

利用冠层三维结构特征改进瑞士阿尔高州星载光子计数激光雷达森林蓄积量估测

孔丹^{1,2}, 庞勇^{1,2}, 汪祖媛³, 李增元^{1,2}

1. 中国林业科学研究院资源信息研究所, 北京 100091;

2. 国家林业和草原局 林业遥感与信息技术重点实验室, 北京 100091;

3. 瑞士联邦森林、雪与景观研究所, 瑞士 比尔门斯多夫 8903

摘要: 星载光子计数激光雷达系统以冰云与陆地高程卫星-2 ICESat-2 (Ice, Cloud and land Elevation Satellite-2) 搭载的先进地形激光高度计系统 ATLAS (Advanced Topographic Laser Altimeter System) 为代表, 能够快速获取大区域植被三维信息, 已广泛用于森林参数反演, 同时也存在普适性差等问题。为此, 本研究以瑞士阿尔高州针阔混交森林为研究对象, 从森林三维结构解析的视角出发, 评估冠层水平与垂直结构特征在提升ICESat-2数据复杂林分蓄积量估测精度中的作用, 并探索适用于该区域的最优模型形式, 同时与仅包含传统高度统计特征的基线模型进行对比。首先, 将去噪后的ICESat-2 ATLAS数据分割为100 m的估测单元, 并通过质量控制, 识别并去除异常单元, 确保数据质量; 然后, 通过特征参数分组预筛选和有规则约束的全子集筛选方法, 综合利用点云高度分布特征、冠高及冠高异质性特征、垂直结构特征进行森林蓄积量估测, 筛选最优特征组合形式。研究结果表明, 瑞士阿尔高州针阔混交森林ATLAS蓄积量估测的最优模型由冠层顶部平均高、65%密度分位数、叶面积加权冠层体积和枝叶剖面的平均值组成。十折交叉验证结果显示, 该模型的精度达到平均精度为 $R^2=0.78$, $RMSE=92.48\text{ m}^3/\text{hm}^2$, $rRMSE=0.24$ 。相比之下, 仅包含传统特征参数的基线模型 $R^2=0.66$, $rRMSE$ 由0.28降低至0.24, 说明综合引入结构特征可将精度提高ICESat-2数据蓄积量估测精度, 改善在冠层异质性较高林分中的估测表现。综上, 综合利用森林三维结构特征改进基于ICESat-2数据森林蓄积量估测精度可以有效提升模型在复杂林分条件下的适用性, 为大区域森林蓄积量与碳储量监测提供方法支撑。

关键词: 森林蓄积量, ICESat-2 ATLAS, 森林三维结构特征, 冠高异质性, 垂直结构, 质量控制, 分组预筛选, 有规则约束的全子集

中图分类号: S771.8/P2

引用格式: 孔丹, 庞勇, 汪祖媛, 李增元. 2025. 利用冠层三维结构特征改进瑞士阿尔高州星载光子计数激光雷达森林蓄积量估测. 遥感学报, 29(10): 2976-2988

Kong D, Pang Y, Wang Z Y and Li Z Y. 2025. Improving forest stock volume estimation using three-dimensional canopy structure characteristics based on ICESat-2 ATLAS data in Aargau, Switzerland. National Remote Sensing Bulletin, 29(10): 2976-2988 [DOI: 10.11834/jrs.20254317]

1 引言

近年来, 随着全球气候变化加剧, 实现“碳中和”和“碳达峰”已经成为全球生态治理的重要议题。森林生态系统对于减缓温室效应、维持气候平衡具有至关重要的作用 (于贵瑞 等, 2022)。森林蓄积量是指在一定面积范围内所有生长树木的木材总体积, 通常用于表示森林中树木的生长量。

作为评估森林资源状况和生产力的重要指标, 森林蓄积量在全球森林资源清查中起着核心作用, 直接关系到森林的固碳能力, 并用于判断森林的健康状态和生态功能 (Chen 等, 2021)。准确测定森林蓄积量对于深入理解和有效管理森林生态系统, 促进森林资源的可持续利用具有重要意义。

传统森林蓄积量估测使用随机或系统抽样的方法, 其精度受限于森林样地的数量和分布

收稿日期: 2024-07-31; 预印本: 2024-10-21

基金项目: 国家重点研发计划 (编号: 2023YFE0105100, 2020YFE200800)

第一作者简介: 孔丹, 研究方向为星载、机载激光雷达信号处理及林业应用。E-mail: ak_4798@126.com

通信作者简介: 庞勇, 研究方向为林业遥感机理模型、激光雷达信号处理及林业应用、森林碳汇计量、森林变化监测。E-mail: pangy@ifrit.ac.cn

(Goldbergs 等, 2018)。激光雷达技术可以快速获取大范围的森林三维信息, 广泛应用于森林蓄积量、生物量估测 (Du 等, 2023)。激光雷达技术按照采集平台可以分为地基、机载 (无人机和有人机)、星载激光雷达, 其中地基和无人机激光雷达受限于数据获取速率, 适用于精细尺度或小区域范围森林蓄积量估测; 通过有人机激光雷达 ALS (Airborne Laser Scanning) 可以获取林场或县市级别的激光雷达数据, 通常作为蓄积量估测模型尺度外推的基准数据 (Coops 等, 2021); 尽管 ALS 在小尺度森林参数提取中具有较高的精度, 其应用往往受到高昂的运行成本和有限的覆盖范围的限制, 难以满足大尺度、长期监测的需求。相比之下星载传感器, 具备全球覆盖和高时间分辨率的优势, 能够以更低的成本实现对大面积森林资源的持续监测。星载激光雷达是获取大尺度激光雷达数据的最佳手段, 目前在轨的星载激光雷达主要包括以下两种: (1) 全波形激光雷达, 例如搭载于国际空间站的全球生态系统动力学调查 GEDI (Global Ecosystem Dynamics Investigation) 激光雷达和中国的陆地生态系统碳监测卫星 TECIS (China's Terrestrial Ecosystem Carbon Inventory Satellite) 等; (2) 光子计数激光雷达, 例如 ICESat-2。美国国家航天局于 2018 年发射了 ICESat-2 卫星, 搭载了新一代多波束微脉冲光子计数的激光雷达系统, ATLAS 的特点是波束多和高脉冲频率, 相比于 GEDI 传感器, ATLAS 较小的光斑直径 (约 11 m) 可以有效减小光斑内部地形的影响; 在相同情况下, 高脉冲频率 (10 kHz) 使得 ATLAS 传感器能够获得更丰富的森林垂直结构信息 (Neumann 等, 2019)。ICESat-2 卫星已持续运行 5 年, 收集全球地表三维信息, 成为大范围森林资源监测的重要数据之一。其在森林蓄积量、生物量估测方面表现出巨大潜力, 常作为光学遥感数据森林参数估测的基准数据 (Narine 等, 2020; Nandy 等, 2021)。

ICESat-2 数据在外观上与传统小光斑点云数据相似, 普遍使用面积法来进行蓄积量、生物量估测, 该方法通过光子数据的特征参数与参考蓄积量的统计学关系进行估测。如 Narine 等 (2019, 2020) 分别基于模拟和真实的 ICESat-2 数据提取高度变量、密度变量特征, 建立了高精度的生物量估测模型, 并结合 Landsat 数据进行空间连续的生物量制图, 证明了 ICESat-2 在生物量估测及作

为尺度外推数据基准的能力。Song 等 (2023) 利用随机森林模型结合冠层参数及地形因子进行 ICESat-2 生物量估测, 并通过克里金的插值的方法进行空间连续制图。Varvia 等 (2022) 评估了积雪深度、太阳噪声、强弱波束对 ICESat-2 数据在森林地上生物量精度的影响, 结果表明使用无雪条件下的夜间强光束数据进行 AGB 估算精度最佳。Liu 和 Popescu (2022) 使用最小绝对收缩和选择算子 (LASSO) 回归建立了 ATL08 LiDAR 指标与参考地上生物量之间的关系, R^2 为 0.77, RMSE 为 56.67 Mg/ha。Varvia 等 (2024) 提出一种分层混合推断方法, 用以评估 ICESat-2 数据估算芬兰努尔梅斯地区平均生物量的准确性, 估算值为 65.7 ± 1.9 Mg/ha, 相对标准误为 2.9%。Neuenschwander 等 (2024) 使用模拟 ICESat-2 数据对北极森林区域 ICESat-2 数据进行生物量估测和制图, 结果表明使用 70% 和 98% 光子点云高度分位数作为特征的模型精度最高 ($R^2=0.74$, RMSE=59.06 Mg/ha, rRMSE=64%)。其他研究主要集中在 ICESat-2 数据的外推能力 (Nandy 等, 2021) 以及联合其他数据 (如 GEDI 数据) 估测森林参数的能力 (Silva 等, 2020; 孟鸽 等, 2024)。

