

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*

祁国华^{1,2}, 马晓双^{1,2**}, 何诗瑜¹, 吴鹏海^{1,2}

(1: 安徽大学资源与环境工程学院, 合肥 230601)

(2: 安徽大学湿地生态保护与修复安徽省重点实验室, 合肥 230601)

摘要: 湖泊富营养化导致的水华问题严重影响了淡水资源的利用和保护, 快速、全面、准确的监测水华信息对于湖泊水环境的治理具有十分重要的意义. 本文以巢湖为研究区域, 利用多源光学遥感影像和时空融合技术, 采用波段融合的方式将 *NDVI* 指数波段加入到遥感影像当中, 并通过监督分类解译水华信息, 以此揭示 2009—2018 年 10 年间巢湖水华的时空变化规律. 结果表明: 巢湖发生的水华以零星和局部水华为主, 2012 年巢湖水体首次出现区域蓝藻水华; 巢湖水华的季变化性强, 夏、冬半年水华变化差异大, 其中 2018 年季节间的差异最为显著; 巢湖水华后五年发生频率明显高于前五年, 西半湖水华较东半湖严重, 且巢湖西北部是水华的高发区域, 2011 年水华开始向沿岸蔓延, 2014 年水华首次出现在西南湖区, 2016 年巢湖水华高发区域新增巢湖东部和中部地区. 另外, 本文根据巢湖的相关气象数据, 构建了 Logistic 水华气象风险概率预测模型, 模型平均预测准确率高达 87.52%. 探究巢湖水华长时序的时空变化规律有助于从宏观上把握其动态趋势, 为后续湖区治理以及周边生态环境建设提供理论依据, 水华气象风险概率预测模型可为巢湖水华的预警和防控提供决策支持.

关键词: 巢湖; 水华; 遥感监测; 监督分类; 时空变化; 气象数据; Logistic 模型

Long-term spatiotemporal variation analysis and probability prediction of algal blooms in Lake Chaohu (2009–2018) based on multi-source remote sensing data *

Qi Guohua^{1,2}, Ma Xiaoshuang^{1,2**}, He Shiyu¹ & Wu Penghai^{1,2}

(1: School of Resources and Environmental Engineering, Anhui University, Hefei 230601, P.R.China)

(2: Anhui Province Key Laboratory of Wetland Ecosystem Protection and Restoration, Anhui University, Hefei 230601, P.R. China)

Abstract: The algal blooms problem caused by eutrophication of lakes has seriously affected the utilization and protection of freshwater resources, hence the rapid, comprehensive and accurate monitoring of algal blooms information is of great significance for the governance of water environment. Taking Lake Chaohu as the research area, this article uses multi-source optical remote sensing images and spatio-temporal fusion technology to reveal the spatio-temporal change trends of algal blooms from the year of 2009 to 2018, by employing band fusion method and supervised classification-based blooms extraction method. The results show that: The algal blooms in Lake Chaohu are mainly sporadic and localized. In 2012, regional cyanobacteria blooms appeared for the first time. Algal blooms in Lake Chaohu has a strong seasonal variation, and there is a big difference between summer and winter which is most significant in 2018; in the last five years, the occurrence frequency of algal blooms was significantly higher than that of the previous five years. The west half lake blooms were more serious than the east half lake blooms, and the northwest part of Lake Chaohu had a high incidence of algal blooms. In 2011, the algal blooms began to spread to the coast; In 2014, the algal blooms first appeared in the southwest lake area; In 2016, the eastern and central parts of Lake Chaohu became the high-occurrence area. Besides, based on the relevant meteorological data of Lake Chaohu, this paper constructs a Logistic forecast model of algal meteorological risk.

* 2020-05-25 收稿; 2020-06-30 收修改稿.

国家自然科学基金项目(41701390)和安徽省科技重大专项(201903a07020014)联合资助.

** 通信作者; E-mail: mxs88@163.com.

logical risk probability with an average prediction accuracy rate of 87.52%. The study of the spatio-temporal variation of cyanobacteria blooms in Lake Chaohu can help us to grasp its dynamic trend from a macro perspective, providing a theoretical basis for subsequent lake area governance and ecological environment construction. Besides, the algal blooms meteorological risk probability prediction model can provide decision support for early warning and prevention of Lake Chaohu's algal blooms.

Keywords: Lake Chaohu; algal blooms; remote sensing monitoring; supervised classification; spatiotemporal changes; meteorological data; Logistic model

淡水资源是我国社会经济发展最重要的资源之一,它的利用和保护关系到国家发展和社会民生.近年来,随着人类活动的增加和经济的快速发展,淡水湖泊的富营养化已经成为一个严重的生态、环境、经济和社会问题,而蓝藻水华则是富营养化湖泊中最常见的环境问题.据报道,我国有不少大型淡水湖泊,如巢湖^[1]、太湖^[2]、滇池^[3]、洱海^[4]等都出现过大规模的水华现象.水华的频繁发生,严重威胁了湖泊周边居民的生产生活用水安全,并且对湖区的生物多样性甚至整个生态系统造成了严重的影响^[5].因此,快速、全面、准确的掌握水华发生的规律,探究水华发生的驱动因素,对后期水华的监测、治理以及预防具有十分重要的意义.

近年来卫星遥感技术的发展和应用愈发成熟,使得大范围、长时序的水华监测得以实现,与常规监测手段相比具有无可比拟的优势,国内外众多学者利用卫星影像在蓝藻水华监测研究中取得了明显的成效.例如李晟铭等利用 Landsat 影像探究了巢湖水华的长时序时空变化特征^[6];唐晓先等^[7]通过 MODIS 数据揭示了 2000—2015 年间巢湖水华的时空变化规律. Landsat 卫星发射时间早、空间分辨率高,虽然能满足长时序研究的要求,但重返周期长,实际研究中数据密度较低,由于水华高动态性的变化特点,数据密度对水华动态分析的影响较大;MODIS 数据时间分辨率较高,每天都可获取到影像,但其空间分辨率较低,最高仅有 250 m,在一定程度上会影响水华提取的精度,并且不应用监督分类的方式解译水华.大部分长时序的研究都受限于单一数据源的缺点,无法兼备高时空数据的优点.因此,要想提高水华提取的精度,更加准确客观地分析水华长时序时空变化趋势,必须获取高时间和高空间分辨率的遥感影像.

如何利用技术手段提高水华解译的精度是水华研究的一个热点课题. Rouse 等^[8]提出了归一化植被指数(normalized difference vegetation index, *NDVI*),用来区分水体和水中蓝藻;Hu^[9]建立了一种浮游藻类指数(floating algae index, *FAI*)来识别水华,有效地减小了大气对水华提取的影响.指数阈值提取法对水华有较好的辨识性,但是对于不同数据源甚至同一数据源不同时相的遥感影像,阈值都会发生变化,因此阈值的选定存在极大的主观性和不确定性;再者,阈值提取法避免不了会将“异物同谱”的物质进行误分,容易导致水华面积偏大.不少学者,如潘琛等^[10]利用监督分类的方法对内陆湖泊的水华信息进行有效提取;Zhang 等^[11]利用监督分类的方法对海洋中的赤潮进行了识别.监督分类的优势在于能够有效利用研究者的经验,摆脱阈值选择带来的困扰,并且在样本选取时还能消除浑浊水体的背景干扰和“异物同谱”物质的影响,但是易于受影像质量问题(如云、雾等)的影响,在样本选择过程中往往会出现误选,导致难以取得较高的解译精度.因此本质上来讲,提高样本选择的准确性以及在监督分类中考虑更多的影像信息特征是提高水华提取精度的有效途径.

