

DOI:10.13275/j.cnki.lykxyj.2021.06.003

气候敏感的落叶松人工林 林分生物量模型研究

何 潇^{1,2}, 徐奇刚^{1,2}, 雷相东^{1,2*}

(1. 中国林业科学研究院资源信息研究所, 北京 100091; 2. 国家林业和草原局森林经营与生长模拟实验室, 北京 100091)

摘要: [目的] 建立林分生物量模型, 分析不同因子对林分生物量的影响, 为区域尺度生物量的估算提供模型和依据。[方法] 以东北和华北地区 7 个省份的落叶松人工林为研究对象, 利用第 8 次一类清查固定样地数据, 采用普通最小二乘方法和稳健回归方法建立林分生物量模型。采用主成分分析和相关分析法筛选气候变量, 建立气候敏感的林分生物量模型 (包括地上生物量 A_{GB} 和总生物量 T_{GB})。使用决定系数 (R^2)、均方根误差 ($RMSE$) 和相对均方根误差 ($rRMSE$) 评价模型。将各因子的解释率分解为独立解释和共变部分, 量化不同因子的解释率。[结果] (1) 最优基础模型形式为变量直接引入, 最优 A_{GB} 和 T_{GB} 模型 R^2 分别为 0.967、0.953, 地上部分的大于总量。普通最小二乘回归和稳健回归结果类似, 稳健回归稍优于普通最小二乘回归, 基于稳健回归的最优 A_{GB} 模型的 $RMSE$ 、 $rRMSE$ 要比对应的普通最小二乘回归模型分别低 0.046 t·hm⁻²、0.085%, 对于 T_{GB} 则分别低 0.059 t·hm⁻²、0.081%。(2) A_{GB} 和 T_{GB} 与湿热指数 (AHM) 相关系数较大, 分别为 -0.350 和 -0.363, 气候敏感的林分生物量模型进一步提升了的模型预测效果, A_{GB} 的 R^2 提高了 0.41%, 而 $RMSE$ 和 $rRMSE$ 降低了 6.85%; T_{GB} 的 R^2 提高了 0.63%, 误差统计量降低了 6.79%。(3) A_{GB} 和 T_{GB} 的林分因子独立解释分别为 87.37%、82.32%, 气候因子独立解释分别为 0.40%、0.60%, 共变部分分别为 9.33%、9.98%, 林分因子的解释率远大于气候因子, 共变部分较大。[结论] 当林分生物量模型的建模数据质量较高时, 稳健回归和普通最小二乘回归建立的模型差异不大, 但气候因子对林分生物量具有显著影响, 需要建立气候敏感的林分生物量模型进行生物量估计。

关键词: 林分生物量模型; 稳健回归; 林分因子; 气候因子; 解释率量化

中图分类号: S758.5

文献标志码: A

文章编号: 1001-1498(2021)06-0020-08

森林生态系统中存储的生物量-碳储量是全球碳循环的基础。研究表明, 全球森林生态系统大约含有 861 ± 66 Gt 碳储量^[1], 这对于全球气候调节服务至关重要^[2]。生物量估计的常用方法是建立回归模型, 目前, 国内外研建了大量的单木生物量模型^[3-7], 这些单木模型是林分生物量计算的基础。由于森林资源调查数据中可能没有单木数据, 仅记录了林分因子, 导致单木生物量模型的应用受限。而在大尺度上更直接有效的方法是建立林分生物量模型, 利

用林分因子估计生物量, 但是林分层次的生物量模型并不多见^[8-10]。此外, 为了制定应对气候变化的适应性经营策略, 迫切需要了解气候对生物量估算的影响。Rudgers 等^[11]发现林木生物量及其分配会受到气候的影响, 使得生物量模型的参数变化, 进而影响生物量估计结果。Fu 等^[12]和 Zeng 等^[13]建立了气候敏感的单木生物量异速生长方程, 发现气候与生物量之间存在显著关系。Usoltsev 等^[14]将气候变量引入可加性林分生物量方程中以确定气候对

收稿日期: 2020-11-15 修回日期: 2021-03-04

基金项目: 中央级公益性科研院所基本科研业务费专项 (CAFYBB2019SY025); 林业公益性行业科研专项项目“我国主要林区林地地地质量和生产力评价研究”(201504303)

* 通讯作者: 雷相东, 研究员, 博士生导师。主要研究方向: 森林生长模型与模拟。Email: xdlei@jifrit.ac.cn

林分生物量的影响。因此, 在生物量估计模型中应考虑气候变量, 以减少模型的不确定性。

常见的模型参数估计方法是最小二乘法, 然而森林生长受到立地质量^[15]和外界干扰^[16]的影响, 可能导致调查数据存在较多的离群点, 加上生物量数据的异方差性^[17], 如果使用最小二乘法求解参数可能会导致模型参数的有效性降低, 引起模型的不确定性增加^[18]。因此, 有学者提出了用稳健回归来尽可能降低离群点对参数估计的影响^[19], 郭清文等^[18]的实验结果证明: 基于稳健回归的生物量模型不仅能有效降低建模数据中离群点带来的影响, 还使得参数估计在理论上接近最优。

