

DOI:10.13275/j.cnki.lykxyj.2021.04.006

我国 3 种针叶林的材积源生物量模型研建

曾伟生*

(国家林业和草原局调查规划设计院, 北京 100714)

摘要: [目的] 林分水平的生物量模型或数表, 是森林资源调查监测的必备计量工具。研建 3 种针叶林的材积源生物量模型, 既是方法学探索, 也为生产实践提供参考成果。[方法] 基于我国 3 种主要针叶林 (落叶松、油松、杉木) 的 3000 个固定样地的实测数据, 分别利用普通回归估计、加权回归估计和分段建模方法, 建立了林分水平的材积源生物量模型, 并与已发表的相关模型进行了比较。[结果] 建立的 3 种针叶林的加权回归生物量模型, 其确定系数 R^2 在 0.912~0.937 之间, 平均预估误差 MPE 在 0.93%~1.58% 之间, 总体相对误差 TRE 在 $\pm 2.0\%$ 以内, 独立样本检验的 TRE 在 $\pm 2.6\%$ 以内; 分段建模生物量模型, 其确定系数 R^2 在 0.915~0.953 之间, 平均预估误差 MPE 在 0.81%~1.55% 之间, 总体相对误差 TRE 在 $\pm 0.3\%$ 以内, 独立样本检验的 TRE 在 $\pm 1.3\%$ 以内。用本研究数据对已发表的 3 种针叶林的生物量模型进行适用性检验, 其总体相对误差 TRE 分别达到 11.62%、-25.19% 和 -6.26%, 按不同生物量等级的估计误差更大, 甚至出现系统偏差, 严重超出了允许误差范围。[结论] 林分每公顷生物量与蓄积量呈线性相关; 建立林分生物量模型, 应当采用加权回归方法, 并保证样本数量足够且结构合理; 当一个模型难以准确估计各个等级的生物量时, 可以采用分段建模方法。本研究所建 3 种针叶林的材积源生物量模型, 预估精度高, 可以在实践中推广应用。

关键词: 生物量; 加权回归; 分段建模; 落叶松; 油松; 杉木

中图分类号: S758

文献标志码: A

文章编号: 1001-1498(2021)04-0049-09

森林生物量与森林蓄积量一样, 既是各级森林资源监测的重要指标, 更是反映森林生态系统生产力的重要参数^[1-3]。对森林生物量的估计, 既可通过建立立木生物量模型来获得^[4-5], 也可通过建立林分生物量或生物量转换因子模型来获得^[2, 4]。根据 Luo 等^[6]的综述, 1978 年至 2013 年间, 我国学者已发表了近 200 个树种的 5924 个立木生物量模型。2014 年以来, 国家林业局有计划地编制了我国主要树种的立木生物量模型, 并颁布实施了系列行业标准^[7-13]。但是, 不论是国外^[5, 14-20]还是国内^[21-28], 发表的林分生物量模型都要显著少于立木生物量模型。

在已有的林分水平模型中, 影响最大的是方精云等^[21-22]发表的 21 种森林类型的材积源生物量模型, 该模型在很多研究中得到了引用^[27, 29-32]。此

外, 王斌等^[24]利用 1266 个不同森林类型样地的数据, 建立了我国 16 种森林类型的生物量与蓄积量之间的双曲线模型。经分析, 这些模型存在 3 个方面的不足: 一是建模样本较少, 大部分模型都是建立在小样本基础上。如方精云等^[22]建立的 21 个模型有 18 个 (仅落叶松、油松、杉木除外) 的建模样地数在 30 以下, 王斌等^[24]建立的 16 个模型有 10 个的建模样地数在 50 以下。二是建模方法简单, 基本都是采用普通最小二乘法, 未考虑生物量和蓄积量数据的异方差性。三是评价指标单一, 仅提供了确定系数 R^2 ^[22] 或相关系数 R ^[24] 这一项评价指标, 未提供其他误差方面的评价指标, 其适用性存疑。因此, 对林分生物量建模方法做进一步研究是非常必要的。

