

DOI: 10.20103/j.stxb.202310162250

田国帅,周小成,郝优壮,谭芳林,王永荣,吴善群,林华章.结合修正后的全球生态系统动态调查冠层高度的森林地上生物量模型优化——以福建省为例.生态学报,2024,44(16):7264-7277.

Tian G S, Zhou X C, Hao Y Z, Tan F L, Wang Y R, Wu S Q, Lin H Z. Optimization model of forest aboveground biomass based on MGEDI canopy height: a case study in Fujian, China. Acta Ecologica Sinica, 2024, 44(16): 7264-7277.

# 结合修正后的全球生态系统动态调查冠层高度的森林地上生物量模型优化

——以福建省为例

田国帅<sup>1</sup>, 周小成<sup>1,\*</sup>, 郝优壮<sup>1</sup>, 谭芳林<sup>2</sup>, 王永荣<sup>3</sup>, 吴善群<sup>4</sup>, 林华章<sup>5</sup>

1 福州大学卫星空间信息技术综合应用国家地方联合工程研究中心, 空间数据挖掘与信息共享教育部重点实验室, 福州 350108

2 福建省林业科学研究院, 福州 350012

3 福建省漳平五一国有林场, 龙岩 364400

4 福建省大田县林业局, 三明 366100

5 福建省大田桃源国有林场, 三明 366199

**摘要:**森林地上生物量(Above Ground Biomass, AGB)是衡量森林生态系统碳存储、能量流动和生物多样性的关键指标,对于气候变化研究和森林资源管理至关重要。福建省地处多云多雨的亚热带,地形和森林类型复杂,森林地上生物量估算难度大。为提升森林地上生物量估算效果,将最新星载激光雷达数据全球生态系统动态调查(GEDI)、Landsat 以及 Sentinel 系列卫星等多源遥感数据进行集成和综合利用,通过 Landsat 影像计算的林龄对 GEDI\_V27 冠层高度产品进行优化,结合优化后的 MGEDI\_V27 冠层高度产品,建立传统遥感特征结合冠层高度的极端梯度提升模型(XGBoost)生物量反演模型,实现了福建省森林地上生物量的有效估算与制图。研究结果表明:(1)通过林龄优化后的 GEDI 冠层高度精度评价结果为  $R^2 = 0.67$ ,  $RMSE = 2.24m$ ;(2)通过递归特征消除算法对三种森林类型进行特征优选,得到 10 个遥感特征,其中,三种森林类型最重要的遥感特征均为森林冠层高度,并且对比评价了在包含传统遥感特征因子的情况下有无冠层高度对于模型精度的影响,结果表明,在冠层高度因子参加特征构建时,森林 AGB 回归分析的精度明显提高,证实了冠层高度在生物量估算中具有显著的重要性;(3)研究得到的福建省森林 AGB 范围为  $0.001\text{—}363.331\text{Mg}/\text{hm}^2$ ,整体精度评价结果为  $R^2 = 0.75$ ,  $RMSE = 17.34\text{Mg}/\text{hm}^2$ ,2020 年全省 AGB 总量为 8.22 亿 Mg,平均值为  $101.24\text{Mg}/\text{hm}^2$ 。通过优化 GEDI 中的森林冠层高度,并且结合传统遥感特征,可以实现对福建省森林地上生物量的精确估算和监测,研究成果有助于区域森林碳汇的评估。

**关键词:**遥感;全球生态系统动态调查(GEDI);冠层高度;森林类型;极端梯度提升模型(XGBoost)回归;森林地上生物量

## Optimization model of forest aboveground biomass based on MGEDI canopy height: a case study in Fujian, China

TIAN Guoshuai<sup>1</sup>, ZHOU Xiaocheng<sup>1,\*</sup>, HAO Youzhuang<sup>1</sup>, TAN Fanglin<sup>2</sup>, WANG Yongrong<sup>3</sup>, WU Shanqun<sup>4</sup>, LIN Huazhang<sup>5</sup>

1 Key Laboratory of Spatial Data Mining and Information Sharing of Ministry of Education, National & Local Joint Engineering Research Center of Satellite Geospatial Information Technology, Fuzhou University, Fuzhou 350108, China

2 Fujian Academy of Forestry, Fuzhou 350012, China

3 Fujian Zhangping Wuyi State-owned Forest Farm, Longyan 364400, China

**基金项目:**福建省科技厅高校产学研合作项目(2022N5008);福建省科技厅对外合作项目(202210007)

**收稿日期:**2023-10-16; **网络出版日期:**2024-06-18

\* 通讯作者 Corresponding author. E-mail: zhouxc@fzu.edu.cn

4 Fujian Datian forestry bureau, Sanming 366100, China

5 Fujian Datian Taoyuan State-owned Forest Farm, Sanming 366199, China

**Abstract:** Above Ground Biomass (AGB) is a key indicator of forest ecosystem carbon storage, energy flow changes and biodiversity, and is crucial for climate change research and forest resource management. Fujian Province, as the largest collective forest area in southern China, has abundant forest resources, accurately estimating forest aboveground biomass can lay the foundation for estimating carbon storage and provide decision-making support for achieving the dual carbon goals. Fujian Province is located in a cloudy and rainy subtropical zone with complex terrain and forest types, making it difficult to estimate forest aboveground biomass, estimating forest aboveground biomass using traditional methods is difficult to meet accuracy requirements. In order to improve the accuracy of aboveground forest biomass, this study integrated and comprehensively utilized multi-source remote sensing data such as the latest spaceborne lidar data GEDI, Landsat and Sentinel series satellites. Above all, the GEDI\_V27 canopy height product was optimized based on the forest age calculated from Landsat. Then combined with the optimized MGEDI\_V27 canopy height product, by establishing an XGBoost biomass inversion model that combined traditional remote sensing features with canopy height, we effectively improved model accuracy, estimated and mapped the aboveground biomass of forests in Fujian Province. The research results showed that: (1) The GEDI canopy height accuracy evaluation result optimized by forest age was  $R^2 = 0.67$ ,  $RMSE = 2.24m$ ; (2) The recursive feature elimination algorithm was used to optimize the features of the three forest types, and 10 remote sensing features were obtained. Among them, the most important remote sensing features of the three forest types were forest canopy height, and a comparative evaluation was performed on the features including traditional remote sensing features. The results showed that when the canopy height factor was included in the feature construction, the accuracy of the forest AGB regression analysis was significantly improved, confirming that canopy height played a significant role in biomass estimation; (3) The studied forest AGB range in Fujian Province was 0.001—363.331Mg/hm<sup>2</sup>, the overall accuracy evaluation result was  $R^2 = 0.75$ ,  $RMSE = 17.34 Mg/hm^2$ , and the total AGB amount in the province in 2020 was 822 million Mg. The average value was 101.24Mg/hm<sup>2</sup>, reflecting the good ecological quality of Fujian Province. By optimizing the forest canopy height in GEDI and combining it with traditional remote sensing features, the accuracy of forest aboveground biomass modeling can be significantly improved, and it is possible to accurately estimate and monitor forest biomass in Fujian Province. The research results are helpful for the high-precision estimation of aboveground biomass in regional forests, and have certain guiding significance for the assessment of carbon sinks.

**Key Words:** remote sensing; global ecosystem dynamics investigation (GEDI); canopy height; forest type; extreme gradient boosting(XGBoost) regression; forest above ground biomass

森林地上生物量(Above Ground Biomass, AGB)包括树干、树枝和树叶生物量,是评估森林在全球碳循环中的作用和分析生态系统生产力的重要变量,高精度的森林 AGB 空间分布数据可直接评估森林生态系统的碳储量,并应用于全球碳收支监测和管理<sup>[1]</sup>。因此,快速获取大范围森林 AGB 信息,分析森林碳汇空间格局及潜力,可以为各级政府的林业碳汇交易、森林碳中和动态监测评估提供科学合理的决策支持。

传统的森林 AGB 估算是通过实地野外调查森林结构参数,需要大量的人力资源,成本高,效率低。遥感技术具有高时效性、宏观监测等优势,可以弥补传统调查方法的不足,当前森林 AGB 建模和估算使用了目前所有能用到的包括机载或星载的可见光<sup>[2-3]</sup>、微波雷达<sup>[4]</sup>和激光雷达<sup>[5-6]</sup>等多源数据。其中,基于传统光学遥感的森林 AGB 估算发展较早,利用植被反射光谱反映叶绿素含量和植被生长状况,构建生物量反演模型<sup>[7-8]</sup>。光学遥感影像提取的预测因子包括反射率、波段计算指数、生物物理变量、图像变换分量以及高分辨率影像纹理特征等<sup>[9-11]</sup>,Moradi 等<sup>[2]</sup>以伊朗北部森林为研究区,提取 Sentinel-2 数据的光谱特征以及植被指数,通过多元回归,人工神经网络,k-最近邻以及随机森林模型进行生物量估算,结果表明,人工神经网络估计生物量的性能最佳,RRMSE = 19.9%。微波遥感技术不受天气影响且具有穿透树冠能力,能与森林生态系统各层林结构和功能发生不同的作用,包含了森林生态系统各层林结构和功能等综合信息<sup>[12]</sup>。基于微波遥感获取的遥感因子主要为后向散射系数和干涉信息,如干涉相位和相干系数,后者多用于反演冠层高度和地表

