

卫星像元尺度的湖泊浮叶/挺水植被覆盖度估算模型及应用*

秦海涛^{1,2,3}, 罗菊花^{2,3**}, 徐颖^{2,3}, 张春雨^{1,2,3}, 徐亚田², 孟迪⁴, 何锋⁴, 鲁露⁴

(1:南京信息工程大学, 南京 210044)

(2:中国科学院南京地理与湖泊研究所湖泊与流域水安全重点实验室, 南京 210008)

(3:中国科学院大学, 南京 211135)

(4:云南省昆明市滇池高原湖泊研究院, 昆明 650228)

摘要: 浮叶/挺水植被是重要的湖泊水生植被类群, 其面积/覆盖度是湖泊生态健康评估及固碳潜力核算的重要参数, 大面积、精确获取湖泊中浮叶/挺水植被面积/覆盖度及其变化信息, 对湖泊生态修复及碳汇核算至关重要。卫星遥感是获取湖泊浮叶/挺水植被面积/覆盖度最有效的手段。然而, 传统的卫星监测方法只能获取卫星像元内是否存在水生植被, 无法定量估算像元内植被的覆盖度, 从而无法定量、精准地获取湖泊中浮叶/挺水植被的面积/覆盖度。围绕该问题, 本研究利用无人机、Sentinel-2 MSI和Landsat 8 OLI遥感数据, 基于XGBoost建模方法, 通过逐步升尺度的思路, 分别构建了基于Sentinel-2 MSI和Landsat 8 OLI像元尺度的浮叶/挺水植被覆盖度定量估算模型, 并成功地应用于四大淡水湖泊。结果表明: 基于Sentinel和Landsat的估算模型测试集 R^2 分别是0.95和0.97, RMSE分别是7.85%和4.80%, MAE分别是5.35%和3.35%。1990-2022年, 鄱阳湖和洞庭湖呈显著的增加趋势($p < 0.01$), 太湖呈先增后减的趋势($p < 0.01$), 洪泽湖有不显著的增加趋势($p = 0.59$)。模型在四大淡水湖的长时序应用, 证明了模型的稳健性和应用潜力, 预期能为湖泊生态系统碳汇核算和固碳潜力评估提供方法和数据支撑。

关键词: 湖泊; 浮叶/挺水植被; 覆盖度; 卫星像元尺度; 遥感

Estimation model and application of satellite pixel-scale floating/emergent aquatic vegetation coverage in lakes

Haitao Qin^{1,2,3}, Juhua Luo^{2,3**}, Ying Xu^{2,3}, Chunyu Zhang^{1,2,3}, Yatian Xu², Di Meng⁴, Feng He⁴, Lu Lu⁴

(1: Nanjing University of Information Science & Technology, Nanjing 210044, China)

(2: Key Laboratory of Lake and Watershed Science for Water Security, Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China)

(3: University of Chinese Academy of Sciences, Nanjing 211135, China)

(4: Kunming Dianchi and Plateau Lakes Institute, Kunming 650228)

Abstract: Floating/emergent aquatic vegetation is an important aquatic vegetation group in lakes, and its area/coverage is an important parameter for lake ecological health assessment and carbon sequestration potential accounting. Accurately obtaining the area/coverage of floating/emergent aquatic vegetation over large lake areas and understanding their changes is crucial for lake ecological restoration and carbon sink accounting. Satellite remote sensing is the most effective means to obtain the area/cover of floating/emergent aquatic vegetation in lakes. However, traditional satellite monitoring methods can only obtain the presence or absence of aquatic vegetation within satellite pixels, and cannot quantitatively estimate the coverage of aquatic vegetation in the pixels. Consequently, it is impossible to quantitatively and accurately obtain the area/coverage of floating/emergent aquatic vegetation in lakes. To address this issue, we utilized UAV, Sentinel-2 MSI, and Landsat 8 OLI remote sensing data. Using the XGBoost modeling method, we developed quantitative estimation models for floating and emergent aquatic vegetation coverage at the Sentinel-2 MSI and Landsat 8 OLI pixel scales through a stepwise upscaling approach, and successfully applied it to China's four largest freshwater lakes. The results showed that the test sets of the two estimation models based on Sentinel and Landsat images had R^2 of 0.95 and 0.97, RMSE of 7.85% and 4.80%, and MAE of 5.35% and 3.35%, respectively. From 1990 to 2022, Lake Dongting and Lake Poyang showed highly significant increasing trends ($p < 0.01$), Lake Taihu showed an increasing and then decreasing trend ($p < 0.01$), and Lake Hongze had a non-significant increasing trend ($p = 0.59$). The long-term application of the model in the four largest freshwater lakes proved the robustness and application potential of the model, which is expected to provide methodological and data support for the accounting of carbon sinks in lake ecosystems and the assessment of carbon sequestration potential.

Keywords: Lakes; Floating/emergent aquatic vegetation; Coverage; Satellite pixel scale; Remote sensing

在湖泊生态系统中, 水生植被扮演着至关重要的角色。其中, 浮叶/挺水植被由于其较高的生物量, 既是湖泊中的重要碳汇^[1], 又在秋季大量腐烂死亡时可能转化为碳源^[2,3], 对湖泊的生态稳态产生深远影响。

* 2024-09-18 收稿; 2025-01-15 收修改稿。

国家自然科学基金项目(42271377); 云南省省市一体化专项(202202AH210006)联合资助。

** 通信作者; E-mail:jhluo@niglas.ac.cn

因此,浮叶/挺水植被的面积和覆盖度是评估湖泊生态系统健康状况的关键生态指标^[4,5]。首先,适度覆盖度的浮叶/挺水植被能够抑制藻类生长、减缓水流速度、稳定沉积物并降低再悬浮,从而提高水体透明度^[6-8]。然而,过高覆盖度的浮叶/挺水植被可能会引起生态负效应^[9],尤其是浮叶植被,其浓密的叶片会阻挡阳光,降低水下光照条件和水体溶解氧,从而对水下的水生植被和其他生物生长产生不利影响,导致湖泊生物多样性减少^[10-12]。因此,精确监测湖泊中浮叶/挺水植被的面积/覆盖度,对评估水生生态系统稳定状态以及制定植被恢复和管理策略具有重要意义。

卫星遥感技术因其大范围、长时序和低成本等优势而成为浮叶/挺水植被监测的重要手段^[13]。浮叶植被和挺水植被位于水面以上,具有相似的植被光谱特征,难以直接区分。因此,在光学遥感分类应用时,大多将二者归为一类,即浮叶/挺水植被^[14]。由于该类植被具有典型的植被光谱特征,因此,大多对植被敏感的植被指数都可用于提取该类植被,包括归一化差异植被指数(normalized difference vegetation index, NDVI)、增强植被指数(enhanced vegetation index, EVI)、增强浮叶/挺水植被指数(enhanced floating/emergent aquatic vegetation index, EAVI)和归一化差异浮叶/挺水植被指数(Normalized Difference Floating/emergent aquatic vegetation Index, NDAVI)^[15-18]。利用这些植被指数和相应的阈值,可以在卫星影像上识别出含有浮叶/挺水植被的像元。目前,覆盖度的计算仍然依赖传统的分类方法,这种方法通过判断像元内是否存在水生植被,获取覆盖有水生植被的像元,再基于像元数量和像元面积得到水生植被面积,进而计算得到水生植被覆盖度。因此,这种方法获取植被像元只代表该像元存在一定覆盖度的水生植被,但无法定量获取该像元内植被的准确覆盖度是多少,无法精确获取湖泊中水生植被的面积和覆盖度。为了能精确估算像元内的湖泊浮叶/挺水植被覆盖比例,进而得到精确的植被面积,最好的方法即构建卫星像元尺度的覆盖度估算模型。首先,Hu等(2019)通过使用MODIS数据结合光谱阈值法在像元内估算了浒苔的面积覆盖和生物量,证明了利用遥感数据中的光谱信息进行水生植物覆盖度估算的可行性^[19]。此外,已有研究证明了构建覆盖度定量估算模型的有效性,即通过少量实地获取的小样方估计的覆盖度与高空间分辨率卫星像元(50cm)匹配,并在小区域的河流段落中利用机器学习方法估算沉水植被覆盖度^[20]。然而,目前仍缺乏关于构建更稳健覆盖度定量估算模型的研究,主要挑战在于:在湖泊等水域中,由于立地条件复杂,获取大量与卫星像元大小相匹配的地面覆盖度样本十分困难,尤其对中高空间分辨率卫星(10 m或30 m)建模。无人机(unmanned aerial vehicle, UAV)遥感作为地面调查和卫星遥感之间的桥梁,具备高灵活性、低成本和高空间分辨率等优势,已逐渐取代了人工采集实地数据^[21],已被广泛应用于复杂立地条件下的森林、湿地的近地样方数据获取和调查^[22],为湖泊中的浮叶/挺水植被的近地大样方调查和覆盖度数据获取提供了良好的契机。

