

基于无人机数据的森林冠层体密度和冠层基高估算

孙浩¹, 郭笑怡^{1,2,3}, 张洪岩^{1,2,3}, 赵建军^{1,2,3}

1. 东北师范大学 地理科学学院, 长春 130024;

2. 长白山地理过程与生态安全教育部重点实验室, 长春 130024;

3. 吉林省遥感信息技术应用创新基地, 长春 130024

摘要: 森林冠层体密度 CBD (Canopy Bulk Density) 和冠层基高 CBH (Canopy Base Height) 是许多火行为模型的关键输入参数。然而, 在中国很少有研究关注这些参数的估算以及在区域的空间分布情况。无人机技术的发展为精细尺度估算 CBD 和 CBH 的空间分布提供了机遇。本研究首先利用野外调查数据计算样地的 CBD 和 CBH; 然后, 利用无人机 LiDAR 点云和多光谱影像, 构建基于面状区域的最优子集和随机森林估算模型, 并对估算结果进行评价; 最后, 绘制研究区的 CBD 和 CBH 空间分布图。研究结果表明: 采用相同数据源和模型时, 估算 CBH 的 R^2 总是高于 CBD。CBD 最优估算方法为融合 LiDAR 和多光谱数据的随机森林模型, R^2 为 0.5142, RMSE 为 0.0773 kg/m³, rRMSE 为 40.73%。CBH 的最优估算方法为仅使用 LiDAR 数据的随机森林模型, R^2 为 0.6477, RMSE 为 1.6245 m, rRMSE 为 31.17%。使用单一数据源时, LiDAR 估算精度明显高于多光谱数据。融合两种数据源不一定提升 CBD 和 CBH 的估算精度。本研究中构建的最优子集模型需要 3—6 个特征变量, 随机森林模型则需要输入 10—52 个特征变量。仅使用多光谱影像估算 CBD 和 CBH 时, 最优子集回归估算精度更好, 但是空间预测结果易受地表覆盖类型的影响。本研究能够为森林冠层可燃物参数估算提供方法参考, 同时也可以为林火行为预测模型提供精细尺度的输入数据。

关键词: 遥感, LiDAR, 多光谱数据, 冠层体密度, 冠层基高, 最优子集模型, 随机森林模型

中图分类号: TP701/P2

引用格式: 孙浩, 郭笑怡, 张洪岩, 赵建军. 2024. 基于无人机数据的森林冠层体密度和冠层基高估算. 遥感学报, 28(12): 3107–3122

Sun H, Guo X Y, Zhang H Y and Zhao J J. 2024. Estimating canopy bulk density and canopy base height using UAV LiDAR and multispectral images. National Remote Sensing Bulletin, 28 (12): 3107–3122 [DOI: 10.11834/jrs.20233094]

1 引言

中国森林生态系统受火干扰频繁 (Adams 和 Shen, 2015)。1987 年—2007 年, 平均每年林火数量超过 8000 次, 年均燃烧面积近 4×10^6 ha (Chang 等, 2015)。2019 年和 2020 年春季, 四川省凉山彝族自治州两次森林大火, 不仅造成森林损失, 并导致严重的人员伤亡 (吕蓓茹 等, 2021)。2022 年 8 月, 受到极端高温天气影响, 重庆市多地先后发生山火。准确预测林火蔓延行为对森林管理、林火扑救和降低受灾损失至关重要 (徐奔奔 等, 2022)。

许多研究采用模型预测林火行为, 常见的火行为模型包括: BEHAVE (Burgan 和 Rothermel, 1984)、FARSITE (Finney, 1998) 和 FlamMap (Finney, 2006) 等。尤其是 FARSITE 模型已经在中国得到了较多的应用 (吴志伟 等, 2012; 钱兰 等, 2020; 徐奔奔 等, 2022)。冠层体密度 CBD (Canopy Bulk Density) 和冠层基高 CBH (Canopy Base Height) 的空间分布数据是利用模型进行火行为预测的关键输入参数, 二者决定了引发林火的阈值范围, 是林分之间潜在火灾蔓延指标 (Riaño 等, 2004; Engelstad 等, 2019)。在火行为模型中, CBD 是指从树冠底部到顶部, 以单位高度为垂直间隔, 一

收稿日期: 2023-03-30; 预印本: 2023-08-25

基金项目: 国家自然科学基金 (编号: 42071359); 吉林省教育厅科学技术研究项目 (编号: JJKH20211291KJ)

第一作者简介: 孙浩, 研究方向为环境遥感。E-mail: sunh699@nenu.edu.cn

通信作者简介: 郭笑怡, 研究方向为环境遥感与时空数据建模。E-mail: guoxy914@nenu.edu.cn

定高度范围内的滑动平均的最大值; CBH是指特定CBD阈值(通常设定为 $30 \text{ lb/acre} \cdot \text{ft} \approx 0.012 \text{ kg/m}^3$)的最低高度(Scott和Reinhardt, 2001),如图1所示。在中国缺少上述参数的空间数据,限制了火行为模型的应用。CBD和CBH可以通过野外调查数据推算,但是耗时长、成本高,且难以进行空间外推,不适用于大范围的数据获取(Chamberlain等, 2021)。遥感技术具有快速、宏观和时效性强的优势,能够反映森林分布的空间异质性,为估算火行为建模所需要输入的森林冠层可燃物参数提供可能(Erdody和Moskal, 2010)。

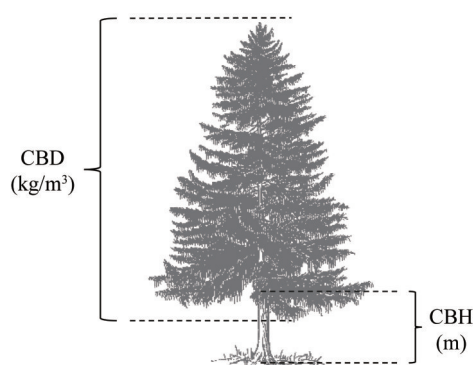


图1 火行为模型中CBD与CBH定义示意图

Fig. 1 Depiction of CBD and CBH used for simulation of fire behavior

光学卫星影像能够以较低的成本,在大范围内估算森林冠层可燃物参数,并持续更新这些参数的变化。以LANDFIRE数据产品为例,它以Landsat影像为基础,采用分类回归树的方法估计30 m空间分辨率的可燃物信息(Rollins, 2009)。然而,该数据对林分冠层可燃物参数的估计精度较低,仅能作为国家和区域尺度的基准参考,限制了其在精细尺度林火行为模拟中的应用(Reeves等, 2009)。利用Landsat 5影像在拉森火山国家公园估算CBD效果较好,决定系数(R^2)为0.63,其中NDVI和穗帽变换后的各变量对其估算的重要性较高,但是对CBH估算的效果较差(Pierce等, 2012)。在希腊沿海森林, Landsat 5影像估算CBH的 R^2 达到了0.51(Palaiologou等, 2013)。在西班牙西北部, Arellano-Pérez等(2018)利用Sentinel-2A卫星影像的2—8、8A、11、12波段数据以及5个植被指数,结合野外实测数据建立非参数的随机森林模型, CBD、CBH模型的 R^2 分别为0.31和0.48。由此可见,利用中分辨率光学遥感数据在不

同区域对CBD和CBH两个参数的估算精度不稳定,作为输入参数可能对火行为模拟产生较大的不确定性。

LiDAR点云数据具有独特的穿透树冠能力,可获取三维冠层结构,具有光学遥感数据无法比拟的优势,可以弥补森林冠层结构信息估算困难的问题(Lim等, 2003; 李增元等, 2016)。Andersen等(2005)使用机载LiDAR数据基于面状区域(area-based)选取最大高度、平均高度和变异系数等8个特征因子,建立森林冠层可燃物参数的回归模型,对CBD和CBH估算的 R^2 分别为0.84和0.77。近年来,机器学习算法也被用于LiDAR点云估算冠层可燃物参数。在美国科罗拉多中部森林地区,利用LiDAR预测冠层可燃物时,随机森林模型对CBD和CBH的变化具有46%和28%的解释能力(Bright等, 2017)。利用随机森林-k最邻近算法对低密度的LiDAR点云森林冠层可燃物建模的结果表明, CBH估算的 R^2 为0.7; CBD估算精度较低, R^2 为0.48(Engelstad等, 2019)。此外, Chamberlain等(2021)利用LiDAR点云进行单木分割,使用最大5%的间隔高度作为特征变量,结合回归方法估算CBH,而后利用树高一胸径关系方程和森林植被模拟器中的火与可燃物扩展模块FVS-FFE(Forest Vegetation Simulator-Fire and Fuels Extension)得到CBD的估计。结果显示CBH的 R^2 为0.4,均方根误差RMSE(Root Mean Square Error)为1.25; CBD的 R^2 达到0.76, RMSE为0.021 kg/m^2 。Luo等(2018)提出一种利用廓线方法直接估算针叶林单树CBH模型,其 R^2 达到0.88, RMSE为1.62 m。上述研究都是利用航空机载LiDAR数据估算冠层可燃物参数,但针对不同参数的估算精度并不稳定,且在各研究区之间存在较大差异。

随着无人机遥感技术的发展,其以小型化、高空间分辨率、高便捷性以及高性价比等特点,在局部地表参数精细化遥感研究中发挥重要作用(李德仁和李明, 2014),已成为航天和航空遥感平台的重要补充。Shin等(2018)研究显示,无人机光谱影像结合运动结构恢复SfM(Structure-from-Motion)方法生成点云数据,而后采用单木分割方法估算CBD和CBH的精度不高。因此需要进一步了解无人机LiDAR和多光谱影像基于面状区域估算方法的适用性。

本文以吉林省蛟河市青背林场为研究区,综合利用无人机获取的LiDAR点云与多光谱影像构建面状区域特征变量,结合地面实测数据,采用最优子集和随机森林模型对CBD和CBH两种森林冠层可燃物参数进行估算,并对估算结果进行评价。以期为中国区域森林CBD和CBH估算提供方法支撑,进一步为林火行为预测模型提供可靠的输入参数。