在利用遥感影像监测到水华信息的基础上,众多学者针对水华时空变化开展了相关的研究.例如马荣华等对太湖蓝藻水华最初在夏季出现时间的变化、覆盖面积、空间分布特点等做了研究^[12],朱利等^[13]从发生频率、起始日期和持续时间三个角度分析探讨了巢湖水华的动态变化规律.以上这些研究虽然都尝试从不同角度探寻水华的时空变化规律,但受限于遥感数据量较少、分辨率较低以及水华提取方法精度不高等的的影响,难以较准确获取长时间尺度、高时空分辨率的水华信息.此外,还有学者通过分析湖区气象条件变化对水华发生的影响,利用 Logistic 回归建立气象风险概率预测模型,将气象因子与遥感监测结果建立直接的关系^[14].气象因素虽不是水华发生的主导因素,但却是诱导因素,不同湖区气象条件差异显著,如何选取气象因子精准建立预测模型以及判定水华气象风险的主导气象因素值得探寻和深入研究.

本文利用 2009—2018 年 10 年间巢湖区域的多源遥感影像以及时空融合技术得到研究区长时序高时空分辨率的遥感数据,采用波段融合的方式将 *NDVI* 指数波段加入到遥感影像当中并通过监督分类提取水华信息,从水华发生程度、四季变化特征、发生频率、空间分布及其原因等角度揭示了巢湖水华的长时序时空

变化特征;另外,通过分析水华发生期的气温、日照、降水、风速等气象因素,构建 Logistic 巢湖水华气象风险预测模型,为巢湖水华的监测、治理和防控提供了重要的理论依据和技术支撑.

1 研究区域与数据

1.1 研究区域概况

巢湖(31°25′~31°42′N,117°17′~117°50′E;图 1)位于安徽省中部合肥市境内,被巢湖市、包河区、肥东县、肥西县、庐江县一市一区三县环抱其中,为合肥市最大的内湖,也是我国的第五大淡水湖. 巢湖东西长 55 km,南北宽 21 km,湖岸线 170 多 km,平均水深 3 m,最大水深 6.78 m,面积约 780 km²,以忠庙—姥山—新河口为界分为东西半湖. 湖水主要靠地表径流补给,水域辽阔,四周河沟渠道 360 多条,纵横交错,曲折回环,其中较大的河流有南淝河、派河、杭埠河、白石天河、兆河、柘皋河等,并通过裕溪河与长江连通^[15].

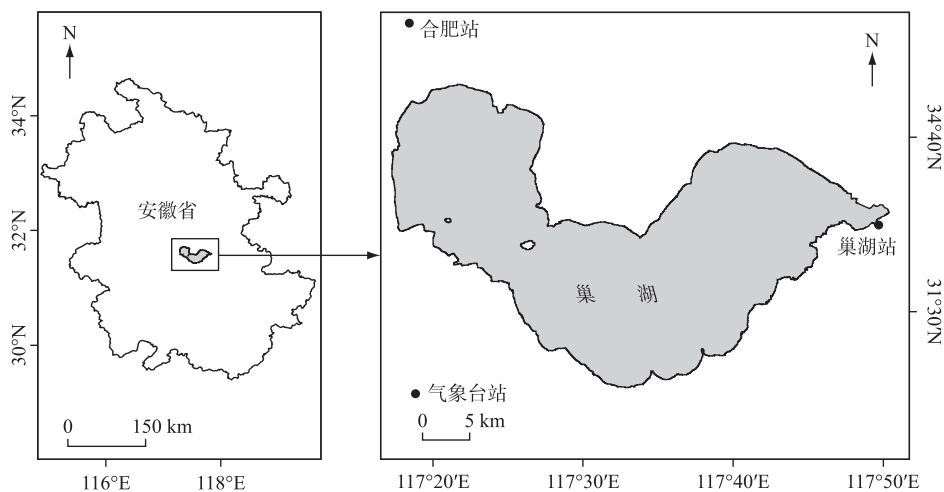


图 1 巢湖地理位置及气象站示意

Fig.1 Location in Lake Chaohu and diagram of meteorological stations

1.2 遥感数据与气象数据

本文实验选取 2009—2018 年 10 年间巢湖区域的高分辨率多源光学遥感数据以及利用时空融合技术得到的时空融合数据,数据基本均分在各个月份,数据源包括高分一号、Landsat、Sentinel-2、MODIS 和融合数据,需要说明的是,MODIS 数据空间分辨率较低,本文只将其作为间接数据源,参与数据的时空融合,各相关数据源使用情况如表 1 所示. 同期气象数据来源于国家气象信息中心(<http://data.cma.cn/>),站点为合肥站(31°47′N,117°18′E)和巢湖站(31°35′N,117°5′E),获取包括气温、日照、降水量、风速、风向等逐日气象数据,气象数据统计采用两个站点气象数值的日平均值.

表 1 卫星遥感数据使用情况

Tab.1 Data usage of satellite remote sensing

数据源	高分一号	Landsat 5、7、8	Sentinel-2	MODIS	融合数据
分辨率/m	8、16	30	10	250	30
数量	139	125	42	331	242
重返周期/d	4	16	10	1	—*
时间范围	2013—2017 年	2009—2018 年	2016—2018 年	2009—2018 年	2009—2018 年
数据来源	遥感集市	RADI	欧空局(ESA)	NASA	时空融合

* “—”表示不具有该项参数.

2 研究方法

2.1 影像预处理与时空融合

受各种因素的影响,遥感影像往往不能直接使用,首先须对影像进行辐射定标、大气校正、正射校正、几何校正等预处理,对部分数据源影像因传感器故障、单幅影像不足以覆盖研究区域等问题,还须进行去条带、图像镶嵌等相关步骤.为获取长时序高时空分辨率的遥感数据以及解决部分时间段数据缺失和分布不均匀问题,本文采用 ESTARFM(enhanced spatial and temporal adaptive reflectance fusion model)增强时空自适应反射率融合模型^[16]对 MODIS 数据和 Landsat 数据进行时空融合.

ESTARFM 模型的融合理论是对已知影像的像素值取相同的权值,然后进行加权平均得到融合影像的像素值^[17].该模型在原有时空融合模型的基础上进行了改进,对赋权方法进行了调整,通过引入一个转换系数,提高了地表情况较为复杂以及空间性质差异较大区域的预测精度,同时在融合过程中不仅充分考虑像元之间的时间和空间差异性,而且考虑光谱差异性^[18],是目前应用最为广泛并且融合精度较高的时空融合模型之一^[19].部分研究者已将该模型成功地应用到温度、垫面监测与热环境分析、土地利用类型变化等动态性变化较高的地物监测中,充分表明该模型对高动态变化地物类型的影像融合同样适用^[20-21].在融合过程中,该模型首先利用权重函数进行卷积运算得到中心像元值,如公式(1)所示:

$$L(x_{w/2}, y_{w/2}, t_p, B) = L(x_{w/2}, y_{w/2}, t_0, B) + \sum_{i=1}^N W_i \cdot V_i \cdot (M(x_i, y_i, t_p, B) - M(x_i, y_i, t_0, B)) \quad (1)$$

式中, L 代表 Landsat 影像, M 代表 MODIS 影像, w 为计算窗口的大小, t_0 、 t_p 分别表示影像获取的两个时刻, B 表示影像波段, N 是中心预测像元的相似像元数目, W_i 是第 i 个相似像元的权重, V_i 是第 i 个相似像元的转换系数, (x_i, y_i) 是第 i 个相似像元的位置.

根据时相 1 (m 时刻)的 Landsat 和 MODIS 与时相 2 (t_p 时刻)的 MODIS 来预测 t_p 时刻 Landsat 影像的反射率,记为 $L_m(x_{w/2}, y_{w/2}, t_p, B)$;然后利用时相 3 (n 时刻)的 Landsat 和 MODIS 与时相 2 (t_p 时刻)的 MODIS 来预测 t_p 时刻 Landsat 影像的反射率,记为 $L_n(x_{w/2}, y_{w/2}, t_p, B)$.通过 2 个时相预测结果的加权组合可以获得更加准确的 t_p 时刻 Landsat 影像的反射率^[22],如公式(2)所示:

$$L(x_{w/2}, y_{w/2}, t_p, B) = T_m \cdot L_m(x_{w/2}, y_{w/2}, t_p, B) + T_n \cdot L_n(x_{w/2}, y_{w/2}, t_p, B) \quad (2)$$

式中, $L(x_{w/2}, y_{w/2}, t_p, B)$ 是 t_p 时刻中心像素点的融合结果, T_m 、 T_n 分别是两个时刻预测结果影像的权重系数.具体的时空融合示意图如图 2 所示.

