

高山松地上生物量遥感估算的不确定性分析

黄屹杰¹, 张加龙¹, 胡耀鹏¹, 程 滔²

(1. 西南林业大学 林学院, 云南 昆明 650233; 2. 国家基础地理信息中心 调查监测部, 北京 100830)

摘要: 【目的】采用遥感数据估算森林地上生物量仍存在一些不确定性问题, 研究估算过程中的误差来源及其占比, 对提高森林地上生物量的估测精度具有重要意义。【方法】从遥感影像提取因子, 结合高山松 *Pinus densata* 外业调查数据, 建立多元线性回归、梯度提升回归树、随机森林等 3 种地上生物量估测模型, 对样地尺度与 3 种模型的不确定性进行分析和度量。【结果】①高山松单株生物量模型不确定性为 16.43%, 样地尺度的不确定性为 7.07%; ②多元线性回归模型残差不确定性为 34.86%, 参数不确定性为 21.30%, 与样地不确定性合成后总不确定性为 41.45%; ③非参数模型中, 梯度提升回归树估测高山松地上生物量的总不确定性为 23.12%, 随机森林为 19.42%。【结论】3 种遥感估算模型中, 非参数模型的不确定性明显低于参数模型。相较于样地尺度, 遥感估算模型的不确定性对地上生物量估算精度的影响较大。图 3 表 3 参 26

关键词: 高山松; 地上生物量; 遥感估测模型; 不确定性

中图分类号: S758.5 文献标志码: A 文章编号: 2095-0756(2022)03-0531-09

Uncertainty analysis of estimating aboveground biomass of *Pinus densata* by remote sensing

HUANG Yijie¹, ZHANG Jialong¹, HU Yaopeng¹, CHENG Tao²

(1. College of Forestry, Southwest Forestry University, Kunming 650233, Yunnan, China; 2. Department of Investigation and Monitoring, National Geomatics Center of China, Beijing 100830, China)

Abstract: [Objective] To improve the estimation accuracy of forest aboveground biomass, this study is aimed to conduct an uncertainty analysis, trying to figure out the percentage error of estimating forest aboveground biomass by remote sensing and the causes behind. [Method] With factors extracted from remote sensing images and combined with the data of *Pinus densata* from field surveys, three types of aboveground biomass estimation model were established, namely Multiple Linear Regression (MLR), Gradient Boost Regression Tree (GBRT), and Random Forest (RF), before the uncertainty of sample plot scale and three models was measured and analyzed. [Result] (1) The uncertainty of tree biomass model for *P. densata* is 16.43%, and the uncertainty of the scale up to the sample plot is 7.07%; (2) The residual uncertainty of the MLR model is 34.86%, the parameter uncertainty is 21.30% whereas the total uncertainty combined with the sample plot uncertainty is 41.45%. (3) In the non-parametric model of the GBRT modeling estimates, the total uncertainty of the aboveground biomass is 23.12%, and the RF is 19.42%. [Conclusion] Among the three remote sensing models, the uncertainty of the non-parametric model is obviously lower than that of the parametric model. Compared with the uncertainty of the sample plot scale, the remote sensing estimation model has a great effect on the

收稿日期: 2021-07-08; 修回日期: 2021-12-24

基金项目: 国家自然科学基金资助项目 (31860207); 2020 年云南省高层次人才培养支持计划“青年拔尖人才”专项 (81210468); 西南林业大学科研启动基金 (111932)

作者简介: 黄屹杰 (ORCID: 0000-0001-6922-1029), 从事森林生物量遥感估测研究。E-mail: huangyijie@swfu.edu.cn。通信作者: 张加龙 (ORCID: 0000-0002-6969-3656), 教授, 博士, 从事资源环境遥感研究。E-mail: jialongzhang@swfu.edu.cn

accuracy of the aboveground biomass estimation. [Ch, 3 fig. 3 tab. 26 ref.]

Key words: *Pinus densata*; aboveground biomass; remote sensing model; uncertainty

森林生物量是森林生态系统的最基本数量特征,是研究林业和生态问题的基础^[1]。地上生物量(aboveground biomass, AGB)是森林生态系统生产力的重要指标和质量的综合体现,对研究全球气候变化具有重要意义。目前,森林地上生物量的调查方法有地面调查和遥感监测2种。传统的地面调查法耗时、成本高,而遥感监测方法快速、无损,具有在更大尺度上应用的优势,已成为地上生物量估测的主流^[2-3]。在森林调查中,常使用统计模型来预测样地内个体树木的生物量,然后对同一树种的单株地上生物量进行汇总,来作为遥感估算生物量的训练数据和精度评估数据^[4]。用于地上生物量估测的遥感数据源包括光学遥感、机载雷达、激光测量数据等^[5]。与成本较高的机载雷达数据相比,光学影像穿透森林冠层获取树干信息的能力较差,但随着新算法的不断改进,使用光学遥感影像进行地上生物量估算的精度也在不断提高^[6]。