2 数据与方法

2.1 研究区概况

研究区位于吉林省蛟河市青背林场(图2),地理位置在 43.2°N — 43.56°N , 127.09°E — 127.62°E 之间,为长白山地向松嫩平原过渡的低山丘陵区。在林场内设定面积约为 1 km^2 区域开展无人机飞行和野外实测数据采集。研究区为中温带大陆性季风气候,年平均气温 3.4°C ,降水量约 700 mm 。优势树种为落叶松(*Larix gmelinii*)和樟子松(*Pinus sylvestris*),伴有少量的蒙古栎(*Quercus mongolica*)、黑桦(*Betula dahurica*)和红松(*Pinus koraiensis* Siebold et Zuccarini)等。数据采集区域地势南高北低,海拔范围在 290 — 340 m 。

2.2 数据来源与处理

野外数据采集工作于2021年8月17日至2021年8月20日进行,天气晴朗,对研究区实施无人机飞行和林木地面测量。

2.2.1 LiDAR数据

使用大疆经纬Matrice 600 Pro无人机,搭载LR1601-IRIS激光雷达点云数据采集系统获取原始LiDAR数据。飞行高度为 80 m ,速度为 4 m/s ,平均点密度约 100 pts/m^2 。为提高无人机飞行期间的定位精度,在数据采集过程中使用GNSS(Global Navigation Satellite Systems)动态后差分方法进行定位。无人机在空中飞行期间同步采集基站静态数据,飞行完成后连接CORS基站接收差分信号,测量静态基站位置信息。地面基站静态数据使用南方S86 RTK测量系统采集。

通过对IMU/GPS惯导数据和地面基站测量数据进行处理,获取差分后的POS数据,并与原始LiDAR数据相结合,获得初始点云数据。采用

TerraSolid软件依次进行航带拼接、裁剪和噪声点去除,进一步将点云分为地面点和植被点,利用地面点生成数字高程模型,并用于点云高度归一化处理,获得消除地形影响的标准点云数据。

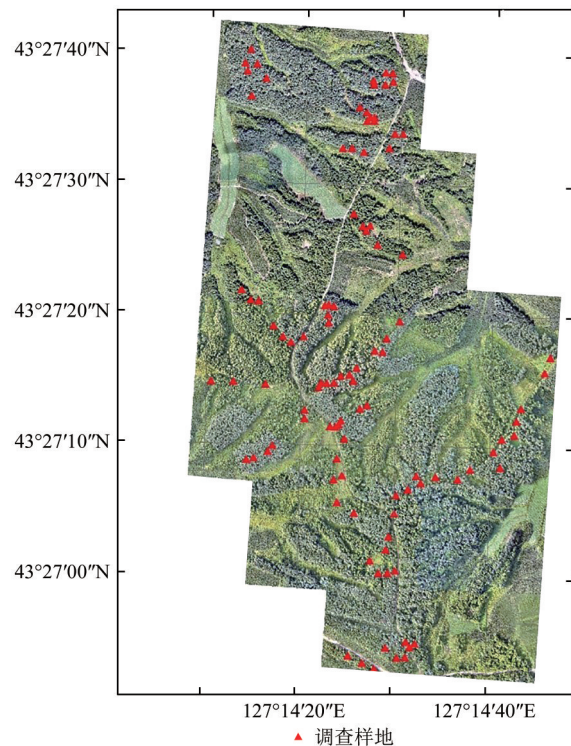


图2 野外获取的无人机真彩色(R:红光波段,G:绿光波段,B:蓝光波段)数字正射影像和地面调查样地分布

Fig. 2 True color UAV multispectral image of the study area and distribution of filed sites

2.2.2 多光谱影像

采用大疆精灵4无人机获取多光谱影像,其传感器集成了蓝光($450\pm 16\text{ nm}$)、绿光($560\pm 16\text{ nm}$)、红光($650\pm 16\text{ nm}$)、红边($730\pm 16\text{ nm}$)和近红外($840\pm 26\text{ nm}$)5个光谱波段。飞行高度设置为 100 m ,航向和旁向重叠度分别为 80% 和 75% ,地面分辨率为 4 cm 。在每次飞行前,利用标准白板进行辐射定标,用于光谱数据反射率计算。利用大疆智图软件,对原始多光谱影像进行拼接和校正,获得研究区的数字正射影像。

2.2.3 地面调查与样地CBD和CBH计算

在对研究区进行全面踏查的基础上,建立 106 块圆形调查样地,每块样地的半径为 5.64 m ,样地范围不重叠(图2)。对每块样地内胸径大于 5 cm 的立木,测量其树高、胸径和枝下高(位于树冠

最底端活枝的高度), 并进行编号和记录树种名称。为保证野外测量的位置精度, 同样采用GNSS后差分方式进行精确定位。

根据实测数据和各树种的树高一胸径关系方程计算单株树木的生物量。而后, 将生物量、树高和枝下高输入FVS-FFE (Scott 和 Reinhardt, 2001; Reinhardt 和 Crookston, 2003) 计算样地 CBD 和 CBH 作为真实值, 为后续建模提供数据。

为了改进建模数据的正态性, 采用Box-Cox变换对样地尺度的CBD和CBH进行计算, 以了解何种变换形式较为合适。Box-Cox变换的一般形式为

$$y^{(\lambda)} = \begin{cases} \frac{y^\lambda - 1}{\lambda}, & \lambda \neq 1 \\ \ln y, & \lambda = 1 \end{cases} \quad (1)$$

式中, y 为预测变量, λ 为变换参数。本研究采用

最大对数似然函数求解 λ 。

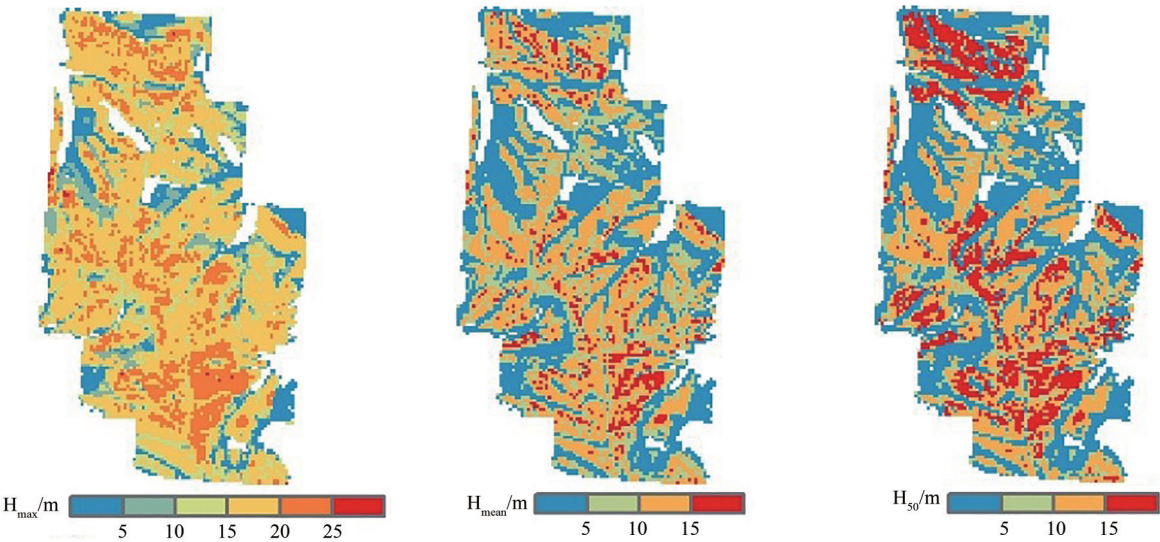
2.3 研究方法

2.3.1 特征变量构建

根据已有研究, 从LiDAR数据的高度、强度和结构3个方面选择33个特征变量, 多光谱数据从光谱反射率、植被指数和纹理特征3个方面选择38个特征变量, 共计71个特征变量, 如表1所示。所有特征均在10 m空间分辨率的水平上进行统计。5个光谱反射率和11个植被指数需计算均值和标准差, 因此形成10个光谱特征和22个植被指数特征。图3和图4分别展示了部分LiDAR和多光谱影像计算的可见光大气阻抗指数VARI (Visible Atmospherically Resistant Index) 特征变量示例。

表 1 LiDAR点云和多光谱影像的特征变量
Table 1 Features of LiDAR and multispectral image

数据源	特征类型	数量	特征变量
LiDAR	高度特征	14	最大点云高度、平均点云高度、点云高度变异系数、点云高度标准差、点云高度的偏度、点云高度的峰度、点云平均高度减点云高度标准差、冠层点云平均高度、冠层点云方差、点云高度百分位数(25%、50%、75%)、点云高度四分位数极差(75%高度分位数减25%高度分位数)、最大间隔分位数的高度(按5%递增)
	结构特征	9	冠层覆盖度、冠层孔隙率、叶面积指数、叶面积密度平均值、叶面积密度方差、叶面积密度百分位数(25%、50%、75%)、垂直复杂度指数
	强度特征	10	点云强度最大值、点云强度平均值、点云强度变异系数、点云强度标准差、点云强度的偏度、点云强度的峰度、点云强度百分位数(25%、50%、75%)、冠层点云强度总和
多光谱影像	光谱反射率	10	红光波段、绿光波段、蓝光波段、红边波段、近红外波段(所有光谱均计算均值和标准差)
	植被指数	22	可见光大气阻抗植被指数、简单比值指数、归一化差值植被指数、归一化差值红边指数、绿色归一化差值植被指数、增强植被指数、优化的土壤调节植被指数、结构不敏感色素植被指数、叶绿素植被指数、红边叶绿素指数、绿色叶绿素指数(所有植被指数均计算均值和标准差)
	纹理特征	6	同质性、相关性、熵、对比度、相异性、角二阶矩