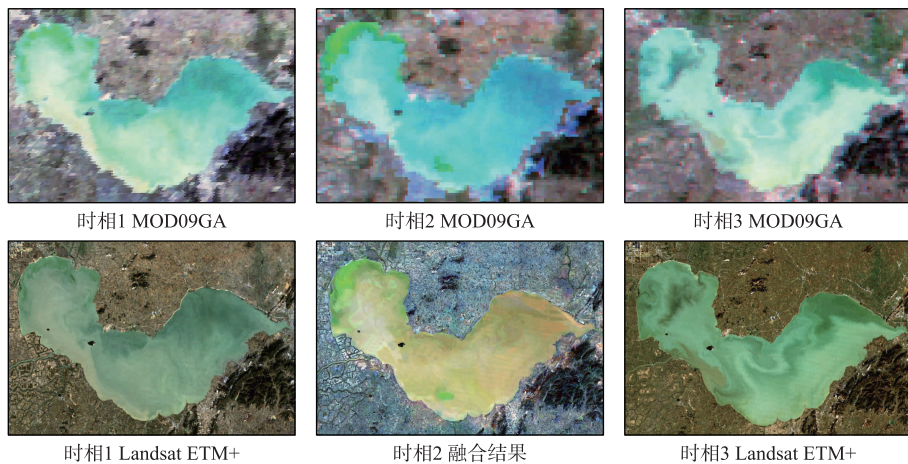


图 2 时空融合示意图

Fig.2 The schematic diagram of spatial and temporal data fusion

2.2 水华解译方法

支持向量机(support vector machine, SVM)是一种有监督的学习模型,通常用来进行模式识别、地物分类和回归分析^[23],特别是在地物分类中有着广泛的应用和较高的分类精度^[24]. SVM 解译水华的理想目标是将水华像元根据其在不同波段的光谱特征准确地进行分类,但是由于水华分布的复杂性、无序性和不规则性,仅仅依靠光谱特征难以获得较为理想的分类结果,因此,在监督分类中充分利用影像上水华的多特征信息,对于提高分类精度具有很大的意义^[25-26]. 本研究通过波段融合的方式将 *NDVI* 指数波段加入预处理之后的遥感影像当中,使影像具备了原本不具有的特征信息,对难以区分的中、低浓度水华具有较好的辨识作用. 此外,通过波段融合还能够综合利用不同波段之间的特征信息^[27],提高样本的选取精度. 对于本研究而言,由于 *NDVI* 能有效描述水体的含植被量并且在水华识别中具有较为广泛的使用,故将其作为一个输入波段,得到一种结合 *NDVI* 与 SVM 的水华提取方法,以下简称 *NDVI*-SVM. 该方法解译水华的基本流程如图 3(仅展示关键过程的影像结果)所示,水华解译基本过程如下:

1) *NDVI* 提取与波段融合:将预处理后的遥感影像先进行归一化 *NDVI* 提取,然后将 *NDVI* 指数影像与预处理后的遥感影像进行波段融合,得到含有 *NDVI* 指数波段的多波段影像.

2) 研究区矢量边界裁剪:利用巢湖矢量边界对影像进行裁剪,目的是增强影像中不同地物的显色差异以及加快后期监督分类速度.

3) 样本选取及评价:在 2% 的拉伸方式下对波段融合后的影像依次进行 *NDVI* 指数波段、红波段、绿波段显色,水华在影像中呈粉红色形态与周围蓝色水体有明显区别;随后进行样本选取,样本共分为两类,即水华和无水华水体样本,样本的选择应遵循代表性、统计性和准确性等基本原则. 为保证样本选择的精度,每类样本的选取数量保证在 50 个以上,同时对所选样本进行评价,通过计算 Jeffries-Matusita 距离和转换分离度,其中大于 1.9 说明样本之间可分离性好,属于合格样本;小于 1.8,需要重新选择样本.

4) SVM 监督分类:采用支持向量机的分类方式对样本进行分类,并将分类结果与预处理结果影像进行叠加,目视判断分类精度.

5) 掩膜:对分类结果进行掩膜获取水华解译信息,并根据像元大小和数量计算水华覆盖面积.

2.3 Logistic 模型

Logistic 回归模型是一种非线性回归概率模型,通过转化问题来分析被预测变量的条件概率同预测变量之间的关系,在二分类因变量问题中应用较为广泛^[28]. 本研究将水华发生与否(分别取值为 1 和 0)作为因变量,将各种气象因素变量值作为自变量,建立各气象因素与水华发生与否的二元概率回归模型.

Logistic 回归模型表达式如下:

$$\rho = \frac{\exp(\beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \cdots + \beta_k \cdot X_k)}{1 + \exp(\beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \cdots + \beta_k \cdot X_k)} \quad (3)$$

或

$$\ln\left(\frac{\rho}{1-\rho}\right) = \beta_0 + \beta_1 \cdot X_1 + \beta_2 \cdot X_2 + \cdots + \beta_k \cdot X_k \quad (4)$$

式中, $\rho = P(y = 1 | X_1, X_2, \cdots, X_k)$ 为事件发生概率, X_1, X_2, \cdots, X_k 为自变量因子值, $\beta_0, \beta_1, \beta_2, \cdots, \beta_k$ 为逻辑回归系数,模型通过 SPSS 软件构建和拟合度检验.

根据中国环境监测总站制定的“太湖、巢湖水华特征临时判别方法”^[29],可对巢湖水华特征进行判别分类,将湖区藻类覆盖面积 $\leq 5\%$ 的情况界定为无水华,此时模型因变量取值为 0;反之,界定为有水华,因变量取值为 1. 本文根据获取的水华信息以及利用上述判别方法得到所有对应日期的因变量值,同时输入筛选的对应日期气象因子的自变量值,便可实现模型的拟合构建.

3 结果与分析

3.1 *NDVI*-SVM 方法精度评价

目前绝大部分利用遥感影像提取水华的研究都无法进行精度评价,往往只能借助目视解译来判断提取结果是否可靠,本文给出其中 2 种数据源影像水华解译前后的结果,如图 4 所示. 此外,为验证 *NDVI*-SVM 方

