

DOI:10.3880/j.issn.1004-6933.2021.04.023

湖泊藻类动态模型数据同化模式的改进

李 港¹, 陈 诚¹, 何欣霞^{1,2}, 何梦男¹, 邓 跃³, 陈求稳^{1,2}

(1. 南京水利科学研究院生态环境研究所, 江苏 南京 210029; 2. 重庆交通大学河海学院, 重庆 400074;
3. 中建四局第五建筑工程有限公司, 广东 深圳 518000)

摘要:以太湖为研究区域,采用 2014—2016 年的水环境生态监测数据,率定了三维水生态动力学模型(3DHED),模拟了太湖蓝藻生物量的时空变化;通过融入遥感数据建立了基于集合卡尔曼滤波(EnKF)的蓝藻生物量预测数据同化(DA)模式,同时提出了一种改进数据同化(mDA)的策略,降低了遥感数据不确定性的影响,显著提升了模型模拟精度。结果表明:相比 3DHED 蓝藻生物量的模拟结果,DA 模拟结果的均方根误差均值降低了 10.4%,IOA 均值增加了 48.8%;mDA 在 DA 基础上对蓝藻生物量的模拟精度进一步提升,其均方根误差均值为 1.16 mg/L,在 DA 基础上降低了 8.6%,IOA 均值为 0.71,在 DA 基础上增加了 10.9%,并有效提升了对蓝藻生物量峰值的捕捉能力,表明提出的 mDA 方法能有效减小原 DA 模式中遥感观测数据误差的影响,提升水华模拟精度。

关键词:太湖;蓝藻;三维水生态动力学模型;集合卡尔曼滤波;数据同化

中图分类号:X524 **文献标志码:**A **文章编号:**1004-6933(2021)04-0156-10

Modify of data assimilation model for lake algae dynamic model // LI Gang¹, CHEN Cheng¹, HE Xinxia^{1,2}, HE Mengnan¹, DENG Yue³, CHEN Qiuwen^{1,2} (1. Center for Eco-Environmental Research, Nanjing Hydraulic Research Institute, Nanjing 210029, China; 2. College of Hehai, Chongqing Jiaotong University, Chongqing 400074, China; 3. CCFED The Fifth Construction & Engineering Co., Ltd., Shenzhen 518000, China)

Abstract: Taking Taihu Lake as the research area, a three-dimensional hydro-ecological dynamics(3DHED) model was developed to simulate the spatiotemporal variation of cyanobacteria biomass in Taihu Lake based on the water environmental ecological monitoring data from 2014 to 2016. A data assimilation (DA) model of cyanobacteria biomass prediction based on ensemble Kalman filter (EnKF) was established by integrating remote sensing data. At the same time, a modified data assimilation (mDA) strategy is proposed to reduce the impact of remote sensing data uncertainty and significantly improve the simulation accuracy of the model. The results showed that compared with 3DHED imulation results, the RMSE of DA simulation results decreased by 10.4%, and the IOA increased by 48.8%. The simulation accuracy of mDA for cyanobacteria biomass based on DA was further improved. The mean RMSE of mDA was 1.16 mg/L, which was 8.6% lower than that of DA, and the mean IOA was 0.71, which was 10.9% higher than that of DA. mDA effectively enhanced the ability to capture the peak biomass of cyanobacteria, which means that the proposed method can effectively reduce the influence of remote sensing data error in the original DA model and improve the accuracy of bloom simulation.

Key words: Taihu Lake; cyanobacteria; three-dimensional hydro-ecological dynamics model; ensemble Kalman filter; data assimilation

湖泊是重要的淡水资源库,近些年来,由于营养盐的大量排放以及气候变化的影响,湖泊富营养化及水华问题日益严重^[1]。据 2018 年中国生态环境状况公报显示,我国约有 1/3 的湖库达到了富营养化状态^[2],其中太湖、滇池、巢湖等重要水源地富营

养化形势严峻,严重威胁饮用水安全和人体健康^[3]。水华的精确模拟预测是富营养化湖泊风险管控的基础,对于富营养化湖泊的治理具有重要意义。

富营养化模型是进行湖泊水华模拟预测的重要

基金项目:国家重点研发计划(2018YFC0830800);江苏省双创团队(SC917001);南京水利科学研究院创新团队(Y917020)

作者简介:李港(1997—),男,硕士研究生,研究方向为湖泊水环境保护。E-mail: gangli_rcee@163.com

通信作者:陈诚(1990—),男,工程师,博士,主要从事生态环境遥感与环境水信息学研究。E-mail: chen Cheng@nhri.cn

工具。国内外学者开发了 EFDC (environmental fluid dynamics code)、WASP (water quality analysis simulation program)、Delft3D-Eco 和三维水生生态动力学 (three-dimensional hydro-ecological dynamics, 3DHED) 模型等一系列湖泊水质模型^[4-7], 已成功应用于滇池^[8]、Langelmavesi-Roime 湖^[9]、维多利亚湖^[10] 等富营养化湖泊的藻类模拟预测。但由于大范围蓝藻生物量的空间分布获取困难, 模型使用站点观测数据作为输入数据时, 会产生较大的初值误差, 同时, 模型运行过程中模拟误差也会不断积累, 使得模拟精度无法得到有效保证^[11]。数据同化 (data assimilation, DA) 技术通过将观测数据融入模型计算过程, 校正模型模拟结果, 同步更新模型参数, 降低了输入数据和模型结构不确定性的影响^[12]。该技术包含连续数据同化和顺序数据同化两大类, 其中连续顺序同化算法利用同化时间窗口内的所有观测数据和模型状态值进行最优估计, 不断调整模型初始场, 最终将模型轨迹拟合至同化窗口周期内的观测值上, 主要以变分算法为主^[13]。相较于变分算法, 以集合卡尔曼滤波为代表的顺序数据同化算法能够提供状态量的均值及其相应的误差协方差, 其预报误差也可以随着模型动态变化, 同时由于引入了集合的思想, 不需要伴随或线性算子, 具有适用于非线性系统、程序设计相对简单以及易实现并行运算等特点^[14], 近年来得到了广泛的应用。Loos 等^[15] 结合 HSFP、EFDC 以及 WQFS (water quality forecast system), 建立了基于 EnKF (ensemble Kalman filter) 的 Yeongsan 河藻类数值模型和数据同化系统。刘卓等^[16] 采用 2009—2011 年太湖站点监测数据对 Delft3D-BLOOM 模型进行 EnKF 数据同化, 提高了 Delft3D-BLOOM 模型对太湖蓝藻的模拟精度。对于大型湖泊而言, 数据同化技术虽然能将初始场的使用时间减少到一个观测周期内, 降低了模型误差的累积, 但加入的同化数据是稀疏站点数据插值得到的蓝藻生物量数据, 忽略了蓝藻生物量的空间异质性, 存在较大的初值误差。遥感数据由于具有较好的空间代表性^[17], 尤其对大型湖区的蓝藻生物量表征能力优于站点观测数据, 所以将其与数据同化技术相结合, 以降低初值误差, 有效提高模型模拟与预测精度。但某些时刻遥感数据质量受云雾等气象条件的影响, 观测误差较大, 此时加入低质量的遥感数据进行数据同化, 反而会造成一个观测周期内的模拟精度降低。