因此,本研究基于无人机、Sentinel-2 MSI和Landsat 8 OLI数据,采用逐步升尺度的方法,分别构建了无人机-Sentinel-2 MSI和Sentinel-2 MSI-Landsat 8 OLI卫星像元尺度的浮叶/挺水植被覆盖度估算模型,并成功应用于中国四大淡水湖的浮叶/挺水植被面积的长时序(1990-2022年)监测。通过定量估算模型精细化估算浮叶/挺水植被的覆盖度和面积,并展示了其在大范围应用的潜力。未来,该模型可为湖泊碳汇核算和固碳潜力评估提供方法和数据支撑。

1 研究区与方法

1.1 研究区概况

太湖、鄱阳湖、洞庭湖和洪泽湖是位于中国东部江淮流域(27.78° -36.10° N, 111.70° -121.67° E)的四大淡水湖(图1(a-e))。该地区属于亚热带季风气候区,水资源丰富。江淮流域的湖泊平均深度在1.1至8.4 m之间,其中太湖的平均水深为1.9 m,鄱阳湖平均水深8.4 m,洞庭湖平均水深6.4 m,洪泽湖平均水深1.77 m,并且四大淡水湖均分布着多种浮叶/挺水植被群落^[23],例如菱角、荇菜、菰、芦苇等,具体如表1。

本研究选择太湖(30°55'40"-31°32'58"N, 119°52'32"-120°36'10"E)作为模型开发的实验湖泊。其东部湖区是典型的草型区,水生植被群落丰富;而西部以藻类为主,全年藻华频繁^[24,25]。尤其是东太湖湾拥有丰富的浮叶/挺水植被群落(图1(b1)),本研究在东太湖湾的三个不同时期和不同区域进行了三次无人机飞行实验,具体见图1A~C。

1.2 数据获取及预处理

1.2.1 无人机数据 无人机飞行实验在东太湖湾进行(图1(A-C)),飞行时间分别为2022年6月22日、2023年6月16日和2023年8月6日(表2)。无人机型号为DJI Phantom 4 RTK,配备一台高分辨率的可见光相机,能够捕捉RGB(红、绿、蓝)图像,手动设置飞行速度为8m/s,飞行高度为120m,旁向重叠率70%,航向重叠率80%。利用Pix4d软件处理无人机数据,包括图像几何校正、配准、融合等,获得无人机飞行区域的正射影像,空间分辨率为0.03 m。

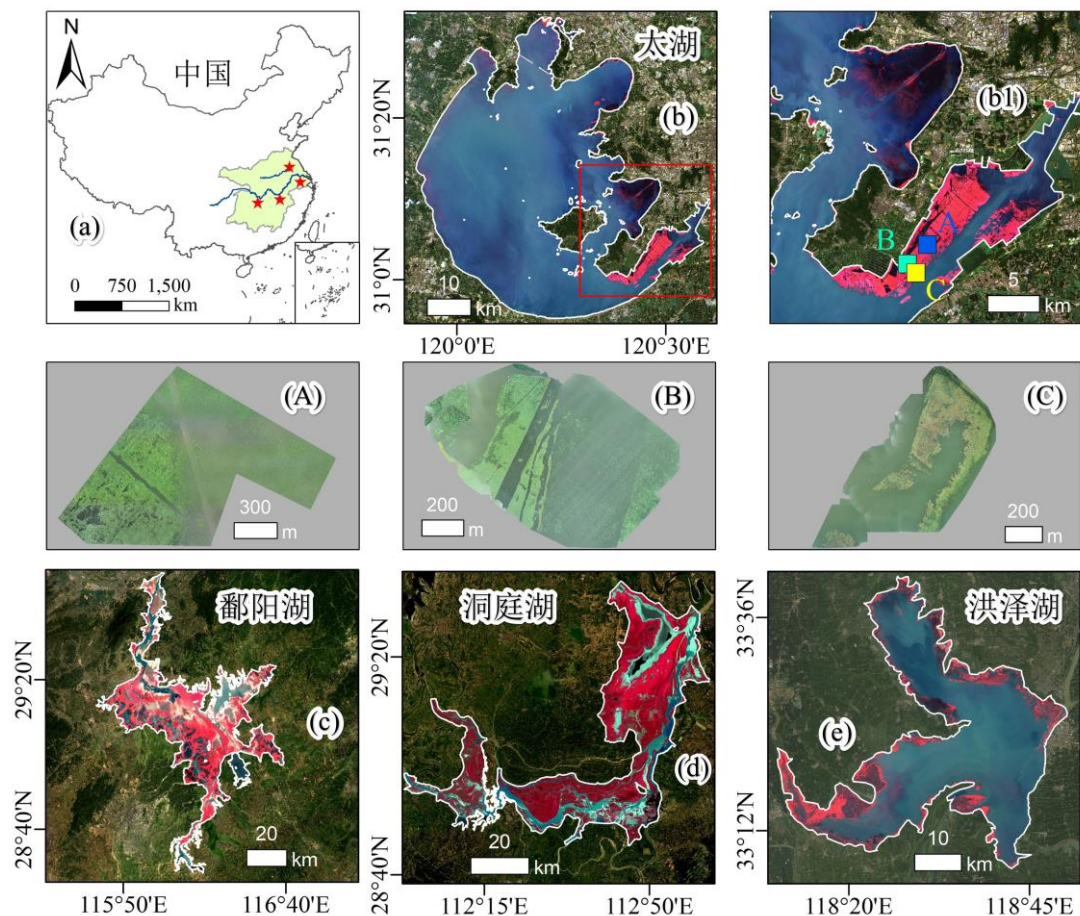


图 1 四大淡水湖的位置 (a); 太湖 (b)、鄱阳湖 (c)、洞庭湖 (d) 和洪泽湖 (e) 的 Landsat 假彩色影像与 Landsat 真彩色背景影像; 东太湖湾 (b1) 的三景无人机真彩色图像 (A-C)。

Fig.1 Location of four major freshwater lakes (a); False RGB Landsat images in Lake Taihu (b), Lake Poyang (c), and Lake Dongting (d), Lake Hongze (e) with true RGB Landsat background images; Three true RGB unmanned aerial vehicle (UAV) images in East Taihu Bay (A-C).