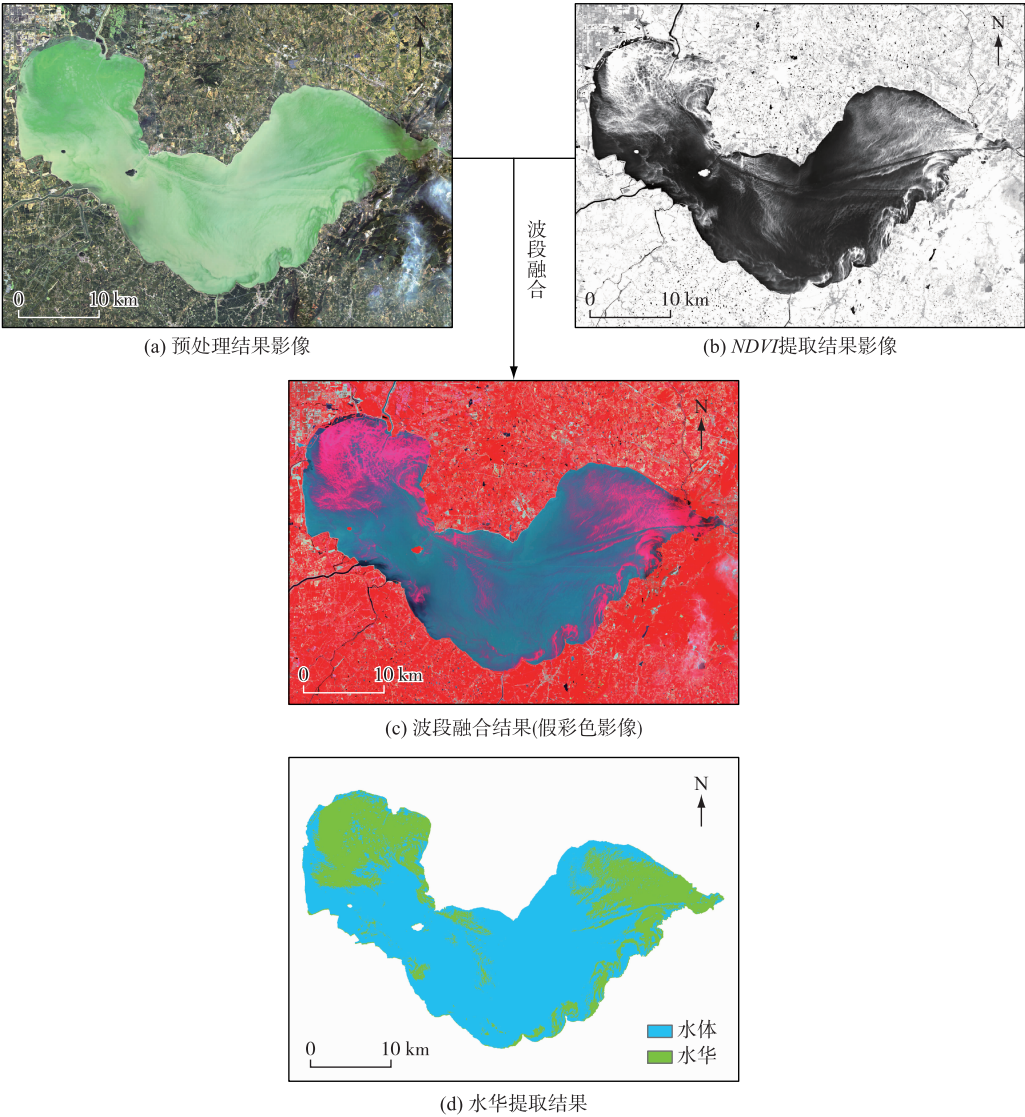


图3 水华提取流程

Fig.3 Algal blooms extraction process

法的可靠性,本文提出一种新的评价方式,该评价方式的基本思想是:对于一种可靠的水华解译方法而言,其在时相极接近的两个不同传感器所获取影像上的水华解译结果应当极为接近,通过在同一影像拉伸(2%拉伸)方式下分别采用 *NDVI* 阈值提取法、传统 *SVM* 监督分类法和 *NDVI-SVM* 方法对影像进行水华提取,如表 2 所示,可以发现两对时相中 *NDVI-SVM* 方法提取面积极为接近(仅分别相差 0.2712 和 0.646 km²). 对于水体总面积约 780 km² 的巢湖而言,上述提取面积的差值微乎其微,而且之所以存在差值主要有以下三个原因:其一,不同分辨率卫星对水华的识别会有略微差别;其二,监督分类本身存在主观意识;其三,卫星影像过境时间虽相差无几,但还是存在时间差,水华可能会在几分钟内发生略微的变化.

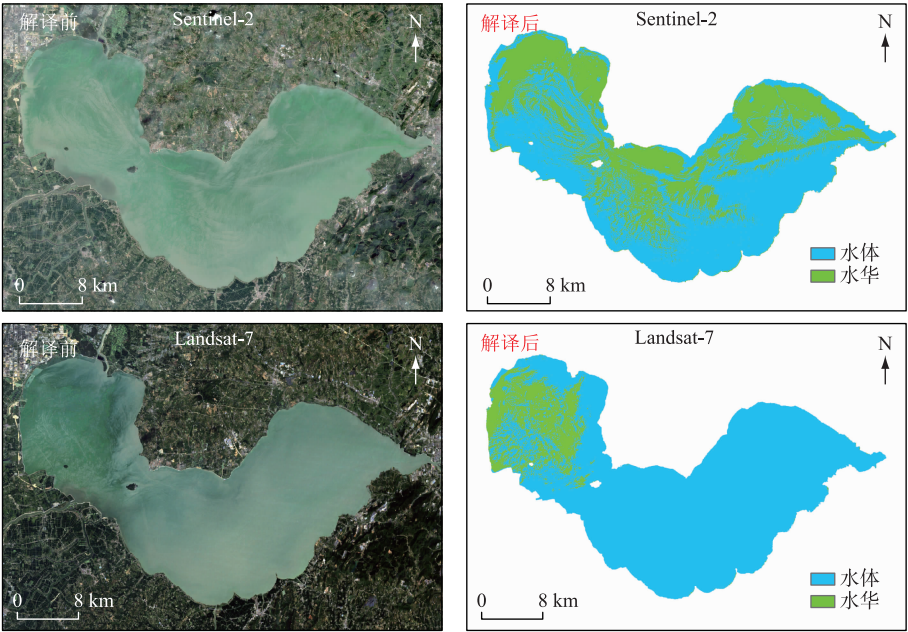


图 4 *NDVI-SVM* 方法水华解译结果
Fig.4 *NDVI-SVM* method to interpret algal blooms results

表 2 水华提取方法误差分析
Tab.2 Error analysis of extracted method of algal blooms

水华提取方法	日期	数据源	卫星过境时间 (北京时间)	水华面积/ km ²	面积差值/ km ²	相对误差/ %
<i>NDVI</i> 阈值法	2018 年 1 月 12 日	Landsat-7	10:45:34	10.3896	1.8013	18.98
		Sentinel-2	10:50:49	8.5883		
	2018 年 9 月 9 日	Landsat-7	10:41:36	117.4158	7.3074	6.04
		Sentinel-2	10:45:39	124.7232		
SVM	2018 年 1 月 12 日	Landsat-7	10:45:34	5.7474	0.9913	18.88
		Sentinel-2	10:50:49	4.7561		
	2018 年 9 月 9 日	Landsat-7	10:41:36	75.1257	6.7763	8.63
		Sentinel-2	10:45:39	81.902		
<i>NDVI-SVM</i>	2018 年 1 月 12 日	Landsat-7	10:45:34	4.9599	0.2712	5.32
		Sentinel-2	10:50:49	5.2311		
	2018 年 9 月 9 日	Landsat-7	10:41:36	101.5281	0.646	0.63
		Sentinel-2	10:45:39	102.1741		

为进一步定量分析 *NDVI-SVM* 方法的可靠性,本文提出一种定量评价指标,其计算表达式如下:

$$\sigma = \frac{|S_1 - S_2|}{(S_1 + S_2)/2} \tag{5}$$

式中, σ 代表相对误差值, S_1 代表 Landsat-7 影像水华提取面积, S_2 代表 Sentinel-2 影像水华提取面积. 误差值越小,表明相近时间不同传感器水华提取结果越为接近,水华解译效果越好,提取精度越高.

根据误差评价结果可以看出,*NDVI-SVM* 的误差值较小,而 *NDVI* 阈值提取法和传统 SVM 分类法相对较大,另外发现 *NDVI-SVM* 的水华提取面积介于 *NDVI* 阈值提取法(提取面积一般会偏大)和传统 SVM 监督分

类法(提取面积一般会偏小)之间,这充分验证了 *NDVI-SVM* 方法的可靠性,此外,该方法对多源高分辨率遥感影像具有普遍的适用性,同时有着较快的监督分类速度. 因此,整体而言,*NDVI-SVM* 方法是一种较为理想并实质有效的蓝藻水华解译方法.