基于此, 本文以太湖为研究区域, 采用 2014—2016 年水环境监测数据率定的 3DHED 模型进行藻类时空模拟, 加入蓝藻生物量遥感反演数据进行模

型的 EnKF 数据同化, 并提出一种数据同化改进策略, 通过比较两种方式模拟结果与实测数据的均方根误差 (RMSE), 进行数据同化时遥感数据的有效筛选, 避免因部分时刻遥感数据误差较大给模型带来的影响, 提升蓝藻生物量模拟精度, 以期富营养化湖泊水华的精确模拟预测和风险管控提供技术支撑。

1 研究区概况与数据资料

1.1 研究区概况

太湖是中国第三大淡水湖泊, 总面积为 2338 km², 平均水深约 2 m^[18], 分为梅梁湾、竺山湖、贡湖、西北湖、湖心区、东部湖区、西南湖和东太湖 8 个湖区 (图 1)。1980 年以来, 随着太湖周边城市化进程的不断加快, 湖泊的不合理利用开发加剧^[19], 如工业生产、围湖造田以及沿岸渔业发展等, 导致大量污染物汇入太湖, 使得太湖生态系统逐步退化, 水体自净能力下降, 水质日趋恶化, 蓝藻水华频发^[20]。2007 年蓝藻水华暴发造成了无锡市近百万人的饮用水危机, 引起全社会广泛关注^[21]。2017 年, 太湖再次暴发大面积蓝藻水华, 严重影响周边群众生产生活^[22], 太湖富营养化形势依然严峻。

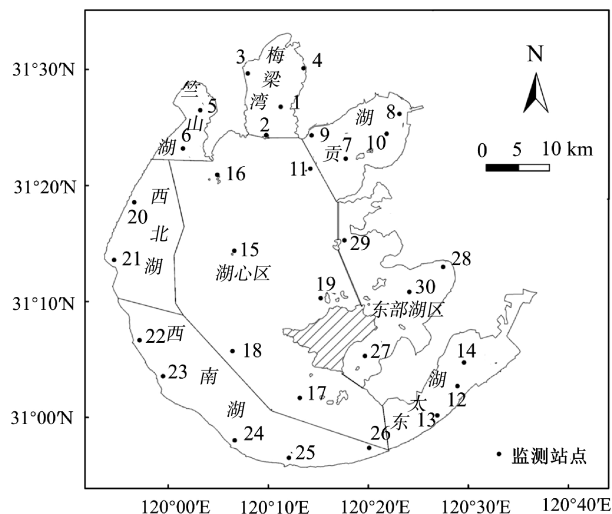


图 1 研究区域及采样点分布

Fig. 1 Study area and sampling point distribution

1.2 数据资料

数据资料包括太湖 2014—2016 年的气象、水文、水质、藻类生物量及遥感反演蓝藻生物量等数据。气象数据来源于中国气象信息中心共享网 (www.cma.gov.cn), 包括太湖湖心区站点无锡站每日监测一次的平均风速、最大风向、光照强度。水文、水质和藻类生物量数据来源于水利部太湖流域管理局水文水资源监测局: 水文数据包括 9 条入湖河流流量、10 条出湖河流水位的日监测数据和 4 个水位监测站数据 (22 夹浦、24 小梅口、27 西山、20

大浦口),水质数据和藻类生物量数据包括太湖8个湖区30个站点(1 三号标、2 拖山、3 间江口、4 小湾里、5 竺山湖、6 龙头、7 大贡山、8 沙墩港、9 渔业村、10 贡湖、11 东太湖、12 庙港、13 钱港、14 乌龟山、15 平台山、16 焦山、17 十四号灯标、18 湖心南、19 横山、20 大浦口、21 伏东、22 夹浦、23 太湖新塘、24 小梅口、25 大钱、26 汤娄、27 西山、28 胥口、29 漫山、30 胥湖)每月监测一次的水温(T)、溶解氧(DO)、光照、pH值、硝态氮(NO₃⁻-N)、总氮(TN)、磷酸盐(PO₄³⁻-P)、总磷(TP)、叶绿素a(Chl-a)、碎屑浓度、蓝藻生物量、绿藻生物量和硅藻生物量等数据。遥感数据来源于国家地球系统科学数据共享服务平台(<http://www.geodata.cn>),包括2014年1月1日至2016年12月30日的MODIS遥感反演蓝藻生物量数据^[23],其空间分辨率为250 m,选取其中与站点实测数据日期对应的36幅遥感数据作为待同化的观测数据。

2 模型与研究方法

2.1 3DHED 模型

3DHED模型通过耦合SELFE模型与SALMO模型,考虑了动力条件对藻类分布的影响,可以有效模拟藻类时空变化^[7]。其耦合原理为:①通过SELFE模型将湖泊划分为若干个三角网格,并将SALMO模型应用到其中的每个网格;②通过SELFE模型计算得到全湖所有网格的水动力数据,然后将其提供给SALMO模型;③根据水动力数据与模型水质初始场信息,利用SALMO模型在每个网格上进行水质模拟。3DHED模型参数率定参考郭静等^[24]的前期研究成果,采用2014—2016年的水环境生态监测数据对模型影响较大的20个参数进行率定。选定的参数及其率定结果如表1所示。

2.2 EnKF 数据同化

EnKF计算过程包括预测和分析两个步骤:设有 N 个集合,在 $k=0$ 时刻对每个集合进行初始化。式(1)为预测阶段,式(2)~(6)为分析阶段。

$$\begin{cases} \mathbf{X}_{fi,k+1} = \mathbf{M}_{k,k+1}(\mathbf{X}_{ai,k}) + \mathbf{w}_{i,k} \\ \mathbf{w}_{i,k} \sim N(0, \mathbf{Q}_k) \end{cases} \quad (1)$$

式中: $\mathbf{X}_{ai,k}$ 为 k 时刻第 i 集合状态变量分析值; $\mathbf{X}_{fi,k+1}$ 为 $k+1$ 时刻状态变量预测值; $\mathbf{M}_{k,k+1}$ 为模型算子,本文选取3DHED模型; $\mathbf{w}_{i,k}$ 为模型结构不确定性引起的模拟误差,服从均值为0、协方差矩阵为 \mathbf{Q}_k 的正态分布。

$$\mathbf{X}_{ai,k+1} = \mathbf{X}_{fi,k+1} + \mathbf{K}_{k+1}[\mathbf{Y}_{0k+1} - \mathbf{H}_{k+1}(\mathbf{X}_{fk+1})] \quad (2)$$

$$\bar{\mathbf{X}}_{ak+1} = \frac{1}{N} \sum_{i=1}^N \mathbf{X}_{ai,k+1} \quad (3)$$