本研究利用东北和华北地区的落叶松 (*Larix* spp.) 人工林生物量数据, 比较稳健回归方法和普通最小二乘法求解林分生物量模型的差异, 并在模型中考虑气候因子, 量化分析林分因子和气候因子对林分生物量 (包括林分地上生物量 A_{GB} 和林分总生物量 T_{GB}) 的解释率, 以期为大尺度生物量精准估计提供基础模型和方法。

1 材料与方法

1.1 数据来源

数据来源于黑龙江、吉林、辽宁、北京、河北、山西和内蒙古等 7 个省份 (34°34'~53°33' N, 97°12'~135°05' E) 第 8 次森林资源连续清查 (2009—2013) 中的落叶松人工纯林固定样地。经逻辑检查后共得到 445 块有效样地数据。样地调查因子包括平均年龄、平均胸径以及平均高等, 其中平均高是在样地中测量 3~5 株平均木的树高取平均值得到。考虑到划分训练集和验证集可能会导致气候因子对林分生物量的解释率量化受到抽样的影响, 因此本研究没有进行样本划分, 各因子统计量见表 1。气候因子来自 ClimateAP^[20], 根据每个样地的地理坐标和海拔下载, 气候因子变量名及其描述见表 2, 样地气候因子为 1981—2010 期间的平均值。

1.2 林分生物量计算

林分生物量使用单木生物量模型计算, 然后累加得到。我国林业行业标准《LY/T 2654—2016: 立木生物量模型及碳计量参数—落叶松》^[4]中就提供了落叶松的单木生物量模型, 对于有行业标准的伴生树种也采用行业标准中给出单木生物量模型进行计算, 剩余树种则使用文献^[3]和^[5]中给出的

表 1 林分因子和林分生物量统计量

Table 1 Summary statistics of stand variables and biomass

变量 Variables	平均值 Mean	标准差 S.D.	最小值 Min.	最大值 Max.
平均年龄 Age/a	28	9	11	60
平均高 H/m	12.0	4.0	4.2	24.0
株数密度 N (株·hm ⁻²)	1021	595	263	3933
平均胸径 D_g /cm	13.4	4.3	6	26.4
林分蓄积 V (m ³ ·hm ⁻²)	82.81	53.22	3.28	282.25
地上生物量 A_{GB} (t·hm ⁻²)	53.98	33.49	2.75	168.07
总生物量 T_{GB} (t·hm ⁻²)	72.05	44.53	3.74	226.50

表 2 气候因子变量名称及其含义

Table 2 The name of climatic variables and its description

变量 Variables	描述 Description
<i>AHM</i>	湿热指数 Annual heat: moisture index
<i>CMD</i>	哈格里夫斯气候水分亏缺 Hargreaves climatic moisture deficit
<i>DD_0</i>	0℃以下的天数 Degree-days below 0℃
<i>DD_18</i>	18℃以下的天数 Degree-days below 18℃
<i>DD18</i>	18℃以上的天数 Degree-days above 18℃
<i>DD5</i>	5℃以上的天数 Degree-days above 5℃
<i>EMT/℃</i>	近30年期间的极端最低气温 Extreme minimum temperature over a 30-year period
<i>EXT/℃</i>	近30年期间的极端最高气温 Extreme maximum temperature over a 30-year period
<i>EREF</i>	哈格里夫斯相对蒸发 Hargreaves reference evaporation
<i>MAP/mm</i>	年平均降水量 Mean annual precipitation (mm)
<i>MAT/℃</i>	年平均温度 Mean annual temperature
<i>MCMT/℃</i>	最冷月份的平均温度 Mean coldest month temperature
<i>MWMT/℃</i>	最暖月份的平均温度 Mean warmest month temperature
<i>NFFD</i>	无霜期天数 The number of frost-free days
<i>PAS/mm</i>	上年8月至当年度7月期间的降雪量 Precipitation as snow between August in previous year and July in current year
<i>TD/℃</i>	最暖月份的平均温度和最冷月份的平均温度之差 Temperature difference between MWMT and MCMT, or continentality

单木生物量模型进行计算。

1.3 最小二乘回归与稳健回归

最小二乘是回归模型参数求解的常用方法, 但如果建模数据中存在离群点且排除是数据录入错误等原因所导致的, 则使用最小二乘回归可能得到的

不是最优估计。而稳健回归是在剔除离群点和保留离群点之间的一个折中，根据观测值的稳健情况对每个样本点进行赋权，对应的优化目标为式(1)。

$$\arg \min : \sum w_i (y_i - \hat{y}_i)^2 \quad (1)$$

式中， w_i 是第 i 个样本对应的权重， y_i 为第 i 个样本的因变量实测值； \hat{y}_i 为第 i 个样本的模型估计值。

本研究使用 Huber 提出的稳健回归方法(式2)^[19]，可以看出权重估计依赖于残差分布，而残差估计又反过来需要权重。因此，稳健回归是一个迭代重复加权最小二乘直到结果收敛，获得稳健的参数估计结果的过程。

$$w_j(u_j) = \begin{cases} 1 & |u_j| \leq 1.5 \\ \frac{1.5}{|u_j|} & |u_j| > 1.5 \end{cases} \quad (2)$$

