

DOI: 10.20103/j.stxb.202412133074

张晓彤, 刘明月, 郎红梅, 张龙丰, 尹轩, 刘玮佳, 张永彬, 吴风华, 李富平, 满卫东. 基于无人机多光谱和机器学习的芦苇地上生物量反演研究. 生态学报, 2025, 45(14): - .

Zhang X T, Liu M Y, Lang H M, Zhang L F, Yin X, Liu W J, Zhang Y B, Wu F H, Li F P, Man W D. *Phragmites australis* aboveground biomass inversion based on UAV and Machine Learning. Acta Ecologica Sinica, 2025, 45(14): - .

# 基于无人机多光谱和机器学习的芦苇地上生物量反演研究

张晓彤<sup>1</sup>, 刘明月<sup>1,2,3,\*</sup>, 郎红梅<sup>4</sup>, 张龙丰<sup>4</sup>, 尹 轩<sup>1</sup>, 刘玮佳<sup>1</sup>, 张永彬<sup>1</sup>, 吴风华<sup>1</sup>, 李富平<sup>1,2</sup>, 满卫东<sup>1,2</sup>

1 华北理工大学矿业工程学院, 唐山 063210

2 河北省矿区生态修复产业技术研究院, 唐山 063210

3 矿产资源绿色开发与生态修复协同创新中心, 唐山 063210

4 唐山市自然资源和规划局, 唐山 063000

**摘要:** 植被地上生物量 (AGB) 是反映湿地生态系统功能状态与健康水平的重要生态参数, 可为生态系统稳定性分析和碳汇能力评估提供关键支撑。以滨海典型芦苇湿地为研究对象, 基于无人机获取的高分辨率多光谱影像和 113 个实测芦苇 AGB 样本, 系统评估不同特征变量组合的建模效果, 并利用随机森林 (RF) 算法的重要性分析对变量进行筛选。随后分别构建了 RF、支持向量机 (SVM) 和极端梯度提升 (XGBoost) 三种机器学习算法模型, 对筛选后的特征变量进行建模与精度对比, 旨在探索基于低空遥感技术开展湿地植被生物量反演的可行路径, 建立适用于湿地生态监测的高效估算框架。研究表明, 联合使用多光谱波段反射率与植被指数的变量组合显著优于单一变量类型, 能够有效提升 AGB 估算精度。在三种建模算法中, XGBoost 模型性能最优,  $R^2$  为 0.731, RMSE 为  $0.184\text{kg}/\text{m}^2$ , RPD 达到 2.431, 表现出较强的稳定性与鲁棒性。空间分布分析结果显示, 研究区芦苇 AGB 具有显著的空间异质性, 高值区主要分布于水体周边的浅水带及地势低洼区域, 而地势较高区域的 AGB 值相对较低。定量统计显示, 研究区芦苇总 AGB 为 3038.78 吨, 芦苇 AGB 值范围为  $0.486\text{—}1.705\text{kg}/\text{m}^2$ , 平均为  $0.88\text{kg}/\text{m}^2$ 。变量贡献分析显示, 红边波段及其衍生植被指数, 特别是修正型叶绿素吸收反射植被指数 (MCARI), 在芦苇 AGB 反演模型中的贡献度较高, 表明其在捕捉湿地植被的叶绿素含量和冠层结构变化方面具有较强的敏感性, 对 AGB 的估算精度有显著提升作用。本研究不仅验证了基于无人机多光谱影像结合机器学习算法进行芦苇 AGB 反演的可行性和有效性, 也为后续开展湿地生态系统功能评价、碳储量估算及关键栖息地监测提供了科学依据。

**关键词:** 机器学习; 地上生物量; 植被指数; 无人机反演; 变量组合

## *Phragmites australis* aboveground biomass inversion based on UAV and Machine Learning

ZHANG Xiaotong<sup>1</sup>, LIU Mingyue<sup>1,2,3,\*</sup>, LANG Hongmei<sup>4</sup>, ZHANG Longfeng<sup>4</sup>, YIN Xuan<sup>1</sup>, LIU Weijia<sup>1</sup>, ZHANG Yongbin<sup>1</sup>, WU Fenghua<sup>1</sup>, LI Fuping<sup>1,2</sup>, MAN Weidong<sup>1,2</sup>

1 College of Mining Engineering, North China University of Science and Technology, Tangshan 063210, China

2 Hebei Industrial Technology Institute of Mine Ecological Remediation, Tangshan 063210, China

3 Collaborative Innovation Center of Green Development and Ecological Restoration of Mineral Resources, Tangshan 063210, China

4 Tangshan Natural Resources and Planning Bureau, Tangshan 063000, China

**基金项目:** 河北省自然科学基金项目 (D2023209008); 河北省中央引导地方科技发展资金项目 (236Z3305G); 国家自然科学基金项目 (41901375, 52274166)

**收稿日期:** 2024-12-13; **网络出版日期:** 2025-00-00

\* 通讯作者 Corresponding author. E-mail: liumy917@ncst.edu.cn

**Abstract:** Vegetation aboveground biomass (AGB) is an important ecological parameter reflecting the functional status and health level of wetland ecosystems, which can provide key support for ecosystem stability analysis and carbon sink capacity assessment. In this paper, we take typical *Phragmites australis* (*P. australis*) wetlands in Binhai as the research object, and based on the high-resolution multispectral images acquired by UAV and 113 measured *P. australis* AGB samples, we systematically evaluate the modelling effect of different combinations of feature variables, and screen the variables by using the importance analysis of the Random Forest (RF) algorithm. Subsequently, three machine learning algorithm models, namely RF, Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost), were constructed respectively to model and compare the accuracy of the screened feature variables, aiming at exploring feasible paths to carry out wetland vegetation biomass inversion based on low-altitude remote sensing technology, and establishing a high-efficiency estimation framework applicable to wetland ecological monitoring. The results show that the joint use of the variable combination of multispectral band reflectance and vegetation index is significantly better than the single variable type, which can effectively improve the accuracy of AGB estimation. Among the three modelling algorithms, the XGBoost model has the best performance, with a  $R^2$  of 0.731, a RMSE of 0.184 kg/m<sup>2</sup>, and a RPD of 2.431, which shows strong stability and robustness. The results of the spatial distribution analysis showed that the *P. australis* AGB in the study area had significant spatial heterogeneity, with the high value areas mainly distributed in the shallow water zone around the water body and in the low-lying areas, while the AGB values in the higher terrain areas were relatively low. Quantitative statistics showed that the total AGB of *P. australis* in the study area was 3038.78t, and the range of *P. australis* AGB values was 0.486—1.705kg/m<sup>2</sup>, with an average of 0.88kg/m<sup>2</sup>. Variable contribution analysis showed that the red edge band and its derived vegetation indexes, especially Modified Chlorophyll Absorption Reflectance Vegetation Index (MCARI) contributed more in the *P. australis* AGB inversion model, indicating its strong sensitivity in capturing changes in chlorophyll content and canopy structure of wetland vegetation, which significantly enhanced the estimation accuracy of AGB. This study not only verifies the feasibility and effectiveness of *P. australis* AGB inversion based on UAV multispectral images combined with machine learning algorithms, but also provides a scientific basis for subsequent wetland ecosystem function evaluation, carbon stock estimation and critical habitat monitoring.