表1 模型参数率定结果

Table 1 Calibration results of model parameters

参数	含义	藻类	取值范围	取值
PFA	浮游动物对藻类的摄食选择参数	硅藻	0.55 ~ 1.0	0.9
		绿藻	0.5 ~ 1.0	0.9
		蓝藻	0.18 ~ 0.5	0.18
TOPTA	藻类最佳生长温度(°C)	硅藻	14 ~ 22	17
		绿藻	20 ~ 25	20
		蓝藻	25 ~ 32	29
PHOMAX	藻类的最大光合作用速率(d ⁻¹)	硅藻	1 ~ 3	3
		绿藻	2 ~ 4	3.3
		蓝藻	1 ~ 3	3
VA	藻类沉降速率(m · d ⁻¹)	硅藻	0.017 ~ 0.35	0.02
		绿藻	0.016 ~ 0.25	0.01
		蓝藻	0.02 ~ 0.025	0.005
KP	藻类对磷酸盐的半饱和常数	硅藻		1.7
		绿藻		2
		蓝藻		1.7
PFD	浮游动物对碎屑的摄食选择参数		0.45 ~ 1.0	0.5
APSFMAX	底泥磷释放的最大速率(mg · m ⁻³ · d ⁻¹)			0.5
APSFMIN	底泥磷释放的最小速率(mg · m ⁻³ · d ⁻¹)			0.45
ANSFMIN	底泥氮释放的最小速率(mg · m ⁻³ · d ⁻¹)			0.08
KDEN	反硝化与有机物之间的关系系数		0.002 ~ 0.045	0.005 25

$$\begin{cases} \mathbf{K}_{k+1} = \mathbf{P}_{fk+1} \mathbf{H}^T (\mathbf{H} \mathbf{P}_{ak+1} \mathbf{H}^T + \mathbf{v}_{i,k})^{-1} \\ \mathbf{v}_{i,k} \sim N(0, \mathbf{R}_k) \end{cases} \quad (4)$$

$$\begin{aligned} \mathbf{P}_{fk+1} \mathbf{H}^T &= \frac{1}{N-1} \sum_{i=1}^N (\mathbf{X}_{fi,k+1} - \bar{\mathbf{X}}_{fi,k+1}) \\ &\left[\mathbf{H}(\mathbf{X}_{fi,k+1}) - \mathbf{H}(\bar{\mathbf{X}}_{fi,k+1}) \right]^T \end{aligned} \quad (5)$$

$$\begin{aligned} \mathbf{H} \mathbf{P}_{ak+1} \mathbf{H}^T &= \frac{1}{N-1} \sum_{i=1}^N \left[\mathbf{H}(\mathbf{X}_{fi,k+1}) - \mathbf{H}(\bar{\mathbf{X}}_{fi,k+1}) \right] \\ &\left[\mathbf{H}(\mathbf{X}_{fi,k+1}) - \mathbf{H}(\bar{\mathbf{X}}_{fi,k+1}) \right]^T \end{aligned} \quad (6)$$

式中: $\mathbf{X}_{ai,k+1}$ 为 $k+1$ 时刻第 i 个集合的状态变量分析值; \mathbf{K}_{k+1} 为卡尔曼增益矩阵; \mathbf{Y}_{0k+1} 为 $k+1$ 时刻观测值; \mathbf{H}_{k+1} 为观测算子,表示 $k+1$ 时刻状态变量值与观测之间的函数关系; $\mathbf{v}_{i,k}$ 为观测误差,服从均值为0协方差矩阵为 \mathbf{R}_k 的高斯分布; $\bar{\mathbf{X}}_{ak+1}$ 为所有集合的分析值的平均值; $\bar{\mathbf{X}}_{fi,k+1}$ 为第 i 个集合预测值; \mathbf{P}_f 为预测误差方差矩阵; \mathbf{P}_a 为分析场误差矩阵。

同时进行状态变量和模型参数同化,对模型模拟精度的提高有明显效果^[25]。本文以太湖优势藻种蓝藻生物量作为状态变量,以敏感性最高的蓝藻最佳生长温度作为模型参数^[16],同时进行状态变量和模型参数数据同化。采用EnKF算法进行数据同化时,合理的同化方案和合适的参数设置是保证同化效果的前提。EnKF算法利用集合思想解决了实际应用中背景协方

差矩阵估计和预报困难的问题,集合数越多,集合均值协方差与真值协方差越接近,但计算负荷会显著增加;集合数过少,则可能导致模型误差协方差估计错误。本研究将集合数设置为 30、50、70、100、200 和 500 进行数值试验,结果发现集合数为 100 时可较好地兼顾模拟精度和计算效率。采用数据同化方法将观测数据与模拟数据相融合,发现融合的优劣性与模拟误差协方差和观测误差协方差紧密相关^[26],所以选择合适的模拟误差和观测误差对提高同化精度至关重要。本文参考 Chen 等^[27-28]的研究结果,将观测误差与模拟误差分别设置为 1%、10%、20%、30%,筛选观测误差和模拟误差的最优组合,进行同化参数设置。

2.3 改进数据同化

同化模型中背景场和观测数据的误差均会导致模型的预报误差,而背景场与观测误差均与实测数据相关,因此本研究提出一种改进的数据同化(modified data assimilation, mDA)策略。首先在有遥感数据的时刻,加入遥感数据进行模型的数据同化计算,同时进行模型模拟计算,直到有站点实测数据的下一时刻,以站点实测数据为参考,计算并比较该时刻同化结果和模型模拟结果的 RMSE,从而进行同化时刻数据源的筛选。提出的 mDA 能够有效剔除低质量的遥感数据,避免遥感数据质量较差所导致的同化结果精度降低的问题,从而提高同化模型的鲁棒性。改进数据同化流程如图 2 所示。

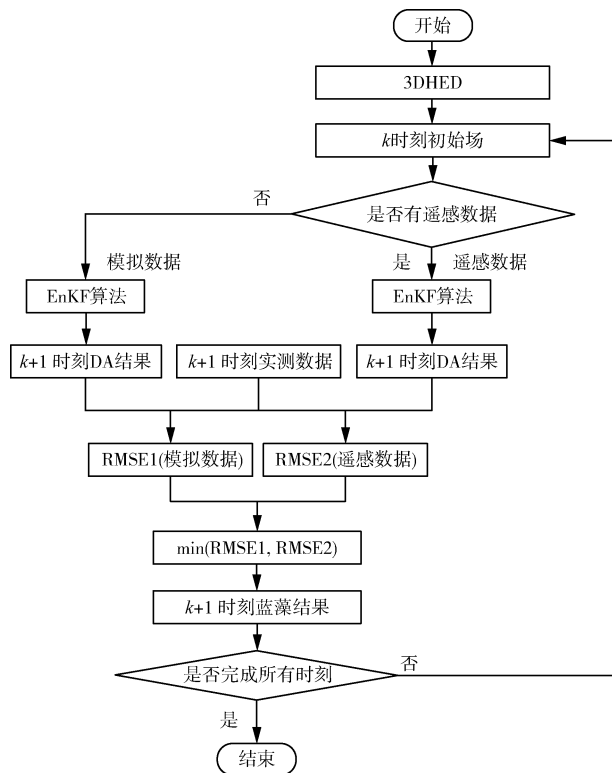


图 2 改进的数据同化流程

Fig. 2 Modified data assimilation process

2.4 模型精度评价指标

本文采用 RMSE 和一致性系数 IOA (index of agreement) 定量评价蓝藻生物量的观测数据与模拟结果的拟合度。RMSE 表示模拟结果与站点实测数据的偏差,其值越小说明模拟结果与站点实测数据越接近。IOA 表示模拟结果与站点实测站点数据的吻合程度,描述了模拟结果与站点实测数据在变化趋势上的一致性,IOA 值介于 0~1 之间,越接近 1 表示模拟结果与实测数据吻合度越高。

$$I_R = \sqrt{\frac{\sum_{i=1}^n (C_{simi} - C_{obsi})^2}{n}} \quad (7)$$

$$I_{OA} = 1 - \frac{\sum_{i=1}^n (C_{obsi} - C_{simi})^2}{\sum_{i=1}^n (|C_{simi} - \bar{C}_{obs}| + |C_{obsi} - \bar{C}_{obs}|)^2} \quad (8)$$