基于面积法建立的 ICESat-2 数据蓄积量估测模型通常在不同区域、树种选择不同的特征参数。其中, 基于点云高度分布的参数是森林蓄积量估测最常用的特征参数, 例如高度、密度百分位数和平均高度 (Næsset, 2002), 然后根据一定的规则进行特征参数筛选, 保留符合要求的特征参数 (Lefsky 等, 2007; Pang 等, 2008; Schneider 等, 2020)。面积法的森林蓄积量估测模型面临两个主要问题: 首先, 大量激光雷达特征参数之间普遍存在很强的共线性, 导致特征参数筛选工作的复杂性, 难以开发适用于大区域、多树种的模型 (Hall 等, 2005); 其次, 常用的特征参数普遍由光子点云高度垂直分布特征计算得出, 未考虑冠层水平及垂直的异质性 (Bouvier 等, 2015)。针对面积法存在的问题, Bouvier 等 (2015) 使用冠层高度模型 CHM (Canopy Height Model) 的平均值、CHM 的方差、间隙率和叶面积密度分布的变异系数构建机载森林参数估测模型。前 3 个指标与平均冠层高度、高度异质性和水平冠层分布有关, 第 4 个指标为垂直结构信息, 体现了林分垂直分布的信息, 在林分属性的估算精度上高于仅使用高度指标的

模型,同时也避免了数据过拟合。目前已有研究表明ICESat-2数据能够准确描绘森林冠层高度分布和部分垂直结构信息(Queinnec等, 2021)。然而,ICESat-2光子点云密度与精度远低于机载数据,这些特征参数是否有助于改进蓄积量估测模型的精度有待研究。

为保证模型评估的可重复性与结论的稳健性,本研究选取瑞士阿尔高州作为研究区,原因如下:阿尔高州提供开放且质量可控的机载激光雷达数据与森林清查资料(Jiang等, 2023),可为ICESat-2 ATLAS的蓄积量建模提供可靠参考与独立验证,降低参考数据误差对结果的影响;研究区以针阔混交林为主,林分在水平与垂直结构上均表现出显著异质性(Jiang等, 2023),适合检验在提升模型普适性与稳健性方面的贡献;区域地形起伏中等、经营历史清晰,便于控制地形与经营干扰等混杂因素;同时,所处纬度带来较好的ICESat-2轨道覆盖与重复观测条件,有利于提升样本代表性。

因此,本研究拟选取瑞士阿尔高州作为研究区,从森林三维结构解析的视角出发,基于ICESat-2数据分别提取了冠层高度、冠高异质性、垂直结构特征,通过分组预筛选和有规则的全子集回归方法进行变量筛选。综合利用ICESat-2提取的森林冠层高度、高度异质性和水平冠层分布等信息进行蓄积量估测,研究旨在评估利用冠层水平和垂直结构特征来改进ICESat-2数据森林蓄积量估测精度的可行性,并探究针对瑞士针阔混交森林蓄积量估测的最优模型形式,以期为大区域森林蓄积量估测提供方法参考。

2 研究方法

2.1 研究区概况

研究区位于瑞士阿尔高州(Canton Aargau),地里位置 $7^{\circ}42'59''\text{E}$ — $^{\circ}27'14.06''\text{E}$, $47^{\circ}8'38''\text{N}$ — $47^{\circ}37'16''\text{N}$,总面积为 1404 km^2 。该州地处阿尔河下游,为典型的温带气候特征,年均温度、降雨量分别为 9°C 、 850 mm 。阿尔高州地区主要地形为平原,海拔高度从 206 — 908 m 。研究区域的森林覆盖类型为混交林,主要树种包括挪威云杉(*Picea abies* L.)、山毛榉(*Fagus sylvatica* L.)、桦树(*Betula pubescens*)、枫树(*Ace pseudoplatanus*)和

欧洲赤松(*Pinus sylvestris*)。

2.2 数据及预处理

2.2.1 ICESat-2数据及预处理

ICESat-2卫星由美国于2018年发射,其搭载的光子计数技术激光雷达系统ATLAS,为大区域森林蓄积量估测提供了林区单光子水平数据。ICESat-2官方提供了从Level-0到Level-3的21种数据产品(ATL00—ATL21)。其中,每个光束包含强光束和弱光束,强光束能量约为弱光束的4倍(Neuenschwander和Magruder, 2016)。其中ATL03提供经过地理定位的光子级别的数据,提供每个光子的纬度、经度和高度等信息。

本研究采用与机载数据飞行时间相近的ATL03产品数据(时间间隔在4个月内),在瑞士研究区分别得到6个条带数据,其分布如图1所示。ATLAS容易受到大气散射和太阳辐射等因素的干扰,因此,有效去除噪声光子对于后续研究和应用至关重要。本研究通过随机抽样一致分段曲线拟合与方向自适应的基于排序点的聚类结构识别算法结合的方式对实验区ATL03数据进行去噪,依次提取精细地表与植被信号,在测试数据信号提取精度(F 值)达到0.97。在此基础上,将去噪后的光子进行高度归一化并沿轨道方向按 100 m 间隔进行分割,作为蓄积量估测单元。

2.2.2 样地实测蓄积量

实测样地数据来自于瑞士第5期国家森林资源清查数据NFI(National Forest Inventory),瑞士NFI5在 500 m^2 的同心圆内,测量了样地每棵活立木(胸径 $>12\text{ cm}$)的胸径、树种等信息。应用在阿尔高州中测量的所有树木训练非线性最小二乘模型计算树木生物量、蓄积量,具体计算公式如下:

$$\text{Vol} = \left(b_0 + b_1 \times \log(D) + b_2 \times (\log(D))^4 + b_3 \times D + t_2 \times T_2 + t_3 \times T_6 + t_4 \times T_7 + t_5 \times T_5 \right) \quad (1)$$

式中,Vol为活立木, D 为每棵树的胸径值, b_0 — b_3 以及 t_2 — t_5 均为树种相关都修改因子,详见表1。 T_n 表示根据树种将计算出的树木蓄积量略有提高(或降低)的系数,其中 $n=2, 6, 7, 12$, t_2 、 t_3 、 t_4 、 t_5 分别对应冷杉、其他针叶树、山毛榉、其他阔叶树。当计算冷杉时 $T_2=1$, T_6 、 T_7 、 $T_5=0$ 。