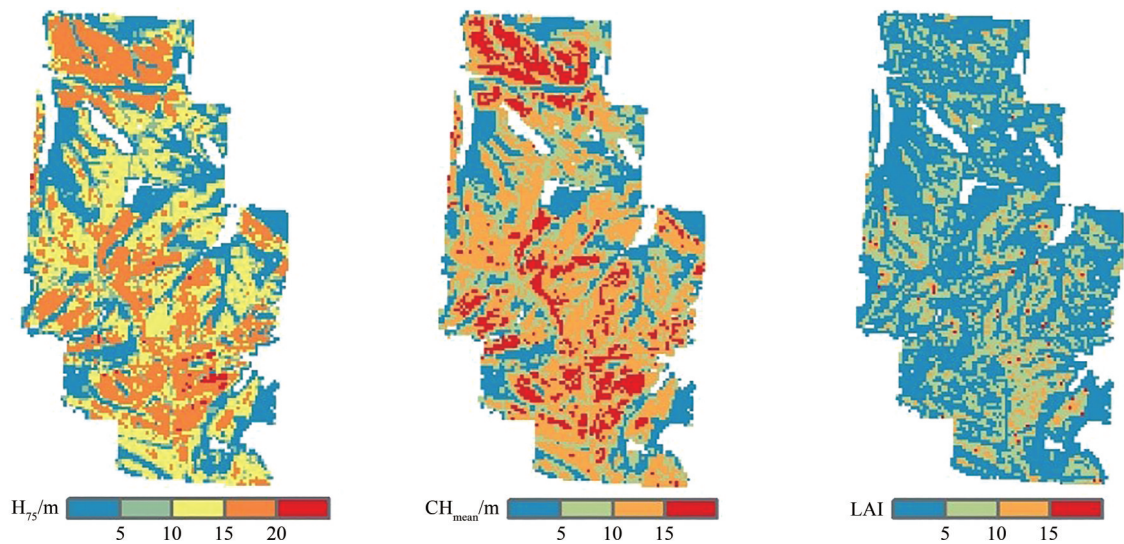


图3 部分LiDAR特征变量空间分布
Fig. 3 Spatial distribution of partial LiDAR features

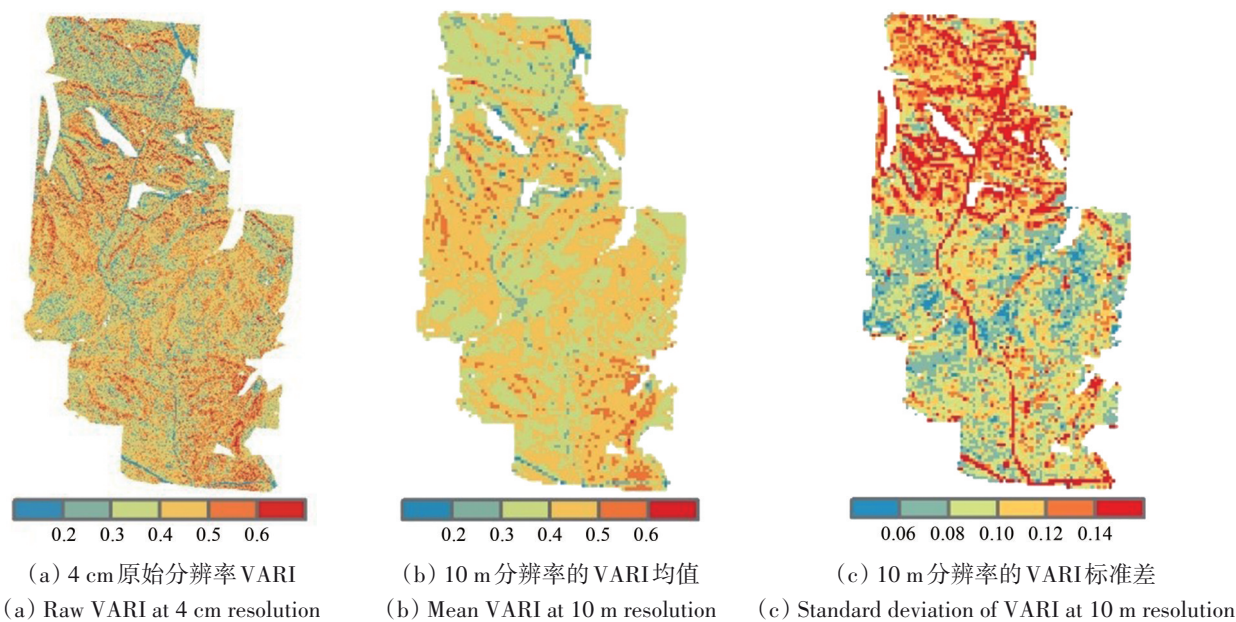


图4 不同空间分辨率的VARI分布
Fig. 4 Distribution of VARI at various spatial resolution

2.3.2 估算模型

(1) 最优子集。最优子集是一种多元线性回归方法，可建立多个特征变量与CBD和CBH之间的线性关系。为了得到有效简洁的线性回归模型，最优子集模型可对所有预测变量的可能组合模型都进行尝试，筛选出最优变量组合，通过所有子集中某一特征变量能够引起 R^2 变化的平均值作为该特征变量的相对重要性指标。然后，利用回归方程的显著性 (p)、 R^2 和贝叶斯信息准则BIC (Bayesian Information Criteria)，以及最优子集回归

方程中各特征变量的膨胀方差VIF (Variance of Inflation) 综合评价符合要求的最优子集模型。模型显著性需达到0.001水平，拥有相对高的 R^2 和相对低的BIC值，同时预测变量间的 $VIF < 10$ ，被认为是最优的拟合模型。

(2) 随机森林。随机森林是一种流行的集成机器学习算法，可以利用一系列特征变量训练多棵相关性较小的决策树，并将多棵决策树的结果进行融合，以提升预测精度。该算法可以处理高维数据的非线性关系，并提供特征变量的重要性

排序,在诸多研究中被证明具有较好的预测精度。本研究利用5次重复十折交叉验证平均误差最小方法确定最终输入到随机森林模型中的特征数量及组合。对随机森林模型影响较大的参数包括决策树的数量、单棵决策树中最大解释变量个数和单棵树的最小叶子节点数3个参数。经过多次试验后,上述3个参数分别设置为500、7和5。采用节点纯度增加量评估各变量的相对重要性,节点纯度越高表示该变量的重要性越大。对于随机森林中每棵树模型,使用最小残差平方和度量节点纯度增加量,在父节点分裂时,力求分裂后残差平方和下降最多。残差平方和的计算方法为

$$i(\tau) = \sum_{m \in \tau} (y_m - \bar{y}(\tau))^2 \quad (2)$$

式中, $i(\tau)$ 表示节点的残差平方和, y_m 为在节点 τ 样本的预测值, $\bar{y}(\tau)$ 为 τ 内 m 个样本的均值。节点纯度增加量为

$$\text{IncNodePurity} = \frac{1}{T} \sum_{x_i \in T} (i_v(\tau) - i_v(\tau_L) - i_v(\tau_R)) \quad (3)$$

式中, T 表示包含特征变量 x_i 的决策树的数量, $i_v(\tau)$ 表示父节点的残差平方和, $i_v(\tau_L)$ 和 $i_v(\tau_R)$ 分别表示左、右子节点的残差平方和。

2.3.3 精度评价

为了定量评估模型估算结果,采用十折交叉验证的 R^2 、RMSE 和相对均方根误差 rRMSE (relative Root Mean Square Error) 3 个指标评估无人机数据

估算的结果。各指标计算公式如下:

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

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (5)$$

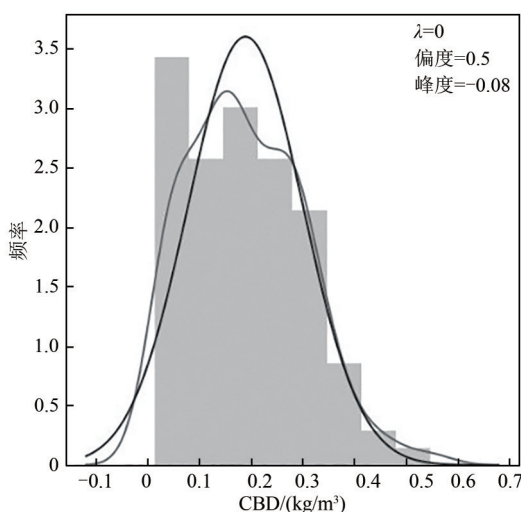
$$\text{rRMSE} = \frac{\text{RMSE}}{\bar{y}} \times 100\% \quad (6)$$

式中, $\hat{y}_i$ 表示估算值, $\bar{y}$ 为平均测量值, y_i 为测量值, n 为样本的数量。

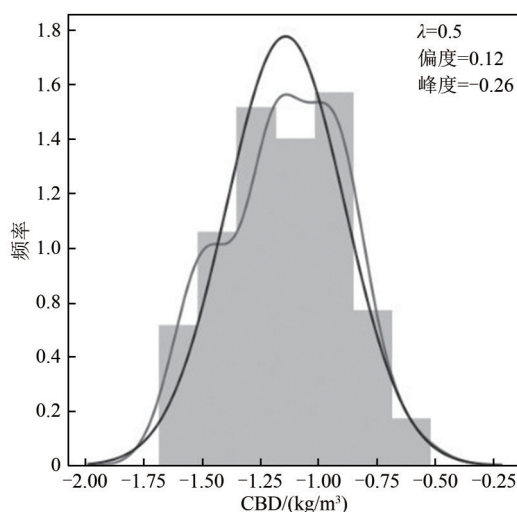
3 结果与分析

3.1 样地水平数据变换

对样地水平的 CBD 和 CBH 数据进行 Box-Cox 变换,通过最大对数似然函数求解得到 λ 值分别为 0.53 和 0.50,为了便于构建回归方程,本研究中 λ 取值为 0.5 较为合适。图 5 为经过 0.5 次幂变换前后,样地水平的 CBD 和 CBH 数值分布特征对比情况。根据直方图可以看出,在变换之前 CBD 和 CBH 峰值都表现出偏向低值的特征。经过变换后样地水平数据 CBD 和 CBH 的正态性、对称性都得到不同程度的改善,且在低值处数据分布的正态性改善更加明显。因此,经变换后样地水平的 CBD 和 CBH 更适合线性回归模型,满足其正态性及方差齐次性的假设。