式中: I_R 为 RMSE 值; I_{OA} 为 IOA 值; C_{simi} 为第 i 个站点的蓝藻生物量模拟数据; C_{obsi} 为第 i 个站点的蓝藻生物量观测数据; \bar{C}_{obs} 为蓝藻生物量观测数据的平均值; n 为观测站点个数。本文计算的生物量值采用基于碳标记的蓝藻生物量,计算时使用常规质量浓度除以碳原子量得到。

3 结果与分析

3.1 数据同化参数确定

表 2 为 16 种不同观测误差和模拟误差组合下模拟结果与观测数据的 RMSE 值。从表 2 中可以看出,当观测误差设置为 1% 且模拟误差设置为 10% 时,数据同化结果的 RMSE 最小,因此本文以观测误差 1%、模拟误差 10% 作为 EnKF 数据同化方法的参数设置。

表 2 不同观测误差与模拟误差下的 RMSE 值

Table 2 RMSE values under different observation and simulation errors

模拟误差/%	不同观测误差下的 RMSE 值/(mg · L ⁻¹)			
	1%	10%	20%	30%
1	0.502	0.502	0.728	0.993
10	0.386	0.502	0.502	0.923
20	0.502	0.502	0.525	0.785
30	0.720	0.507	0.562	0.800

3.2 数据同化结果分析

针对太湖 8 个湖区各湖区选择一个代表性站点进行 3DHED 模拟结果和 DA 模拟结果的比较,图 3 为 8 个代表性站点蓝藻生物量观测值、3DHED 模拟结果以及 DA 模拟结果的对比。从图 3 可以看出,3DHED 模拟结果虽能在一定程度上反映蓝藻生物

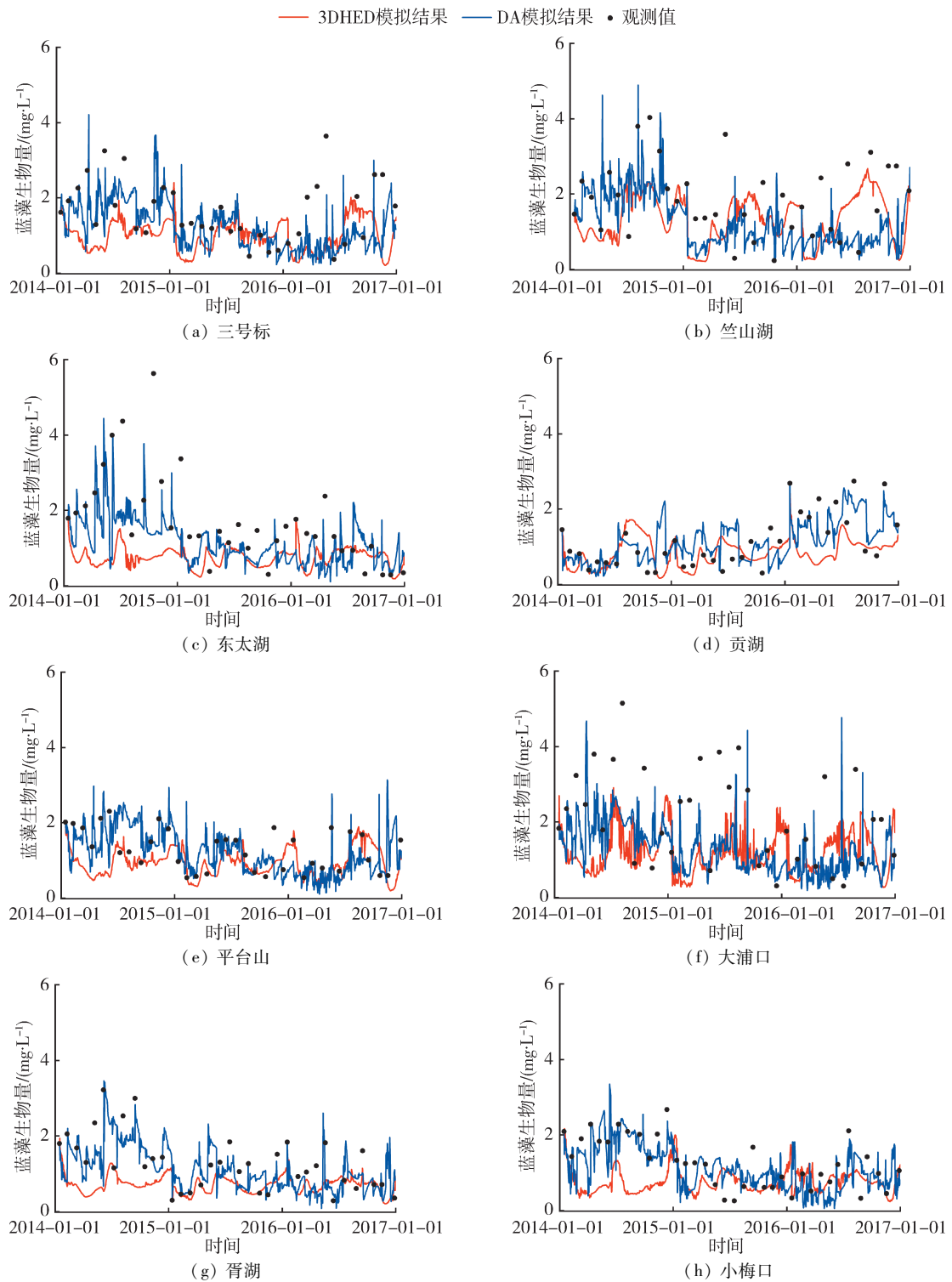


图3 蓝藻生物量观测值、3DHED 模拟结果和 DA 模拟结果对比

Fig. 3 Comparison of observation values, 3DHED simulation results and DA simulation results of cyanobacteria biomass

量的变化趋势,但对蓝藻生物量峰值的捕捉能力欠佳,而加入遥感数据进行同化后,同化时刻初值误差有所降低,模型误差累积周期缩短,使得 DA 模拟结果的蓝藻生物量变化趋势与峰值的模拟效果均较 3DHED 模拟结果有显著提高。8 个代表性站点的 3DHED 模拟结果、DA 模拟结果与站点实测数据的 RMSE 和 IOA 及其平均值见表 3。从表 3 可以看出,8 个代表性站点 3DHED 模拟结果的 RMSE 均值为

1.44 mg/L, IOA 均值为 0.43; DA 模拟结果的 RMSE 均值为 1.34 mg/L, IOA 均值为 0.60。相较于 3DHED 模拟结果,8 个代表性站点 IOA 平均值提升了 39.5%, RMSE 平均值降低了 6.9%,表明数据同化方法能够有效提高模型的模拟精度。