目前,遥感估算地上生物量在因子选取、建模方法、数据饱和等方面存在较多不确定性问题^[7-8]。区域尺度森林地上生物量遥感估算包含测量和抽样、单株生物量模型、遥感因子和地形因子等不确定性来源^[9-10]。其中,与测量和抽样误差相比,模型的不确定性对地上生物量估算的影响较大,如SHETTLES等^[11]利用激光雷达和地面数据估测地上生物量发现:模型不确定性占总不确定性的55%。同时,用于建模的生物量数据是使用该区域树种的异速生长方程计算得出的。因此,单株生物量模型的误差会传播到样地级的生物量估测,从而影响遥感估算生物量的不确定性^[12-13]。现阶段不确定性的量化方法主要分为3种^[14]:过程模型分析法、随机误差传递法和Monte Carlo模拟法。

本研究基于云南省香格里拉市 Landsat 8 影像和外业调查数据,建立了生物量遥感估测模型,分析了遥感估算样地高山松 *Pinus densata* 地上生物量的总不确定性,以期为提高森林地上生物量估算精度提供参考。

1 研究地区与数据

1.1 研究区概况

研究区位于云南省西北部的迪庆藏族自治州香格里拉市(26°52'~28°52'N, 99°20'~100°19'E),该区总面积为11 613 km²,平均海拔为3 459 m。香格里拉市的森林覆盖率较高,达75%,主要植被类型为寒温性针叶林,优势树种有云杉 *Picea asperata*、冷杉 *Abies fabri*、高山松、云南松 *Pinus yunnanensis*、高山栎 *Quercus semicarpifolia* 等。

1.2 数据来源

于2015年11月和2016年3月,在每个乡镇高山松纯林区域随机布设60块样地,大小为30 m×30 m,且每块样地间隔3 km以上。样地高山松林分的龄级主要为近熟林和成熟林。外业调查记录胸径大于5 cm样木的树高、胸径,同时对每块样地的林分密度进行了计算(表1)。

Landsat 8 影像数据来源于美国地质调查局(<http://glovis.usgs.gov/>),3景数据为2015年11—12月的成像,云量均小于2%。

表1 外业调查样地统计

Table 1 Basic statistics of the sample plots of field surveys			
指标	平均胸径/cm	平均树高/m	林分密度/(株·hm ⁻²)
最大值	22.76	14.18	3 100.00
最小值	6.85	4.48	222.22
平均值	14.80	8.70	1 181.30
标准差	3.68	2.11	714.14

2 研究方法

2.1 样地地上生物量计算

采用每木检尺,通过胸径、树高和单株生物量模型^[15]计算每株地上生物量,进而得出每块样地的地上生物量。单株生物量模型(W)如下:

$$W = 0.073 \times D^{1.739} \times H^{0.880} \quad (1)$$

式 (1) 中: D 为胸径, H 为树高, 模型 R^2 为 0.992, 均方根误差为 30.778 kg。

2.2 Landsat 数据处理

对影像数据进行辐射定标, 采用 FLAASH 方法进行大气校正。以研究区 SPOT-5 影像为参考数据, 选取 100 个地面控制点, 采用二项式方法对影像进行几何校正, 使用双线性内插法将影像重新采样为 $30\text{ m} \times 30\text{ m}$, 误差控制在 1 个像元内^[8]; 接着, 采用坡度匹配模型^[16]进行地形校正。具体操作方法参考文献^[17]。

对预处理后的影像数据提取 4 类遥感因子: ①原始波段因子, 分别为 C、B1、B2、B3、B4、B5、B7; ②植被指数因子, 分别为 B43、B42、B53、B54、B57、B73、B74、B3Albedo、B473、NDVI、ND32、ND53、ND54、ND57、ND452、DVI; ③信息增强因子, 分别为 VIS123、Albedo、MID57; ④纹理信息因子^[18], 分别为均值 (ME)、方差 (VA)、同质性 (HO)、反差 (CO)、相异 (DI)、熵 (EN)、角二阶矩 (SM)、相关性 (CR)、偏度 (SK)。用 R5 和 R9 分别代表 5×5 和 9×9 窗口, 如 R9B5CR 代表 9 窗口第 5 波段的相关性纹理。

岳彩荣^[19]提取了 5×5 、 7×7 、 9×9 、 15×15 的 4 种窗口纹理信息因子对香格里拉森林生物量进行了遥感估测; 张加龙等^[20]基于遥感影像和连续清查固定样地对高山松地上生物量进行估算时, 使用了 5×5 、 9×9 的 2 种窗口纹理信息因子。以上研究表明: 高山松地上生物量与 5×5 和 9×9 窗口的纹理信息因子相关性更强。因此, 本研究的纹理信息因子选用 5×5 和 9×9 窗口。4 类遥感因子的提取和计算参考文献^[8, 17, 20]。