(a) CBD 变换前分布
(a) Distribution of raw CBD



(b) CBD 变换后分布
(b) Distribution of transformed CBD

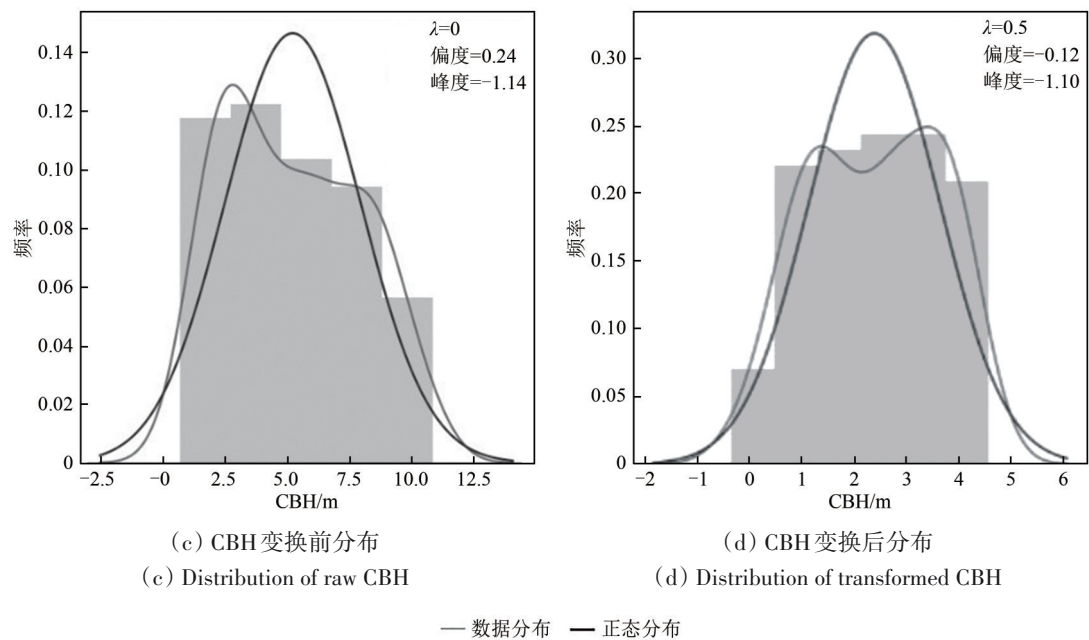


图 5 样地 CBD 和 CBH 变换前后数据分布对比

Fig. 5 Comparison between distribution of raw CBD and CBH and transformed CBD and CBH at plot level

3.2 估算模型及精度评价

利用最优子集模型构建不同遥感数据源特征变量与样地水平 0.5 次幂的 CBD 和 CBH 之间的多元线性回归模型，从符合显著性要求的全部方程中筛选出各特征变量 $VIF < 10$ ，并且具有最小 BIC 值

的多元线性回归方程，结果如表 2 所示。所有被选择的回归方程显著性均达到 0.0001 水平，各种数据源经筛选后的最优特征子集包括 3—6 个变量。相同数据源估算 CBD 和 CBH 时，所使用的特征变量存在明显差异。

表 2 估算 CBD 和 CBH 的最优子集模型

Table 2 The best subset regression of estimation of CBD and CBH for various data sources

数据源	<i>P</i>	BIC	多元线性回归方程
LiDAR	<0.0001	-199.85	$CBD^{0.5} = 0.029H_{mean} + 0.017H_{std} - 0.280VCI + 0.235$
多光谱影像	<0.0001	-189.44	$CBD^{0.5} = 0.783EVI_{mean} - 0.00005Rededge_{std} - 0.630Homogeneity - 1.331VARI_{std} - 2.790VARI_{mean} + 0.511$
LiDAR+多光谱	<0.0001	-207.29	$CBD^{0.5} = 0.019H_{mean} + 0.001Blue_{mean} - 1.342OSAVI_{mean} - 1.414NDVI_{std} - 1.763Correlation + 1.969$
LiDAR	<0.0001	99.97	$CBH^{0.5} = 0.112H_{mean} + 0.095H_{max} - 0.038H_{kurt} - 0.069LAI - 2.842VCI + 1.802$
多光谱影像	<0.0001	131.25	$CBH^{0.5} = 4.994Cl_{rededge_{std}} - 0.006Red_{std} + 0.019Blue_{std} - 0.140SR_{mean} + 1.712$
LiDAR+多光谱	<0.0001	95.96	$CBH^{0.5} = 0.117H_{max} + 0.051H_{50} + 0.004Blue_{std} - 0.00003I_{CRS} - 0.051LAI - 1.918VCI + 0.672$

为了评估最优子集和随机森林模型估算 CBD 与 CBH 的精度，利用十折交叉验证的方法对建模结果进行评价，计算了 R^2 、RMSE 和 rRMSE，结果如图 6 所示。根据散点图趋势线可以发现，所有结果总体上呈现低值高估、高值低估的现象，且在多光谱数据中尤为明显。应用相同的数据源和模型时，CBD 的估算精度总是低于 CBH。精度最高的 CBD 估算方法为融合数据源的随机森林模型， R^2 为 0.5142，RMSE 为 0.0773 kg/m³，rRMSE 为

40.73%。精度最高的 CBH 的估算方法为 LiDAR 数据源的随机森林模型， R^2 为 0.6477，RMSE 为 1.6245 m，rRMSE 为 31.17%。使用相同的模型估算 CBD 和 CBH 时，LiDAR 数据的精度比多光谱数据更好。单独使用 LiDAR 数据估算结果最低 R^2 为 0.4777，多光谱数据估算结果的最大 R^2 为 0.4864。最优子集回归在多光谱影像上表现的比随机森林模型更好。随机森林模型针对不同的估算参数和不同的数据源，精度各有优劣。

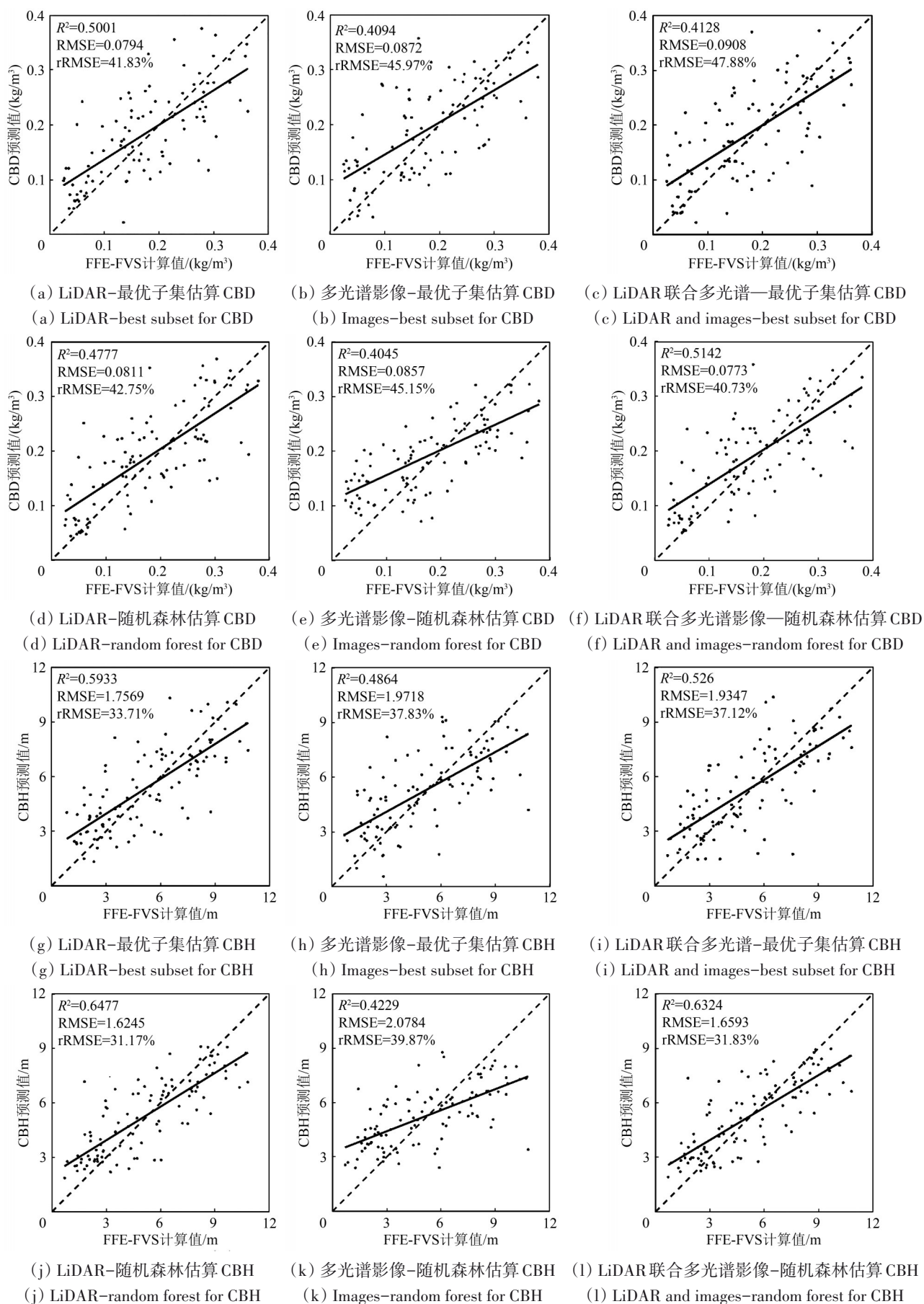


图6 最优子集和随机森林模型估算CBD与CBH精度评估

Fig. 6 Accuracy of estimates of CBD and CBH using the best subset regression and random forest