3.3 改进数据同化结果分析

图 4 为 8 个湖区代表性站点蓝藻生物量观测值、DA 模拟结果以及 mDA 模拟结果的对比。从图 4

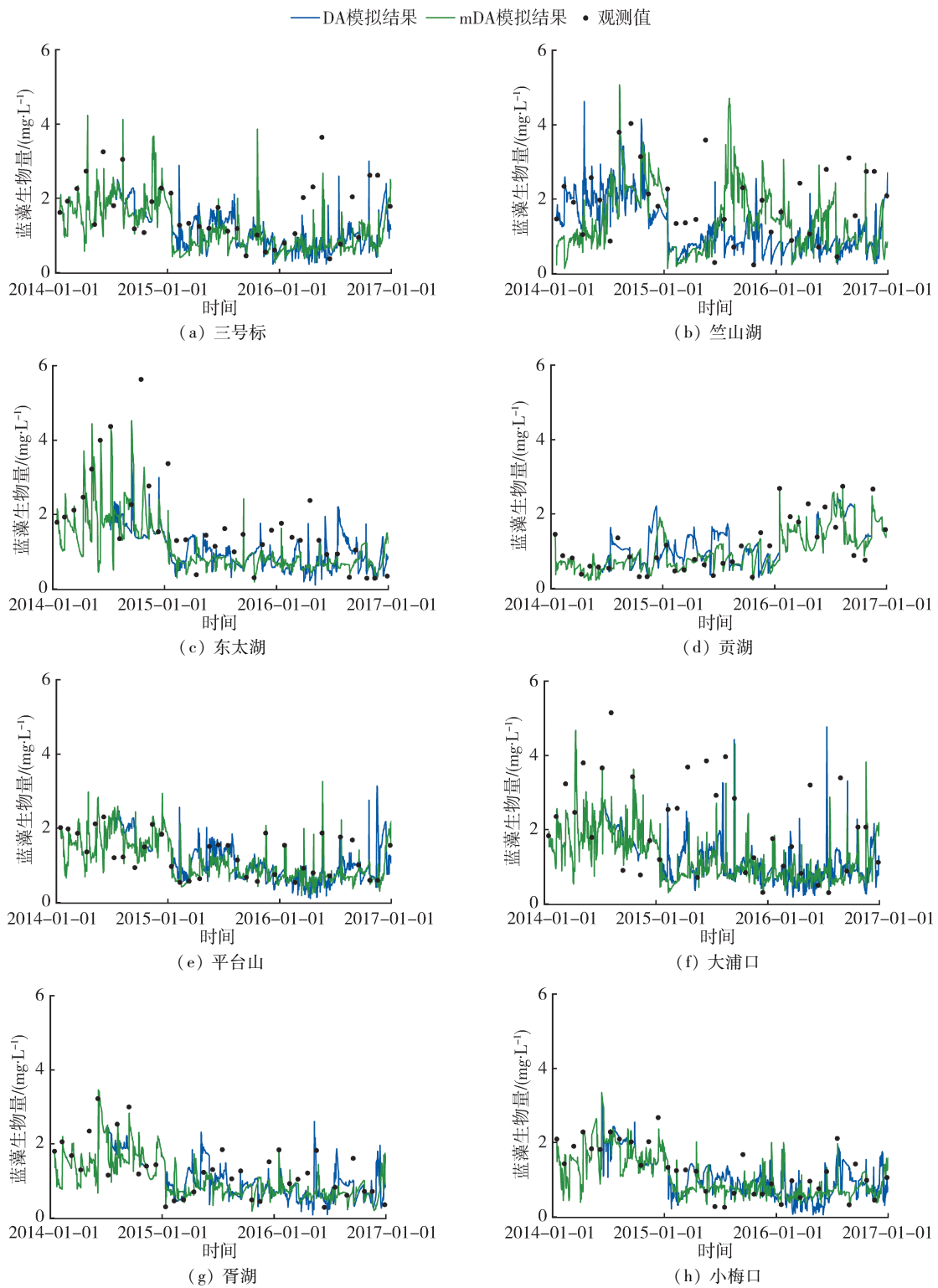


图4 蓝藻生物量观测值、DA 模拟结果和 mDA 模拟结果对比

Fig. 4 Comparison of observation values, DA simulation results and mDA simulation results of cyanobacteria biomass

可以看出,DA 模拟结果和 mDA 模拟结果均能较为准确地模拟蓝藻生物量的变化趋势,但 mDA 捕捉到蓝藻生物量峰值的次数有一定增加,表明采用 DA 方法进行数据同化时,部分同化时刻遥感数据存在较大误差,mDA 方法可进行高质量遥感数据的筛选,有效提升模拟精度。mDA 模拟结果的 RMSE 值介于 0.79 ~ 1.68 mg/L,8 个站点 RMSE 均值为

1.21 mg/L,在 DA 模拟结果基础上降低了 9.7%,mDA 模拟结果的 IOA 值介于 0.53 ~ 0.89,均值为 0.68,在 DA 模拟结果基础上提高了 13.3%。

图 5 为太湖 30 个采样点 3DHED 模拟结果、DA 模拟结果和 mDA 模拟结果的 IOA 值及 RMSE 值对比。3 种模拟方式下 30 个站点的 IOA 均值分别为 0.43、0.64 和 0.71, RMSE 均值为 1.42 mg/L、

1.27 mg/L和1.16 mg/L。各站点 IOA 值由大到小顺序为 mDA、DA、3DHED, RMSE 值由小到大顺序为 mDA、DA、3DHED。总体而言, mDA 的蓝藻生物量模拟精度最高, 但少部分站点存在 mDA 模拟结果的 RMSE 高于 DA 模拟结果的现象。

表3 8个湖区3DHED、DA和mDA模拟结果的IOA值和RMSE值的对比

Table 3 Comparison of IOA and RMSE values of 3DHED simulation results, DA simulation results and mDA simulation results in 8 lake zones

湖区	IOA 值			RMSE 值/(mg · L ⁻¹)		
	3DHED	DA	mDA	3DHED	DA	mDA
梅梁湾	0.46	0.67	0.71	1.23	1.1	1.07
竺山湖	0.39	0.52	0.56	2.17	2.1	2.01
西北湖	0.41	0.65	0.66	1.55	1.21	1.14
贡湖	0.40	0.60	0.66	1.27	1.17	1.14
湖心区	0.47	0.64	0.69	1.09	1.09	0.95
西南湖	0.45	0.67	0.75	1.36	1.14	0.98
东部湖区	0.40	0.72	0.82	1.31	1.05	0.86
东太湖	0.44	0.66	0.80	1.39	1.26	1.09

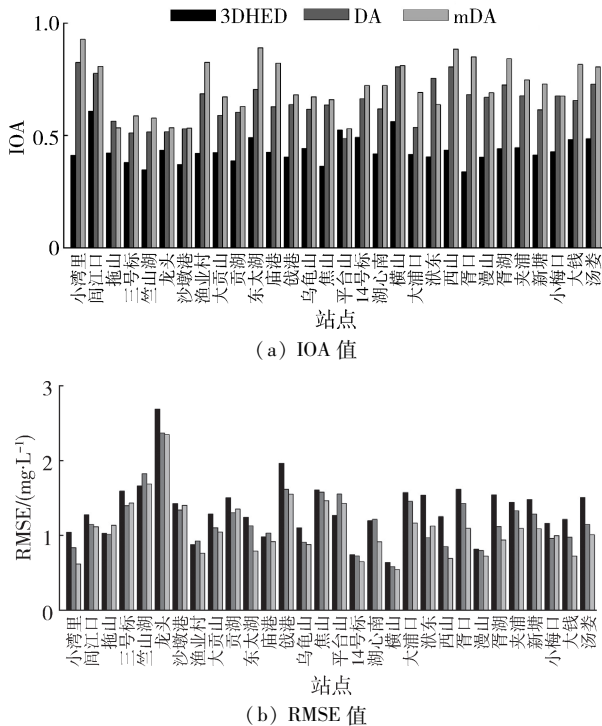


图5 30个站点3DHED、DA和mDA模拟结果的IOA值及RMSE值对比

Fig.5 Comparison of IOA and RMSE values of 3DHED simulation results, DA simulation results, and mDA simulation results in 30 sites