2.3 因子筛选方法

基于 SPSS 软件对样地地上生物量与备选遥感因子进行相关性分析, 结果按照 Pearson 系数进行排序。本研究因子入选的显著性水平为 $P \leq 0.05$, 因子剔除的水平为 $P \geq 0.10$, 得到相关性排前 14 个的遥感因子。为避免解释变量之间的多重共线性问题, 将各个遥感因子之间相关性较高的因子剔除, 最终筛选 9 个因子参与模型建立, 分别为 B74、9 窗口第 5 波段相关性 (R9B5CR)、9 窗口第 6 波段方差 (R9B6VA)、9 窗口第 3 波段熵 (R9B3EN)、5 窗口第 5 波段角二阶矩 (R5B5SM)、9 窗口第 6 波段相关性 (R9B6CR)、5 窗口第 3 波段熵 (R5B3EN)、5 窗口第 4 波段熵 (R5B4EN)、5 窗口第 4 波段偏度 (R5B4SK)。

2.4 建模方法

遥感估算高山松地上生物量的 3 种方法分别为多元线性回归 (MLR)、梯度提升回归树 (GBRT) 以及随机森林 (RF)。从 60 个样地调查数据中随机选取 48 个 (80%) 进行建模, 剩余 12 个 (20%) 用来独立性检验。建模和检验样本的地上生物量统计结果见表 2。

2.4.1 多元线性回归

多元线性回归模型可以同时

对多个解释变量与 1 个因变量进行拟合, 用回归方程来表示拟合关系。本研究多元线性回归建模在 SPSS 软件中实现。

2.4.2 梯度提升回归树 该方法通过构建多个弱分类器, 经过多次迭代最终组合成 1 个强分类器^[21]。本研究基于 Python 语言的 Sklearn 工具包提供的梯度提升回归树算法, 对高山松地上生物量与遥感因子进行建模分析。在建模分析前需要对弱分类器的最大迭代次数 (n_estimators)、学习速率 (learning_rate)、最大深度 (max_depth) 等参数进行确定。参数确定往往需要遵循一定的经验, 比如最大迭代次数通常在预测值收敛的情况下越小越好, 学习速率通常小于 0.10, 最大深度通常不应大于 15。依据这些经验设置上述参数的范围, 比较参数值组合下模型学习的效果, 以得到模型输出预测结果^[8]。本研究根据拟合优度的最佳参数组合, 选取最大迭代次数为 60, 子采样为 0.5, 学习速率为 0.05, 决策树最大深度为 7, 叶子节点最少样本数为 3。

2.4.3 随机森林 随机森林适用于多数分类与回归的问题, 由一系列决策树组成。基于 Python 语言的 Sklearn 工具包进行高山松地上生物量建模分析。程序使用 bootstrap 重采样方法从样本集中提取多个重

表 2 建模和检验样本的地上生物量实测值

Table 2 Measured values of aboveground biomass of modeling and testing samples

样本集	样地数/个	地上生物量/(t·hm ⁻²)			
		最大值	最小值	平均值	标准差
建模样本	48	169.75	12.95	56.50	33.85
检验样本	12	94.97	15.44	62.56	26.48

采样样本进行建模,之后对多个模型值进行预测并组合,再通过投票得出最终值,一般会在随机产生的分类树中选出重叠次数最多的决策树作为最终模型^[6]。本研究初步设置决策树数量为 50,逐步增加决策树数量模拟建模过程,回归误差趋于稳定,最终确定决策树数量为 300。

2.5 精度评价方法

采用的评价指标包括决定系数 (R^2)、均方根误差 (R_{MSE})、预测精度 (F)。各指标计算如下:

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

$$R_{MSE} = \sqrt{\frac{\sum_{i=1}^s (y_i - \hat{y}_i)^2}{s}}; \quad (3)$$

$$F = \frac{1}{s} \sum_{i=1}^s \left(1 - \left| \frac{y_i - \hat{y}_i}{\hat{y}_i} \right| \right) \times 100\%。 \quad (4)$$

式 (2)~(4) 中: y_i 表示第 i 个样本地上生物量的实测值; \hat{y}_i 表示第 i 个样本地上生物量的预测值; \bar{y} 表示所有样本地上生物量实测值的均值; s 表示样本总数量。

2.6 不确定性分析与度量

2.6.1 模型残差变异不确定性计算 样地尺度上单株生物量模型的残差不确定性 (σ_w) 计算如下:

$$\sigma_w = \frac{R_{MSEw}}{\bar{M}} \times 100\%。 \quad (5)$$

式 (5) 中: R_{MSEw} 为单株生物量模型的均方根误差; \bar{M} 为高山松地上生物量实测值的平均值^[22]。

采用六步法^[23] 计算地上生物量遥感估算模型残差变异产生的误差: 第 1 步升序排列样地地上生物量实测值 (y); 第 2 步利用生物量模型的预测值 (\hat{y}) 计算残差 (M_ε), 残差为样地地上生物量实测值与预测值的差值; 第 3 步将 60 块样地进行分组, 每 6 块地为 1 组, 共 10 组; 第 4 步计算每组样地地上生物量预测值的平均值 ($\hat{\bar{y}}$) 及残差标准差 (σ_ε)。每组样地地上生物量预测值的平均值、残差、残差标准差计算如下:

$$\hat{\bar{y}} = \frac{1}{n} \sum_{j=1}^n \hat{y}_j; \quad (6)$$

$$M_\varepsilon = y - \hat{y}; \quad (7)$$

$$\sigma_\varepsilon = \sqrt{\frac{\sum_{j=1}^n (M_{\varepsilon j} - \bar{M}_\varepsilon)^2}{n-1}}。 \quad (8)$$

