

基于遥感藻总量和气象因子的巢湖不同湖区藻华预测^{*}

刘东^{1**}, 张民², 沈明¹, 曹志刚¹, 齐天赐¹, 邱银国¹, 段洪涛^{1,3}

(1: 中国科学院南京地理与湖泊研究所, 中国科学院流域地理学重点实验室, 南京 210008)

(2: 中国科学院南京地理与湖泊研究所, 湖泊与环境国家重点实验室, 南京 210008)

(3: 西北大学城市与环境学院, 西安 710127)

摘要: 湖泊能为人类提供不可或缺的资源, 而全球普遍存在的湖泊富营养化导致的藻华频繁暴发正不断损害湖泊生态环境服务功能。为合理保护湖泊环境和防治藻华危害, 需预测藻华暴发。以我国富营养巢湖为研究区, 本文构建了一种基于遥感藻总量和气象因子的不同湖区藻华暴发概率预测方法。基于 MODIS/Aqua 数据, 研究首先反演了 2003—2019 年日尺度的藻华分布和考虑垂向结构的水柱藻总量。然后, 统计了西、中和东巢湖的藻华面积, 判别了藻华/非藻华日, 并匹配日平均藻总量和气象因子。最后, 筛选出藻华形成的关键影响因子——藻总量、气温和水汽压, 并构建了不同湖区日藻华暴发概率的 Logistic 预测模型。不同湖区月平均藻总量基本一致, 但藻华暴发日占比呈“西高东低”特征。对西、中和东巢湖的藻华/非藻华检验样本, 模型精度分别为 90%、85% 和 89.5%, 模型也适用于 2020 年夏秋季和冬春季藻华预测。湖泊藻华暴发是藻类大量增殖并在一定气象条件下的产物, 故基于遥感藻总量和气象因子的藻华暴发概率预测科学合理, 可推广应用到太湖等其他富营养湖泊。

关键词: 巢湖; 水柱藻总量; 遥感; 藻华概率; MODIS

Predicting algal blooms in different zones of Lake Chaohu based on satellite-derived algal biomass and meteorological factors^{*}

Liu Dong^{1**}, Zhang Min², Shen Ming¹, Cao Zhigang¹, Qi Tianci¹, Qiu Yinguo¹ & Duan Hongtao^{1,3}

(1: Key Laboratory of Watershed Geographic Sciences, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, P.R.China)

(2: State Key Laboratory of Lake Science and Environment, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, P.R.China)

(3: College of Urban and Environmental Sciences, Northwest University, Xi'an 710127, P.R.China)

Abstract: Lakes provide indispensable resources for humans, but the frequent outbreaks of algal blooms caused by eutrophication in lakes worldwide are continuously damaging the ecological services of lakes. In order to reasonably protect the lake resources and prevent the harm of algal blooms, it is necessary to predict the outbreak of algal blooms. Taking the typical eutrophic Lake Chaohu as the study area, this study constructed a method to predict the probability of algal blooms based on the satellite-derived water column-integrated algal biomass and meteorological factors. Based on MODIS/Aqua data, this study firstly retrieved the algal bloom distributions and water column-integrated algal biomass on different dates during 2003–2019. Then, we calculated the algal blooms area in the West Chaohu, Middle Chaohu and East Chaohu, identified the dates of algal blooms or non-algal blooms, and matched the corresponding satellite-derived water column-integrated algal biomass and meteorological factors data. Finally, we screened out the key factors affecting the formation of algal blooms including algal biomass, air temperature and vapor pressure, and constructed Logistic models for predicting daily algal blooms probability in different lake zones of Lake Chaohu. We found the monthly average algal biomass in different zones were almost the same in Lake Chaohu, but the proportion of days with algal blooms was high in the

* 2021-05-28 收稿; 2021-12-21 收修改稿。

国家自然科学基金项目(41901299, 42101433, 41971309)、中国科学院青年创新促进会项目(2021313)、江苏省自然科学基金项目(BK20181102)和中国科学院南京地理与湖泊研究所青年科学家小组项目(E1SL002)联合资助。

** 通信作者; E-mail: dliu@niglas.ac.cn.

west and low in the east. The accuracy of the model was 90%, 85% and 89.5% for West Chaohu, Middle Chaohu and East Chaohu, respectively, and the model was also applicable to the prediction of algal bloom in winter-spring and summer-autumn seasons in 2020. Algal blooms in lakes are the results of algal proliferation along with certain meteorological conditions, so the algal bloom probability prediction method based on remote sensing algal biomass and meteorological data is scientific and reasonable, and can be applied to other eutrophic lakes such as Lake Taihu.

Keywords: Lake Chaohu; water-column integrated algal biomass; remote sensing; algal bloom probability; MODIS

湖泊为人类生产生活提供饮用水、娱乐场所和栖息地等,是人类赖以生存的基础^[1]。然而,受人类排污和气候变化的影响,全球越来越多的湖泊呈现富营养化,一个典型结果就是藻华暴发呈增加态势^[2]。藻类大量繁殖后,藻颗粒腐烂分解会消耗溶解氧造成水体缺氧,蓝藻等还会释放藻毒素,因此藻含量是评价湖泊富营养化程度或水质好坏的一个关键指标,并通常用叶绿素a浓度(Chl.a)表征藻含量^[3-5]。关于富营养湖泊藻含量、藻华时空变异等,目前已有大量研究报道^[2,6],但为合理防治藻华带来的各种生态环境危害,还需提前预测藻华暴发情况,相关方面的报道还较少。湖泊藻华暴发是藻增殖至一定生物量后,在适宜气象条件下的产物^[6-7]。对富营养太湖,Shi等^[7]报道气温和总磷是藻类增殖的主要影响因素,而气压和风速是藻华暴发的主控因子。因此,在已知藻含量和气象条件时,可进行藻华暴发概率预测。基于耦合水动力藻含量物理模型,Li等^[8]先模拟出藻含量,再结合气象因子预测太湖未来3天的藻华概率。基于物理模型的藻华预测机理清晰,但模型输入参数繁多、构建难度极大,目前大多研究是利用不同气象因子构建藻华概率预测的经验模型,选用的气象因子包括气温、气压、降雨量和风速等^[9-12]。这些研究要么显式限定在藻华暴发频繁的夏秋季^[12],要么通过输入气温隐式限定藻华暴发于高温季节^[9-11],不适用于预测不同季节尤其是冬季藻华,究其原因是缺乏藻含量输入。

卫星遥感具有高时空分辨率优势,越来越多被应用于湖泊藻含量动态监测。针对单个湖泊或湖泊群,已经发展了波段比值、波段差值和分段拟合等多种表层藻含量的遥感算法^[2,6-7]。然而,富营养湖泊藻颗粒会随风速等环境条件在水柱内垂向迁移:低风速时会上浮至表层,高风速时在水柱内呈垂向均一分布^[13]。Xue等^[14]报道了富营养巢湖水柱藻含量剖面随风速等变化呈现均一、幂函数衰减、指数衰减或高斯分布类型。因此,当水柱藻含量分布不均时,遥感监测到的表层或真光层藻含量并不能真实反映湖泊富营养状况。为更准确监测湖泊藻含量,近年来一些研究开始考虑藻类在水柱内的不均一分布,并进行水柱藻总量遥感,通常用单位面积水域内的Chl.a垂向总含量表征水柱藻总量^[14-16];所采用方法大体分为两类:一类是先构建表层藻含量与水柱藻总量的经验关系式,再由遥感的表层藻含量估算藻总量,即面向结果方法^[15-16],该方法已应用于我国滇池和巢湖^[15-16];另一类是先实现水柱不同深度藻含量的函数化描述,再不同深度积分计算藻总量,即面向过程方法,该方法已应用于我国太湖和巢湖^[17]。