**Key Words:** machine learning; aboveground biomass; vegetation index; UAV inversion; variable combinations

湿地作为自然界最富生物多样性的生态景观和人类最重要的自然资源<sup>[1-2]</sup>,在全球碳循环中起着极其重要的作用,具有极高的生态和研究价值<sup>[3-4]</sup>。芦苇(*Phragmites australis*, *P. australis*)是一种全球广泛分布的湿地植物,具有显著生态适应性,尤其在盐碱地等逆境条件下展现出强大的适生性和生态改良潜力。作为湿地生态系统的主要植被类型之一,不仅能为人类的生产活动提供较高的经济价值,如制浆造纸、药用及保健、出口创汇等<sup>[5-6]</sup>,其光合作用还是“蓝碳”的重要驱动力。芦苇地上生物量(Aboveground Biomass, AGB)是评估湿地生态系统健康状况的重要指标<sup>[7]</sup>,也是评估湿地生态系统碳循环和初级生产力的关键参数<sup>[8]</sup>。在全球环境变化背景下,沼泽湿地固碳研究已成为碳循环研究的关键课题。芦苇 AGB 空间分布格局多集中于典型区域的大范围分布研究,但面向湿地生态系统精细化管理需求,生境破碎和碳动态高度异质性的典型群落仍缺乏关注<sup>[9]</sup>。因此,如何快速准确对区域芦苇 AGB 估算并进行评估分析,对湿地生态系统管理、碳储量监测和群落动态变化等多方面具有重要意义<sup>[10-11]</sup>。

传统人工收割植被 AGB 测量方法费时费力,对植物具有破坏性,且湿地水文条件复杂、可达性差,难以获得足量有效样本<sup>[12-13]</sup>。遥感技术可以低成本收集多时相信息,快速实现植被参数的无损检测,目前已成为植被生态遥感参量反演的重要手段<sup>[14-15]</sup>。然而,湿地的破碎化和高度空间异质性使得卫星遥感数据与野外采样之间的时空差异难以弥补,从而影响数据的时效性和准确性<sup>[16-18]</sup>。相比于卫星遥感,无人机遥感探测周期短,影像空间分辨率高,能够近实时获取与研究区采样时间点相吻合的影像,尤其在需要时效性和精细尺度生

物量反演方面独具优势<sup>[19-20]</sup>。另外有研究发现,无人机影像相比 Landsat 卫星影像,在湿地生物量反演中可有效提高反演精度,尤其在反映湿地植被生物量空间变异性方面优势更大<sup>[21]</sup>。由此可见,无人机遥感在湿地植被 AGB 估算中的应用前景广阔,尤其是在提高反演精度和及时性方面具有明显优势。

植被 AGB 遥感反演方法从统计模型逐渐发展到非参数化反演算法,前者主要应用在早期植被 AGB 反演研究中,但其无法充分捕捉复杂的非线性关系<sup>[22-23]</sup>,而机器学习等非参数化算法更适用于高维和非结构化的遥感数据,能够更好的识别数据间的多重共线性,提高模型预测精度,尤其在湿地植被生物量估算中展现了明显优势<sup>[24-25]</sup>。为探究滨海湿地生态系统的固碳能力,相关研究采用随机森林(Random Forest, RF)模型对湿地地上生物量进行评估<sup>[26]</sup>。多数研究表明,RF 模型在建模效果上要优于多元线性回归等传统统计方法<sup>[27]</sup>。此外,支持向量机(Support Vector Machine, SVM)和极端梯度提升(eXtreme Gradient Boosting, XGBoost)在湿地植被 AGB 反演中也表现出了优异的性能,尤其是在红树林等湿地生物量的估算中,精度得到了显著提升<sup>[28]</sup>。虽然 RF、SVM 和 XGBoost 等方法在湿地植被 AGB 估算中具有潜力,但当前针对芦苇湿地生物量估算的相关研究仍然较少,尤其是结合多光谱遥感数据和机器学习建模方法,仍有很大的探索空间。

本文提出了一种适用于群落尺度的低空遥感空间量化反演 AGB 框架,以典型芦苇湿地作为研究对象,利用实测芦苇 AGB 和无人机多光谱影像,提取波段反射率及植被指数,通过 RF 模型提供的重要性进行变量筛选,结合 RF、SVM 和 XGBoost 三种机器学习算法,构建芦苇地上生物量估算模型,并绘制其空间分布图。通过开展湿地芦苇 AGB 反演研究,为湿地生态系统的精细化管理、资源合理利用以及湿地碳汇能力评估提供理论依据和数据支持,同时可跨尺度拓展到区域尺度 AGB 估算,提供与卫星遥感影像尺度匹配的足量样本。

## 1 研究区及数据

### 1.1 研究区概况

本文选择唐山曹妃甸典型芦苇湿地作为研究区,研究区长宽分别为 560m 和 380m(图 1)。该地区属于东部季风区温带半湿润地区,年平均气温 11℃,年平均降水量 636mm。研究区主要植被类型为芦苇,东西两侧地势较高;南侧相对北侧地势较高,水源相对匮乏,芦苇生长较为稀疏;而北侧湿地水文条件良好,芦苇生长茂盛,沿水系形成了优良的湿地生态系统。研究区芦苇湿地因其自然状态的保持、生物多样性的丰富以及对水鸟和湿地物种的重要栖息地作用,对区域生态系统功能和生物多样性保护具有关键意义。

### 1.2 数据来源与处理

#### 1.2.1 无人机多光谱数据获取与处理

于 2023 年 5 月 31 日收集了无人机多光谱影像,选择晴空无云、风速低的条件下完成影像采集任务。由搭载 CMOS 镜头的大疆精灵 4 拍摄获取无人机多光谱影像,包括 1 个用于可见光成像的彩色传感器以及蓝光((450±16)nm)、绿光((560±16)nm)、红光((650±16)nm)、近红外((840±16)nm)和红边((730±16)nm)传感器。数据采集前,在研究区内设置 10 个地面控制点,架设一台 GPS 基站,采用实时动态定位 RTK 方法精确测定每个控制点的经纬度;数据采集过程中,无人机在距地面高度 200m 以 6m/s 的速度飞行,拍照间隔为 2.0s,航向和旁向重叠度设置为 80%,通过内置的 TimeSync 系统获得高精度位置信息,拼接得到地面分辨率为 6cm 的原始多光谱影像和 5 个单波段的正射影像<sup>[29]</sup>;然后使用地面控制点坐标对影像进行几何精校正;最后将多光谱数据重采样至 20cm,取样方中心像素值作为遥感变量与样点数据进行匹配,以确保影像与样点的空