表3为各个湖区3DHED模拟结果、DA模拟结果以及mDA模拟结果的IOA值和RMSE值的对比情况。从表3可以看出, DA模拟结果的IOA均高于3DHED模拟结果的IOA, 而mDA模拟结果的IOA值在DA模拟结果的基础上又得到了进一步提升。对于3种模拟结果的RMSE, mDA模拟结果的RMSE值最低, 模拟效果最好, 整体上太湖各湖区RMSE值表现为西北部湖区大于东南部湖区。

4 讨论

在真实的环境下, 与蓝藻相关的一些物理参数(如蓝藻沉降率和被捕食率等)是难以精确模拟和确定的, 并且参数会随着时间而变化, 其过程远比模型模拟复杂得多^[29]。本文采用3DHED模型模拟2014—2016年太湖全湖蓝藻生物量的时空变化时, 虽然能够在一定程度上反映太湖蓝藻的变化趋势, 但总体模拟效果欠佳, 峰值模拟效果有待进一步提高(图3)。蓝藻从生长到形成水华会经历下沉—休

眠—复苏—上浮4个阶段^[30], 冬季沉于底泥中的藻类细胞在春季适宜条件下迅速生长, 于夏季聚集上浮形成水华^[31]。水动力条件也是决定水华暴发的关键因子^[32], 当蓝藻生物量聚集到一定程度后, 风浪的扰动使蓝藻细胞通过碰撞形成大群体, 风浪作用一旦减弱, 便上浮形成表面可见水华^[33]。结合本文模拟结果, 3DHED模拟过程中未能考虑冬季沉底部分的蓝藻, 同时对于风力作用下蓝藻的迁移过程描述不够清晰, 这可能是导致模型对于蓝藻生物量峰值捕捉不够理想的重要原因^[7]。采用遥感数据对3DHED进行EnKF数据同化后, 由于各个观测时刻的蓝藻生物量得到及时更新, 使得DA模式下的蓝藻生物量模拟精度得到明显提升(图3), 其IOA均值较3DHED模拟结果增加了48.8%, RMSE均值降低了10.4%。

模拟误差和观测误差是数据同化的关键因素^[16], 对同化性能有重要影响。设置较小的观测误差时, 观测数据精度更高, 同化结果会更接近观测值, 反之, 同化结果会更接近模拟值。一般情况下, 观测数据的误差更小, 更接近于真实值。李渊等^[34-35]在进行不同同化试验时采用了较小的观测误差, 取得了较为理想的试验结果。但遥感数据受到云覆盖等因素的影响, 在某些时刻、某些区域的数据误差较大, 在进行数据同化时, 由于模型采用的仍是较小的观测误差, 会导致遥感数据误差被低估, 同化结果趋近于遥感数据, 反而会更加偏离真实值。本文采用改进的数据同化模式后, 结果显示mDA的蓝藻生物量模拟效果较DA整体上有了一定的提升, 对蓝藻生物量峰值的捕捉频次也有所增加(图4), 全湖IOA均值在DA基础上增加10.9%, RMSE均值降低了8.6%, 说明未改进前遥感数据在某些时刻存在误差被低估的现象, 模型模拟结果趋向于误差被低估的遥感数据, 从而偏离真实值。

太湖蓝藻水华总体呈现东南低西北高的趋势,这与入湖河流污染负荷输入密切相关。西南部的天目山水系,城镇较少,人口密度低,水质较好。西北部的南溪水系水流多来自工农业发达的平原和城镇地区,入湖污染负荷较大^[36]。同时太湖夏季盛行东南风,受到太湖风浪的影响,蓝藻水华易在西北部聚集和堆积^[37],因此,重度水华主要集中在太湖西北部沿岸(如竺山湖、梅梁湾等)^[38]。本文采用 mDA 方法进行太湖蓝藻生物量的模拟时,除了在一定程度上减小遥感数据误差对同化结果的影响外,还能较好地模拟蓝藻生物量的空间分布。结合图 1 和表 3 可以看出, RMSE 值总体表现为西北部湖区大于东南部湖区,这是由于蓝藻生物量越大的区域,其 RMSE 值也越大。

湖泊藻类动态模型采用常微分方程描述湖泊生态系统的食物链动态和营养物质循环导致的自身结构不确定性^[25],使得模拟误差会随着模拟时间的增加而不断积累。数据同化技术通过实时加入观测数据,可及时更新初始场数据^[39],将一个初始场的使用时间缩短到一个观测周期内,有效降低了模拟误差的积累,此时初始场数据的误差对模型模拟精度起到决定性作用。遥感数据具有时间上较为连续、监测范围广等特点^[40],但若在观测时刻受到云雾等因素影响,提供的初始场数据误差可能大于模型模拟得到的初始场数据。本文以两种模式下提供的初始场数据模拟到下一观测时刻,比较得到的模拟结果与实测数据的 RMSE,再筛选出与真实情况误差较小的初始场数据进行模型模拟,从而进一步提高模拟精度。近年来,国内学者针对太湖蓝藻生物量模拟开展了一系列研究。张成成^[41]利用三维水生态模型模拟了蓝藻生物量的变化,验证站点的 IOA 值介于 0.23 ~ 0.50。Huang 等^[28]采用 EnKF 方法通过更新模型状态变量和模型参数以提高太湖富营养化模型对浮游植物的模拟精度,其最佳模拟效果下 IOA 值达到了 0.67。本研究 mDA 模式下的蓝藻生物量模拟结果改善效果明显(图 5),全湖平均 IOA 值为 0.71,表明本文提出的 mDA 同化方法具有较高的模拟精度,尤其对于监测站点稀疏的大型湖库水华的精确模拟预测具有广阔的应用前景。值得注意的是,存在少部分站点 mDA 结果精度低于 DA 的现象(图 5 中拖山、汴东等站点),这主要是由于本研究是通过计算整个湖区 36 个站点处同化结果与实测数据的 RMSE 值,来判断是否加入遥感数据进行同化,少部分站点遥感数据质量较差,但在整个湖区范围内精度较高,使得部分站点模拟效果不佳。因此后续的研究可以通过划分不同湖区来进行

遥感数据分区的筛选,从而进一步提高数据同化精度。同时,太湖面积广阔,不同湖区的蓝藻水华严重程度亦有差别,进行不同湖区模型参数的分区率定对模型精度提高具有一定的效果。

5 结 论

a. 采用遥感数据进行蓝藻生物量 EnKF 数据同化后,同化结果的 IOA 均值较 3DHED 模拟结果的 IOA 均值提升了 48.8%, RMSE 均值降低了 10.4%, DA 模拟精度较 3DHED 模型模拟精度有显著提高。

b. 本文提出的 mDA 方法对蓝藻生物量的模拟精度最高,总体精度由大到小为 mDA、DA、3DHED,其中 mDA 的全湖 IOA 均值达到 0.71, RMSE 均值降低至 1.16 mg/L,不仅能较好地反映蓝藻生物量时空变化趋势,还显著提高了对蓝藻生物量的峰值捕捉能力。