本研究利用我国 3 种主要针叶林 (落叶松

收稿日期: 2020-08-16 修回日期: 2020-12-02

基金项目: 国家自然科学基金项目 (31770676)

* 通讯作者: 曾伟生, 博士, 教授级高工。主要研究方向: 森林资源清查与林业数学建模。E-mail: zengweisheng0928@126.com

Larix spp.、油松 *Pinus tabulaeformis* Carr.、杉木 *Cunninghamia lanceolata* (Lamb.) Hook.) 的 3000 个样地的地面实测数据, 采用加权回归估计^[33]和分段建模方法^[34], 建立林分水平的材积源生物量模型, 既为这 3 种森林类型的生物量调查提供计量依据, 也为规范林分生物量建模方法提供科学参考。

1 数据与方法

1.1 数据资料

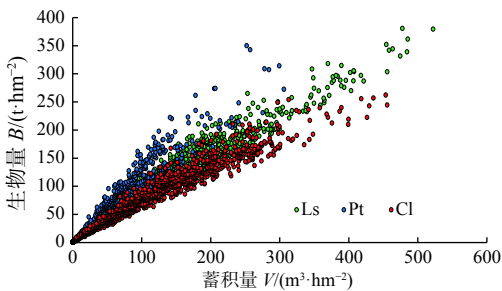
本研究所用数据为第九次全国森林资源清查的固定样地调查资料, 涉及我国 3 种主要的针叶林类型, 即: 落叶松林、油松林和杉木林。按优势树种 (占 65% 以上) 确定的这 3 种类型的针叶林, 全国的有效样地数 (蓄积量大于 0) 分别为 2490、1185 和 3152 个, 每个样地都基于每木胸径测量数据,

采用一元立木材积模型和立木生物量模型计算出蓄积量和生物量 (包括地上生物量和地下生物量)。由于样地数主要集中在每公顷蓄积量和生物量较小的区段, 为了保证所建模型具有广泛适用性, 将全部样地按每公顷生物量大小用上限排外法分为 4 级 (< 50, 50~100, 100~150, ≥ 150 t·hm⁻²), 按每级样本量尽量均等的原则^[35-36]选取建模样本, 每公顷生物量 150 t 以上的样地数相对较少, 尽可能多选一些用作建模样本, 剩下的样地作为检验样本。经综合考虑, 最后确定 3 种林分类型选取建模样地分别为 1200、800 和 1000 个。表 1 是 3 种针叶林分的建模样地数和检验样地数按每公顷生物量等级的分布情况, 图 1 是根据全部 3000 个建模样地数据绘制的散点图。

表 1 3 种针叶林分的建模样本数和检验样本数

Table 1 The number of modeling plots and validation plots for three coniferous forest types

生物量等级 Biomass class/ (t·hm ⁻²)	落叶松 <i>Larix</i> spp.		油松 <i>Pinus tabulaeformis</i>		杉木 <i>Cunninghamia lanceolata</i>	
	建模样地数 Number of modeling plots	检验样地数 Number of validation plots	建模样地数 Number of modeling plots	检验样地数 Number of validation plots	建模样地数 Number of modeling plots	检验样地数 Number of validation plots
< 50	300	431	300	229	300	1 326
50~100	300	550	300	83	300	646
100~150	300	259	120	59	250	154
≥ 150	300	50	80	14	150	26
合计 Total	1 200	1 290	800	385	1 000	2 152



Ls--*Larix* spp. 落叶松; Pt--*Pinus tabulaeformis* 油松; Cl--*Cunninghamia lanceolata* 杉木