以富营养巢湖为研究区,本文首先分析了巢湖不同湖区藻华暴发与否日的藻总量和气象因子差异,然后构建了不同湖区藻华暴发概率的Logistic预测模型,用于计算不同气温、气压、风速和藻总量情形下的藻华暴发概率。综合考虑藻总量和气象因子,所构建的Logistic模型可用于预测巢湖不同季节、不同湖区的藻华暴发概率,对其藻华防治具有重要现实意义。

1 研究区与数据

1.1 研究区概况

巢湖位处长江中下游,湖心经纬度分别为117°34'E和31°32'N(图1),水面积约770 km²,平均水深约3.0 m,是我国第五大淡水湖^[15]。巢湖是合肥市的重要水源地,养育着流域约1080万人^[17]。环巢湖入湖河流主要有南淝河、派河、柘皋河、白石天河和杭埠河等,湖水能通过裕溪河排出并流入长江干流;其中,南淝河和派河水污染严重,携带大量营养盐入湖,杭埠河入湖流量最大、水质较好^[18]。受河流入湖污染物影响,巢湖水体富营养化程度由西至东逐渐降低,根据湖体富营养状态和生态环境特征,通常将巢湖分为西巢湖、中巢湖和东巢湖^[15,17]。

1.2 卫星遥感和站点实测数据

为获得历史时期的巢湖藻华分布和水柱藻总量,研究使用了搭载在美国Aqua卫星上的中分辨率成像

光谱仪(MODerate-resolution imaging spectroradiometer, MODIS/Aqua)扫描获取的2003—2020年日尺度多光谱影像,数据来源于NASA(<https://oceancolor.gsfc.nasa.gov/>)。原始数据利用SeaDAS软件空间重采样为250 m,并进行瑞利校正以去除大气分子的吸收和散射信号,从而获得瑞利校正后的遥感反射率(R_{rc})。 R_{rc} 数据共包含7个波段,波长范围分别为459~479 nm($R_{\text{rc}}(469)$)、545~465 nm($R_{\text{rc}}(555)$)、620~672 nm($R_{\text{rc}}(645)$)、841~891 nm($R_{\text{rc}}(859)$)、1230~1250 nm($R_{\text{rc}}(1240)$)、1628~1652 nm($R_{\text{rc}}(1640)$)和2105~2155 nm($R_{\text{rc}}(2130)$)。为避免内陆浑浊水体精确大气校正而带来更大误差,研究直接使用 R_{rc} 遥感数据进行藻华识别^[19]和水柱藻总量反演^[17]。

为判断水柱藻剖面类型和构建藻华预测模型,研究使用了湖区附近合肥站和巢湖站(图1)的气象日值数据,指标包括风速、气温、水汽压、气压、降水量和日照时数。数据来源于国家气象科学数据中心(<http://www.cma.gov.cn/>),时间范围为2003—2020年。为反映巢湖不同湖区水质,研究使用了湖体20个水质站点(图1)的月尺度总磷数据,水样通常在每月上旬采集,时间跨度为2003—2018年,数据来源于安徽省巢湖管理局(<http://chglj.hefei.gov.cn/>)。

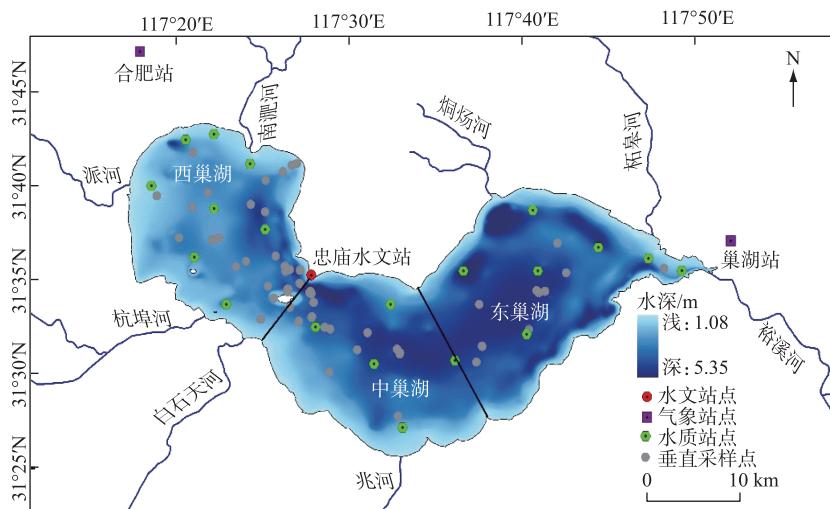


图1 巢湖及不同类型实测站点地理区位

Fig.1 Geographic locations of Lake Chaohu and different kinds of field stations

此外,研究还使用了2013—2017年在巢湖开展的7个航次获取的106个Chl.*a*剖面(图1)。站点均匀分布于全湖,但在藻华水域会加测以更多获取藻华水域的Chl.*a*剖面;此外,在一些点位还进行了时间序列观测以获取水柱Chl.*a*剖面时间动态特征;由于采样船定位不准、随水漂移等,站点位置以GPS记录为准,且不同航次点位会出现重合(图1)。对每个站位,采集表层、0.1、0.2、0.4、0.7、1.0、1.5、2.0和3.0 m共9个深度的水样,然后参照叶绿素测定国家行业标准(SL 88—2012)用Whatman GF/F(孔径为0.7 mm)滤膜过滤水样得到Chl.*a*样品^[17]。实验室内,用90%的丙酮萃取Chl.*a*,并用Shimadzu UV2600分光光度计测量Chl.*a*浓度。更多关于剖面Chl.*a*的获取过程,请参考Liu等^[17]。

2 研究方法

2.1 藻华遥感识别和暴发日占比计算

对于MODIS/Aqua获取的日尺度 R_{rc} 数据(1.2节),将 $R_{\text{rc}}(1640) > 0.0215$ 的水体判断为云覆盖区域^[20]。对无云覆盖水体,研究计算了浮游藻类指数(floating algae index, FAI)^[21]:

$$\begin{aligned} \text{FAI} &= R_{\text{rc}}(859) - R_{\text{rc}}'(859) \\ R_{\text{rc}}'(859) &= R_{\text{rc}}(645) + \frac{(859 - 645)}{(1240 - 859)}(R_{\text{rc}}(1240) - R_{\text{rc}}(645)) \end{aligned} \quad (1)$$

式中, $R_{\text{re}}(645)$ 、 $R_{\text{re}}(859)$ 和 $R_{\text{re}}(1240)$ 分别表示中心波长为 645、859 和 1240 nm 3 个波段的反射率, “859–645”和“1240–859”为对应中心波长的差值^[21]. FAI 反映的是一个像元覆盖的 250 m×250 m 水域的藻华综合情况, 为避免混合像元造成藻华面积高估, 参照 Duan 等^[6]使用阈值 0.02 提取纯藻华像元, 即: $FAI > 0.02$ 的像元对应藻华水体, 否则为非藻华水体. 根据 2003–2019 年日尺度藻华遥感结果, 研究进一步计算了巢湖不同像元、不同月份藻华暴发日占比(藻华暴发天数/总无云观测天数×100%). 由于云覆盖存在时空变异, 藻华暴发日占比(%)比藻华暴发天数能更好地反映不同月份藻华暴发情况.