表 1 四大淡水湖浮叶/挺水植被优势种群落

Tab.1 Dominant species communities of floating/emergent aquatic vegetation in four major freshwater lakes

湖名	浮叶植被优势种群落	挺水植被优势种群落	文献
太湖	荇菜、菱角等	菰、芦苇、香蒲、莲等	Luo 等 (2017)
鄱阳湖	荇菜、芡实、菱角、水鳖等	菰、芦苇、南荻等	胡振鹏 等 (2019)
洞庭湖	野菱等	芦苇、南荻、苔草等	余蓉 等 (2016)
洪泽湖	荇菜等	芦苇、菰等	刘伟龙 等 (2009)

1.2.2 卫星数据 本研究使用 Sentinel-2 MSI 地表反射率数据集 (COPERNICUS/S2_SR)、Landsat-5 TM 地表反射率数据集 (LANDSAT/LT05/C02/T1_L2) (数据可用时间: 1990-2011) 和 Landsat-8 OLI 地表反射率数据集 (LANDSAT/LT08/C02/T1_L2) (数据可用时间: 2013-2022)。上述所有卫星数据集均通过 Google Earth Engine (GEE) 平台进行访问、获取和处理^[26]。这些数据集均已进行预处理, 包括辐射定标和大气校正。研究过程中将空间分辨率 20 m 的波段重采样为 10 m。云掩膜是分别利用 Quality Assurance 60 (QA60) 波段和 Quality Assurance Pixel (QA_PIXEL) 波段完成。此外, 本研究还将藻华区域分类并掩膜, 以确保分析的准确性^[15]。

在星地同步实验区-太湖, 本研究选择了与无人机近同步的高质量 Sentinel-2 MSI 卫星影像, 日期分别为 2022 年 6 月 27 日、2023 年 6 月 12 日和 2023 年 8 月 6 日。在升尺度构建模型阶段, 为了保证不同数

据源的获取时间在 5 天之内且均具备高质量影像,研究选用了 2020 年 9 月 8 日的 Landsat 8 OLI 卫星影像,以及近同步的 2020 年 9 月 5 日的 Sentinel-2 MSI 卫星影像(表 2)。此外,在四大淡水湖泊的应用中,采用了 1990-2022 年间 6-9 月份的 Landsat 系列卫星影像。

表 2 用于匹配建模的影像数据源及其获取时间表

湖泊	获取时间		
	无人机	Sentinel 2 MSI	Landsat 8 OLI
太湖	2022 年 6 月 22 日	2022 年 6 月 27 日	/
	2023 年 6 月 16 日	2023 年 6 月 12 日	/
	2023 年 8 月 6 日	2023 年 8 月 6 日	/
	/	2020 年 9 月 5 日	2020 年 9 月 8 日

1.3 研究方法

本论文提出的建模流程如图 2 所示,主要是通过两次升尺度(UAV-Sentinel 2 MSI 和 Sentinel 2 MSI-Landsat 8 OLI),分别构建基于 Sentinel 2 MSI 和 Landsat 8 OLI 卫星像元的浮叶/挺水植被覆盖度估算模型。

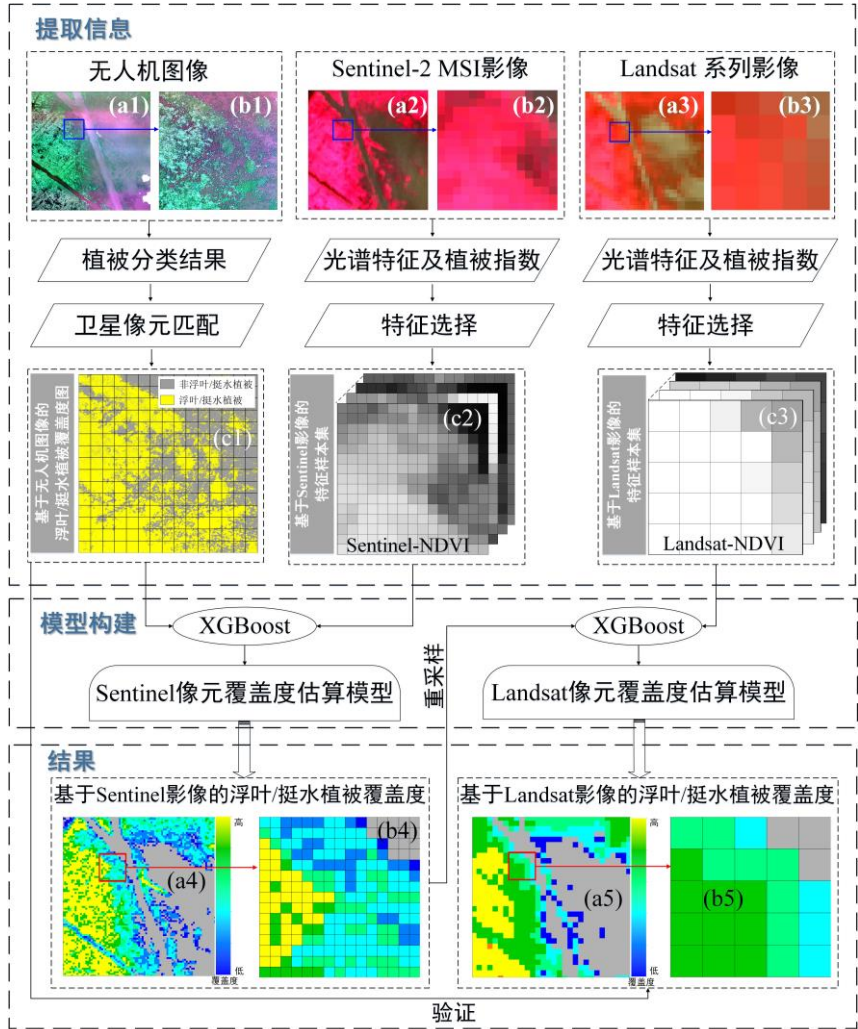


图 2 逐步升尺度构建像元尺度浮叶/挺水植被覆盖度模型的工作流程图

Fig. 2 Workflow for developing models for estimating pixel-scale floating/emergent aquatic vegetation coverage by stepwise upscaling method

1.3.1 无人机提取浮叶/挺水植被覆盖度方法 可见光波段差异植被指数 (visible-band difference vegetation index, $VDVI$) 已广泛应用于无人机获取的数字正射影像中, 以提取植被信息, 并证明其具有高精度和良好的适用性^[27]。本研究基于获得的高空间分辨率无人机图像 (3 cm) 和 $VDVI$ 直方图分布, 采用大津法 (OTSU) 结合目视解译选取合适的水生植被阈值^[28]。随后, 依据选取的阈值对 $VDVI$ 图像进行二值化处理, 将大于阈值的像元划分为浮叶/挺水植被类别, 小于或等于阈值的像元划分为非浮叶/挺水植被类别。接下来, 使用 QGIS Desktop 软件, 基于无人机图像提取的浮叶/挺水植被结果, 匹配 Sentinel-2 MSI 卫星像元大小 (10m×10m) 并进行切割, 计算每个像元中浮叶/挺水植被的比例。此外, 本研究对每个像元的解译结果进行了详细的目视评估, 并剔除了可能存在较大误差的样本, 特别是受到条带效应影响导致分类结果不确定的像元。最终, 生成浮叶/挺水植被覆盖度样本集 (图 2 (c1))。

$$VDVI = \frac{2 \times \rho_{Green} - \rho_{Red} - \rho_{Blue}}{2 \times \rho_{Green} + \rho_{Red} + \rho_{Blue}} \quad (1)$$

式中: ρ_{Red} 、 ρ_{Green} 、 ρ_{Blue} 分别为无人机图像红、绿、蓝波段的反射率值。

1.3.2 Sentinel 像元浮叶/挺水植被覆盖度估算模型构建 通过文献搜索筛选与浮叶/挺水植被敏感的卫星特征^[13], 本研究选择了 14 个 Sentinel-2 MSI 遥感特征变量, 包括光谱波段和植被指数 (表 3)。研究使用 Python 语言的 “scikit-learn” 库, 采用递归特征消除算法 (Recursive Feature Elimination, RFE) 对特征变量筛选。首先, 根据重要性得分对特征变量进行排序, 删除得分最低的若干个特征, 重复以上步骤直至找到最优的特征变量组合。将选择的卫星特征样本匹配无人机的覆盖度样本 (图 2 (c1) 和 (c2)), 最终得到输入模型的总样本集, 分为 70% 训练集和 30% 测试集。其中, 无人机图像解译的覆盖度数据在 2023 年 8 月 6 日有 1117 个样本, 2023 年 6 月 16 日有 890 个样本, 2022 年 6 月 22 日有 2002 个样本, 共计 4009 个样本。

