

# 基于机载激光雷达的落叶松组分生物量反演

洪奕丰<sup>1,2</sup>, 张守攻<sup>1</sup>, 陈伟<sup>2</sup>, 陈东升<sup>1</sup>, 项伟波<sup>1</sup>, 庞勇<sup>3\*</sup>

(1. 中国林业科学研究院林业研究所 国家林业和草原局林木培育重点实验室, 北京 100091;

2. 国家林业和草原局华东调查规划设计院, 浙江 杭州 310019; 3. 中国林业科学研究院资源信息研究所, 北京 100091)

**摘要:** [目的] 构建基于机载 LiDAR 的落叶松组分生物量反演模型, 讨论不同方法对模型构建的影响。 [方法] 以地面实测样地数据和同步获取的机载 LiDAR 点云数据为数据源, 分别采用多元线性回归 (MLR) 和随机森林 (RF) 方法, 估测了长白落叶松的组分生物量, 利用“刀切法”评价了模型的泛化能力。 [结果] 表明: (1) MLR 筛选得到的  $H_{interval}$ 、 $H_{80}$ 、 $D_{10}$ 、 $D_{20}$  与各组分生物量普遍表现为显著 ( $P < 0.05$ ) 或极显著水平 ( $P < 0.01$ )。 (2) MLR 模型的  $R^2$  高于 0.82 (枝、叶除外); RF 模型的  $R^2$  均高于 0.91, 且均拥有较小  $rRMSE$ 、 $TRE$  值。 (3) MDI 和 MDA 方法的变量相对重要值排序均能较好地体现 LiDAR 变量与生物量之间的关系, MDI 在趋势性判断和阈值设定方面更具优势。 [结论] LiDAR 变量与组分生物量具有显著的相关性。 RF 拥有更好的拟合效果和泛化能力, MLR 则对 LiDAR 和组分生物量的关系有更明确的解释能力。 反演模型能较好地反映林分的现势特征, 生物量被低估的现象会随着林龄的增加而逐步增多。

**关键词:** 生物量反演, 机载激光雷达, 组分生物量, 落叶松

**中图分类号:** S771.8

**文献标识码:** A

**文章编号:** 1001-1498(2019)05-0083-08

传统的生物量估测主要运用模型估计法, 过去的几十年中建立的生物量模型超过 2 000 个, 涉及 100 多个树种<sup>[1-3]</sup>。虽精度高、适用性强, 却耗时费力, 破坏性较大, 且难以进行大范围、长时间跨度的生物量动态监测<sup>[4]</sup>。遥感技术的飞速发展, 使得快速获取区域尺度森林生物量的空间分布成为了可能。许多学者运用光学和雷达数据构建生物量反演模型, 取得了一定的效果, 但易饱和、不稳定等缺陷使其在区域范围应用时难以保证足够的精度<sup>[5-9]</sup>。

激光雷达 (LiDAR) 是一项主动遥感技术, 因其能精准地获取森林结构参数, 大幅提高了生物量的估测精度<sup>[10-11]</sup>。利用 LiDAR 估测生物量一般可以分为直接法和间接法: 直接法是将 LiDAR 提取的高度、密度、强度等变量作为输入, 直接建立与生物量的回归关系; 间接法是先利用 LiDAR 变量估计胸径、郁闭度、叶面积指数、蓄积量等森林参数, 再通过林木生长方程或生物量转换因子 (BEF) 等计算生物量<sup>[12-13]</sup>。已有的成果大都针对森林生物量或地上

部分生物量, 而对组分生物量的研究较少。庞勇等<sup>[14]</sup>利用 LiDAR 点云数据, 采用逐步回归法估测了小兴安岭森林组分生物量, 取得了较高的精度。然而, 遥感参数与生物量存在复杂的非线性关系, 传统的回归方法并不能进行有效和完整的表达, 且推导的关系一般仅适用于研究区, 模型的泛化能力不强<sup>[15]</sup>。本研究以黑龙江省长白落叶松 (*Larix olgensis* Henry.) 为研究对象, 以机载 LiDAR 提取的高度、密度变量组为基础, 分别采用多元线性回归 (MLR) 和随机森林 (RF) 方法, 构建组分生物量反演模型, 讨论与评价不同方法对组分生物量估测精度的影响。

## 1 研究区概况与数据获取

### 1.1 研究区概况

研究区位于黑龙江省孟家岗林场 (46°20'~46°30' N, 130°32'~130°52' E), 地处完达山西麓余脉, 以低山丘陵为主, 地势东北高, 西南低, 海拔 170~

575 m,平均海拔为250 m,坡度位于 $10^{\circ}\sim 20^{\circ}$ 之间,较为平缓。土壤种类以典型暗棕壤为主,其次为白浆化暗棕壤,另有少量的潜育暗棕壤、原始暗棕壤、草甸暗棕壤等。

根据2016年森林资源二类调查,林场经营面积 $15\,503\text{ hm}^2$ ,林地面积 $13\,671\text{ hm}^2$ ,森林覆盖率达86.3%,活立木总蓄积量 $1\,464\,508\text{ m}^3$ 。主要树种以长白落叶松、樟子松(*Pinus sylvestris* L. var. *mongolica* Litv.)和红松(*Pinus koraiensis* Sieb. et Zucc.)等人工林为主,面积约占 $2/3$ ,其他天然次生林约占 $1/3$ 。

### 1.2 遥感数据获取

2016年6—7月,采用LiChy(LiDAR, CCD and Hyperspectral)系统在研究区进行LiDAR点云和高清影像的采集,具体参数见文献[16]。使用“运-5”飞机作为飞行平台,平均飞行高度 $2\,500\text{ m}$ ,飞行相对地面速度约 $200\text{ km}\cdot\text{h}^{-1}$ 。总覆盖面积约 $300\text{ km}^2$ ,共获取83条航带数据,航带平均扫描宽度约 $1\,000\text{ m}$ ,航带重叠率60%。