参考文献:

- [1] 叶许春,孟元可,张永生,等. 三峡库区香溪河回水区营养状态变化特征与驱动因子[J]. 水资源保护,2018,34(4): 80-85. (YE Xuchun, MENG Yuanke, ZHANG Yongsheng, et al. Variation characteristics of trophic states of Xiangxi River backwater area in Three Gorges Reservoir and its driving factors [J]. Water Resources Protection, 2018, 34(4): 80-85. (in Chinese))
- [2] 中华人民共和国生态环境部. 2018 年《中国生态环境状况公报》(摘录二)[J]. 环境保护,2019,47(12): 50-55. (Ministry of Ecology and Environment of the People's Republic of China. China ecological environment status bulletin in 2018 (excerpt 2) [J]. Environmental Protection, 2019, 47(12): 50-55. (in Chinese))
- [3] 王伟琴,金永堂,吴斌,等. 水源水中微囊藻毒素的遗传毒性与健康风险评估[J]. 中国环境科学,2010,30(4): 468-476. (WANG Weiqin, JIN Yongtang, WU Bin, et al. Assessment on genotoxicity and health risks of microcystin in drinking water sources [J]. China Environmental Science, 2010, 30(4): 468-476. (in Chinese))
- [4] HAMRICK J M. A three-dimensional environmental fluid dynamics computer code: theoretical and computational aspects[J]. Virginia Institute of Marine Science, 1992, 317:1-63.
- [5] AMBROSE R B, WOOL T A, MARTIN J L. The water quality analysis and simulation program, WASP5: part A: model documentation[J]. Athens Environmental Research Laboratory, 1993, 1: 201-210.
- [6] LOS F J, VILLARS M T, VAN DER T, et al. A 3-dimensional primary production model (BLOOM/GEM)

- and its applications to the South North Sea; coupled physical-chemical-ecological model[J]. *Journal of Marine Systems*, 2008, 74(1/2): 259-294.
- [7] CHEN C, HUANG J, CHEN Q, et al. Assimilating multi-source data into a three-dimensional hydro-ecological dynamics model using Ensemble Kalman Filter [J]. *Environmental Modelling & Software*, 2019, 117:188-199.
- [8] 张洪. 基于 EFDC 模型滇池外海水环境仿真研究 [D]. 杭州:浙江大学, 2017.
- [9] VENALAINEN A, SAHLGREN V, PODSECHIN V, et al. Small-scale variability of the wind field over a typical Scandinavian lake [J]. *Boreal Environment Research*, 2003, 8:71-81.
- [10] SEKADENDE B C, LYIMO T J, KURMAYER R. Microcystin production by cyanobacteria in the Mwanza Gulf(Lake Victoria, Tanzania) [J]. *Hydrobiologia*, 2005, 543(1) :299-304.
- [11] 陈鹤, 杨大文, 刘钰, 等. 集合卡尔曼滤波数据同化方法改进土壤水分模拟效果 [J]. *农业工程学报*, 2016, 32(2) :99-104. (CHEN He, YANG Dawen, LIU Yu, et al. Data assimilation technique based on ensemble Kalman filter for improving soil water content estimation [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2016, 32(2) :99-104. (in Chinese))
- [12] 王卫光, 李进兴, 魏建德, 等. 基于蒸散发数据同化的径流过程模拟 [J]. *水科学进展*, 2018, 29(2) :159-168. (WANG Weiguang, LI Jinxing, WEI Jiande, et al. Runoff simulation by hydrological model based on the assimilated evapotranspiration [J]. *Advances in Water Science*, 2018, 29(2) :159-168. (in Chinese))
- [13] 李渊, 李云梅, 吕恒, 等. 基于数据同化的太湖叶绿素多模型协同反演 [J]. *环境科学*, 2014, 35(9) :3389-3396. (LI Yuan, LI Yunmei, LYU Heng. et al. Multi-model collaborative retrieval of Chlorophyll a in Taihu Lake based on data assimilation [J]. *Environmental Science*, 2014, 35(9) :3389-3396. (in Chinese))
- [14] 马建文, 秦思娴. 数据同化算法研究现状综述 [J]. *地球科学进展*, 2012, 27(7) :747-757. (MA Jianwen, QING Sixian. Recent advances and development of data assimilation algorithms [J]. *Advances in Earth Science*, 2012, 27(7) :747-757. (in Chinese))
- [15] LOOS S, SHIN C M, SUMIHAR J, et al. Ensemble data assimilation methods for improving river water quality forecasting accuracy [J]. *Water Research*, 2019, 115:343.
- [16] 刘卓, 李志杰, 胡柳明, 等. 基于集合卡尔曼滤波的湖泊富营养化模型 Delft3D-BLOOM 数据同化 [J]. *湖泊科学*, 2017, 29(5) :1070-1083. (LIU Zhuo, LI Zhijie, HU Liuming, et al. Ensemble Kalman filter based data assimilation in the Delft3D-BLOOM lake eutrophication model [J]. *Journal of Lake Science*, 2017, 29(5) :1070-1083. (in Chinese))
- [17] 王寿兵, 徐紫然, 张洁. 大型湖库富营养化蓝藻水华防控技术发展述评 [J]. *水资源保护*, 2016, 32(4) :88-99. (WANG Shoubing, XU Ziran, ZHANG Jie. A review of technologies for prevention and control of cyanobacteria blooms in large-scale eutrophicated lakes and reservoirs [J]. *Water Resources Protection*, 2016, 32(4) :88-99. (in Chinese))
- [18] 赵静, 胡庆芳, 王腊春, 等. 基于 MSWEP 数据的太湖流域降水特性分析 [J]. *水资源保护*, 2020, 36(2) :27-33. (ZHAO Jing, HU Qingfang, WANG Lachun, et al. Analysis of precipitation characteristics in Taihu Lake Basin based on MSWEP [J]. *Water Resources Protection*, 2020, 36(2) :27-33. (in Chinese))
- [19] 王华, 陈华鑫, 徐兆安, 等. 2010—2017 年太湖总磷浓度变化趋势分析及成因探讨 [J]. *湖泊科学*, 2019, 31(4) :919-929. (WANG Hua, CHEN Huaxin, XU Zhaoan, et al. Variation trend of total phosphorus and its controlling factors in Lake Taihu, 2010-2017 [J]. *Journal of Lake Science*, 2019, 31(4) :919-929. (in Chinese))
- [20] 王寿兵, 屈云芳, 徐紫然. 基于生物操纵的富营养化湖库蓝藻控制实践 [J]. *水资源保护*, 2016, 32(5) :1-4. (WANG Shoubing, QU Yunfang, XU Ziran. Algal bloom control in eutrophic lakes and reservoirs based on biomanipulation [J]. *Water Resources Protection*, 2016, 32(5) :1-4. (in Chinese))
- [21] 刘聚涛, 杨永生, 高俊峰, 等. 太湖蓝藻水华灾害灾情评估方法初探 [J]. *湖泊科学*, 2011, 23(3) :334-338. (LIU Juntao, YANG Yongsheng, GAO Junfeng, et al. Comprehensive evaluation method of cyanobacteria bloom hazard in Lake Taihu [J]. *Journal of Lake Science*, 2011, 23(3) :334-338. (in Chinese))
- [22] 朱伟, 陈怀民, 王若辰, 等. 2017 年太湖水华面积偏大的原因分析 [J]. *湖泊科学*, 2019, 31(3) :621-632. (ZHU Wei, CHEN Huaimin, WANG Ruochen, et al. Analysis on the reasons for the large bloom area of Lake Taihu in 2017 [J]. *Journal of Lake Science*, 2019, 31(3) :621-632. (in Chinese))
- [23] LI Z, CHEN Q, XU Q. Modeling algae dynamics in Meiliang Bay of Taihu Lake and parameter sensitivity analysis [J]. *Journal of Hydro-environment Research*, 2015, 9(2) :216-225.
- [24] 郭静, 陈求稳, 李伟峰. 湖泊水质模型 SALMO 在太湖梅梁湾的应用 [J]. *环境科学学报*, 2012, 32(12) :3119-3172. (GUO Jing, CHEN Qiwen, LI Weifeng. Application of lake model SALMO to the Meiliang Bay of Taihu Lake [J]. *Acta Scientiae Circumstantiae*, 2012, 32(12) :3119-3127. (in Chinese))
- [25] 李志杰. 湖泊水生态模型多元数据同化 [D]. 北京:中国科学院大学, 2015.
- [26] RENS E G. Improving the algal bloom prediction in the North Sea by Ensemble Kalman filtering in the GEM/

