

基于遥感反射率的太湖优势藻识别方法^{*}

朱雨新¹, 李云梅^{1,2**}, 张 玉¹, 王怀警¹, 蔡小兰¹, 成 鑫¹, 吕 恒^{1,2}

(1: 南京师范大学虚拟地理环境教育部重点实验室, 南京 210023)

(2: 江苏省地理信息资源开发与利用协同创新中心, 南京 210023)

摘 要: 我国淡水湖库频发水华, 不同类群形成的水华特征、危害及其治理方法差异显著, 因此, 如何区分不同藻种的遥感反射率特征, 获取湖泊优势种信息是一个亟待解决的科学问题。研究基于室内藻种培养实验, 培养了富营养化湖泊中的典型蓝藻和绿藻藻种, 其中, 蓝藻包括铜绿微囊藻(*Microcystis aeruginosa*)、假鱼腥藻(*Pseudanabaena* sp.) 和束丝藻(*Aphanizomenon* sp.), 绿藻包括小球藻(*Chlorella* sp.) 以及四尾栅藻(*Scenedesmus quadricauda*); 基于实测的遥感反射率, 经归一化处理后, 分析了不同藻种的遥感反射率特征, 构建了 *DI* (difference index) 指数以及 *ADI* (algae distinguish index) 指数, 建立了藻种分类模型, 利用验证集数据进行检验, 整体识别精度达 77.55%, Kappa 系数为 0.7178。将分类方法应用于太湖野外实测遥感反射率数据集中, 结果与实测的生物量数据有较好的匹配; 将模型应用于太湖 OLCI (ocean and land colour instrument) 影像数据, 获得了 2019 年 12 月和 2020 年 8 月太湖冬季和夏季的优势藻种分布, 总体而言, 太湖蓝藻占比在夏季及冬季均高于绿藻, 其中微囊藻显示出较明显的优势, 且夏季优势地位明显高于冬季。从季节及空间分布上看, 冬季太湖微囊藻主导区域分布在北部及南部湖湾, 假鱼腥藻主导区域主要位于湖心, 极少部分束丝藻及栅藻零散分布; 夏季, 太湖大部分水域的优势藻种为微囊藻, 假鱼腥藻主导区域分布于湖心及南部, 其余 3 种藻种的主导区域仍只占一小部分并零散分布。优势藻种遥感识别模型的构建可为湖泊水环境遥感监测和水华的预测预防提供技术支持。

关键词: 太湖; 遥感反射率; 优势藻种; 分类模型

Identification of dominant algae in Lake Taihu based on remote sensing reflectance^{*}

Zhu Yuxin¹, Li Yunmei^{1,2**}, Zhang Yu¹, Wang Huaijing¹, Cai Xiaolan¹, Cheng Xin¹ & Lv Heng^{1,2}

(1: Key Laboratory of Virtual Geographic Environment of Education Ministry, Nanjing Normal University, Nanjing 210023, P.R. China)

(2: Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, P.R. China)

Abstract: Algal blooms occur frequently in the freshwater lakes and reservoirs in China, but the characteristics, hazards, and treatment methods of different bloom-forming algae species are significantly different, consequently, a critical scientific problem that must be solved is how to distinguish the reflectance characteristics of different algal species and identify dominant species in lakes using remote sensing reflectance. Based on laboratory algae culture experiments, the research produced typical cyanobacteria species (including *Microcystis aeruginosa*, *Pseudanabaena* sp. and *Aphanizomenon* sp.) and green algae species (including *Chlorella* sp. and *Scenedesmus quadricauda*), which are prevalent in eutrophic lakes. Subsequently, the normalized *in-situ* remote sensing reflectance characteristics of different algal species were analyzed, the *DI* (difference index) and *ADI* (algae distinguish index) were constructed based on spectral differences, and then the algal species identification model was established. Validated by independent verification data set, the overall recognition accuracy and Kappa coefficient were 77.55% and 0.7178, respectively. The results of classification algorithm match well with the measured biomass data when applied to the field measured remote sensing reflectance data set of Lake Taihu. Further model tests based on the OLCI images of Lake Taihu in December 2019 and August

^{*} 2022-05-05 收稿; 2022-06-15 收修改稿。

国家自然科学基金项目 (42071299)、国家自然科学基金云南联合基金重点项目 (U2102207) 和江苏省林业科技创新与推广项目 (LYKJ[2019]32) 联合资助。

^{**} 通信作者; E-mail: liyunmei@njnu.edu.cn。

2020 show that the proportion of cyanobacteria in Lake Taihu is higher than that of green algae in summer and winter, during which *Microcystis* displayed obvious advantages and a stronger dominant position in summer than in winter. Furthermore, in terms of seasonal and geographical distributions, the dominant areas of *Microcystis* were in the northern and southern bays of Lake Taihu in winter, while *Pseudoanabaena* was in the central area of the lake, and scattered with a tiny fraction of *Aphanizomenon* and *Scenedesmus quadricauda* at the same time; compared to summer, *Microcystis* was dominant in most areas of Lake Taihu, *Pseudoanabaena* was mainly distributed in the center and south of the lake, and the other three species only scattered in a small part of the lake. The development of a remote sensing identification model of the main algal species can provide technological assistances to remote sensing monitoring of the lake environment, as well as algal bloom prediction and prevention.

Keywords: Lake Taihu; remote sensing reflectance; dominant algae species; classification model

近年来,我国淡水湖库频发水华,水华暴发不仅仅会造成水体污染、水生系统紊乱,还会直接威胁到当地的经济的发展,我国已经俨然成为了世界上水华暴发最严重、分布最广的国家之一^[1]。湖泊中形成水华的藻类主要为蓝藻、绿藻及硅藻,其中蓝藻水华以微囊藻、长孢藻、束丝藻等为主导^[2]、绿藻以小球藻和栅藻为多,常见硅藻主要为小环藻及舟形藻,我国沿海海域也频发水华,主要由硅藻及甲藻引发。各个湖泊的优势种不尽相同,太湖以微囊藻为绝对优势^[2],而微囊藻和鱼腥藻在巢湖占主导地位^[3],滇池则以微囊藻和束丝藻为主^[4]。不同类群的危害、环境适应性、时空分布特征、物种演替规律以及处理方法并不相同,因此,准确了解优势种信息,对湖泊生态系统修复、水华预警预测具有重要意义。

目前在藻类识别方面各国学者做了大量的研究,除了传统的显微镜方法^[5]之外,主要还有图像法、吸光度法^[6]、荧光光度法^[7]、高压液相色谱法^[8]、流式细胞仪法、化学发光流动射法和分子探针法^[9],这些研究大都是利用藻类的细胞形态特征、藻种 DNA 等实现藻类识别。随着遥感技术的发展,基于遥感的藻类识别方法应运而生。藻类的遥感识别主要依据水体固有光学特性或者表观光学特性建模,因此,藻类的吸收和散射特性是开展藻类遥感监测的基础^[10]。Mahoney^[11]在固有光学特性的基础上建立了鞭毛藻的生物光学定量反演模型。黄昌春等^[12]利用四阶微分、标准化系数等方法对太湖水体吸收光谱进行处理从而对藻类组成进行计算分析。赵巧华和秦伯强^[13]采用实验室培养的方式,测算五种藻类的标准化吸收光谱,分析了同种藻类吸收光谱特征及不同藻类的标准化吸收谱,最终实现了混合藻吸收谱的分离。戴红亮等^[14-15]和王瑜^[16]基于室内控制试验及模拟试验,结合色素吸收特征、藻细胞特征和反射光谱特征等,探索性地建立了藻种门类间比例遥感估算方法。褚乔等^[10]在3种主要水华蓝藻固有光学特性的基础上,通过甄别不同水华蓝藻的吸收、散射和后向散射光谱的特征波段,构建了基于吸收和散射特性的5种水华蓝藻类群的非线性最优化定量识别模型,实现了巢湖主要水华蓝藻类群的定量监测,初步分析了巢湖主要水华蓝藻类群的时空分布特性。Xi等^[17]基于经验正交函数(EOF)检索各种浮游植物色素,并采用广泛的全球原位色素测量数据集,检索多种藻类的叶绿素 *a* 浓度,进而在全球范围内反演不同浮游植物的叶绿素 *a* 浓度。Sun等^[18]利用多年累积的浮游植物色素和浮游植物吸收光谱数据集,获得了8类常见浮游植物类群组成信息,进一步采用矩阵反演方法,构建了各浮游植物类群的比吸收系数,实现了基于浮游植物吸收系数的浮游植物类群组成反演。此类研究多是通过获取不同藻种的标准化吸收、散射曲线,分析其特征差异并配合一定的数学方法进行藻类研究,大多聚焦于不同门类之间的差异,对于同门类不同藻种的光学特性差异研究较少。此外,国内外学者初步开展了基于藻类遥感识别的相关研究。崔廷伟等^[19]利用687~728 nm波段的特征反射峰对藻类进行了识别。Xi等^[20]以代表5种浮游植物光谱的标准遥感反射率光谱为基础,建立了一个广泛的数据集,提出了一种基于相似指数法的浮游植物群落识别方法,将6个分类群的128个浮游植物藻类谱分为5个谱群,并对该方法的性能进行了测试。Shen^[21]基于绿-红光谱斜率,提出了一种利用MERIS数据识别东海藻华优势种的方法。以上研究证明利用遥感反射率光谱进行藻类分类是可行的,但是,现有研究大多只针对海洋中的藻类,且大都是针对不同门类获取的分类结果,针对内陆湖泊可应用于卫星影像的蓝藻或绿藻种间识别的模型鲜有报道。此外,内陆水体的光学特性复杂,对光谱信息的干扰因素较多,增大了利用遥感信息进行藻类分类的困难,需要针对内陆水体的藻类及其生存环境,开展藻类遥感识别方法的研究。