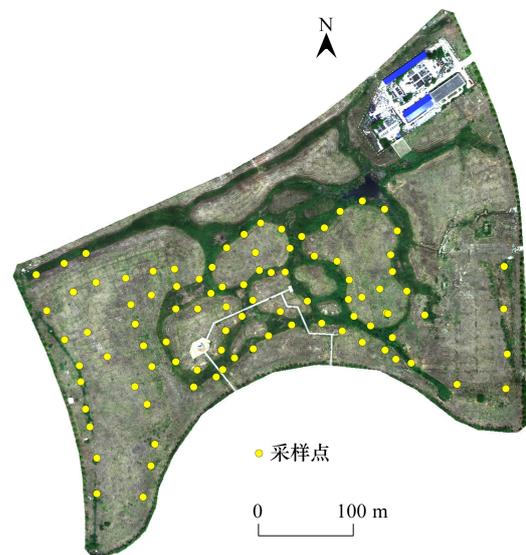


图 1 研究区及采样点分布

Fig.1 Distribution of study area and sampling sites

间分布更加一致。采用随机森林算法对研究区多光谱影像进行分类,划分为芦苇、道路、建筑用地、水体和裸地五大类,总体分类精度可达 94.69%,kappa 系数为 0.93,其中芦苇分类精度为 95.68%。

### 1.2.2 样点数据获取及处理

芦苇 AGB 于 2023 年 5 月 31 日同步进行采集。依据全面性、代表性和典型性的样点选取原则<sup>[30]</sup>,在研究区内随机设置均质样点,样点之间间隔 30—40m,内设 1 个 0.5m×0.5m 的样方,使用高精度 GNSS 记录样方中心位置的经纬度坐标、植株密度,齐地收割样方内的芦苇地上部分,随机选取 5 株,称取鲜重。迅速保鲜带回实验室,经 105℃ 条件下杀青 2h,之后 85℃ 条件下烘干 24h 至恒重,称取干重(精度 0.01g)。根据每个样点获得的 5 株芦苇的 AGB 计算出每个样方芦苇 AGB,再将其换算为单位面积生物量,本文共采集有效生物量样本 113 个。单位面积芦苇 AGB 计算公式如下:

$$AGB = AGB_{avg} \times D$$

式中, $AGB_{avg}$  表示样方内随机选取的 5 株芦苇样本 AGB 的均值; $D$  为植株密度,是单位面积内芦苇总株数<sup>[31]</sup>。

### 1.3 遥感变量提取

植被指数是除光谱反射率外构建生物量估测模型的重要变量<sup>[32]</sup>,为探究无人机多光谱提取的光谱指数在芦苇生物量反演中作用,本文参考相关文献,提取了 5 个原始光谱反射率及 27 种与湿地植被地上生物量密切相关的植被指数(表 1)。

## 2 研究方法

本文基于研究区域内的无人机多光谱影像、芦苇实地调查样点与实测芦苇 AGB;采用随机森林分类方法获取芦苇的空间分布;基于波段反射率和植被指数单一类型变量及二者集成的变量组合的随机森林重要性筛选三种不同变量组合,分别结合 RF、SVM 和 XGBoost 三种机器学习算法,共构建九个芦苇 AGB 估算模型,通过精度评价指标对模型进行评估,以最优模型对芦苇 AGB 进行空间反演,实现对湿地生态系统中芦苇 AGB 的精确估算(图 2)。

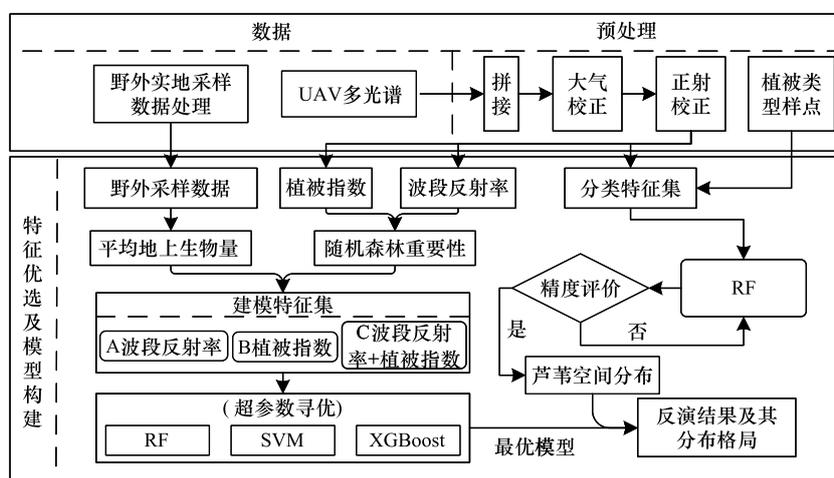


图 2 技术流程图

Fig.2 Flow chart

### 2.1 芦苇地上生物量描述性分析

对本文获取的 113 个有效芦苇地上生物量样本进行统计发现(表 2):芦苇地上生物量数值范围为 0.211—2.189kg/m<sup>2</sup>,平均值为 0.931kg/m<sup>2</sup>,中位数为 0.828kg/m<sup>2</sup>,标准差为 0.447kg/m<sup>2</sup>。

表 1 遥感变量  
Table 1 Remote sensing variables

变量类型 Variable type	数据名称 Data name	全称 Full name	说明 Introduction	参考文献 References
反射率 Reflectance	单波段反射率:蓝、绿、红、红边和近红 修正型叶绿素吸收反射植被指数(MCARI)	BLUE、GREEN、RED、REDEEDGE、NIR Modified Chlorophyll Absorption Reflectance Vegetation Index	$B_{band, i=1,2,3,4,5}$ $MCARI = [(B4-B3) - 0.2 * (B4-B2)] * (B4/B3)$	[33] [34]
植被指数 Vegetation Index	比值植被指数(RVI) 归一化植被指数(NDVI) 增强型植被指数(EVI) 优化土壤调整植被指数(OSAVI) 绿色比值植被指数(GRVI) 归一化差异红边指数(NDRE) 绿色归一化植被指数(GNDVI) 绿色差值植被指数(GDVI) 绿红外植被指数(GIPVI) 权重差值植被指数(WDVI) 修正型土壤调节植被指数(MSAVI) 红植被指数(IPVI) 改进非线性植被指数(MNLI) 垂直植被指数(PVI) 差值植被指数(DVI) 归一化绿蓝差值植被指数(NGBDI) 抗大气指数(VARI) 叶绿素指数(GI) 叶绿素吸收比值指数(CARI) 绿色差异指数(GDI) 绿叶指数(GLI) 红边归一化差值植被指数(RENDVI) 结构不敏感色素指数(SIPI) 归一化绿红差值植被指数(NGRDI) 特定色素简单比值植被指数(PSSRa) 归一化红色指数(RED)	Ratio Vegetation Index Normalized Difference Vegetation Index Enhanced Vegetation Index Optimized Soil-Adjusted Vegetation Index Green-Red Vegetation Index Normalized Difference Red Edge Index Green Normalized Difference Vegetation Index Green Difference Vegetation Index Green Infrared Percentage Vegetation Index Weighted Difference Vegetation Index Modified Soil-Adjusted Vegetation Index Infrared Percentage Vegetation Index Modified Non-Linear Vegetation Index Perpendicular Vegetation Index Difference Vegetation Index Normalized Green-Blue Difference Index Visible Atmospherically Resistant Index Greenness Index Chlorophyll Absorption Ratio Index Green Difference Index Green Leaf Index Red-Edge Normalized Difference Vegetation Index Structure Insensitive Pigment Index Normalized Green-Red Difference Index Pigment-Specific Simple Ratio for Chlorophyll-a Redness Index	$RVI = B3/B5$ $NDVI = (B5-B3)/(B5+B3)$ $EVI = 2.5 * ((B5-B3)/(B5+6 * B3 - 7.5 * B1 + 1))$ $OSAVI = 1.16 * (B5-B3)/(B5+B3+0.16)$ $GRVI = B5/B2$ $NDRE = (B5-B4)/(B5+B4)$ $GNDVI = (B5-B2)/(B5+B2)$ $GDVI = B5-B2$ $GIPVI = B5/(B5+B2)$ $WDVI = B5 - 0.5 * B3$ $MSAVI = (2 - NDVI * WDWI) * (B5-B3)/8 * (B5+B3 + 1 - NDVI * WDWI)$ $IPVI = B5/(B5+B3)$ $MNLI = (1.5 * B5^2 - 1.5 * B2)/(B5^2+B3+0.5)$ $PVI = B5/B3$ $DVI = B5-B3$ $NGBDI = (B2-B1)/(B2+B1)$ $VARI = (B2-B3)/(B2+B3-B1)$ $GI = (B3-B2)/(B3+B2)$ $CARI = (B4-B3)/(0.2 * (B4-B2))$ $GDI = B5-B3+B2$ $GLI = (2 * B2-B3-B1)/(2 * B2+B3+B1)$ $RENDVI = (B4-B3)/(B4+B3)$ $SIPI = (B5-B1)/(B5+B1)$ $NGRDI = (B2-B3)/(B2+B3)$ $PSSRa = B4/B3$ $RED = B3/(B1+B2+B3)$	[34] [34] [34] [34] [34] [35] [36] [34] [35] [35] [36] [36] [36] [35] [35] [34] [34] [37] [38] [35] [35] [38] [34] [35] [36] [35]