2.2 水柱藻总量遥感估算

富营养湖泊水柱内的藻会随气象条件等发生垂向迁移, 低风速时会上浮聚集在水表形成藻华, 而风速大于一定阈值时会呈均匀混合^[13]. 由于遥感只能观测到水表或真光层内藻含量, 为遥感反演水柱藻总量可通过两种途径:面向结果或面向过程方法^[15–17]. 基于实测数据, Liu 等^[17]构建了巢湖水柱藻总量遥感的面向过程方法, 精度优于面向结果方法, 尤其是对水表藻类聚集水域. 因此, 本文采用 Liu 等^[17]提出的面向过程方法遥感反演巢湖藻总量, 主要步骤如下:

1) 表层 Chl.*a* 浓度遥感反演. 基于富营养太湖的表层 Chl.*a* 遥感算法^[7], 并用巢湖的星地同步数据进行重新率定, Liu 等^[17]给出了巢湖的 Chl.*a* 遥感算法($N = 196, R^2 = 0.59$). 算法均方根误差和平均相对误差分别为 25.97 mg/m³ 和 -5.61%, 可同时反演藻华/非藻华水域 Chl.*a*^[7], 公式如下:

$$\begin{aligned} \text{Chl.}a &= -5399.99 \text{Ratio} + 187.17 \\ \text{Ratio} &= (\exp(R_{\text{re}}(645)) - \exp(R_{\text{re}}(859))) / (\exp(R_{\text{re}}(645)) + \exp(R_{\text{re}}(859))) \end{aligned} \quad (2)$$

式中, Ratio 是一个自定义的归一化指数, 用以表征水体藻含量, 藻含量越多 Ratio 值越低.

2) 藻类剖面分布函数定义与参数化. 为减少待定系数量, Liu 等^[17]通过分析巢湖 106 个实测 Chl.*a* 剖面, 认为可将 Chl.*a* 剖面概化为垂向均一和幂函数衰减两类(图 2), 如式(3):

$$\text{Chl.}a(z) = \begin{cases} C_0, & \text{均一分布剖面} \\ n_1 \times z^{n_2}, & \text{幂函数衰减剖面} \end{cases} \quad (3)$$

式中, Chl.*a*(*z*) 为水深 *z* 处的 Chl.*a* 浓度; *C*₀ 为遥感得到的表层 Chl.*a* 浓度; *n*₁ 和 *n*₂ 为待定系数. 系数 *n*₂ 与表层 Chl.*a* 浓度存在显著对数关系, 故可由遥感的表层 Chl.*a* 浓度计算 *n*₂; 在已知 *n*₂ 和表层(水深为 0.01 m) Chl.*a* 浓度时, 可计算得到系数 *n*₁; 更多关于 *n*₁ 和 *n*₂ 的计算过程, 请参考 Liu 等^[17].

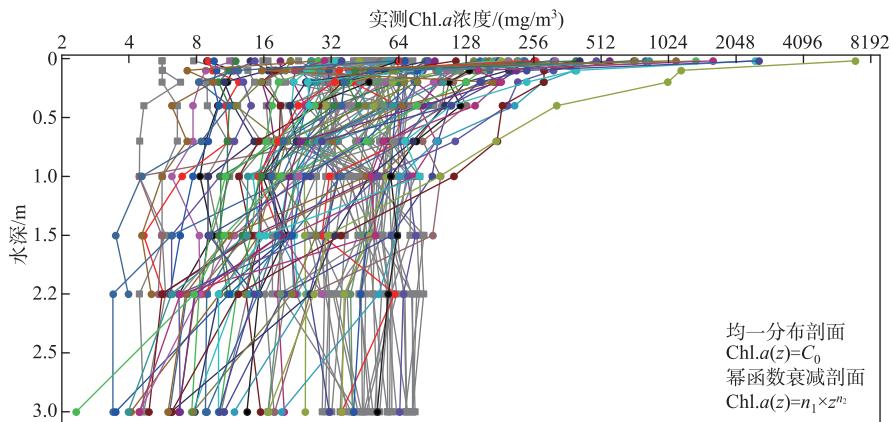


图 2 巢湖不同站点的实测 Chl.*a* 剖面(均一分布剖面用灰色表示, 幂函数衰减剖面用不同颜色表示)

Fig.2 The *in-situ* Chl.*a* profiles at different stations in Lake Chaohu
(Uniform profiles are shown in grey, Power decay profiles are shown in different colors)

3) 藻类剖面分布类型及水柱藻总量遥感. 对任一遥感像元, 首先用二叉决策树判断藻类剖面分布类型: 如果 FAI > 0.02 则为幂衰减分布; 如果 FAI ≤ 0.02 且表层 Chl.*a* ≤ 10 mg/m³, 则为均一分布; 当 FAI ≤ 0.02 且 Chl.*a* > 10 mg/m³ 时, 如果风速 ≤ 2.75 m/s 则为幂衰减分布, 否则为均一分布. 然后, 通过步骤(2)参数化

藻类剖面分布函数,并结合水深(*depth*,水位减湖泊 DEM 得到)积分得到单位水域面积的水柱藻总量:

$$[\text{水柱藻总量}] = \int_0^{\text{depth}} \text{Chl.a}(z) dz \quad (4)$$

式中,对不同深度 Chl.a 浓度积分得到的单位面积水柱藻总量单位为 mg/m². 对富营养巢湖,上述方法遥感估算水柱藻总量的平均绝对百分比误差为 31.03%,更多信息请参考 Liu 等^[17].

2.3 湖区藻华暴发概率预测模型构建

Logistic 模型又称 Logistic 回归分析,以事件发生与否为因变量,以连续或离散型因子为自变量,将二元因变量与自变量之间的联系转换为因变量条件概率同自变量之间的关系问题,在医疗卫生和灾害预测等领域应用较多^[10-12]. 本文以 2003–2019 年的遥感水柱藻总量和气象因子日平均值为自变量,以藻华发生与否为因变量,构建了巢湖西、中、东不同湖区藻华暴发概率的 Logistic 预测模型,主要步骤如下:

1)湖区藻华/非藻华暴发日确定. 参照已发表的藻华判别方法^[11],研究将 MODIS/Aqua 遥感监测到的藻华面积占湖区总面积比 > 5% 的日期确定为藻华暴发日;同时,为避免云覆盖影响,将湖区无云覆盖面积比 > 70% 且藻华面积比 ≤ 5% 的日期确定为非藻华暴发日. 理论上,可以用上述确定的所有样本构建模型,但由于遥感判断藻华存在不确定性,藻华像元少的藻华暴发日和藻华像元多的非藻华暴发日可能是由于遥感误判. 为规避遥感误判的影响,同时确保足够样本量,研究只选取藻华像元多的前 30% 作为藻华暴发日样本,赋值为 1;同时,只选取藻华像元少的前 30% 作为非藻华暴发日样本,赋值为 0.

2)湖区藻华/非藻华暴发日气象因子匹配. 根据步骤(1)确定的藻华/非藻华暴发日,提取对应的日平均水柱藻总量、气温和水汽压等气象因子值. 其中,平均水柱藻总量为各湖区不同像元遥感结果的平均值;对气象因子,西巢湖用合肥站的值,东巢湖用巢湖站的值,而中巢湖用两个站的均值(图 1).

3)湖区藻华暴发概率 Logistic 模型构建. 统计分析藻华/非藻华暴发日对应的藻总量和各气象因子值,判断藻华暴发与否日的差异,筛选出湖区藻华暴发概率估算的自变量最优组合,然后基于 SPSS 构建湖区藻华暴发概率 Logistic 模型,形式如下:

$$P = \frac{\exp(\beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \cdots + \beta_k \times x_k)}{1 + \exp(\beta_0 + \beta_1 \times x_1 + \beta_2 \times x_2 + \cdots + \beta_k \times x_k)} \quad (5)$$

式中,P 为藻华暴发概率,用 0~1 区间值表示;x₁、x₂…x_k为输入变量;β₀、β₁、β₂…β_k为拟合系数. 采用 Cox & Snell 和 Nagelkerke 伪决定系数 R²为模型拟合优度的检验标准(越临近 0.5 效果越好),并采用 Wald 统计量检验各自变量对藻华暴发的影响程度(值越大相对影响越大).