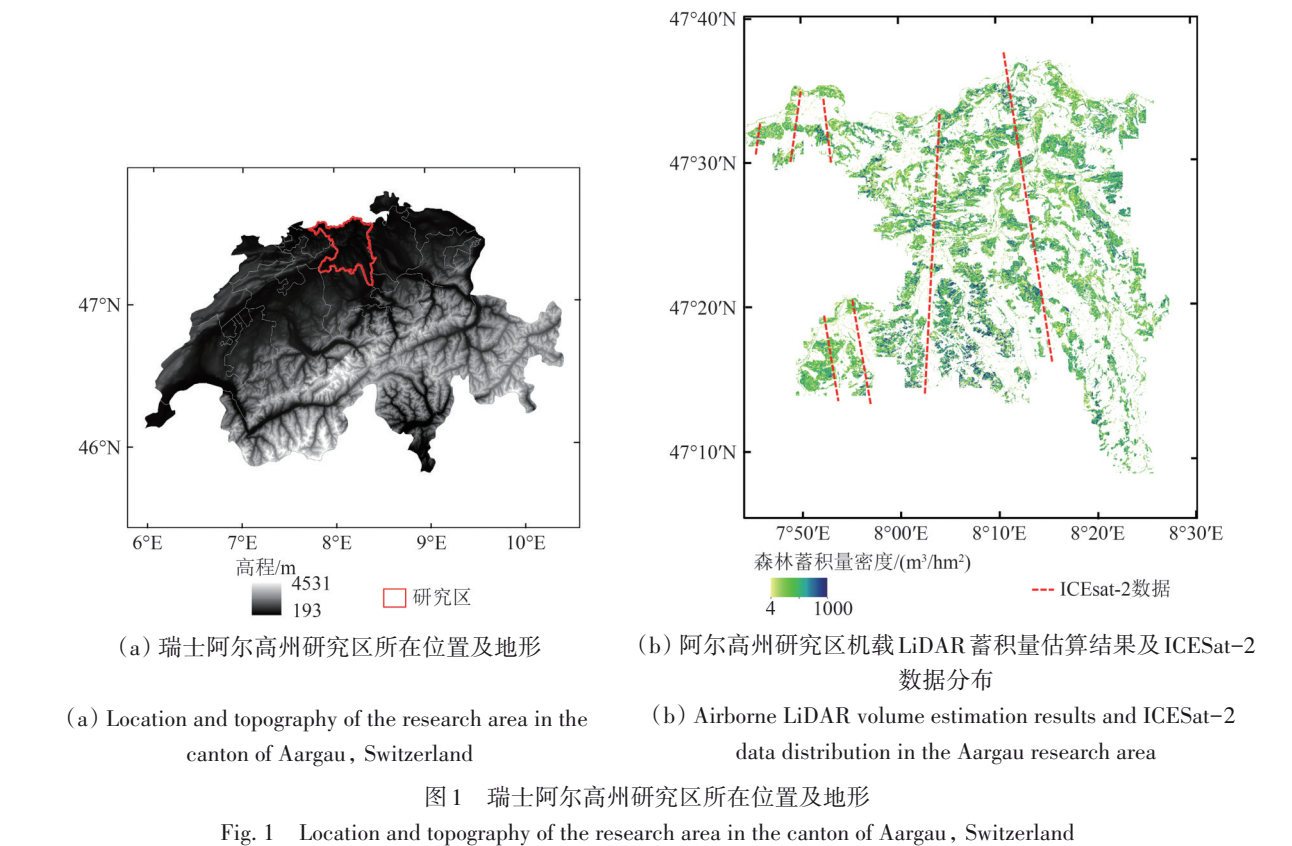


表 1 蓄积量计算公式系数对应表	
Table 1 Coefficient of volume calculation formulas	
参数	蓄积量参数值
b0	-1.121E+01
b1	3.484E+00
b2	-5.656E-03
b3	-9.111E-05
t2	3.631E-02
t3	-3.496E-02
t4	7.011E-03
t5	-1.938E-01

2.2.3 ALS数据及蓄积量估测

ALS数据由瑞士联邦政府地形办公室（Office Fédéral de Topographie Swisstopo）瑞士三维表面（swiss SURFACE3D）项目提供，采集时间为2019年

3—4月和2020年2—3月，数据下载地址为<https://www.swisstopo.admin.ch/en/geodata/height/surface3d.html> [2024-07-31]。数据平均点云密度为15—20 pts/m²，并经过分类与高度归一化处理。该数据具有较高的水平和垂直精度，水平位置误差在±20 cm范围内，垂直误差在±10 cm范围内。

将瑞士NFI5中实测蓄积量作为基准数据，构建ALS数据的蓄积量估测模型。其中，ALS数据提取包括高度、密度和垂直结构等共92个特征参数。通过逐步回归的方法进行变量筛选，并采用最小二乘法构建线性蓄积量估测模型，在瑞士实验区针对针叶林、阔叶林和混交林的分别构建了不同的ALS数据蓄积量估测模型。模型形式详见表2，特征参数的描述详见表3。

表 2 基于 ALS 特征参数的蓄积量估测模型	
Table 2 Accumulation estimation model based on ALS feature parameters	
林地类型	模型
针叶林	$\text{Vol} = -99.47 \times H_{\text{stdv}} - 47.39 \times \text{RH}_{10} + 7.31 \times \text{Zpcum5} + 1801.69 \times \text{Dp}_{35} + 3041.56 \times \text{Dp}_{90} - 10394.62 \times \text{VFP}_{\text{mean}} + 67.29 \times \text{TCH}_{25} - 126.04$
阔叶林	$\text{Vol} = -12.60 \times \text{CHM}_{\text{mean}} - 0.50 \times \text{CHM}_{\text{var}} + 9.47 \times \text{RH}_{15} - 30.52 \times \text{Zpcum9} - 13.87 \times \text{Zpcum7} - 3245.00 \times \text{Dp}_{70} - 6618.66 \times \text{Dp}_{95} - 1724.83 \times \text{VFP}_{\text{stdv}} + 66.70 \times \text{TCH}_{25} - 10.82 \times \text{TCH}_1 + 4405.55$
混交林	$\text{Vol} = -64.41 \times \text{CHM}_{\text{mean}} - 37.18 \times H_{\text{stdv}} + 153.86 \times \text{MCH} - 194.4897$

表3 ICESat-2数据特征参数
Table 3 ICESat-2 derived feature parameters

类别	特征参数	参数描述
基于点云高度分布	RH_x	点云高度分布的高度百分位数
	$H_{(min,max,\cdots, stddev)}$	点云高度分布的统计学参数,包括最小值、最大值、标准差等
	Dp_x	分位数高度的所有植被点占全部点的比例
	$Zpcum1-Zpcum9$	从最低回波高度到10%—90%分位数高度植被点占全部点的比例
	CC	冠层覆盖度
表征冠高及异质性	$CHM_{(mean, var, gini, cv)}$	CHM的平均值、方差、Gini系数和变异系数
	CH	平均冠层剖面高
	TCH	冠层顶部平均高度
	LAVH	分层叶面积指数加权冠层体积
表征垂直结构	$LAD_{(mean, stdev, cv)}$	叶面积密度剖面的变异系数、平均值和标准差
	$VFP_{(mean, stdev, cv)}$	枝叶剖面(由CHM计算)的均值、方差和变异系数
	$RH_{(slope, intercept, residuals)}$	高度分位数与高度值的斜率、截距、

瑞士阿尔高州示范区针叶林、阔叶林和混交林 ALS 蓄积估测模型精度如图2所示,其中 R^2 分别为0.93、0.82和0.84, RMSE 分别为69.13、81.87和76.82 m^3/hm^2 。并按照 ICESat-2 光斑直径将100 m的

估测单元扩展为100 m×11 m的范围(ICESat-2光斑直径约为11 m),裁剪高度归一化后的ALS数据,结合林地类型数据,估测每个100 m×11 m的范围森林蓄积量,作为ICESat-2估测单元的参考蓄积量。

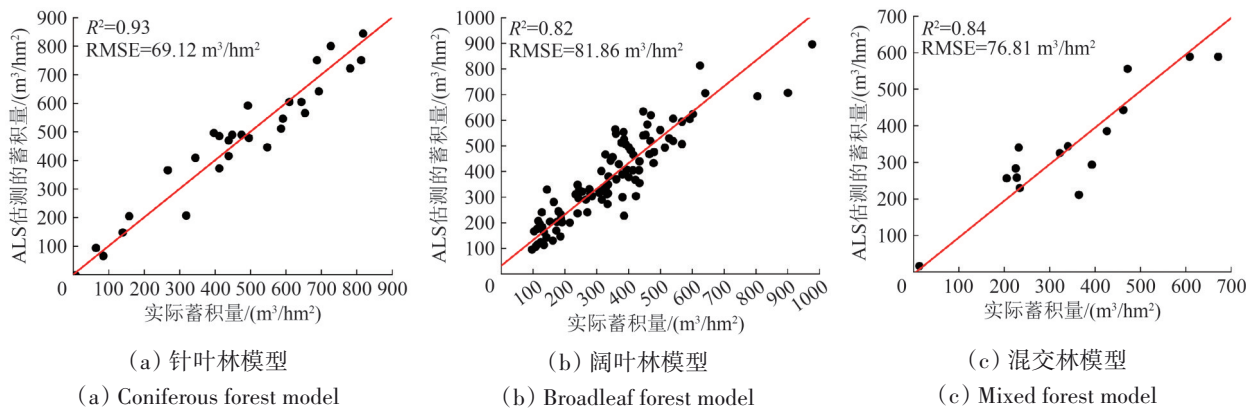


图2 不同林地类型 ALS 估算蓄积量的模型精度
Fig. 2 Model accuracy of ALS-estimated stand volume for different forest types

2.2.4 辅助数据

本研究主要辅助数据为瑞士阿尔高州2018年10 m分辨率的林地类型(数据下载地址: <https://map.geo.admin.ch/> [2024-07-31])。该数据每个像元表示阔叶树占比(0—100),并按照10%和90%的阈值分为针叶林、阔叶树和混交林3种类型,对应2.2.3中ALS蓄积量估测的3种模型。

2.3 基于光子技术激光雷达数据森林蓄积量估测方法

2.3.1 研究思路

研究思路如图3所示:首先对预处理后的ICESat-2蓄积量估测单元进行质量控制,去除异

常单元;然后,计算每个估测单元内ICESat-2数据的特征参数(基于点云高度分布、冠层及冠高异质性、垂直结构的特征参数);最后,结合逐步回归、岭回归等方法进行分组特征筛选并通过有规则约束的全子集回归进行组合,以十折交叉验证的平均精度作为评价指标,选择最优模型。

2.3.2 蓄积量估测单元质量控制

本研究将去噪后的ICESat-2数据沿轨道方向分割为100 m尺度的估测单元,为确保数据的准确性,并在Wu和Shi(2023)研究的基础上,设计了质量控制的流程,以去除不适合蓄积量估测的ICESat-2单元,主要包括以下步骤:

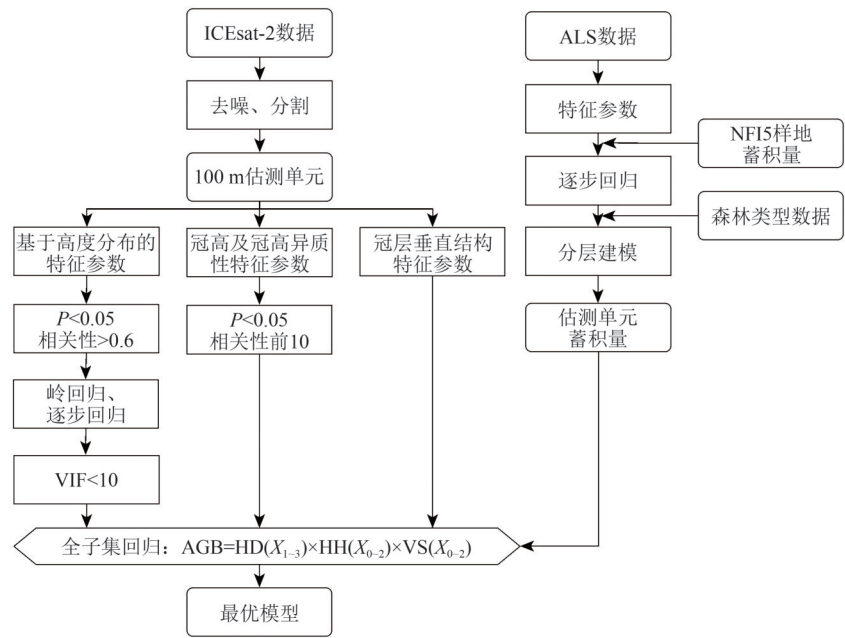


图3 利用森林三维结构特征改善森林蓄积量估测模型的研究思路

Fig. 3 Flowchart of improving forest volume estimation using three-dimensional forest structure features

- (1) 筛选非森林区域的单元：提取估测单元范围内森林分类图的像元值，若估测单元内 80% 以上的像元值为-1（非林地像元），则去除该单元。
- (2) 筛选纯单元：由于处理单元为近似于 100 m×11 m 的条带，容易包含道路、林班边界等情况，导致蓄积量模型的偏差，因此我们按照对森林定义（郁闭度>0.2），去除间隙（由 ALS 数据计算）率>0.8 的单元，确保估测模型中只包含纯森林单元。
- (3) 筛选点云密度不足的单元：ICESat-2 数据点云密度越高，对于描绘森林冠层数据的准确性越高。由于本研究将激光雷达变量计算最小窗口设置为 5 m，去除冠层点数量<20 的单元（Wu 和 Shi, 2023）。
- (4) 筛选去噪质量差的单元：ICESat-2 数据中可能仍存在部分离散的噪声点。这些噪声点相对离散且距离冠层相对较远，因此根据最高值和点云连续性进行筛选，本研究根据区域内数据情况确定阈值：点最高值>40 m；5%—98% 高度分位数任意相邻间隔>10 m。满足这两个条件的单元可能存在异常噪声，不参与蓄积量建模估测。

2.3.3 森林三维结构特征参数提取

ICESat-2 数据外观上类似于传统小光斑点云获取的数据，可以参考小光斑激光点云数据构造特征参数，如高度分位数、密度分位数、冠层覆盖

度等。此外，由于 ICESat-2 数据中光斑高度重叠且相邻光斑排列紧密，可以从中提取冠层点云并计算其结构特征。表 3 描述了综合构建的适于表征森林三维结构信息的 ICESat-2 数据特征参数。将其分为 3 类：（1）基于点云高度分布的特征参数；（2）表征冠高及异质性的特征参数；（3）表征垂直结构的特征参数。

在使用野外实测数据进行林分蓄积量估测时，一般通过式（2）计算：

$$\text{Vol} = \alpha_0 * \bar{H}^{\beta_1} \text{BA}^{\beta_2} + \varepsilon_1 \tag{2}$$

式中， $\bar{H}$ 是林分平均高，BA 为林分胸高断面面积， α_0 、 β_1 、 β_2 为参数项， ε_1 为误差项。

由于机载、星载激光雷达难以准确观测林木胸径指数，胸高断面面积常用冠层密度或覆盖度相关的指数代替，因此基于激光雷达计算蓄积量的公式如下：

$$\text{Vol} = \alpha_0 * H^{\beta_3} D^{\beta_4} + \varepsilon_2 \tag{3}$$

式中， H 为点云高度特征、 D 为密度相关特征， β_3 、 β_4 为参数项， ε_2 为误差项。

高度与密度相关特征普遍源于光子点云高度的分布，然后根据一定的规则进行特征参数筛选，保留符合要求的特征参数。这些特征构建的估测蓄积量估测模型在结构简单的人工林有较好的估测效果。但天然林或混交林大部分冠层结构复杂，林分高度变异性较大，且存在复层林的情况，因

此需要加入表征冠层高度变异以及垂直结构的特征参数, 如式 (4)。

$$\text{Vol} = \alpha_0 * H^{\beta_5} D^{\beta_6} \text{HH}^{\beta_7} \text{VS}^{\beta_8} + \varepsilon_3 \quad (4)$$

式中, HH为表征冠层高度变异的特征参数, VS为表征垂直结构的特征参数, β_5 、 β_6 、 β_7 、 β_8 为参数项, ε_3 为误差项。

2.3.4 分组特征预筛选及有规则约束的全子集筛选

本研究共计算了92个激光雷达特征, 为减少变量之间的共线性、避免模型过拟合, 设计了一个优化的变量筛选流程 (图3)。全子集回归通过穷举所有变量组合并根据精度指标选择最优模型, 其精度高, 但计算时间成本高, 且容易过拟合。因此, 本研究提出了一种结构化的预筛选策略: 首先, 将变量划分为3类: 基于点云高度分布的特征参数、表征冠层高度及其异质性的特征参数, 以及反映垂直结构的特征参数。随后, 针对不同类别特征对森林蓄积量的贡献差异, 设计了相应的预筛选策略。最后, 将经过初步筛选的特征参数输入全子集回归框架, 以蓄积量估测精度为评价指标, 筛选出最优特征组合并构建最终模型。

基于点云高度分布的特征参数能够表征林分蓄积量的主体特征, 但存在共线性问题, 本研究设计了严格的变量筛选方案: 首先, 根据特征与林分参考蓄积量的相关性、显著性进行前向筛选; 在此基础上使用逐步回归、岭回归的方法进行后向筛选, 筛选后的变量还需进行方差膨胀因子VIF (Variance Inflation Factor) 的筛选, 确保所有变量VIF值<10。表征冠高及异质性的特征参数, 主要采用相关性、显著性进行筛选。垂直结构的特征参数通常与高度特征参数协同使用, 不能直接表征林分蓄积量, 因此不在初步筛选中考虑。最终, 通过全子集回归确定的模型变量包括1—3个基于点云高度分布的特征参数、0—2个表征树冠高度及异质性的特征参数、以及0—2个垂直结构特征参数。

2.3.5 蓄积量估测模型的精度评价

研究采用决定系数 (R^2)、均方根误差 (RMSE, m^3/hm^2) 和相对均方根误差 (rRMSE) 3个指标评估基于ICESat-2蓄积量估测模型精度。采用100次随机十折交叉验证, 提升评估结果的可靠性和普适性。3个指标的计算公式如下:

$$R^2 = 1 - \left(\sum (y_i - \hat{y}_i)^2 \right) / \left(\sum (y_i - \bar{y})^2 \right) \quad (5)$$

$$\text{RMSE} = \sqrt{\left(\sum (y_i - \hat{y}_i)^2 / n \right)} \quad (6)$$

$$\text{rRMSE} = (\text{RMSE} / \bar{y}) \times 100\% \quad (7)$$

式中, y_i 是第*i*个观测值 (ALS数据估测的蓄积量), $\hat{y}_i$ 是第*i*个模型预测值 (ICESat-2数据估测的蓄积量), $\bar{y}$ 是所有观测值的平均值, n 是观测值的总数。 R^2 的值越接近1表示模型的解释能力越强。

在本研究中, 为评估各类激光雷达特征在森林蓄积量估测中的统计贡献, 我们采用基于*P*值的显著性检验, 并在多特征同时检验时对*P*值进行多重比较校正, 统一以经校正的*P*值<0.05作为显著性标准。具体做法为: 将预筛选的候选特征纳入多元线性回归模型, 估计各特征系数并进行显著性检验, 得到相应的*P*值; 在此基础上对*P*值进行校正, 以降低假阳性率。为便于直观比较不同特征的证据强度, 如式 (8), 我们将*P*值转换为以10为底的负对数 $-\log(P)$, 并据此对特征进行排序和可视化, 以保证结果的可读性与可比性。