3.2 巢湖水华时空变化规律分析

3.2.1 水华发生程度年际变化 根据水华面积占湖泊总面积的比例(*P*),将巢湖水华发生程度划分为 4 级. 当 $P \leq 5\%$ 时,无蓝藻水华;当 $5\% < P \leq 10\%$ 时,有零星蓝藻水华;当 $10\% < P \leq 40\%$ 时,有局部蓝藻水华;当 $40\% < P \leq 60\%$ 时,有区域蓝藻水华;当 $P > 60\%$ 时,有全湖蓝藻水华^[30]. 本研究对 2009—2018 年间的巢湖水华发生情况进行了统计(表 3).

表 3 巢湖蓝藻水华发生次数
Tab.3 Occurrence times of cyanobacteria blooms in Lake Chaohu

时间	影像 总数	零星蓝藻水华 发生次数	局部蓝藻水华 发生次数	区域蓝藻水华 发生次数	全湖蓝藻水华 发生次数
2009 年	49	18	10	0	0
2010 年	55	10	18	0	0
2011 年	53	12	22	0	0
2012 年	49	13	15	2	0
2013 年	54	14	16	0	0
2014 年	49	10	18	1	0
2015 年	62	8	32	3	0
2016 年	59	14	19	1	0
2017 年	61	18	14	0	0
2018 年	57	14	21	1	0

如图 5 所示,2009—2018 年期间,巢湖发生的蓝藻水华以零星和局部水华为主,区域水华占比较低,全湖水华维持在零状态,2012 年首次出现区域水华,除 2009 和 2017 年外,其它年份局部水华发生次数所占的比例都高于零星水华,10 年间局部水华发生次数的比例变化趋势最为显著.

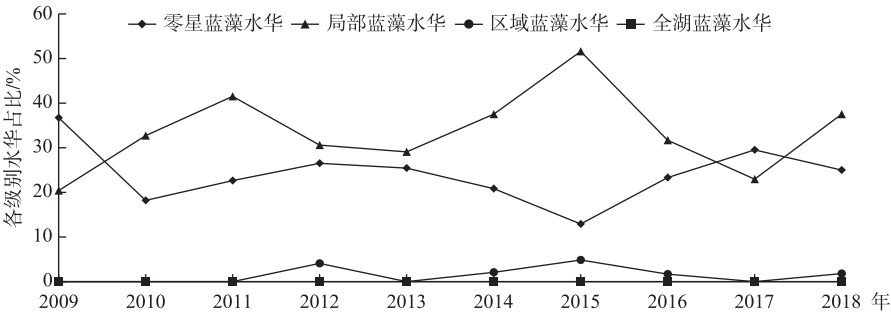


图 5 巢湖各等级蓝藻水华发生次数占总次数的比例

Fig.5 The percentages of various degrees number of algal blooms to total number occurred in Lake Chaohu

3.2.2 水华季节变化趋势 根据气象学划分的季节(3、4、5 月为春季,6、7、8 月为夏季,9、10、11 月为秋季,12 和次年 1、2 月为冬季)对巢湖水华进行四季变化特征分析(增加 2019 年 1、2 月份的水华信息作为 2018 年的冬季). 如图 6 所示,水华在夏、秋季最为严重,而冬、春季处于相对低迷状态;10 年间,冬春季的水华面积基本保持平稳状态,而夏秋季整体呈增长趋势;另外,考虑到巢湖所处的安徽省中部春季和秋季较短,本文将一年分为夏半年(5—10 月)和冬半年(其余月份),夏半年水华普遍比冬半年严重,并且不同年份夏、冬半年水华面积差异较大,其中 2018 年差异最大,水华覆盖面积差值高达 95.687 km².

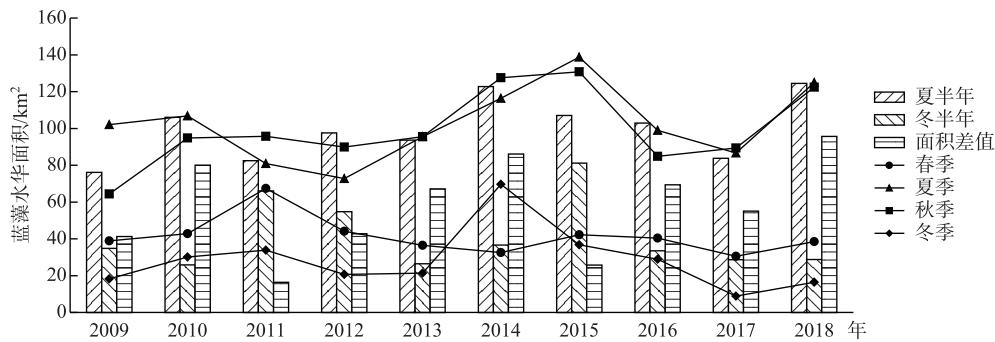


图 6 巢湖蓝藻水华季节变化

Fig.6 Seasonal variation of algal blooms in Lake Chaohu

3.2.3 水华空间变化特征 从各年的巢湖水华频率分布图以及 10 年间整体水华频率分布图可以发现(图 7),西巢湖水华发生频率明显高于东巢湖,且西巢湖的高发区域主要在西北部,其主要原因之一是巢湖西北部是合肥市市区排污河(南淝河、十五里河、塘西河)的入湖口;后五年水华发生频率明显高于前五年,东中部湖区发生频率较之前有所增加。尽管近十几年来对巢湖治理的措施和政策相继出台,投资也不断加大,但湖泊富营养化仍然保持在较高水平。湖区水华发生频率的变化也较为复杂,2011 年水华开始向沿岸蔓延;2014 年水华首次出现在西南湖区;2016 年巢湖水华高发区域新增巢湖东部和中部地区,这与当年巢湖泄洪换水有关,换水使得湖区营养化水平明显下降,蓝藻优势种群发生改变,冬季鱼腥藻频繁出现在中东湖区^[31];2017、2018 年相比前几年,高频暴发区域有所减少,表明近几年巢湖的治理初见成效,并且随着合肥滨湖新区的逐步建设与发展,巢湖西北岸乃至整个巢湖区域的水华都会得到有效治理。