式中， w_j 表示第 j 次迭代的权函数； u_j 表示第 j 次迭代的标准化残差。

1.4 基础林分生物量模型与气候敏感的林分生物量模型的建立

董利虎等^[9]指出落叶松林分生物量模型的误差结构是相乘的，本研究以式(3)~(6)作为基础林分生物量模型，分别采用最小二乘法 and 稳健回归法求解模型参数。因为使用对数回归来消除生物量数据的异方差，在预测时会产生偏差，曾伟生等^[21]详细讨论了非线性模型对数回归的偏差修正原理与方法，本研究采用 Baskerville^[22]提出的校正因子 CF ，对模型进行修正。

$$\log(B) = \beta_1 + \beta_2 \times \log(Dg) + \beta_3 \times \log(N) + \varepsilon \quad (3)$$

$$\log(B) = \beta_1 + \beta_2 \times \log(Dg^2H) + \beta_3 \times \log(N) + \varepsilon \quad (4)$$

$$\log(B) = \beta_1 + \beta_2 \times \log(Dg^2N) + \beta_3 \times \log(H) + \varepsilon \quad (5)$$

$$\log(B) = \beta_1 + \beta_2 \times \log(Dg) + \beta_3 \times \log(H) + \beta_4 \times \log(N) + \varepsilon \quad (6)$$

式中， B 为林分生物量，包括 A_{GB} 和 T_{GB} ； Dg 为林分平均胸径； N 为株数密度； H 为林分平均高； β_i ($i=1 \sim 4$) 为模型参数； ε 为误差项。

向最优基础林分生物量模型中添加气候变量，为了降低气候变量的共线性，本研究采用主成分分析和相关分析方法筛选气候变量，步骤如下：

(1) 对所有气候变量进行主成分分析；(2) 选择因子载荷较大的气候变量；(3) 通过相关分析筛选出气候变量之间相关性较低但和林分生物量相关性较高的气候变量。最后利用筛选出来的气候变量

建立气候敏感的林分生物量模型。

1.5 模型评价

模型评价包括两方面：首先确保模型参数在 0.05 水平下显著；其次使用决定系数 (R^2)、均方根误差 ($RMSE$)、相对均方根误差 ($rRMSE$) 等指标对模型进行评价，公式见文献 [23]。

1.6 各因子解释率量化

林分生物量的总解释率可被划分为：(1) 独立解释的部分 (a ：林分因子的独立解释率， c ：气候因子的独立解释率)；(2) 共变的部分 (b)，即无法确定是由哪个因子所解释的部分；(3) 模型不可解释的部分 (d)。计算方法如下：含林分因子和气候因子的全模型的 R^2 即为总解释率 $a + b + c$ ，且存在 $a + b + c = 1 - d$ 的关系，最优基础林分生物量模型的 R^2 为 $a + b$ ，仅以气候因子为自变量的林分生物量模型的 R^2 为 $b + c$ 。由上述关系可计算 a 、 b 、 c 和 d 。

2 结果与分析

2.1 基础林分生物量模型优选

所有模型参数在 0.05 水平下显著(表3、表4)。对于同一形式的模型，2种方法的参数估计值略有差异，从参数符号可以看出 A_{GB} 和 T_{GB} 均随 Dg 、 N 和 H 的增加而增加，但 CF 没有明显的区别。从基础模型的评价指标发现， A_{GB} 模型优于 T_{GB} 模型。对于 A_{GB} 而言，采用直接引入变量的模型比变量组合形式的模型拟合效果要好，除 5 号和 6 号模型的评价指标显示出普通最小二乘回归略优于稳健回归外，其余模型均显示出稳健回归略优于普通最小二乘回归，其中模型 8 的 R^2 最高，可达 0.967，稳健回归的 $RMSE$ 和 $rRMSE$ 要比普通最小二乘回归分别低 0.046 t·hm⁻² 和 0.085%，二者差异不大。对于 T_{GB} 而言，其规律和 A_{GB} 相似，所有 T_{GB} 模型中，15 号和 16 号模型的 R^2 可达 0.953，稳健回归的误差统计量略好一些，但比 A_{GB} 模型低 0.014。由于建模数据的质量较好，两种方法建立的模型没有太大差异，稳健回归的优越性没有得到充分体现，甚至普通最小二乘回归略优于稳健回归。不论是 A_{GB} 还是 T_{GB} ，最优基础模型形式相同，采用稳健回归的方法求得的参数略优。因此，本研究分别以模型 8 和模型 16 作为 A_{GB} 和 T_{GB} 的最优基础林分生物量模型，用于解释率的量化分解。