3 结果与分析

3.1 基于有规则约束的全子集回归方法模型筛选结果

本研究对所有符合条件的变量组合进行了穷举, 并采用十折交叉验证的平均 R^2 作为评价指标, 在36068个子集模型中筛选最优模型。表4为采用有规则约束的全子集回归方法筛选出的10个最优模型的平均精度。其 R^2 为0.776, RMSE为92.482 m^3/hm^2 , rRMSE为0.239。在特征选择方面, 最优模型在高度分布特征组中选取了Dp₆₅ (65%高度分位数), 在冠高及冠高异质性特征组中选择了TCH和LAVH, 在垂直结构特征组中选取了VFP_{mean}。这4个特征参数在前10个最优模型中的选择频率最高, 这不仅反映了这些特征参数的重要性, 也表明了模型在不同数据子集中的稳定性。最优模型方程及系数如下:

$$\text{Vol} = 92.93 \times \text{Dp}_{65} + 22.20 \times \text{TCH} + 7.224 \times \text{LAVH} - 85.113 \times \text{VFP}_{\text{mean}} - 47.167 \quad (8)$$

3.2 最优特征参数显著性

采用有规则约束的全子集回归方法, 从92个激光雷达特征中筛选出4个关键变量 (TCH、VFP_{mean}、

LAVH、 Dp_{65}), TCH 为冠层顶部平均高、 VFP_{mean} 为枝叶剖面的平均值、LAVH 为叶面积加权冠层体积、 Dp_{65} 为 65% 密度分位数, 构建林分蓄积量估测模型。入模后在多元线性回归框架下进行系数检验, 基于 P 值判断变量与蓄积量之间关联的统计显著性。图 4 展示了基于 $-\log(P)$ 的特征排序与可视化结果, 该指标用于量化统计显著性。结果表明, TCH 的 $-\log(P)$ 为 24.98 ($P < 0.0001$), 其与蓄积量的关联在

统计上最为可靠; TCH 通常表征林分的垂直高度与冠层覆盖, 是蓄积量估算的核心结构指标。 VFP_{mean} 与 LAVH 的 $-\log(P)$ 分别为 2.77 ($P \approx 0.002$) 和 2.27 ($P \approx 0.006$), 均达到统计显著; VFP_{mean} 主要反映植被垂直分布特征, LAVH 与叶面积的垂直结构相关。 Dp_{65} 的 $-\log(P)$ 为 0.50 ($P \approx 0.32$), 其刻画林木冠层的水平分布, 能为垂直结构信息提供补充, 故在模型中予以保留。

表 4 前 10 个 ICESat-2 蓄积量估测最优模型及其精度
Table 4 Top ten optimal models of ICESat-2 and their accuracy

序号	基于点云高度分布的特征参数	表征冠高及异质性的特征参数	表征垂直结构的特征参数	R^2	RMSE/(m^3/hm^2)	rRMSE
1	Dp_{65}	TCH, LAVH	VFP_{mean}	0.776	92.481	0.239
2	Dp_{65}	TCH	VFP_{mean} , VFP_{stdev}	0.776	92.616	0.240
3	Dp_{65} , H_{max}	TCH, LAVH	VFP_{mean}	0.776	92.657	0.240
4	Dp_{65}	TCH, LAVH	VFP_{mean} , LAD_{mean}	0.775	92.771	0.240
5	Dp_{65}	TCH, LAVH	VFP_{mean} , LAD_{cv}	0.775	92.999	0.241
6	Dp_{65}	TCH, LAVH	VFP_{mean} , VFP_{cv}	0.775	92.760	0.240
7	Dp_{65}	TCH, LAVH	VFP_{mean} , VFP_{stdev}	0.775	92.771	0.240
8	Dp_{65} , Dp_{50}	TCH, LAVH	VFP_{mean} , VFP_{stdev}	0.775	92.999	0.241
9	Dp_{65}	TCH, CHM_{gini}	VFP_{mean}	0.775	92.999	0.241
10	Dp_{65}	TCH, LAVH	VFP_{mean} , RH_{slope}	0.775	92.840	0.240

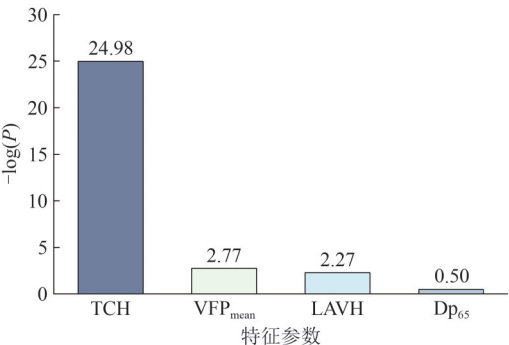


图 4 主要特征参数的 P 值
Fig. 4 P of main feature parameters

3.3 冠高异质性和垂直结构特征对蓄积量估测模型精度的提升

本研究通过对蓄积量估测结果进行 100 次十折交叉验证, 评估了加入冠高及冠高异质性和垂直结构特征后基于 ICESat-2 数据蓄积量估测模型的精度提升效果。图 5 展示了各模型的精度对比。传统模型分别使用逐步回归 ST (Stepwise Regression) 和岭回归 LMR (Linear Model with Ridge regularization) 方法进行变量筛选, 结果显示, 引入冠层高度及其异质性以及表征垂直结构的特征后, 仅使用传

统特征的蓄积量估测模型精度明显提升: 平均 R^2 由 0.68 提高至 0.78, 平均 rRMSE 由 0.28 降至 0.24。因此本研究模型较传统模型具有更高的估测精度。此外, R^2 由 0.20—0.89 改进至 0.47—0.94, rRMSE 由 0.16—0.43 改进至 0.12—0.34, 表明本研究模型的普适性高于传统模型。

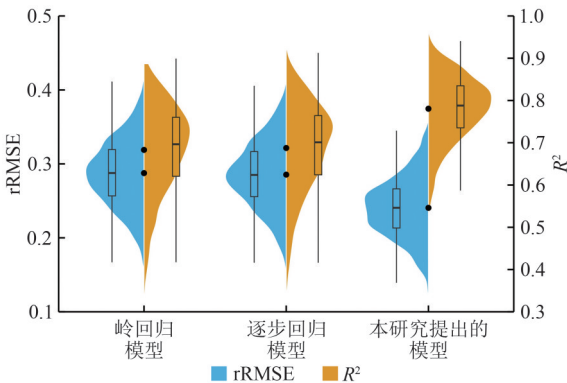


图 5 不同模型蓄积量估测精度
Fig. 5 Accuracy of volume estimation across different models

为准确评估冠层高度异质性变量与垂直结构变量对 ICESat-2 森林蓄积量估测模型的优化效果, 本研究对各估测单元开展了 100 次交叉验证, 并将

预测值与实测值进行对比, 其结果如图6所示。结果表明, 引入上述变量后, 模型整体精度显著提升: R^2 平均值由 0.66 提高至 0.76, 提升幅度约 10%; rRMSE 则由 0.24 降低至 0.22, 相对误差下降约 2%。此外, 为进一步验证新模型的有效性, 我们对参考蓄积量与估测蓄积量进行了拟合分析, 结果显示模型在回归斜率与截距方面均有明显改善: 斜

率由 0.66 提升至 0.76, 而截距由 $125.15 \text{ m}^3/\text{hm}^2$ 降至 $87.24 \text{ m}^3/\text{hm}^2$ 。引入冠层高度异质性与垂直结构变量后, 模型的饱和效应得到一定程度优化。同时, 在图6中标记的样地 (①—④) 处, 本研究提出的模型有效修正了传统基于高度特征模型中部分偏差较大的估测结果。