图 1 全部建模样地生物量与蓄积量散点图

Fig. 1 The scatterplot of biomass vs. volume for all modeling plots

1.2 建模方法

将基于前述 3 种针叶林 3000 个样地的蓄积量、生物量实测数据, 首先分别普通回归和加权回归, 建立林分生物量模型, 并分析其建模效果, 最

后再用分段建模方法, 建立估计效果更好的生物量模型。

1.2.1 回归估计方法 林分生物量主要与林分蓄积量有关, 基于蓄积量的生物量模型应用最为广泛^[21-22, 24, 27, 29-32]。根据方精云等^[22] 对全国 21 种森林类型的研究结果, 林分生物量与蓄积量之间呈线性相关。从 3000 个样地的每公顷生物量与蓄积量数据的散点图分析, 这种线性相关规律也是非常明显的 (图 1)。因此, 本研究确定采用如下线性形式的林分生物量模型:

$$B = a_0 + b_0 V + \varepsilon_1 \quad (1)$$

式中: B 为每公顷生物量 (t·hm⁻²), V 为每公顷蓄积量 (m³·hm⁻²), a_0 、 b_0 为模型参数, ε_1 为误差项, 假定其服从均值为 0 的正态分布。将 (1) 式两边除以 V , 可得到如下林分生物量转换因子模型:

$$B_{CF} = c_0 + d_0/V + \varepsilon_2 \quad (2)$$

式中: B_{CF} 为生物量转换因子 ($= B/V$), c_0 、 d_0 为模型参数, ε_2 为误差项。式 (2) 为非线性模型, 如果设定 $y = B_{CF}$, $x = 1/V$, 则可以转为以下线性模型:

$$y = c_0 + d_0x + \varepsilon_2 \quad (3)$$

上述模型 (1)、(3) 的参数可采用普通线性回归估计方法求解。根据对模型的结构分析, 模型参数之间理论上应该存在以下关系:

$$c_0 = b_0, d_0 = a_0 \quad (4)$$

事实上, 模型 (2)、(3) 的拟合结果, 就相当于模型 (1) 的加权回归结果, 其权函数为 $w = 1/V$ 。因此, 根据实际的拟合结果, 式 (4) 必然是不成立的。由于生物量数据与蓄积量数据类似, 都具有异方差性, 模型 (1) 应该采用加权回归估计才是合适的^[33, 37]。参照有关权函数的研究结论^[37], 权函数 $w = 1/V$ 效果不一定最好, 更通用的权函数应为 $w = 1/V^k$, 其中 k 一般在 0.5~1.0 之间。

为了区别, 这里将生物量模型 (1) 的拟合方法称为普通回归, 生物量转换因子模型 (2) 或 (3) 的拟合方法称为加权回归 1 (权函数 $w_1 = 1/V$), 以模型 (1) 为基础进行的加权回归估计方法称为加权回归 2 (权函数 $w_2 = 1/V^k$)。

1.2.2 分段建模方法 当变量的取值范围很大时, 用一个模型通常难以对各个区段都作出准确估计, 解决这一问题的有效方法就是分段建模。在建立单木水平的生物量模型时, 就已经有人用到了这一方法^[34]。林分生物量的建立, 同样可能碰到这一问题。假设最小的区段 (如每公顷蓄积量 $50 \text{ m}^3 \cdot \text{hm}^{-2}$ 以下) 存在明显偏估, 就可以将自变量 $V = 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 置为两个模型的链接点, 并将适用于 $V < 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 的模型参数设定为 a_1 和 b_1 , 适用于 $V \geq 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 的模型参数设定为 a_2 和 b_2 。为了保证两个模型在链接点的估计值一致, 先拟合其中一个模型的 2 个参数后, 另一个模型 2 个参数的估计就要受到这一条件的约束, 其中只有一个参数是独立估计的, 另一个参数直接根据 (5) 式由其他 3 个参数推出。

$$a_2 + b_2 \times 50 = a_1 + b_1 \times 50 \quad (5)$$

根据两个分段模型拟合的先后顺序, 可以得出 2 组分段模型:

$$\begin{cases} B = a_1 + b_1V (V < 50) \\ B = a_1 + (b_1 - b_2) \times 50 + b_2V (V \geq 50) \end{cases} \quad (6)$$

$$\begin{cases} B = a_2 + (b_2 - b_1) \times 50 + b_1V (V < 50) \\ B = a_2 + b_2V (V \geq 50) \end{cases} \quad (7)$$