表3 基础林分地上生物量模型拟合与评价结果

Table 3 Fitting and evaluation results of basic stand aboveground biomass model

序号 Number	模型 Model	方法 Method	模型参数 Model parameters				R^2	$RMSE/(t \cdot hm^{-2})$	$rRMSE/\%$	CF
			β_1	β_2	β_3	β_4				
1	式3	OLS	-9.279 0 (0.097 9)	2.472 5 (0.018 4)	0.992 7 (0.010 5)	—	0.940	8.203	15.198	1.058 28
2		Huber	-9.274 9 (0.099 8)	2.471 8 (0.018 8)	0.992 3 (0.010 7)	—	0.940	8.197	15.186	1.058 28
3	式4	OLS	-8.337 2 (0.116 6)	0.805 2 (0.007 6)	0.889 6 (0.012 9)	—	0.927	9.050	16.766	1.073 51
4		Huber	-8.324 0 (0.114 3)	0.808 2 (0.007 4)	0.884 0 (0.012 6)	—	0.928	8.997	16.669	1.073 57
5	式5	OLS	-9.066 0 (0.095 0)	0.986 7 (0.009 9)	0.453 2 (0.019 1)	—	0.961	6.639	12.299	1.055 24
6		Huber	-9.069 3 (0.098 3)	0.984 4 (0.010 3)	0.466 5 (0.019 8)	—	0.959	6.764	12.531	1.055 28
7	式6	OLS	-9.044 7 (0.085 8)	2.184 7 (0.027 6)	0.295 6 (0.023 3)	0.960 5 (0.009 3)	0.966	6.145	11.384	1.049 74
8		Huber	-9.001 8 (0.090 6)	2.153 2 (0.029 1)	0.326 1 (0.024 6)	0.955 1 (0.009 9)	0.967	6.099	11.299	1.049 84

注: 1. 括号中的数值为参数标准差; 2. OLS代表普通最小二乘回归, Huber代表稳健回归。下同。

Notes: 1. The value in parentheses is the standard deviation of the parameter. 2. OLS stands for general least square regression, Huber stands for robust regression. The same as below.

表4 基础林分总生物量模型拟合与评价结果

Table 4 Fitting and evaluation results of basic stand total biomass model

序号 Number	模型 Model	方法 Method	模型参数 Model parameters				R^2	$RMSE/(t \cdot hm^{-2})$	$rRMSE/\%$	CF
			β_1	β_2	β_3	β_4				
9	式3	OLS	-8.959 7 (0.115 5)	2.468 9 (0.021 7)	0.989 4 (0.012 4)	—	0.914	13.083	18.158	1.069 22
10		Huber	-8.959 2 (0.120 1)	2.469 9 (0.022 6)	0.989 0 (0.012 9)	—	0.913	13.112	18.199	1.069 22
11	式4	OLS	-8.051 4 (0.120 0)	0.807 1 (0.007 8)	0.888 0 (0.013 2)	—	0.926	12.109	16.807	1.075 74
12		Huber	-8.040 9 (0.117 9)	0.810 6 (0.007 6)	0.882 2 (0.013 0)	—	0.926	12.066	16.747	1.075 80
13	式5	OLS	-8.695 1 (0.106 2)	0.972 5 (0.011 1)	0.488 2 (0.021 4)	—	0.950	9.902	13.744	1.061 94
14		Huber	-8.681 5 (0.110 0)	0.967 5 (0.011 5)	0.508 0 (0.022 1)	—	0.949	10.032	13.924	1.062 00
15	式6	OLS	-8.677 2 (0.100 6)	2.122 0 (0.032 4)	0.356 4 (0.027 3)	0.950 6 (0.011 0)	0.953	9.687	13.445	1.058 58
16		Huber	-8.634 2 (0.104 4)	2.089 2 (0.033 6)	0.390 6 (0.028 4)	0.944 4 (0.011 4)	0.953	9.628	13.364	1.058 69

2.2 气候敏感的林分生物量模型

主成分分析表明, 前3个主成分的贡献率分别为56.0%、22.9%、16.7%, 其中第一主成分中载荷大于0.3的气候变量有: MAT , $NFFD$, EMT , 主要是和温度有关的气候变量; 第二主成分中载荷大于0.3的气候变量有: CMD , AHM , 主要是和湿热特征有关的气候变量; 第三主成分中载荷大于0.33的气候变量有: TD , EXT , $MWMT$, 主要是和极端天气有关的气候变量。从相关分析(表5)的结果看出: A_{GB} 和 T_{GB} 的结果相似, 都与 AHM 和 CMD 的相关系数较高, 而与其他气候因子的较低。由于 AHM 和 CMD 之间的相关系数为0.856, 而 AHM 和林分生物量的关系更密切, 因此将 AHM 添加到最优基础林分生物量模型中, 得到气候敏感的林分生物量模型。