本研究选择逐步多元线性回归 (multiple linear regression, MLR)^[29]、支持向量回归 (support vector regression, SVR)^[30] 与极端梯度提升 (extreme gradient boosting, XGBoost)^[31] 的机器学习方法来构建模型。利用 Python 语言中的 “scikit-learn” 库进行模型构建、网格搜索以找到最优参数组合, 并进行交叉验证以评估模型性能。

1.3.3 Landsat 像元浮叶/挺水植被覆盖度估算模型构建 本研究选取了另外两景近同步且无云的 Sentinel-2 MSI 和 Landsat 8 OLI 卫星影像, 分别为 2020 年 9 月 5 日的 Sentinel-2 MSI 影像和 2020 年 9 月 8 日的 Landsat 8 OLI 影像, 利用 Sentinel 像元浮叶/挺水植被覆盖度估算模型, 生成 Sentinel-2 MSI 像元尺度的浮叶/挺水植被覆盖度图 (图 2 (b4))。接下来, 研究对比了三种常见的重采样方法: 最邻近插值法 (nearest neighbor interpolation)、双线性插值法 (bilinear interpolation) 和三次卷积插值法 (cubic convolution interpolation)。研究使用了 ArcMap 软件中的重采样工具, 将 Sentinel 像元尺度的浮叶/挺水植被覆盖度图重采样至 Landsat 卫星像元大小 (30 m×30 m), 以匹配 Landsat 8 OLI 数据的空间分辨率。为评估不同重采样方法对覆盖度估算结果的影响, 本研究通过可视化效果和模型性能指标 (R^2 、RMSE 和 MAE) 进行对比分析 (附图 I 和附表 I)。结果表明, 不同重采样方法对结果的影响较小, 但双线性插值法在模型精度方面表现最佳。因此, 最终采用双线性插值法重采样的新像元值作为每个 Landsat 8 OLI 卫星像元的实测覆盖度, 从而构建了 Landsat 卫星像元尺度的浮叶/挺水植被覆盖度样本集。

特征选择过程中, 同样使用递归特征消除算法提取并筛选最优 Landsat 8 OLI 卫星特征变量组合, 将 Landsat 8 OLI 卫星特征样本 (图 2 (c3)) 与重采样后的覆盖度样本匹配, 得到输入模型的总样本集, 共有 11146 个样本, 样本集被分为 70% 训练集和 30% 测试集。研究继续使用性能表现最佳的 XGBoost 机器学习算法构建模型。同时, 本研究也使用同步无人机图像生成的浮叶/挺水植被覆盖度 (图 2 (c1)) 来验证 Landsat 像元浮叶/挺水植被覆盖度估算模型反演的像元尺度浮叶/挺水植被覆盖度结果 (图 2 (a5))。

1.3.4 性能评估指标 模型精度通过以下指标来评价: 决定系数 (coefficient of determination, R^2)、均方根误差 (root mean square error, RMSE) 和平均绝对误差 (mean absolute error, MAE)。 R^2 值越高表示模型性能越好, 而 RMSE 和 MAE 值越小表示精度越高。

$$R^2 = \frac{(\sum (x_i - \bar{x})(y_i - \bar{y}))^2}{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2} \quad (2)$$

表 3 Sentinel 2 MSI 卫星影像的特征变量
Table 3 Feature variables of Sentinel 2 MSI satellite image

特征变量	变量名称	描述
光谱特征	Band 2 - Blue	490nm - (10m)
	Band 3 - Green	560nm - (10m)
	Band 4 - Red	665nm - (10m)
	Band 5 - Red Edge	705nm - (20m)
	Band 6 - Red Edge	740nm - (20m)
	Band 7 - Red Edge	783nm - (20m)
	Band 8 - NIR	842nm - (10m)
		$(\rho_{NIR} - \rho_{Red})/(\rho_{NIR} + \rho_{Red})$
植被指数	NDVI	
	GNDVI	$(\rho_{NIR} - \rho_{Green})/(\rho_{NIR} + \rho_{Green})$
	NDAVI	$(\rho_{NIR} - \rho_{Blue})/(\rho_{NIR} + \rho_{Blue})$
	NDWI	$(\rho_{Green} - \rho_{NIR})/(\rho_{Green} + \rho_{NIR})$
	DVI	$\rho_{NIR} - \rho_{Red}$
	EVI	$\frac{2.5 \times (\rho_{NIR} - \rho_{RED})}{\rho_{NIR} + 6 \times \rho_{RED} - 7.5 \times \rho_{BLUE} + 1}$
	RVI	ρ_{NIR}/ρ_{Red}

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - x_i)^2} \quad (3)$$

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - x_i| \quad (4)$$

式中： N 为样本数量； x_i 表示以无人机数据作为实测值的第 i 个覆盖度值（或者表示以精度更高反演结果的重采样参考值的第 i 个覆盖度值）； y_i 表示模型估计的第 i 个覆盖度值； $\bar{x}$ 和 $\bar{y}$ 分别表示被作为实测覆盖

度值的平均值和模型估计覆盖度值的平均值。

2 结果与分析

2.1 基于无人机图像的浮叶/挺水植被提取结果

本研究基于三次不同日期的无人机图像，采用 $VDVI$ 提取浮叶/挺水植被信息，结合目视解译和阈值法进行二值化分类，白色为浮叶/挺水植被，黑色为非浮叶/挺水植被。由于无人机图像具有厘米级的高空间分辨率，结合目视解译和地面调查，基于该方法可以准确提取浮叶/挺水植被（图 3）。

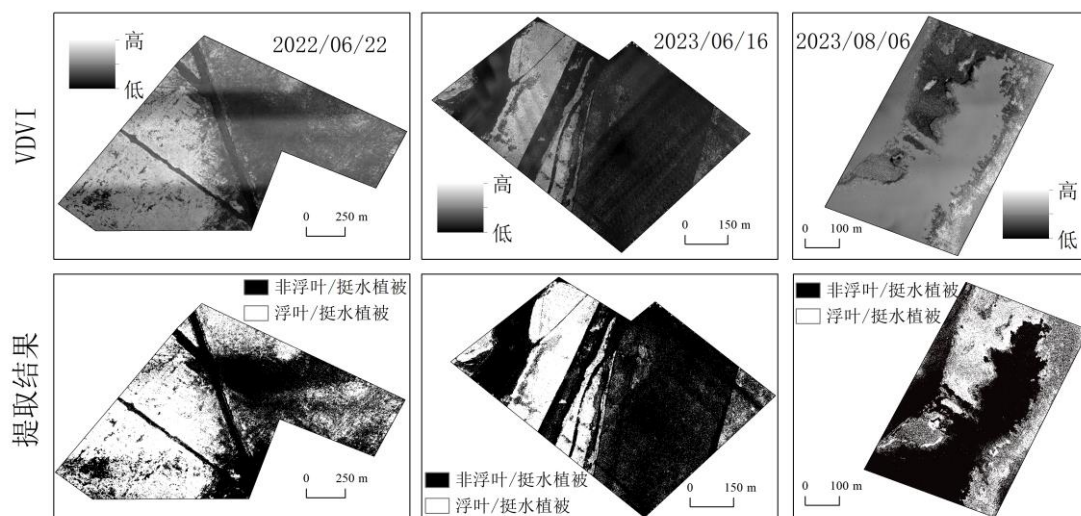


图3 浮叶/挺水植被的提取结果

Fig.3 Extraction results of floating/emergent aquatic vegetation

2.2 Sentinel 像元浮叶/挺水植被覆盖度估算模型与验证

通过递归特征消除算法的特征选择，本研究最终输入模型的特征变量分别是 Band3、Band4、Band8、NDVI、GNDVI、NDAVI、NDWI。分别基于 MLR、SVR 和 XGBoost 机器学习算法构建浮叶/挺水植被覆盖度估算模型，性能评价结果如表 4 所示。XGBoost 模型表现最佳，训练集和测试集的 R^2 分别为 0.98 和 0.95， $RMSE$ 分别为 5.41% 和 7.85%， MAE 分别为 3.75% 和 5.35%（图 4a）。因此，选用 XGBoost 构建的 Sentinel 像元浮叶/挺水植被覆盖度估算模型对整个实验区（太湖）进行浮叶/挺水植被覆盖度反演（图 4（b-d））。结果显示，浮叶/挺水植被主要分布在太湖的东部草型湖区，与实地调查结果相符。模型估算结果成功识别了覆盖度超过 20% 的像元，不同年份的 2022 年 6 月 27 日的覆盖度反演结果略大于 2023 年 6 月 12 日，同年 8 月 6 日的覆盖度反演结果也大于 6 月 12 日，符合物候生长规律。