式 (6) 是先拟合适用于 $V < 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 的模型参数 a_1 和 b_1 , 再拟合适用于 $V \geq 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 的模型; 式 (7) 是先拟合适用于 $V \geq 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 的模型参数 a_2 和 b_2 , 再拟合适用于 $V < 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 的模型。通过对比其评价指标的优劣, 选定拟合效果较好的模型。

1.2.3 模型评价方法 用于模型评价的指标包括以下 6 项: 确定系数 R^2 、估计值的标准差 (也称剩余标准差) SEE 、总体相对误差 TRE 、平均系统误差 ASE 、平均预估误差 MPE 和平均百分标准误差 $MPSE$ ^[38-39]。其中 MPE 和 $MPSE$ 的计算公式如下:

$$MPE = t_\alpha \cdot (SEE/\bar{y}) / \sqrt{n} \times 100 \quad (8)$$

$$MPSE = \sum |y_i - \hat{y}_i| / \hat{y}_i / n \times 100 \quad (9)$$

式中: y_i 为实际调查值, \hat{y}_i 为模型预估值, \bar{y} 为样本平均值, n 为样地数, t_α 为置信水平 α 时的 t 值。对建立的回归模型, 计算以上 6 项指标值, 根据指标大小进行模型评价。

从实用性角度考虑, 一般要求模型的 TRE 和 ASE 均在 $\pm 3\%$ 以内, MPE 小于 3% , $MPSE$ 小于 15% 。另外, 残差图也是评价模型的重要参考依据。一个好的模型, 残差应当呈随机分布。也就是说, 模型每个区段的总体相对误差 TRE 都应该相差不大, 一般应在 $\pm 5\%$ 以内。为了评价模型的广泛适用性, 还采用检验样本进行独立交叉检验, 计算模型的总体相对误差 TRE 是否在允许误差范围内。

2 结果与分析

利用 3 种针叶林的 3000 个样地的每公顷蓄积量和生物量数据, 分别采用普通回归、加权回归 1 (权函数 $w_1 = 1/V$) 和加权回归 2 (权函数 $w_2 = 1/V^k$) 拟合线性生物量模型 (1), 其拟合结果和评价指标见表 2。

从表 2 可以明显看出, 不论是 2 个参数的估计值还是 6 项评价指标, 加权回归 2 的结果都居于普通回归和加权回归 1 之间, 且更接近加权回归 1 的结果, 唯有 TRE 和 ASE 这 2 项指标比较特殊: 普

表2 林分生物量模型的参数估计值和模型评价指标

Table 2 The parameter estimates and evaluation indices of stand-level biomass models

林分类型 Forest type	估计方法 Method	参数估计值 Parameter estimate		评价指标 Evaluation indices					
		a_0	b_0	R^2	SEE/m^3	$TRE/\%$	$ASE/\%$	$MPE/\%$	$MPSE/\%$
落叶松 <i>Larix</i> spp.	普通回归 OR	10.4923	0.7409	0.953	15.00	0.00	-4.84	0.81	15.26
	加权回归1 WR1	0.0811	0.8659	0.919	19.64	-5.02	-0.02	1.06	12.74
	加权回归2 WR2	0.2827	0.8342	0.937	17.33	-1.59	2.86	0.93	12.90
油松 <i>Pinus tabulaeformis</i>	普通回归 OR	3.2800	1.0706	0.914	16.57	0.00	-5.28	1.56	18.71
	加权回归1 WR1	0.0476	1.1300	0.911	16.85	-0.91	0.00	1.58	15.43
	加权回归2 WR2	0.0653	1.1243	0.912	16.79	-0.44	0.36	1.58	15.44
杉木 <i>Cunninghamia lanceolata</i>	普通回归 OR	10.1919	0.6172	0.938	14.04	0.00	-5.79	1.01	16.85
	加权回归1 WR1	0.0551	0.7461	0.886	18.94	-6.22	-0.01	1.36	13.36
	加权回归2 WR2	0.2562	0.7121	0.914	16.44	-1.98	3.58	1.18	13.76