高程,这两个参数常作为有效预测因子参与森林生物量建模与估算<sup>[13-14]</sup>,Kumar 等<sup>[14]</sup>使用 PALSAR 数据对印度 Dudhwa 国家公园进行 AGB 反演,通过线性回归完成建模,其结果显示  $R^2$  为 0.78, RMSE = 59.77Mg/hm<sup>2</sup>。光学遥感和微波遥感各有优缺点,在一定程度上可以互补。因此,已有研究结合两种数据,提取光谱、纹理、雷达等特征,运用多元回归分析和机器学习方法建立森林 AGB 反演模型,如 Su 等<sup>[15]</sup>通过结合 Landsat 时序影像以及 PALSAR 数据,以广东省为研究区进行 AGB 反演建模,结果表明该模型可有效进行森林 AGB 制图,其  $R^2$  为 0.81, RMSE = 6.54Mg/hm<sup>2</sup>。虽然光学遥感与微波遥感可以实现森林 AGB 的估算,但参数容易达到饱和,难以提升精度<sup>[16]</sup>。激光雷达以其独特的优势可弥补不足,获取与森林 AGB 密切相关的垂直结构信息。机载 LiDAR 点云可以提取冠层高度、冠幅等参数,引入传统的异速方程<sup>[17]</sup>、卷积神经网络模型<sup>[18]</sup>或利用回归模型<sup>[19]</sup>构建森林 AGB 估算模型,但覆盖范围有限且成本较高。大范围研究需要结合星载 LiDAR 和其他遥感数据,提取冠层高度、郁闭度等特征进行估算。新一代星载 LiDAR 如 GEDI 和 ICESat-2 因其可提供的冠层高度数据为研究森林 AGB 估算提供了更多潜力<sup>[20-21]</sup>,Qi 等<sup>[22]</sup>模拟了 GEDI 的激光雷达观测,结合 TanDEM-X 数据对三种森林类型进行了中等分辨率下的森林生物量估计,结果表明  $R^2$  可达到 0.87, RMSE 为 24.6Mg/hm<sup>2</sup>。在当前数据源愈加丰富的情况下,有研究开展了将 LiDAR 数据获取的冠层高度与传统数据结合进行 AGB 建模的研究,Yue 等<sup>[23]</sup>研究发现,若把冠层高度参与到 AGB 建模中,精度可取得有效提升, $R^2$  从 0.53 提升至 0.78, RMSE 从 1.62Mg/hm<sup>2</sup> 降至 1.08Mg/hm<sup>2</sup>,证实结合冠层高度进行森林 AGB 的确实可以提高估算精度。

福建省作为中国南方最大集体林区,森林覆盖率全国第一,准确估算森林 AGB 对于生态系统、碳汇和生物多样性保护具有重要意义<sup>[24]</sup>。但福建省地形以山地为主,地势复杂多变,造成了显著的环境梯度差异,对森林 AGB 估算带来了一定难度。本文以福建省为研究区,利用多源遥感数据以及改进后的冠层高度产品建立森林 AGB 高精度估算模型,实现福建省 2020 年森林 AGB 高分辨率估算。研究成果将反映福建省森林 AGB 的空间分布格局,为森林资源管理、碳汇计算和生态保护提供科学依据。

## 1 研究区域及数据

### 1.1 研究区域

福建省地处中国东南沿海地区,地理位置为 115°50'E—120°40'E, 23°33'N—28°20'N (图 1),属亚热带海洋性季风气候,地形复杂,山地丘陵占全省面积 80% 以上。全省面积约 12.4 万 km<sup>2</sup>。福建省雨水充沛,森林资源十分丰富,是中国南方重点集体林区,森林覆盖率达 65% 以上,稳居全国第一。由于良好的气候条件,福建省森林以常绿森林为主,树木生长速度快,物种丰富,森林生态系统健康。

### 1.2 数据

#### 1.2.1 多源遥感数据

本文所用的多源遥感数据主要有光学遥感影像、合成孔径雷达影像、无人机数据以及星载 LiDAR 数据。光学遥感影像(表 1)为长时序 Landsat 系列卫星影像和 Sentinel-2 卫星影像;合成孔径雷达数据(表 2)为 Sentinel-1 卫星数据;无人机数据包括可见光数据和机载 LiDAR 数据。其中光学遥感影像和 SAR 影像主要提供光谱、纹理、雷达因子等信息,其预处理步骤均在 GEE 云平台完成。

表 1 可见光卫星参数

Table 1 Optical satellite parameters

传感器 Sensor	影像数据集 ID Image dataset ID	影像时间范围 Image time range	空间分辨率/m Spatial resolution
Landsat 8 OLI	LNADSAT/LC08/C01/T1_SR	2013-04-11—2020-11-01	30
Landsat 7 ETM+	LNADSAT/LE07/C01/T1_SR	2000-01-01—2013-01-15	30
Landsat 5 TM	LNADSAT/LT05/C01/T1_SR	1986-01-01—2001-01-01	30
Sentinel-2	COPERNICUS/S2	2020-06-01—2020-09-01	10

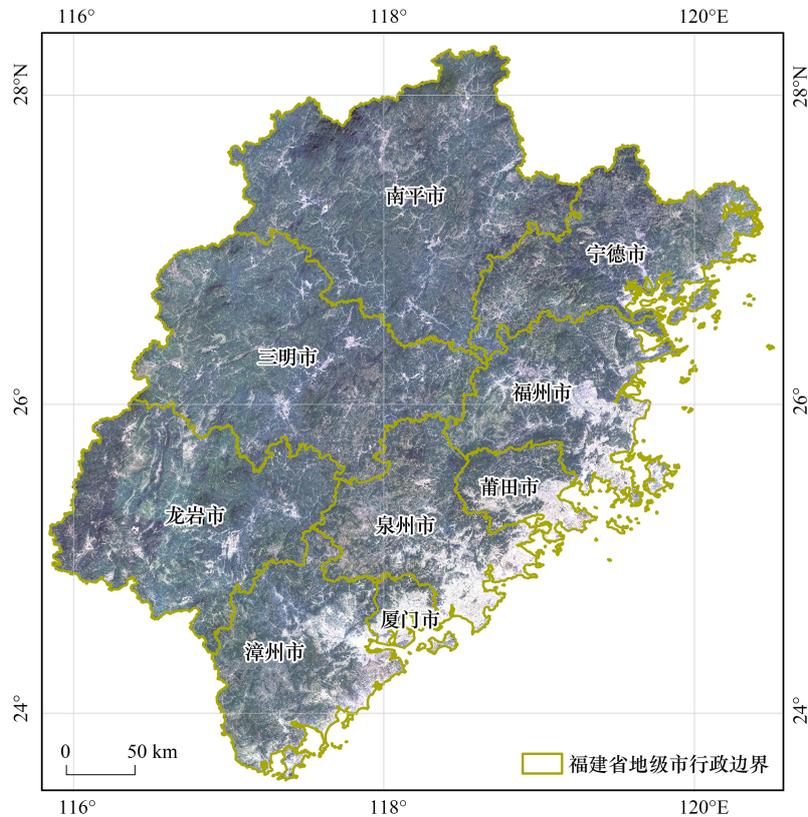


图 1 研究区

Fig.1 Study area

表 2 雷达卫星参数

Table 2 Radar satellite parameters

成像模式 Acquisition mode	幅宽 Width	极化方式 Polarization	空间分辨率 Spatial resolution
条带模式 Stripmap	80km	HH+HV、HH、VH+VV、VV	5m×5m
干涉宽幅模式 Interferometric Wide swath	250km	HH+HV、HH、VH+VV、VV	5m×20m
波模式 Wave	20km×20km	HH/VV	5m×5m
超宽幅模式 Extra Wide swath	400km	HH+HV、HH、VH+VV、VV	20m×40m

HH:水平水平;HV:水平垂直;VH:垂直水平;VV:垂直垂直

### 1.2.2 GEDI 数据

GEDI 卫星<sup>[38]</sup>自 2018 年 12 月发射以来,以高重复频率和 25m 直径的光斑采集数据,如图 2 所示。GEDI 对地观测覆盖范围为 51.6°S—51.6°N,福建省位于其监测区域。本文采用的为 GEDI\_2A 级数据产品及其衍生的冠层高度产品 GEDI\_V27。