因此,本研究拟通过内陆湖泊代表性藻种培养实验,建立不同藻种遥感反射率数据集,进而分析太湖优

势蓝藻、绿藻藻种的反射率特征,并构建基于特征波段的藻种分类模型,应用于卫星影像,分析太湖优势藻种的空间分布特征。

1 材料和方法

1.1 研究区概况

研究区太湖是中国的第三大淡水湖泊,位于长江三角洲($30^{\circ}55'40''\sim 31^{\circ}32'58''\text{N}$, $119^{\circ}52'32''\sim 120^{\circ}36'10''\text{E}$)(图1),北临江苏无锡,东近江苏苏州,是周边城市的重要水源地。湖泊面积为 2339 km^2 ,平均水深 1.9 m 。近年来太湖富营养化问题严重,是我国典型的内陆富营养化湖泊,水华主要发生在夏季至冬季。太湖藻类主要由蓝藻、绿藻、硅藻、隐藻、裸藻和甲藻 6 大门类组成,其中蓝藻占绝对优势,其次为绿藻^[2]。

分别于 2019 年 12 月 8 日和 2020 年 8 月 16 日在太湖进行了野外观测,两次共获取 29 个表层样点数据,并取水样分析不同浮游植物生物量。

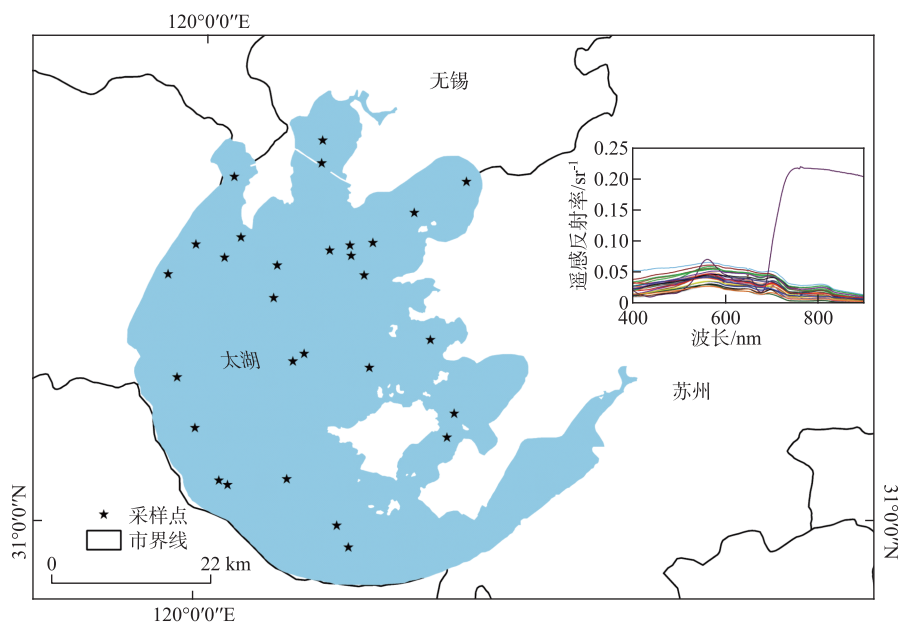


图1 研究区位置示意图及样点分布(右上角为2019年12月8日及2020年8月16日在太湖野外实测获取的29条遥感反射率数据)

Fig.1 Location of study area and distribution map of sample sites (The upper right corner shows 29 remote sensing reflectance data obtained from field measurements in Lake Taihu on December 8, 2019 and August 16, 2020)

1.2 藻种培养

由太湖藻种调查报告以及生物量数据报告可知^[22],太湖蓝藻中的微囊藻、束丝藻、假鱼腥藻以及绿藻中的小球藻和栅藻生物量占总生物量的80%以上,可视为代替优势种。因此,本研究选取铜绿微囊藻(*Microcystis aeruginosa*)、束丝藻(*Aphanizomenin* sp.)、假鱼腥藻(*Pseudanabaena* sp.)、小球藻(*Chlorella* sp.)和四尾栅藻(*Scenedesmus quadricauda*)进行室内培养实验,藻种均为纯藻,购于中国科学院淡水藻种库,编号分别为FACHB-1174、FACHB-1399、FACHB-1277、FACHB-8及FACHB-44,藻种培养基均为BG-11。藻种储存在 25°C 、2000 lx光强的光照恒温培养箱中。实验前,将在光照培养箱中驯化一周的藻种转移至事先灭菌的50 mL锥形瓶中,生长7~10天后进行逐级扩大培养^[23]。培养期间每隔2~4 h人工摇瓶一次,直到达到实验所需藻量(约5 L)。

1.3 遥感反射率测量

纯藻的实验过程为:于地面放置 10 m×10 m 纯黑尼龙黑布,并将纯黑的玻璃桶(四壁喷上黑漆)放置其上,玻璃桶尺寸为底部直径 15 cm,高 15 cm。向桶中加入 1 L 纯水,然后以 200 mL 为梯度分 3 次加入藻的母液并搅拌均匀,待水面平静后测量一次遥感反射率,测完取 200 mL 供室内试验使用,每一梯度的数据以藻种名+1/2/3 命名记录。3 次实验共获得 40 个有效数据,选取第 1 次实验的 15 个数据点进行模型的构建,剩余 25 个数据点作为验证集。

混藻实验数据用以验证算法在混合情况的适用性,此阶段实验进行了两种藻种之间的混合,以铜绿微囊藻和束丝藻的混合为例,向纯黑的有机玻璃桶中加入 1 L 纯水、50 mL 铜绿微囊藻和 100 mL 的束丝藻,分两次再加入 50 mL 束丝藻,与纯藻实验相同,每加入一次藻液搅拌均匀,待水面平静后测量一次遥感反射率。经实验共获得 24 个混藻有效实验样点,具体体积配比如表 1 所示,实验中默认所有藻种生物量均为 1×10^6 mg/L。

表 1 混藻实验体积配比
Tab.1 Volume ratio of mixed algae experiment

样点编号	藻液体积/mL				
	铜绿微囊藻	束丝藻	假鱼腥藻	小球藻	四尾栅藻
1	50	100	—	—	—
2	50	150	—	—	—
3	50	200	—	—	—
4	50	—	100	—	—
5	50	—	150	—	—
6	50	—	200	—	—
7	50	—	—	100	—
8	50	—	—	150	—
9	50	—	—	200	—
10	100	50	—	—	—
11	150	50	—	—	—
12	200	50	—	—	—
13	—	50	100	—	—
14	—	50	150	—	—
15	—	50	200	—	—
16	50	—	—	—	100
17	50	—	—	—	150
18	50	—	—	—	200
19	—	100	50	—	—
20	—	150	50	—	—
21	—	200	50	—	—
22	100	—	—	50	—
23	150	—	—	50	—
24	200	—	—	50	—