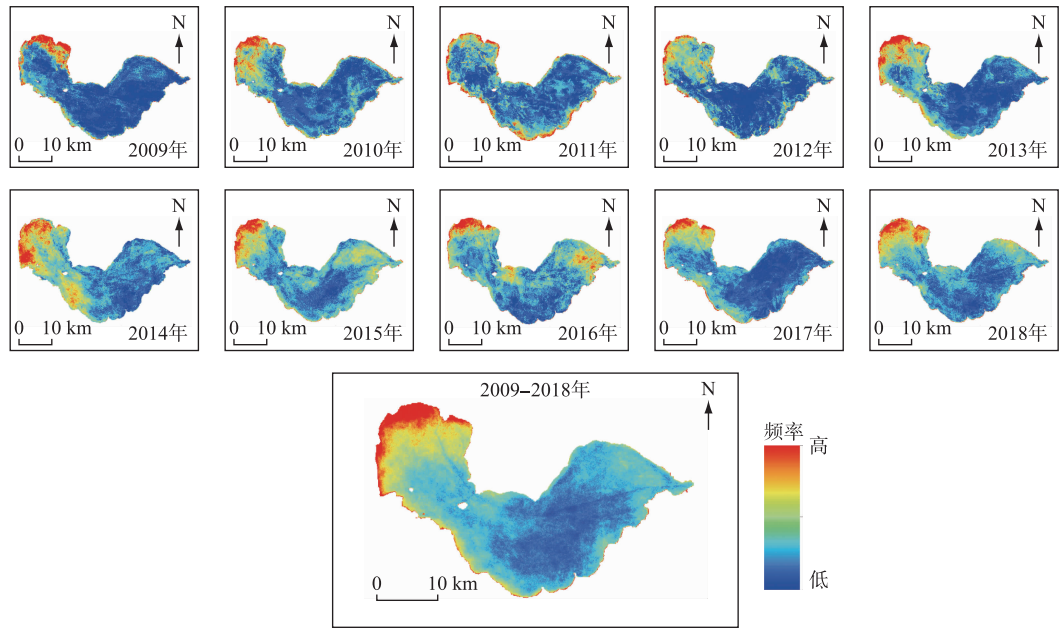


图 7 巢湖蓝藻水华年度发生频率

Fig.7 Annually occurrence frequency of algal blooms in Lake Chaohu

3.2.4 水华空间分布原因 巢湖水华发生的空间分布原因除了与排污河入湖口有关外,也受限于风向这一气

象因素的影响. 如图 8 所示,为了研究这一问题,以 2018 年为例,将夏季风(5—10 月)风向、冬季风(其余月份)风向、全年风向的风向频率进行统计制成风向玫瑰图;与此同时,给出相应风向日期的水华频率分布图,可以发现,以主导风向为东风和东南风的夏季风,水华在巢湖西北岸的发生频率较高;以主导风向为西北风的冬季风,水华在巢湖南岸、东南岸的发生频率较高;从主导风向为东风和东南风的全年风玫瑰图来看,西北岸发生频率较高. 由此可见,巢湖水华发生的空间分布与风向有直接的关系,这主要是因为风向会使湖泊表层水体向下风向流动,从而带动蓝藻聚集形成大规模水华.

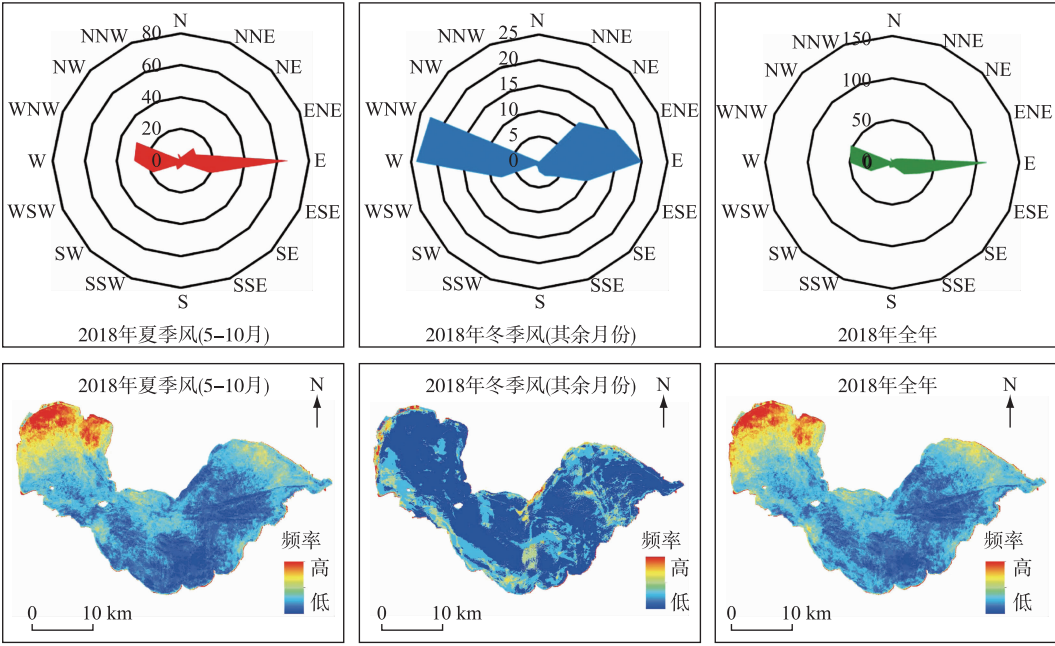


图 8 主导风向与水华高发区域对比

Fig.8 Comparison of dominant wind direction and algal blooms high frequency outbreak area

3.3 水华发生期气象条件分析

气象数据采用日平均值(8:00—20:00),通过分析卫星监测当日及前几日气象条件变化,确定模型的自变量因子,此外,从所有卫星监测日期中随机选取部分日期,统计水华发生情况与气象数据的显著相关性,来进一步验证自变量选取的合理性.

1) 气温:藻类的生长需要适宜的温度条件,统计水华发生日的日最高气温和日最低气温,发现两者变化幅度较大,但是根据水华发生日平均气温变化情况来看(图 9),最高 30.4℃,最低 6.0℃,平均 20.6℃,平均气温与水华的 Pearson 相关系数为 0.69(在 0.01 水平上双侧显著相关),由此可见,较高或较低的温度都会抑制藻类的生长. 因此,平均气温可能是诱发水华发生的重要气象条件,故选取平均气温作为模型因变量因子之一.

2) 日照:从水华发生日的日照时数变化情况来看(图 9),水华发生日日照时数普遍较长,基本都在 6.5 h 以上,平均日照时数 9.0 h,最长日照时数高达 11.0 h,日照时数与水华的 Pearson 相关系数为 0.79(在 0.01 水平上双侧显著相关),同时发现每次水华发生的前几日,日照时数普遍偏低,这是因为水华发生前普遍会有阴雨天气出现,而天气转晴之后突增的强日照可能是水华发生的诱导原因. 因此,较长的日照时数对水华的发生具有一定的促进作用.

3) 降水量:降水对水华的影响主要有两方面:其一,降水会导致温度降低和日照时数骤降,从而抑制水华的发生;其二,降水会对水体产生强大的搅拌作用,从而影响藻类上浮聚集形成水华,但是统计发现水华发生前均有明显降水或连续性阴雨天气,天气转晴后,气温的升高和日照时数的突增,则会诱发水华的发生.

根据水华发生日的前五日日均降水量变化情况(图 10)来看,其与水华发生有明显的关系,并且前五日日均降水量与水华的 Pearson 相关系数为 0.69(在 0.01 水平上双侧显著相关)。因此,将前五日日均降水量作为构建模型的气象指标之一。

4) 风速:统计水华发生日的最大风速、最小风速和平均风速,发现最大风速和最小风速的风速变化范围较大,不能很好地拟合与水华发生的关系,但是根据水华发生日的平均风速变化情况(图 10)来看,平均风速都在 3.2 m/s 以内,由此看来,较低的风速对水体的搅拌作用小,有利于藻类聚集上浮。

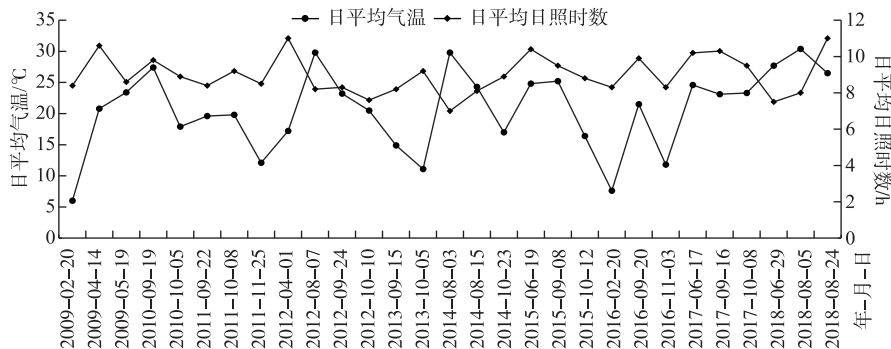


图 9 巢湖水华发生期平均气温和平均日照时数变化

Fig.9 Changes in the average temperature and sunshine hours during the algal blooms period of Lake Chaohu

