 $\mathbf{Y}$ 对应的模型输出变量(即观测站点上的模拟径流)纳入 $\mathbf{X}$ 中。但在 EnKF 法更新过程中不对观测站点的模拟径流进行更新处理,因为只有对状态变量的更新才有意义^[13]。此时, $\mathbf{X}$ 中包括 n_{st} 个模型状态变量和被更新的参数与 n_{ob} 个观测站点的模型模拟径流,这样 $\mathbf{X}$ 与 $\mathbf{Y}$ 中存在对应变量,二者的状态转移矩阵 $\mathbf{H}^*$ 可以写成:

表 1 SWAT 模型径流数据同化中被更新的模型状态变量与参数

Table 1 Updated state variables and parameters in streamflow assimilation of SWAT model

次序	模型状态变量/参数	描述	单位	单元	采样区间	参数数目
1	Q_{stor}	地表径流滞留量	mm	HRU		
2	SW_{ly}	ly 土层的土壤含水量	m^3/m^3	$HRU * N_{lay}$		
3	V_{stored}	河道蓄水量	m^3	子流域		
1	CN_2	水文条件 II 下 SCS 曲线数		HRU	(60,98)	27
2	SOL_AWC	田间持水量时土壤孔隙度	mm/mm	$HRU * N_{lay}$	(0.01,0.4)	7
3	esco	土壤水蒸发补偿系数		HRU	(0.01,1.0)	7
4	α_{gw}	地下径流退水系数	d	HRU	(0,1.0)	7
5	aq _{shthr,q}	地下径流产流阈值	mm	HRU	(0,500)	7
6	surlag	地表径流滞留系数		流域	(1,12)	2
7	n	河道曼宁系数		子流域	(0,0.3)	2

注:HRU 表示 SWAT 模型中的水文响应单元; N_{lay} 为 HRU 土壤剖面上的土壤分层数; $HRU * N_{lay}$ 表示变量或参数在 SWAT 模型中离散的基本单元为 HRU 上的 N_{lay} 层土壤。

$$\boldsymbol{H}^* = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

$[0]_{n_{\text{ob}} \times n_{\text{st}}}$

$I_{n_{\text{ob}} \times n_{\text{ob}}}$

(12)

式中状态转移矩阵 $\boldsymbol{H}^*$ 由 $n_{\text{ob}} \times n_{\text{st}}$ 维的 $\mathbf{0}$ 矩阵和 $n_{\text{ob}} \times n_{\text{ob}}$ 维的单位矩阵构成。

基于 EnKF 法和 SWAT 模型的站点观测径流数据同化方案的实现是将 EnKF 法耦合到 SWAT 模型模拟过程中,该同化方案的流程如图 1 所示。

2 研究区与数据

研究区为淮河上游淮滨水文站以上流域(简称淮滨流域),流域控制面积为 16005 km²,整个流域除西部和西南部地区为山区外,大部分地区为平地,高程范围 25 ~ 1 117 m。土地利用类型主要为农业用地(包括旱地 34.2% 和水田 37.6%)和林地(23.6%)。土壤类型主要分为水稻土、黄褐土、黄棕壤、沙姜黑土等七类土壤。该流域冬春干旱少雨,夏秋闷热多雨,多年平均降水量约 900 mm,其中有 50% ~ 80% 发生在 6—9 月,多年平均气温约 15℃。

SWAT 模型的输入数据主要包括下垫面数据和气象数据两类(见表 2)。降雨数据主要由流域内共 106 个雨量站提供。最高/最低气温、太阳辐射、风速和相对湿度数据从研究区内(信阳)及其附近(固始、广水、驻马店和枣阳)共 5 个气象站获得,来自中国气象科学数据共享网。另外,流域内有 6 个水文站,分别为大坡岭、长台关、竹竿铺、息县、潢川和淮滨,提供流域的逐日径流量数据(因潢川站流量数据质量有问题,故未采纳)。研究区气象、水文站点分布见图 2,图中数字 1 ~ 37 为子流域代码。