为了精确提取森林垂直结构参数,须对LiDAR点云数据进行预处理。航带拼接时对重叠区域点云予以剔除,以保证点云在研究区密度水平的一致性。采用改进的渐进加密三角网(IPTD)滤波算法<sup>[17]</sup>,对点云进行分类,即区分地面点和非地面点;利用高程阈值法(Elevation > 2 m)结合人工编辑,将非地面点区分为植被点和其他类型点云。分类后的地面点通过TIN插值方法生成数字高程模型(DEM),对植被点(DSM)与DEM取差值得到冠层高度模型(CHM),由此对植被点的高度进行了归一化处理,消除了地形变化的影响。

以样地为统计单元提取LiDAR点云数据的高度、密度变量组作为落叶松生物量反演模型的输入

变量<sup>[18]</sup>。高度变量组(46个)包括:高度百分位数( $H_1, H_5, H_{10}, \dots, H_{95}, H_{99}$ )15个、累计高度百分位数( $AIH_1, AIH_5, AIH_{10}, \dots, AIH_{95}, AIH_{99}$ )15个、最大高度( $H_{max}$ )、最低高度( $H_{min}$ )、平均高度( $H_{mean}$ )、高度中位数( $H_{median}$ )、高度百分位数四分位数间距( $H_{interval}$ )以及高度的峰度( $H_{kurtosis}$ )、平均偏差( $H_{mab}$ )、变异系数( $H_{cv}$ )、标准差( $H_{stdev}$ )等。密度变量组(10个)为将点云(> 2 m)从低到高分成十个相同的高度层,每层回波数的比例即为相应的密度变量( $D_5, D_{10}, D_{20}, \dots, D_{80}, D_{90}$ )。

### 1.3 样地数据获取

2016年6—7月,与飞行实验同步进行地面样地调查。根据《黑龙江省森林资源规划设计调查技术操作细则》的龄组划分标准,按幼龄林( $\leq 20\text{ a}$ )、中龄林(21~30 a)、近熟林(31~40 a)、成熟林( $\geq 41\text{ a}$ ,成熟林与过熟林合并处理)选择典型长白落叶松人工林,设置 $0.06\text{ hm}^2$ ( $20\text{ m}\times 30\text{ m}$ )样地40块。对样地进行每木检尺,起测胸径为5 cm。使用DGPS对样地4个角进行定位,精度控制在cm级(10 cm以内),地面调查与机载实验的差分基站相同,保证外业调查数据与机载飞行数据的位置有可靠的一致性。

在样地调查基础上,分别4个林龄各选取4块标准地,每块标准地按优势木、劣势木各1株,标准木2株共选取64株解析样木(表1)。地上部分采用Monsic分层切割法测定样木的干、皮、树、叶的鲜质量。地下部分采用全根挖掘法,分别测定小根( $\leq 2\text{ cm}$ )、粗根(2~5 cm)、根茎( $\geq 5\text{ cm}$ )的鲜质量。各器官分别取样后,在 $105\text{ }^{\circ}\text{C}$ 下进行30 min的杀青处理,然后将烘箱的温度调到 $80\text{ }^{\circ}\text{C}$ 烘干至恒质量,测出各组分样品的含水率。根据样品含水率乘以总鲜质量推算出各组分生物量<sup>[19]</sup>。

表1 样木数据概况

Table 1 Summary statistics of sampling trees

林龄 Ages/a	样木数量 Numbers	胸径 D/cm	树高 Height/m	树干 Stem/kg	树皮 Bark/kg	树枝 Branch/kg	树叶 Leaf/kg	树根 Under/kg
幼龄林(4~19) Young stand	16	3.4~9.2	3.7~13.8	1.7~15.9	0.3~2.9	0.7~4.8	0.6~2.0	0.8~5.2
中龄林(21~29) Middle-age stand	16	9.2~13.1	6.5~17.1	10.1~48.6	2.1~7.1	3.7~9.0	1.5~3.1	4.7~11.9
近熟林(31~39) Near-mature stand	16	12.9~16.7	11.3~18.4	26.9~106.4	4.0~12.3	5.4~12.3	1.8~3.5	7.4~20.1
成熟林(41~55) Mature stand	16	17.2~23.0	11.7~20.8	64.0~185.7	7.9~19.2	8.7~19.4	2.5~5.6	15.8~49.9

## 2 研究方法

### 2.1 基础模型构建

经对数转换的异速生长模型是生物量估测的常用方法(式1),本研究以64株解析木数据为基础,利用似然无关法构建研究区落叶松人工林生物量相容性模型(式2)。

$$\ln M = \ln(a) + b \times \ln DBH \quad (1)$$

$$\begin{cases} \ln M_{stem} = \ln(a_1) + b_1 \times \ln D \\ \ln M_{Bark} = \ln(a_2) + b_2 \times \ln D \\ \ln M_{Branch} = \ln(a_3) + b_3 \times \ln D \\ \ln M_{Leaf} = \ln(a_4) + b_4 \times \ln D \\ \ln M_{Under} = \ln(a_5) + b_5 \times \ln D \\ \ln M_{Above} = \ln(a_6) + b_6 \times \ln D \\ \ln M_{Total} = \ln(a_7) + b_7 \times \ln D \end{cases} \quad (2)$$

式中,  $M_{stem}$ ,  $M_{Bark}$ ,  $M_{Branch}$ ,  $M_{Leaf}$ ,  $M_{Under}$ ,  $M_{Above}$ ,  $M_{Total}$  分别代表树干、树皮、树枝、树叶、树根、地上生物量、全株生物量,  $D$  为胸径,  $a_i$ ,  $b_i$  为模型参数。