表 4 不同模型精度评估

Table 4 Accuracy evaluation of different models

影像	方法	训练集			测试集		
		决定系数 R^2	均方根误差 $RMSE$	平均绝对误差 MAE	决定系数 R^2	均方根误差 $RMSE$	平均绝对误差 MAE
Sentinel-2 MSI	MLR	0.93	9.07%	7.08%	0.93	9.18%	6.80%
	SVR	0.92	9.80%	7.47%	0.92	9.82%	7.28%
	XGBoost	0.98	5.41%	3.75%	0.95	7.85%	5.35%
Landsat 8 OLI	MLR	0.96	5.70%	4.35%	0.95	5.86%	4.43%
	SVR	0.94	6.83%	4.90%	0.93	6.91%	5.00%
	XGBoost	0.98	3.90%	2.77%	0.97	4.80%	3.35%

3.3 Landsat 像元浮叶/挺水植被覆盖度估算模型与验证

通过递归特征消除算法完成特征选择，本研究最终输入模型的特征变量分别是 Band3、Band4、Band5、NDVI、GNDVI。利用 XGBoost 机器学习算法构建 Landsat 像元浮叶/挺水植被覆盖度估算模型，性能评价结果如表 4 所示。XGBoost 模型训练集和测试集的 R^2 分别为 0.98 和 0.97， $RMSE$ 分别为 3.90% 和 4.80%， MAE 分别为 2.77% 和 3.35%。在空间上，使用该模型生成的 2020 年 9 月 8 日太湖的浮叶/挺水植被反演结果成功识别了覆盖度超过 30% 的像元。与基于 Sentinel 像元浮叶/挺水植被覆盖度估算模型从 2020 年 9 月 5 日近同步 Sentinel-2 MSI 影像中获取的浮叶/挺水植被覆盖度相比，表现出良好的空间一致性（图 5（a-b））。本研究进行了两个尺度的精度验证。首先，再次将基于 Sentinel-2 MSI 卫星影像的覆盖度数据重采样到 30 m，与 Landsat 8 OLI 卫星像元匹配，保证了覆盖度数据在空间尺度上的一致性，同时避免了覆盖度范围不匹配的问题。结果显示，在低覆盖度时，Landsat 像元浮叶/挺水植被覆盖度估算模型存在高估，在高覆盖度时存在低估，但总体而言，模型性能依然具备一定的准确性（ $R^2 = 0.57$ ， $RMSE = 11.36\%$ ， $MAE = 8.59\%$ ）

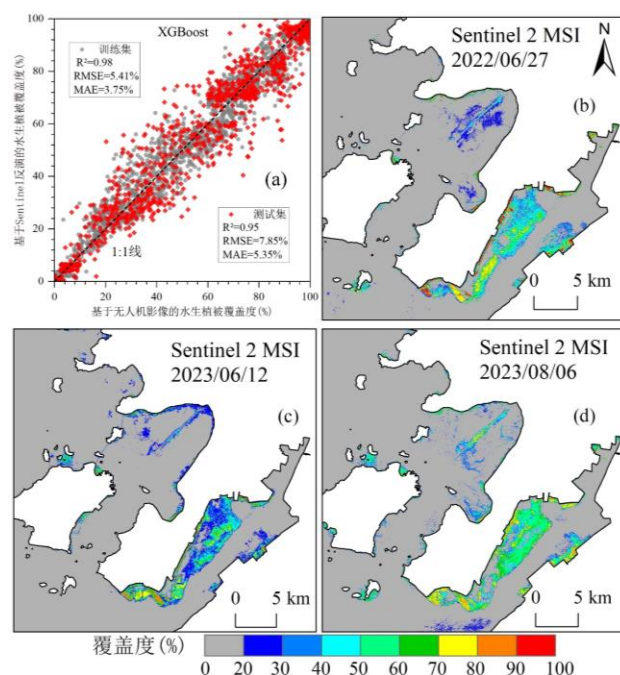


图4 基于 Sentinel-2 MSI 数据和 XGBoost 模型的浮叶/挺水植被覆盖度性能评估 (a)；

Sentinel 像元覆盖度估算模型反演的浮叶/挺水植被覆盖度 (b-d)

Fig.4 Performance evaluation of floating/emergent aquatic vegetation coverage derived from Sentinel-2 MSI data using XGBoost-based model (a); maps of floating/emergent aquatic vegetation coverage derived from the Sentinel-pixel coverage estimation model (b-d).

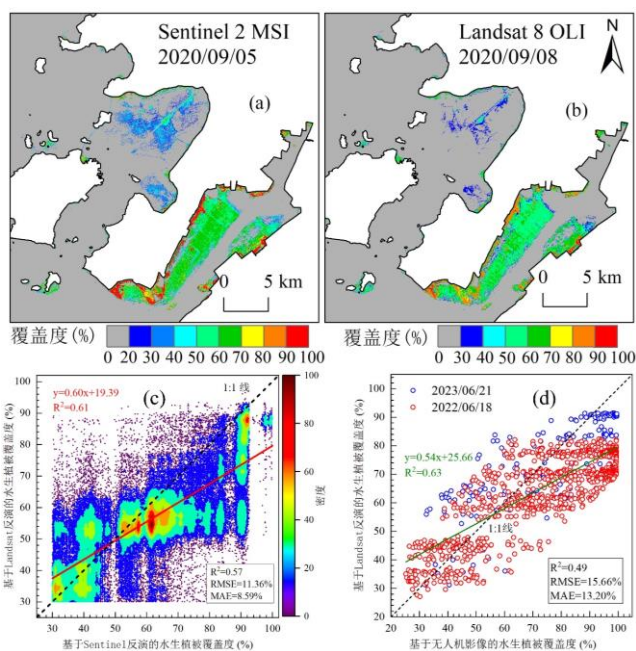


图5 基于 Sentinel 像元覆盖度估算模型和 Landsat 像元覆盖度估算模型反演的浮叶/挺水植被覆盖度 (a-b)；基于太湖所有样本的 Sentinel 像元和 Landsat 像元覆盖度估算模型反演的浮叶/挺水植被覆盖度的验证 (c)；使用无人机的浮叶/挺水植被覆盖度验证 Landsat 像元覆盖度估算模型的浮叶/挺水植被覆盖度 (d)

Fig.5 Floating/emergent aquatic vegetation coverage map derived from Sentinel-pixel coverage estimation model and Landsat-pixel coverage estimation model; validation of floating/emergent aquatic vegetation coverage between Sentinel-pixel and Landsat-pixel coverage estimation models based on all samples in Lake Taihu (c); validation of floating/emergent aquatic vegetation coverage by Landsat-pixel coverage estimation model using UAV floating/emergent aquatic vegetation coverage (d)

(图 5(c))。其次,使用了 2022 年 6 月 18 日和 2023 年 6 月 21 日获得的同步无人机图像进一步验证 Landsat 像元浮叶/挺水植被覆盖度估算模型的估算结果。结果表明,在低覆盖度时, Landsat 像元浮叶/挺水植被覆盖度估算模型存在高估,在高覆盖度 ($> 80\%$) 时存在低估。尽管如此,模型的整体估算性能具有一定的合理性 ($R^2 = 0.49$, $RMSE = 15.66\%$, $MAE = 13.20\%$) (图 5 (d))。