3.3 特征变量的相对重要性

表3显示了利用最优子集模型筛选的CBD和CBH特征变量。当仅使用LiDAR数据时，筛选出的特征均为高度特征和结构特征。平均点云高度(H_{mean})和最大点云高度(H_{max})两个特征分别对估算CBD和CBH显示出最大的重要性，前者对CBD估算的相对重要性超过80%，且在估算CBH时也体现了较高的重要性。CBD和CBH的最优特征子集也同时包含了垂直复杂指数VCI (Vertical complexity Index)。多光谱影像估算CBD和CBH筛

选出的最优特征子集所包含的变量完全不同。CBD的最优特征子集主要由植被指数特征构成，CBH的最优特征子集则包含了两个光谱特征。联合使用LiDAR和多光谱影像时，CBD和CBH所选择的特征来源差异明显。CBD的最优特征子集仅有一个LiDAR特征(H_{mean})，相对重要性超过50%，其他特征均来自多光谱影像。相反，CBH的特征子集以LiDAR数据的高度和结构特征为主，另外包含冠层点云反射强度总和(I_{CRS})和蓝光波段的标准差(B_{std})。

表 3 最优子集模型选择的特征变量重要性
Table 3 The importance of selected features using the best subset regression

数据源	CBD				CBH			
	特征变量	<i>P</i>	VIF	相对重要性/%	特征变量	<i>P</i>	VIF	相对重要性/%
LiDAR	H_{mean}	<0.01	1.10	82.84	H_{max}	<0.01	2.50	34.17
	VCI	<0.01	1.09	14.26	H_{mean}	<0.01	4.18	29.72
	H_{std}	0.01	1.06	2.90	VCI	<0.01	1.66	24.47
	—	—	—	—	LAI	<0.01	3.00	6.90
	—	—	—	—	H_{kurt}	<0.01	2.60	4.74
多光谱影像	$\text{VARI}_{\text{mean}}$	<0.01	4.68	38.81	SR_{mean}	<0.01	1.23	39.24
	Homogeneity	<0.01	1.12	31.81	Blue_{std}	<0.01	5.79	32.95
	EVI_{mean}	<0.01	4.69	15.98	$\text{Cl.rededge}_{\text{std}}$	<0.01	1.23	14.31
	VARI_{std}	<0.01	1.13	9.81	Red_{std}	<0.01	5.40	13.51
	$\text{Rededge}_{\text{std}}$	<0.01	1.09	3.59				
LiDAR+多光谱	H_{mean}	<0.01	1.68	53.28	H_{max}	<0.01	1.94	30.16
	$\text{Blue}_{\text{mean}}$	<0.01	2.20	20.38	H_{50}	<0.01	2.24	24.07
	Correlation	<0.01	1.32	14.88	VCI	<0.01	1.12	18.61
	$\text{OSAVI}_{\text{mean}}$	<0.01	2.75	6.92	Blue_{std}	<0.01	1.32	15.61
	NDVI_{std}	<0.01	1.93	4.55	I_{CRS}	<0.01	1.38	7.58
	—	—	—	—	LAI	<0.01	1.69	3.97

注： H_{75} 、 H_{50} 、 H_{25} 表示点云高度百分位数。

与最优特征子集方法筛选结果相比，随机森林模型需要更多的特征变量以获得较好的估算精度。经随机森林模型特征选择后，利用LiDAR数据估算CBD和CBH分别需要10和32个特征变量，多光谱影像需要15和28个特征变量，两种数据联合使用时则需要32和52个特征变量。由于随机森林模型最少仅需要10个特征变量，以及篇幅限制，因此图7中仅显示了相对重要性排在前十位的特征变量。利用随机森林模型得到具有较高重要性的LiDAR变量以高度特征为主，尤其是估算CBD时所采用的全部为高度特征。对二者估算重要性较高的相同变量包括冠层高度平均值(CH_{mean})、点云高度百分位数(H_{75} 、 H_{50} 、 H_{25})、 H_{max} 和点云高度

四分位数极差(H_{QD})。在估算CBH时，LiDAR的VCI、叶面积密度分位数(LAD_{75})和叶面积密度变异系数(LAD_{cv})3个结构特征变量也具有相对高的重要性。多光谱影像在估算CBD和CBH时，所选择的特征变量较为复杂，估算CBD时纹理特征变量的重要性排名较高，而CBH则以植被指数和光谱特征为主。其中，可见光大气阻抗植被指数均值($\text{VARI}_{\text{mean}}$)、归一化差值红边指数均值($\text{NDRE}_{\text{mean}}$)和简单比值指数均值(SR_{mean})对CBD和CBH估算都具有较高的重要性。两种数据联合估算CBD和CBH时，重要性较高的特征仍然以LiDAR变量占主导位置，并且排名前5位的重要特征变量与仅使用LiDAR数据时保持一致。

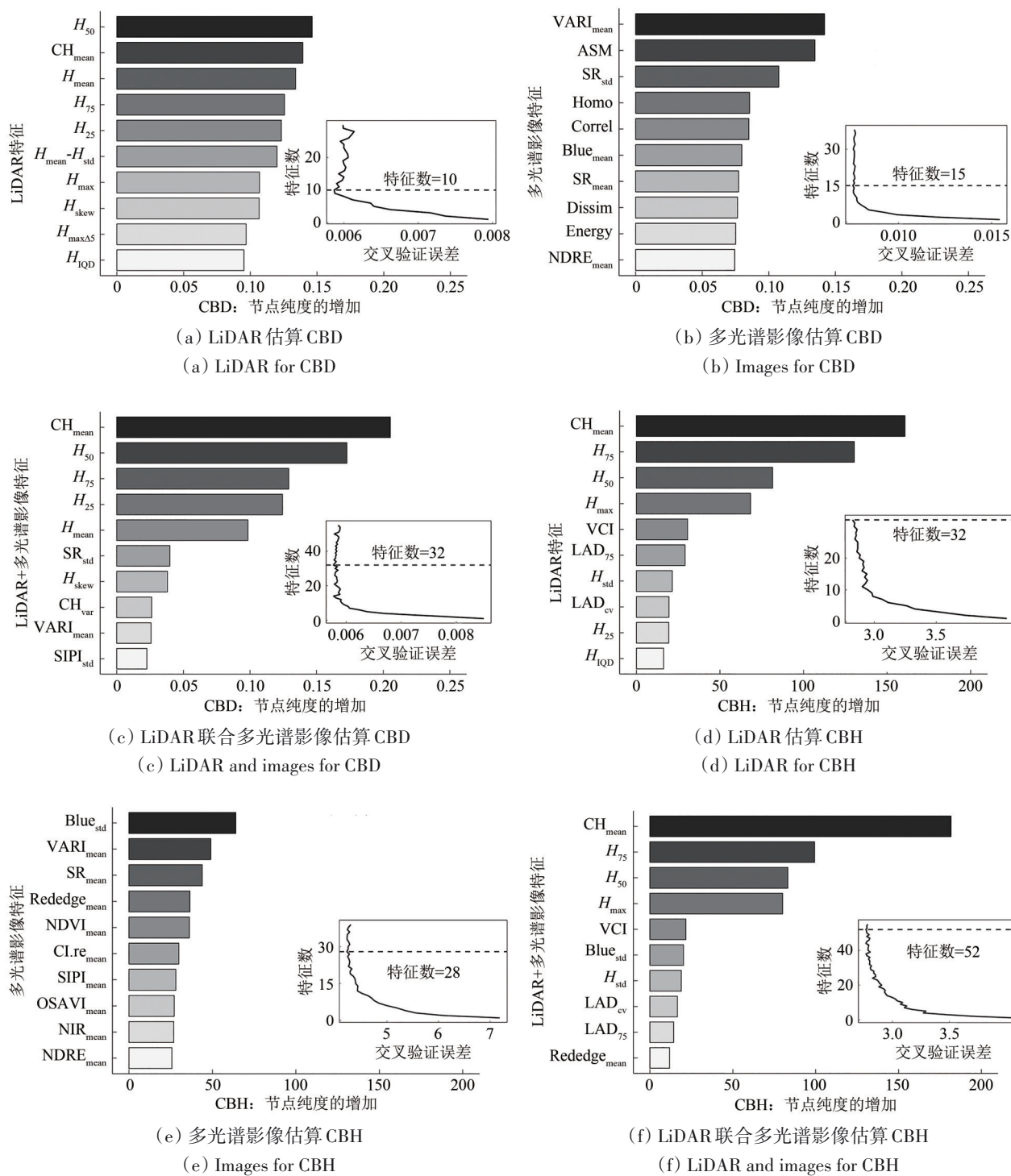


图7 随机森林模型筛选的特征数量以及重要性排名前10位的特征变量

Fig. 7 The number and the top ten importance of selected features using random forest