注：加权回归1的评价指标是针对生物量模型（1）而不是生物量转换因子模型（2），加权回归2的权函数为 $w_2 = 1/V^{0.7}$ 。

Note: The values of evaluation indices in WR1 are for the biomass model (1), not for the biomass conversion factor model (2), and the weighted function in WR2 is $w_2 = 1/V^{0.7}$; WR means weighted regression, and OR means ordinary regression.

通回归 TRE 为 0, ASE 较大; 加权回归 1 则 ASE 接近于 0, TRE 较大; 而加权回归 2 则处于折中状态, TRE 和 ASE 都与 0 相差不大, 尽可能同时控制在预定的误差范围内 (如 $\pm 3\%$ 以内)。尽管从 R^2 、 SEE 和 MPE 这 3 项指标看, 普通回归模型要好些, 但从 ASE 和 $MPSE$ 看, 则普通回归模型要显著差些。最后, 再来看另外一项重要指标总体相

对误差 TRE , 为了更深入了解模型在不同生物量等级的拟合效果, 表 3 分别落叶松、油松和杉木按建模样本和检验样本列出了总体和各个生物量等级的 TRE 。

从表 3 可以看出, 不论是考虑建模样本还是检验样本, 加权回归模型的结果都要优于普通回归模型, 而加权回归模型 2 又要略优于加权回归模型 1。

表3 林分生物量模型的总体相对误差

Table 3 The total relative errors of stand-level biomass models

林分类型 Forest type	生物量等级 Biomass class/ ($t \cdot hm^{-2}$)	建模样本 Modeling plots			检验样本 Validation plots		
		普通回归 OR	加权回归1 WR1	加权回归2 WR2	普通回归 OR	加权回归1 WR1	加权回归2 WR2
落叶松 <i>Larix</i> spp.	< 50	-19.95%	2.02%	5.00%	-19.53%	2.90%	5.90%
	50~100	-0.82%	-1.36%	2.10%	-0.69%	-1.20%	2.27%
	100~150	2.09%	-4.36%	-0.89%	1.83%	-4.64%	-1.17%
	≥ 150	2.28%	-7.50%	-4.08%	4.40%	-5.24%	-1.74%
	合计 Total	0.00%	-5.02%	-1.59%	-2.07%	-2.37%	1.05%
油松 <i>Pinus tabulaeformis</i>	< 50	-13.22%	-6.59%	-6.18%	-12.86%	-5.81%	-5.40%
	50~100	-1.83%	-2.71%	-2.24%	0.76%	0.01%	0.49%
	100~150	2.45%	-0.22%	0.27%	5.40%	2.75%	3.25%
	≥ 150	7.71%	3.89%	4.41%	9.91%	6.34%	6.87%
	合计 Total	0.00%	-0.91%	-0.44%	-0.68%	0.06%	0.53%
杉木 <i>Cunninghamia lanceolata</i>	< 50	-19.85%	2.48%	6.35%	-18.97%	4.34%	8.26%
	50~100	1.71%	-2.00%	2.37%	1.11%	-2.63%	1.71%
	100~150	3.06%	-6.71%	-2.43%	3.70%	-6.03%	-1.72%
	≥ 150	1.67%	-10.76%	-6.61%	2.52%	-9.68%	-5.49%
	合计 Total	0.00%	-6.22%	-1.98%	-5.50%	-1.64%	2.59%