根据40个样地每木检尺数据,利用建立的单木水平生物量模型计算得到各单木的总生物量和各组分生物量,再分别汇总得到各样地的总生物量和各组分生物量,见表2。

表2 模型参数估计值及评价统计指标

Table 2 Parameter estimates and evaluation statistics of biomass model

组分 Components	参数估计值 Parameter estimates			统计指标 Evaluation statistics				
	$a$	$b$	$R^2_{adj}$	$SEE$	$TRE$	$MPE$	$MSE$	$MPSE$
				/t	/%	/%	/%	/%
树干 Stem	0.055 0	2.512 5	0.968 4	14.59	9.03	7.03	6.31	18.47
树皮 Bark	0.028 5	1.993 0	0.954 5	1.78	7.70	7.21	6.52	17.34
树枝 Branch	0.161 6	1.456 5	0.917 3	1.69	5.91	5.40	5.98	16.81
树叶 Leaf	0.235 6	0.905 5	0.842 9	0.49	5.69	4.73	5.69	16.21
树根 Under	0.043 4	2.102 0	0.957 7	4.20	10.85	8.01	7.50	17.15
地上 Above			0.962 2	18.01	8.41	6.58	6.38	16.74
全株 Total			0.963 6	21.02	8.80	6.44	6.53	16.38

注:  $M_{Above} = M_{Stem} + M_{Bark} + M_{Branch} + M_{Leaf}$ ;  $M_{Total} = M_{Above} + M_{Under}$ . 指标说明见参考文献[20]。

### 2.2 遥感反演模型构建

2.2.1 多元线性回归 本研究采用经对数转换的线性回归来构建生物量反演模型<sup>[21]</sup>。

$$\ln M_{components} = a + b_1 \ln x_1 + b_2 \ln x_2 + \dots + b_n \ln x_n + \varepsilon \quad (3)$$

其中,  $M_{components}$  为全株及各组分生物量,  $x_1, x_2,$

$\dots, x_n$  为 LiDAR 建模变量,  $a$  为常数项,  $b_1, b_2, \dots, b_n$  为回归参数,  $\varepsilon$  为误差项。

采用逐步回归法(Stepwise)筛选变量,自变量使统计量  $F$  值过小并且  $t$  检验达不到显著水平( $p > 0.1$ ),则予以剔除;  $F$  值较大且  $t$  检验达到显著水平( $p < 0.05$ )则可以进入。重复进行直到回归方程中所有的自变量均符合进入模型的要求,方程外的自变量均不符合进入模型的要求为止。采用自变量相关稀疏矩阵的条件数  $k$  检查变量的多重共线性:  $k < 100$ , 多重共线性程度很小;  $100 < k < 1\,000$ , 共线性程度较强;  $k > 1\,000$ , 共线性程度严重。

2.2.2 随机森林回归 随机森林(RF)作为一种改进的决策树方法,自 Leo Breiman 提出以来,被广泛地应用于数据挖掘、生物信息学等研究中<sup>[22]</sup>。

模型变量重要性程度通常使用平均不纯度减少(MDI)或平均精确率减少(MDA)来评估。

平均不纯度减少计算方法为:

$$GI_m = 1 - \sum p_{mk}^2 \quad (5)$$

$$VIM_{jm} = GI_m - GI_{left} - GI_{right} \quad (6)$$

$$VIM_{gini} = \sum_{i=1}^N \left( \sum_{j=1}^M VIM_{jm} \right) \quad (7)$$

其中,  $GI_m$  为基尼指数,  $p_{mk}$  为节点  $m$  中类别  $k$  所占的比例,  $VIM_{jm}$  为特征  $X_j$  在节点  $m$  的重要性,  $GI_{left}$  和  $GI_{right}$  分别表示节点分枝后两个新节点的基尼指数,  $M$  为节点总数,  $N$  为决策树总数。  $VIM_{gini}$  为平均不纯度减少方法计算得到的重要性程度。

平均精确率减少计算方法为:

$$VIM_{err} = \sum (err_{OOB2} - err_{OOB1}) / N \quad (8)$$

其中,  $VIM_{err}$  为平均精确率减少计算方法得到的重要性程度;  $err_{OOB1}$  为每一棵决策树,使用 OOB 数据计算得到预测值与观测值的差值;  $err_{OOB2}$  为随机地对所有 OOB 样本特征加入噪声干扰,再次计算得到预测值与观测值的差值;  $N$  为决策树总数。

使用 R 软件的 Random Forest 包进行生物量模型拟合需要确定  $mtry$  和  $ntree$  两个重要参数,本研究参考 Leo Breiman 的推荐默认值<sup>[22]</sup>,结合调试结果,最终确定  $mtry$  取 18,  $ntree$  取 1 000。

### 2.3 模型评价

采用决定系数( $R^2$ )、均方根误差(RMSE)、相对均方根误差( $rRMSE$ )、总相对误差(TRE)作为模型的评价统计指标,指标含义及公式见文献[20]。本研究拟合模型使用全部数据,采用“刀切法”计算评价统计指标,每次取 39 个样本拟合,1 个样本测试,

如此循环直至遍历所有样本,以测试结果的平均值作为反映模型泛化能力的依据。生物量反演模型构建及统计分析均采用 R 软件完成。

### 3 结果与分析

表3、表4分别表达了 MLR 和 RF 模型拟合的结果, LiDAR 变量与生物量有较强的相关性,经逐步回归得到的  $H_{interval}$ 、 $H_{80}$ 、 $D_{10}$ 、 $D_{20}$  与各组分生物量的

相关性普遍表现为显著 ( $P < 0.05$ ) 或极显著水平 ( $P < 0.01$ )。MLR 方法构建的生物量模型除了树枝、树叶外,其余组分模型的  $R^2$  均高于 0.82, 树干、地上及全株模型的  $R^2$  均高于 0.90; RF 方法构建的生物量模型  $R^2$  均高于 0.91, 同时拥有较小的  $rRMSE$  和  $TRE$  值,表明两种方法构建的模型均有较好的拟合效果,可解释的变异占总变异比例较高。