### 1.2.3 辅助数据

本研究中使用的辅助数据主要包括地形数据以及森林类型数据,其中福建省地形数据为空间分辨率 30m 的 SRTMGL1\_003 数字产品。森林类型数据是基于中国科学院研发的全球 30m 地表覆盖精细分类产品(GLC\_FCS30-2020)<sup>[5]</sup>,并利用 2020 年的 10m 分辨率 Sentinel 卫星影像、重点区域 1—2m 中国高分系列卫星遥感数据,结合样地数据和实地勘察数据,进行验证和修正,获得包括针叶林、阔叶林、混交林的福建省 2020 年森林类型图(图 3),选取小班数据以及伐区实测数据共计 18.6 万个验证森林类型精度,三种森林类型的平均精度为 66.58%。

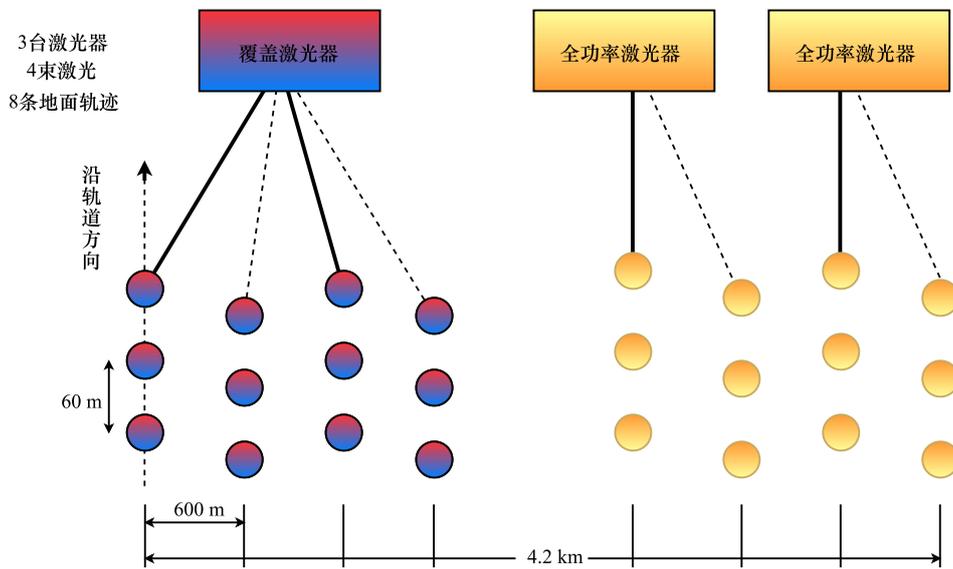


图2 全球生态系统动力学调查(GEDI)轨道参数

Fig.2 Global ecosystem dynamics investigation (GEDI) orbital parameters

#### 1.2.4 样本数据

样本数据(图4)主要是用来对森林地上生物量进行建模,主要包括无人机数据、森林资源二类调查小班数据、伐区设计调查数据以及野外实测数据集;

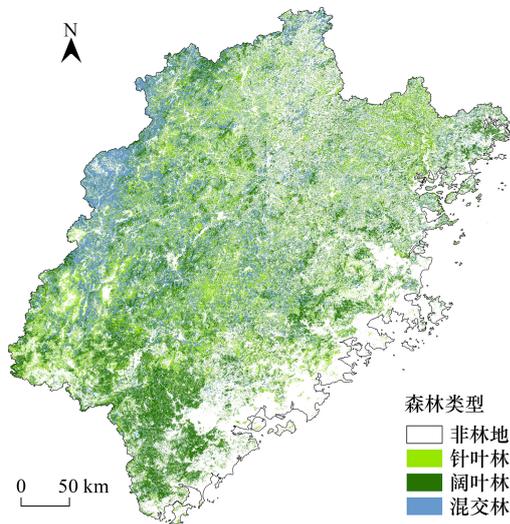


图3 福建省森林类型分布

Fig.3 Distribution of forest types in Fujian Province

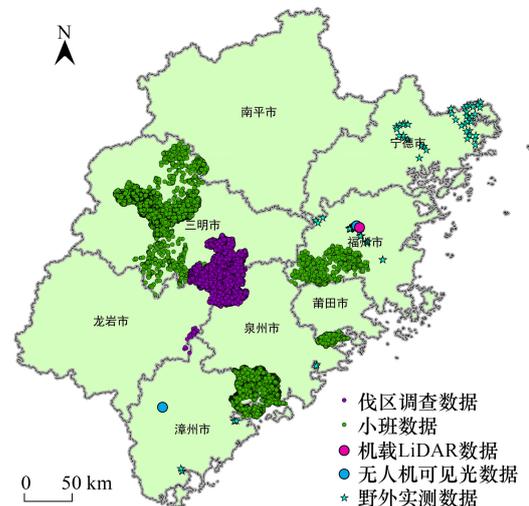


图4 样本点分布

Fig.4 Sample point distribution

选择间接制作生物量样本集,通过小班数据和伐区调查设计数据中的平均冠层高度以及平均胸径,根据树木异速方程估算出树木的地上生物量,样本个数共计91015个,其中训练集和验证集按照8:2划分。参考我国南方森林异速生长的研究总结<sup>[25]</sup>,在其基础上加入实测数据及伐区调查数据共计8.6万个,得到了福建省针叶林和阔叶林的生物量计算异速方程,混交林取其平均值。这三种森林类型的异速方程是以冠层高度、胸径为因子估算树木的叶子、树枝以及树干的生物量,具体公式如下所示:

$$W_1 = e^{-3.446+0.930 \times \ln(D^2H)} + e^{-6.501+1.044 \times \ln(D^2H)} + e^{-6.974+0.956 \times \ln(D^2H)} \quad (1)$$

$$W_2 = e^{-3.505+0.984 \times \ln(D^2H)} + e^{-2.335+0.74 \times \ln(D^2H)} + e^{-4.258+0.833 \times \ln(D^2H)} \quad (2)$$

$$W_3 = e^{-3.476+0.957 \times \ln(D^2H)} + e^{-4.418+0.892 \times \ln(D^2H)} + e^{-5.616+0.895 \times \ln(D^2H)} \quad (3)$$

式中,  $W_1$  为针叶林地上生物量,  $W_2$  为阔叶林地上生物量,  $W_3$  为混交林地上生物量;  $D$  为树木的胸径,  $H$  为冠层高度。

## 2 研究方法

研究的总体路线如图 5 所示。首先在 GEE 云平台对遥感影像进行预处理之后使用 LandTrendr<sup>[26]</sup> 森林干扰与恢复监测算法将研究区划分为森林干扰区(即时序遥感时间区间内, LandTrendr 算法检测到因受自然灾害或人为干扰森林状态发生改变的区域)和非干扰区(即时序遥感时间区间内, LandTrendr 算法检测到森林未受干扰呈现持续森林特征的区域), 根据森林恢复时间节点计算干扰区森林林龄, 非干扰区的森林年龄标记为大于干扰区的最大林龄, 之后使用海量森林小班数据构建福建省的平均森林冠层高度与森林年龄之间的拟合曲线, 结合计算的林龄获得福建省冠层高度, 并与原始的 GEDI\_V27 产品进行融合改进, 获得优化后的福建省冠层高度产品, 最终结合 Sentinel 影像提取的各类特征, 经特征优选后使用 XGBoost 模型构建森林 AGB 估算模型, 实现福建省 2020 年森林 AGB 的估算与分布制图。