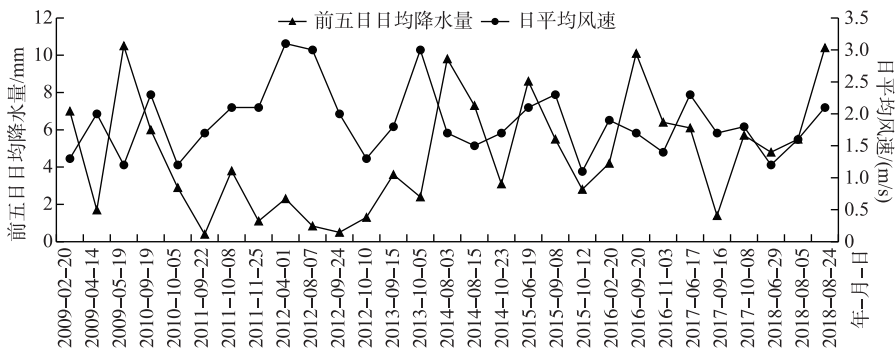


图 10 巢湖水华发生期前五日日均降水量和平均风速变化

Fig.10 Changes of daily precipitation and average wind speed for the first five days of algal blooms in Lake Chaohu

需要说明的是,本研究所使用的光学遥感数据只能在晴空下监测巢湖蓝藻水华发生的情况,有云或阴雨天时则监测不到,这可能会导致气象数据分析稍有偏差,但是目前有学者已经研究利用雷达遥感数据对太湖蓝藻水华进行监测,取得了明显的成效^[32]。后续将尝试利用雷达遥感数据监测巢湖蓝藻水华,从而能进一步提高巢湖水华气象风险的预测准确率。

3.4 Logistic 模型检验与结果

为检验模型的精度和稳定性,采用交叉验证法,将全部数据随机均分为 5 组,其中 4 组作为训练样本构建模型,1 组作为验证组与其对应日期的模型预测结果进行对比。实验发现,Logistic 模型对于 5 组数据的预测准确率分别为 88.0%、85.3%、92.0%、85.5%、86.8%,平均预测准确率 87.52%。模型的卡方值为 54.291,自由度为 4,显著性为 0.001,小于 0.05 的显著水平,表明所构建模型的显著性较高。在模型拟合优度 Hosmer & Lemeshow 检验中^[33],卡方统计值为 3.148,自由度为 7,显著性为 0.871 > 0.05,说明模型能较准确地拟合整

体。以上数据均表明本研究所选气象因子较为恰当,也说明模型在巢湖水华气象风险预测中的适用性。

在最终的模型变量检验中,采用 Wald 统计量检验自变量影响的显著程度^[34]。所有自变量因子均满足较强的统计显著性,其中日照时数的显著性为 0.38,明显高于其他气象因子,表明日照时数是影响水华发生的主导气象因素;另外,所有自变量因子的 S.E.值均大于 0.05,充分说明各气象因子与水华发生的相关性具有统计学意义^[35]。

根据最高预测准确率数据得到的最佳回归方程如下:

$$\ln\left(\frac{\rho}{1-\rho}\right) = 3.663 + 0.176 X_1 - 0.279 X_2 - 2.762 X_3 + 3.05 X_4 \quad (6)$$

式中, ρ 表示预测概率, X_1 表示日平均气温, X_2 表示日平均日照时数, X_3 表示日平均风速, X_4 表示前五日均降水量。

模型预测结果表明,巢湖水华一般在阴雨天之后日照时数较长的情况下发生概率较大,另外,较低的风速和适宜的温度对水华发生概率的影响也不容忽视。进一步分析预测结果可知,错误预测的情况一般出现在水华发生衰退期,因为水华发生期蓝藻在湖面大规模集聚,即使气象条件发生骤变,大规模的水华不可能短期内从湖面消失,水华仍旧持续一定的时间,所以水华发生后期不仅与气象条件有关,还与水华发生前期的规模和程度有很大的关系。

在富营养化湖泊中,蓝藻的暴发是因为水体中氮磷等营养物质偏高,当营养盐浓度稳定在恒定范围内,气象、水文等因素就成了诱导水华发生的关键条件。基于气象数据构建的 Logistic 水华气象风险概率预测模型,重点分析气象因素对水华暴发的影响,实现了通过气象因素对水华发生概率的预测,并且将该模型成功地运用到了巢湖水华气象风险的预测中,为后期巢湖甚至其它富营养化湖泊的治理和防控提供了技术理论和参考价值。

4 结论

卫星遥感技术在蓝藻水华监测中的广泛应用,使水华的治理与防控变得高效化和便捷化。本文基于高时空分辨率的遥感数据,利用 NDVI-SVM 水华提取方法,从水华发生程度、四季变化特征、发生频率、空间分布及其原因等多角度揭示了巢湖水华的长时序时空变化规律;同时通过同期气象数据,构建了 Logistic 巢湖水华气象风险概率预测模型。研究的主要创新点及结论如下:

1) 高时空分辨率的遥感数据:本研究通过多源卫星以及时空融合技术获取了巢湖区域长时序高时空分辨率的遥感数据,解决了因恶劣天气导致的部分时间段数据量不足和分布不均匀问题,为基于遥感数据的水华长时序研究提供了可靠的数据支撑。

2) 较高精度的水华提取方法:NDVI-SVM 水华提取方法避免了阈值选择的主观性和不确定性,减小了“异物同谱”物质的影响,同时在 SVM 监督分类时综合利用了多波段影像的多特征信息,提高了水华解译的精度,特别是中、低浓度水华的提取精度。

3) 多角度的时空分析:巢湖发生的蓝藻水华以零星和局部水华为主,全湖水华维持在零状态,10 年间局部水华发生次数的比例变化趋势最为显著;巢湖水华季节变化性强,夏冬半年变化差异大;后五年水华发生频率明显高于前五年,东中部湖区水华发生频率较之前有所增加;巢湖水华的空间分布与主导风向及排污河入湖口位置有很大的关系,整体来看,西巢湖水华较东巢湖严重,巢湖西北部是水华的高发区域。

4) 稳定的水华气象风险预测模型:本研究通过水华遥感监测信息和气象数据构建了巢湖水华气象风险预测模型,平均预测准确率高达 87.52%。预测结果表明,在阴雨天之后,日照时数较长的情况下,巢湖水华发生的概率最大,除此之外,较低的风速和适宜的温度也是影响水华发生概率的重要条件,其中日照时数是影响水华发生概率的主导气象因素。

需要说明的是,基于遥感影像的水华提取方法并不唯一,水华提取方法还有待改进,精度还有待提高,而且水华发生的内在机理复杂,未来应该探究多个因素,并考虑彼此之间的联系,另外,Logistic 水华气象风险概率预测模型还有待优化。