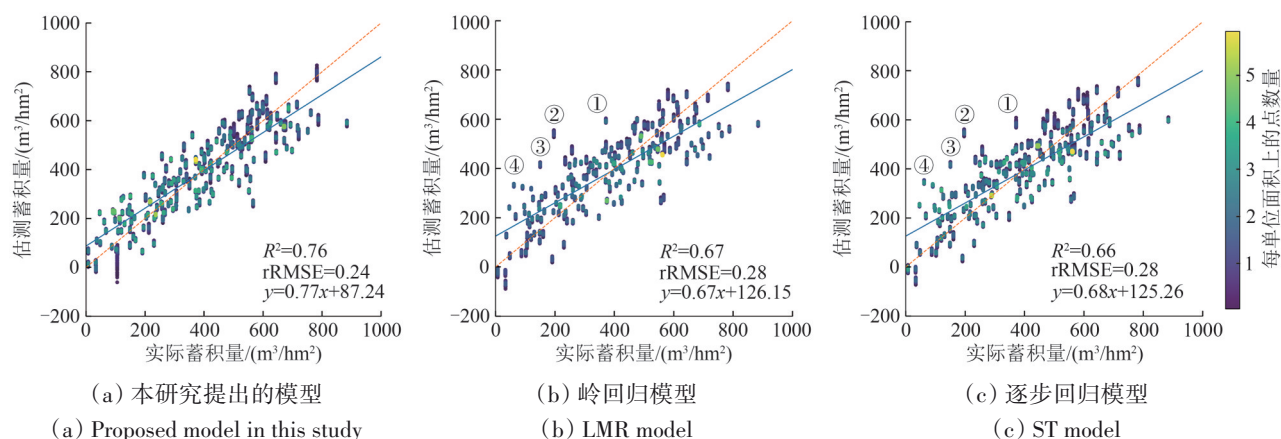


图6 引入冠层高度异质性与垂直结构变量前后森林蓄积量估测模型精度对比

Fig. 6 Comparison of forest volume estimation model accuracy before and after incorporating canopy height heterogeneity and vertical structure variables

4 讨论

4.1 分析传统模型限制性

在传统的 ICESat-2 数据蓄积量估测模型中, 常用的基于点云高度分布的特征未能充分表征森林冠层的异质性 (Bouvier 等, 2015)。在基于 ALS 数据的蓄积量与生物量估测研究中, 冠层高度、冠高异质性以及垂直结构变量已被广泛应用于提升估测精度 (Knapp 等, 2020), 并改善模型在不同树种与区域间的普适性 (Bouvier 等, 2015; Van Ewijk 等, 2020; Tompalski 等, 2019)。基于此, 本研究评估了利用 ALS 数据开发的冠层水平与垂直结构特征改进 ICESat-2 数据森林蓄积量估测精度的可行

性, 结果见图6。可见加入这些特征后, 模型的 R^2 提高了约 10%, 而 rRMSE 的降低幅度相对有限。这可能归因于传统高度变量已能解释蓄积量的主体部分, 但在冠层异质性较高的林分中表现不足。此外, 有 4 个估测单元 (标记为①—④) 在加入冠层高度异质性与垂直结构变量后表现出明显的精度提升。为进一步探讨其原因, 本研究提取了这估测单元对应的机载 LiDAR 点云侧视图, 详见图7, 并分析其林分结构。结果表明, 这些区域大多包含稀疏分布的大树、较高的冠层异质性, 或位于林区边界。在此类条件下, 传统模型往往导致蓄积量的高估, 而引入冠层高度异质性与垂直结构变量后, 模型在对高异质性林分的蓄积量估测方面得到了有效改善。

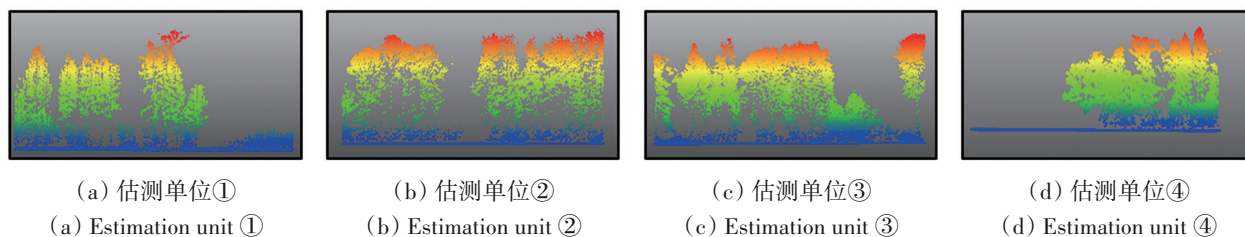


图7 本研究模型精度提升明显的4个估测单元对应的林分情景及机载LiDAR点云侧视图

Fig. 7 Stand scenarios and side views of airborne LiDAR point clouds for estimation units with significant accuracy improvements under the proposed model

4.2 分析最优模型中特征参数对蓄积量估测的贡献

TCH是林分冠层高度的综合指标,相较于传统的高度特征如高度百分位数,TCH受点云密度和传感器功率变化的影响较小(Asner和Mascaro, 2014)。叶面积加权冠层体积LAHV是通过将不同高度层的叶面积密度与冠层体积进行加权累加来计算的(Almeida等, 2019)。TCH和LAHV没有直接显示冠层高度的异质性,但它们通过与胸高断面面积的相关性间接反映了此特性。VFP是根据Harding等(2001)描述的方法从5 m CHM的垂直剖面导出的,描述了叶面剖面的垂直分层(Knapp等, 2020),它不包含地面成分,但会增加较低高度的剖面部分的权重,以补偿高大树木的遮挡。 Dp_{65} 为密度百分位数,体现了点云高度的密度分布。

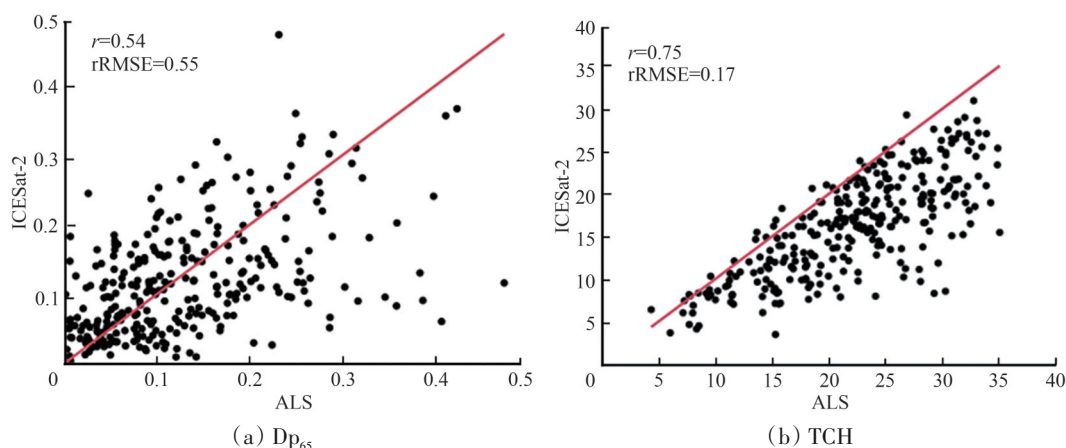
4.3 分析基于ALS与ICESat-2数据计算的特征参数一致性

本研究将基于ALS数据开发的特征参数迁移到ICESat-2数据中,由于两者在点云密度上的显著差异,增加了2—5倍的计算单元尺度,再以皮尔逊相关系数(r)、相对均方根误差($rRMSE$)为评价指标,将分别由ALS和ICESat-2数据计算的特征参数进行一致性分析,结果见图8。可见除TCH之外的其他特征一致性较差(TCH: $r=0.75$, $rRMSE=0.17$; 其他: $r=0.45—0.55$, $rRMSE=0.35—0.55$)。主要原因可能是ALS的LiDAR传感器与ATLAS激光束发射能量、功率以及接收器的灵敏度存在差异(Næsset, 2009)。此外ICESat-2数据噪声的影响也会导致冠层及内部点云分布的随机性。由于点云密度的影响,由ICESat-2数据计算的TCH普遍小于基于ALS计算的TCH,随着点云密度的降低,LiDAR数据记录冠顶的概率随之降低,计算的TCH值也随之降低(Roussel等, 2017),这也是对

估测单元设置点云密度限制的原因之一。其他特征参数与ALS计算的结果差异较大,但是基本分布于1:1线两侧,说明以ALS的方法计算ICESat-2数据特征并不完全一致,但其对蓄积量估测模型的贡献表明,这些变量能够一定程度解释蓄积量的变异性。因此,后续研究中基于ICESat-2数据开发能够准确描述森林冠层异质性和垂直结构的特征参数是提高估测精度的关键。这需要针对ICESat-2数据的特点调整特征提取方法,确保所得特征能够有效反映森林的三维结构和生物量。

4.4 本研究提出方法的应用前景及局限性

ICESat-2的主要优势在于其全球观测能力。本研究旨在评估垂直结构和冠层高度异质性变量对基于ICESat-2数据的林分蓄积量估测模型的改进效果。研究结果显示,在瑞士阿尔高州,这种方法取得了较好的成效。然而,目前研究中的模型系数和结果主要适用于瑞士阿尔高州或具有相似森林条件的区域。在其他地区,模型需要重新拟合参数或进行校准。研究方法的适用性与区域森林条件密切相关。在森林结构简单的林分,如人工林,仅高度变量已足以代表大部分林分蓄积量,而加入垂直结构等变量可能意义不大;反之,在垂直结构复杂的林分,冠层高度异质性和垂直结构变量可为冠层复杂度提供重要信息,从而提高模型的精度。目前研究的样本数量和森林类型的多样性还不足以构建一个普适性模型。影响生物量和蓄积量估测模型普适性的因素众多,包括森林类型、垂直结构复杂性和地形等。Bouvier等(2015)引入垂直结构和冠层异质性变量可以提高模型在不同区域的适用性。然而,由于ICESat-2与机载激光雷达数据在变量上的差异,该模型是否适用于其他区域还需要通过更多的生态分区数据和样本数据来验证。