2.3 模型在四大淡水湖泊的应用

基于 Landsat 系列影像,利用 Landsat 像元浮叶/挺水植被覆盖度估算模型,本研究选取江淮流域的四大淡水湖泊作为研究对象,分析其浮叶/挺水植被面积的长时序变化。在 1990 年到 2022 年期间,将每年 6 至

9 月份（丰水期）无云的所有 Landsat 系列影像（Landsat 5 TM 和 Landsat 8 OLI）进行反演，从中选取浮叶/挺水植被面积最大的一景结果来做年际变化分析。尽管如此，依然有部分年份缺少高质量影像，但整体趋势仍具备可靠性。通过定量计算浮叶/挺水植被的面积，发现太湖的浮叶/挺水植被面积从 1992 年的 53.75 km² 增加到 2004 年的 212.23 km²，逐渐减少至 2022 年的 101.75 km² ($R^2 = 0.40$, $p < 0.01$)，整体呈先增后减的趋势。鄱阳湖的浮叶/挺水植被的面积从 1991 年的 714.86 km² 增加到 2022 年的 1651.84 km²，总体上呈显著增加趋势 ($R^2 = 0.27$, $p < 0.01$)。洞庭湖的浮叶/挺水植被的面积从 1990 年的 535.60 km² 增加到 2022 年的 1614.96 km²，总体上呈显著增加趋势 ($R^2 = 0.36$, $p < 0.01$)。洪泽湖的浮叶/挺水植被面积从 1991 年的 141.68 km² 增加到 2022 年的 176.36 km²，呈不显著的增加趋势 ($R^2 = 0.01$, $p < 0.59$) (图 6)。

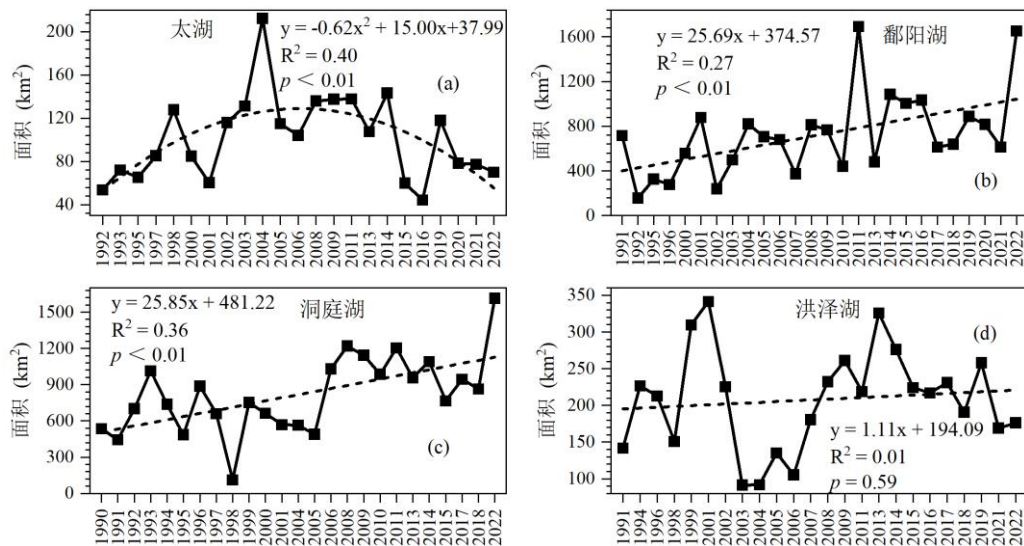


图 6 太湖 (a)、鄱阳湖 (b)、洞庭湖 (c) 和洪泽湖 (d) 的浮叶/挺水植被面积变化趋势
Fig.6 Change trend of floating/emergent aquatic vegetation area in Lake Taihu (a), Lake Poyang (b), Lake Dongting (c), and Lake Hongze (d).

3 讨论

本研究采用了逐步升尺度构建样本集的方法，结合无人机、Sentinel-2 MSI 和 Landsat 8 OLI 的数据，提出了一种像元尺度估算浮叶/挺水植被覆盖度的新思路。无人机作为地面和卫星之间的桥梁，不仅解决了在复杂立地条件下获取地面覆盖度样本的困难，还很好地完成了与卫星像元大小匹配的挑战，促进了模型的成功构建。本研究再次证明了无人机能够替代人类获取高精度和大规模的地面样本数据，特别是在湿地和湖泊等人类难以进入的立地环境中^[22,32,33]。此外，不同遥感数据源相匹配过程中，获取准同步的数据源构建的模型是最理想的实验方案。然而，由于实验环境、条件等限制，在实际实验中获取与卫星准同步无人机数据往往具有一定挑战。本研究的研究对象为浮叶/挺水植被，其物候变化表明，在几天内其生长位置稳定，面积、覆盖度等植被参数变化很小^[34,35]。因此，在未能获取到准同步的匹配数据源时，使用几天内的数据匹配开展覆盖度模型构建是可行的。本研究用于匹配建模的不同数据源获取时间误差控制在 5 天内，对年尺度的浮叶/挺水植被物候变化来说，覆盖度、面积的变化可以被忽略。

虽然本研究构建的像元植被覆盖度模型具有较高的精度和鲁棒性，然而，在将模型应用于富营养化湖泊之前，需要先识别和掩膜藻华爆发区，以避免因为浮叶/挺水植被与藻华的光谱特征相似而导致的植被覆盖度高估。不仅是藻华，少量的长势较好且长到水面的沉水植被可能也会被模型归类为浮叶/挺水植被。在湖泊水生植被监测研究中，这一问题普遍存在，目前传统的分类方法也同样难以解决这种异物同谱的问题。本研究的模型使用了分类方法中相同的植被光谱和指数，因此结果也可能导致浮叶/挺水植被的高估。然而，由于本文的研究湖泊主要由浮叶/挺水植被组成，因此受到的影响较小；但在生长旺盛期以沉水植被为主导的湖泊，模型反演结果预计会受到较大影响。

此外，通过模型测试，本研究发现 Sentinel-2 MSI 影像上植被覆盖度低于 20% 和 Landsat 8 OLI 影像上覆盖度低于 30% 的像元，模型表现出显著的不确定性，因此被排除在研究范围之外。可能是由于在富营养化湖泊中，在低覆盖度区域，一些具有高浓度有机质和叶绿素 a 的水体像元光谱与低覆盖度植被的像元光谱相似，表现出异物同谱的特征，会将这些水体像元误判为植被像元而导致植被覆盖度高估^[36,37]，具体原因还需要进一步的做定点和定量实验。在高覆盖度区域，由于模型中使用的典型特征为红光及 NDVI 等植被指数，这些特征本身在植被高覆盖度时就会出现饱和现象^[38,39]，导致高覆盖度像元的反演结果低估。目前这种低估和高估的问题很难被优化校正，未来可以只针对模型高估和低估的区域，使用更多的光谱指标构建更为精准的模型，即通过覆盖度分段建模解决和优化这一问题。

最后，由于浮叶植被和挺水植被具有相似的植被光谱特征，在光学遥感分类时通常将它们归为一类。本研究中，同样使用光学特征变量输入模型，模型估算的结果是浮叶植被和挺水植被的总覆盖度和面积。然而，考虑到浮叶植被和挺水植被在湖泊生态系统中扮演不同角色且生物量存在差异，进一步准确区分它们具有重要意义。未来的研究可以结合水生植被分类方法，以实现湖泊中浮叶植被和挺水植被的区分，并

分别准确估算浮叶植被和挺水植被的面积。

4 结论

本研究提出了一种升尺度方法,利用无人机、Sentinel-2 MSI 和 Landsat 8 OLI 遥感数据,基于 XGBoost 机器学习算法,分别构建了两个基于 Sentinel-2 MSI 和 Landsat 8 OLI 卫星像元尺度的浮叶/挺水植被覆盖度定量估算模型。