式(5)计算得到的藻华暴发概率表示湖区藻华暴发的可能性,值越大表示藻华暴发的可能性越大. 基于历史数据对不同湖区构建了如式(5)的藻华暴发概率 Logistic 模型后,假设不同湖区藻总量在短期内稳定(蓝藻生长周期是 8~10 d^[22]),则可以将遥感监测到的近几日平均藻总量和气象预报资料输入模型,计算得到未来几日的藻华暴发概率预测值.

3 结果分析

3.1 巢湖环境特征

巢湖环境因子存在明显的季节变异特征. 气温和水汽压夏季(6~8 月)高、冬季(12~2 月)低,年内呈明显单峰变异特征:7 月气温为(28.86±2.83)℃,水汽压为(30.97±2.67) hPa;1 月气温为(3.10±3.13)℃,水汽压为 5.69±11.88 hPa(图 3a~b). 风速季节差异较小,整体呈冬春季高、夏秋季低,但同一月份内风速变异很大,这反映出巢湖风速日变化剧烈(图 3c). 总磷浓度在夏秋季高、冬春季低;西巢湖 8 月总磷浓度((0.2±0.05) mg/L)能达 1 月((0.1±0.04) mg/L)的 2 倍,中巢湖总磷浓度季节变异弱于西巢湖,东巢湖季节变异更弱;值得注意的是,总磷浓度高的夏秋季其变异也大. 不同环境因子的季节变异共同驱动巢湖藻华的季节变异.

不同环境因子的湖区差异不一致. 不同湖区的气候态月平均气温和水汽压等差异很小;虽然风速湖区差异也不明显,但西部湖区的整体略低于东部湖区(图 3a~c). 西巢湖的气候态月平均总磷浓度明显高于中巢湖,中巢湖比东巢湖高;在 4 月,西巢湖总磷浓度((0.16±0.06) mg/L)是中巢湖((0.1±0.04) mg/L)的 1.6 倍,是东巢湖((0.08±0.03) mg/L)的 2 倍. 值得注意的是:虽然不同湖区的气候态月平均环境因子空间

差异不明显,但在一些具体日期仍差异明显,尤其是风速.

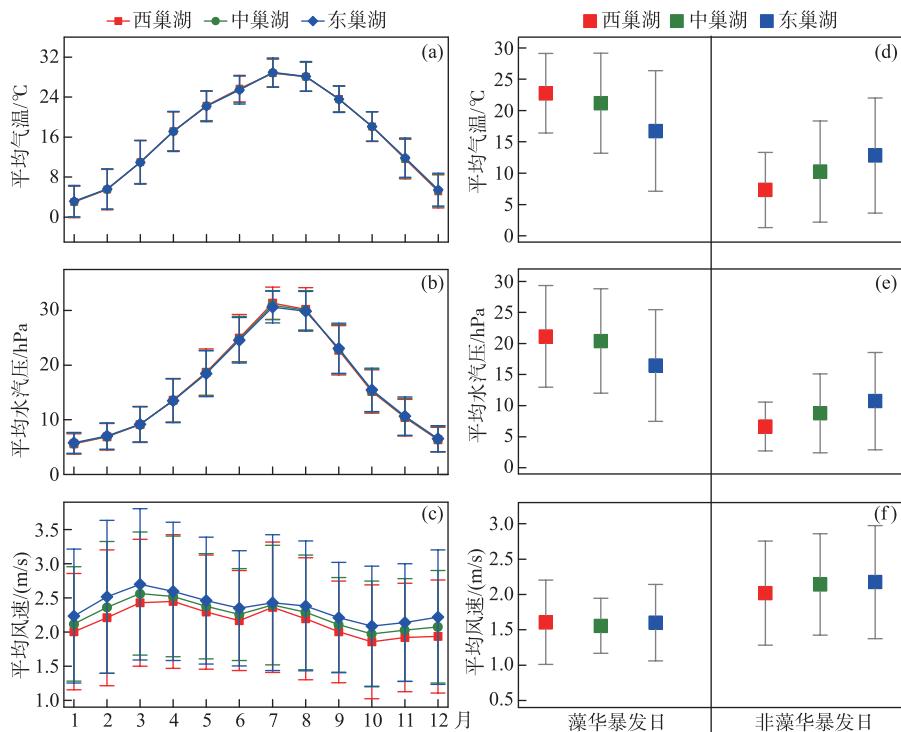


图3 不同湖区气温、水汽压和风速气候态平均值:(a~c)按月份;(d~f)按藻华或非藻华日

Fig.3 Climatological mean values of air temperature, water vapor pressure and wind speed in different lake zones according to (a-c) months and (d-f) algal blooms or not

3.2 巢湖水柱藻总量遥感结果

图4为巢湖气候态月平均水柱藻总量. 水柱藻总量在巢湖南部近岸浅水区呈明显低值, 藻总量低的冬季月份甚至低于 10 mg/m^2 . 而在西巢湖北部和东巢湖东部近岸水域, 呈现出较高藻总量, 如2月和4月(图4). 此外的其他水域, 气候态月平均水柱藻总量并未表现明显空间差异, 这可能与巢湖强水动力导致的水藻横向输运和强烈混合有关. 由遥感反演的2003—2019年不同日期的水柱藻总量计算的不同湖区气候态月平均值也未表现明显差异: 西巢湖平均藻总量为 $(66.95 \pm 27.6) \text{ mg/m}^2$, 中巢湖为 $(63.07 \pm 26.84) \text{ mg/m}^2$, 东巢湖为 $(67.11 \pm 25.38) \text{ mg/m}^2$. 虽然不同湖区气候态月平均水柱藻总量差异小, 但一些日期的不同湖区水柱藻总量仍表现出明显差异. 如图5a所示, 当平均藻总量高于 100 mg/m^2 时, 不同湖区的藻总量表现出明显差异, 这可能是不同湖区藻华暴发差异的一个重要原因.

不同湖区水柱藻总量季节变异显著. 平均藻总量在冬季低值时小于 10 mg/m^2 , 而在夏季高值时能达 200 mg/m^2 (图5a). 由2003—2019年日尺度藻总量可知: 高藻总量不会持续很久, 大多时间湖区平均藻总量为 60 mg/m^2 上下. 利用Census X-12方法对2003—2019年的月平均藻总量进行时间序列分解, 可得: 巢湖不同湖区藻总量年内呈“双峰、双谷”的季节变异特征, 主次双峰出现于8月和3月前后, 主次双谷出现于12月和6月前后(图5b). 对西巢湖, 8月和3月的气候态平均藻总量分别为 (90.12 ± 44.92) 和 $(78.24 \pm 45.8) \text{ mg/m}^2$, 12月和6月的气候态平均藻总量分别为 (51.96 ± 21.7) 和 $(61.7 \pm 25.7) \text{ mg/m}^2$. 巢湖不同湖区藻总量的季节变异规律与气温、总磷的季节变化特征存在一定相似性(3.1节).