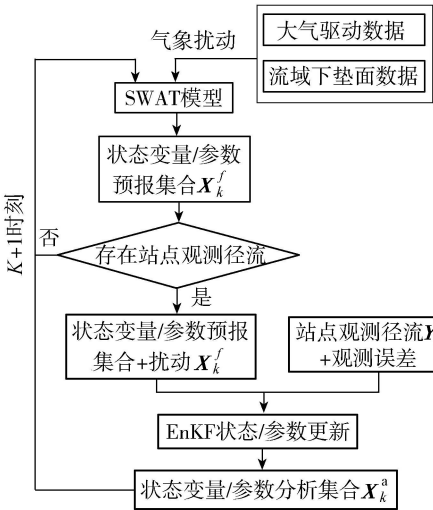


图 1 基于 EnKF 法与 SWAT 模型耦合的站点观测径流数据同化方案流程

Fig.1 Flow chart of the streamflow data assimilation scheme based on the coupling of EnKF and SWAT

表 2 淮滨流域 SWAT 模型构建中采用的数据

Table 2 Data utilized in SWAT model construction in Huaibin River Basin

数据	数据项	尺度	数据来源
下垫面数据	数字高程模型(DEM)	90 m	中国科学院国际科学数据服务平台
	土地利用图	1 000 m	寒区旱区科学数据中心
	土壤图	2 200 m	河南省土壤手册
气象观测数据	最高/最低气温、太阳辐射、风速、相对湿度	逐日数据	中国气象科学数据共享服务网
	降雨	逐日数据	

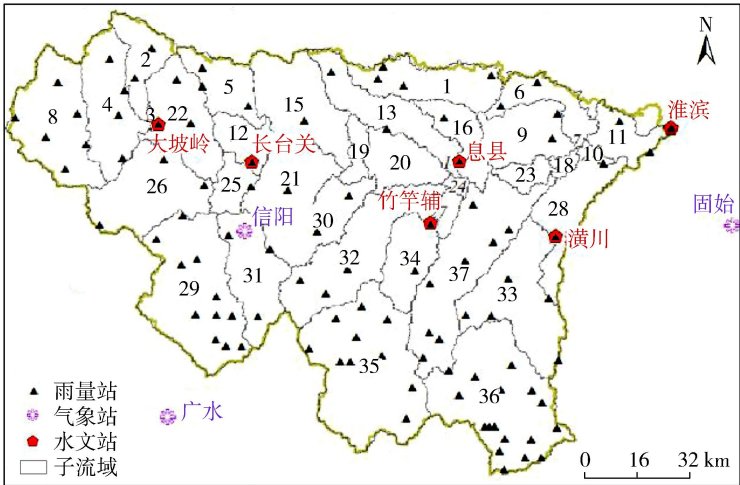


图 2 淮滨流域气象、水文站站点分布及子流域划分

Fig.2 Map of the meteorological and hydrological stations, and the subbasin division in Huaibin River Basin

基于上述数据资料,将淮滨流域划分为 37 个子流域,并在此基础上,进一步将整个流域离散成 146 个水文响应单元(HRU)(图 2)。由于无法对每个计算单元(HRU 或子流域)上的参数一一优化,为此,将各个计算单元上的参数进行归类。具体而言,将水文条件Ⅱ下 SCS 曲线数按土地利用和土壤类型的组合共分成 27 个参数;将田间持水量时土壤孔隙度、土壤水蒸发补偿系数、地下径流退水系数和地下径流产流阈值按土壤类型均分为 7 个参数;将地表径流滞留系数和河道曼宁系数根据河流分级(是否一级)分为 2 个参数(表 1)。在模型参数敏感性分析的基础上,采用自动率定与人工调整相结合的方法,对模型参数进行调整,并对调整后的模型进行模型与观测误差参数的估算。

3 径流数据同化实现与结果分析

3.1 模型模拟与径流观测误差量化

EnKF 法可以显式地处理模型与观测误差,并假定其都服从高斯分布。由 EnKF 法可以看出,对模型误差(模型输入、参数及结构)和径流观测误差的合理量化在一定程度上决定了数据同化的效果。本次基于径流数据同化的 SWAT 模型参数优化研究中,主要考虑降雨和最高/最低气温输入误差、模型状态误差、模型参数误差以及径流观测误差,误差均方差分别记为 σ_p 、 σ_T 、 σ_s 、 σ_{par} 和 σ_{obs} 项。其中,模型状态误差又由模型状态变量误差和与观测变量对应的模型输出变量误差组成,分别记为 σ_{state} 和 σ_{sflux} 。由于最高/最低气温输入误差对降雨径流模拟的影响较为敏感,借鉴以往研究^[24],设 σ_T 为 1℃。而其他因子误差均方差的设置是在模型参数敏感性分析的基础上,采用自动率定与人工调整相结合的方法对模型参数进行调整后,基于调整后的模型采用最大后验估计(MAP)法并结合全局优化 SCE-UA 法^[25]对其进行量化的。