基于 Sentinel 和 Landsat 像元的两个估算模型测试集 R^2 分别是 0.95 和 0.97, $RMSE$ 分别是 7.85% 和 4.80%, MAE 分别是 5.35% 和 3.35%, 模型表现出较高的精度。将构建的模型应用于四大淡水湖泊,结果表明:1990-2022 年,鄱阳湖和洞庭湖呈显著的增加趋势 ($p < 0.01$),太湖呈先增后减的趋势 ($p < 0.01$),洪泽湖有不显著的增加趋势 ($p = 0.59$)。本研究构建的卫星像元尺度浮叶/挺水植被覆盖度遥感定量估算模型,应用在四大淡水湖泊的长时序监测,证明了模型的稳健性,表现出巨大的应用潜力,预期能为湖泊生态系统碳汇核算和固碳潜力评估提供方法和数据支撑。

5 附录

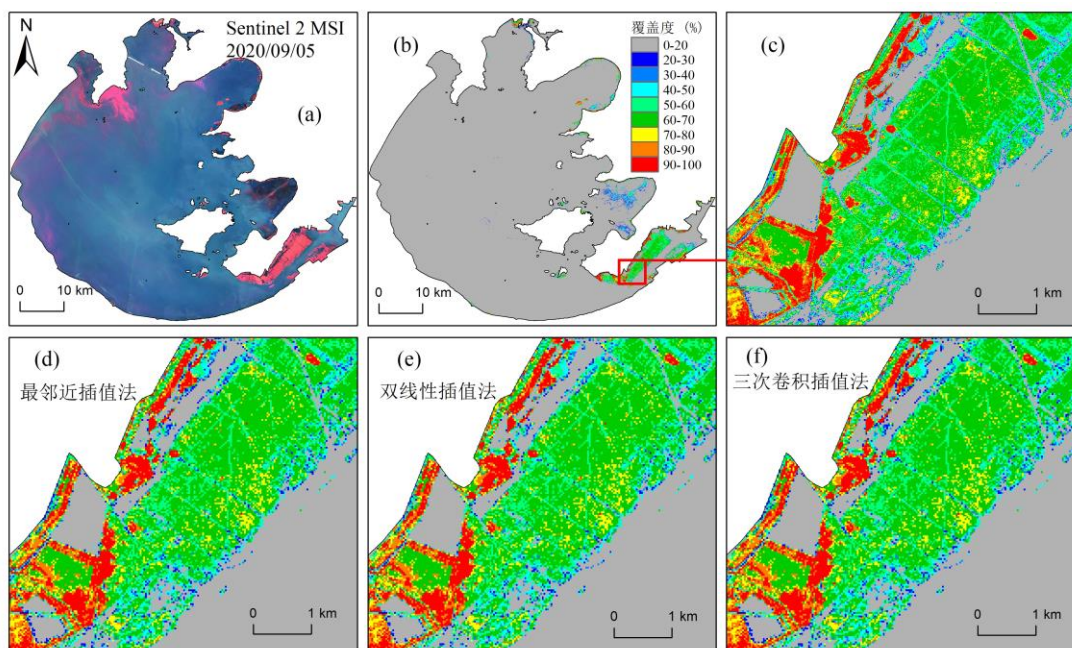
附图附表见 DOI: 10.18307/2025.0633.

6 参考文献

- [1] Peixoto R B, Marotta H, Bastviken D, et al. Floating Aquatic Macrophytes Can Substantially Offset Open Water CO₂ Emissions from Tropical Floodplain Lake Ecosystems. *Ecosystems*, 2016, 19(4): 724-736. DOI: 10.1007/s10021-016-9964-3
- [2] Zhao Z, Zhang D, Shi W, et al. Understanding the Spatial Heterogeneity of CO₂ and CH₄ Fluxes from an Urban Shallow Lake: Correlations with Environmental Factors. *Journal of Chemistry*, 2017, 2017(1): 8175631. DOI: 10.1155/2017/8175631
- [3] Audet J, Carstensen M V, Hoffmann C C, et al. Greenhouse gas emissions from urban ponds in Denmark. *Inland Waters*, 2020, 10(3): 373-385. DOI: 10.1080/20442041.2020.1730680
- [4] Lu L, Luo J, Xin Y, et al. A novel strategy for estimating biomass of submerged aquatic vegetation in lake integrating UAV and Sentinel data. *Science of The Total Environment*, 2023: 169404. DOI: 10.1016/j.scitotenv.2023.169404
- [5] Villa P, Dalla Vecchia A, Piasek E, et al. Assessing PROSPECT performance on aquatic plant leaves. *Remote Sensing of Environment*, 2024, 301: 113926. DOI: 10.1016/j.rse.2023.113926
- [6] Quan W, Shen X, Yan L. Advances in research of biological purification of eutrophic water body. *The Journal of Applied Ecology*, 2003, 14(11): 2057-2061. [全为民, 沈新强, 严力蛟. 富营养化水体生物净化效应的研究进展. 应用生态学报, 2003 (11): 2057.]
- [7] Cazzanelli M, Warming T P, Christoffersen K S. Emergent and floating-leaved macrophytes as refuge for zooplankton in a eutrophic temperate lake without submerged vegetation. *Hydrobiologia*, 2008, 605(1): 113-122. DOI: 10.1007/s10750-008-9324-1
- [8] Ji G, Xu Z X, Wang L Q. Effects of floating-leaved macrophytes on water quality and phytoplankton: an in situ experiment in a Chinese shallow lake. *Desalination and Water Treatment*, 2016: 1-12. DOI: 10.1080/19443994.2016.1180641
- [9] Zhu Y, Zhang S, Zhao S, et al. Analysis of shallow water lake bogging and causes in Lake Wuliangsu. *Water Resources Protection*, 2017, 33(5): 69-74. [朱永华, 张生, 赵胜男, 等. 乌梁素海浅水湖泊沼泽化现状及成因分析. 水资源保护, 2017, 33(5): 69-74. DOI: 10.3880/j.issn.1004-6933.2017.05.012]
- [10] Dube T, Mutanga O, Sibanda M, et al. Testing the detection and discrimination potential of the new Landsat 8 satellite data on the challenging water hyacinth (*Eichhornia crassipes*) in freshwater ecosystems. *Applied Geography*, 2017, 84: 11-22. DOI: 10.1016/j.apgeog.2017.04.005
- [11] Jia Q, Cao L, Yésou H, et al. Combating aggressive macrophyte encroachment on a typical Yangtze River lake: lessons from a long-term remote sensing study of vegetation. *Aquatic Ecology*, 2017, 51(1): 177-189. DOI: 10.1007/s10452-016-9609-9
- [12] Wang W H, Wang Y, Sun L Q, et al. Research and application status of ecological floating bed in eutrophic landscape water restoration. *Science of The Total Environment*, 2020, 704: 135434. DOI: 10.1016/j.scitotenv.2019.135434
- [13] Luo J, Yang J, Duan H, et al. Research progress of aquatic vegetation remote sensing in shallow lakes. *Natl. Remote Sens. Bull.*, 2022, 26: 68-76. [罗菊花, 杨井志成, 段洪涛, 等. 浅水湖泊水生植被遥感监测研究进展. 遥感学报, 2022, 26(1): 68-76. DOI: 10.11834/jrs.20221208]
- [14] Luo J, Li X, Ma R, et al. Applying remote sensing techniques to monitoring seasonal and interannual changes of aquatic vegetation in Taihu Lake, China. *Ecological Indicators*, 2016, 60: 503-513. DOI: 10.1016/j.ecolind.2015.07.029
- [15] Luo J, Ni G, Zhang Y, et al. A new technique for quantifying algal bloom, floating/emergent and submerged vegetation in eutrophic shallow lakes using Landsat imagery. *Remote Sensing of Environment*, 2023, 287: 113480. DOI: 10.1016/j.rse.2023.113480
- [16] Shi L, Jia Y, Zuo A, et al. Dynamic change of vegetation cover and productivity of Poyang Lake wetland based on MODIS EVI time series. *Biodiversity Science*, 2018, 26(8): 828. [史林鹭, 贾亦飞, 左奥杰, 等. 基于 MODISEVI 时间序列的鄱阳湖湿地植被覆盖和生产力的动态变化. 生物多样性, 2018, 26(8): 828-837. DOI: 10.17520/biods.2018089]
- [17] Villa P, Mousivand A, Bresciani M. Aquatic vegetation indices assessment through radiative transfer modeling and linear mixture simulation. *International Journal of Applied Earth Observation and Geoinformation*, 2014, 30: 113-127. DOI: 10.1016/j.jag.2014.01.017
- [18] Wang H, Li Y, Zeng S, et al. Recognition of aquatic vegetation above water using shortwave infrared baseline and phenological features. *Ecological Indicators*, 2022, 136: 108607. DOI: 10.1016/j.ecolind.2022.108607
- [19] Hu L, Zeng K, Hu C, et al. On the remote estimation of Ulva prolifera areal coverage and biomass[J]. *Remote Sensing of Environment*, 2019, 223: 194-207. DOI: 10.1016/j.rse.2019.01.014
- [20] Espel D, Courty S, Auda Y, et al. Submerged macrophyte assessment in rivers: An automatic mapping method using Pléiades imagery. *Water Research*, 2020, 186: 116353. DOI: 10.1016/j.watres.2020.116353
- [21] Yang L I, Chunbo W, Bin L I, et al. Surveying the composition, coverage, and biomass of aquatic vegetation by unmanned aerial vehicle and sampling survey: Lake Dapokou of lake Dianchi as an example. *Ecology and Environment*, 2020, 29(8): 1625. DOI: 10.16258/j.cnki.1674-5906.2020.08.014
- [22] Lu L, Luo J, Xin Y, et al. How can UAV contribute in satellite-based Phragmites australis aboveground biomass estimating? *International Journal of Applied Earth Observation and Geoinformation*, 2022, 114: 103024. DOI: 10.1016/j.jag.2022.103024
- [23] Duan H, Ma R, Xu X, et al. Two-decade reconstruction of algal blooms in China's Lake Taihu. *Environmental Science & Technology*, 2009, 43(10): 3522-3528. DOI: 10.1021/es8031852
- [24] Qin B, Xu P, Wu Q, et al. Environmental issues of Lake Taihu, China. *Hydrobiologia*, 2007, 581(1): 3-14. DOI: 10.1007/s10750-006-0521-5
- [25] Luo J, Duan H, Ma R, et al. Mapping species of submerged aquatic vegetation with multi-seasonal satellite images and considering life history information. *International Journal of Applied Earth Observation and Geoinformation*, 2017, 57: 154-165. DOI: 10.1016/j.jag.2016.11.007
- [26] Kumar L, Mutanga O. Google Earth Engine Applications Since Inception: Usage, Trends, and Potential. *Remote Sensing*, 2018,