因此，从模型本身特性、6项评价指标及独立检验结果综合考虑，应当采用加权回归模型2的拟合结果。

然而，如果再仔细查看表3中加权回归模型2在各个生物量等级的TRE（加粗的部分），发现还是存在一些不足，如：生物量小的区段总体上表现为正偏，而生物量大的区段总体上表现为负偏（油松相反）；部分区段TRE较大，超出了 $\pm 5\%$

的范畴。根据对残差图所作的分析，生物量小的区段容易出现较大偏差，因此，如果采用分段建模方法，应该能提高预估精度。综合考虑样本量的支撑程度和规范统一性，本研究将自变量 $V = 50 \text{ m}^3 \cdot \text{hm}^{-2}$ 设置为分段建模的链接点，同时建立了式（6）和式（7）两套模型，拟合结果见表4，基于建模样本和检验样本计算的各个生物量等级的总体相对误差TRE见表5。

表4 分段建模的参数估计值和模型评价指标

Table 4 The parameter estimates and evaluation indices of segmented biomass models

林分类型 Forest type	模型 Model	参数估计值 Parameter estimate				评价指标 Evaluation indices					
		a_1	b_1	a_2	b_2	R^2	SEE/ m^3	TRE/%	ASE/%	MPE/%	MPSE/%
落叶松 <i>Larix</i> spp.	(6)	0.1162	0.9292	8.2572	0.7664	0.953	15.07	-0.10	1.15	0.81	11.56
	(7)	0.0676	0.9797	11.7874	0.7453	0.955	14.74	-0.52	-0.49	0.79	11.39
油松 <i>Pinus tabulaeformis</i>	(6)	0.0413	1.1434	3.0528	1.0832	0.915	16.54	0.07	0.08	1.55	15.41
	(7)	0.0300	1.1575	4.3150	1.0718	0.915	16.50	-0.24	-0.55	1.55	15.42
杉木 <i>Cunninghamia lanceolata</i>	(6)	0.1004	0.7867	6.9799	0.6491	0.936	14.23	0.23	2.46	1.02	11.90
	(7)	0.0170	0.8653	12.4800	0.6160	0.941	13.63	-0.57	-0.63	0.98	11.57

注：用黑体表示的参数估计值，是根据其他3个参数推算的。

Note: The black parameter estimates are derived from other three.

表5 分段生物量模型的总体相对误差

Table 5 The total relative errors of segmented biomass models

林分类型 Forest type	生物量等级 Biomass class/ $(\text{t} \cdot \text{hm}^{-2})$	建模样本 Modeling plots				检验样本 Validation plots			
		模型 (6) Model (6)	模型 (7) Model (7)	模型 (6) Model (6)	模型 (7) Model (7)				
落叶松 <i>Larix</i> spp.	< 50	-4.53%	-9.10%	-3.72%	-8.32%				
	50~100	-0.77%	-2.98%	-0.63%	-2.85%				
	100~150	0.81%	0.47%	0.54%	0.21%				
	≥ 150	0.19%	1.05%	2.35%	3.09%				
	合计 Total	-0.10%	-0.52%	-0.31%	-1.94%				
油松 <i>Pinus tabulaeformis</i>	< 50	-5.01%	-10.32%	-4.20%	-9.69%				
	50~100	-1.61%	-2.98%	1.16%	-0.42%				
	100~150	0.64%	1.27%	3.65%	4.19%				
	≥ 150	4.65%	6.49%	7.16%	8.66%				
	合计 Total	0.11%	-0.56%	1.23%	-0.51%				
杉木 <i>Cunninghamia lanceolata</i>	< 50	-1.76%	-9.61%	-0.17%	-8.25%				
	50~100	1.75%	-1.27%	1.13%	-1.85%				
	100~150	1.03%	1.26%	1.70%	1.88%				
	≥ 150	-1.34%	0.55%	-0.39%	1.31%				
	合计 Total	0.23%	-0.57%	0.77%	-3.11%				