3.3 藻华暴发日占比时空变异

虽然巢湖不同湖区藻总量差异较小(3.2节), 但不同月份的藻华暴发日占比存在明显空间差异. 整体而

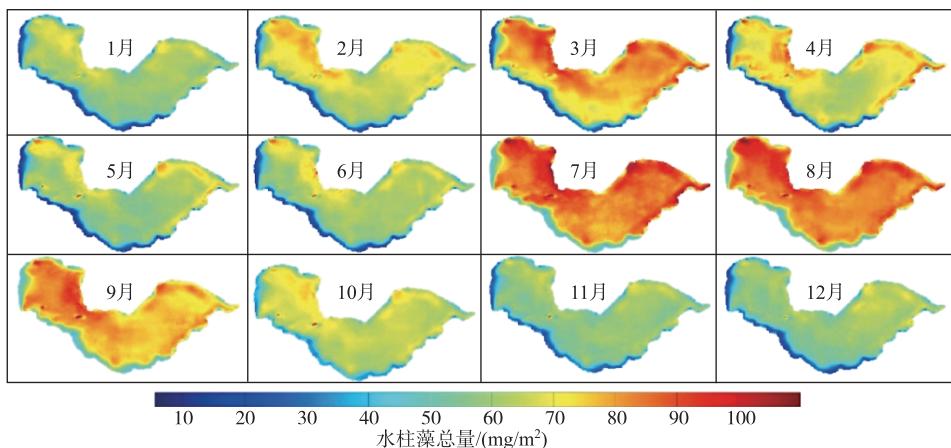


图 4 巢湖 2003—2019 年气候态月平均水柱藻总量

Fig.4 The climatological monthly mean algal biomass of Lake Chaohu during 2003–2019

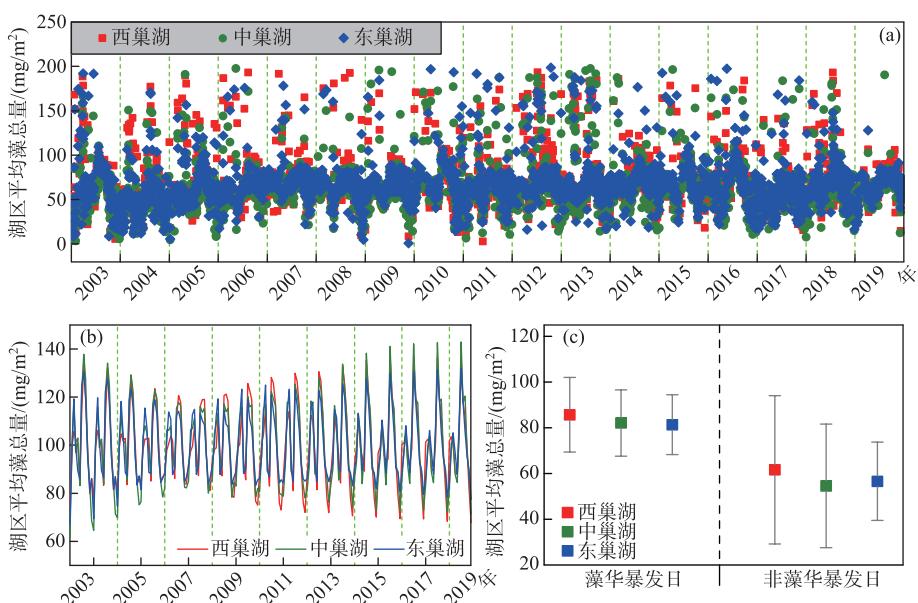


图 5 不同湖区平均值:(a) 日藻总量;(b) 时间序列分解后的月藻总量;(c) 藻华或非藻华日气温

Fig.5 Statistical mean results for different lake zones: (a) daily algal biomass, (b) monthly algal biomass after time-series decomposition, and (c) air temperature when algal blooms occurred or not

言,巢湖藻华暴发日占比呈明显“西高东低”特征,尤其是藻华暴发严重的夏秋季(图 6)。由 2003—2019 年不同月份的藻华暴发日占比可知:冬春季藻华暴发少,不同湖区藻华暴发日占比基本都<10%,且空间差异不明显;夏秋季藻华暴发明显多,尤其是西巢湖,8—9 月份一些水域的藻华暴发日占比甚至>30%;此外,西巢湖夏秋季不同月份藻华暴发日占比高值区存在不一致,夏季 6—8 月高值区集中在北部水域,而秋季 9—11 月在西南水域也会出现高值区(图 6)。

不同于水柱藻总量的“双峰、双谷”季节特征(3.2 节),巢湖不同湖区、月份的藻华暴发日占比均呈明显“单峰、单谷”特征,尤其是藻华暴发严重的西巢湖(图 6)。由气候态月平均结果可知:不同湖区藻华暴发日占比 9 月份达到峰值,1 月达到谷值;在 9 月和 1 月,西巢湖的藻华暴发日占比分别为 $27.9\% \pm 6.62\%$ 和

$1.85\% \pm 3.37\%$, 中巢湖分别为 $13.28\% \pm 6.09\%$ 和 $1.58\% \pm 1.35\%$, 东巢湖分别为 $3.47\% \pm 2.16\%$ 和 $1.37\% \pm 0.75\%$ (图 6). 此外, 不同月份藻华暴发日占比变异大(标准差大), 这反映出巢湖不同年份的藻华暴发情况存在显著差异, 尤其是藻华暴发严重的西巢湖和中巢湖(图 6).

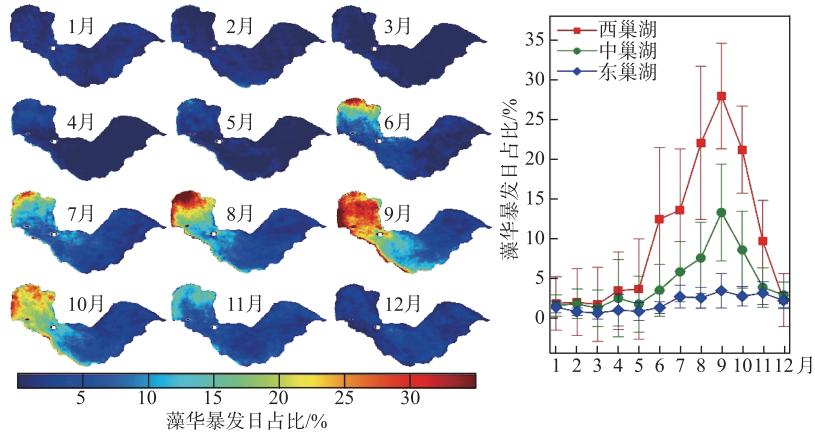


图 6 2003—2019 年气候态月平均藻华暴发日占比

Fig.6 The climatological monthly mean proportions of algal blooms days during 2003–2019