于 2021 年 9 月和 12 月选取晴朗无云的天气使用美国 ASD 公司生产的野外光谱仪进行遥感反射率测量。仪器波段范围为 350~1050 nm,光谱分辨率为 2 nm。测量观测几何参照 Cimoli 等^[24]、Nam 等^[25]的研究,具体观测如图 2 所示。

测量的参数包括:水体总的辐亮度 L_{sw} 、仪器对着天空测量的天空漫散射光值 L_{sky} 、标准灰板辐亮度 L_p (反射率为 30%)。仪器对着水面测量时,每次连续采集 15 条光谱曲线,处理时剔除异常数据后取均值为水面反射率。遥感反射率的计算公式为:

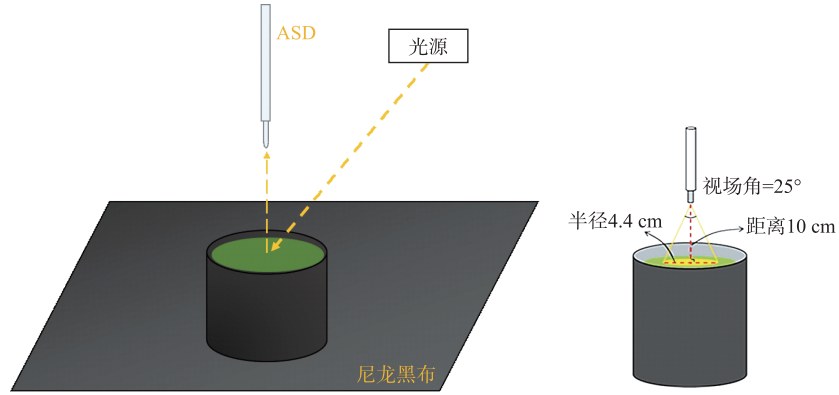


图 2 光谱测量示意图

Fig.2 Schematic diagram of spectral measurement

$$R_{rs} = (L_{sw} - r \cdot L_{sky}) \rho_p / (\pi \cdot L_p) \quad (1)$$

式中, ρ_p 为标准参考版的反射率, r 取 0.022。最终取不同藻种实测遥感反射率减去同装置测量的纯水反射率(背景值)做分析。

1.4 室内分析数据

藻类颗粒物吸收系数采用 Tassan 和 Ferrari^[26] 建立的定量滤膜技术测定。首先,利用直径 25 mm、孔径 0.7 μm 的 Whatman GF/F 滤膜过滤 25 mL 水样,然后,在岛津 UV3600plus 型分光光度计下用积分球在透射率模式与反射率模式下分别测量,得到颗粒物吸收系数 $a_p(\lambda)$, 非色素颗粒物吸收系数 ($a_d(\lambda)$) 测量过程与上述步骤相同,只需在测量之前先将水样用 0.1 mol/L 的次氯酸钠溶液漂白 3 遍,最后一遍至少停留 30 min。色素颗粒物吸收系数为总颗粒物吸收系数与非色素颗粒物吸收系数之差,即 $a_{ph}(\lambda) = a_p(\lambda) - a_d(\lambda)$ 。

浮游植物生物量采用细胞体积转换法测定,通过显微镜检查浮游植物群落组成,并采用随机视野法对每一种属浮游藻类进行计数,得到各物种的浮游植物丰度,并求和得到总浮游植物丰度。由于浮游植物细胞密度接近于水的密度,1 mm^3 约等于 1 mg 鲜重生物量,因此生物量 (mg/L) 为体积 (μm^3) 与丰度 (cells/L) 的乘积。

1.5 影像及数据处理

OLCI(ocean and land colour instrument) 传感器基于中分辨率成像光谱仪 MERIS (medium resolution imaging spectrometer) 开发,优化后相比较 MERIS 传感器改进了 7 个光谱波段,可观测到沿海地区的海洋水色,在 400~1020 nm 波段范围内共设置 21 个光谱通道,其中包含 16 个水色波段,并具有高信噪比的优势。空间分辨率为 300 m,能反映精细尺度下湖泊水质状况。结合现场数据采集时间,使用了 2019 年 12 月 8 日以及 2020 年 8 月 17 日 OLCI 影像。采用的数据为 Level-1B 产品,在 SNAP 中使用 per-pixel geo-coding 打开,并对太湖周围地区进行裁剪、重投影、辐亮度转反射率和瑞利散射校正等操作,将预处理后的影像输出为 ENVI 可读的 DAT 格式进行后续操作,大气校正方法选用 6S 模型。使用均方根误差 ($RMSE$) 及平均绝对百分比误差 ($MAPE$) 进行大气校正精度评价,精度评价公式为:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{y}_i - y_i)^2} \quad (2)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\bar{y}_i - y_i}{y_i} \right| \quad (3)$$

式中, n 为样点数, $\bar{y}_i$ 为 OLCI 影像大气校正后的遥感反射率, y_i 为实测遥感反射率。

2 藻种分类算法的构建

2.1 遥感反射率特征

如图3所示,实验室观测的遥感反射率校正前后数值略有差别,对光谱的波谷特征影响不大。总体来看,随着浓度的升高,藻种遥感反射率绝对值呈上升趋势,且曲线特征愈加明显,浓度较低时,曲线较为平滑,随着浓度的升高,遥感反射率绝对值上升的同时波谷特征也愈加明显。遥感反射率光谱与藻类色素吸收和藻颗粒散射具有明显的共变关系。铜绿微囊藻(图3A)在440~490 nm处藻类叶绿素和类胡萝卜素的强吸收造成该波段遥感反射率低,光谱上表现出一个谷^[27]。620和675 nm处藻蓝蛋白和叶绿素 a 主导的强吸收使得遥感反射率光谱出现明显的谷^[27]。而550和700 nm附近,藻类色素普遍吸收微弱且藻类颗粒物散射较强^[28],同时,由于675 nm处的强吸收,700 nm出现一个反射峰。这些光谱特征与乐成峰等^[29]的研究结果一致。

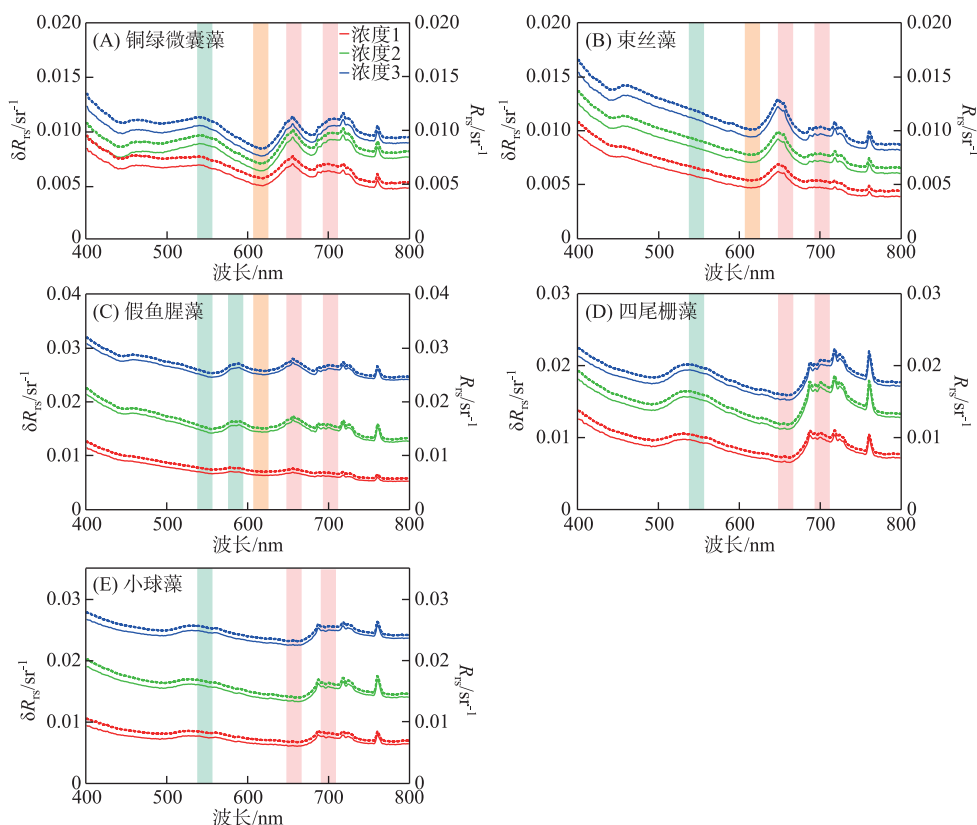


图3 不同藻种遥感反射率曲线(实线表示校正后遥感反射率(δR_{rs});虚线表示校正前遥感反射率(R_{rs})。绿、橙、红阴影部分表示特征波段:绿色表示560和590 nm;橙色表示620 nm;红色表示660和700 nm)