式 (6)~(8) 中: \hat{y}_j 表示第 j 块样地地上生物量预测值; $M_{\varepsilon j}$ 、 \bar{M}_ε 分别表示第 j 块样地的残差和残差平均值; n 表示 6 块样本集样地数量。

第 5 步对 $\hat{\bar{y}}$ 与 σ_ε 进行相关性拟合, 自变量为地上生物量模型预测值, 因变量为残差标准差。模型残差变异与样地分组地上生物量预测值的平均值关系可表达为:

$$\sigma_\varepsilon = f(\hat{\bar{y}})。 \quad (9)$$

第 6 步将各块样地地上生物量预测值代入拟合后得出的公式, 对所有样地的残差标准差求和, 除以地上生物量实测值的和, 从而得到模型残差的不确定性。

2.6.2 模型参数不确定性计算 对于单株生物量模型和多元线性回归模型的参数不确定性, 可通过泰勒级数一阶展开式进行量化计算。生物量模型 $B = f(x, \hat{\alpha})$ 经泰勒级数一阶展开如下:

$$B = f(x, \hat{\alpha}) \approx f(x, \alpha) + \frac{\partial f(x, \alpha)}{\partial \alpha_j} (\hat{\alpha} - \alpha) \quad (10)$$

式 (10) 中: $f(x, \alpha)$ 为地上生物量预测值; x 为解释变量; $\hat{\alpha}$ 为参数模拟值; α 表示参数真实值, $\frac{\partial f(x, \alpha)}{\partial \alpha_j}$ 为模型参数 α_j 的偏导数。 $\frac{\partial f(x, \alpha)}{\partial \alpha_j} (\hat{\alpha} - \alpha)$ 为模型参数引起的误差 (σ_p), σ_p 可近似表示为:

$$\sigma_p^2 \approx \mathbf{Z}_{jk} \text{var}(\alpha) \mathbf{Z}_{jk}^T \quad (11)$$

式 (11) 中: \mathbf{Z}_{jk} 表示 \mathbf{Z} 矩阵为 $j \times k$ 矩阵; \mathbf{Z}_{jk}^T 为 \mathbf{Z}_{jk} 的转置矩阵; $\text{var}(\alpha)$ 为生物量方程中估计参数 α 的协方差矩阵。

2.6.3 不确定性合成方法 假设间接测量值与直接测量值之间存在某种函数关系, 那么间接测量的不确定性可通过特定函数关系, 用直接测量的不确定性表示为合成不确定性^[24]。误差可以表示为:

$$N_R = \sqrt{N_{U_1}^2 + N_{U_2}^2 + N_{U_3}^2 + \cdots + N_{U_i}^2} \quad (12)$$

式 (12) 中: N_R 表示总误差; N_{U_i} 表示第 i 个变量的误差。根据此式, 计算单株生物量模型的不确定性 (σ_{tree}):

$$\sigma_{\text{tree}} = \sqrt{\sigma_w^2 + \sigma_p^2} \quad (13)$$

式 (13) 中: σ_w 为单株生物量模型残差不确定性; σ_p 为参数不确定性。

2.6.4 结合遥感模型估测的地上生物量不确定性 单株生物量的误差传播到样地尺度的不确定性 (σ_{plot}) 由下式计算:

$$\sigma_{\text{plot}} = \sqrt{\sum_{i=1}^N \sigma_{\text{tree}i}^2 / S} \quad (14)$$

式 (14) 中: N 为单位样地内样木株数; $\sigma_{\text{tree}i}$ 为样地内第 i 株样木的单株生物量不确定性; S 为样地面积。

得到样地尺度不确定性和地上生物量遥感估测模型地上生物量的不确定性后, 根据式 (12), 结合遥感信息数据, 估测高山松地上生物量的总不确定性 (σ):

$$\sigma = \sqrt{\sigma_{\text{plot}}^2 + \sigma_R^2} \quad (15)$$

式 (15) 中: σ_{plot} 为样地尺度不确定性; σ_R 为遥感估测模型不确定性。

3 结果与分析

3.1 样地尺度上地上生物量的不确定性

单株生物量模型参数的协方差矩阵为:

$$\begin{pmatrix} 0.002\ 700 & -0.000\ 760 & 0.000\ 089 \\ -0.000\ 450 & 0.006\ 100 & -0.040\ 000 \\ 0.000\ 089 & -0.040\ 000 & 0.006\ 200 \end{pmatrix}.$$