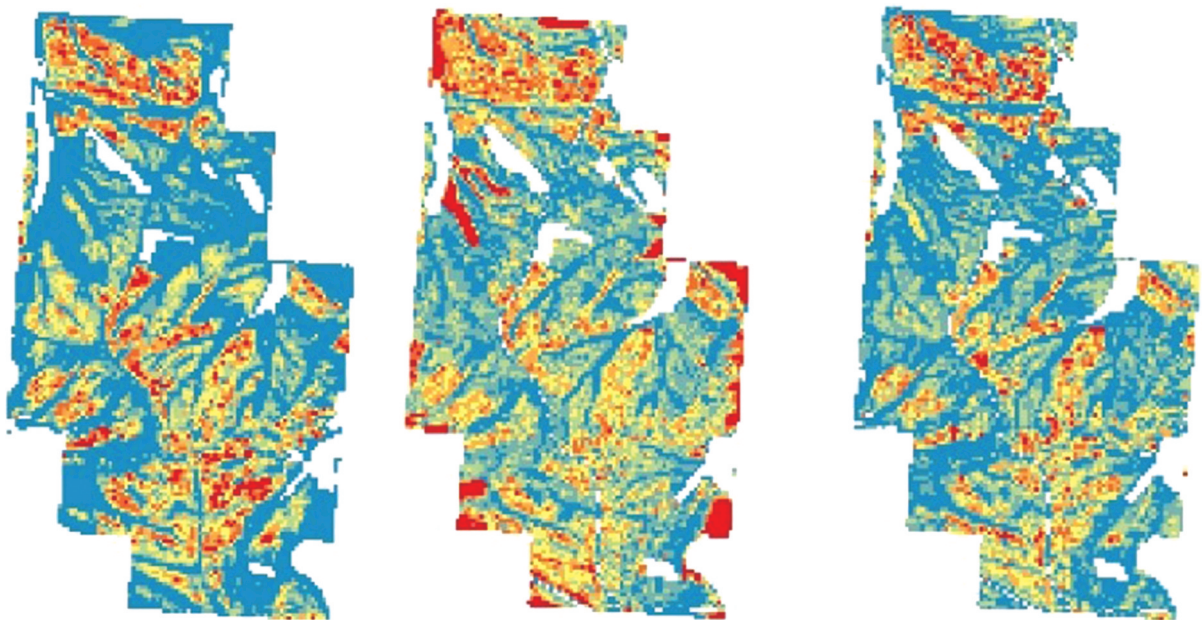
3.4 CBD与CBH制图

利用针对不同数据集构建的最优子集和随机森林模型，结合前者筛选出的3—6个特征变量和后者筛选的10—52个特征变量估算研究区的CBD

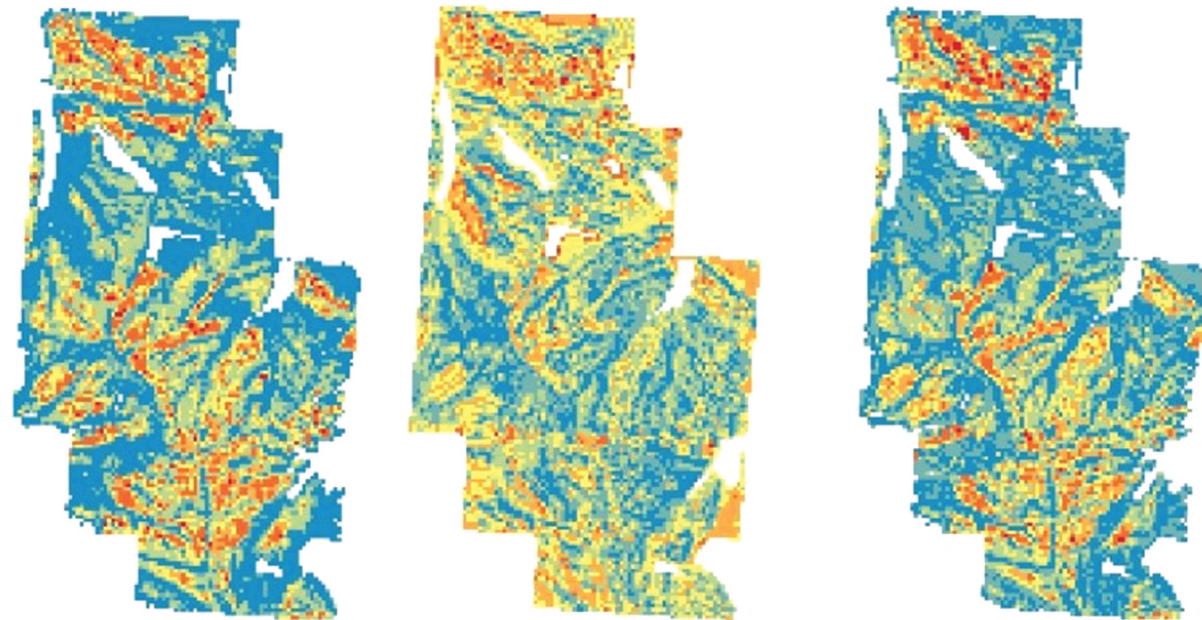
和CBH，结果如图8所示。其中缺失数据区域为10 m格网内的标准LiDAR点云高度低于2 m，被认为是非森林区（Jarron等，2020）。不同数据和模型估算CBD的最大值范围在0.38—0.73 kg/m³，平

均值范围在 0.13—0.18 kg/m³；CBH 的最大值在 9.12—14.86 m，平均值在 2.87—4.44 m。从空间格局来看，不同的数据源预测结果分布总体较为相似。但是多光谱数据在研究区边缘区域通常出现

高值现象，利用最优子集回归估算 CBD 时尤为明显。随机森林模型估算 CBD 时，高值区域少于最优子集回归模型；而在估算 CBH 时，估算数值总体上高于最优子集模型。



(a) LiDAR—最优子集估算 CBD (b) 多光谱影像—最优子集估算 CBD (c) LiDAR 联合多光谱—最优子集估算 CBD
(a) LiDAR – Best Subset for CBD (b) Images – Best Subset for CBD (c) LiDAR and images – Best Subset for CBD



(d) LiDAR—随机森林估算 CBD (e) 多光谱影像—随机森林估算 CBD (f) LiDAR 联合多光谱影像—随机森林估算 CBD
(d) LiDAR – Random Forest for CBD (e) Images – Random Forest for CBD (f) LiDAR and images – Random Forest for CBD

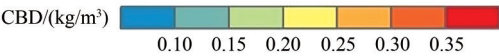




图8 利用最优子集回归和随机森林模型估算 CBD 与 CBH 的空间分布
Fig. 8 Spatial distribution of estimated CBD and CBH using the best subset regression and random forest , respectively

4 讨 论

目前，国内对森林可燃物的研究集中在可燃物分类和可燃物含水率估算（全兴文 等，2019；李晓彤 等，2022），对林冠结构中的林分平均高

度、冠层高度和枝下高关注较多（汪垚 等，2018；李梅 等，2022；Wang 等，2022），利用遥感数据估算林分冠层可燃物参数仍然较少。本研究利用无人机遥感数据构建了基于面状区域的林分 CBD 和 CBH 估算方法。

与已有的大部分研究不同 (Erdody 和 Moskal, 2010; Pierce 等, 2012; González-Ferreiro 等, 2014; Maguya 等, 2015; Arellano-Pérez 等, 2018; Engelstad 等, 2019), 本文选择了能够充分体现泛化误差的十折交叉验证方法对模型进行评价。虽然十折交叉验证比直接利用训练集或留一法计算的 R^2 要低, 但其充分考虑了模型的过拟合程度。尽管如此, 本文中估算 CBD ($R^2>0.4045$, $RMSE<0.0908 \text{ kg/m}^3$, $rRMSE<47.88\%$) 的精度依然优于前人使用中分辨率的卫星光学影像估算结果, 如 Landsat 5 (CBD, $R^2=0.33$) (Palaiologou 等, 2013) 和 Sentinel-2 (CBD, $R^2=0.31$) (Arellano-Pérez 等, 2018), 影像空间分辨率在 10—30 m。这些研究估算冠层可燃物参数的空间分辨率与原始影像相同, 每个像元反映冠层、地表和阴影等混合体的综合光谱特征, 无法反映内部的异质性, 并且中分辨率光谱影像无法捕捉冠层可燃物复杂的三维结构特征 (Reeves 等, 2009)。相比之下, 本研究使用的无人机多光谱影像空间分辨率达到 0.04 m, 估算的面状区域格网空间分辨率为 10 m, 每个格网中的特征统计量能够相对好的反映内部异质性。尽管 Erdody 和 Moskal (2010) 使用高分辨率的航空影像 (0.61 m) 估算林分冠层可燃物参数, CBH 的精度 ($R^2=0.24$) 明显低于本文的估算结果 ($R^2>0.4229$, $RMSE<2.0784 \text{ m}$, $rRMSE=39.87\%$)。该研究使用的影像中仅包含近红、红和绿 3 个光谱波段, 本研究使用了 5 个光谱波段信息。因此, 精细空间分辨率的多光谱影像和丰富的光谱波段信息能够提高林分冠层可燃物参数的估算精度。

一些研究在点云单木分割的基础上估算 CBD 和 CBH。Shin 等 (2018) 使用了无人机高分辨率影像, 通过 SfM 方法构建点云, 而后进行单木分割, 该研究结果显示 CBH 与点云特征仅存在较弱的相关性 ($R^2=0.4$, $RMSE=2.52 \text{ m}$), 与 CBD 则没有相关性。Chamberlain 等 (2021) 利用机载 LiDAR 数据进行单木分割, 然后估算冠层可燃物参数, 其 CBH 的估算精度也较低 ($R^2=0.4$, $RMSE=1.25 \text{ m}$)。上述研究属于一种间接估算冠层可燃物参数的思想, 需要将单株树木从点云中分割出来, 然后在独立的树木基础上汇总至特定的格网内。CBD 和 CBH 的估算精度受到单木分割的精度影响, 尤其是树木相对茂密的区域, 低矮树木难以被检测到 (Chamberlain 等, 2021)。在单木尺度估算 CBD

和 CBH 时, 还需要引入树种, 并根据点云估算树高和胸径, 多个中间参数估算时误差累积也是导致精度不高的重要原因。Luo 等 (2018) 提出一种直接在单木尺度上估算 CBH 方法, 且具有较高的精度 ($R^2=0.88$, $RMSE=1.62 \text{ m}$)。但是, 单木尺度结果无法直接作为参数输入到火行为模型, 仍然需要转换到一定尺寸的格网内。由于每一棵单木都需要汇总至特定格网, 位于格网边缘的树木冠层可能被分配至不同的格网 (González-Ferreiro 等, 2014)。