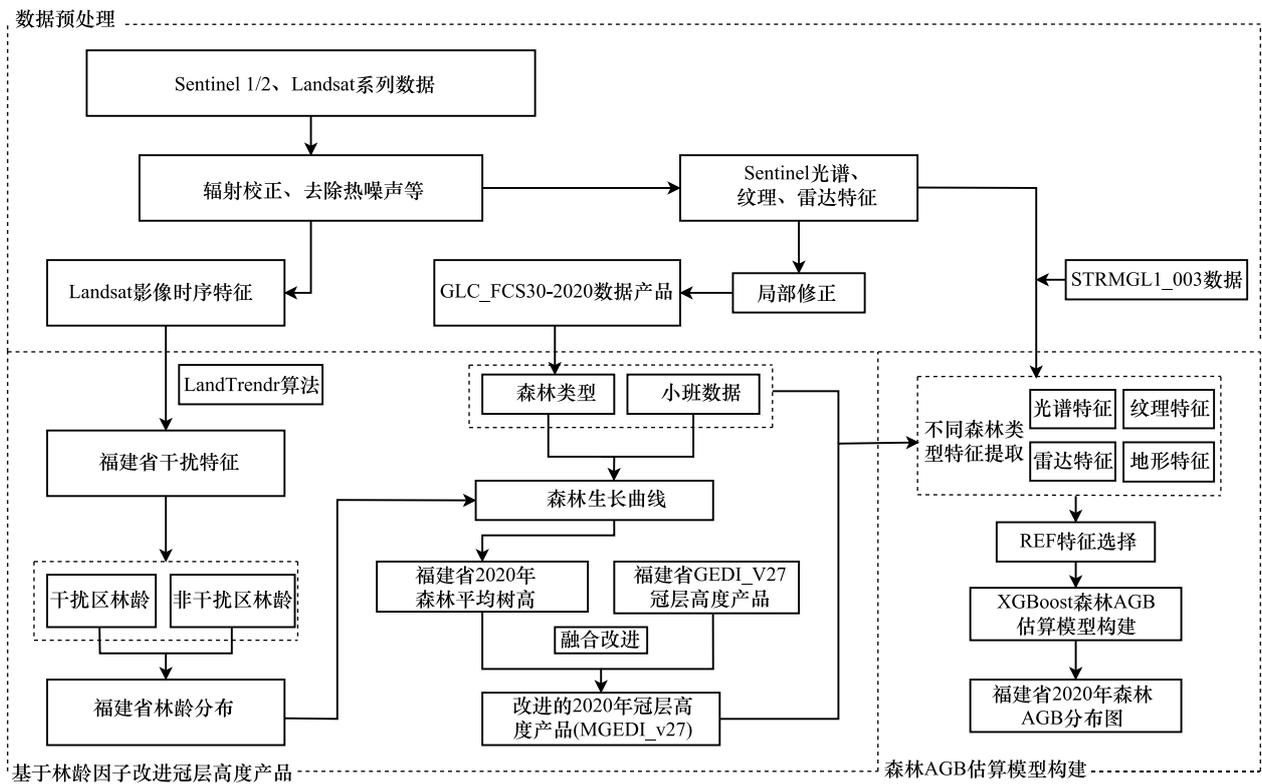


图 5 福建省 AGB 估算流程图

Fig.5 Flow chart of AGB estimation in Fujian Province

REF:递归特征消除;AGB:地上生物量

### 2.1 基于林龄因子改进冠层高度产品

采用了 LandTrendr 森林扰动检测算法,对 1986—2020 年的 Landsat 系列卫星数据进行分析,以获得福建省 2020 年的林龄数据。使用卫星时序数据来检测林地的历史干扰,包括采伐、火烧和造林等,以估算现有林分的林龄。这个方法不仅能够确定干扰发生的年份,还能识别干扰的持续时间,这有助于更准确地估算林龄。为了获得福建省不同类型的森林(针叶林、阔叶林和混交林)的生长曲线,利用 61414 个森林小班数据,其中

包括平均冠层高度、林龄和森林类型等信息。通过统计分析和曲线拟合,我们得到了每个年龄段对应的平均冠层高度数据,并且发现福建森林在相同林龄时的平均树高波动范围约 2 米,只有一小部分会因特殊树种或地形而产生较大的误差。最后,我们将林龄因子和森林类型生长曲线结合,生成了 2020 年福建省 30m 分辨率的平均冠层高度分布图。然后,将这个分布图与 GEDI\_V27 数据相结合,形成了改进的冠层高度模型 MGEDI\_V27。在这个过程中,设置了一个 2 米的阈值,用于判断何时使用平均冠层高度数据来改进 GEDI\_V27 冠层高度。在未受干扰的林区,当 GEDI\_V27 高度值大于该阈值时,选择 GEDI\_V27 值作为 MGEDI\_V27 的冠层高度;在森林干扰区,如果平均冠层高度与 GEDI\_V27 高度值之差大于 2 米,我们使用平均冠层高度值。通过这种方法,能够更精确地估算福建省的冠层高度产品(图 6)。

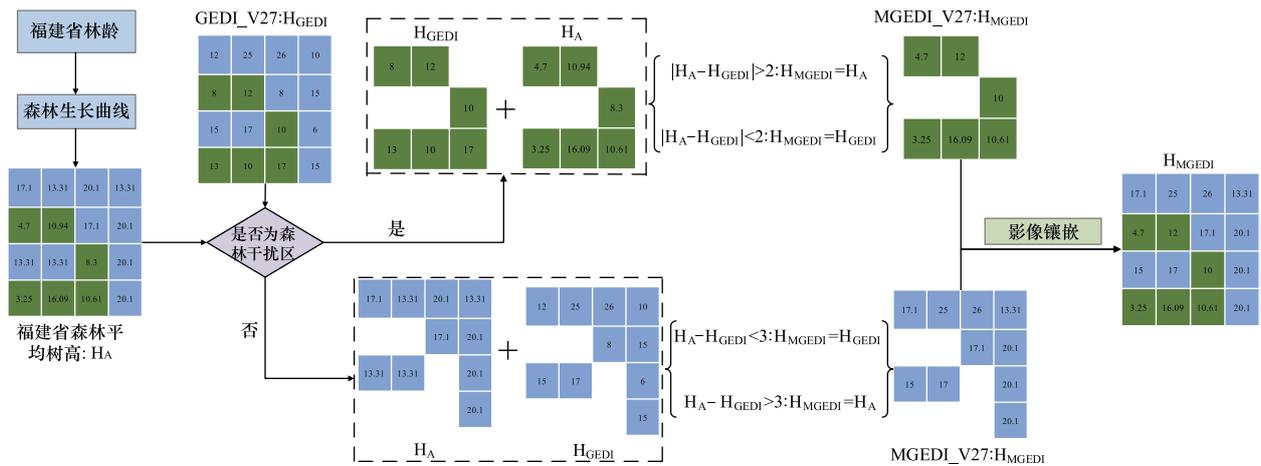


图 6 GEDI\_V27 冠层高度产品改进策略

Fig.6 GEDI\_V27 Canopy Height Product Improvement Strategy

$H_{GEDI}$ : GEDI\_V27 产品获取的冠层高度;  $H_A$ : 森林生长曲线获取的平均树高;  $H_{MGEDI}$ : 修正后的冠层高度; MGEDI\_V27: 改进后的冠层高度产品

## 2.2 结合优化后冠层高度的福建省森林 AGB 估算模型构建

### 2.2.1 遥感影像特征构建

遥感特征提取在定量遥感精确性方面具有关键作用。不同特征侧重不同方面,例如植被指数反映森林覆盖和生长状况,地形特征揭示生长环境,纹理特征反映森林密度等。因此,在森林 AGB 估算中需要综合多种特征。在传统的遥感特征因子,例如光谱、纹理、地形和雷达遥感因子的基础上,进一步结合冠层高度构建特征集以提高生物量反演的准确性,构建的特征集如表 3 所示。

表 3 遥感特征集

Table 3 Remote sensing feature set

数据源 Data source	遥感特征 Features	特征名称 Feature name	特征数目 Feature number
Sentinel-2	光谱特征	B1, B2, B3, B4, B5, B6, B7, B8, B8A, B9, B10, B11, B12, NDVI, EVI, RVI, DVI, LSWI, GNDVI, TVI, SAVI, NDWI, MSAVI, GCVI, Hue, Saturation, Value	27
	纹理特征	NIR 近红外波段的 16 个纹理特征	16
Sentinel-1	雷达特征	VV, VH, VV_stdDev, VH_stdDev, VV 和 VH 的 20 个纹理特征	24
SRTM	地形特征	Elevation, Slope, Aspect, Hillshade	4
GEDI	冠层高度特征	MGEDI_V27	1
总计			72

B1—B12 为 Sentinel 卫星采集的 13 个原始波段; NDVI: 归一化植被指数; EVI: 增强型植被指数; RVI: 比值植被指数; DVI: 差值植被指数; LSWI: 地表植被指数; GNDVI: 绿光归一化差值植被指数; TVI: 三角植被指数; SAVI: 土壤调节植被指数; NDWI: 归一化差分水体指数; MSAVI: 修正型土壤调节植被指数; GCVI: 叶绿素指数

为了避免各特征值之间取值的大小差别过大造成的误差,采用了 python 语言中 scikit-learn 库的预处理库对该特征集进行标准化处理,使得特征值统一分布在 [0, 1] 之间。

### 2.2. 特征优选算法

遥感特征优选是构建森林 AGB 估算的重要步骤之一。在研究过程中文提取了众多与森林 AGB 相关联的变量。然而,同时将这些变量构建估算模型会导致特征之间的冗余,影响模型训练的效率,并增加模型的复杂性。此外,特征优选还可以减少特征集数量,使得机器学习模型更易于解释,减少维度并且避免过拟合。本文选取递归特征消除算法 (Recursive Feature Elimination, RFE)<sup>[27]</sup> 进行特征优选,其中底层模型采用随机森林,通过交叉验证的抽样方法进行特征优选。采用 scikit-learn 库中的特征选择 RFE 库进行特征优选。

### 2.3 精度评价标准

采用不同的机器学习模型构建了森林 AGB 估算模型,并利用测试集与验证集测试了模型的精度,以评估模型的准确性。计算了两个统计参数:决定系数  $R^2$  与均方根误差 (Root Mean Squared Error, RMSE)。