3.4 不同湖区藻华暴发概率预测模型

为确定巢湖不同湖区藻华暴发概率 Logistic 预测模型的自变量, 研究首先分析了藻华暴发与否日不同因子的差异. 由藻华/非藻华暴发日与环境因子的历史匹配数据可得(2.3 节), 不同湖区藻华/非藻华暴发日的平均气温、平均水汽压和平均藻总量存在明显差异, 尤其是藻总量(图 3d~f, 5c):对西巢湖, 藻华暴发日的平均气温、水汽压和藻总量分别为 22.75°C 、 21.14 hPa 和 85.73 mg/m^2 , 而非藻华暴发日的值分别为 7.33°C 、 6.63 hPa 和 61.65 mg/m^2 ; 对中巢湖, 藻华暴发日的平均气温、水汽压和藻总量分别为 21.18°C 、 20.41 hPa 和 82.1 mg/m^2 , 而非藻华暴发日的值分别为 10.27°C 、 8.76 hPa 和 54.59 mg/m^2 ; 对东巢湖, 藻华暴发日的平均气温、水汽压和藻总量分别为 16.74°C 、 16.47 hPa 和 81.37 mg/m^2 , 而非藻华暴发日的值分别为 12.82°C 、 10.72 hPa 和 56.62 mg/m^2 . 因此, 本研究以日平均气温、水汽压和水柱藻总量为自变量, 通过随机选取 2/3 历史匹配样本(西巢湖 $N=220$, 中巢湖 $N=336$, 东巢湖 $N=322$)构建了不同湖区的藻华暴发概率 Logistic 预测模型, 如式(6) :

$$\begin{aligned}
 P_1 &= \frac{\exp(-5.237 + 0.13 \times [\text{平均气温}] + 0.187 \times [\text{平均水汽压}] + 0.012 \times [\text{平均藻总量}])}{1 + \exp(-5.237 + 0.13 \times [\text{平均气温}] + 0.187 \times [\text{平均水汽压}] + 0.012 \times [\text{平均藻总量}])} \times 100\% \\
 P_2 &= \frac{\exp(-5.069 - 0.058 \times [\text{平均气温}] + 0.206 \times [\text{平均水汽压}] + 0.028 \times [\text{平均藻总量}])}{1 + \exp(-5.069 - 0.058 \times [\text{平均气温}] + 0.206 \times [\text{平均水汽压}] + 0.028 \times [\text{平均藻总量}])} \times 100\% \\
 P_3 &= \frac{\exp(-10.5 - 0.133 \times [\text{平均气温}] + 0.13 \times [\text{平均水汽压}] + 0.125 \times [\text{平均藻总量}])}{1 + \exp(-10.5 - 0.133 \times [\text{平均气温}] + 0.13 \times [\text{平均水汽压}] + 0.125 \times [\text{平均藻总量}])} \times 100\%
 \end{aligned} \quad (6)$$

式中, P_1 、 P_2 和 P_3 分别为西、中和东巢湖的藻华暴发概率, 乘以 100% 则取值范围为 0~100%, 值越大表示湖区藻华暴发概率越大. 利用模型训练之外的剩余 1/3 匹配样本检验模型精度: 如果式(6)得到的值 $> 50\%$ 则预测为藻华暴发, 反之为不暴发; 对西、中和东巢湖, 式(6)预测藻华暴发与否的精度分别为 90% ($N=110$)、85% ($N=168$) 和 89.5% ($N=161$).

模型拟合优度检验结果表明: 式(6)很好地拟合了建模数据, 且水汽压和水柱藻总量对不同湖区藻华暴发均有显著影响($P<0.05$). 对西巢湖, Cox & Snell 和 Nagelkerke R^2 分别为 0.55 和 0.74, 日平均气温、水汽压和水柱藻总量 3 个自变量影响均显著($P<0.05$); Wald 值分别为 6.84、9.02 和 6.59, 即水汽压影响最大. 对中巢湖, Cox & Snell 和 Nagelkerke R^2 分别为 0.32 和 0.49, 水汽压和藻总量影响显著($P<0.05$); Wald 值分别为 22.69 和 39.03, 即藻总量影响更大. 对东巢湖, Cox & Snell 和 Nagelkerke R^2 分别为 0.23 和 0.45, 3 个自变量影

响均显著 ($P < 0.05$) ; Wald 值分别为 7.1、6.24 和 53.01, 即藻总量影响明显大.

3.5 模型应用案例分析

虽然巢湖藻华暴发主要发生在夏秋季, 但冬春季也时有藻华暴发(图 6). 为检验式(6)对不同季节藻华暴发的预测能力, 本文对比分析了 2020 年冬春季和夏秋季巢湖藻华遥感结果和预测的藻华暴发概率. 基于 MODIS/Aqua R_{re} 数据, 研究得到了 2020 年冬春和夏秋季少云覆盖的 4 天的藻华和水柱藻总量分布, 如图 7 所示. 巢湖在这 4 天内均暴发了藻华, 但不同日期藻华空间分布差异大: 冬春季的 2 月 9 日和 12 日, 藻华主要分布在中、东巢湖, 总面积分别为 419.00 和 424.69 km^2 ; 夏秋季的 6 月 17 日和 7 月 25 日, 藻华主要分布在西巢湖, 总面积分别为 97.31 和 175.56 km^2 (图 7). 不同日期遥感反演的全湖平均藻总量分别为 89.29、87.43、67.50 和 108.64 mg/m^2 , 整体表现为高藻华面积、高藻总量, 且藻华暴发水域水柱藻总量通常也较高(图 7).

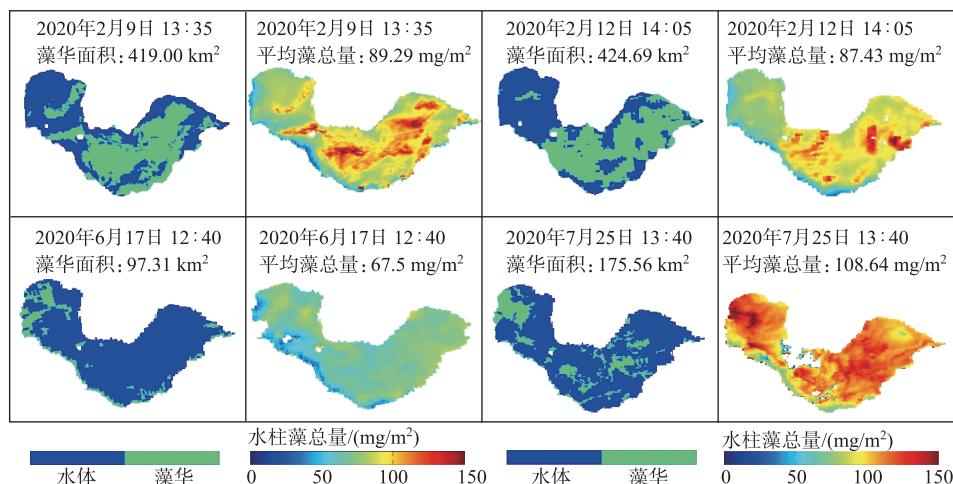


图 7 2020 年不同日期遥感藻华与藻总量

Fig.7 The satellite-derived algal blooms and algal biomass on different dates in 2020

将日平均气温、水汽压和藻总量代入式(6), 本文进一步计算了上述不同日期、不同湖区的藻华暴发概率, 如表 1 所示. 冬春季 2 月 9 日和 12 日的藻华暴发概率由西湖区向东湖区递增, 对藻华严重的东巢湖计算的暴发概率分别为 86.6% 和 85.0%, 而藻华少的西巢湖分别为 7.7% 和 19.5%; 夏秋季 6 月 17 日和 7 月 25 日的藻华暴发概率由西湖区向东湖区递减, 对藻华严重的西巢湖计算的暴发概率分别为 95.3% 和 98.2% (表 1). 因此, 不同湖区藻华暴发概率计算结果和遥感监测结果整体吻合, 在中巢湖存在一定误差(图 7, 表 1). 中巢湖藻华暴发概率预测误差偏大可能有两方面原因: 一是模型输入的气象因子采用的是合肥站和巢湖站的平均值, 与实际值可能存在一定偏差; 二是模型输入没有考虑风速, 而风速也是巢湖藻华暴发与否的一个重要影响因子^[14,17].