从表4的6项评价指标看，模型（6）和（7）之间的差异不明显；但从表5的对比可以看出，模

型（6）明显优于模型（7），每个生物量等级的误差基本都在 $\pm 5\%$ 以内。除油松的分段模型改进甚

微外, 落叶松和杉木的分段模型有显著改进, 杉木分段模型各生物量等级的误差甚至达到了 $\pm 2\%$ 以内。因此, 综合考虑模型的各项评价指标及检验结果, 最终选定分段模型(6)作为3种针叶林分的生物量估计模型。

3 讨论

本研究针对我国在林分生物量建模方面存在的样本数量偏少、建模方法简单、评价指标单一等问题, 基于第九次全国森林资源清查3000个固定样地的实测数据, 综合利用加权回归方法和分段建模方法, 建立了落叶松、油松、杉木3种主要针叶林的每公顷生物量模型。最终确定的分段回归模型如下:

$$\begin{aligned} \text{落叶松: } B &= 0.1162 + 0.9292V (V < 50 \text{ m}^3 \cdot \text{hm}^{-2}) \\ B &= 8.2572 + 0.7664V (V \geq 50 \text{ m}^3 \cdot \text{hm}^{-2}) \end{aligned} \quad (10)$$

$$\begin{aligned} \text{油松: } B &= 0.0413 + 1.1434V (V < 50 \text{ m}^3 \cdot \text{hm}^{-2}) \\ B &= 3.0528 + 1.0832V (V \geq 50 \text{ m}^3 \cdot \text{hm}^{-2}) \end{aligned} \quad (11)$$

$$\begin{aligned} \text{杉木: } B &= 0.1004 + 0.7867V (V < 50 \text{ m}^3 \cdot \text{hm}^{-2}) \\ B &= 6.9799 + 0.6491V (V \geq 50 \text{ m}^3 \cdot \text{hm}^{-2}) \end{aligned} \quad (12)$$

其确定系数 R^2 在 0.915~0.953 之间, 平均预估误差 MPE 在 0.81%~1.55% 之间, 平均百分标准误差 $MPSE$ 在 11.56%~15.41% 之间。林分生物量与蓄积量呈线性相关, 这与方精云等^[22] 的研究结论是一致的。但是, 由于在样本数量、建模方法等方面存在的差异, 模型的适用性肯定会有很大不

同。在引言中提到方精云等^[22] 建立的 21 个林分类型的生物量模型, 建模样本数量在 30 以上的仅有以下 3 个模型:

$$\text{落叶松: } B = 33.8060 + 0.6096V (n = 34, R^2 = 0.82) \quad (13)$$

$$\text{油松: } B = 5.0928 + 0.7554V (n = 82, R^2 = 0.96) \quad (14)$$

$$\text{杉木: } B = 22.5410 + 0.3999V (n = 56, R^2 = 0.95) \quad (15)$$

表6列出了利用本研究所收集的全部样地对这3个模型的检验结果。可以看出, 3个模型不仅总体的 TRE 远远超出了允许误差范围, 不同生物量等级的估计值更是存在非常大的系统偏差。通过与本研究所建模型参数进行对比, 发现式(13)~(15)的截距参数都要偏大, 而斜率参数都要偏小, 这正是对蓄积量小的样地会产生正偏而对蓄积量大的样地会产生负偏的直接原因。之所以其参数估计值出现大的偏差, 主要原因应该是参数估计方法不恰当, 采用的是普通回归而不是加权回归。其次, 样本量的大小及样本结构的好坏也是影响因素之一。笔者曾试图系统抽取表1中全部样本的2/3建模、1/3检验, 尽管建模样本数量大幅增加, 但因为样本结构不理想, 建模结果并未达到预期要求。因此, 建模成功的要素, 一是样本数量足够; 二是样本结构合理; 三是建模方法科学。表6也列出了利用全部样本对本研究所建模型的检验结果, 总相对误差都在 $\pm 1\%$ 以内, 各个生物量等级的估计误差大都在 $\pm 5\%$ 以内, 最大的也未超出 $\pm 10\%$ 的范围。这样的模型, 才是适用性广的模型。