从表6看出, 气候敏感的林分生物量模型(模型17~20)的参数均显著, 2种算法的参数估计结果略有差异, A_{GB} 和 T_{GB} 均随 AHM 增加而降低, 但矫正因子 CF 基本没有差异。根据评价指标, 稳健回归稍优于普通最小二乘回归, 而气候敏感的模型明显优于最优基础模型: 对于 A_{GB} , 模型18与4相比, R^2 提高了0.41%, 而 $RMSE$ 和 $rRMSE$ 降低了6.85%; 对于 T_{GB} , 模型20与16相比, R^2 提高了0.63%, 而 $RMSE$ 和 $rRMSE$ 降低了6.79%。这说明建立气候敏感的林分生物量在统计上是合理的, 气候因子对林分生物量有显著影响, 加入 AHM 后, 模型的拟合优度提升。本研究分别以模型18和模型20作为 A_{GB} 和 T_{GB} 的最优气候敏感的林分生物量模型, 用于解释率的量化解。

表 5 林分生物量与部分气候变量的相关系数矩阵

Table 5 Correlation coefficient matrix between stand biomass and some climatic variables

变量 Variables	A_{GB}	T_{GB}	MAT	$NFFD$	EMT	CMD	AHM	TD	EXT	$MWMT$
A_{GB}	1	0.999	-0.085	-0.090	-0.082	-0.285	-0.350	0.074	-0.025	-0.059
T_{GB}		1	-0.082	-0.078	-0.084	-0.306	-0.363	0.094	-0.001	-0.037
MAT			1	0.956	0.937	0.053	0.376	-0.555	0.461	0.698
$NFFD$				1	0.870	-0.076	0.296	-0.373	0.545	0.802
EMT					1	0.095	0.418	-0.758	0.179	0.458
CMD						1	0.856	-0.249	-0.102	-0.114
AHM							1	-0.389	0.040	0.142
TD								1	0.430	0.202
EXT									1	0.915
$MWMT$										1

表 6 气候敏感的林分生物量模型拟合与评价结果

Table 6 Fitting and evaluation results of climate sensitive stand biomass basic model

序号 Number	模型 Model	方法 Method	模型参数 Model parameters					R^2	RMSE/ (t·hm ⁻²)	rRMSE/%	CF
			β_1	β_2	β_3	β_4	β_5				
17	$\log(A_{GB}) = \beta_1 + \beta_2 \times \log(Dg) + \beta_3 \times \log(H) + \beta_4 \times \log(N) + \beta_5 \times \log(AHM)$	OLS	-8.105 4 (0.101 2)	2.202 1 (0.023 4)	0.225 69 (0.020 4)	0.957 8 (0.007 9)	-0.257 7 (0.019 4)	0.971	5.730	10.616	1.041 89
18		Huber	-8.090 6 (0.106 0)	2.191 8 (0.024 5)	0.238 6 (0.021 3)	0.955 1 (0.008 3)	-0.258 3 (0.020 3)	0.971	5.708	10.575	1.041 92
19	$\log(T_{GB}) = \beta_1 + \beta_2 \times \log(Dg) + \beta_3 \times \log(H) + \beta_4 \times \log(N) + \beta_5 \times \log(AHM)$	OLS	-7.581 4 (0.118 9)	2.142 3 (0.027 5)	0.274 8 (0.023 40)	0.947 6 (0.009 3)	-0.300 6 (0.022 8)	0.959	9.037	12.543	1.049 41
20		Huber	-7.561 2 (0.124 3)	2.130 2 (0.028 7)	0.290 3 (0.025 0)	0.944 0 (0.009 7)	-0.301 8 (0.023 8)	0.959	9.016	12.514	1.049 44

2.3 林分因子和气候因子对林分生物量的影响

分别以 A_{GB} 和 T_{GB} 为对象，量化气候因子和林分因子对林分生物量的贡献，结果如图 1 所示。对于 A_{GB} ：气候因子的总解释率 ($b + c$) 为 9.73%，其中独立解释部分 (c) 为 0.40%，而林分因子的总解释率 ($a + b$) 为 96.70%，独立解释部分 (a) 为 87.37%，共变部分 (b) 为 9.33%，二者的总解释率 ($a + b + c$) 为 97.10%。对于 T_{GB} ：气候因子的总解释率 ($b + c$) 为 10.58%，其中独立解释部

分 (c) 为 0.60%，而林分因子的总解释率 ($a + b$) 为 95.30%，独立解释部分 (a) 为 85.32%，共变部分 (b) 为 9.98%，二者共同解释率为 95.90%。总的来看，林分因子对林分生物量的解释率很高，加入气候因子后模型解释率明显提升，但其中共变部分较大，气候因子的独立解释率不到 1%。气候因子对 A_{GB} 的独立解释率小于对 T_{GB} 的独立解释率，说明不同组分的生物量对于气候的响应程度有所不同。