表 2 芦苇样本的统计性描述

Table 2 Statistical description of *Phragmites australis* samples

芦苇 <i>Phragmites australis</i>	样本数量 Number of samples	平均值 Mean/ (kg/m <sup>2</sup> )	最大值 Maximum/ (kg/m <sup>2</sup> )	最小值 Minimum/ (kg/m <sup>2</sup> )	中位数 Median/ (kg/m <sup>2</sup> )	标准差 Standard deviation/ (kg/m <sup>2</sup> )
芦苇 <i>Phragmites australis</i>	113	0.931	2.189	0.211	0.828	0.447

## 2.2 遥感变量筛选

为探讨不同变量组合对芦苇 AGB 反演的影响,遥感变量被分为三类:波段反射率(A)、植被指数(B)以及波段反射率+植被指数(C)的组合。为了提高建模效率并增强模型的可解释性<sup>[39]</sup>,通过 RF 算法对这些变量的重要性进行排序,自动评估各变量对模型预测的贡献。通过 RF 算法的自动化变量筛选能力、强抗噪性、适应非线性与高维数据的特点<sup>[40-41]</sup>,以及其易于解释与高效建模的优势,从 A、B、C 三种变量组合中挑选出对模型贡献较大的变量,并根据重要性逐步增加变量数量进行建模。最后,以模型均方根误差(RMSE)最小的变量组合作为最终特征变量,建立非参数模型<sup>[33]</sup>。

## 2.3 模型构建

本文选择 RF、SVM 和 XGBoost 三种机器学习模型,在变量筛选基础上分别构建不同变量组合下的湿地芦苇 AGB 估算模型。三种建模方法是通过 R 语言中“caret”包的“train”函数实现的,并采用网格搜索法对各模型中的主要参数进行优化。

RF 是一种基于集成学习思想的算法。该算法由多个决策树构成,通过决策树来对数据进行建模,同时将预测的结果进行集成,以提高模型的稳定性和性能。在 RF 中,每棵决策树在进行随机选择子样本后,在子样本上独立的进行训练,增加了模型的随机性,提高了组合决策树模型对 AGB 预测的精度,具有较强的抗噪性,可以防止过拟合现象的产生。最终的 AGB 预测结果是由多个决策树的预测结果进行加权得到的。

SVM 是一种监督学习方法,主要用于二元分类。它通过求解最大边距超平面来确定决策边界,使样本点到超平面的距离最大化<sup>[42]</sup>。SVM 能有效解决小样本学习问题,并通过核函数克服维度灾难和非线性可分问题。在高维空间映射时,SVM 保持较低的计算复杂度。为提高拟合能力,本研究使用 sigmoid 核函数,将数据映射到高维空间以处理复杂的非线性关系。

XGBoost 是一种高效的基于决策树的预测模型。与传统的梯度提升决策树(GBDT)不同,XGBoost 通过泰勒展开式对损失函数进行二阶近似优化,加速了梯度下降过程,提升了模型的收敛速度和预测精度。它通过递归划分特征空间,并依据残差调整和优化模型,在每轮迭代中逐步逼近最优解。为提高精度,XGBoost 结合 boosting 方法并通过并行计算加速训练,弥补单棵回归树的局限性。

## 2.4 模型精度评价指标

采用十折交叉验证方法来保证模型的稳定性,减少划分训练和验证样本对最后输出结果导致的偶然性误差。同时使用决定系数(Determination Coefficient,  $R^2$ )、均方根误差(Root Mean Square Error, RMSE)以及相对百分比差异(Relative Percentage Difference, RPD)3 个指标评价模型的估算效果。 $R^2$ 表示自变量和因变量的拟合效果, $R^2$ 的值在 0—1 之间, $R^2$ 越接近于 1,表示模型的拟合效果越好;RMSE 用于衡量预测值与真实值之间的误差,RMSE 越小,表示模型的预测结果越好;RPD 是标准差与均方根误差的比值,证明模型的预测能力,当  $RPD < 1.4$  时,表示模型无法估计 AGB;当  $1.4 < RPD \leq 2.0$  时,模型的预测能力相对可靠;当  $RPD > 2.0$ ,认为该模型的可靠性较高<sup>[43]</sup>。 $R^2$ 、RMSE 和 RPD 的计算公式如下:

$$R^2 = \frac{\sum (Y'_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2}$$

$$RMSE = \sqrt{\frac{\sum (Y_i - Y'_i)^2}{N}}$$

$$RPD = \frac{SD}{RMSE}$$

式中,  $Y_i$  为实测地上生物量,  $Y'_i$  为模型估算的地上生物量,  $\bar{Y}$  为平均实测地上生物量,  $N$  为样本量,  $SD$  为样本的标准偏差。