本研究表明, 仅使用多光谱数据估算 CBD 和 CBH 的精度低于 LiDAR。多光谱影像由于缺少冠层表面以下的信息, 不能对三维可燃物参数有很好的估计 (Erdody 和 Moskal, 2010; Palaiologou 等, 2013)。Erdody 和 Moskal (2010) 进一步研究发现 LiDAR 与航空影像联合使用可以提升冠层可燃物估算精度。Bright 等 (2017) 在 LiDAR 的特征变量基础上, 增加 Landsat 时间序列变量可以降低冠层可燃物估算模型的相对误差。然而本文结果显示, 无人机 LiDAR 和多光谱影像融合不一定提升 CBD 和 CBH 的估算精度。这可能与 LiDAR 点云密度有关, 前人研究中主要使用低密度的机载 LiDAR 数据, 点密度一般在 $0.4\text{--}20 \text{ pts/m}^2$, 本研究使用的无人机点云密度达到 100 pts/m^2 。高密度 LiDAR 点云可较好的描述的森林结构 (Tinkham 等, 2011; González-Ferreiro 等, 2014), 低密度点云则需要光谱信息作为补充。

CBH 本质上是一个相对简单且与高度相关的参数, CBD 是一个更加复杂的参数, 与冠层可燃物载量关系密切。因此, 本文中使用相同数据源和相同模型估算 CBD 和 CBH 时, 前者的精度总是低于后者。随机森林模型的估算结果并不总是高于最优子集模型, 未能充分体现其识别特征变量与估算参数之间隐含非线性关系的能力, 可能是由于缺乏足够的训练样本数量所导致。González-Ferreiro 等 (2014) 认为简单的线性模型足以估算大多数冠层可燃物参数, 不需要复杂的机器学习模型。这一点在无人机多光谱影像上体现的尤其明显。

特征重要性评价结果显示, 当使用 LiDAR 数据估算林分 CBD 和 CBH 时, 高度变量在最优子集模型和随机森林模型中都是最重要的特征因子。高度变量作为 LiDAR 数据最基本的特征因子, 是

不同研究中估算森林冠层可燃物参数时使用频率最高的特征变量。LiDAR的结构变量也在各模型中体现了一定的重要性。由于本文中使用LiDAR数据密度较高,因此比已有研究中LiDAR特征变量更加多样化。但是,LiDAR的强度特征重要性较低,可能是由于无人机飞行高度较低,以及地物表面特征相似,导致点云强度的空间异质性不强。在涉及光谱影像的模型中,特征的选择和重要性表现出了复杂性和多样性的特点,同一模型中可能包括了光谱、植被指数和纹理特征,因此在使用无人机多光谱影像估算林分冠层可燃物参数时需要根据具体模型进行特征选择。不同数据源估算CBD和CBH时,在研究区边缘其数值存在明显差异。当格网内存在高度大于2 m的LiDAR点时,则该格网被认为有冠层存在(Jarron等,2020)。这种方法不能很好的区分接近成熟的玉米和低矮的森林冠层。多光谱数据利用最优子集模型估计CBD时,同质性(Homogeneity)具有较高的重要性。在研究区边缘域存在玉米地,而玉米地的纹理同质性较高,导致在研究区边缘部分区域存在异常高值。

5 结 论

本研究以吉林省蛟河市青背林场为研究区,以无人机LiDAR和多光谱影像为数据源,结合野外样地调查数据,基于面状区域方法估算了林分CBD和CBH。在构建最优子集模型时,采用了0.5次幂的变换对样地水平的CBD和CBH数据进行处理,有效的改善了数据分布特征。本研究表明,无人机遥感数据可以用于估算CBD和CBH。LiDAR数据估算CBD和CBH的精度明显高于多光谱数据。仅使用多光谱影像时,应当采用形式较为简单的最优子集回归方法估算CBD和CBH。将LiDAR和多光谱影像联合使用时不一定能够提升模型精度。

CBD和CBH的估算主要针对地形复杂的山区,地形起伏对无人机多光谱影像反射特征带来了不确定性。在后续研究中,将尝试如基于朗伯反射地形校正方法,以及考虑散射辐射地形校正方法对无人机多光谱影像进行校正。并考虑直接将地形变量作为额外变量输入模型,以期更加准确的估算CBD和CBH。此外,利用SfM方法能够以较低的成本获取点云数据,未来研究将考虑SfM与面

状区域估算方法相结合,分析其估算CBD和CBH的适用性。

参考文献(References)

- Adams M A and Shen Z H. 2015. Introduction to the characteristics, impacts and management of forest fire in China. *Forest Ecology and Management*, 356: 1 [DOI: 10.1016/j.foreco.2015.09.019]
- Andersen H E, McGaughey R J and Reutebuch S E. 2005. Estimating forest canopy fuel parameters using LIDAR data. *Remote Sensing of Environment*, 94(4): 441-449 [DOI: 10.1016/j.rse.2004.10.013]
- Arellano-Pérez S, Castedo-Dorado F, López-Sánchez C A, González-Ferreiro E, Yang Z Q, Díaz-Varela R A, Álvarez-González J G, Vega J A and Ruiz-González A D. 2018. Potential of Sentinel-2A data to model surface and canopy fuel characteristics in relation to crown fire hazard. *Remote Sensing*, 10(10): 1645 [DOI: 10.3390/rs10101645]
- Bright B C, Hudak A T, Meddens A J H, Hawbaker T J, Briggs J S and Kennedy R E. 2017. Prediction of forest canopy and surface fuels from Lidar and satellite time series data in a Bark Beetle-Affected forest. *Forests*, 8(9): 322 [DOI: 10.3390/f8090322]
- Burgan R E and Rothermel R C. 1984. BEHAVE: Fire BEHAVIOR Prediction and Fuel Modeling System-FUEL Subsystem. Rocky Mountain Research Station [DOI: 10.2737/INT-GTR-167]
- Chamberlain C P, Meador A J S and Thode A E. 2021. Airborne Lidar provides reliable estimates of canopy base height and canopy bulk density in Southwestern Ponderosa Pine Forests. *Forest Ecology and Management*, 481: 118695 [DOI: 10.1016/j.foreco.2020.118695]
- Chang Y, Zhu Z L, Bu R C, Li Y H and Hu Y M. 2015. Environmental Controls on the characteristics of mean number of forest fires and mean forest area burned (1987—2007) in China. *Forest Ecology and Management*, 356: 13-21 [DOI: 10.1016/j.foreco.2015.07.012]
- Engelstad P S, Falkowski M, Wolter P, Poznanovic A and Johnson P. 2019. Estimating canopy fuel attributes from low-density LiDAR. *Fire*, 2(3): 38 [DOI: 10.3390/fire2030038]
- Erdody T L and Moskal L M. 2010. Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sensing of Environment*, 114(4): 725-737 [DOI: 10.1016/j.rse.2009.11.002]
- Finney M A. 1998. FARSITE: Fire Area Simulator-Model Development and Evaluation. Rocky Mountain Research Station [DOI: 10.2737/RMRS-RP-4]
- Finney M A. 2006. An overview of flammability fire modeling capabilities// Fuels Management - How to Measure Success: Conference Proceedings. USDA Forest Service, Rocky Mountain Research Station
- González-Ferreiro E, Diéguez-Aranda U, Crecente-Campo F, Barreiro-Fernández L, Miranda D and Castedo-Dorado F. 2014. Modelling canopy fuel variables for *Pinus Radiata* D. Don in NW Spain with low-density LiDAR data. *International Journal of Wildland Fire*, 23(3): 350-362 [DOI: 10.1071/WF13054]
- Jarron L R, Coops N C, Mackenzie W H, Tompalski P and Dykstra P.