在 MAP 法中,需要事先给定各个因子误差均方差的先验分布。借鉴已有研究^[8],各因子误差均方差的先验分布设置如下:① σ_{obs} 的先验分布设置为均值 9.25%、标准差 0.925% 的高斯分布;②其他因子 σ_p 、 σ_{state} 、 σ_{sflux} 和 σ_{par} 的先验分布设置为给定阈值内的均匀分布,其给定的阈值分别为 0~0.5、0~0.3、0~0.5 和 0~0.3,另外 σ_{obs} 给定的阈值为 0~0.5。确定各个因子误差均方差的先验分布后,即可获得抽样得到的某组误差均方差的先验概率。基于该组误差均方差对模型降雨输入、状态变量、参数以及输出变量进行扰动,得到率定期内的模型模拟径流输出系列集合。以流域出口控制站(淮滨站)的实测径流为验证数据,计算整个率定期内该组误差参数的似然函数。根据以上误差均方差的先验概率与似然函数计算得到其后验概率。通过 SCE-UA 方法搜索使后验概率达到最大时的误差参数组,作为 σ_{obs} 、 σ_p 、 σ_{state} 、 σ_{sflux} 和 σ_{par} 的最优估计,即最大后验估计。

对以上误差均方差的估计,采用了淮滨站 2002—2004 年共 3 年的日径流量数据。在 SCE-UA 方法中,模型截断运行次数为 60 000 次,而实际运行过程中,当模型运行到 34 860 次时已收敛,最终获得 σ_{obs} 、 σ_p 、 σ_{state} 、 σ_{sflux} 和 σ_{par} 的最优估计结果,分别为 0.087、0.305、0.028、0.29、0.113。

图 3 为上述均方差参数条件下,2005—2008 年的模型模拟径流(95%的置信区间)与站点实测径流过程,图中黑色实线为实测径流,灰色区域为模拟径流 95% 的置信区间。从图 3 可以看出,4 年内绝大多数观测径流数据都落在 95% 的置信区间内,说明该误差参数估计具有一定的合理性。同时也注意到,低流量时($Q<100\text{ m}^3/\text{s}$),径流模拟存在一定程度的低估,这与观测误差参数估算前模型的参数率定有一定关系。

3.2 模型参数同化更新过程

根据流域内 5 个站点(大坡岭、长台关、竹竿铺、息县和淮滨站)2003—2005 年的日径流实测数据,基于 EnKF 法实时更新 SWAT 模型敏感参数与状态变量,以期获得模型参数的最优估计。从被同化更新的 7 类共 59 个敏感参数中(表 1),随机选取 2 个代表性参数,即水文条件Ⅱ下 SCS 曲线数和表征土壤有效持水能力参数(SOL_AWC),给出其在数据同化过程中的演变过程,见图 4。从图 4 可以看出,随着观测信息的不断加入,这两个参数的集合分布越来越窄,集合均值逐渐趋于稳

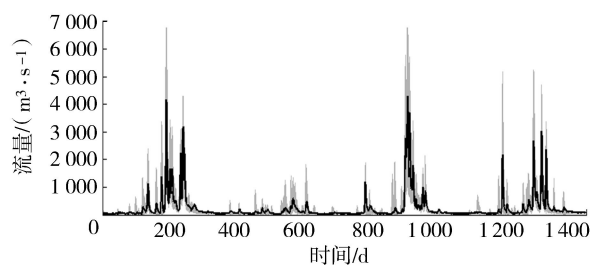


图 3 最优误差均方差条件下 2005—2008 年的模型模拟(95%置信区间)与站点实测径流过程

Fig. 3 Modeling(with 95% confidence interval) and observed runoff process based on the optimal error parameters over 2005—2008

定。从参数集合在某一天的概率密度分布(图 4 小图)可以看出,参数集合由初始时的均匀分布,通过实时的同化更新,逐渐演变为高斯分布;同时,集合的方差/标准差逐渐变小,说明参数估计的不确定性在逐渐减小。