### 3 结果与分析

#### 3.1 变量筛选结果

对比 A、B、C 三组不同变量的 RF 模型变量重要性排序和 RMSE 变化趋势, 发现随着变量的逐渐增加, RMSE 的波动逐渐变小, 当变量数目分别为 4、19 和 15 时, A、B、C 三组的 RMSE 达到最小(图 3), 因此 A、B、C 三组分别选择前 4、19、15 个变量来构建机器学习模型对芦苇 AGB 进行估测。总体来看, 植被指数的重要性高于波段反射率。仅采用波段反射率建模时, 红边波段反射率对芦苇 AGB 的重要性最高, 可达 30.2%, 使用植被指数或波段反射率与植被指数组合建模时, MCARI 的重要性最高, 且两组变量中重要性排名前五的变量多由红边和近红波段组成(图 3)。这一结果表明, 红边及近红波段对植被生长、光合作用以及叶绿素含量变化非常敏感; MCARI 作为一种改良叶绿素吸收反射指数, 对叶绿素的变化非常敏感, 也能够反映植被的生物量变化, 同时能够减少冠层参数, 如叶面积指数(LAI)和土壤背景等变化的影响, 这一点在高植被覆盖区尤为重要, 因为它有助于减少饱和现象的发生, 从而提供更为准确的生物量估计。

#### 3.2 模型精度比较

不同变量组合和建模方法在芦苇 AGB 估算性能上存在显著差异(表 3)。结合波段反射率和植被指数的模型表现出最佳估算效果。相较于仅使用单一波段反射率, 引入植被指数能够显著提高模型的预测精度。这可能是由于植被指数能够综合多个波段的信息, 如红边波段及红边指数, 对植被生长状态和生物量信息更为敏感, 同时能够减少冠层参数和土壤背景等变化的影响, 从而增强了模型的预测能力。在波段反射率的基础上, 加入植被指数后, 三种模型的建模精度  $R^2$  分别提升了 2.80%、6.89% 和 12.59%。不同变量组合构建的 XGBoost 模型预测精度均高于 SVM 和 RF 模型; 其中联合波段反射率和植被指数构建的 XGBoost 估测精度最高,  $R^2$  为 0.731, RMSE 为 0.184 kg/m<sup>2</sup>, RPD 为 2.431。

表 3 模型精度验证结果

Table 3 Model accuracy validation results

变量类型 Variable	模型 Model	决定系数 ( $R^2$ ) Determination coefficient	均方根误差 (RMSE) Root mean squared error/(kg/m <sup>2</sup> )	相对百分比差异 (RPD) Relative percentage difference
A: 波段反射率	RF	0.382	0.204	2.188
A: Reflectance	SVM	0.622	0.230	1.943
	XGBoost	0.639	0.220	2.032
B: 植被指数	RF	0.393	0.198	2.255
B: Vegetation Index	SVM	0.660	0.195	2.296
	XGBoost	0.681	0.187	2.380
C: 波段反射率+植被指数	RF	0.393	0.197	2.262
C: Reflectance+Vegetation Index	SVM	0.668	0.191	2.338
	XGBoost	0.731	0.184	2.431

RF: 随机森林 Random Forest; SVM: 支持向量机 Support Vector Machine; XGBoost: 极端梯度提升 eXtreme Gradient Boosting

采用十折交叉验证对芦苇 AGB 估测模型进行验证, 通过分析不同模型拟合效果与 1:1 线的偏离程度, 结果表明 XGBoost 模型的拟合曲线更接近 1:1 线, 且实测值与预测值呈现出较好的线性关系, 表明该模型更适用于芦苇 AGB 的预测。3 种模型均存在高值低估和低值高估现象, 尤其在高植被覆盖区域, 光学信号趋于饱和, 导致模型难以捕捉进一步增加的植被信息, 从而低估了芦苇 AGB 大于 1.5 kg/m<sup>2</sup> 的区域; 而在低植被覆盖



区域,产生了高估现象,特别是对于 AGB 低于  $1.0\text{kg}/\text{m}^2$  的区域(图 4)。联合波段反射率和植被指数构建的 XGBoost 模型的离散程度较低,结果更为稳定,是本文芦苇 AGB 估测的最优模型。

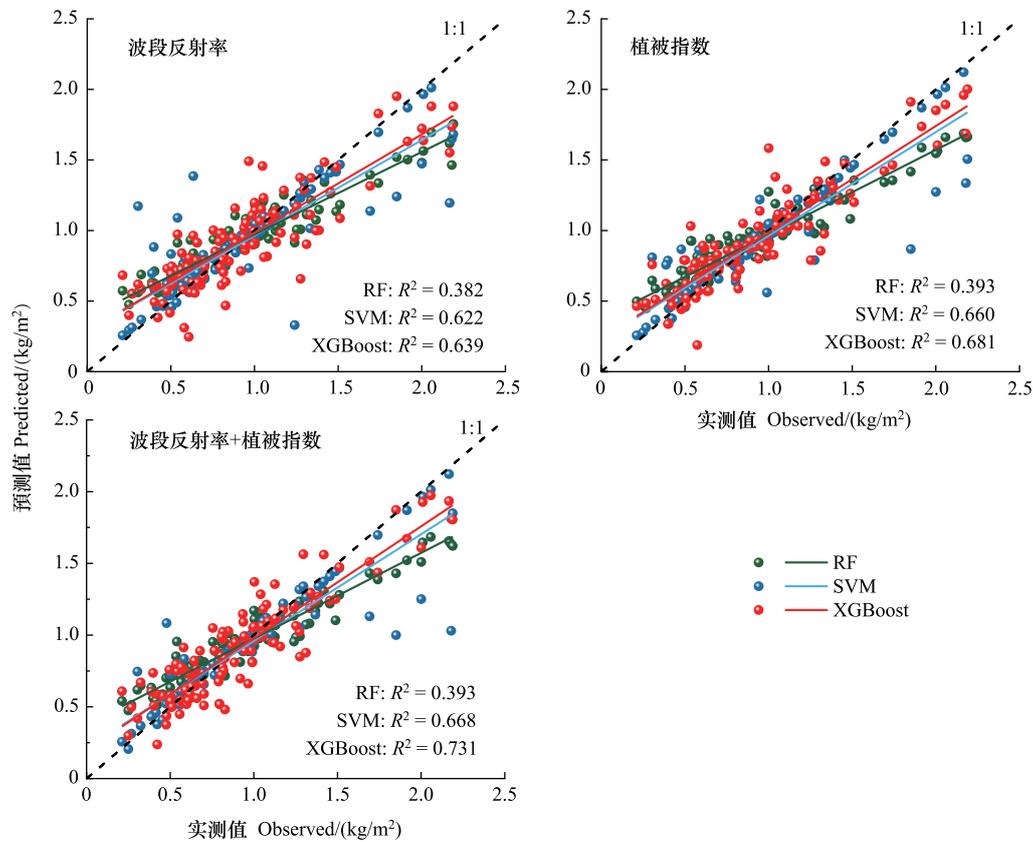


图 4 不同组合的芦苇地上生物量拟合图

Fig.4 AGB fitting plots for different combinations of *Phragmites australis*

AGB: 地上生物量 Aboveground Biomass; RF:随机森林 Random Forest; SVM:支持向量机 Support Vector Machine.; XGBoost: 极端梯度提升 eXtreme Gradient Boosting