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{4}$$

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (x_i - y_i)^2} \tag{5}$$

式中,  $x$  表示模型预测的森林 AGB,  $y$  是样本数据中的森林 AGB,  $n$  是样本个数,  $\bar{y}$  是样本数据集中的森林 AGB 的平均值。

## 3 数据结果处理与分析

### 3.1 福建省林龄估算结果

根据森林受到干扰的时间、干扰时长结合森林恢复的时间推算出林分林龄信息。因此基于干扰区特征,可以计算出福建省干扰区的具体林龄。森林干扰区域内,时间跨度有 34 年,从检测到森林变化开始,得到了最大林龄为 32 年。而非干扰区的树龄以干扰区的最大林龄也就是 32 年为最低阈值,最终绘制了 2020 年福建省森林林龄空间分布图并进行了精度验证 (图 7)。

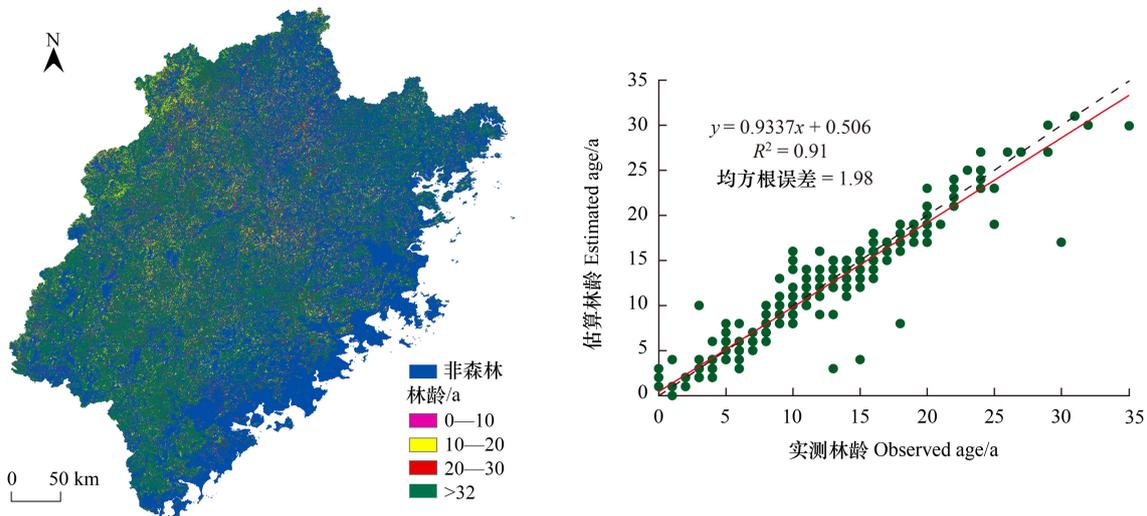


图 7 2020 年福建省森林林龄空间分布图及精度验证

Fig.7 Spatial distribution map and accuracy verification of forest age in Fujian Province in 2020

### 3.2 林龄与冠层高度平均曲线拟合

为了获得福建省三种森林类型(针叶林、阔叶林和混交林)的生长曲线,使用大量森林小班数据,共 61414 个林分斑块。这些林分斑块包括平均冠层高度、林龄以及森林类型等林分参数。本文对该数据集进行了统计分析,并绘制了与林分斑块中每个年龄相对应的平均冠层高度数据的拟合曲线(图 8)。

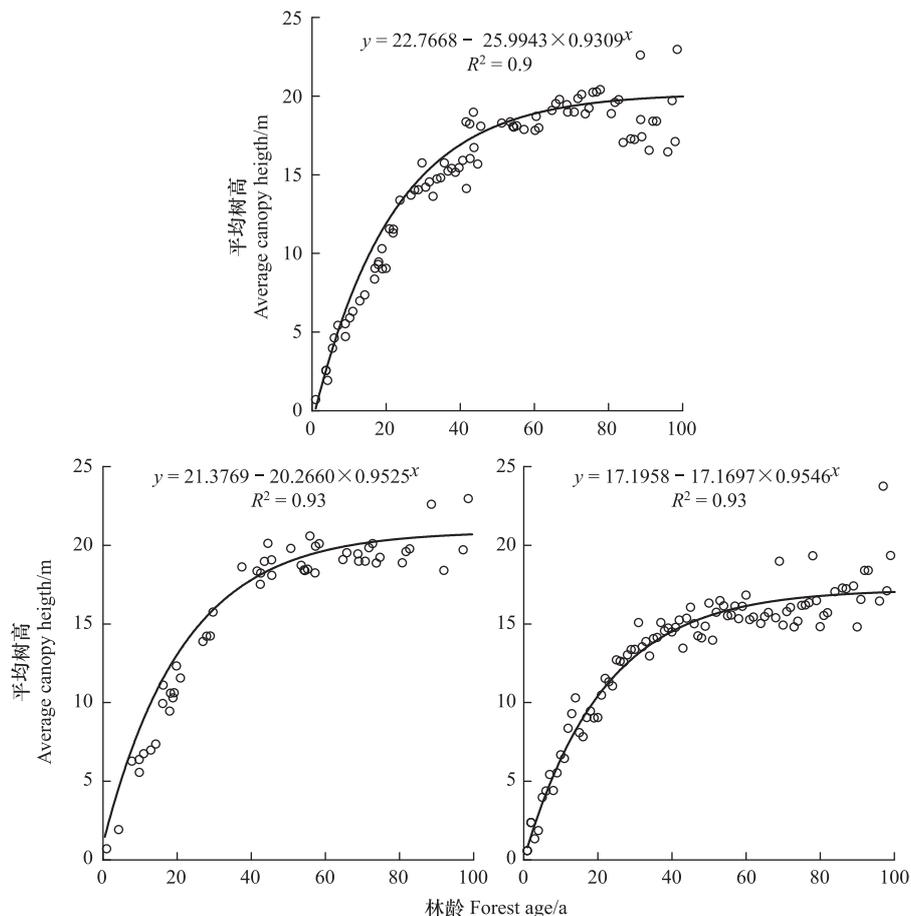


图 8 平均冠层高度数据拟合曲线

Fig.8 Average canopy height data fitting curve

### 3.3 福建省 GEDI 冠层高度产品优化

为了评估本研究使用的冠层高度产品改进方法的有效性,使用若干样本点对比了 GEDI\_V27 和 MGEDI\_V27 冠层高度产品的准确性并绘制了优化后的福建省冠层高度分布(图 9)。评估结果中, GEDI\_V27 冠层高度模型在福建省区域的准确度较低,与样本集的一致性较差,而 MGEDI\_V27 更加集中于实测冠层高度附近,这意味着 MGEDI\_V27 来自外部因素引起的干扰减少了,更加接近实际的森林高度,即福建省森林冠层高度的精度得到有效提高。将林龄估算的平均冠层高度与 GEDI\_V27 冠高模型合并,得到了空间分辨率为 30m 的福建省 2020 年冠层高度 MGEDI\_V27 空间分布图。

### 3.4 遥感影像特征优选结果

通过 RFE 算法优化遥感特征集,最终选出 10 个关键遥感特征(图 10)用于构建福建省森林地上生物量估算模型。三种类型森林优选特征集中最重要的特征都是冠层高度,表明冠层高度是与树木生物量相关性最高的因素。此外,不同的森林类型在特征排序方面存在差异,而地面高程也在特征重要性排序中占有重要地位,这可能是由于不同海拔高度的森林受环境因素影响较大,从而与森林生物量存在较高的相关性。

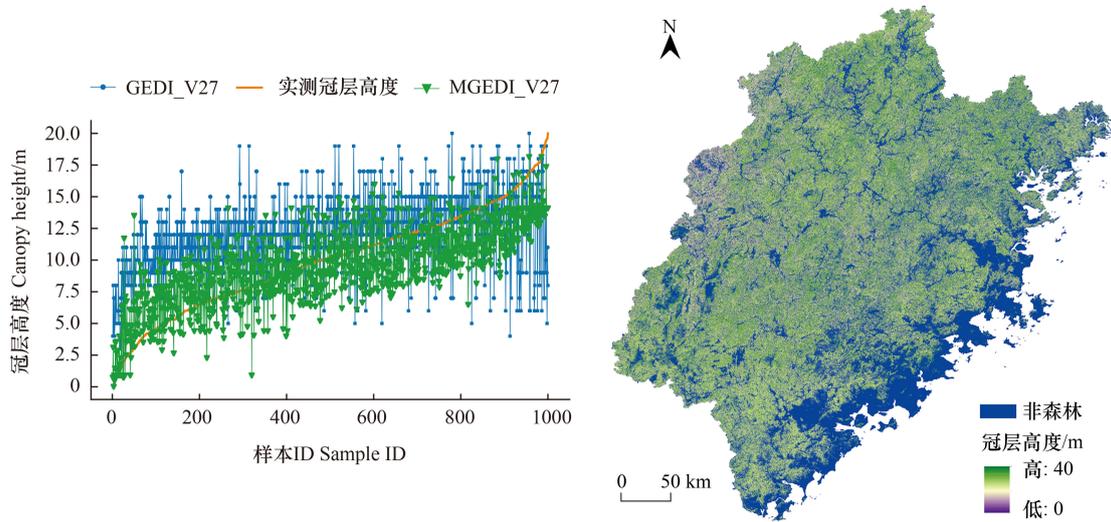


图 9 改进前后的 GEDI 冠层高度对比及改进后冠层高度分布

Fig.9 Comparison of GEDI canopy height before and after improvement and canopy height distribution after improvement