Fig.3 Remote sensing reflectance spectra of different algae species (The solid line in the figure represents the remote sensing reflectance after correction (δR_{rs}), and the dotted line represents the remote sensing reflectance before correction (R_{rs}). The shaded parts are characteristic bands: green represent 560 and 590 nm; orange represent 620 nm; red represent 660 and 700 nm)

束丝藻(图3B)曲线峰谷特征与铜绿微囊藻大致相同,只是在440~600 nm较铜绿微囊藻更为平滑,没有出现峰谷特征,620和675 nm处藻蓝蛋白和叶绿素 a 主导的强吸收使得遥感反射率光谱出现反射谷^[30],但是反射谷没有铜绿微囊藻明显,同时,由于675 nm处的强吸收因而700 nm出现一个反射峰,这些特征为

后续藻种区分奠定了基础。

假鱼腥藻(图 3C)的遥感反射率数值较其他两种蓝藻更高,其曲线特征与其余两种蓝藻大致相同,但其峰谷趋势却没有其余两种蓝藻明显,620 和 675 nm 处藻蓝蛋白和叶绿素 *a* 主导的强吸收使得遥感反射率光谱出现明显的谷^[30]。而 700 nm 附近,藻类色素普遍吸收微弱且藻类颗粒物散射较强^[28],同时,叠加了 675 nm 处的强吸收,使得 700 nm 出现一个反射峰。在 590 nm 处较其他两种蓝藻多出一个峰,可能是由假鱼腥藻含有的藻红素和藻蓝素分别在 580 和 620 nm 附近的吸收共同造成的^[31],假鱼腥藻生长过程呈现棕红色也正是由藻红素这种色素引起的。

四尾栅藻(图 3D)与小球藻(图 3E)在 540 和 700 nm 附近存在明显的特征反射峰。前者是藻类色素物质的弱吸收和藻颗粒的散射造成的,后者则主要是由于 675 nm 的强吸收形成的反射峰。总体来看,随着藻浓度的升高,反射峰的峰高逐渐增加。与 3 种蓝藻不同的是,小球藻和四尾栅藻遥感反射率曲线均没有表现出 620 nm 附近藻蓝蛋白的反射谷,因此,620 nm 附近的反射谷或许可作为区分蓝绿藻的遥感识别特征。而四尾栅藻与小球藻的曲线几乎没有区别,这也为进一步区分绿藻增加了难度。

2.2 藻种分类算法的构建

不同藻种的遥感反射率数值大小存在一定的差别,大小的差别来源于很多方面:光源强度、浓度高低、色素组分等。仅从绝对值比较会带来很多的不确定性,且一些系统误差也会对数值造成影响,难以构建稳健的内陆湖泊藻种分类算法。为避免数值高低不一致对方法造成的误差,对 5 种藻种的实测遥感反射率曲线进行了归一化处理^[20],将所有波段的遥感反射率除以 560 nm 处的遥感反射率数值得到归一化结果,如图 4 所示。

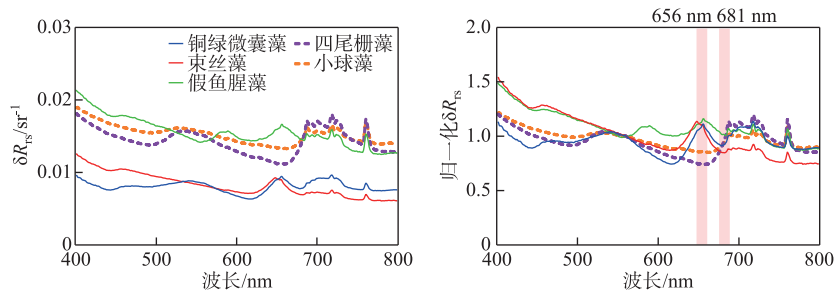


图 4 5 种藻种归一化前(a)、后(b)的遥感反射率曲线

Fig.4 Remote sensing reflectance spectra of five algae species before (a) and after (b) normalization

由于蓝绿藻本身形态、所含色素的差异,蓝藻反射率曲线在 660 nm 左右呈现出峰值,而绿藻却为谷值;681 nm 处蓝藻呈现一个小的波谷,而绿藻在此处仍处于快速上升阶段,此特征由 620 nm 处藻蓝蛋白的吸收和 675 nm 叶绿素的强吸收共同造成。基于此特征,定义蓝绿藻峰谷差异指数 *DI* (difference index) 进行蓝藻和绿藻的区分,此外,由于两种绿藻的 *DI* 值差异较大,根据 *DI* 阈值,还可进一步区分四尾栅藻和小球藻。*DI* 的计算方法为:

$$DI = R_{rs}(\lambda_1) - R_{rs}(\lambda_2) \quad (4)$$

式中, $\lambda_1 = 656 \text{ nm}$, $\lambda_2 = 681 \text{ nm}$ 。

蓝藻中 3 种藻的光谱差异主要体现在 500~700 nm 波段,因此,构建了参数 *ADI* (algae distinguish index) 计算该波段范围反射谷的深度(图 5),通过 *ADI* 值的大小区分不同蓝藻藻种。*ADI* 的计算公式为:

$$ADI = R_{rs}(\lambda_1) - R_{rs}(\lambda_2) + \frac{(R_{rs}(\lambda_3) - R_{rs}(\lambda_2))(\lambda_2 - \lambda_1)}{\lambda_3 - \lambda_1} \quad (5)$$

式中, $\lambda_1 = 560 \text{ nm}$, $\lambda_2 = 620 \text{ nm}$, $\lambda_3 = 656 \text{ nm}$ 。

综上,利用第一次实验获取的遥感反射率实测数据率定 *DI* 及 *ADI* 的阈值,结果如图 6 所示,即两种绿藻的 *DI* 值均小于 0,而 3 种蓝藻的 *DI* 值均大于 0,两种绿藻的 *DI* 值差异较大,将其阈值界定为 -0.0015 以进行两种绿藻的区分;3 种蓝藻间的阈值,首先通过确定不同种类之间的分界线,然后经多次试错最终确定,算法流程如图 7 所示。

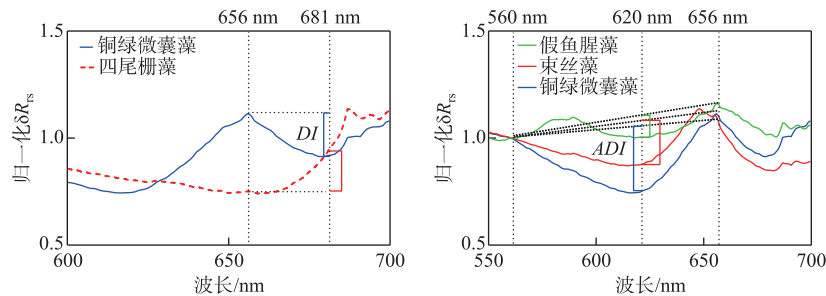


图 5 DI 及 ADI 构建示意图 (DI 示意图以铜绿微囊藻和四尾栅藻为例)
Fig.5 Schematic of DI and ADI (The DI schematic diagram takes *Microcystis aeruginosa* and *Scenedesmus quadricauda* as examples)