表3 MLR 参数估计值及评价统计指标

Table 3 Parameter estimates and evaluation statistics of MLR

组分 Components	参数估计值 Parameter estimates				统计指标 Evaluation statistics			
	$a$	$b^1$	$b^2$	$b^3$	$R^2$	$RMSE/t$	$rRMSE$	$TRE/\%$
树干 Stem	-2.762 3**	-0.632 4**	1.702 8**	11.850 6**	0.914 3	0.409 3	0.077 7	0.60
树皮 Bark	-3.121 8**	-0.882 2**	1.187 9**	10.783 8**	0.825 1	0.048 9	0.081 7	0.66
树枝 Branch	-0.872 6*	-0.788 1**	0.442 0**	7.347 6*	0.541 4	0.069 3	0.096 1	0.91
树叶 Leaf	0.114 1	-1.100 8**	-0.169 2	5.864 6	0.636 1	0.152 2	0.678 6	0.96
树根 Under	-2.574 5**	-0.597 2**	1.149 7**	8.376 5**	0.873 3	0.083 7	0.068	0.45
地上 Above	-1.586 2**	-0.651 3**	1.403 5**	10.583 6**	0.907 7	0.458 6	0.067 3	0.45
全株 Total	-1.304 4**	-0.641 9**	1.362 6**	10.217 7**	0.905 0	0.537 0	0.066 8	0.45

注: \* :  $P < 0.05$ ; \*\* :  $P < 0.01$ 。  $\ln M_{Branch} = a_0 + b_1 \ln H_{interval} + b_2 \ln H_{80} + b_3 \ln D_{20}$ 。其余组分模型形式均为:  $\ln M_{components} = a_0 + b_1 \ln H_{interval} + b_2 \ln H_{80} + b_3 \ln D_{10}$ 。

表4 RF 评价统计指标

Table 4 Evaluation statistics of RF

组分 Components	$R^2$	$RMSE/t$	$rRMSE$	$TRE/\%$
树干 Stem	0.973 4	0.536 7	0.107 7	1.066 5
树皮 Bark	0.964 8	0.051 4	0.087 5	0.725 9
树枝 Branch	0.919 4	0.071 3	0.098 9	0.944 9
树叶 Leaf	0.912 8	0.032 5	0.143 0	1.941 5
树根 Under	0.964 6	0.104 0	0.086 7	0.713 3
地上 Above	0.971 1	0.623 7	0.095 5	0.853 2
全株 Total	0.970 6	0.720 6	0.093 2	0.814 7

对比两种方法, RF 拥有更好的拟合效果。  $M_{stem}$ 、 $M_{above}$  和  $M_{total}$  组分生物量模型方面, RF 的  $R^2 \approx 0.97$ , MLR 的  $R^2 \approx 0.91$ ,  $rRMSE$  和  $TRE$  相近。树枝和树叶模型方面, MLR 的  $R^2$  仅有 0.5~0.6, RF 则仍维持在 0.9 以上,表明 RF 有更好的拟合效果。此外,无论是 MLR 还是 RF,生物量  $M_{total} > 8 t$  的区间低估情况明显增多(图1),说明随着林龄的增加,林分郁闭度逐渐降低(自然稀疏或人为间伐), LiDAR

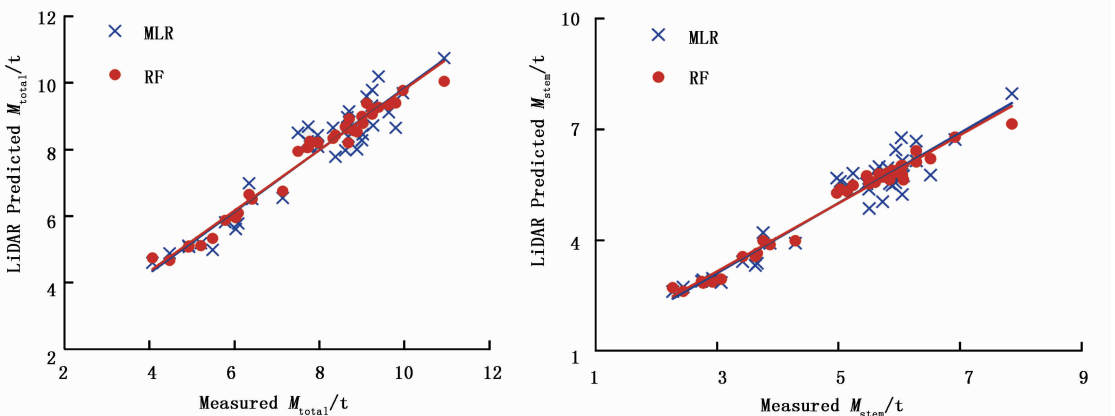


图1 样地生物量实测值与 LiDAR 估测值对比。

Fig. 1 Comparison of biomass from field measured and airborne LiDAR.