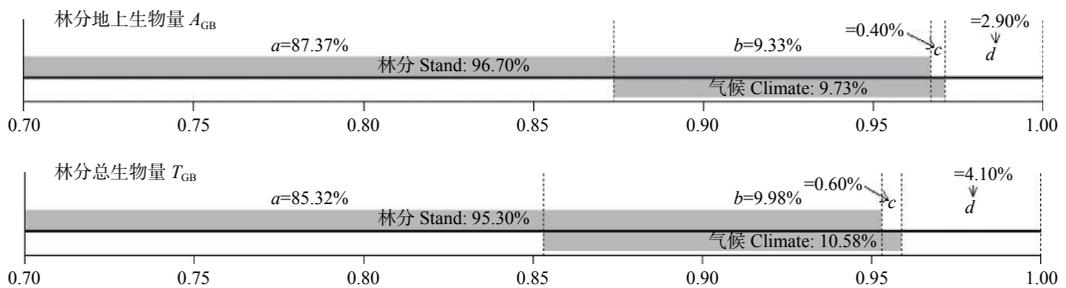


图 1 林分因子和气候因子对林分生物量的解释率分解

Fig. 1 Explanation rate decomposition of stand and factors to stand biomass

3 讨论

本研究采用普通最小二乘法和稳健回归方法建立落叶松人工林林分生物量模型。从模型形式上看, 不论是 A_{GB} 还是 T_{GB} , 最优基础模型的自变量均为 Dg 、 H 和 N 同时直接引入的方式, 而采用组合变量的形式, 模型精度略低, 这与单木层次类似^[24]。其原因是组合变量相当于对参数进行了限制, 使得自变量的作用减弱。洪奕丰等^[25]指出建立林分生物量模型需要多维度的自变量, 而在本研究中, Dg 、 H 和 N 较好的描述了林分生物量, R^2 在 0.95 以上, 该结果和徐奇刚等^[26]的结果一致。从生物量组分上看, 林分地上生物量优于总生物量, 主要是树根生物量的分布不均匀、变异性较大所导致的^[27]。从方法上看, 对数转换虽然是消除生物量模型异方差的有效手段^[21], 但没有解决离群点对模型参数估计的影响, 稳健回归则对残差值较大的样本进行了多次限制, 使得参数估计更加有效, 但这与建模数据的质量有很大关系, 如果建模数据中基本不存在离群点, 稳健回归则会把每一个样本点赋予接近 1 的权重, 那么稳健回归和普通最小二乘回归的差异不大, 对于本研究而言, 由于建模数据的质量较高, 虽然基于稳健回归的模型的误差统计量略小, 但这种差异还不到 1%, 稳健回归的优势就难以发挥。另外, 当前仅开发了基于线性形式的稳健回归方法, 本研究没有对加性误差进行实验, 更有效的手段是基于非线性形式开发稳健回归方法, 对生物量模型的误差结构进行深入分析。

了解林分生物量与气候变量之间的关系才能量化气候因子对林分生物量的影响, 同时有必要在大尺度范围内进行研究, 才能获得较为可信的模拟结

果, 如果尺度太小, 则气候变量的变异较小, 生物量对气候的响应不能完全表达。为了避免这种不确定性, 本研究以我国北方地区 7 个省份的落叶松人工林作为研究对象, 在大尺度上对林分生物量和气候因子之间的关系进行了探索。为了降低共线性的风险, 结合主成分分析和相关分析, 即能筛选出具有统计学意义的气候变量加入到林分生物量模型中, 是一套合理、方便的方法。从理论上讲, 气候敏感的林分生物量模型比传统的模型更好, 而本研究的实验结果也证明: 不论是 A_{GB} 还是 T_{GB} , 考虑气候影响的模型性能优于最优基础模型, 这和以前的研究结果一致^[12-14]。 AHM 是将气温和降水量相结合的一个综合指标, 可用来比较温度和降水量均不同的环境下的气候对林分生物量的影响程度。本研究发现单独使用温度或降水量模拟气候对生物量的影响并不合适, 温度和降水存在共同变化, 从图 2 可以看出生物量与气候因子之间的复杂作用关系, 比如温度较高而降水量过低时, AHM 比较大, 此时林分生物量比较小, 但是当温度和降水量同时变高或变低时, AHM 可能增加也可能会降低, 此时林分生物量的变化波动就比较厉害, 无法用单一气候变量说明林分生物量的变化趋势, 这也印证了与利用主成分分析和相关分析筛选出具有统计学意义的气候变量的重要性。气候因子的变化将会影响林木的生理过程 (如光合和呼吸作用), 导致生物量的积累和分配发生变化^[28]。通过解释率的分解, 本研究发现林分因子对林分生物量的独立解释率在 87% 以上, 气候因子对林分生物量的独立解释率不超过 0.6%, 有大约 10% 的解释率无法确定是由哪一类因子所解释的, 虽然气候因子对林分生物量的解释率不高, 但对林分生物量有显著影响。