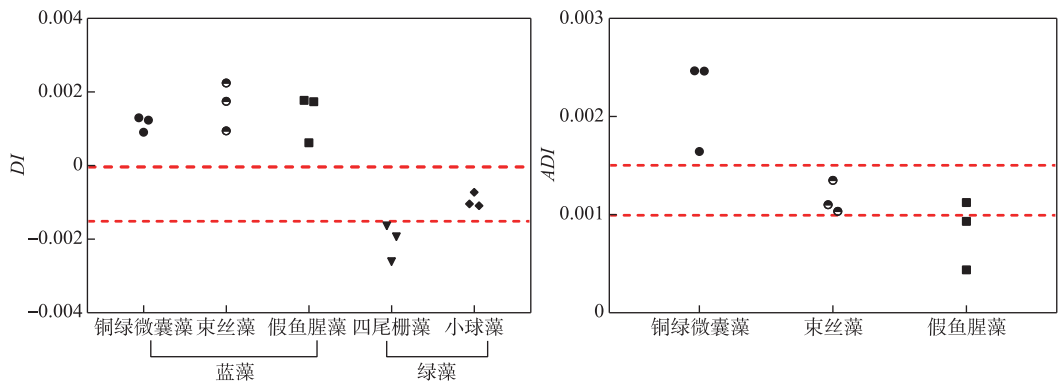


图 6 模型阈值选取 (虚线表示阈值)
Fig.6 Model threshold selection (The dotted line indicates the threshold)

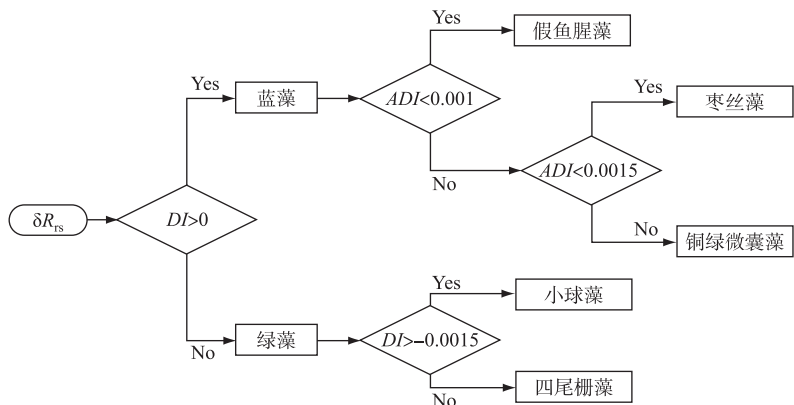


图 7 算法流程图
Fig.7 Flow chart of algorithm

2.3 算法精度评价

2.3.1 基于室内观测数据的精度评价 使用 1.3 节所述的纯藻及混藻所测得的遥感反射率作为验证集 (混藻验证集中, 当某藻种生物量占比 > 60% 时认定该藻为优势藻种), 实验共获得 49 个验证样本数据, 模型的精度评价使用混淆矩阵和 Kappa 系数 (K 值)^[32] 来表示, 由于利用 DI 指数区分蓝藻和绿藻精度为 100%, 因此

评价只基于 5 种藻种类别来计算。得出的混淆矩阵及精度评价结果如表 2 所示。

表 2 藻种分类模型混淆矩阵及精度评价结果
Tab.2 Confusion matrix and accuracy evaluation results of algal species identification model

		实际结果				
		铜绿微囊藻	束丝藻	假鱼腥藻	小球藻	四尾栅藻
识别结果	铜绿微囊藻	11	0	0	0	0
	束丝藻	0	3	0	0	0
	假鱼腥藻	0	8	11	0	0
	小球藻	0	0	0	5	0
	四尾栅藻	0	0	0	3	8
评价结果		整体精度 = 77.55%			Kappa 系数 = 0.7178	

总体来看,整体识别精度达 77.55%,Kappa 系数为 0.7178,说明数据具有较高的一致性,且精度较好,从类别识别来看,在验证中表现出了较高的精度,利用 *DI* 指数区分蓝藻和绿藻是可行的,但对两种绿藻的区分效果相对差些;利用 *ADI* 指数区分铜绿微囊藻和假鱼腥藻的精度较高,但是,当束丝藻浓度较低时,可能被误识别为假鱼腥藻,小球藻浓度较低时,可能被误识别为四尾栅藻,混藻中束丝藻与假鱼腥藻混合容易产生误分,与铜绿微囊藻混合时,浓度占比小于 67%时也产生了误分,混藻被错分的样点情况如表 3 所示。

表 3 错误分类的混藻样品
Tab.3 Misclassified samples of mixed algae

样品编号	藻液体积/mL				
	铜绿微囊藻	束丝藻	假鱼腥藻	小球藻	四尾栅藻
1	—	100	50	—	—
2	—	150	50	—	—
3	—	200	50	—	—
4	50	100	—	—	—

2.3.2 基于野外实测数据的精度评价 使用 2.1 节所述太湖野外观测遥感反射率数据及生物量数据进行验证,实验室分析生物量结果表明数据集共包含 45 种浮游植物,蓝藻和绿藻占绝对主导地位,实测生物量范围为 0.07~41.78 mg/L。数据中共有 27 个由蓝藻主导的样点,其中优势藻为微囊藻、假鱼腥藻及束丝藻的样点数分别为 20、4 和 3 个;绿藻主导的样点 2 个,优势藻为小球藻及栅藻的样点各 1 个(表 4)。经验证,总体精度达 79.31%。算法准确地区分了蓝藻、绿藻主导的样点,并能准确识别由铜绿微囊藻主导的样点,由于湖泊中藻类组成复杂,导致部分点位被误判为微囊藻,后续需增加实验补充更多的藻种类型。

表 4 野外实测数据识别结果
Tab.4 Identification results of field measured data

优势藻	微囊藻	假鱼腥藻	束丝藻	小球藻	四尾栅藻
总个数	20	4	3	1	1
正确识别个数	18	2	2	0	1

2.3.3 基于卫星影像的算法评估 将算法应用于 2019 年 12 月 8 日和 2020 年 8 月 17 日的 OLCI 影像上,获取冬、夏季太湖优势藻的时空分布。所使用的 OLCI 数据为 Level-1B 产品,大气校正方法选用 6S 模型。选取 560、620、665 以及 681 nm 波段进行计算,相关波段的实测值与 6S 大气校正值的散点图如图 8 所示。总体来说,大气校正精度尚可,数据大都集中于 1:1 线附近。

所得优势藻结果分布如图 9 所示。在冬季,微囊藻主导的区域主要分布在太湖北部及南部沿湾;束丝

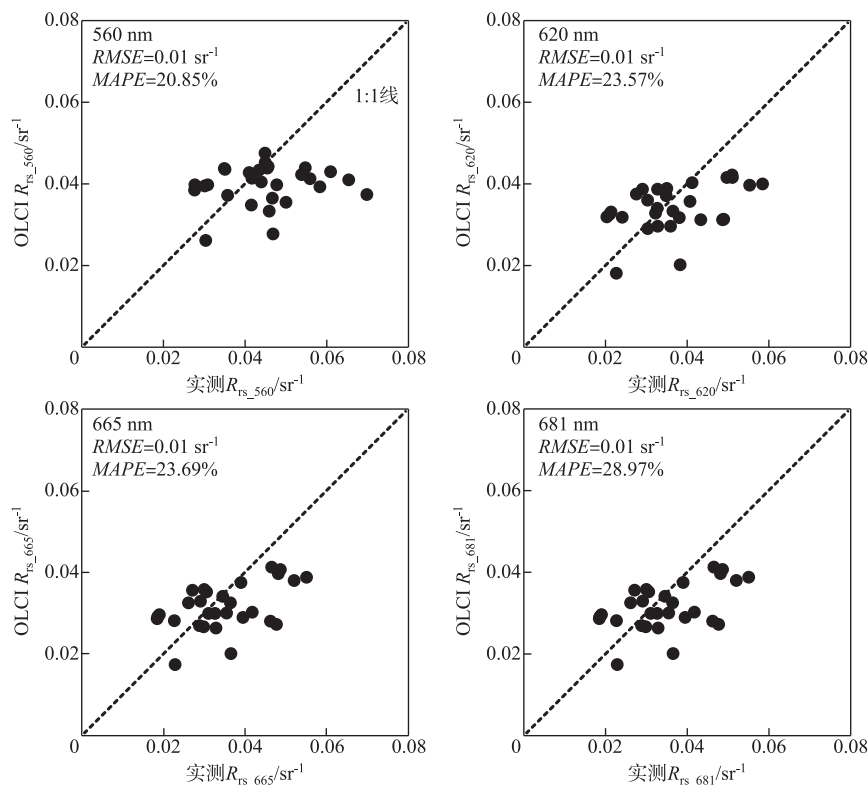


图 8 OLCI-6S 大气校正后特征波段遥感反射率实测值与校正值散点图
Fig.8 Scatter plot of remote sensing reflectance of characteristic bands after OLCI-6S