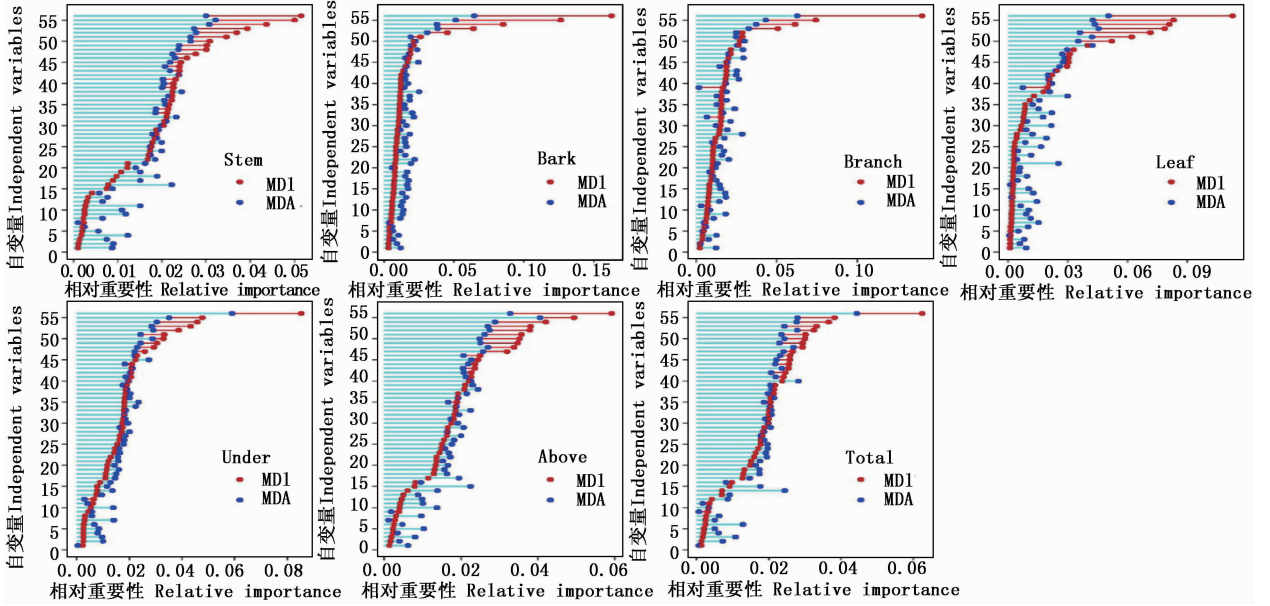
点云密度不高的情况下,冠层回波信号不够,容易树顶及树冠边缘错失现象,导致生物量被低估<sup>[18]</sup>。

图2表达了 RF 组分生物量模型中各 LiDAR 变

量归一化后的相对重要性程度。相同的组分模型中,MDI 和 MDA 各变量的相对重要性排序总体趋势基本一致,尤其是前 15 个变量,重要程度几乎完全

一致,即表明两种方法虽然选择指标和算法有所区别,但均是反映自变量对因变量的重要性影响。MDI 与 MDA 相比,各变量的相对重要性表现出更大的差异性,一般前 5~10 个变量的重要性程度要显

著高于其余变量;MDA 虽然排序趋势与 MDI 一致,但各变量重要性程度差异较小,变化平缓。因此,若需要根据相对重要性程度筛选变量优化 RF 模型,MDI 方法在趋势性判断和阈值设定方面更具优势。



注:图中纵坐标自变量代码与变量名称的对应关系见附表 1

Note: The corresponding relationship between the ordinate independent variable code and the variable name in the figure is shown in attached list 1

图 2 组分 RF 生物量模型各变量相对重要值图

Fig. 2 Relative importance of factors of biomass components

从表 5 可以看出,4 种林分状态下,林分密度、平均高均与生物量有显著的相关性。其中,类型 II 反映了同一水平林分间伐前后的生物量变化情况,表明 LiDAR 变量能较好地反映林分的现势性,除了能表达林分自然生长变化,也能表达人为干扰的情

况。因此,利用 LiDAR 的高度变量和密度变量进行森林生物量的反演是十分有效的<sup>[10]</sup>。树枝、树叶的拟合效果较差,一方面因冠层交错导致生物量低估,另一方面可能与外业样本采集中容易造成树枝、树叶生物量的耗损有关<sup>[23]</sup>。

表 5 不同林分特征的反演效果比较

Table 5 Comparison of inversion effects of different forest characteristics

类型 Types	样地号 Plots	M <sub>Total</sub> 实测值 M <sub>Total</sub> Measured Value/t	MLR 预测值 MLR Predicted Value/t	RF 预测值/t RF Predicted Value	林分密度 Density/(株·hm <sup>-2</sup> )	平均胸径 D/cm	平均树高 Height/m
I	17_02	9.270 2	8.713 9	9.172 4	1 217	18.2	22.2
	17_711	7.974 9	8.069 1	8.191 4	547	23.9	22.8
II	17_20	4.939 3	5.062 0	5.092 9	1 650	12.4	13.6
	17_16	7.144 1	6.534 9	6.740 0	1 950	13.2	13.7
III	17_704	10.943 1	10.738 3	10.040 0	420	31.5	28.9
	17_12	7.504 5	8.494 7	7.936 3	483	23.7	20.7
IV	16_F2_1	8.664 2	8.963 7	8.617 2	1 256	17.2	19.3
	17_22	4.078 8	4.587 5	4.733 4	1 283	13.0	14.0

类型 I: 相同树高不同密度(高树高水平);类型 II: 相同树高不同密度(低树高水平);类型 III: 相同密度不同树高(低密度水平);类型 IV: 相同密度不同树高(高密度水平)。Type I: Same tree height and different density when tree height is high level; Type II: Same tree height and different density when tree height is low level; Type III: Same density and different tree height when density is low level; Type IV: Same density and different tree height when density is high level.

## 4 讨论

LiDAR 变量与组分生物量均有显著的相关性,

构建的模型具有较好的拟合效果和较高的预估精度, $R^2$  普遍高于 0.82,树干、地上和全株等组分生物量模型的  $R^2$  更是优于 0.90。利用“刀切法”对模型

的泛化能力进行评价,  $rRMSE$  值介于 6% ~ 10%,  $TRE$  值介于 0.4% ~ 1.0%, 结果令人满意。这一方面由于 LiDAR 数据能高精度提取林分结构参数, 有利于生物量反演精度提升; 另一方面可能与研究区落叶松全部为人工林有关, 林相整齐、均质度高, LiDAR 数据能较好地代表林分整体状态<sup>[14]</sup>。