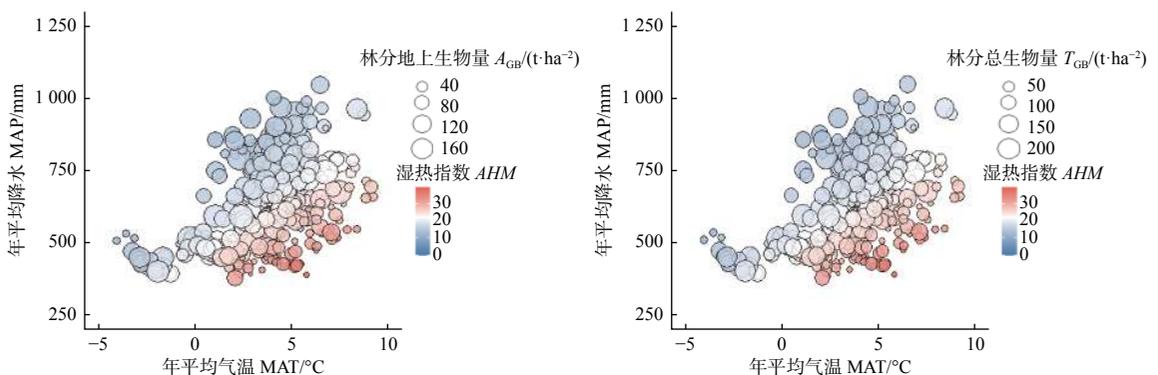


图 2 林分生物量与气候因子的综合作用

Fig. 2 Comprehensive effect between stand biomass and climatic factors

同时,不同生物量组分对于气候的响应程度有所差异,说明气候等非生物因子对于生物量的影响不能被忽视,应该深入探究林分生物量多元驱动机制。

4 结 论

在建模数据质量高的情况下,稳健回归方法和普通最小二乘方法的差异不大,均可用来估计林分生物量的参数。最优基础林分地上生物量和总生物量的模型 R^2 分别为 0.967、0.953,可以对林分生物量进行很好的估计,而通过相关方法筛选气候变量,建立气候敏感的林分生物量模型能够进一步提升模型性能,模型 R^2 分别达到 0.971、0.959,可用于林分生物量精准估计。进一步分析发现:林分因子对林分生物量的独立解释率大于 88%,气候因子对林分生物量的独立解释率不超过 0.6%,但是二者共同解释率大于 95%,因此建立林分生物量模型时不能忽视气候因子的影响,应在大尺度范围内建立气候敏感的生物量模型,以减少生物量模型的不确定性。

参 考 文 献:

- [1] Pan Y, Birdsey R A, Fang J, *et al.* A large and persistent carbon sink in the world's forests[J]. *Science*, 2011, 333(6045): 988-993.
- [2] De Groot R S, Alkemade R, Braat L, *et al.* Challenges in integrating the concept of ecosystem services and values in landscape planning, management and decision making[J]. *Ecological Complexity*, 2010, 7(3): 260-272.
- [3] 陈传国,朱俊凤.东北主要林木生物量手册[M].北京:中国林业出版社,1989.
- [4] 国家林业局.立木生物量模型及碳计量参数落叶松[S].北京:中国标准出版社,2016.
- [5] Wang C. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests[J]. *Forest Ecology and Management*, 2006, 222(1-3): 9-16.
- [6] Návár J. Biomass component equations for Latin American species and groups of species[J]. *Annals of Forest Science*, 2009, 66: 208-208.
- [7] Ter-Mikaelian M T, Korzukhin M D. Biomass equations for sixty-five north American tree species[J]. *Forest Ecology and Management*, 1997, 97(1): 1-24.
- [8] 雷相东,张会儒,牟惠生.东北过伐林区蒙古栎林分相容性生物量模型研究[J].第四纪研究,2010,30(3): 559-565.
- [9] 董利虎,李凤日.大兴安岭东部主要林分类型乔木层生物量估算模型[J].应用生态学报,2018,29(9): 2825-2834.