- BLOOM model [D]. TU Delft; Delft University of Technology, 2013.
- [27] CHEN W, HUANG C, SHEN H, et al. Comparison of ensemble-based state and parameter estimation methods for soil moisture data assimilation[J]. *Advances in Water Resources*, 2015, 86: 425-438.
- [28] HUANG J, GAO J, LIU J, et al. State and parameter update of a hydrodynamic-phytoplankton model using ensemble Kalman filter[J]. *Ecological Modelling*, 2013, 263: 81-91.
- [29] 丁建丽, 陈文倩, 王璐. HYDRUS 模型与遥感集合卡尔曼滤波同化提高土壤水分监测精度[J]. *农业工程学报*, 2017, 33(14): 166-172. (DING Jianli, CHEN Wenjing, WANG Lu. Improving monitoring precision of soil moisture by assimilation of HYDRUS model and remote sensing-based data by ensemble Kalman filter[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2017, 33(14): 166-172. (in Chinese))
- [30] 孔繁翔, 高光. 大型浅水富营养化湖泊中蓝藻水华形成机理的思考[J]. *生态学报*, 2005, 25(3): 589-595. (KONG Fanxiang, GAO Guang. Hypothesis on cyanobacteria bloom-forming mechanism in large shallow eutrophic lakes[J]. *Acta Ecologica Sinica*, 2005(3): 589-595. (in Chinese))
- [31] 孔繁翔, 马荣华, 高俊峰, 等. 太湖蓝藻水华的预防、预测和预警的理论与实践[J]. *湖泊科学*, 2009, 21(3): 314-328. (KONG Fanxiang, MA Ronghua, GAO Junfeng, et al. The theory and practice of prevention, forecast and warning on cyanobacterial bloom in Lake Taihu [J]. *Journal of Lake Science*, 2009, 21(3): 314-328. (in Chinese))
- [32] 张艳会, 李伟峰, 陈求稳. 太湖水华程度及其生态环境因子的时空分布特征[J]. *生态学报*, 2016, 36(14): 4337-4345. (ZHANG Yanhui, LI Weifeng, CHEN Qiuwen, et al. Spatial-temporal variance of the intensity of algal bloom and related environmental and ecological factors in Lake Taihu[J]. *Acta Ecologica Sinica*, 2016, 36(14): 4337-4345. (in Chinese))
- [33] 秦伯强, 杨桂军, 马健荣, 等. 太湖蓝藻水华“暴发”的动态特征及其机制[J]. *科学通报*, 2016, 61(7): 759-770. (QIN Boqiang, YANG Guijun, MA Jianrong, et al. Dynamics of variability and mechanism of harmful cyanobacteria bloom in Lake Taihu [J]. *China Science Bulletin*, 2016, 61(7): 759-770. (in Chinese))
- [34] 李渊, 李云梅, 吕恒, 等. 太湖叶绿素 a 同化系统敏感性分析[J]. *湖泊科学*, 2015, 27(1): 163-174. (LI Yuan, LI Yunmei, LYU Heng, et al. Sensitivity analysis on Lake Taihu data assimilation scheme of chlorophyll-a concentration[J]. *Journal of Lake Science*, 2015, 27(1): 163-174. (in Chinese))
- [35] 师春香, 谢正辉, 钱辉, 等. 基于卫星遥感资料的中国区域土壤湿度 EnKF 数据同化[J]. *中国科学: 地球科学*, 2011, 41(3): 375-385. (SHI Chunxiang, XIE Zhenghui, QIAN Hui, et al. China land soil moisture EnKF data assimilation based on satellite remote sensing data [J]. *Scientia Sinica Terrae*, 2011, 41(3): 375-385. (in Chinese))
- [36] 朱广伟. 太湖水质的时空分异特征及其与水华的关系[J]. *长江流域资源与环境*, 2009, 18(5): 439-445. (ZHU Guangwei. Spatio-temporal distribution pattern of water quality in Lake Taihu and its relation with cyanobacterial blooms[J]. *Resources and Environment in the Yangtze Basin*, 2009, 18(5): 439-445. (in Chinese))
- [37] 金颖薇, 朱广伟, 许海, 等. 太湖水华期营养盐空间分异特征与赋存量估算[J]. *环境科学*, 2015, 36(3): 936-945. (JIN Yinwei, ZHU Guangwei, XU Hai, et al. Spatial distribution pattern and stock estimation of nutrients during bloom season in Lake Taihu [J]. *Environmental Science*, 2015, 36(3): 936-945. (in Chinese))
- [38] 张艳会, 李伟峰, 陈求稳. 太湖蓝藻水华发生风险区划[J]. *湖泊科学*, 2015, 27(6): 1133-1139. (ZHANG Yanhui, LI Weifeng, CHEN Qiuwen. Algae bloom risk zoning of Lake Taihu [J]. *Journal of Lake Science*, 2015, 27(6): 1133-1139. (in Chinese))
- [39] 邹逸航, 马旭林, 姜胜, 等. COSMIC 掩星资料同化对台风“天兔”预报影响的试验[J]. *海洋学研究*, 2017, 35(3): 9-19. (ZHOU Yihang, MA Xulin, JIANG Sheng, et al. Effect of COSMIC occultation data assimilation on prediction of typhoon Usage [J]. *Journal of Marine Sciences*, 2017, 35(3): 9-19. (in Chinese))
- [40] 陈元鹏, 任佳, 王力. 基于多源遥感数据的生态保护修复项目区监测方法评述[J]. *生态学报*, 2019, 39(23): 8789-8797. (CHEN Yuanpeng, REN Jia, WANG Li. Review on monitoring method of ecological conservation and restoration project area based on multi-source remote sensing data [J]. *Acta Ecologica Sinica*, 2019, 39(23): 8789-8797. (in Chinese))
- [41] 张成成. 藻类动力学模型和三维水流模型结合及其应用[D]. 北京: 中国科学院大学, 2014.

(收稿日期: 2020-07-07 编辑: 王芳)