### 3.3 芦苇 AGB 空间分布

联合波段反射率和植被指数,结合 XGBoost 模型对研究区芦苇 AGB 进行空间反演制图,结果表明芦苇 AGB 在研究区内呈现明显的空间异质性(图 5)。芦苇 AGB 估测值范围为  $0.486\text{—}1.705\text{kg}/\text{m}^2$ ,平均值为  $0.881\text{kg}/\text{m}^2$ ,总量为  $3038.78\text{t}$ ,空间分布与实际调查情况基本一致,验证了该模型在芦苇 AGB 反演中的有效性。中部低洼湿地及浅水区域芦苇 AGB 较高,东西两侧大面积芦苇 AGB 呈低值分布,南北两侧 AGB 相对较高。北侧低洼湿地或浅水区域芦苇 AGB 值较高,水源充沛可为芦苇生长提供良好的生态水文条件;中部偏东位置分布研究区最大范围的水域,此地区芦苇 AGB 值明显高于其他地区,证实水文条件对芦苇生长具有重要影响;而南侧芦苇相较于北侧 AGB 值较低,这可能与南侧地势较高以及修建栈道、观光等人类活动阻隔或影响芦苇生长有关。

## 4 讨论

综合利用无人机多光谱影像的波段反射率及其衍生的植被指数,可以显著提高芦苇 AGB 的建模精度。与单独使用波段反射率或植被指数相比,XGBoost 模型的  $R^2$  分别提高了 14.40% 和 7.34%。研究表明,尽管植被指数作为建模变量能够提升模型的预测精度,但存在一定的局限性<sup>[44]</sup>。然而,采用多光谱影像的波段反射率和植被指数的组合,可以显著提高模型对植被变化的敏感性。植被指数通过综合多个波段的信息,能够更

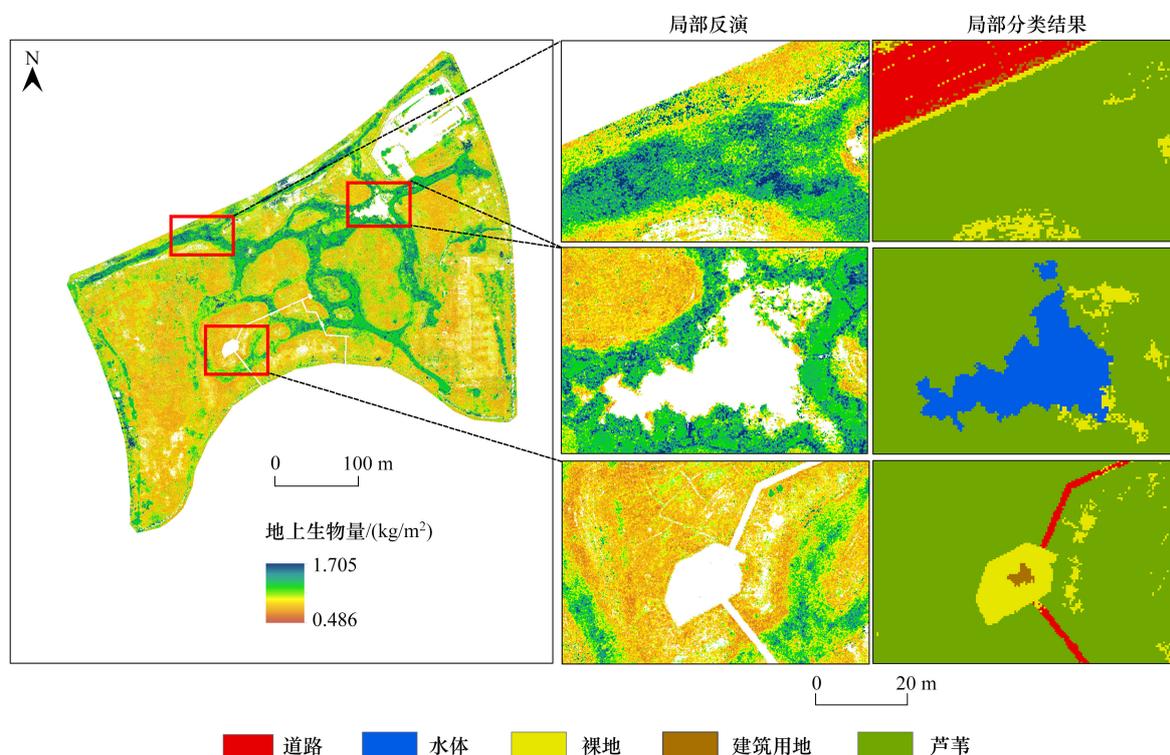


图5 芦苇地上生物量反演图和部分局部细节图和分类结果

Fig.5 *Phragmites australis* AGB inversion map and some parts of local details and classification results

准确地捕捉芦苇特有的光谱特征,减少非植被因素的干扰,从而增强模型对芦苇 AGB 估算的精度,提升其对芦苇 AGB 生理和生化变化的响应能力<sup>[45]</sup>。在变量筛选结果中,红边波段和红边指数对芦苇 AGB 估算贡献较大,这一发现与过去对草地生物量估算的研究成果相一致,强调了红边指数在湿地生物量反演中具有更高的灵敏度<sup>[46-47]</sup>。

不同建模算法在芦苇 AGB 预测中的表现存在差异,而建模效果受不同植被类型或生态系统特性的影响显著。选择合适的建模方法应根据数据类型、数据质量以及研究区域的植被覆盖等特征来决定。通过对比, XGBoost 模型的估算精度最高,  $R^2$  达 0.731, 优于 SVM 和 RF。有研究表明, RF 模型在估算草地等生态系统中的生物量反演时<sup>[48]</sup>, 能够通过构建多个决策树并取其平均值来减少过拟合, 并在面对复杂多样的数据集或样本量较小的情况下依然保持较好的稳定性, 但在面对噪声较大的样本集上容易产生过拟合现象。对于 SVM 模型, 由于芦苇 AGB 存在较强的空间异质性, 若对于所有样本使用统一核函数, 可能无法捕获变量间的多重共线性, 从而限制了模型的估算精度。相较之下, XGBoost 通过简化模型复杂度、提高学习效率并防止过拟合, 适合处理小样本数据<sup>[49]</sup>, 特别的在湿地植被生物量估算中具有较强的优势。

基于 XGBoost 模型的芦苇 AGB 空间分布预测结果与实际调查结果相符, 具有显著的空间异质性, 验证了模型的合理性。芦苇 AGB 估测值范围为 0.486—1.705 kg/m<sup>2</sup>, 与 Lu 等人<sup>[31]</sup> 对 7 月南大港芦苇 AGB 的估测结果相比, 本文芦苇 AGB 估算值相对较低, 这一差异可能归因于本研究的采样时间为 5 月底, 并非芦苇生物量峰值阶段。但本文形成的无人机低空遥感反演芦苇 AGB 方法框架可保障早期芦苇 AGB 估算精度。