10(10): 1509. DOI: 10.3390/rs10101509

- [27] Xiaoqin W, Miaomiao W, Shaoqiang W, *et al.* Extraction of vegetation information from visible unmanned aerial vehicle images. *Transactions of the Chinese Society of Agricultural Engineering*, 2015, 31(5). [汪小钦, 王苗苗, 王绍强, 等. 基于可见光波段无人机遥感的植被信息提取. 农业工程学报, 2015, 31(5): 152-159. DOI: 10.3969/j.issn.1002-6819.2015.05.022]
- [28] Otsu N. A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 1979, 9(1): 62-66. DOI: 10.1109/TSMC.1979.4310076
- [29] Purevdorj TS, Tateishi R, Ishiyama T, *et al.* Relationships between percent vegetation cover and vegetation indices. *International Journal of Remote Sensing*, 1998, 19(18): 3519-3535. DOI: 10.1080/014311698213795
- [30] Ge J, Meng B, Liang T, *et al.* Modeling alpine grassland cover based on MODIS data and support vector machine regression in the headwater region of the Huanghe River, China. *Remote Sensing of Environment*, 2018, 218: 162-173. DOI: 10.1016/j.rse.2018.09.019
- [31] Chen T, Guestrin C. Xgboost: A scalable tree boosting system[C]//Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016: 785-794. DOI: 10.1145/2939672.2939785
- [32] Villoslada M, Bergamo T F, Ward R D, *et al.* Fine scale plant community assessment in coastal meadows using UAV based multispectral data. *Ecological Indicators*, 2020, 111: 105979. DOI: 10.1016/j.ecolind.2019.105979
- [33] Zhang Y, Liu T, Tong X, *et al.* Inversion of vegetation coverage based on multi-source remote sensing data and machine learning method in the Horqin Sandy Land, China. *Journal of Desert Research*, 2022, 42(3): 187. [张亦然, 刘廷玺, 童新, 等. 基于多源遥感和机器学习方法的科尔沁沙地植被覆盖度反演. 中国沙漠, 2022, 42(3): 187-195. DOI: 10.7522/j.issn.1000-694X.2021.00149]
- [34] Ma R, Duan H, Gu X, *et al.* Detecting Aquatic Vegetation Changes in Taihu Lake, China Using Multi-temporal Satellite Imagery. *Sensors*, 2008, 8(6): 3988-4005. DOI: 10.3390/s8063988
- [35] Alagialoglou L, Manakos I, Katsikis E, *et al.* Machine Learning for Identifying Emergent and Floating Aquatic Vegetation from Space: A Case Study in the Dniester Delta, Ukraine. *SN Computer Science*, 2024, 5(5): 597. DOI: 10.1007/s42979-024-02873-7
- [36] Li Y, Xie P, Zhao D, *et al.* Eutrophication strengthens the response of zooplankton to temperature changes in a high-altitude lake. *Ecology and Evolution*, 2016, 6(18): 6690-6701. DOI: 10.1002/ece3.2308
- [37] Zhao L, Zhu R, Zhou Q, *et al.* Trophic status and lake depth play important roles in determining the nutrient-chlorophyll a relationship: Evidence from thousands of lakes globally. *Water Research*, 2023, 242: 120182. DOI: 10.1016/j.watres.2023.12018
- [38] Camps-Valls G, Campos-Taberner M, Moreno-Martínez Á, *et al.* A unified vegetation index for quantifying the terrestrial biosphere. *Science Advances*, 2021, 7(9): eabc7447. DOI: 10.1126/sciadv.abc7447
- [39] Feng X, Tian J, Wang Y, *et al.* Spatio-Temporal Variation and Climatic Driving Factors of Vegetation Coverage in the Yellow River Basin from 2001 to 2020 Based on kNDVI. *Forests*, 2023, 14(3): 620. DOI: 10.3390/f14030620



附图 I 2020 年 9 月 5 日的 Sentinel-2 MSI 影像假彩色合成图 (a); 基于 Sentinel 像元覆盖度估算模型的覆盖度图和放大区域 (b 和 c); 最邻近插值法、双线性插值法和三次卷积插值法重采样生成的覆盖度图 (d-e)

Attached Fig I False RGB Sentinel-2 MSI satellite image on September 5, 2020 (a); maps coverage and zoom based on Sentinel pixel coverage estimation model (b and c); coverage map generated by nearest neighbor interpolation, bilinear interpolation and cubic convolution interpolation resampling (d-e)

附表 I 基于 Landsat 像元浮叶/挺水植被覆盖度估算模型的不同重采样方法比较

Appendix I Comparison of different resampling methods based on Landsat pixel floating/emergent vegetation coverage estimation model

方法	训练集			测试集		
	决定系数 R^2	均方根误差 RMSE	平均绝对误差 MAE	决定系数 R^2	均方根误差 RMSE	平均绝对误差 MAE
最邻近插值法	0.98	3.89%	2.77%	0.97	4.90%	3.36%
双线性插值法	0.98	3.90%	2.77%	0.97	4.80%	3.35%
三次卷积插值法	0.98	3.91%	2.79%	0.97	4.85%	3.37%