将协方差矩阵带入式 (11), 得出单株生物量模型的参数不确定性为 12.42%。

由式 (5) 计算模型的残差不确定性为 10.76%。对残差不确定性与参数不确定性合成, 可得出单株生物量模型的不确定性为 16.43%。当模型误差叠加到样地尺度时, 通过式 (14) 得出样地尺度不确定性为 7.07%。

3.2 多元线性回归 (MLR) 模型的不确定性

本研究建立了地上生物量与遥感因子的多元线性回归模型, 根据拟合优度选出最佳模型为:

$$y = 2\ 820.989 - 520.296x_{R5B4EN} + 73.833x_{R9B5CR} - 259.225x_{R9B3EN}.$$

其中: y 表示地上生物量预测值, x_{R5B4EN} 为 5 窗口第 4 波段熵, x_{R9B5CR} 为 9 窗口第 5 波段相关性, x_{R9B3EN} 为 9 窗口第 3 波段熵。模型 R^2 为 0.326, R_{MSE} 为 26.12, 预测精度为 64.91%(图 1A)。

对残差标准差 (y) 和样地地上生物量预测分组平均值 (x) 进行相关性拟合发现: 拟合效果最佳的模型

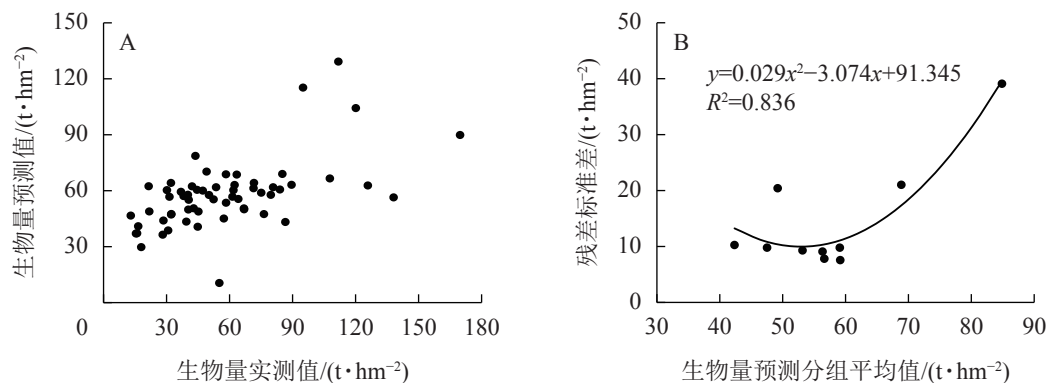


图1 基于多元线性回归的地上生物量预测效果和残差不确定性

Figure 1 Prediction effect and the residual uncertainty of aboveground biomass based on multiple linear regression

为 $y = 91.345 + 0.029x^2 - 3.074x$, $R^2 = 0.836$, 计算后得出多元线性回归模型的残差不确定性为 34.86% (图 1B)。

多元线性回归模型参数的协方差矩阵为:

$$\begin{pmatrix} 0.038\ 00 & -0.004\ 50 & 0.000\ 94 \\ -0.007\ 80 & 0.068\ 00 & -0.800\ 00 \\ 0.000\ 94 & -0.800\ 00 & 0.044\ 00 \end{pmatrix}.$$

将协方差矩阵代入式 (11) 可得出参数变异不确定性为 21.30%。对残差不确定性和参数不确定性合成, 通过多元线性回归模型, 估算出高山松地上生物量的不确定性为 40.84%。

3.3 梯度提升回归树 (GBRT) 模型的不确定性

图 2A 所示: 梯度提升回归树模型的 R^2 为 0.815, R_{MSE} 为 14.24, 预测精度为 74.72%。图 2B 所示: 残差标准差 (y) 与样地地上生物量预测分组平均值 (x) 拟合的最佳模型为 $y = 0.260\ 1x - 2.127\ 3$, $R^2 = 0.670$, 得出残差不确定性为 22.01%。非参数建模方法不考虑参数不确定性, 基于梯度提升回归树模型估算得出高山松地上生物量的不确定性为 22.01%。