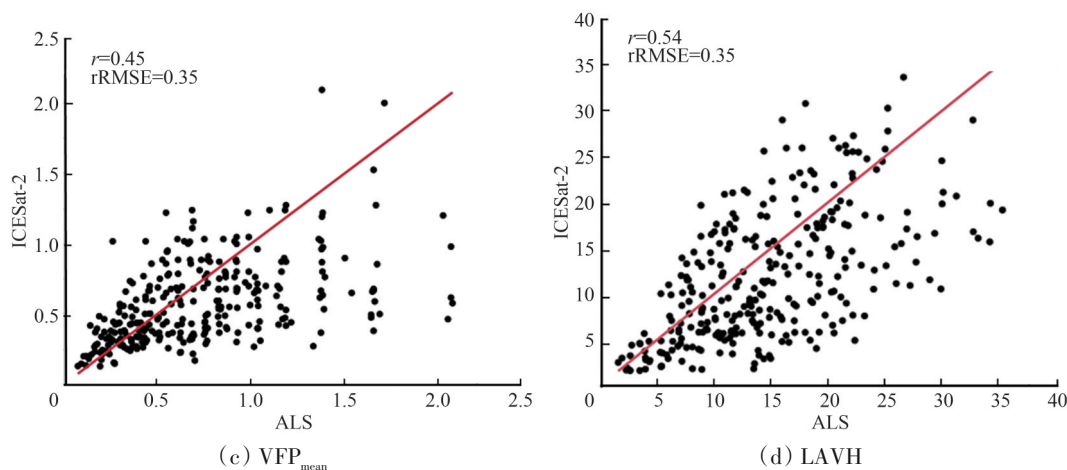


图8 星载光子计数激光雷达与机载激光雷达数据计算的特征参数一致性

Fig. 8 Consistency of feature parameters calculated from ICESat-2 and ALS data

本研究使用 ICESat-2 数据在基于省级尺度（瑞士阿尔高州）进行蓄积量估测方法研究。作为卫星传感器，ICESat-2 具备更广的时间覆盖能力，能够对同一地区进行重复观测，这对监测森林的动态变化具有重要意义。与此相比，机载激光雷达的数据获取不仅需要复杂的飞行计划，还有较高的运行成本，这在大面积或地形复杂的区域尤为明显。此外，ICESat-2 提供的是全球性、免费的数据资源，极大地降低了研究成本，提高了观测效率。在更大的区域尺度上，如国家或洲际级别，ICESat-2 的全球覆盖能力和成本效益将更加突出，能够支撑全球范围内的森林资源监测，为全球森林蓄积量估测和碳循环研究提供重要数据支撑。

5 结 论

本研究以提升基于 ICESat-2 数据的森林蓄积量估测精度为目标，结合冠层高度异质性与垂直结构特征，构建并验证了改进的估测模型。主要结论如下：

（1）冠层高度异质性和垂直结构特征能够显著改善传统基于高度分布特征的 ICESat-2 蓄积量估测模型，对高异质性林分的适应性更强。

（2）基于有规则约束的特征筛选方法能够在保证数据稳定性的前提下有效提取最优特征组合，为 ICESat-2 数据在森林蓄积量估测中的应用提供了可靠途径。

（3）本研究证明了 ICESat-2 数据在省州尺度森林蓄积量监测中的可行性与优势，未来在大尺度森林资源调查和动态监测中具有广阔的推广潜力。

参考文献 (References)

- Almeida D R A, Stark S C, Chazdon R, Nelson B W, Cesar R G, Meli P, Gorgens E B, Duarte M M, Valbuena R, Moreno V S, Mendes A F, Amazonas N, Gonçalves N B, Silva C A, Schiatti J and Brancalion P H S. 2019. The effectiveness of lidar remote sensing for monitoring forest cover attributes and landscape restoration. *Forest Ecology and Management*, 438: 34-43 [DOI: 10.1016/j.foreco.2019.02.002]
- Asner G P and Mascaro J. 2014. Mapping tropical forest carbon: calibrating plot estimates to simple LiDAR metric. *Remote Sensing of Environment*, 140: 614-624 [DOI: 10.1016/j.rse.2013.09.023]
- Bouvier M, Durrieu S, Fournier R A and Renaud J P. 2015. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment*, 156: 322-334 [DOI: 10.1016/j.rse.2014.10.004]
- Chen L, Ren C Y, Zhang B, Wang Z M, Liu M Y, Man W D and Liu J F. 2021. Improved estimation of forest stand volume by the integration of GEDI LiDAR data and multi-sensor imagery in the Changbai Mountains Mixed forests Ecoregion (CMMFE), Northeast China. *International Journal of Applied Earth Observation and Geoinformation*, 100: 102326 [DOI: 10.1016/j.jag.2021.102326]
- Coops N C, Tompalski P, Goodbody T R H, Queinnec M, Luther J E, Bolton D K, White J C, Wulder M A, van Lier O R and Hermosilla T. 2021. Modelling lidar-derived estimates of forest attributes over space and time: a review of approaches and future trends. *Remote Sensing of Environment*, 260: 112477 [DOI: 10.1016/j.rse.2021.112477]
- Du L M, Pang Y, Wang Q, Huang C Q, Bai Y, Chen D S, Lu W and Kong D. 2023. A LiDAR biomass index-based approach for tree- and plot-level biomass mapping over forest farms using 3D point clouds. *Remote Sensing of Environment*, 290: 113543 [DOI: 10.1016/j.rse.2023.113543]
- Goldbergs G, Levick S R, Lawes M and Edwards A. 2018. Hierarchi-