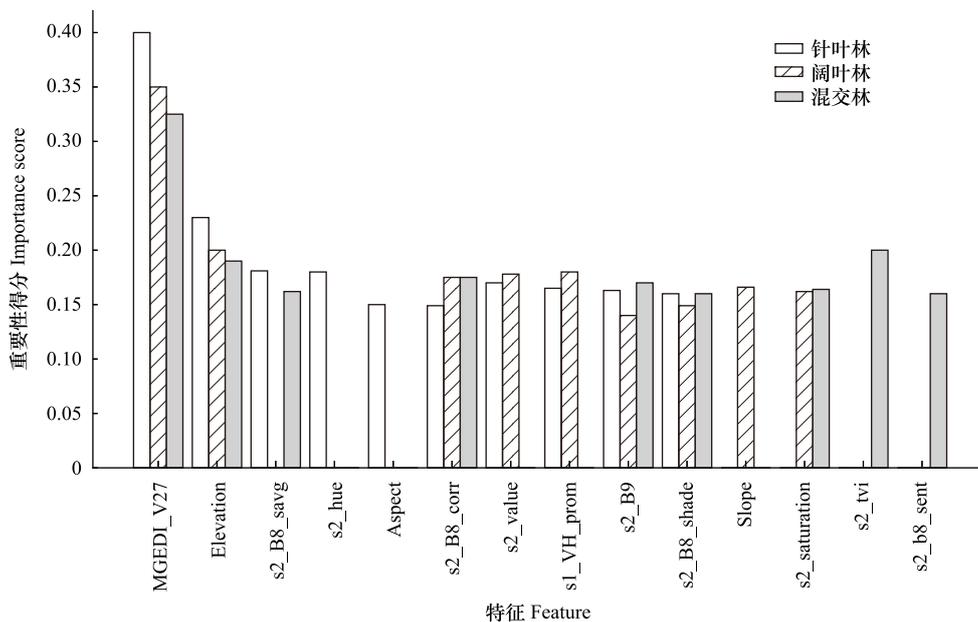


图 10 特征优选结果

Fig.10 Feature optimization results

MGEDI\_V27: 优化后的冠层高度; Elevation: 高程; s2\_B8\_savg: B8 的和平均; s2\_hue: 色相; Aspect: 坡向; s2\_B8\_corr: B8 的相关性; s2\_value: 亮度; s1\_VH\_prom: VH 的聚类突; s2\_B9: B9 波段; s2\_B8\_shade: B8 的聚类荫; Slope: 坡度; s2\_saturation: 饱和度; s2\_tvi: 三角植被指数; s2\_B8\_sent: B8 的和熵

### 3.5 结合优化冠层高度信息的森林 AGB 估算模型精度对比分析

使用 XGBoost 模型建立了森林 AGB 回归模型, 比较了使用传统遥感特征与结合冠层高度分布产品的模型精度。结果如图 11 所示, 只使用传统遥感特征时, 模型在不同森林类型的  $R^2$  值分别为 0.74、0.73 和 0.59, RMSE 值分别为  $19.53\text{Mg}/\text{hm}^2$ 、 $24.29\text{Mg}/\text{hm}^2$  和  $16.81\text{Mg}/\text{hm}^2$ 。但当加入冠层高度信息后, 模型精度显著提高,  $R^2$  值分别增至 0.87、0.84 和 0.71, RMSE 值分别降至  $14.07\text{Mg}/\text{hm}^2$ 、 $18.71\text{Mg}/\text{hm}^2$  和  $14.11\text{Mg}/\text{hm}^2$ 。总体而言,

模型在高生物量区域的估算偏低,因为在福建省的大部分区域,森林比较密集,仅使用光学遥感影像难以准确估算 AGB,需要冠层高度信息来提升模型精度。加入优化后的冠层高度显著提升了模型精度。对不同森林类型的比较显示,混交林模型的精度最低,可能因为该类型复杂,遥感特征难以明确区分。阔叶林的 RMSE 较高,但  $R^2$  值相对较高,可能因为阔叶林具有大的枝叶 AGB,导致估算结果波动较小但个别样本间差异较大。

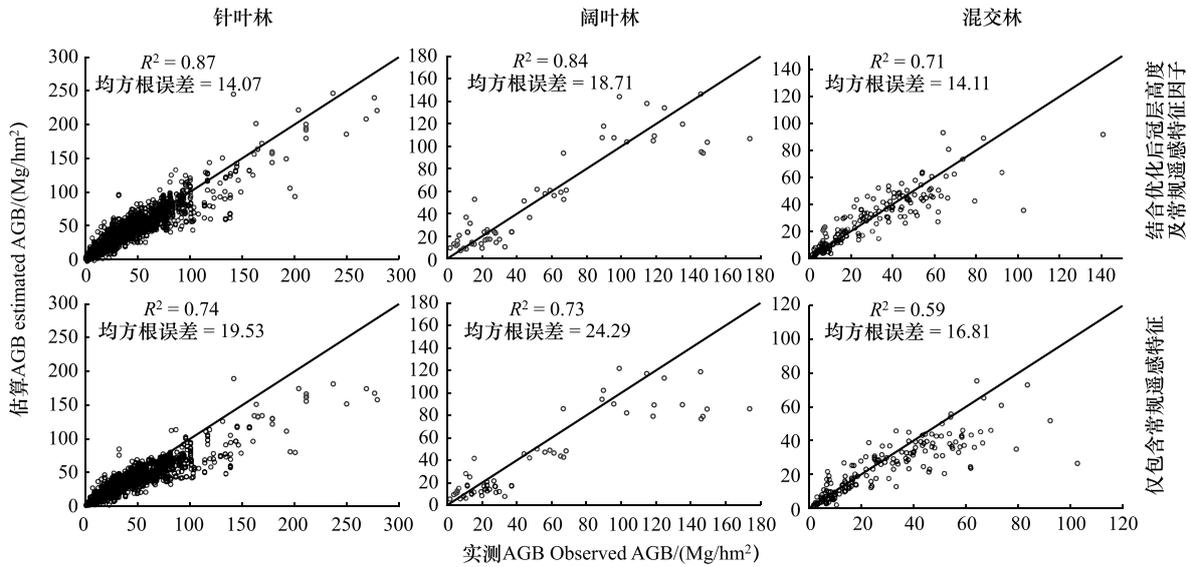


图 11 是否包含冠层高度的 XGBoost 森林 AGB 建模精度对比

Fig.11 Comparison of modeling accuracy of XGBoost forest AGB with or without canopy height

### 3.6 福建省森林 AGB 制图结果

基于所构建的最优森林 AGB 估算模型,利用福建省 2020 年森林区域的多种数据源提取了针叶林、阔叶林和混交林三种森林类型的遥感特征集。利用对应不同森林类型的森林 AGB 反演模型与遥感特征集,我们估算了针叶林、阔叶林以及混交林的地上生物量,且绘制了 2020 年福建省 30m 空间分辨率的森林 AGB 空间分布图,如图 12 所示。

图中展示了福建省的森林 AGB 分布情况。根据统计结果显示,福建省森林 AGB 的范围在 0.001—363.331Mg/hm<sup>2</sup> 之间,大部分森林 AGB 位于 100—150Mg/hm<sup>2</sup> 之间,2020 年全省 AGB 总量为 8.22 亿 Mg,平均值为 101.24Mg/hm<sup>2</sup>,表明森林整体的地上生物量较高,这说明福建省森林繁茂且生态系统健康。生物量较低的区域主要位于经济繁华的城镇中,这主要是由于人为活动干扰与森林覆盖面积的原因。然而,尽管这些区域的森林 AGB 较低,但城镇附近的森林种植区域的地上生物量依然较高,这表明在城镇附近的森林得到了良好的保护。从福建省森林 AGB 的整体空间格局上看,发现福建省东北部区域的森林 AGB 相对较低,而西北以及西南区域的地上生物量整体较高。