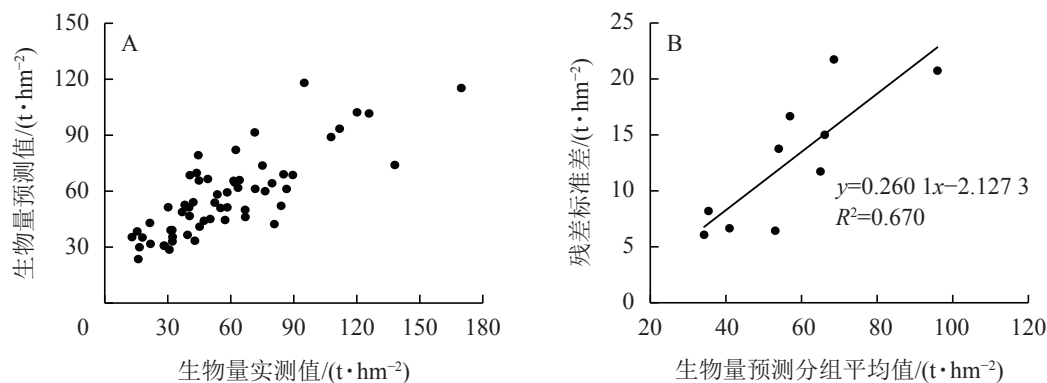


图2 基于梯度提升回归树的地上生物量预测效果和残差不确定性

Figure 2 Prediction effect and the residual uncertainty of aboveground biomass based on gradient boost regression tree

3.4 随机森林 (RF) 模型的不确定性

图 3A 所示: 随机森林模型的 R^2 为 0.889, R_{MSE} 为 11.02, 预测精度为 76.47%。图 3B 所示: 残差标准差 (y) 与样地生物量预测分组平均值 (x) 拟合后的最佳模型为线性关系, 模型为 $y = 0.214\ 3x - 1.475\ 7$, $R^2 = 0.863$ 。与多元线性回归和梯度提升回归树模型的残差不确定性计算方法相同, 在不考虑参数不确定性的情况下, 得出随机森林估算高山松地上生物量的不确定性为 18.09%。

3.5 遥感估测生物量总不确定性

由表 3 可见: 在样地尺度上, 单株生物量模型的不确定性为 16.43%, 总不确定性为 7.07%; 在 3 种遥感估算模型中, 多元线性回归模型的不确定性为 40.84%, 梯度提升回归树模型的不确定性为

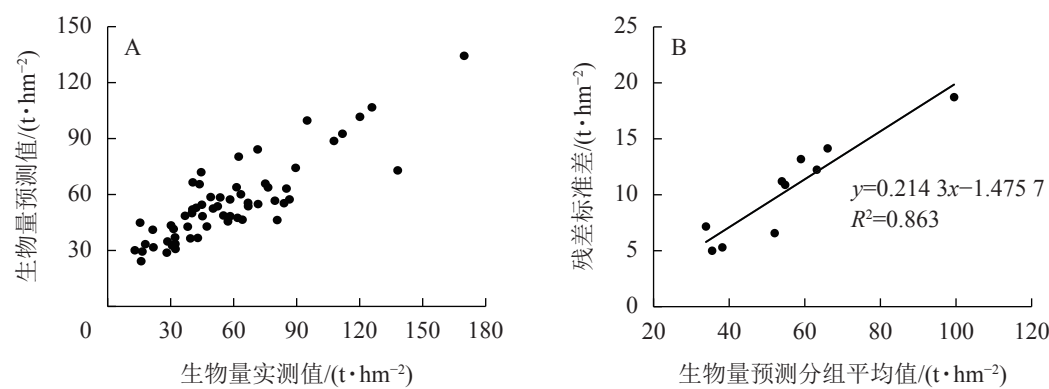


图 3 基于随机森林的地上生物量预测效果和残差不确定性
Figure 3 Prediction effect and the residual uncertainty of aboveground biomass based on random forest

表 3 3 种地上生物量估测模型的不确定性

模型	遥感估算地上生物量的不确定性			样地尺度地上生物量的不确定性				总不确定性/%
	模型残差	模型参数	遥感估测模型	模型残差	模型参数	单株生物量模型	样地尺度不	
	变异/%	变异/%	不确定性/%	变异/%	变异/%	不确定性/%	确定性/%	
多元线性回归	34.86	21.30	40.84					41.45
梯度提升回归树	22.01		22.01	10.76	12.42	16.43	7.07	23.12
随机森林	18.09		18.09					19.42

说明：梯度提升回归树与随机森林为非参数模型，不考虑模型参数变异的不确定性

22.01%，随机森林模型的不确定性为 18.09%。基于多元线性回归、梯度提升回归树、随机森林模型，估算出高山松地上生物量的总不确定性分别为 41.45%、23.12% 和 19.42%。

4 结论与讨论

本研究结合误差传递的方法，量化了 3 种生物量遥感估测模型与样地尺度单株生物量模型产生的误差，并将各个来源的不确定性合成，得出遥感估算高山松地上生物量的总不确定性。遥感估算地上生物量的不确定性主要来源于遥感估测模型的不确定性，其中多元线性回归为 40.84%，梯度提升回归树为 22.01%，随机森林为 18.09%。样地尺度不确定性相对较小，为 7.07%。从预测效果来看，非参数的随机森林和梯度提升回归树模型的不确定性要优于参数的多元线性回归模型，随机森林模型的不确定性低于梯度提升回归树模型。单株生物量模型参数变异为 12.42%，残差不确定性为 10.76%，模型的不确定性主要来源于参数变异，残差不确定性相对较低。