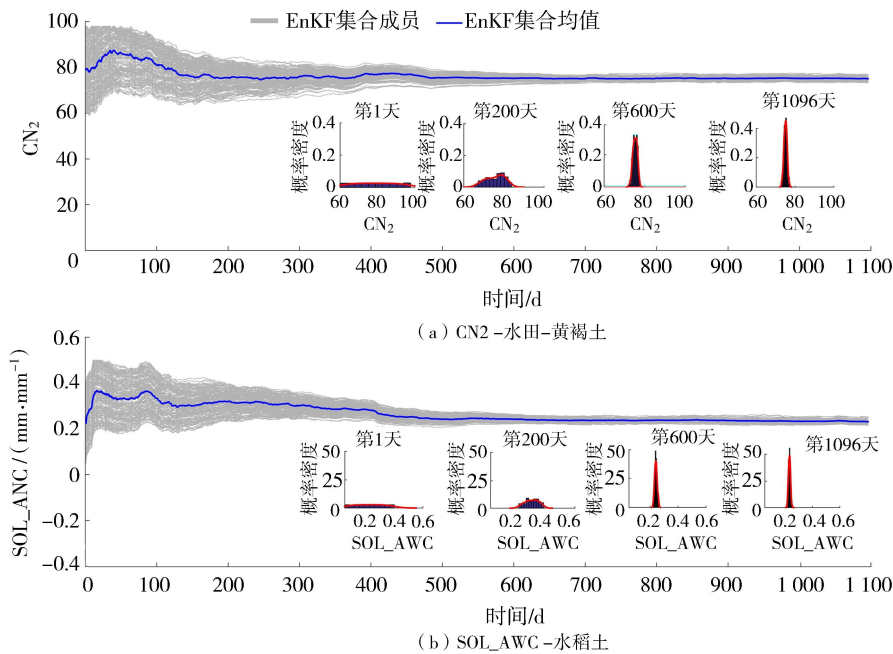


图 4 基于 EnKF 法径流数据同化的 SWAT 模型参数 CN_2 和 SOL_AWC 的演变过程

Fig. 4 Evolution process of CN_2 and SOL_AWC in EnKF based streamflow assimilation

图 5 为 SWAT 模型中被优化的七类共 59 个敏感参数其集合均方差随径流数据同化更新的变化过程。从图 5 可以看出,七类参数的集合均方差随同化更新过程的继续逐渐减小,表明参数集合的分布越来越窄,即参数估计的不确定性逐渐减小。最终,参数集合的均方差都收敛到初始集合均方差的 5% 以内,说明基于 EnKF 法的径流数据同化对 SWAT 模型参数具有一定的优化更新能力。