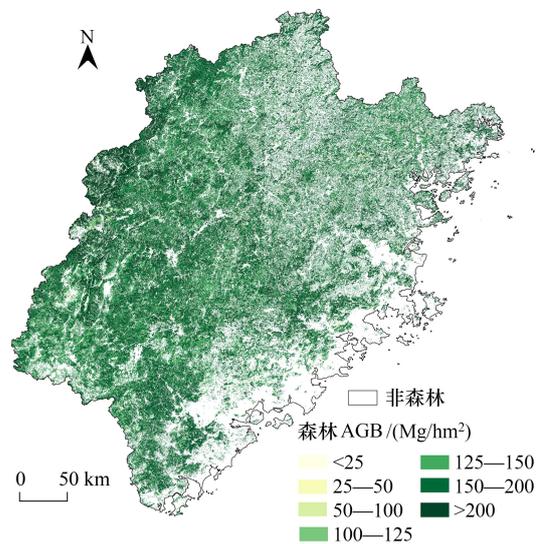


图 12 福建省 30m 空间分辨率森林 AGB 空间分布图

Fig.12 Spatial distribution map of forest AGB at 30m spatial resolution in Fujian Province

### 3.7 精度评估

采用 18203 个样本点的验证数据集来评估福建省森林 AGB 的估算结果。如图 13 所示,整体评估结果显示决定系数  $R^2$  为 0.75,均方根误差 RMSE 为 17.34Mg/hm<sup>2</sup>,这表明本研究在森林 AGB 估算方面具有较高的精度。数据点主要分布在 1:1 线附近,表现出相对较小的波动性。在验证过程中,福建省森林 AGB 主要集中在 200Mg/hm<sup>2</sup>以下的范围。然而,在生物量较低的样本区域,本研究的反演结果存在一定程度的高估。这可能与影像特征的 30m 分辨率有关,这使得森林特征在一定程度上被放大,从而导致生物量较低区域的估算值偏高。

此外,利用样本集验证了三种森林类型的结果精度,评估结果如表 4 所示,针叶林的  $R^2$  为 0.8, RMSE 为 14.43;阔叶林的  $R^2$  为 0.81, RMSE 为 19.26;混交林的  $R^2$  为 0.65, RMSE 为 18.33。

## 4 讨论

通过 Landsat 时序卫星影像计算的林龄优化 GEDI\_V27 冠层高度产品,将其作为特征因子纳入森林 AGB 的回归模型中,并通过 REF 算法构建福建省三种森林类型优选特征集,其中最重要的特征变量均为冠层高度,采用 XGBoost 模型在传统遥感特征因子的基础上加入冠层高度因子,不仅显著提高了模型在森林 AGB 估算中的精度,而且相较于先前在福建开展的仅依靠传统遥感特征的研究<sup>[28]</sup>,实现了精度的显著提升。此外,模型估算结果与实测数据相比有较高的精度,为本研究提供了一定的可信度与理论支撑,但是本文依然存在一些问题值得讨论与分析:

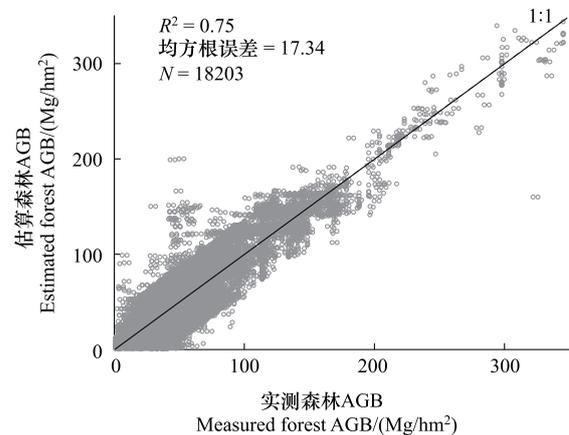


图 13 福建省森林 AGB 精度验证

Fig.13 Verification of AGB accuracy of forests in Fujian Province

表 4 三种森林类型精度验证结果

Table 4 Accuracy verification results of three forest types

森林类型 Forest type	相关系数 $R^2$	均方根误差 RMSE	样本个数 N
针叶林 Coniferous forest	0.8	14.43	15525
阔叶林 Broad-leaved forest	0.81	19.26	1212
混交林 Mixed forest	0.65	18.33	1466

### 4.1 特征集合之间的耦合效应

基于 RFE 算法对提取的遥感特征进行优化,针对福建省针叶林、阔叶林以及混交林进行地上生物量估算的遥感变量重要性排序。发现冠层高度在所有森林类型中均为最重要的特征因子,而地形因子中的地面高程也在特征重要性排序中具有较高的相关性,这与之前的研究有着相同的结论,宋涵玥等<sup>[29]</sup>研究了星载激光雷达在森林资源调查中的重要作用,张宁等<sup>[30]</sup>探讨了生物量随海拔变化规律。构建的优选特征集涵盖了植被指数、纹理特征以及地形因子,但是这些特征之间似乎还是存在一定的共线性,后续能否通过进一步减少特征个数以实现近似的建模精度有待进一步研究。

### 4.2 森林生物量估算模型的不确定性与误差分析

由于福建省地形复杂,森林郁闭度整体较高,因此在 Xiang 生物量计算异速生长方程<sup>[25]</sup>的基础上通过结合大量省内实测数据以及伐区调查数据以修正模型使其符合福建省的情况,但福建树种组成十分复杂<sup>[31]</sup>,除了文中提到的三种森林类型,还含有大量的竹林、桉树以及果园等,因此仅使用这三类树种模型可能会对精度造成一定的影响,后续研究可以尝试对福建省森林类型进行更加详细的分类<sup>[32-33]</sup>,得到各类树种的生长曲

线,以此获得更加准确的 AGB 估算模型。

#### 4.3 冠层高度信息的可优化空间

使用 GEDI 数据主要为冠层高度信息,虽然其 B 级产品与森林 AGB 有很密切的关系,但是由于福建省数据采集过于稀少,同时数据质量并不高,我们很难将 GEDI 的 B 级产品扩展到福建省全域,因而难以利用该数据提升森林 AGB 反演模型。但在同一年,星载光子计数 LiDAR 卫星 ICESat-2<sup>[34]</sup>也开始数据获取工作,且有良好的周期获取数据,虽然 ICESat-2 在福建省森林数据区域的精度较低,但如果能通过一定方法优化该数据,两者结合可以有效地扩充福建省的森林垂直信息数据,弥补 GEDI 数据不足的问题,以有效提升福建省森林冠层高度以及生物量的估算精度。

#### 4.4 XGBoost 模型在森林资源调查中的泛化性

使用 XGBoost 模型用于反演森林 AGB,其原因是已有多名学者证明在多个机器学习模型中,XGBoost 模型的强大性能<sup>[35-36]</sup>,即在遥感定量反演的研究中,XGBoost 往往是表现最佳的模型,基于此,已有学者直接采用 XGBoost 进行森林资源调查<sup>[37]</sup>,上述研究均表明了 XGBoost 模型在遥感数据分析中的优异性能,其原因可能是归因于其出色的非线性建模能力以及对噪声和复杂空间-光谱关系的鲁棒性,使其在处理遥感数据时具有显著优势。

### 5 结论

(1) 结合长序列光学遥感影像以及福建省森林生长曲线,可以较好地提高 GEDI 冠层高度产品的精度,在一定程度上弥补了复杂地形区域 GEDI 数据采集过程中出现的较大误差。经过本文的方法获得了改进后精度更高的福建省冠层高度模型 MGEDI\_V27,精度评价结果为  $R^2$  为 0.67, RMSE 为 2.24m

(2),利用多源遥感数据构建的特征集共有 72 个特征,最终特征优选出 10 个遥感特征用于构建福建省森林 AGB 反演模型。三种森林类型最重要的遥感特征均为森林冠层高度,表明冠层高度作为树木立体结构的关键指标,能够反映森林生物量的空间分布。另外评估了森林冠层高度对模型精度的改进效果,结果表明只使用传统遥感特征时,模型在不同森林类型的  $R^2$  值在 0.59 至 0.74 之间, RMSE 值在 16.81Mg/hm<sup>2</sup> 至 24.29Mg/hm<sup>2</sup> 之间,加入冠层高度信息后,模型精度显著提高,  $R^2$  值提高到 0.71 至 0.87, RMSE 值分别降到 14.07Mg/hm<sup>2</sup> 至 18.71Mg/hm<sup>2</sup>,表明冠层高度信息可有效提高森林 AGB 估算的准确性,尤其适用于地形和森林类型复杂的福建省。

(3) 通过 18203 个样本点对福建省的森林 AGB 结果进行了精度验证。反演的森林 AGB 结果以验证样本集为参考,得到的精度评价结果为  $R^2 = 0.75$ , RMSE = 17.34 Mg/hm<sup>2</sup>。各森林类型的精度评估结果为:针叶林的  $R^2$  为 0.8, RMSE 为 14.43 Mg/hm<sup>2</sup>;阔叶林的  $R^2$  为 0.81, RMSE 为 19.26Mg/hm<sup>2</sup>;混交林的  $R^2$  为 0.65, RMSE 为 18.33Mg/hm<sup>2</sup>。福建省森林 AGB 的范围在 0.001—363.331Mg/hm<sup>2</sup> 之间,2020 年全省 AGB 总量为 8.22 亿 Mg,平均值为 101.24Mg/hm<sup>2</sup>。