3 种遥感估算模型的不确定性研究表明：多元线性回归通过回归方程直观反映遥感因子和地上生物量的关系，但不能有效描述它们之间复杂的非线性关系，要提高参数模型的预测精度则需要大量样地观测数据。张加龙等^[8] 基于外业调查和 Landsat 8 OLI 影像，建立了多元线性回归、偏最小二乘、随机森林和梯度提升回归树 4 种模型，估测了香格里拉市高山松地上生物量，结果表明：随机森林和梯度提升回归树 2 种非参数模型的预估精度优于其他 2 种参数模型；陈蜀蓉等^[25] 运用 4 种模型对浙江缙云县公益林生物量进行建模估算的结果同样表明：随机森林和 Erf-BP 神经网络 2 种非参数模型优于参数模型。本研究建立的梯度提升回归树和随机森林 2 种非参数模型有较好的拟合优度和预测精度，不确定性优于参数的多元线性回归模型。遥感估测模型的不确定性对森林地上生物量的估算精度占主要的影响，因此，提升模型精度仍是今后生物量估算研究的重要方向。

已有不少学者进行了单株生物量模型不确定性的研究。秦立厚等^[10] 分别用 30、42、48 株杉木 *Cunninghamia lanceolata* 进行单株生物量建模时发现：一元生物量模型的残差变异引起的不确定性分别为 15.2%、12.3%、11.7%，二元生物量模型的不确定性为 13.3%、9.4%、8.7%，表明二元单株生物量模型的不确定性要低于一元模型，且建模样本的增加可以显著降低模型残差变异不确定性。与其相比，本

研究在对样地尺度生物量不确定性进行度量时,引用了基于幂函数形式的二元单株生物量估算模型^[15]。由于缺少实际伐倒的样木生物量数据,运用模型预测精度和均方根误差间接计算模型不确定性,结果可能存在偏差。CHAVE等^[26]计算的热带雨林地区生物量模型残差变异误差为31.3%,而本研究模型残差变异误差为10.76%。这可能是热带树种之间的特征差异大,导致产生的残差变异较大,而本研究建模对象仅为亚热带高山松单一树种,未对抽样、测量误差以及遥感影像坐标校正误差进行分析。这在后续研究中还需进一步完善。

5 参考文献

- [1] 唐守正, 张会儒, 胥辉. 相容性生物量模型的建立及其估计方法研究[J]. 林业科学, 2000, **36**(专刊 1): 19–27.
TANG Shouzheng, ZHANG Huiru, XU Hui. Study on establish and estimate method of compatible biomass model[J]. *Sci Silv Sin*, 2000, **36**(spec 1): 19–27.
- [2] 王维枫, 雷渊才, 王雪峰, 等. 森林生物量模型综述[J]. 西北林学院学报, 2008, **23**(2): 58–63.
WANG Weifeng, LEI Yuancai, WANG Xuefeng, et al. A review of forest biomass models [J]. *J Northwest For Univ*, 2008, **23**(2): 58–63.
- [3] 徐婷, 曹林, 余光辉. 基于 Landsat 8 OLI 的特征变量优化提取及森林生物量反演[J]. 遥感技术与应用, 2015, **30**(2): 226–234.
XU Ting, CAO Lin, SHE Guanghui. Feature extraction and forest biomass estimation based on Landsat 8 OLI [J]. *Remote Sensing Technol Appl*, 2015, **30**(2): 226–234.
- [4] MCROBERTS R E, WESTFALL J A. Effects of uncertainty in model predictions of individual tree volume on large area volume estimates [J]. *For Sci*, 2014, **60**(1): 34–42.
- [5] LU Dengsheng, CHEN Qi, WANG Guangxing, et al. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems [J]. *Int J Digital Earth*, 2016, **9**(1): 63–105.
- [6] BAO Rui, ZHANG Jialong, LU Chi, et al. Estimating above-ground biomass of *Pinus densata* Mast. using best slope temporal segmentation and Landsat time series[J/OL]. *J Appl Remote Sensing*, 2021, **15**(2)[2021-03-25]. doi: 10.1117/1.JRS.15.024507.
- [7] ZHAO Panpan, LU Dengsheng, WANG Guangxing, et al. Examining spectral reflectance saturation in landsat imagery and corresponding solutions to improve forest aboveground biomass estimation [J]. *Remote Sensing*, 2016, **8**(6): 469–495.
- [8] 张加龙, 胥辉, 陆驰. 应用 Landsat 8 OLI 和 GBRT 对高山松地上生物量的估测[J]. 东北林业大学学报, 2018, **46**(8): 25–30.
ZHANG Jialong, XU Hui, LU Chi. Estimating above ground biomass of *Pinus densata* based on Landsat 8 OLI and gradient boost regression [J]. *J Northeast For Univ*, 2018, **46**(8): 25–30.
- [9] COHEN R, KAINO J, OKELLO J A, et al. Propagating uncertainty to estimates of above-ground biomass for Kenyan mangroves: a scaling procedure from tree to landscape level [J]. *For Ecol Manage*, 2013, **310**: 968–982.
- [10] 秦立厚, 张茂震, 钟世红, 等. 森林生物量估算中模型不确定性分析[J]. 生态学报, 2017, **37**(23): 7912–7919.
QIN Lihou, ZHANG Maozhen, ZHONG Shihong, et al. Model uncertainty in forest biomass estimation [J]. *Acta Ecol Sin*, 2017, **37**(23): 7912–7919.
- [11] SHETTLES M, TEMESGEN H, GRAY A N, et al. Comparison of uncertainty in per unit area estimates of aboveground biomass for two selected model sets [J]. *For Ecol Manage*, 2015, **354**: 18–25.
- [12] CLARK D B, KELLNER J R. Tropical forest biomass estimation and the fallacy of misplaced concreteness [J]. *J Veg Sci*, 2012, **23**(6): 1191–1196.
- [13] CHEN Qi, LAURIN G V, VALENTINI R. Uncertainty of remotely sensed aboveground biomass over an African tropical forest: propagating errors from trees to plots to pixels [J]. *Remote Sensing Environ*, 2015, **160**: 134–143.
- [14] 秦立厚. 浙江省森林碳储量估算不确定性研究[D]. 杭州: 浙江农林大学, 2017.
QIN Lihou. *Uncertainty of Forest Carbon Storage Estimation in Zhejiang Province*[D]. Hangzhou: Zhejiang A&F University, 2017.
- [15] 孙雪莲. 基于 Landsat 8-OLI 的香格里拉高山松林生物量遥感估测模型研究[D]. 昆明: 西南林业大学, 2016.
SUN Xuelian. *Biomass Estimation Model of Pinus densata Forests in Shangri-La City Based on Landsat 8-OLI by Remote*