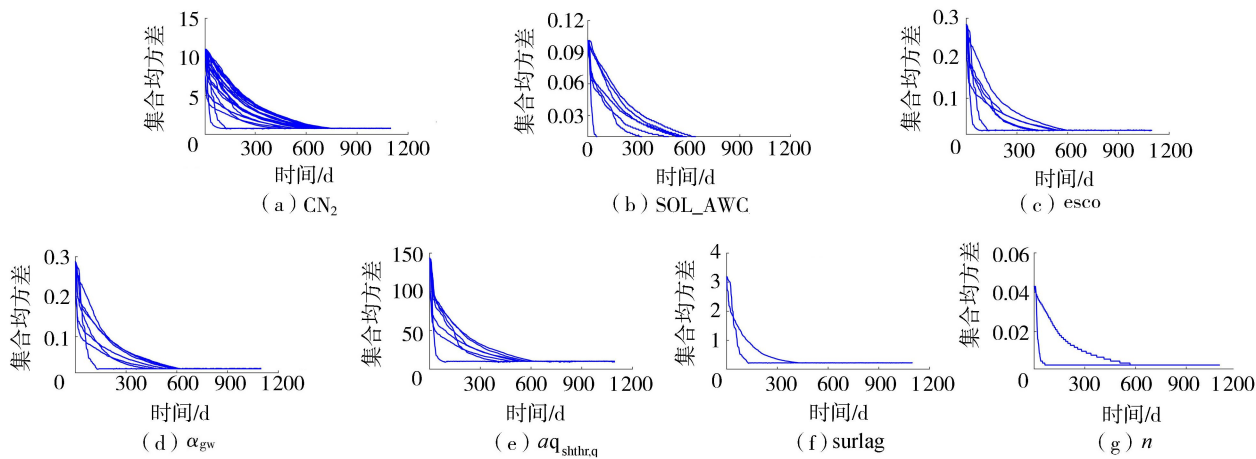


图 5 SWAT 模型中七类参数的集合均方差随径流数据同化的变化过程

Fig. 5 Variations of the ensemble mean square errors for 7 types of parameters in SWAT updated with the streamflow assimilation

3.3 模型参数优化效果评估

基于实测径流数据同化进行 SWAT 模型参数优化时,由于真实(或最优)的模型参数值未知且很多参数无法直接量测,很难直接对参数估计效果进行评估。本文通过基于优化后的参数模拟获得的流域内站点径流过程间接地对参数优化效果进行评估。具体方案为:基于对准滨流域内 5 个站点 2003—2005 年(率定期)实测径流数据同化获得的参数估计结果运行模型,得到 2006—2008 年(验证期)的径流模拟过程(记为

EnKF);将该径流模拟过程与站点实测径流过程进行比较,通过模拟径流与站点实测径流过程线的接近/贴合程度以及相应的评估指标,对基于 EnKF 法径流数据同化获得的模型参数优化效果进行评估。同时,将未经率定的参数(即在给定模型输入条件下,SWAT 模型默认的参数)在验证期的径流模拟过程(记为 SIM)作为参照。

表 3 为基于 EnKF 法优化的参数及未经率定的参数获得的流域内 5 个水文站点在验证期的径流模拟结果。从表 3 可以看出,基于 EnKF 法获得的参数估计结果所对应的流域内 5 个站点在验证期的径流模拟精度相较于未经率定的情况有大幅度的提高。EnKF 法对应 5 个站点径流模拟的相对偏差(除竹竿辅站外)都保持在 15% 以内,均方根误差显著减小,尤其是流域出口淮滨站的均方根误差减小了 34.5%,而纳什效率系数和相关系数 r 均明显增加,5 个站点的纳什效率系数都在 0.6 以上,其中大坡岭、息县、淮滨 3 站在 0.8 以上,尤其流域出口淮滨站达到了 0.88,表明基于 EnKF 法在 SWAT 模型中同化站点径流数据、更新模型参数与状态变量可以得到较好的模型参数估计结果。

表 3 基于 EnKF 法优化及未经率定的参数条件下流域内 5 个站点 2006—2008 年的径流模拟结果

Table 3 Simulated streamflow results of five hydrological stations over 2006—2008 base on the EnKF optimized parameters and the uncalibrated model parameters

水文站	经 EnKF 法优化的参数				未经率定的参数			
	相对偏差/%	均方根误差/ ($\text{m}^3 \cdot \text{s}^{-1}$)	纳什效率系数	r	相对偏差/%	均方根误差/ ($\text{m}^3 \cdot \text{s}^{-1}$)	纳什效率系数	r
大坡岭	3.44	29.46	0.81	0.90	0.64	32.34	0.77	0.88
长台关	-10.97	57.80	0.71	0.86	-7.22	69.04	0.59	0.79
竹竿辅	-38.04	52.18	0.62	0.81	-42.56	55.99	0.56	0.78
息 县	-14.95	132.20	0.80	0.90	-22.30	161.50	0.71	0.87
淮 滨	-12.94	150.40	0.88	0.94	-22.64	229.40	0.72	0.91

图 6 为基于 EnKF 法优化的参数(EnKF)及未经率定参数(SIM)条件下,流域出口淮滨站在验证期(2006-01-01—2008-12-31)的径流模拟过程。从图 6 可以看出,基于 EnKF 法参数估计结果的验证期径流(EnKF),其洪峰过程(图 6(a))和基流过程(图 6(b),图 6(b)纵坐标为以 10 为底的对数坐标)都明显优于未经参数率定条件下的径流模拟结果(SIM)。虽然基于 EnKF 法估计的参数所对应的径流过程相较于实测径流过程还存在一定程度的低估,但其相对于未率定过程而言其低估程度大为减小,表明通过 EnKF 法同化站点实测径流对 SWAT 模型参数具有一定的优化能力,可以作为未来 SWAT 模型参数率定的手段。