表 1 2020 年不同日期藻华暴发概率模型计算结果

Tab.1 The modeled results of algal blooms probability on different dates in 2020

日期	西巢湖				中巢湖				东巢湖			
	气温/ ℃	水汽压/ hPa	藻总量/ mg/m^2	藻华暴发 概率/%	气温/ ℃	水汽压/ hPa	藻总量/ mg/m^2	藻华暴发 概率/%	气温/ ℃	水汽压/ hPa	藻总量/ mg/m^2	藻华暴发 概率/%
2020-02-09	4.2	6.7	79.1	7.7	4.7	7.2	91.2	21.2	5.1	7.6	96.5	86.6
2020-02-12	8.7	9.6	74.7	19.5	8.9	10.3	90.4	28.2	9.1	11.0	96.1	85.0
2020-06-17	26.5	21.5	63.9	95.3	27.0	21.3	65.9	40.2	27.4	21.1	72.5	8.8
2020-07-25	24.2	25.4	110.1	98.2	24.0	25.6	101.1	83.7	23.8	25.7	113.4	97.9

4 讨论

4.1 巢湖藻华暴发概率预测

本文所构建的公式(6)可应用于巢湖不同湖区次日($t+1$)藻华暴发概率预测。巢湖藻华暴发的主控藻种是蓝藻,其生长周期为8~10 d^[22],即可以假设短期内湖区藻总量不变,然后将今日(t)的遥感藻总量作为 $t+1$ 日的藻总量输入模型。值的注意的是,由于卫星遥感易受云雨覆盖影响,经常会导致 t 日的藻总量没有遥感结果;此种情况下,为预测 $t+1$ 日的藻华暴发概率,可以使用前几日的湖区藻总量遥感平均值为输入。对于 $t+1$ 日的气温和水汽压,可以使用欧洲中期天气预报中心(European Centre for Medium-Range Weather Forecasts, ECMWF)发布的模式预测数据集(时间分辨率为3 h,空间分辨率为0.125度),数据集由中央气象局统一分发至各省市气象局,巢湖区域数据可以从安徽省气象局获取。

根据富营养状态,本文只将巢湖粗略分为3个湖区,而实际应用中会关注更小区域的藻华暴发情况,如自来水厂取水口、临近市镇的近岸水域等。为实现这些重点区域的藻华暴发概率预测,可以对巢湖进行更精细化划分,然后基于历史气象因子和遥感藻总量,根据2.3节方法,分别构建各重点水域藻华暴发概率Logistic预测模型。此外,当假设某一水域藻总量不变,也可预测未来几日的藻华暴发概率,但不同区域水藻交换导致的藻颗粒水平迁移会使模型预测精度随区域变小和预测时间变长而降低。

4.2 巢湖藻华预测精度影响因素

水柱藻总量是巢湖藻华暴发的关键影响因素,尤其是对东巢湖,藻总量遥感精度会直接影响藻华预测结果(3.4节)。本文以近几日的遥感藻总量输入模型计算 $t+1$ 日的藻华概率,一个重要依据是巢湖水动力强,进而可以假设三大湖区的日平均藻总量一致、短期内湖区藻总量也不变(3.3节)。由图8可知,虽然三大湖区平均藻总量在大部分时间基本一致,但仍有一些日期存在明显差异;如果这些日期不同湖区水藻交换明显,则会导致输入模型的藻总量与实际情况不符,进而影响藻华预测结果。因此,为提高藻华预测精度,需要保证模型输入藻总量的真实性:一方面可通过算法优化提高藻总量遥感精度^[17];另一方面可使用静止卫星数据以缩短藻总量遥感与藻华预测的时间差。

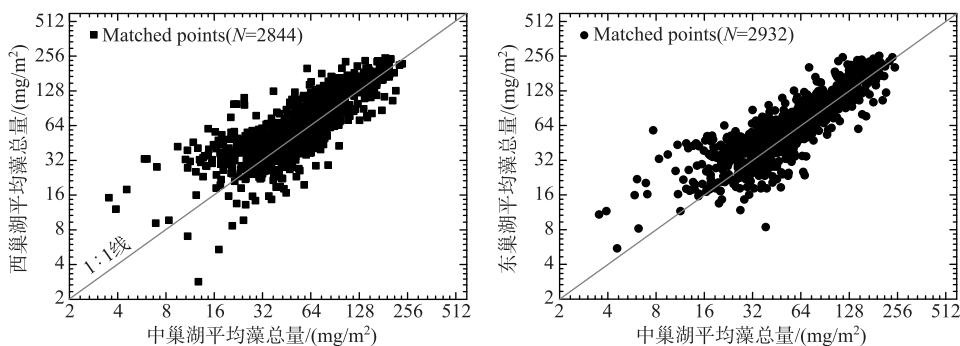


图8 2003—2019年巢湖不同湖区遥感日平均藻总量对比

Fig.8 Comparisons of satellite-derived daily mean algal biomass in different lake zones during 2003–2019

营养盐、日照时长和气温等会影响藻类增殖^[6,7],风速、水汽压和降雨等会影响藻类迁移^[7,14],因此这些环境因子也会影响藻华暴发。对富营养化湖泊,Webster等^[13]报道存在风速阈值,低于该阈值时藻类会大量上浮至表层形成藻华,而高于该阈值时水柱内藻类会趋于均匀混合。陈莉琼等^[12]报道日照时长、前三日气压差、风速和降雨量是影响富营养滇池藻华形成的关键气象因素,可用于预测滇池藻华暴发概率。然而,营养盐缺乏日尺度资料,风速和降雨在一天内不同小时会剧烈变动,藻华暴发与否日照时长没明显差异,因此本文的藻华概率预测模型只输入了气温和水汽压。不同湖区模型通过调整气温和水汽压系数,充分利用两者在藻华/非藻华暴发日差异,以更好预测藻华暴发概率。此外,由于缺乏湖区气象数据,使用的是附近合肥站和巢湖站的气象资料,但不同位置的气象因子可能存在明显时空差异。因此,为进一步提高巢湖藻华暴发

预测精度,一方面需使用湖体高频气象资料;另一方面,模型构建时需同时考虑影响藻类生长和分布的风速和日照时长等。

本文是以水柱积分 Chl.*a* 表征藻总量,进而构建 Logistic 模型预测巢湖不同季节的藻华暴发概率。但不同季节藻华暴发时藻总量存在差异(表 1)、优势藻种会不一致^[23],藻类不同生长阶段 Chl.*a* 含量也会不一致,这些都会影响用 Chl.*a* 表征藻总量的准确度,进而影响藻华预测模型的通用性。为提高巢湖不同季节的藻华暴发概率预测精度,有必要考虑季节差异,分季节构建藻华预测模型。

4.3 对其他湖泊藻华预测的启示

藻华广泛发生于全球湖泊^[2],为合理防治藻华危害,需对不同区位、不同类型湖泊藻华提前预测。本文研究的是富营养巢湖的藻华预测,蓝藻是主控性藻种,在藻总量、水汽压和气温满足一定条件后会上浮至表层形成藻华。因此,对本文构建的藻华暴发概率预测模型进行参数重新率定,可推广应用到全球其他蓝藻主控的富营养湖泊,如太湖。而对非蓝藻主控的富营养湖泊,由于藻华暴发的控制机制不同,藻华暴发预测的敏感因子也不同,不能简单通过参数率定将本文构建的模型应用于其藻华暴发预测。但湖泊藻华暴发都是藻类增殖至一定生物量后,在适宜气象条件下的产物^[6-7]。因此,对非蓝藻主控的其他富营养湖泊,仍可参照 2.3 节方法,以藻总量和气象因子为输入构建相应藻华暴发概率 Logistic 预测模型。