#### 参考文献 (References):

- [1] Ar'evalo P, Baccini A, Woodcock C E, Olofsson P, Walker W S. Continuous mapping of aboveground biomass using Landsat time series. *Remote Sensing of Environment*, 2023, 288: 113483.
- [2] Moradi F, Darvishsefat A A, Pourrahmati M R, Deljouei A, Borz S A. Estimating aboveground biomass in dense Hyrcanian forests by the use of sentinel-2 data. *Forests*, 2022, 13(1): 104.
- [3] 吴培强,任广波,张程飞,王浩,刘善伟,马毅. 无人机多光谱和 LiDAR 的红树林精细识别与生物量估算. *遥感学报*, 2022, 26(6): 1169-1181.
- [4] Tamimnia H, Salehi B, Mahdianpari M, Beier C M, Johnson L. Evaluating pixel-based and object-based approaches for forest above-ground biomass estimation using a combination of optical, sar, and an extreme gradient boosting model. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2022, : 485-492.
- [5] Zhang X, Liu L Y, Chen X D, Gao Y, Xie S, Mi J. GLC\_FCS30: global land-cover product with fine classification system at 30? m using time-series Landsat imagery. *Earth System Science Data*, 2021, 13(6): 2753-2776.
- [6] Zolkos S G, Goetz S J, Dubayah R. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, 2013, 128: 289-298.
- [7] Main-Knorn M, Cohen W B, Kennedy R E, Grodzki W, Pflugmacher D, Griffiths P, Hostert P. Monitoring coniferous forest biomass change using

- a Landsat trajectory-based approach. *Remote Sensing of Environment*, 2013, 139: 277-290.
- [ 8 ] 柳钦火, 仲波, 吴纪桃, 肖志强, 王桥. 环境遥感定量反演与同化: 科学出版社, 2011.
- [ 9 ] 苟睿坤, 陈佳琦, 段高辉, 杨瑞, 卜元坤, 赵君, 赵鹏祥. 基于 GF-2 的油松人工林地上生物量反演. *应用生态学报*, 2019, 30(12): 4031-4040.
- [ 10 ] 菅永峰, 韩泽民, 黄光体, 王熊, 李源, 周靖靖, 佃袁勇. 基于高分辨率遥感影像的北亚热带森林生物量反演. *生态学报*, 2021, 41(6): 2161-2169.
- [ 11 ] 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.
- [ 12 ] Pham T D, Yoshino K, Le N N, Bui D T. Estimating aboveground biomass of a mangrove plantation on the Northern coast of Vietnam using machine learning techniques with an integration of ALOS-2 PALSAR-2 and Sentinel-2A data. *International Journal of Remote Sensing*, 2018, 39(22): 7761-7788.
- [ 13 ] Jan A, Henrik P, Lars U. Biomass growth from multi-temporal TanDEM-X interferometric synthetic aperture radar observations of a boreal forest site. *Remote Sensing*, 2018, 10(4): 603.
- [ 14 ] Kumar S, Garg R D, Govil H, Kushwaha S P S. PolSAR-decomposition-based extended water cloud modeling for forest aboveground biomass estimation. *Remote sensing*, 2019, 11(19): 2287.
- [ 15 ] Su H, Shen W, Wang J, Ali A, Li M. Machine learning and geostatistical approaches for estimating aboveground biomass in Chinese subtropical forests. *Forest Ecosystems*, 2020, 7(04): 851-870.
- [ 16 ] Lu D, Chen Q, Wang G X, Liu L J, Li G Y, Moran E. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*, 2016, 9(1): 63-105.
- [ 17 ] 闻馨, 刘凯, 曹晶晶, 朱远辉, 王子予. 基于森林冠层高度和异速生长方程的中国红树林地上生物量估算. *热带地理*, 2023, 43(1): 1-11.
- [ 18 ] Du C Y, Fan W Y, Ma Y, Jin H I, Zhen Z. The effect of synergistic approaches of features and ensemble learning algorithm on aboveground biomass estimation of natural secondary forests based on ALS and landsat 8. *Sensors*, 2021, 21(17): 5974.
- [ 19 ] 巨一琳, 姬永杰, 黄继茂, 张王菲. 联合 LiDAR 和多光谱数据森林地上生物量反演研究. *南京林业大学学报: 自然科学版*, 2022, 46(1): 58-68.
- [ 20 ] Nandy S, Srinet R, Padalia H. Mapping forest height and aboveground biomass by integrating ICESat-2, sentinel-1 and sentinel-2 data using random forest algorithm in northwest Himalayan foothills of India. *Geophysical Research Letters*, 2021, 48(14): e2021GL093799.
- [ 21 ] Silva C A, Duncanson L, Hancock S, Neuenschwander A, Thomas N, Hofton M, Fatoyinbo L, Simard M, Marshak C Z, Armston J, Lutchke S, Dubayah R. Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. *Remote Sensing of Environment*, 2021, 253: 112234.
- [ 22 ] Qi W L, Saarela S, Armston J, Ståhl G, Dubayah R. Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI lidar data. *Remote Sensing of Environment*, 2019, 232: 111283.
- [ 23 ] Yue J B, Yang G J, Li C C, Li Z H, Wang Y J, Feng H K, Xu B. Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sensing*, 2017, 9(7): 708.
- [ 24 ] Quegan S, Le Toan T, Chave J, Dall J, Exbrayat J, Minh D H T, Lomas M, D'Alessandro M M, Paillou P, Pathanassiou K, Rocca F, Saatchi S, Scipal K, Shugart H, Smallman T L, Soja M J, Tebaldini S, Ulander L, Villard L, Williams M. The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space. *Remote Sensing of Environment*, 2019, 227: 44-60.
- [ 25 ] Xiang W H, Zhou J, Ouyang S, Zhang S L, Lei P F, Li J X, Deng X W, Fang X, Forrester D I. Species-specific and general allometric equations for estimating tree biomass components of subtropical forests in Southern China. *European Journal of Forest Research*, 2016, 135(5): 963-979.
- [ 26 ] Kennedy R E, Yang Z Q, Cohen W B. Detecting trends in forest disturbance and recovery using yearly Landsat time series; 1. LandTrendr-temporal segmentation algorithms. *Remote Sensing of Environment*, 2010, 114(12): 2897-2910.
- [ 27 ] Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. *Machine Learning*, 2002, 46(1): 389-422.
- [ 28 ] 苏华, 张明慧, 李静, 陈修治, 汪小钦. 基于光学与 SAR 因子的森林生物量多元回归估算——以福建省为例. *遥感技术与应用*, 2019, 34(4): 847-856.
- [ 29 ] 宋涵玥, 舒清态, 席磊, 邱霜, 魏治越, 杨泽至. 基于星载 ICESat-2/ATLAS 数据的森林地上生物量估测. *农业工程学报*, 2022, 38(10): 191-199.
- [ 30 ] 张宁, 王金牛, 石凝, 王丽华, 朱牛, 田炳辉, 张林, 盖艾鸿. 岷江源区两种优势针叶树当年生小枝性状与生物量分配随海拔的分异规律. *生态学报*, 2023, 43(23): 9814-9826.
- [ 31 ] 兰思仁. 福建省森林景观类型及地理分布概述. *林业资源管理*, 2002(1): 55-59.
- [ 32 ] Adugna T, Xu W B, Fan J L. Comparison of random forest and support vector machine classifiers for regional land cover mapping using coarse resolution FY-3C images. *Remote Sensing*, 2022, 14(3): 574.
- [ 33 ] Decuyper M, Chávez R O, Lohbeck M, Lastra J A, Tsendbazar N, Hackländer J, Herold M, Vågen T G. Continuous monitoring of forest change dynamics with satellite time series. *Remote Sensing of Environment*, 2022, 269: 112829.
- [ 34 ] Liu A B, Cheng X, Chen Z Q. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sensing of Environment*, 2021, 264: 112571.
- [ 35 ] Li Y C, Li M Y, Li C, Liu Z Z. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. *Scientific Reports*, 2020, 10: 9952.
- [ 36 ] Luo M, Wang Y F, Xie Y H, Zhou L, Qiao J J, Qiu S Y, Sun Y J. Combination of feature selection and CatBoost for prediction: the first application to the estimation of aboveground biomass. *Forests*, 2021, 12(2): 216.
- [ 37 ] 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.
- [ 38 ] Dubayah R, Blair J B, Goetz S, Fatoyinbo L, Hansen M, Healey S, Hofton M, Hurt G, Kellner J, Luthcke S, Armston J, Tang H, Duncanson L, Hancock S, Jantz P, Marselis S, Patterson P L, Qi W, Silva C. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, 2020, 1: 100002.