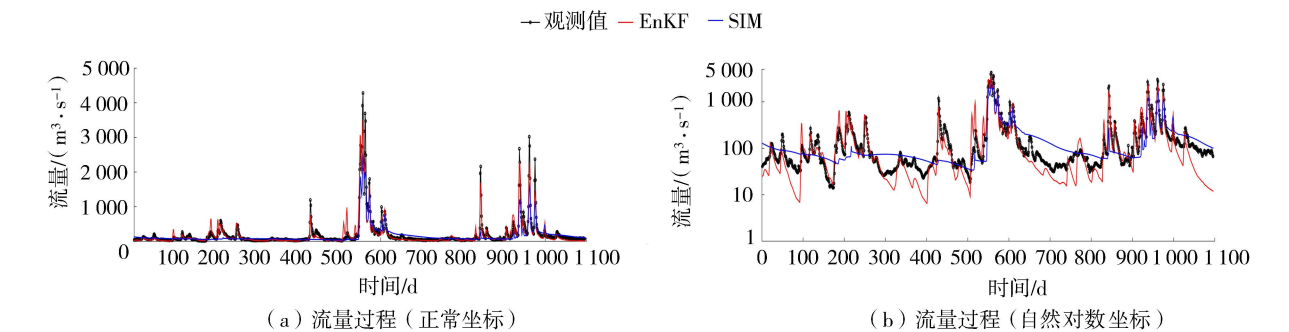


图 6 基于 EnKF 法及未经率定的参数条件下,流域出口淮滨站在验证期的径流过程

Fig. 6 Runoff process at the outlet of Huaibin River Basin based on the EnKF optimized parameters and uncalibrated model parameters

4 结 语

本文基于状态扩展方式的 EnKF 法,在对 SWAT 模型误差(模型输入、模型结构、模型参数)和站点径流观测误差采用最大后验估计方法合理量化的基础上,以条件协方差法并引入平滑因子有效刻画参数随时间的演变过程,构建了基于径流数据同化的水文模型参数优化方案。通过站点实测径流数据的同化对 SWAT 模型参数及状态变量进行实时更新,以淮河上游淮滨水文站以上流域为研究区,对该方案中模型参数的优化效果进行了评估。

结果显示,在站点实测径流数据同化过程中,被优化更新的参数集合逐渐收敛并趋于稳定。通过对比分析基于稳定后的参数获得的模拟径流和实测径流过程,发现通过径流数据同化获得的参数具有较好的模拟预测精度,基于数据同化方法得到的参数进行降雨径流模拟,流域出口淮滨站径流模拟的纳什效率系数可达 0.88,间接证明了基于 EnKF 法的径流数据同化对 SWAT 模型参数具有一定的优化估计能力,且采用数据同化方式进行水文模型参数的率定具有一定的可行性。

相较于目前常用的参数率定方法(如全局优化方法)将模型模拟误差全部归咎于模型参数误差这一不足^[26-27],通过同化站点径流进行模型参数优化,可综合考虑模型输入、模型结构和模型参数的不确定性,理论上,基于数据同化进行模型参数的优化可获得更优的参数估计。从研究实例来看,当同化进行到 600 ~ 800 时间步长后,模型参数就可收敛且趋于稳定,说明相较于一般的参数率定方法(一般需要一个丰、平、枯水文周期的实测数据),基于数据同化的参数优化要求更短的实测数据系列,这对于资料匮乏地区(通常数据系列较短)降雨径流模型中参数的优化具有重要意义。然而,基于数据同化的模型参数优化必须建立在对模型模拟误差与径流观测误差合理量化的基础上,而模型模拟误差与径流观测误差的估计一直以来都是水文界的一个难题^[8,28],这也是当前该方法在应用中面临的一大挑战。另外,由于水文模型,尤其具有物理机制的分布式模型的变量与参数众多^[29-30],导致径流数据同化过程中不可避免地存在高维度现象,这样同化效果的不确定性及同化实施的难度都将大大增加,因此,基于径流数据同化进行水文模型参数的优化还需要发展新的、更优的数据同化方法与策略。

参考文献:

- [1] LIU Yuqiong, GUPTA H V. Uncertainty in hydrologic modeling: toward an integrated data assimilation framework[J]. *Water Resources Research*, 2007, 43(7): W07401.
- [2] REICHLER R H. Data assimilation methods in the earth sciences[J]. *Advances in Water Resources*, 2008, 31(11): 1411-1418.
- [3] FJELD M, AAM S. An implementation of estimation techniques to a hydrological model for prediction of runoff to a hydroelectric power station[J]. *IEEE Transactions on Automatic Control*, 1980, 25(2): 151-163.
- [4] CHEN He, YANG Dawen, HONG Yang, et al. Hydrological data assimilation with the ensemble Square-Root-Filter: use of streamflow observations to update model states for real-time flash flood forecasting[J]. *Advances in Water Resources*, 2013, 59: 209-220.
- [5] MORADKHANI H, SOROOSHIAN S, GUPTA H V, et al. Dual state-parameter estimation of hydrological models using ensemble Kalman filter[J]. *Advances in Water Resources*, 2005, 28(2): 135-147.
- [6] LEE H, SEO D J, LIU Y, et al. Variational assimilation of streamflow into operational distributed hydrologic models: effect of spatiotemporal scale of adjustment[J]. *Hydrology and Earth System Sciences*, 2012, 16(7): 2233-2251.
- [7] LI Yuan, RYU D, WESTERN A W, et al. Assimilation of stream discharge for flood forecasting: the benefits of accounting for routing time lags[J]. *Water Resources Research*, 2013, 49(4): 1887-1900.
- [8] LI Yuan, RYU D, WESTERN A W, et al. An integrated error parameter estimation and lag-aware data assimilation scheme for real-time flood forecasting[J]. *Journal of Hydrology*, 2014, 519: 2722-2736.
- [9] ZILIANI M G, GHOSTINE R, AIT-EL-FQUIH B, et al. Enhanced flood forecasting through ensemble data assimilation and joint state-parameter estimation[J]. *Journal of Hydrology*, 2019, 577: 15.
- [10] 李漫漫, 石朋, 尚艳丽等. 基于集合卡尔曼滤波的新安江模型状态变量实时修正方法[J]. *河海大学学报(自然科学版)*, 2019, 47(3): 209-214. (LI Manman, SHI Peng, SHANG Yanli, et al. Real-time correction method for state variables of the Xinanjiang model based on ensemble Kalman filter[J]. *Journal of Hohai University(Natural Sciences)*, 2019, 47(3): 209-214. (in Chinese))
- [11] 李致家, 霍文博, 张珂. 格林-安普特降雨径流模型改进及初步应用[J]. *河海大学学报(自然科学版)*, 2020, 48(5): 385-391. (LI Zhijia, HUO Wenbo, ZHANG Ke. Improvement and preliminary application of Green-Ampt rainfall-runoff model[J]. *Journal of Hohai University(Natural Sciences)*, 2020, 48(5): 385-391. (in Chinese))
- [12] LIU Yongwei, WANG Wen, HU Yiming, et al. Improving the distributed hydrological model performance in upper Huai River Basin: using streamflow observations to update the basin states via the ensemble Kalman Filter[J]. *Advances in Meteorology*, 2016: 921616.
- [13] CLARK M P, RUPP D E, WOODS R A, et al. Hydrological data assimilation with the ensemble Kalman filter: use of streamflow observations to update states in a distributed hydrological model[J]. *Advances in Water Resources*, 2008, 31(10): 1309-1324.