尽管机器学习算法和多变量组合在提升湿地芦苇地上生物量模型估算精度方面取得了进展, 但模型仍存在高值低估和低值高估现象。未来研究可以通过整合多源遥感数据, 如无人机高光谱影像和雷达点云数据, 来进一步提高预测精度。本文采用的低空遥感技术具备较高的成本效益和分辨率, 提供了一种经济且实用的湿地芦苇 AGB 估算方案, 有助于提升湿地生态精细化监测的科学性, 为湿地资源的可持续管理提供技术

保障。

## 5 结论

本文基于无人机多光谱影像、实测芦苇 AGB 数据和机器学习算法,提出一种适用于群落尺度的盐沼植被 AGB 反演框架,旨在获取空间异质性显著的芦苇湿地 AGB 精细化空间分布格局。结果表明:(1)红边波段及其衍生的红边指数对芦苇 AGB 估算的贡献度较大;(2)引入植被指数可显著提升芦苇 AGB 的模型精度,以波段反射率和植被指数组合构建的芦苇 AGB 反演模型效果最好;三种算法对比,XGBoost 模型的反演精度最高, $R^2$  为 0.731, RMSE 为 0.184kg/m<sup>2</sup>, RPD 为 2.431;(3)研究区芦苇 AGB 总量为 3038.78t, 值域范围介于 0.486—1.705kg/m<sup>2</sup>, 平均值为 0.881kg/m<sup>2</sup>;但芦苇 AGB 空间异质性明显,东西两侧大面积芦苇 AGB 呈低值分布,北侧 AGB 相对较高,特别的在低洼湿地及浅水区域,芦苇 AGB 值较高但分布较破碎,与实地调查数据吻合,可为湿地芦苇生物量的快速预测及生态管理提供数据支持。

## 参考文献(References):

- [ 1 ] 梁建平, 马大喜, 毛德华, 王宗明. 双台河口国际重要湿地芦苇地上生物量遥感估算. 国土资源遥感, 2016, 28(3): 60-66.
- [ 2 ] Barbier E B, Hacker S D, Kennedy C, Koch E W, Stier A C, Silliman B R. The value of estuarine and coastal ecosystem services. *Ecological Monographs*, 2011, 81(2): 169-193.
- [ 3 ] 唐娜, 崔保山, 赵欣胜. 黄河三角洲芦苇湿地的恢复. 生态学报, 2006, 26(8): 2616-2624.
- [ 4 ] 程嘉伟, 邓昶身, 鲁长虎. 苏州太湖湖滨人工种植和原生芦苇湿地鸟类群落. 动物学杂志, 2014, 49(3): 347-356.
- [ 5 ] 邹乐, 李欢, 章家保, 陈加银, 杨华韬, 龚政. 滨海湿地植被地上生物量遥感反演. 生态学报, 2023, 43(20): 8532-8543.
- [ 6 ] Shen X J, Jiang M, Lu X G, Liu X T, Liu B, Zhang J Q, Wang X W, Tong S Z, Lei G C, Wang S Z, Tong C, Fan H Q, Tian K, Wang X L, Hu Y M, Xie Y H, Ma M Y, Zhang S W, Cao C X, Wang Z C. Aboveground biomass and its spatial distribution pattern of herbaceous marsh vegetation in China. *Science China Earth Sciences*, 2021, 64(7): 1115-1125.
- [ 7 ] 田艳林, 刘贤赵, 毛德华, 王宗明, 李延峰, 高长春. 基于 MODIS 数据的松嫩平原西部芦苇湿地地上生物量遥感估算. 生态学报, 2016, 36(24): 8071-8080.
- [ 8 ] Jensen D, Cavanaugh K C, Simard M, Okin G S, Castañeda-Moya E, McCall A, Twilley R R. Integrating imaging spectrometer and synthetic aperture radar data for estimating wetland vegetation aboveground biomass in coastal Louisiana. *Remote Sensing*, 2019, 11(21): 2533.
- [ 9 ] 章莹, 卢剑波. 外来入侵物种互花米草(*Spartina alterniflora*)及凤眼莲(*Eichhornia crassipes*)的遥感监测研究进展. 科技通报, 2010, 26(1): 130-137.
- [ 10 ] Lyu X, Li X B, Gong J R, Li S K, Dou H S, Dang D L, Xuan X J, Wang H. Remote-sensing inversion method for aboveground biomass of typical steppe in Inner Mongolia, China. *Ecological Indicators*, 2021, 120: 106883.
- [ 11 ] Han M, Pan B, Liu Y B, Yu H Z, Liu Y R. Wetland biomass inversion and space differentiation: a case study of the Yellow River Delta Nature Reserve. *PLoS One*, 2019, 14(2): e0210774.
- [ 12 ] 刘润红, 梁士楚, 赵红艳, 漆光超, 李丽香, 姜勇, 牛振国. 中国滨海湿地遥感研究进展. 遥感技术与应用, 2017, 32(6): 998-1011.
- [ 13 ] 戎荣, 孙斌, 武志涛, 高志海, 杜自强, 滕思翰. 灌丛化草原小叶锦鸡儿灌丛地上生物量测量方法研究. 草业学报, 2023, 32(1): 36-47.
- [ 14 ] Kalaeska M, Lalonde M, Moore T R. Estimation of foliar chlorophyll and nitrogen content in an ombrotrophic bog from hyperspectral data: Scaling from leaf to image. *Remote Sensing of Environment*, 2015, 169: 270-279.
- [ 15 ] Chen C, Ma Y, Ren G B, Wang J B. Aboveground biomass of salt-marsh vegetation in coastal wetlands: Sample expansion of in situ hyperspectral and Sentinel-2 data using a generative adversarial network. *Remote Sensing of Environment*, 2022, 270: 112885.
- [ 16 ] Klemas V. Remote sensing of coastal wetland biomass: an overview. *Journal of Coastal Research*, 2013, 290: 1016-1028.
- [ 17 ] González-Jaramillo V, Fries A, Bendix J. AGB estimation in a tropical mountain forest (TMF) by means of RGB and multispectral images using an unmanned aerial vehicle (UAV). *Remote Sensing*, 2019, 11(12): 1413.
- [ 18 ] Zhou R, Yang C, Li E H, Cai X B, Wang X L. Aboveground biomass estimation of wetland vegetation at the species level using unoccupied aerial vehicle RGB imagery. *Frontiers in Plant Science*, 2023, 14: 1181887.
- [ 19 ] 刘琳, 郑兴明, 姜涛, 李雷, 丁艳玲. 无人机遥感植被覆盖度提取方法研究综述. 东北师大学报: 自然科学版, 2021, 53(4): 151-160.
- [ 20 ] 杨正道, 舒清态, 黄金君, 周文武, 胥丽, 罗绍龙, 王书伟. 基于无人机多光谱的高山松地上生物量估测. 广西林业科学, 2024, 53(1): 10-17.
- [ 21 ] Doughty C L, Ambrose R F, Okin G S, Cavanaugh K C. Characterizing spatial variability in coastal wetland biomass across multiple scales using UAV and satellite imagery. *Remote Sensing in Ecology and Conservation*, 2021, 7(3): 411-429.
- [ 22 ] Wan R R, Wang P, Wang X L, Yao X, Dai X. Modeling wetland aboveground biomass in the Poyang Lake National Nature Reserve using machine