藻则在北部及南部沿湾小部分零碎区域占优;假鱼腥藻主导的区域位于湖心;栅藻为优势藻的区域只占一小部分并位于南部。此季节并未识别出小球藻主导的区域,在生物量实测结果中也几乎没有出现小球藻为主导地位的点位。在夏季,微囊藻的分布范围更广,全湖近一半水域优势藻种识别为微囊藻;束丝藻为优势的区域呈零散分布;假鱼腥藻主导区域则分布于湖心及南部;栅藻及小球藻只占一小部分,于北部南部零碎

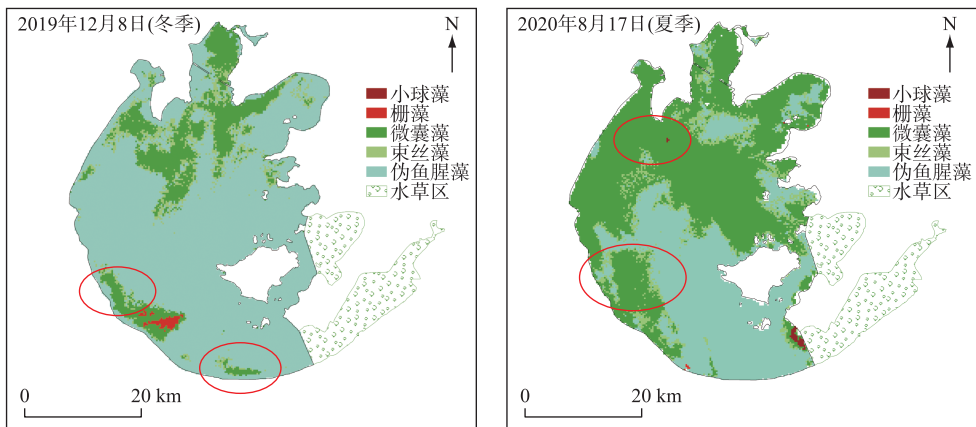


图 9 2019 年 12 月 8 日 (冬季) 和 2020 年 8 月 17 日 (夏季) 太湖优势藻种分布
Fig.9 Distribution of dominant algae species in Lake Taihu on December 8, 2019 (winter) and August 17, 2020 (summer)

分布。总体而言,蓝藻占比高于绿藻,由优势种分布结果可以看出,夏季微囊藻的优势地位明显高于冬季;在水华暴发区(红圈处),大部分区域优势藻种均识别为微囊藻,这与前人研究认为太湖中水华主要由微囊藻引起的结果一致^[33]。

利用有同步生物量数据的 29 个样点进行星地同步验证,精度如表 5 所示,总体精度达到 51.72%。

表 5 野外实测数据识别结果
Tab.5 Identification results of field measured data

优势藻	微囊藻	假鱼腥藻	束丝藻	小球藻	四尾栅藻
总个数	20	4	3	1	1
正确识别个数	12	1	1	0	1

同时,与湖泊一流域科学数据中心(<http://lake.geodata.cn/>)下载的 2017 年 0.5 m 水深藻生物量数据进行对比,为方便与本研究提出的方法进行比较,将甲藻、硅藻等藻相加作为其他,只列出了蓝藻及绿藻占总藻量的比例,数据显示 8 月和 12 月的站点数据优势藻均为蓝藻,对比站点经纬度对应的 OLCI 影像像元(表 6),识别结果也均为蓝藻,二者结果一致。

3 讨论

色素组成是影响反射和吸收光谱特性的主要因素^[30],在实验中发现,蓝藻门和绿藻门的色素颗粒物吸收系数曲线存在差异。图 10 为室内实验测量所得的 5 种藻种 3 个梯度的色素颗粒物吸收系数,属于蓝藻门的 3 种藻种(图 10A,B,C)主要特征为在 450、620 以及 675 nm 左右出现吸收峰,在 550 及 660 nm 左右出现吸收谷。440 nm 左右的吸收峰是由于叶绿素及类胡萝卜素的强吸收,620 nm 的特征则是由藻蓝蛋白引起的,675 nm 处同样是由于叶绿素 *a* 引起的强吸收,550 nm 处藻类吸收微弱且散射较强,660 nm 左右处的特征则是由 620 nm 处藻蓝蛋白强吸收及 675 nm 叶绿素 *a* 强吸收共同造成的。此外,假鱼腥藻在 580 nm 处较其余两种蓝藻多一个吸收峰,这是由藻红素^[31]引起;属于绿藻门的两种藻种(图 10D,E)主要特征为 440、490 及 675 nm 左右出现吸收峰,550 nm 左右出现吸收谷,440 及 675 nm 出现吸收峰及 550 nm 左右出现吸收谷的原因与蓝藻相同,480 nm 出现峰值是因为叶绿素 *b* 及 α -胡萝卜素强吸收,因为绿藻中不含藻蓝素^[34],绿藻并未出现 620 nm 处的峰值,也正因此,660 nm 左右处并未出现谷值。正是由于这些色素差异,才有可能利用 560、620、660 及 680 nm 进行藻类识别算法的构建,色素组成及其浓度差异,最终体现到遥感反射率的差异中,是遥感识别优势藻类的物理基础。

相较于色素颗粒物吸收系数曲线,遥感反射率的峰谷特征会相对弱化,并且随着优势藻种占比发生变化,如图 11A 所示,以微囊藻和小球藻的混合为例(蓝绿混藻),当小球藻的占比较小时(混藻 1),曲线显示为蓝藻的特征:620 nm 出现反射谷,660 nm 附近出现反射峰,随着小球藻占比的逐渐增大(混藻 2、混藻 3),优势藻逐渐变为小球藻,620 nm 处的反射谷不再明显,660 nm 处的反射峰也趋于平缓。野外数据集中记录为表面有明显藻颗粒漂浮的代表性点位(图 11B),其遥感反射率曲线大多表征为两种形态,即点 201912THL11_1 和 202008THL_1(两条加粗线),若呈现出红色线的特征,则藻种信号较强,可以正确识别为微囊藻主导,而呈现出紫色线特征的点位则较为复杂,藻种反射率特征不明显,判别精度不高。相较于室内实验,野外环境复杂,悬浮物及 CDOM 等水体组分的存在使得遥感反射率呈现多种形态^[35],室内培养与野外藻种的生存状态、藻种的多样性等差异性都会影响模型的识别精度,这些问题都是后续研究需要解决的重点问题。

此外,由于大气环境、水体的反射较低等因素影响,卫星遥感传感器获取的水面信息非常微弱,给优势藻类的卫星遥感识别造成了困难。为了提高藻类识别的精度,后续研究需进一步补充验证集以验证方法精度,且在大气校正、多种类型藻类光学特性方面,还需进行更进一步的研究。本研究基于 OLCI 影像进行了初步的方法试验,后期可推广至有 560、620、660 及 680 nm 相近波段的其他影像。

4 结论

本研究基于室内实验培养藻种的遥感反射率,提出了能够区分 5 种不同蓝绿藻的方法。研究中培养了

表 6 2017 年 0.5 m 水深藻生物量数据
Tab.6 Algae biomass data in water depth of 0.5 m in 2017