2020. Detection of sub-canopy forest structure using airborne LiDAR. *Remote Sensing of Environment*, 244: 111770 [DOI: 10.1016/j.rse.2020.111770]
- Li D R and Li M. 2014. Research advance and application prospect of unmanned aerial vehicle remote sensing system. *Geomatics and Information Science of Wuhan University*, 39(5): 505-513, 540 (李德仁, 李明. 2014. 无人机遥感系统的研究进展与应用前景. 武汉大学学报(信息科学版), 39(5): 505-513, 540) [DOI: 10.13203/j.whugis.20140045]
- Li M, Liu Q W, Feng Y M and Li Z Y. 2022. Analysis of estimation models of plantation stand heights using UAV LiDAR. *National Remote Sensing Bulletin*, 26(12): 2665-2678 (李梅, 刘清旺, 冯益明, 李增元. 2022. 无人机激光雷达人工林分高估测模型分析. 遥感学报, 26(12): 2665-2678) [DOI: 10.11834/jrs.20210246]
- Li X T, Liu Q, Qin X L, Liu S C and Wang C Y. 2022. Method for national fuel types classification based on multi-source data. *National Remote Sensing Bulletin*, 26(3): 480-492 (李晓彤, 刘倩, 覃先林, 刘树超, 王崇阳. 2022. 基于多源数据的全国可燃物类型划分方法. 遥感学报, 26(3): 480-492) [DOI: 10.11834/jrs.20219208]
- Li Z Y, Liu Q W and Pang Y. 2016. Review on forest parameters inversion using LiDAR. *Journal of Remote Sensing (in Chinese)*, 20(5): 1138-1150 (李增元, 刘清旺, 庞勇. 2016. 激光雷达森林参数反演研究进展. 遥感学报, 20(5): 1138-1150) [DOI: 10.11834/jrs.20165130]
- Lim K, Treitz P, Wulder M, St-Onge B and Flood M. 2003. LiDAR remote sensing of forest structure. *Progress in Physical Geography: Earth and Environment*, 27(1): 88-106 [DOI: 10.1191/0309133303pp360ra]
- Luo L P, Zhai Q P, Su Y J, Ma Q, Kelly M and Guo Q H. 2018. Simple method for direct crown base height estimation of individual conifer trees using airborne LiDAR data. *Optics Express*, 26(10): A562-A578 [DOI: 10.1364/OE.26.00A562]
- Lv B R, Peng L, Chen J H, Chen R N and Ge X T. 2021. Analysis of public opinion information pulsation of forest fire on social media. *Geomatics World*, 28(3): 61-66 (吕蓓茹, 彭玲, 陈嘉辉, 陈若男, 葛星彤. 2021. 基于社交媒体的森林火灾舆情信息脉动分析. 地理信息世界, 28(3): 61-66) [DOI: 10.3969/j.issn.1672-1586.2021.03.011]
- Maguya A S, Tegel K, Junttila V, Kauranne T, Korhonen M, Burns J, Leppanen V and Sanz B. 2015. Moving voxel method for estimating canopy base height from airborne laser scanner data. *Remote Sensing*, 7(7): 8950-8972 [DOI: 10.3390/rs70708950]
- Palaiologou P, Kalabokidis K and Kyriakidis P. 2013. Forest mapping by geoinformatics for landscape fire behaviour modelling in coastal forests, Greece. *International Journal of Remote Sensing*, 34(12): 4466-4490 [DOI: 10.1080/01431161.2013.779399]
- Pierce A D, Farris C A and Taylor A H. 2012. Use of random forests for modeling and mapping forest canopy fuels for fire behavior analysis in Lassen Volcanic National Park, California, USA. *Forest Ecology and Management*, 279: 77-89 [DOI: 10.1016/j.foreco.2012.05.010]
- Qian L, Zhang Q X and Zhang Y M. 2020. Correction effect of ensemble Kalman filter algorithm on FARSITE prediction. *Fire Safety Science*, 29(1): 32-41 (钱兰, 张启兴, 张永明. 2020. 集合卡曼滤波算法对FARSITE模型林火蔓延预测的修正效果研究. 火灾科学, 29(1): 32-41) [DOI: 10.3969/j.issn.1004-5309.2020.01.04]
- Quan X W, He B B, Liu X Z, Liao Z M, Qiu S and Yin C M. 2019. Retrieval of fuel moisture content by using radiative transfer models from optical remote sensing data. *Journal of Remote Sensing (in Chinese)*, 23(1): 62-77 (全兴文, 何彬彬, 刘向苗, 廖展芒, 邱实, 殷长明. 2019. 多模型耦合下的植被冠层可燃物含水率遥感反演. 遥感学报, 23(1): 62-77) [DOI: 10.11834/jrs.20197422]
- Reeves M C, Ryan K C, Rollins M G and Thompson T G. 2009. Spatial fuel data products of the LANDFIRE project. *International Journal of Wildland Fire*, 18(3): 250-267 [DOI: 10.1071/WF08086]
- Reinhardt E and Crookston N L. 2003. The Fire and Fuels Extension to the Forest Vegetation Simulator. Rocky Mountain Research Station [DOI: 10.2737/RMRS-GTR-116]
- Riaño D, Chuvieco E, Condés S, González-Matesanz J and Ustin S L. 2004. Generation of crown bulk density for *Pinus Sylvestris* L. from lidar. *Remote Sensing of Environment*, 92(3): 345-352 [DOI: 10.1016/j.rse.2003.12.014]
- Rollins M G. 2009. LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. *International Journal of Wildland Fire*, 18(3): 235-249 [DOI: 10.1071/WF08088]
- Scott J H and Reinhardt E D. 2001. Assessing Crown Fire Potential By Linking Models of Surface and Crown Fire Behavior. Rocky Mountain Research Station [DOI: 10.2737/RMRS-RP-29]
- Shin P, Sankey T, Moore M M and Thode A E. 2018. Evaluating unmanned aerial vehicle images for estimating forest canopy fuels in a Ponderosa Pine Stand. *Remote Sensing*, 10(8): 1266 [DOI: 10.3390/rs10081266]
- Tinkham W T, Huang H Y, Smith A M S, Shrestha R, Falkowski M J, Hudak A T, Link T E, Glenn N F and Marks D G. 2011. A comparison of two open source LiDAR surface classification algorithms. *Remote Sensing*, 3(3): 638-649 [DOI: 10.3390/rs3030638]
- Wang Y, Ni W J, Zhang Z Y, Liu J L, Yu H Y and Zhang D F. 2018. Retrieval of forest canopy heights by using large-footprint waveform data assisted by the LiDAR model over hillsides. *Journal of Remote Sensing (in Chinese)*, 22(3): 466-477 (汪垚, 倪文俭, 张志玉, 刘见礼, 于浩洋, 张大凤. 2018. 激光雷达回波模型辅助的坡地森林冠层高度反演. 遥感学报, 22(3): 466-477) [DOI: 10.11834/jrs.20187152]
- Wang Y T, Xi X H, Wang C, Yang X B, Wang P, Nie S and Du M. 2022. A novel method based on kernel density for estimating crown base height using UAV-Borne LiDAR data. *IEEE Geoscience and Remote Sensing Letters*, 19: 7004105 [DOI: 10.1109/LGRS.2022.3171316]
- Wu Z W, He H S, Liang Y, Luo X and Cai L Y. 2012. Spatial distribution characteristics of potential fire behavior in Fenglin Nature Reserve based on FARSITE Model. *Acta Ecologica Sinica*, 32(19): 6176-6186 (吴志伟, 贺红士, 梁宇, 罗旭, 蔡龙炎. 2012. 基于FARSITE模型的丰林自然保护区潜在林火行为空间分布特征.

生态学报, 32(19): 6176-6186 [DOI: 10.5846/stxb201109111333]
Xu B B, Wang W Y, Chen L F, Tao J H, Ji X Y, Zhang C J and Fan M.
2022. Forest fire spread simulation based on VIIRS active fire data and FARSITE model. National Remote Sensing Bulletin, 26(8):

1575-1588 (徐奔奔, 王伟烨, 陈良富, 陶金花, 纪轩禹, 张成杰, 范萌. 2022. 基于 VIIRS 火点数据和 FARSITE 系统的森林火灾蔓延模拟. 遥感学报, 26(8): 1575-1588) [DOI: 10.11834/jrs.20219427]

Estimating canopy bulk density and canopy base height using UAV LiDAR and multispectral images

SUN Hao¹, GUO Xiaoyi^{1,2,3}, ZHANG Hongyan^{1,2,3}, ZHAO Jianjun^{1,2,3}

1. Schol of Geographical Sciences, Northeast Normal University, Changchun 130024, China;

2. Key Laboratory of Geographical Processes and Ecological Security in Changbai Mountains, Ministry of Education, Changchun 130024, China;

3. Application Innovation Center of Remote sensing information technology in Jilin Province, Changchun 130024, China

Abstract: Wildfire behavior modeling programs require spatial layers of Canopy Bulk Density (CBD) and Canopy Base Height (CBH) to predict fire spread. However, the two canopy fuel metrics have been investigated by only a few studies in China. Inaccurate spatial estimates may result from the utilization of traditional field-based estimates, which assume averages across spatial extents. Recently, unmanned aerial vehicles (UAVs) have emerged as valuable tools that provide LiDAR point clouds and multispectral images for estimating CBD and CBH at fine resolution. The main objective of this study is to develop an area-based approach to estimate CBD and CBH and evaluate the accuracy of various UAV datasets at 10 m resolution at the local scale in China. A case study area is set up in Jiaohe City, Jilin Province, which is predominantly covered by coniferous forests in low mountains and hills. Field data, species, crown base height, total tree height, and diameter at breast height are obtained from 106 circular plots and served as modeling and validation datasets. The Fire and Fuels Extension of Forest Vegetation Simulator is used to calculate CBD and CBH for each plot. Best subset regression and random forest models are employed to establish relationships between the 106 field data points collected and the predictive variables derived from UAV LiDAR and multispectral imagery. Given the nonlinearity of the data, the Box – Cox procedure is utilized and shows that 0.5 power transformation is appropriate for best subset regression. The R^2 value of CBD is always lower than that of CBH when the same models and input dataset are used. The fusion of LiDAR with multispectral imagery produces the best accurate estimation of CBD when random forest is employed ($R^2 = 0.5142$, root mean squared error [RMSE] = 0.0773 kg/m³, relative RMSE [rRMSE] = 40.73%). LiDAR achieves the most accurate estimation for CBH ($R^2 = 0.6477$, RMSE = 1.6245 m, rRMSE = 31.17%). For the best subset regression and random forest models, the use of LiDAR point clouds alone has higher accuracy in estimating CBD and CBH compared with the use of multispectral imagery. The best subset regression models have R^2 values that are greater than those of the random forest models for multispectral imagery alone. This finding indicates that the CBD and CBH values estimated using multispectral imagery are higher than those estimated using LiDAR at a margin of the study area because of crop land. For the various models, fusing LiDAR with multispectral imagery does not necessarily improve estimation accuracy compared with using LiDAR and multispectral imagery alone. Therefore, we recommend using the random forest model that fuses LiDAR and multispectral imagery and LiDAR alone to map CBD and CBH in the study area, respectively, because they have the lowest RMSE. The best subset regression model involves 3 to 6 variables, and the random forest models have 10 to 52 predictive variables. Among the original LiDAR predictor variables, height features are the most important, and structure features have considerable importance. The selected multispectral imagery features of both models exhibit diversity in various canopy flue metrics. This study provides clear evidence that UAV LiDAR and multispectral imagery can be used to derive fine-resolution CBD and CBH, which are crucial for fire behavior modeling at the landscape scale and for forest management activities and decision-making.

Key words: remote sensing, LiDAR, multispectral images, canopy bulk density, canopy base height, the best subset regression, random forest

Supported by National Natural Science Foundation of China (No. 42071359); Science and Technology from Education Department of Jilin Province (No. JJKH20211291KJ)