- Sensing[D]. Kunming: Southwest Forestry University, 2016.
- [16] NICHOL J, HANG L K, SING W M. Empirical correction of low sun angle images in steeply sloping terrain: a slope-matching technique [J]. *Int J Remote Sensing*, 2006, **27**(3): 629 – 635.
- [17] 陆驰. 基于 Landsat 的香格里拉市高山松地上生物量及其动态变化建模研究[D]. 昆明: 西南林业大学, 2017.
- LU Chi. *The Estimation and Dynamic Modeling on the Aboveground Biomass of Pinus densata in Shangri-La Based on Landsat*[D]. Kunming: Southwest Forestry University, 2017.
- [18] HARALICK R M, SHANMUGAM K, DINSTEN I. Textural features for image classification [J]. *Stud Media Commun*, 1973, **3**(6): 610 – 621.
- [19] 岳彩荣. 香格里拉县森林生物量遥感估测研究[D]. 北京: 北京林业大学, 2011.
- YUE Cairong. *Forest Biomass Estimation in Shangri-La County Based on Remote Sensing*[D]. Beijing: Beijing Forestry University, 2011.
- [20] 张加龙, 胥辉. 基于遥感的高山松连清固定样地地上生物量估测模型构建[J]. *北京林业大学学报*, 2020, **42**(7): 1 – 11.
- ZHANG Jialong, XU Hui. Establishment of remote sensing based model to estimate the aboveground biomass of *Pinus densata* for permanent sample plots from national forestry inventory [J]. *J Beijing For Univ*, 2020, **42**(7): 1 – 11.
- [21] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine [J]. *Ann Stat*, 2001, **29**(5): 1189 – 1232.
- [22] OU Guanglong, LÜ Yanyu, XU Hui, *et al.* Improving forest aboveground biomass estimation of *Pinus densata* forest in Yunnan of southwest China by spatial regression using Landsat 8 images [J]. *Remote Sensing*, 2019, **11**(23): 738 – 762.
- [23] HOSMER D W, LEMESHOW S. *Applied Logistic Regression* [M]. New York: John Wiley & Sons, 2000.
- [24] 李金海. 误差理论与测量不确定度评定[M]. 北京: 中国计量出版社, 2003.
- LI Jinhai. *Error Theory and Measurement Uncertainty Assessment*[M]. Beijing: China Metrology Publishing House, 2003.
- [25] 陈蜀蓉, 张超, 郑超超, 等. 公益林生物量估算方法研究——以浙江省缙云县公益林为例[J]. *浙江林业科技*, 2015, **35**(5): 20 – 28.
- CHEN Shurong, ZHANG Chao, ZHENG Chaochao, *et al.* Estimation methods for biomass of ecological forest in Jinyun [J]. *J Zhejiang For Sci Technol*, 2015, **35**(5): 20 – 28.
- [26] CHAVE J, CONDIT R, AGUILAR S, *et al.* Error propagation and scaling for tropical forest biomass estimates [J]. *Philos Trans R Soc Lond Series B Biol Sci*, 2004, **359**(1443): 409 – 420.