调查年份	调查月份	观测站代码	绿藻占总藻量比例/%	蓝藻占总藻量比例/%	其他/%
2017	8	THL13	1.26	98.11	0.63
2017	8	THL15	2.34	92.38	5.28
2017	8	THL16	3.89	81.73	14.38
2017	8	THL17	2.32	94.58	3.10
2017	8	THL18	0.00	96.42	3.58
2017	8	THL20	0.10	97.72	2.18
2017	8	THL21	1.98	97.03	0.99
2017	8	THL22	7.68	87.48	4.83
2017	8	THL24	0.36	98.75	0.89
2017	8	THL25	4.14	89.31	6.55
2017	8	THL26	5.50	61.50	33.00
2017	8	THL27	33.93	62.05	4.02
2017	8	THL28	2.44	94.15	3.41
2017	8	THL29	0.14	97.42	2.44
2017	8	THL30	0.41	98.59	1.00
2017	8	THL31	0.46	99.35	0.19
2017	8	THL00	1.49	91.67	6.85
2017	8	THL05	0.00	98.35	1.65
2017	8	THL06	11.59	73.33	15.08
2017	8	THL07	0.00	99.75	0.25
2017	8	THL08	0.20	98.62	1.18
2017	8	THL09	2.88	90.31	6.81
2017	8	THL10	5.66	81.62	12.73
2017	8	THL11	9.29	50.44	40.27
2017	8	THL12	6.63	80.87	12.50
2017	12	THL00	1.20	89.43	9.37
2017	12	THL01	5.37	85.78	8.85
2017	12	THL03	0.15	99.04	0.81
2017	12	THL04	0.99	98.24	0.77
2017	12	THL05	0.08	99.25	0.68
2017	12	THL06	1.55	95.51	2.94
2017	12	THL07	0.00	99.23	0.77
2017	12	THL08	0.13	99.68	0.19
2017	12	THL13	0.26	99.17	0.57
2017	12	THL14	4.10	90.26	5.64
2017	12	THL32	0.06	99.28	0.66

蓝藻门及绿藻门中的 5 种藻种,经实验获取遥感反射率,通过分析其特征并经归一化处理后,基于特征波段构建了 *DI* 指数和 *ADI* 指数进行藻种分类识别,结果表明:在室内培养藻种验证集中精度较好,总体精度为 77.55%,利用 *DI* 指数可以较好地区分蓝藻和绿藻,利用 *ADI* 指数可以有效区分 3 种蓝藻,能够准确识别出铜绿微囊藻和假鱼腥藻。将模型应用于太湖野外观测数据,总体精度达 79.31%,能够识别出样点的主导藻种。尝试将方法应用于卫星影像,并依据结果分析太湖优势藻种的空间分布特征,星地同步结果总体精度为 51.72%。本研究为内陆湖泊藻种识别提供了新思路,但要真正实现种间识别的应用要求,下一步还需对遥感反射率数据集做进一步的补充,将内陆湖泊 CDOM 及悬浮物的影响考虑进来,并提高大气校正精度,将方法应用于长时序的卫星影像,为相关部门开展生态修复工作提供依据。

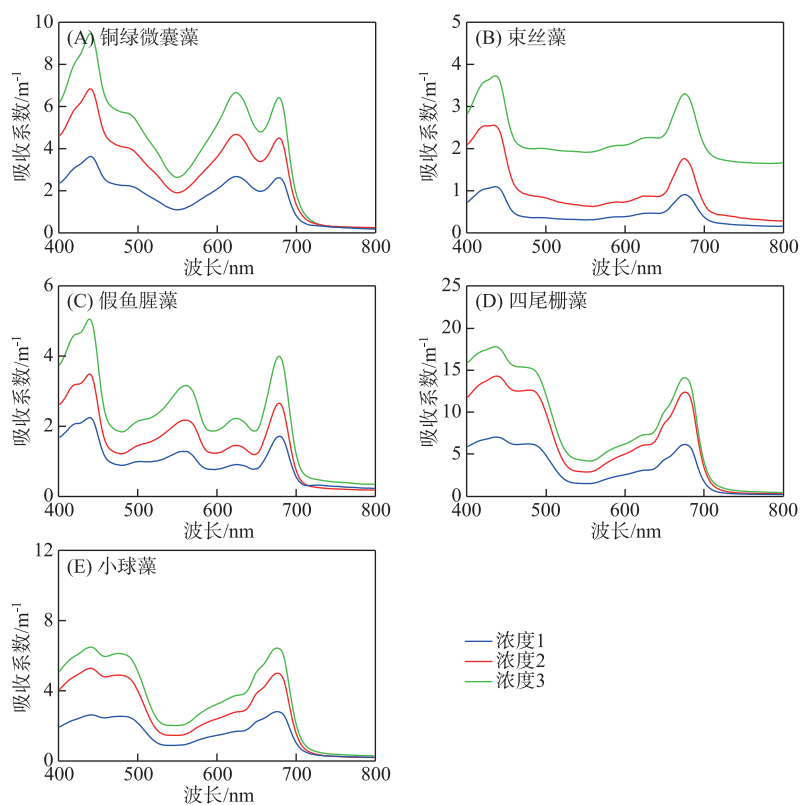


图 10 5 种藻种色素颗粒物吸收系数曲线(1、2、3 分别表示实验配比的 3 个浓度梯度)

Fig.10 Absorption coefficient curve of pigment particles of five algae species
(1, 2 and 3 represent three concentration gradients of the experimental ratio, respectively)