致谢:感谢马晓双副教授在论文写作中给予的指导和帮助。

5 参考文献

- [1] Wang SG, Diao XJ, He LS. Effects of algal bloom formation, outbreak, and extinction on heavy metal fractionation in the surficial sediments of Lake Chaohu. *Environmental Science & Pollution Research International*, 2015, **22** (18) : 14269-14279.
- [2] Duan HT, Ma RH, Xu XF *et al.* Two-decade reconstruction of algal blooms in China's Lake Taihu. *Environment Science & Technology*, 2009, **43**(10) : 3522-3528.
- [3] Lu WK, Yu LX, Ou XK *et al.* Relationship between occurrence frequency of cyanobacteria bloom and meteorological factors in Lake Dianchi. *J Lake Sci*, 2017, **29**(3) : 534-545. DOI: 10.18307/2017.0302. [鲁韦坤, 余凌翔, 欧晓昆等. 滇池蓝藻水华发生频率与气象因子的关系. 湖泊科学, 2017, **29**(3) : 534-545.]
- [4] Ji NG, Wang SR, Zhang L. Characteristics of dissolved organic phosphorus inputs to freshwater lakes: A case study of Lake Erhai, southwest China. *Science of the Total Environment*, 2017, **601-602** : 1544-1555.
- [5] Liu EF, Shen J. A comparative study of metal pollution and potential eco-risk in the sediment of Lake Chaohu (China) based on total concentration and chemical speciation. *Environmental Science & Pollution Research International*, 2014, **21** (12) : 7285-95.
- [6] Li SM, Liu JP, Song KS *et al.* Analysis on spatial and temporal character of algae bloom in Lake Chaohu and its driving factors based on Landsat imagery. *Resources and Environment in the Yangtze Basin*, 2019, **28**(5) : 1205-1213. [李晟铭, 刘吉平, 宋开山等. 基于 Landsat 影像巢湖蓝藻水华暴发时空变化特征及其驱动因素分析. 长江流域资源与环境, 2019, **28**(5) : 1205-1213.]
- [7] Tang XX, Shen M, Duan HT. Temporal and spatial distribution of algal blooms in Lake Chaohu (2000-2015). *J Lake Sci*, 2017, **29**(2) : 276-284. DOI: 10.18307/2017.0203. [唐晓先, 沈明, 段洪涛. 巢湖蓝藻水华时空分布(2000—2015 年). 湖泊科学, 2017, **29**(2) : 276-284.]
- [8] Rouse JWJ, Haas RH, Schell JA *et al.* Monitoring vegetation systems in the Great Plains with erts. *Nasa Special Publication*, 1974, **351** : 309.
- [9] Hu C. A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment*, 2009, **113** (10) : 2118-2129.
- [10] Pan C, Lin Y, Chen YY. Decision tree classification of remote sensing images based on multi-feature. *Journal of Optoelectronics Laser*, 2010, **21**(5) : 731-736. [潘琛, 林怡, 陈映鹰. 基于多特征的遥感影像决策树分类. 光电子·激光, 2010, **21**(5) : 731-736.]
- [11] Zhang T, Su FZ, Yang XM *et al.* A method for detecting red tide information from MODIS data and its application in Pearl River Estuary. *Journal of Geo-Information Science*, 2009, **11**(2) : 244-249.
- [12] Ma RH, Kong FX, Duan HT *et al.* Spatio-temporal distribution of cyanobacteria blooms based on satellite imageries in Lake Taihu, China. *J Lake Sci*, 2008, **20**(6) : 687-694. DOI: 10.18307/2008.0605. [马荣华, 孔繁翔, 段洪涛等. 基于卫星遥感的太湖蓝藻水华时空分布规律认识. 湖泊科学, 2008, **20**(6) : 687-694.]
- [13] Zhu L, Wang Q, Wu CQ *et al.* Monitoring and annual statistical analysis of algal blooms in Chaohu based on remote sensing. *Environmental Monitoring in China*, 2013, **29**(2) : 162-166. [朱利, 王桥, 吴传庆等. 巢湖水华遥感监测与年度统计分析研究. 中国环境监测, 2013, **29**(2) : 162-166.]
- [14] 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.]
- [15] 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.]
- [16] Niu Z. Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model. *Journal of Applied Remote Sensing*, 2012, **6**(1) : 063507.
- [17] Dao PD, Mong NT, Chan HP. Landsat-MODIS image fusion and object-based image analysis for observing flood inundation in a heterogeneous vegetated scene. *Glscience & Remote Sensing*, 2019, **56**(8).

- [18] Wu MQ, Niu Z, Wang CY. Assessing the accuracy of spatial and temporal image fusion model of complex area in South China. *Journal of Geo-information Sciences*, 2014, **16**(5): 776-783. [邬明权, 牛铮, 王长耀. 多源遥感数据时空融合模型应用分析. 地球信息科学学报, 2014, **16**(5): 776-783.]
- [19] Zhu X, Chen J, Gao F *et al.* An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions. *Remote Sensing of Environment*, 2010, **114**(11): 2610-2623.
- [20] Wei R, Shan J. Spatial and temporal fusion for urban land surface temperature image mapping. *Geomatics and Information Science of Wuhan University*, 2018, **43**(3): 428-435. [魏然, 单杰. 城市地表温度影像时空融合方法研究. 武汉大学学报: 信息科学版, 2018, **43**(3): 428-435.]
- [21] Dong SY, Zhang WJ, Xu JY. Comparative analysis of spatial-temporal fusion algorithm based on weight filter. *Beijing Surveying and Mapping*, 2019, **33**(4): 375-381. [董世元, 张文娟, 许君一. 基于权重滤波的时空融合算法对比分析. 北京测绘, 2019, **33**(4): 375-381.]
- [22] Bai L, Cai J, Liu Y *et al.* Responses of field evapotranspiration to the changes of cropping pattern and groundwater depth in large irrigation district of Yellow River basin. *Agricultural Water Management*, 2017, **188**: 1-11.
- [23] Cortes C, Vapnik V. Support-vector networks. *Machine Learning*, 1995, **20**(3): 273-297.
- [24] He YX, Zhou J. Comparison of accuracy of several supervised classification methods in ENVI. *Technological Innovation and Application*, 2018, (13): 22-24. [何予霄, 周江. ENVI 中几种监督分类方法精度比较. 科技创新与应用, 2018, (13): 22-24.]
- [25] Pan C. Recognition and change detection of cyanobacteria bloom using the improved SVM model based on the wavelet kernel and dual weight. *Shanghai Land & Resources*, 2013, **34**(4): 38-43. [潘琛. 基于小波核双重加权 SVM 模型的蓝藻水华识别与变化检测. 上海国土资源, 2013, **34**(4): 38-43.]
- [26] Mou DD, Liu L. Comparative study of ELM and SVM in hyperspectral image supervision classification. *Remote Sensing Technology and Application*, 2019, **34**(1): 115-124. [牟多铎, 刘磊. ELM 与 SVM 在高光谱遥感图像监督分类中的比较研究. 遥感技术与应用, 2019, **34**(1): 115-124.]
- [27] Zhao JC, Lin SZ, Li DW *et al.* A comparative study of intuitionistic fuzzy sets in Multi-band image fusion. *Infrared Technology*, 2018, **40**(9): 881-886. [赵竞超, 蔺素珍, 李大威等. 多波段图像融合的直觉模糊化处理方法比较. 红外技术, 2018, **40**(9): 881-886.]
- [28] Liu Y, Wang Z, Guo H *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.
- [29] Zhou XC, Xu S. Preliminary study of remote sensing monitor Lake Chaohu cyanophyte bloom. *China Environmental Monitoring*, 2010, **26**(3): 36-39. [周先传, 徐升. 巢湖蓝藻水华遥感监测初探. 中国环境监测, 2010, **26**(3): 36-39.]
- [30] Sha LW, Liu G, Wen ZD *et al.* Spatio-temporal variation rules of cyanobacteria blooms in Taihu Lake using MODIS Data. *Wetland Science*, 2018, **16**(3): 432-437. [沙林伟, 刘阁, 温志丹等. 基于 MODIS 数据的太湖蓝藻水华时空变化规律研究. 湿地科学, 2018, **16**(3): 432-437.]
- [31] 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.]
- [32] Zhang TT, Hu H, Ma XS *et al.* Long-Term spatiotemporal variation and environmental driving forces analyses of algal blooms in Taihu Lake based on multi-source satellite and land observations. *Water*, 2020, **12**(4): 1035.
- [33] Yu HK, Lu JB, Ye BS. Discriminant analysis of meteorological conditions and planning burning in forest edge based on Logistic regression. *Fire Science*, 2017, **26**(4): 214-219. [余会康, 陆继彬, 叶本闪. 基于 Logistic 回归的林缘计划烧除气象条件判别分析. 火灾科学, 2017, **26**(4): 214-219.]
- [34] Du Q, Fan W, Li K *et al.* Geohazard susceptibility assessment by using binary logical regression and information value model. *Disaster Science*, 2017, **32**(2): 220-226. [杜谦, 范文, 李凯等. 二元 Logistic 回归和信息量模型在地质灾害分区中的应用. 灾害学, 2017, **32**(2): 220-226.]
- [35] Duan YC. Logistic regression analysis of meteorological factors of dry heavy freezing damage of tea plant. *Chinese Agricultural Science Bulletin*, 2016, **32**(7): 152-156. [段永春. 茶树干燥型大冻害气象因素的 Logistic 回归分析. 中国农学通报, 2016, **32**(7): 152-156.]