比较而言, RF 拥有更好的拟合效果。利用“刀切法”评价 RF 模型的泛化能力, 与穆喜云等<sup>[24]</sup>的研究结论一致。这一方面归功于 RF 自身强大的学习能力; 另一方面, 与森林生物量与 LiDAR 数据之间复杂的非线性关系有关<sup>[25]</sup>。MLR 筛选出的变量  $H_{interval}$ 、 $H_{80}$ 、 $D_{10}$ 、 $D_{20}$  与各组分生物量的相关性普遍表现为显著 ( $P < 0.05$ ) 或极显著水平 ( $P < 0.01$ ), 说明 LiDAR 变量拥有对组分生物量很好的解释性<sup>[15]</sup>。虽然 RF 的变量相对重要性排序也可以表达出各变量的贡献度, 但由于 LiDAR 变量自身拥有的多重相关性, 导致各组分生物量模型的变量相对重要性排序差异较大, 下一步研究中, 应当尝试深入分析变量相量相对重要性程度, 并通过交互进一步优化 RF 模型。

## 5 结论

本研究结果表明:

(1) 机载 LiDAR 提取的自变量与生物量均具有显著的相关性, 普遍表现为显著 ( $P < 0.05$ ) 或极显著水平 ( $P < 0.01$ )。利用 LiDAR 变量构建的生物量反演模型能够较好地反映林分的现势特征, 除了能表达林分的自然生长过程, 还能很好的体现人为干扰的信息。

(2) MLR 与 RF 方法均适用于生物量反演模型构建, 除了树枝、树叶外, 模型的  $R^2$  均高于 0.8。其中, MLR 方法优势是模型更加简单明确, LiDAR 自变量对生物量有更明确的解释能力。RF 模型拥有更好的拟合效果和预估精度。

(3) 低林分密度水平下, 由于 LiDAR 点云与冠顶的错失机率增大, 会出现对生物量被低估的现象, 且这一情况会随着林龄的增加而逐步增大。

## 参考文献:

[1] Lambert M, Ung C, Raulier F. Canadian national tree aboveground biomass equations[J]. Canadian Journal of Forest Research, 2005, 35(8):1996-2018.

[2] 沈亚洲, 孙晓梅, 张江涛, 等. 甘肃小陇山林区日本落叶松人工林单株生物量的研究[J]. 林业科学研究, 2011, 24(4):517-522.

[3] 曾伟生, 唐守正, 夏忠胜, 等. 利用线性混合模型和哑变量模型方法建立贵州省通用性生物量方程[J]. 林业科学研究, 2011, 24

(3):285-291.

[4] Haara A, Leskinen P. The assessment of the uncertainty of updated stand-level inventory data[J]. Silva Fennica, 2009, 43(1):87-112.

[5] 徐天蜀, 岳彩荣. 基于印度遥感卫星 IRS-P6 的森林生物量估测模型研究[J]. 中南林业调查规划, 2008, 27(1):42-45.

[6] 郑光, 田庆久, 陈镜明, 等. 结合树龄信息的遥感森林生态系统生物量制图[J]. 遥感学报, 2006, 10(6):932-940.

[7] Smith-Jonforsen G, Folkesson K, Hallberg B, et al. Effects of Forest Biomass and Stand Consolidation on P-Band Backscatter[J]. IEEE Geoscience and Remote Sensing Letters, 2007, 4(4):669-673.

[8] Paloscia S, Macelloni G, Pampaloni P, et al. The potential of C- and L-band SAR in estimating vegetation biomass: the ERS-1 and JERS-1 experiments[J]. IEEE Transactions on Geoscience and Remote Sensing, 1999, 37(4):2107-2110.

[9] 曹庆先, 徐大平, 鞠洪波. 基于 TM 影像纹理与光谱特征和 KNN 方法估算 5 种红树林群落生物量[J]. 林业科学研究, 2011, 24(2):144-150.

[10] Næsset E, Gobakken T. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser[J]. Remote Sensing of Environment, 2008, 112(6):3079-3090.

[11] 刘清旺, 李增元, 陈尔学, 等. 机载 LiDAR 点云数据估测单株木生物量[J]. 高技术通讯, 2010, 20(7):765-770.

[12] 曹林, 余光辉, 代劲松, 等. 激光雷达技术估测森林生物量的研究现状及展望[J]. 南京林业大学学报:自然科学版, 2013, 37(3):163-169.

[13] 尹惠妍, 李海奎. 基于蓄积的森林生物量估算方法的对比分析[J]. 林业科学研究, 2014, 27(6):848-853.

[14] 庞勇, 李增元. 基于机载激光雷达的小兴安岭温带森林组分生物量反演[J]. 植物生态学报, 2012, 36(10):1095-1105.

[15] 刘茜, 杨乐, 柳钦火, 等. 森林地上生物量遥感反演方法综述[J]. 遥感学报, 2015, 19(1):62-74.

[16] Yong P, Li Z, Ju H, et al. LiCHy: The CAF's LiDAR, CCD and hyperspectral integrated airborne observation system[J]. Remote Sensing, 2016, 8(5):398-414.

[17] Zhao X, Guo Q, Su Y, et al. Improved progressive TIN densification filtering algorithm for airborne LiDAR data in forested areas[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2016, 117:79-91.

[18] 李增元, 刘清旺, 庞勇. 激光雷达森林参数反演研究进展[J]. 遥感学报, 2016, 20(5):1138-1150.

[19] 莫德祥. 桂东南不同密度柳杉人工林生态系统生物量及碳氮格局研究[D]. 南宁:广西大学, 2013.

[20] 曾伟生, 唐守正. 立木生物量方程的优度评价和精度分析[J]. 林业科学, 2011, 47(11):106-113.

[21] 穆喜云. 森林地上生物量遥感估测方法研究[D]. 呼和浩特:内蒙古农业大学, 2015.

[22] Breiman L. Random forests[J]. Machine Learning, 2001, 45:5-32.

[23] 黄兴召, 孙晓梅, 张守攻, 等. 辽东山区日本落叶松生物量相容性模型的研究[J]. 林业科学研究, 2014, 27(2):142-148.

[24] 穆喜云, 刘清旺, 庞勇, 等. 基于机载激光雷达的森林地上碳储量估测[J]. 东北林业大学学报, 2016, 44(11):52-56.