- cal integration of individual tree and area-based approaches for savanna biomass uncertainty estimation from airborne LiDAR. *Remote Sensing of Environment*, 205: 141-150 [DOI: 10.1016/j.rse.2017.11.010]
- Hall S A, Burke I C, Box D O, Kaufmann M R and Stoker J M. 2005. Estimating stand structure using discrete-return lidar: an example from low density, fire prone ponderosa pine forests. *Forest Ecology and Management*, 208(1/3): 189-209 [DOI: 10.1016/j.foreco.2004.12.001]
- Harding D J, Lefsky M A, Parker G G and Blair J B. 2001. Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, 76(3): 283-297 [DOI: 10.1016/S0034-4257(00)00210-8]
- Jiang Y, Ruetschi M, Garnot V S F, Marty M, Schindler K, Ginzler C and Wegner J D. 2023. Accuracy and consistency of space-based vegetation height maps for forest dynamics in alpine terrain. *Science of Remote Sensing*, 8: 100099 [DOI: 10.1016/j.srs.2023.100099]
- Knapp N, Fischer R, Cazcarra-Bes V and Huth A. 2020. Structure metrics to generalize biomass estimation from lidar across forest types from different continents. *Remote Sensing of Environment*, 237: 111597 [DOI: 10.1016/j.rse.2019.111597]
- Lefsky M A. 2010. A global forest canopy height map from the moderate resolution imaging spectroradiometer and the geoscience laser altimeter system. *Geophysical Research Letters*, 37(15): L15401 [DOI: 10.1029/2010GL043622]
- Lefsky M A, Keller M, Pang Y, De Camargo P B and Hunter M O. 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. *Journal of Applied Remote Sensing*, 1(1): 013537 [DOI: 10.1117/1.2795724]
- Liu M and Popescu S. 2022. Estimation of biomass burning emissions by integrating ICESat-2, Landsat 8, and Sentinel-1 data. *Remote Sensing of Environment*, 280: 113172 [DOI: 10.1016/j.rse.2022.113172]
- Meng G, Zhao D, Xu C, Chen J H, Li X W, Zheng Z J and Zheng Y. 2024. Forest aboveground biomass estimation in Zhejiang province combining ICESat-2 and GEDI spaceborne lidar data. *National Remote Sensing Bulletin*, 28(6): 1632-1647 (孟鸽, 赵旦, 许聪, 陈俊华, 李秀纹, 郑朝菊, 曾源. 2024. 联合 ICESat-2 和 GEDI 星载激光雷达数据的浙江省森林地上生物量估算. *遥感学报*, 28(6): 1632-1647) [DOI: 10.11834/jrs.20222120]
- Næsset E. 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment*, 80(1): 88-99 [DOI: 10.1016/S0034-4257(01)00290-5]
- Næsset E. 2009. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*, 113(1): 148-159 [DOI: 10.1016/j.rse.2008.09.001]
- Nandy S, Srinet R and Padalia H. 2021. 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*, 48(14): e2021GL093799 [DOI: 10.1029/2021GL093799]
- Narine L L, Popescu S C and Malambo L. 2019. Synergy of ICESat-2 and landsat for mapping forest aboveground biomass with deep learning. *Remote Sensing*, 11(12): 1503 [DOI: 10.3390/rs11121503]
- Narine L L, Popescu S C and Malambo L. 2020. Using ICESat-2 to estimate and map forest aboveground biomass: a first example. *Remote Sensing*, 12(11): 1824 [DOI: 10.3390/rs12111824]
- Neuenschwander A, Duncanson L, Montesano P, Minor D, Guenther E, Hancock S, Wulder M A, White J C, Purslow M, Thomas N, Mandel A, Feng T, Armston J, Kellner J R, Andersen H E, Boschetti L, Fekety P, Hudak A, Pisek J, Sánchez-López N and Stereńczak K. 2024. Towards global spaceborne lidar biomass: developing and applying boreal forest biomass models for ICESat-2 laser altimetry data. *Science of Remote Sensing*, 10: 100150 [DOI: 10.1016/j.srs.2024.100150]
- Neuenschwander A L and Magruder L A. 2016. The potential impact of vertical sampling uncertainty on ICESat-2/ATLAS terrain and canopy height retrievals for multiple ecosystems. *Remote Sensing*, 8(12): 1039 [DOI: 10.3390/rs8121039]
- Neumann T A, Martino A J, Markus T, Bae S, Bock M R, Brenner A C, Brunt K M, Cavanaugh J, Fernandes S T, Hancock D W, Harbeck K, Lee J, Kurtz N T, Luers P J, Luthcke S B, Magruder L, Pennington T A, Ramos-Izquierdo L, Rebold T, Skoog J and Thomas T C. 2019. The ice, cloud, and land elevation satellite-2 mission: a global geolocated photon product derived from the advanced topographic laser altimeter system. *Remote Sensing of Environment*, 233: 111325 [DOI: 10.1016/j.rse.2019.111325]
- Pang Y, Lefsky M, Sun G Q, Miller M E and Li Z Y. 2008. Temperate forest height estimation performance using ICESat GLAS data from different observation periods. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXVII(Part B7): 777-782
- Queinnec M, White J C and Coops N C. 2021. Comparing airborne and spaceborne photon-counting LiDAR canopy structural estimates across different boreal forest types. *Remote Sensing of Environment*, 262: 112510 [DOI: 10.1016/j.rse.2021.112510]
- Roussel J R, Caspersen J, Béland M, Thomas S and Achim A. 2017. Removing bias from LiDAR-based estimates of canopy height: accounting for the effects of pulse density and footprint size. *Remote Sensing of Environment*, 198: 1-16 [DOI: 10.1016/j.rse.2017.05.032]
- Schneider F D, Ferraz A, Hancock S, Duncanson L I, Dubayah R O, Pavlick R P and Schimel D S. 2020. Towards mapping the diversity of canopy structure from space with GEDI. *Environmental Research Letters*, 15(11): 115006 [DOI: 10.1088/1748-9326/ab9e99]
- Silva C A, Duncanson L, Hancock S, Neuenschwander A, Thomas N, Hofton M, Fatoyinbo L, Simard M, Marshak C Z, Armston J, Luthcke S and Dubayah R. 2021. Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. *Remote Sensing of Environment*, 253: 112234 [DOI: 10.1016/j.rse.2020.112234]

- Song H Y, Xi L, Shu Q T, Wei Z Y and Qiu S. 2023. Estimate forest aboveground biomass of mountain by ICESat-2/Atlas data interacting cokriging. *Forests*, 14(1): 13 [DOI: 10.3390/f14010013]
- Tompalski P, White J C, Coops N C and Wulder M A. 2019. Demonstrating the transferability of forest inventory attribute models derived using airborne laser scanning data. *Remote Sensing of Environment*, 227: 110-124 [DOI: 10.1016/j.rse.2019.04.006]
- van Ewijk K, Tompalski P, Treitz P, Coops N C, Woods M and Pitt D. 2020. Transferability of ALS-derived forest resource inventory attributes between an eastern and western Canadian boreal forest mixedwood site. *Canadian Journal of Remote Sensing*, 46(2): 214-236 [DOI: 10.1080/07038992.2020.1769470]
- Varvia P, Korhonen L, Bruguère A, Toivonen J, Packalen P, Maltamo M, Saarela S and Popescu S C. 2022. How to consider the effects of time of day, beam strength, and snow cover in ICESat-2 based estimation of boreal forest biomass? *Remote Sensing of Environment*, 280: 113174 [DOI: 10.1016/j.rse.2022.113174]
- Varvia P, Saarela S, Maltamo M, Packalen P, Gobakken T, Næsset E, Ståhl G and Korhonen L. 2024. Estimation of boreal forest biomass from ICESat-2 data using hierarchical hybrid inference. *Remote Sensing of Environment*, 311: 114249 [DOI: 10.1016/j.rse.2024.114249]
- Wu Z C and Shi F L. 2023. Mapping forest canopy height at large scales using ICESat-2 and Landsat: an ecological zoning random forest approach. *IEEE Transactions on Geoscience and Remote Sensing*, 61: 4400116 [DOI: 10.1109/TGRS.2022.3231926]
- Yu G R, Zhu J X, Xu L and He N P. 2022. Technological approaches to enhance ecosystem carbon sink in China: nature-based solutions. *Bulletin of Chinese Academy of Sciences*, 37(4): 490-501 (于贵瑞, 朱剑兴, 徐丽, 何念鹏. 2022. 中国生态系统碳汇功能提升的技术途径: 基于自然解决方案. *中国科学院院刊*, 37(4): 490-501) [DOI: 10.16418/j.issn.1000-3045.20220121002]

Improving forest stock volume estimation using three-dimensional canopy structure characteristics based on ICESat-2 ATLAS data in Aargau, Switzerland

KONG Dan^{1,2}, PANG Yong^{1,2}, WANG Zuyuan³, LI Zengyuan^{1,2}

1. Institute of Forest Resource Information Technique, Chinese Academy of Forestry, Beijing 100091, China;

2. Key Laboratory of Forestry Remote Sensing and Information System, National Forestry and Grassland Administration, Beijing 100091, China;

3. Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Birmensdorf 8903, Switzerland

Abstract: Satellite-based photon-counting LiDAR systems, represented by the Advanced Topographic Laser Altimeter System (ATLAS) onboard the Ice, Cloud, and land Elevation Satellite-2 (ICESat-2), enable rapid acquisition of large-scale three-dimensional vegetation information and have been widely applied to forest parameter retrieval. However, their general applicability remains limited. To address this issue, this study focused on mixed coniferous and broadleaved forests in the canton of Aargau, Switzerland, and evaluated the role of canopy horizontal and vertical structural features in improving the accuracy of stand volume estimation from ICESat-2 data under complex forest conditions. Furthermore, we explored the optimal model form for this region and compared it against a baseline model using only conventional height-based statistical metrics. First, the denoised ICESat-2 ATLAS data were segmented into 100 m estimation units, and quality control was performed to identify and remove anomalous units, ensuring stable data quality. Next, feature grouping pre-screening combined with rule-constrained all-subset selection was applied to integrate point cloud height distribution metrics, canopy height and heterogeneity indices, and vertical structural features for stand volume estimation, yielding the optimal feature subset. Results indicated that the best-performing model for stand volume estimation in the Aargau mixed forests comprised the mean top-of-canopy height, the 65% height percentile, the leaf area - weighted canopy volume, and the mean value of the foliage profile. Ten-fold cross-validation demonstrated that this model achieved high accuracy, with an average $R^2=0.78$, $RMSE=92.48 \text{ m}^3/\text{hm}^2$, and $rRMSE=0.24$. By comparison, the baseline model using only traditional metrics yielded an $R^2=0.66$ and an $rRMSE$ reduced from 0.28 to 0.24, confirming that the incorporation of structural features substantially improved the accuracy of ICESat-2-based stand volume estimation, particularly in forests with high canopy heterogeneity. In conclusion, integrating three-dimensional structural attributes significantly enhances the applicability of ICESat-2 data for forest stand volume estimation under complex stand conditions, thereby providing methodological support for large-scale forest volume and carbon stock monitoring.

Key words: forest stock volume, ICESat-2 ATLAS, forest structural features, canopy height heterogeneity, vertical structure, quality control, preliminary grouping feature selection, rule-constrained all subset

Supported by National Key Research and Development Program (No. 2023YFE0105100, 2020YFE0200800)