- learning algorithms and Landsat-8 imagery. *Journal of Applied Remote Sensing*, 2018, 12: 046029.
- [23] Hansen E, Ene L, Mauya E, Patočka Z, Mikita T, Gobakken T, Næsset E. Comparing empirical and semi-empirical approaches to forest biomass modelling in different biomes using airborne laser scanner data. *Forests*, 2017, 8(5): 170.
- [24] 任晓琦, 侯鹏, 陈妍. 森林地上生物量遥感反演研究进展. *林草资源研究*, 2023(6): 146-158.
- [25] Wu, Bethel, Mishra, Hardy. Model selection in Bayesian framework to identify the best WorldView-2 based vegetation index in predicting green biomass of salt marshes in the northern Gulf of Mexico[J]. *GIScience & Remote Sensing*, 2018, 55(6): 880-904.
- [26] Sharma P, Leigh L, Chang J, Maimaitijiang M, Caffè M. Above-ground biomass estimation in oats using UAV remote sensing and machine learning. *Sensors*, 2022, 22(2): 601.
- [27] Niu X M, Chen B J, Sun W W, Feng T, Yang X D, Liu Y Y, Liu W W, Fu B L. Estimation of coastal wetland vegetation aboveground biomass by integrating UAV and satellite remote sensing data. *Remote Sensing*, 2024, 16(15): 2760.
- [28] 郝君, 吕康婷, 胡天祺, 王云阁, 徐刚. 基于机器学习的红树林生物量遥感反演研究. *林草资源研究*, 2024(1): 65-72.
- [29] 李贵鑫, 安东, 于应文, 沈禹颖. 基于无人机多光谱与纹理特征的饲用大豆地上生物量估算研究. *草地学报*, 2024, 32(10): 3262-3271.
- [30] 杨晓芳, 满卫东, 刘明月, 张永彬, 郑浩, 宋敬茹, 康志强. 浙江滨海湿地互花米草生物量遥感估算模型研究. *遥感技术与应用*, 2023, 38(6): 1445-1454.
- [31] Lu L R, Luo J H, Xin Y H, Duan H T, Sun Z, Qiu Y G, Xiao Q T. How can UAV contribute in satellite-based *Phragmites australis* aboveground biomass estimating *International Journal of Applied Earth Observation and Geoinformation*, 2022, 114: 103024.
- [32] 郭芮, 伏帅, 侯蒙京, 刘洁, 苗春丽, 孟新月, 冯琦胜, 贺金生, 钱大文, 梁天刚. 基于 Sentinel-2 数据的青海门源县天然草地生物量遥感反演研究. *草业学报*, 2023, 32(4): 15-29.
- [33] 宋柯馨, 蒋馥根, 胡宗达, 吕云兵, 龙依, 邓目丽, 陈松, 孙华. 西藏自治区草地上生物量遥感反演研究. *生态学报*, 2023, 43(13): 5600-5613.
- [34] 敖登, 杨佳慧, 丁维婷, 安韶山, 何海龙. 54 种植被指数研究进展综述. *安徽农业科学*, 2023, 51(1): 13-21, 28.
- [35] 左萍萍, 付波霖, 蓝斐菲, 解淑毓, 何宏昌, 范冬林, 娄佩卿. 基于无人机多光谱的沼泽植被识别方法. *中国环境科学*, 2021, 41(5): 2399-2410.
- [36] 李敏, 陈利, 李静泰, 闫丹丹, 刘垚, 吴翠玲, 栾兆擎. 基于 Sentinel-2 数据的互花米草地上生物量反演. *海洋环境科学*, 2024, 43(3): 386-397.
- [37] 刘忠, 万炜, 黄晋宇, 韩已文, 王佳莹. 基于无人机遥感的农作物长势关键参数反演研究进展. *农业工程学报*, 2018, 34(24): 60-71.
- [38] 李亚丹, 杜华强, 周国模, 谷成燕, 徐小军, 孙少波, 高国龙. 雷竹叶绿素与高光谱植被指数关系及其反演模型. *浙江农林大学学报*, 2015, 32(3): 335-345.
- [39] 蒋馥根, 孙华, 李成杰, 马开森, 陈松, 龙江平, 任蓝翔. 联合 GF-6 和 Sentinel-2 红边波段的森林地上生物量反演. *生态学报*, 2021, 41(20): 8222-8236.
- [40] Zhang L J, Zhao Y Y, Chen C, Li X J, Mao F J, Lv L J, Yu J C, Song M X, Huang L, Chen J J, Zheng Z D, Du H Q. UAV-LiDAR integration with sentinel-2 enhances precision in AGB estimation for bamboo forests. *Remote Sensing*, 2024, 16(4): 705.
- [41] Gao Y K, Lu D S, Li G Y, Wang G X, Chen Q, Liu L J, Li D Q. Comparative analysis of modeling algorithms for forest aboveground biomass estimation in a subtropical region. *Remote Sensing*, 2018, 10(4): 627.
- [42] 赵晓伟, 黄杨, 汪永强, 储鼎. 基于无人机多光谱数据的玉米苗株估算. *自然资源遥感*, 2022(1): 106-114.
- [43] Fang H, Man W D, Liu M Y, Zhang Y B, Chen X T, Li X, He J N, Tian D. Leaf area index inversion of *Spartina alterniflora* using UAV hyperspectral data based on multiple optimized machine learning algorithms. *Remote Sensing*, 2023, 15(18): 4465.
- [44] 王来刚, 贺佳, 郑国清, 郭燕, 张彦, 张红利. 基于无人机多光谱遥感的玉米 FPAR 估算. *农业机械学报*, 2022, 53(10): 202-210.
- [45] Dong J W, Xiao X M, Wagle P, Zhang G L, Zhou Y T, Jin C, Torn M S, Meyers T P, Suyker A E, Wang J B, Yan H M, Biradar C, Moore B. Comparison of four EVI-based models for estimating gross primary production of maize and soybean croplands and tallgrass prairie under severe drought. *Remote Sensing of Environment*, 2015, 162: 154-168.
- [46] 李淑贞, 徐大伟, 范凯凯, 陈金强, 佟旭泽, 辛晓平, 王旭. 基于无人机与卫星遥感的草原地上生物量反演研究. *遥感技术与应用*, 2022, 37(1): 272-278.
- [47] 赵翊含, 侯蒙京, 冯琦胜, 高宏元, 梁天刚, 贺金生, 钱大文. 基于 Landsat 8 和随机森林的青海门源天然草地地上生物量遥感估算. *草业学报*, 2022, 31(7): 1-14.
- [48] 谭雨欣, 田义超, 黄卓梅, 张强, 陶进, 刘虹秀, 杨永伟, 张亚丽, 林俊良, 邓静雯. 北部湾茅尾海无瓣海桑红树林地上生物量反演——基于 XGBoost 机器学习算法. *生态学报*, 2023, 43(11): 4674-4688.
- [49] 贾元, 张琳, 吴冬秀, 宋创业, 袁伟影, 李凌浩. 基于无人机多光谱实测数据的草地生物量反演模型比较. *生态学报*, 2024, 44(15): 6854-6864.