[25] Chen Q. Modeling aboveground tree woody biomass using national-scale allometric methods and airborne lidar[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2015, 106(3):95-106.

附表 1:相对重要值排序代码与变量对应表  
attached list 1:Relative importance value sorting code and variable correspondence

顺序 Order	$M_{stem}$	$M_{bark}$	$M_{branch}$	$M_{leaf}$	$M_{under}$	$M_{above}$	$M_{total}$
1	$AIH_5$	$D_{50}$	$AIH_{interval}$	$AIH_{interval}$	$D_{50}$	$H_{max}$	$D_{50}$
2	$D_{50}$	$D_{40}$	$H_{mab\_mean}$	$H_{mab\_mean}$	$D_{40}$	$D_{50}$	$H_{99}$
3	$H_{95}$	$H_{ratio}$	$H_1$	$H_{max}$	$AIH_{95}$	$AIH_5$	$AIH_{95}$
4	$H_{max}$	$H_{skewness}$	$H_{interval}$	$AIH_{95}$	$H_{99}$	$AIH_{95}$	$H_{ratio}$
5	$H_{99}$	$D_{70}$	$D_{70}$	$H_{interval}$	$H_{ratio}$	$H_{99}$	$H_{skewness}$
6	$AIH_{99}$	$AIH_{95}$	$H_{max}$	$D_5$	$H_{max}$	$AIH_{99}$	$AIH_5$
7	$H_{media}$	$H_{99}$	$AIH_{30}$	$H_{95}$	$H_{skewness}$	$H_{20}$	$H_{max}$
8	$AIH_{95}$	$H_{20}$	$AIH_{95}$	$D_{20}$	$AIH_5$	$H_{95}$	$AIH_{99}$
9	$AIH_{10}$	$D_{30}$	$D_{50}$	$AIH_{90}$	$AIH_{99}$	$D_{40}$	$D_{40}$
10	$H_{20}$	$H_{stddev}$	$D_5$	$H_{90}$	$AIH_{30}$	$AIH_{10}$	$H_{70}$
11	$AIH_{40}$	$H_{kurtosis}$	$AIH_{40}$	$AIH_{99}$	$H_{25}$	$H_{mean}$	$AIH_{40}$
12	$AIH_{50}$	$AIH_{30}$	$H_{95}$	$H_{99}$	$D_{80}$	$H_{90}$	$H_{mean}$
13	$AIH_{75}$	$H_{95}$	$H_5$	$H_1$	$AIH_{10}$	$H_{sqr}$	$AIH_{30}$
14	$AIH_{90}$	$H_{variation}$	$D_{40}$	$AIH_{80}$	$AIH_{75}$	$H_{25}$	$AIH_{75}$
15	$H_{90}$	$AIH_5$	$D_{20}$	$H_{variation}$	$H_{curt}$	$AIH_{40}$	$H_{20}$
16	$H_{mean}$	$D_{80}$	$AIH_{25}$	$H_{stddev}$	$H_{95}$	$AIH_{90}$	$AIH_{70}$
17	$AIH_{60}$	$AIH_{60}$	$H_{90}$	$H_{mab}$	$H_{media}$	$H_{50}$	$D_{80}$
18	$H_{40}$	$AIH_{99}$	$AIH_1$	$H_{cv}$	$H_{40}$	$H_{skewness}$	$H_{75}$
19	$H_{skewness}$	$AIH_{25}$	$AIH_{90}$	$AIH_{75}$	$AIH_{70}$	$H_{ratio}$	$AIH_{90}$
20	$H_{60}$	$H_{70}$	$D_{60}$	$H_{min}$	$H_{75}$	$H_{media}$	$H_{sqr}$
21	$H_{ratio}$	$D_{60}$	$H_{variation}$	$D_{90}$	$AIH_{90}$	$H_{80}$	$H_{60}$
22	$H_{25}$	$AIH_{10}$	$H_{stddev}$	$H_{80}$	$H_{variation}$	$AIH_{70}$	$AIH_{25}$
23	$AIH_{70}$	$H_{media}$	$H_{40}$	$D_{10}$	$D_{70}$	$H_{curt}$	$H_{variation}$
24	$H_{70}$	$H_{50}$	$H_{99}$	$AIH_1$	$H_{90}$	$D_{80}$	$H_{90}$
25	$H_{80}$	$H_{40}$	$H_{cv}$	$D_{60}$	$D_{30}$	$AIH_{50}$	$AIH_{60}$
26	$H_{50}$	$AIH_{20}$	$H_{min}$	$H_5$	$AIH_{20}$	$H_{30}$	$AIH_{20}$
27	$AIH_{20}$	$H_{80}$	$H_{mab}$	$D_{70}$	$H_{60}$	$H_{stddev}$	$AIH_{10}$
28	$D_{40}$	$AIH_{80}$	$H_{50}$	$D_{30}$	$AIH_{25}$	$AIH_{30}$	$H_{media}$
29	$AIH_{25}$	$AIH_{40}$	$D_{90}$	$AIH_{70}$	$H_{80}$	$H_{75}$	$H_{80}$
30	$H_{curt}$	$AIH_{70}$	$H_{media}$	$D_{80}$	$H_{stddev}$	$H_{variation}$	$H_{30}$
31	$AIH_{80}$	$H_{25}$	$AIH_{80}$	$H_{sqr}$	$AIH_{60}$	$AIH_{60}$	$H_{curt}$
32	$H_{30}$	$H_{75}$	$H_{ratio}$	$D_{40}$	$H_{20}$	$AIH_{75}$	$H_{95}$
33	$H_{75}$	$AIH_1$	$H_{skewness}$	$H_{skewness}$	$H_{sqr}$	$AIH_{20}$	$AIH_{80}$
34	$AIH_{30}$	$AIH_{90}$	$AIH_{99}$	$AIH_{25}$	$H_{50}$	$AIH_{80}$	$AIH_{50}$
35	$H_{sqr}$	$D_{20}$	$AIH_{20}$	$AIH_5$	$H_{mean}$	$H_{60}$	$D_{70}$
36	$H_{stddev}$	$H_1$	$D_{80}$	$D_{50}$	$H_{30}$	$H_{70}$	$H_{40}$
37	$H_{10}$	$H_{10}$	$AIH_{70}$	$H_{ratio}$	$AIH_{40}$	$H_{10}$	$H_{25}$
38	$D_{70}$	$H_{60}$	$D_{30}$	$H_{kurtosis}$	$AIH_{80}$	$H_{40}$	$H_{50}$
39	$D_{30}$	$H_{30}$	$AIH_{75}$	$H_{50}$	$AIH_{50}$	$AIH_{25}$	$H_{stddev}$
40	$H_{variation}$	$H_{90}$	$H_{70}$	$AIH_{10}$	$H_{kurtosis}$	$D_{70}$	$H_{10}$
41	$D_{80}$	$AIH_{75}$	$H_{kurtosis}$	$H_{10}$	$H_{70}$	$D_{90}$	$D_{90}$
42	$D_{90}$	$AIH_{50}$	$H_{60}$	$H_{40}$	$D_{90}$	$D_{60}$	$D_{30}$
43	$AIH_1$	$H_{max}$	$AIH_{50}$	$AIH_{40}$	$H_{mab}$	$D_{30}$	$D_{60}$
44	$H_5$	$D_{90}$	$AIH_{60}$	$H_{curt}$	$H_{10}$	$H_{mab}$	$H_{mab}$
45	$H_{mab}$	$H_{mean}$	$D_{10}$	$AIH_{30}$	$AIH_{interval}$	$D_{10}$	$AIH_{interval}$
46	$D_{60}$	$H_{sqr}$	$AIH_{10}$	$H_{20}$	$H_{mab\_mean}$	$H_5$	$D_{10}$
47	$AIH_{interval}$	$H_{curt}$	$H_{80}$	$AIH_{50}$	$D_{60}$	$H_{kurtosis}$	$H_{mab\_mean}$
48	$D_{10}$	$D_5$	$H_{30}$	$H_{mean}$	$H_{interval}$	$AIH_1$	$AIH_1$
49	$H_1$	$H_5$	$H_{75}$	$AIH_{20}$	$H_5$	$AIH_{interval}$	$H_5$
50	$D_{20}$	$H_{mab}$	$H_{10}$	$H_{25}$	$H_{cv}$	$H_{interval}$	$H_{interval}$
51	$H_{mab\_mean}$	$H_{cv}$	$H_{curt}$	$H_{media}$	$AIH_1$	$H_{mab\_mean}$	$H_{kurtosis}$
52	$H_{kurtosis}$	$D_{10}$	$H_{25}$	$H_{60}$	$H_{min}$	$H_1$	$D_5$
53	$D_5$	$H_{mab\_mean}$	$H_{20}$	$H_{30}$	$D_5$	$H_{min}$	$H_{min}$
54	$H_{interval}$	$AIH_{interval}$	$AIH_5$	$H_{75}$	$D_{20}$	$D_5$	$H_1$
55	$H_{cv}$	$H_{min}$	$H_{mean}$	$AIH_{60}$	$H_1$	$H_{cv}$	$H_{cv}$
56	$H_{min}$	$H_{interval}$	$H_{sqr}$	$H_{70}$	$D_{10}$	$D_{20}$	$D_{20}$