5 结论

以我国长江中下游典型富营养巢湖为研究区,本文首先分析了不同湖区水柱藻总量、藻华暴发日占比及藻华/非藻华暴发日各气象因子差异,然后构建了巢湖不同湖区藻华暴发概率的 Logistic 预测模型,模型输入参数包括遥感的水柱藻总量、日平均气温和日平均水汽压。巢湖不同湖区藻华暴发存在显著时空变异特征,模型预测结果与遥感历史监测结果基本一致;模型也能预测 2020 年冬春季和夏秋季藻华暴发,基本满足藻华预测的应用需求。虽然,基于遥感获取的前几日水柱藻总量和预测的日平均气温及水汽压可以实现巢湖不同湖区藻华暴发概率预测,但藻华暴发和气象因子的高时空动态导致的多源数据时空匹配难会降低模型预测精度,今后还需进一步结合湖体高时空分辨的多源气象因子进行模型优化和藻华预测。

致谢:野外实验数据的获取得到了中国科学院南京地理与湖泊研究所湖泊环境遥感团队各位老师和研究生的支持,在此表示衷心感谢。

6 参考文献

- [1] Liu D, Duan HT, Loiselle S et al. Observations of water transparency in China's lakes from space. *International Journal of Applied Earth Observation and Geoinformation*, 2020, **92**: 102187. DOI: 10.1016/j.jag.2020.102187.
- [2] Ho JC, Michalak AM, Pahlevan N. Widespread global increase in intense lake phytoplankton blooms since the 1980s. *Nature*, 2019, **574**(7780): 667-670. DOI: 10.1038/s41586-019-1648-7.
- [3] Carlson RE. A trophic state index for lakes. *Limnology and Oceanography*, 1977, **22**(2): 361-369. DOI: 10.4319/lo.1977.22.2.0361.
- [4] Liu D, Du YX, Yu SJ et al. Human activities determine quantity and composition of dissolved organic matter in lakes along the Yangtze River. *Water Research*, 2020, **168**: 115132. DOI: 10.1016/j.watres.2019.115132.
- [5] Zhang M, Shi XL, Yang Z et al. The variation of water quality from 2012 to 2018 in Lake Chaohu and the mitigating strategy on cyanobacterial blooms. *J Lake Sci.*, 2020, **32**(1): 11-20. DOI: 10.18307/2020.0102. [张民, 史小丽, 阳振等. 2012—2018 年巢湖水质变化趋势分析和蓝藻防控建议. 湖泊科学, 2020, **32**(1): 11-20.]
- [6] Duan HT, Tao M, Loiselle SA et al. MODIS observations of cyanobacterial risks in a eutrophic lake: Implications for long-term safety evaluation in drinking-water source. *Water Research*, 2017, **122**: 455-470. DOI: 10.1016/j.watres.2017.06.022.
- [7] Shi K, Zhang YL, Zhou YQ et al. Long-term MODIS observations of cyanobacterial dynamics in Lake Taihu: Responses to nutrient enrichment and meteorological factors. *Scientific Reports*, 2017, **7**: 40326. DOI: 10.1038/srep40326.
- [8] Li W, Qin BQ, Zhu GW. Forecasting short-term cyanobacterial blooms in Lake Taihu, China, using a coupled hydrodynamic-algal biomass model. *Ecohydrology*, 2014, **7**(2): 794-802. DOI: 10.1002/eco.1402.
- [9] Ndong M, Bird D, Nguyen-Quang T et al. Estimating the risk of cyanobacterial occurrence using an index integrating meteor-

- ological factors: Application to drinking water production. *Water Research*, 2014, **56**: 98-108. DOI: 10.1016/j.watres.2014.02.023.
- [10] Liu Y, Wang Z, Guo HC et al. Modelling the effect of weather conditions on cyanobacterial bloom outbreaks in Lake Dianchi: A rough decision-adjusted logistic regression model. *Environmental Modeling & Assessment*, 2013, **18**(2): 199-207. DOI: 10.1007/s10666-012-9333-3.
- [11] Qi GH, Ma XS, He SY et al. Long-term spatiotemporal variation analysis and probability prediction of algal blooms in Lake Chaohu (2009–2018) based on multi-source remote sensing data. *J Lake Sci*, 2021, **33**(2): 414-427. DOI: 10.18307/2021.0204. [祁国华, 马晓双, 何诗瑜等. 基于多源遥感数据的巢湖水华长时序时空变化(2009—2018年)分析与发生概率预测. 湖泊科学, 2021, **33**(2): 414-427.]
- [12] Chen LQ, Zhang J, Chen XL et al. Research on meteorological factors and Logistic prediction model of cyanobacterial blooms in Erhai Lake. *Journal of Central China Normal University: Natural Sciences*, 2016, **50**(4): 606-611. [陈莉琼, 张娇, 陈晓玲等. 基于气象数据的洱海蓝藻水华驱动因子及预警研究. 华中师范大学学报: 自然科学版, 2016, **50**(4): 606-611.]
- [13] Webster IT, Hutchinson PA. Effect of wind on the distribution of phytoplankton cells in lakes revisited. *Limnology and Oceanography*, 1994, **39**(2): 365-373. DOI: 10.4319/lo.1994.39.2.0365.
- [14] Xue K, Zhang YC, Duan HT et al. A remote sensing approach to estimate vertical profile classes of phytoplankton in a eutrophic lake. *Remote Sensing*, 2015, **7**(11): 14403-14427. DOI: 10.3390/rs71114403.
- [15] Li J, Zhang YC, Ma RH et al. Satellite-based estimation of column-integrated algal biomass in nonalgae bloom conditions: A case study of lake Chaohu, China. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2017, **10**(2): 450-462. DOI: 10.1109/JSTARS.2016.2601083.
- [16] Bi S, Li YM, Lyu H et al. Quantifying spatiotemporal dynamics of the column-integrated algal biomass in nonbloom conditions based on OLCI data: A case study of Lake Dianchi, China. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, **57**(10): 7447-7459. DOI: 10.1109/TGRS.2019.2913401.
- [17] Liu D, Yu SJ, Cao ZG et al. Process-oriented estimation of column-integrated algal biomass in eutrophic lakes by MODIS/Aqua. *International Journal of Applied Earth Observation and Geoinformation*, 2021, **99**: 102321. DOI: 10.1016/j.jag.2021.102321.
- [18] Zhang M, Kong FX. The process, spatial and temporal distributions and mitigation strategies of the eutrophication of Lake Chaohu (1984–2013). *J Lake Sci*, 2015, **27**(5): 791-798. DOI: 10.18307/2015.0505. [张民, 孔繁翔. 巢湖富营养化的历程、空间分布与治理策略(1984—2013年). 湖泊科学, 2015, **27**(5): 791-798.]
- [19] Hu CM, Lee ZP, Ma RH et al. Moderate Resolution Imaging Spectroradiometer (MODIS) observations of cyanobacteria blooms in Taihu Lake, China. *Journal of Geophysical Research Atmospheres*, 2010, **115** (C4): C04002. DOI: 10.1029/2009jc005511.
- [20] Wang M, Shi W. Cloud masking for ocean color data processing in the coastal regions. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, **44**(11): 3196-3105. DOI: 10.1109/TGRS.2006.876293.
- [21] Hu CM. A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, 2009, **113**(10): 2118-2129. DOI: 10.1016/j.rse.2009.05.012.
- [22] Wu SC, Chen WM. The periodic change of phytoplankton biomass in Taihu Lake. *China Environmental Science*, 2004, **24**(2): 151-154. [吴生才, 陈伟民. 太湖浮游植物生物量的周期性变化. 中国环境科学, 2004, **24**(2): 151-154.]
- [23] Zhang M, Zhang YC, Yang Z et al. Spatial and seasonal shifts in bloom-forming cyanobacteria in Lake Chaohu: Patterns and driving factors. *Phycological Research*, 2016, **64**(1): 44-55. DOI: 10.1111/pre.12112.