- [10] 赵嘉诚,李海奎.杉木单木和林分水平地下生物量模型的构建[J].林业科学,2018,54(2): 81-89.
- [11] Rudgers J A, Hallmark, A, Baker S R, *et al.* Sensitivity of dryland plant allometry to climate[J]. *Functional Ecology*, 2019, 33(12): 2290-2303.
- [12] Fu L Y, Lei X D, Hu Z D, *et al.* Integrating regional climate change into allometric equations for estimating tree aboveground biomass of Masson pine in China[J]. *Annals of Forest Science*, 2017, 74(42): 42.
- [13] Zeng W S, Duo H R, Lei X D, *et al.* Individual tree biomass equations and growth models sensitive to climate variables for *Larix* spp. in China[J]. *European Journal of Forest Research*, 2017, 136(2): 233-249.
- [14] Usoltsev V A, Shobairi S O R, Tsepordey I S, *et al.* Modeling the additive structure of stand biomass equations in climatic gradients of Eurasia[J]. *Environmental Quality Management*, 2018, 28(2): 55-61.
- [15] 雷相东,符利勇,李海奎,等.基于林分潜在生长量的立地质量评价方法与应用[J].林业科学,2018,54(12): 116-126.
- [16] 何 潇,李海奎,曹 磊,等.退化森林生态系统中林分碳储量的驱动因素—以内蒙古大兴安岭为例[J].林业科学研究,2020,33(2): 69-76.
- [17] 胥 辉.生物量模型方差非齐性研究[J].西南林学院学报,1999,19(2): 73-77.
- [18] 郭清文,冯仲科,张彦林,等.单木生物量模型误差分析与定权方法探讨[J].中南林业调查规划,2006,25(1): 5-9.
- [19] Huber P J. Robust Regression: Asymptotics, Conjectures, and Monte Carlo[J]. *Annals of Statistics*, 1973, 1(5): 799-821.
- [20] Wang T L, Wang G Y, Innes J L, *et al.* ClimateAP: An application for dynamic local downscaling of historical and future climate data in Asia Pacific[J]. *Frontiers of Agricultural Science and Engineering*, 2017, 4(4): 448-458.
- [21] 曾伟生,唐守正.非线性模型对数回归的偏差校正及与加权回归的对比分析[J].林业科学研究,2011,24(2): 137-143.
- [22] Baskerville G L. Use of logarithmic regression in the estimation of plant biomass[J]. *Canadian Journal of Forest Research*, 1972, 2(1): 49-53.
- [23] 茆诗松,王静龙,史定华,等.统计手册[M].北京:科学出版社,2003.
- [24] 薛春泉,徐期瑚,林丽平,等.广东主要乡土阔叶树种含龄和胸径的单木生物量模型[J].林业科学,2019,55(2): 97-108.
- [25] 洪奕丰,陈东升,申佳朋,等.长白落叶松人工林单木和林分水平的相容性生物量模型研究[J].林业科学研究,2019,32(4): 33-40.
- [26] 徐奇刚,雷相东,国 红,等.基于多层感知机的长白落叶松人工林林分生物量模型[J].北京林业大学学报,2019,41(5): 97-107.
- [27] 黄建辉,韩兴国,陈灵芝.森林生态系统根系生物量研究进展[J].生态学报,1999,19(2): 270-277.
- [28] Emmingham W H, Waring R H. An index of photosynthesis for comparing forest sites in western Oregon[J]. *Canadian Journal of Forest Research*, 1977, 7: 165-174.

Climate-sensitive Stand Biomass Model for *Larix* spp. Plantation

HE Xiao^{1,2}, XU Qi-gang^{1,2}, LEI Xiang-dong^{1,2}

(1. Research Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China;

2. Key Laboratory of Forest Management and Growth Modeling, National Forestry and Grassland Administration, Beijing 100091, China)

Abstract: [Objective] To Establish stand biomass model and analyze the effects of stand factors and climatic factors on stand biomass, so as to provide a model for the biomass estimation for regional scale. [Method] Taking *Larix* spp. plantations from 7 provinces in Northeast and North China as samples, the data from 8th national forest inventory of permanent sample plots were used to establish the basic stand biomass models of aboveground and total base on log-transformed by least square regression and robust regression, respectively. The climate variables were screened by principal component analysis and correlation analysis, and climate-sensitive stand biomass models (including aboveground biomass A_{GB} and total biomass T_{GB}) were established. The determination coefficient (R^2), root mean square error ($RMSE$) and relative root mean square error ($rRMSE$) were used to evaluate the models. The interpretation rate of each factor was divided into independent interpretation and covariant parts and the interpretation rate of different factors was quantified. [Result] (1) The form of the optimal basic model was the direct introduction of variables. The R^2 of the optimal A_{GB} and T_{GB} models were 0.967 and 0.953, respectively, and the A_{GB} were larger than the T_{GB} biomass. The result of ordinary least square regression was similar to that of robust regression, and robust regression was slightly better than ordinary least square regression. The $RMSE$ and $rRMSE$ of the optimal model based on robust regression were lower than the corresponding ordinary least square regression for A_{GB} by $0.046 \text{ t}\cdot\text{hm}^{-2}$ and 0.085%, respectively, and for T_{GB} by $0.059 \text{ t}\cdot\text{hm}^{-2}$ and 0.081%, respectively. (2) The correlation coefficients between A_{GB} , T_{GB} and moisture-heat index (AHM) were -0.350 and -0.363 , respectively. The climate-sensitive stand biomass model further improved the prediction effect of the model. The R^2 of A_{GB} increased by 0.41%, while $RMSE$ and $rRMSE$ decreased by 6.85%. The R^2 of T_{GB} increased by 0.63%, and the error statistics decreased by 6.79%. (3) The independent interpretations of stand factor of A_{GB} and T_{GB} were 87.37% and 82.32% respectively, the independent interpretations of climate factor were 0.40% and 0.60% respectively, and the covariant parts were 9.33% and 9.98%. The interpretation rate of stand factor was much higher than that of climate factor and the covariant part was large. [Conclusion] When the modeling data quality of the stand biomass model is high, there is little difference between the models established by robust regression and ordinary least square regression, but climate factors have a significant impact on stand biomass. It is necessary to establish a climate-sensitive stand biomass model to estimate the biomass.

Keywords: stand biomass model; robust regression; stand factor; climatic factor; quantitative interpretation rate

(责任编辑: 彭南轩)