# Inversion of Biomass Components for *Larix olgensis* Plantation Using Airborne LiDAR

HONG Yi-feng<sup>1,2</sup>, ZHANG Shou-gong<sup>1</sup>, CHEN Wei<sup>2</sup>, CHEN Dong-sheng<sup>1</sup>, XIANG Wei-bo<sup>1</sup>, PANG Yong<sup>3</sup>

(1. Research Institute of Forestry, Chinese Academy of Forestry, Key Laboratory of Tree Breeding and Cultivation, National Forestry and Grassland Administration, Beijing 100091, China; 2. East China Inventory and Planning Institute, National Forestry and Grassland Administration, Hangzhou 310019, Zhejiang, China; 3. Research Institute of Forest Resources Information Technique, Chinese Academy of Forestry, Beijing 100091, China)

**Abstract:** [ **Objective** ] To establish the inversion models of biomass components for *Larix olgensis* plantation using Airborne LiDAR. [ **Method** ] Compatible models were established by combining dummy variable and nonlinear seemingly unrelated regression based on the biomass data of 64 trees in 40 sample plots of *L. olgensis* plantation. The canopy height indices and density indices were calculated from LiDAR point cloud data. Then the models based on these biomass components from field data and LiDAR indices were built using multivariate linear regression (MLR) and random forest regression (RF). The biomass model validation was accomplished by Jackknifing technique. [ **Result** ] (1) There was a significant ( $P < 0.05$ ) or extremely significant ( $P < 0.01$ ) correlation between biomass components and  $H_{interval}$ ,  $H_{80}$ ,  $D_{10}$ , and  $D_{20}$  which were screened using MLR. (2) The models had good estimation precision with  $R^2 > 0.82$  and  $R^2 > 0.91$  for MLR and RF, respectively. (3) The relationship between LiDAR variables and biomass could be reflected well by relative importance ranking of variables by both MDI and MDA method. [ **Conclusion** ] There is a good correlations between biomass components and LiDAR indices. RF shows stronger ability about fitting goodness and generalization, while MLR can make clearer interpretation of the relationship between LiDAR indices and biomass components. The present situation of *L. olgensis* plantation can be accurately reflected using inversion models, and the underestimation of component biomass will increase with the increase of stand age.

**Keywords:** biomass inversion, airborne LiDAR, biomass component, *Larix olgensis*

(责任编辑:彭南轩)