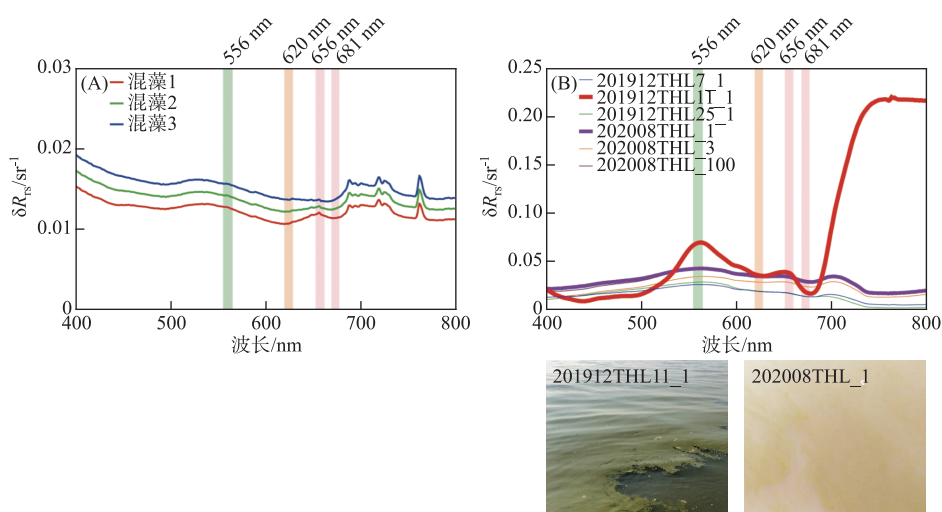


图 11 混藻及野外特征点位遥感反射率

Fig.11 Remote sensing reflectance of mixed algae and field feature points

5 参考文献

- [1] Wu QL, Xie P, Yang LY *et al.* Ecological consequences of cyanobacterial blooms in lakes and their countermeasures. *Advances in Earth Science*, 2008, **23**(11): 1115-1123. [吴庆龙, 谢平, 杨柳燕等. 湖泊蓝藻水华生态灾害形成机理及防治的基础研究. 地球科学进展, 2008, **23**(11): 1115-1123.]
- [2] 张军毅, 朱冰川. 太湖(五里湖)蓝藻水华种类及其演替规律研究. 2012 中国环境科学学会学术年会论文集(第二卷). 南宁: 中国环境科学学会, 2012: 5.
- [3] Zhang M, Yu Y, Yang Z *et al.* The distribution of phytoplankton along trophic gradients and its mediation by available light in the pelagic zone of large eutrophic lakes. *Canadian Journal of Fisheries and Aquatic Sciences*, 2012, **69**(12): 1935-1946. DOI: 10.1139/f2012-113.
- [4] 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.
- [5] Qiao XY. Automatic recognition method of microscopic image of harmful algae. *Journal of Shandong University: Engineering Science*, 2016, **46**(3): 1-6. [乔小燕. 赤潮藻显微图像自动识别方法. 山东大学学报: 工学版, 2016, **46**(3): 1-6.]
- [6] 李春丽. 浮游植物吸收光谱分析方法研究[学位论文]. 青岛: 中国海洋大学, 2007.
- [7] 胡序朋. 基于加权非负最小二乘法的近海浮游植物荧光识别测定技术[学位论文]. 青岛: 中国海洋大学, 2008.
- [8] Zapata M, Rodríguez F, Garrido JL. Separation of chlorophylls and carotenoids from marine phytoplankton: A new HPLC method using a reversed phase C8 column and pyridine-containing mobile phases. *Marine Ecology Progress Series*, 2000, **195**: 29-45. DOI: 10.3354/meps195029.
- [9] Gao YH, Yang JX, Luo QQ *et al.* Automatic identification and analysis techniques of marine phytoplankton. *Journal of Xiamen University: Natural Science*, 2006, **45**(S2): 40-45. [高亚辉, 杨军霞, 骆巧琦等. 海洋浮游植物自动分析和识别技术. 厦门大学学报: 自然科学版, 2006, **45**(S2): 40-45.]
- [10] Chu Q, Zhang YX, Zhang YC *et al.* Quantitative identification methods of bloom-forming cyanobacterial groups of inland lakes based on inherent optical properties. *J Lake Sci*, 2021, **33**(1): 74-85. DOI: 10.18307/2021.0122. [褚乔, 张壹萱, 张玉超等. 基于水华蓝藻固有光学特性的主要类群定量识别方法. 湖泊科学, 2021, **33**(1): 74-85.]
- [11] Mahoney KL. Backscattering of light by *Karenia brevis* and implications for optical detection and monitoring [Dissertation]. Mississippi: The University of Southern Mississippi, 2003.
- [12] Huang CC, Li YM, Sun DY *et al.* Identification of algae using photosensitive pigment absorption signals in Taihu Lake. *Acta Hydrobiologica Sinica*, 2010, **34**(3): 582-590. [黄昌春, 李云梅, 孙德勇等. 基于感光色素吸收信号的太湖藻类识别. 水生生物学报, 2010, **34**(3): 582-590.]
- [13] Zhao QH, Qin BQ. Spectral absorption characteristics of algae and discrimination of the absorption spectrum of mixed algae. *Acta Scientiae Circumstantiae*, 2008, **28**(2): 313-318. DOI: 10.13671/j.hjkxxb.2008.02.012. [赵巧华, 秦伯强. 藻类的光谱吸收特征及其混合藻吸收系数的分离. 环境科学学报, 2008, **28**(2): 313-318.]
- [14] 戴红亮. 淡水湖泊主要藻种的散射特性和理论模拟研究[学位论文]. 南京: 南京师范大学, 2013.
- [15] Dai HL, Lv H, Li YM *et al.* A theoretical modeling of light absorption and scattering properties about *Microcystis aeruginosa*. *Spectroscopy and Spectral Analysis*, 2013, **33**(6): 1462-1467. [戴红亮, 吕恒, 李云梅等. 铜绿微囊藻吸收和散射特性理论模拟. 光谱学与光谱分析, 2013, **33**(6): 1462-1467.]
- [16] 王瑜. 基于吸收光谱的太湖铜绿微囊藻比例估算研究[学位论文]. 南京: 南京师范大学, 2014.
- [17] Xi HY, Losa SN, Mangin A *et al.* Global retrieval of phytoplankton functional types based on empirical orthogonal functions using CMEMS GlobColour merged products and further extension to OLCI data. *Remote Sensing of Environment*, 2020, **240**: 111704. DOI: 10.1016/j.rse.2020.111704.
- [18] Sun YX, Youn SH, Kim Y *et al.* Interannual variation in phytoplankton community driven by environmental factors in the northern East China Sea. *Frontiers in Marine Science*, 2022, **9**: 769497. DOI: 10.3389/fmars.2022.769497.
- [19] Cui TW, Zhang J, Ma Y *et al.* In situ spectra identification of dominating red tide species. *Oceanologia et Limnologia Sinica*, 2005, **36**(3): 277-283. [崔廷伟, 张杰, 马毅等. 基于地物光谱的赤潮优势种识别研究. 海洋与湖沼, 2005, **36**(3): 277-283.]
- [20] Xi H, Hieronymi M, Krasemann H *et al.* Phytoplankton group identification using simulated and in situ hyperspectral remote sensing reflectance. *Frontiers in Marine Science*, 2017, **4**: 272.
- [21] Shen F, Tang RG, Sun XR *et al.* Simple methods for satellite identification of algal blooms and species using 10-year time series data from the East China Sea. *Remote Sensing of Environment*, 2019, **235**: 111484. DOI: 10.1016/j.rse.2019.111484.
- [22] Zhou Y, Shi JH, Qian W *et al.* Distribution of algae and its toxicity to fish in Taihu Lake. *Jiangsu Agricultural Sciences*, 2013, **41**(4): 311-313. [周毅, 石建华, 钱伟等. 太湖藻类分布及其对鱼类的毒性试验. 江苏农业科学, 2013, **41**(4): 311-313.]
- [23] Metsamaa L, Kutser T, Strömbeck N. Recognising cyanobacterial blooms based on their optical signature: A modelling study. *Boreal Envi-*

- ronment Research, 2006, **11**: 493-506.
- [24] Cimoli E, Lucieer V, Meiners KM *et al.* Mapping the *in situ* microspatial distribution of ice algal biomass through hyperspectral imaging of sea-ice cores. *Scientific Reports*, 2020, **10**: 21848. DOI: 10.1038/s41598-020-79084-6.
- [25] Nam G, Shin H, Ha R *et al.* Quantification of phycocyanin in inland waters through remote measurement of ratios and shifts in reflection spectral peaks. *Remote Sensing*, 2021, **13**(16): 3335. DOI: 10.3390/rs13163335.
- [26] Tassan S, Ferrari GM. An alternative approach to absorption measurements of aquatic particles retained on filters. *Limnology and Oceanography*, 1995, **40**(8): 1358-1368. DOI: 10.4319/lo.1995.40.8.1358.
- [27] Bricaud A, Bédhomme AL, Morel A. Optical properties of diverse phytoplanktonic species: Experimental results and theoretical interpretation. *Journal of Plankton Research*, 1988, **10**(5): 851-873. DOI: 10.1093/plankt/10.5.851.
- [28] Gitelson A, Szilagyi F, Mitenzwey KH. Improving quantitative remote sensing for monitoring of inland water quality. *Water Research*, 1993, **27**(7): 1185-1194. DOI: 10.1016/0043-1354(93)90010-F.
- [29] Yue CF, Li YM, Zha Y *et al.* Study on partitioning of phytoplankton absorption coefficient from total spectral absorption in Taihu Lake. *Acta Optica Sinica*, 2010, **30**(11): 3090-3096. [乐成峰, 李云梅, 查勇等. 太湖水体浮游藻类吸收系数分离方法研究. 光学学报, 2010, **30**(11): 3090-3096.]
- [30] Bricaud A, Claustre H, Ras J *et al.* Natural variability of phytoplanktonic absorption in oceanic waters: Influence of the size structure of algal populations. *Journal of Geophysical Research: Oceans*, 2004, **109**(C11): C11010. DOI: 10.1029/2004JC002419.
- [31] Zhou L, Wang MW, Gao JS *et al.* Effects of nitrogen and phosphorus concentrations on growth and photosynthetic physiology of *Pseudanabaena* sp. *Environmental Engineering*, 2020, **38**(11): 110-116, 134. DOI: 10.13205/j.hjgc.202011018. [周利, 王茂伟, 高静思等. 氮、磷浓度对伪鱼腥藻生长及光合生理的影响. 环境工程, 2020, **38**(11): 110-116, 134.]
- [32] Brennan RL, Prediger DJ. Coefficient kappa: Some uses, misuses, and alternatives. *Educational and Psychological Measurement*, 1981, **41**(3): 687-699. DOI: 10.1177/001316448104100307.
- [33] Duan HT, Loisel SA, Zhu L *et al.* Distribution and incidence of algal blooms in Lake Taihu. *Aquatic Sciences*, 2015, **77**(1): 9-16. DOI: 10.1007/s00027-014-0367-2.
- [34] Dembitsky VM, Rozentsvet OA, Pechenkina EE. Glycolipids, phospholipids and fatty acids of brown algae species. *Phytochemistry*, 1990, **29**(11): 3417-3421. DOI: 10.1016/0031-9422(90)85249-F.
- [35] Yin L, Tang JW, Song QJ. A method of classification for algal species based on remote sensing reflectance spectra. *Acta Oceanologica Sinica*, 2011, **33**(3): 55-62. [殷蕾, 唐军武, 宋庆君. 基于遥感反射比光谱的一种藻类识别方法. 海洋学报: 中文版, 2011, **33**(3): 55-62.]