- [14] LÜH, HOU T, HORTON R, et al. The streamflow estimation using the Xinanjiang rainfall runoff model and dual state-parameter estimation method[J]. *Journal of Hydrology*, 2013, 480: 102-114.
- [15] XIE Xianhong, ZHANG Dongxiao. Data assimilation for distributed hydrological catchment modeling via ensemble Kalmanfilter [J]. *Advances in Water Resources*, 2010, 33(6): 678-690.
- [16] LI Ning, KINZELBACH W, LI Haitao, et al. Improving parameter and state estimation of a hydrological model with the ensemble square root filter[J]. *Advances in Water Resources*, 2021, 147: 103813.
- [17] MAZZOLENI M, NOH S J, LEE H, et al. Real-time assimilation of streamflow observations into a hydrological routing model: effects of model structures and updating methods[J]. *Hydrological Sciences Journal-Journal Des Sciences Hydrologiques*, 2018, 63(3): 386-407.
- [18] NEITSCH S L, ARNOLD J G, KINIRY J R, et al. Overview of Soil and Water Assessment Tool(SWAT) Model[J]. *Tier B(All Countries/Territories not Listed Here Are in Tier A)*, 2009, 8: 3.
- [19] EVENSEN G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics[J]. *Journal of Geophysical Research-All Series*, 1994, 99b: 10.
- [20] BRAS J, VAN OOIJ C P, VAN DEN AKKER H P. Mandibular atrophy and metabolic bone loss; mandibular ridge augmentation by combined sandwich-visor osteotomy and resorption related to metabolic bone state[J]. *International Journal of Oral Surgery*, 1985, 14(1): 16-21.
- [21] ANNAN J D, HARGREAVES J C, EDWARDS N R, et al. Parameter estimation in an intermediate complexity earth system model using an ensemble Kalmanfilter[J]. *Ocean Modelling*, 2005, 8(1): 135-154.
- [22] YANG Xiaosong, DELSOLE T. Using the ensemble Kalman filter to estimate multiplicative model parameters[J]. *Tellus A*, 2009, 61(5): 601-609.
- [23] CROW W T, VAN DEN BERG M J, HUFFMAN G J, et al. Correcting rainfall using satellite-based surface soil moisture retrievals; the Soil Moisture Analysis Rainfall Tool(SMART)[J]. *Water Resources Research*, 2011, 47(8): W08521.
- [24] CHEN Fan, CROW W T, STARKS P J, et al. Improving hydrologic predictions of a catchment model via assimilation of surface soil moisture[J]. *Advances in Water Resources*, 2011, 34(4): 526-536.
- [25] DUAN Qingyun, SOROOSHIAN S, GUPTA V. Effective and efficient global optimization for conceptual rainfall-runoff models [J]. *Water Resources Research*, 1992, 28(4): 1015-1031.
- [26] 董洁平, 李致家, 戴健男. 基于 SCE-UA 算法的新安江模型参数优化及应用[J]. *河海大学学报(自然科学版)*, 2012, 40(5): 485-490. (DONG Jieping, LI Zhijia, DAI Jiannan. Optimization and application of Xin'anjiang model parameters based on SCE-UA algorithm[J]. *Journal of Hohai University(Natural Sciences)*, 2012, 40(5): 485-490. (in Chinese))
- [27] 崔东文. 多组群教学优化算法-神经网络-支持向量机组合模型在径流预测中的应用[J]. *水利水电科技进展*, 2019, 39(4): 41-48. (CUI Dongwen. Application of multi-group teaching optimization algorithm-neural network-support vector machine combination model in runoff prediction[J]. *Progress in Water Resources and Hydropower Science and Technology*, 2019, 39(4): 41-48. (in Chinese))
- [28] 贺新月, 曾献奎, 王栋. 融合 AR 模型和 MCMC 方法的水文模拟不确定性分析[J]. *河海大学学报(自然科学版)*, 2020, 48(2): 116-122. (HE Xinyue, ZENG Xiankui, WANG Dong. Uncertainty analysis of hydrological simulation combining AR model and MCMC method[J]. *Journal of Hohai University(Natural Sciences)*, 2020, 48(2): 116-122. (in Chinese))
- [29] 曾家俊, 麦叶鹏, 李志威, 等. 广州天河智慧城 SWMM 参数敏感性分析[J]. *水资源保护*, 2020, 36(3): 15-21. (ZENG Jiajun, MAI Yepeng, LI Zhiwei, et al. Sensitivity analysis of SWMM parameters in Guangzhou Tianhe Wisdom City[J]. *Water Resources Protection*, 2020, 36(3): 15-21. (in Chinese))
- [30] 李永坤, 薛联青, 邸苏闯, 等. 基于 Infoworks ICM 模型的典型海绵措施径流减控效果评估[J]. *河海大学学报(自然科学版)*, 2020, 48(5): 398-405. (LI Yongkun, XUE Lianqing, DI Suchuang, et al. Evaluation of runoff reduction effect of typical sponge measures based on infoworks ICM model[J]. *Journal of Hohai University(Natural Sciences)*, 2020, 48(5): 398-405. (in Chinese))