表6 不同生物量模型总体相对误差的对比

Table 6 The comparison of total relative errors of different biomass models

生物量等级 Biomass class/ (t·hm ⁻²)	落叶松 <i>Larix</i> spp.		油松 <i>Pinus tabulaeformis</i>		杉木 <i>Cunninghamia lanceolata</i>	
	模型(13) Model (13)	模型(10) Model (10)	模型(14) Model (14)	模型(11) Model (11)	模型(15) Model (15)	模型(12) Model (12)
< 50	101.71%	4.23%	-6.84%	7.90%	49.42%	0.46%
50~100	16.91%	0.68%	-24.67%	2.02%	-14.17%	-1.31%
100~150	1.17%	-0.67%	-29.48%	-2.37%	-24.12%	-1.27%
≥ 150	-7.01%	-0.47%	-33.23%	-6.46%	-27.40%	1.22%
合计Total	11.62%	0.19%	-25.19%	-0.20%	-6.26%	-0.52%

图2展示了落叶松生物量模型(10)和模型(13)的残差分布, 可以看出, 因为模型(13)的截距参数 $a_0 = 33.8060$, 会得出每公顷蓄积量为0的落叶松林其生物量高达 $33.8 \text{ t} \cdot \text{hm}^{-2}$ 的结果, 从而导致每公顷蓄积量较小的林分, 其生物量估计结

果出现正偏; 每公顷蓄积量较大的林分, 其生物量估计结果出现负偏。其他2个树种的生物量模型(11)、(12)与模型(14)、(15)的残差分布对比情况也类似, 为省篇幅, 不再列出。

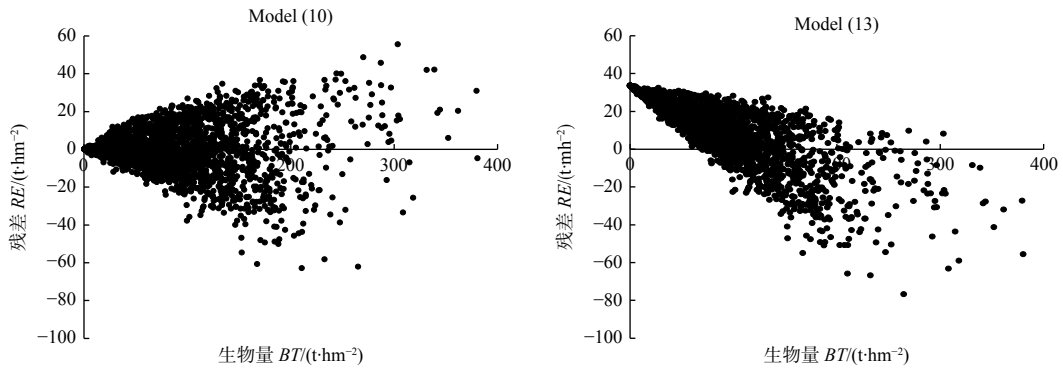


图2 落叶松生物量模型(10)和模型(13)的残差分布对比

Fig. 2 Comparison of residual errors between model (10) and (13) for larch

4 结论

根据本研究的相关结果,可以得出以下结论:

- (1) 林分每公顷生物量与蓄积量呈线性相关。
- (2) 建立林分生物量模型,应当采用加权回归方法;当一个模型难以准确估计各个等级的生物量时,可以采用分段建模方法。
- (3) 样本数量和样本结构是除建模方法之外影响建模效果的另外两个重要因素。
- (4) 本研究所建3种针叶林的生物量模型,预估精度高,可以在实践中推广应用。

最后需要补充的一点是,本研究只是基于优势树种划分的林分类型分别建立材积源生物量模型,没有再分树种组成按绝对纯林(占90%以上)和相对纯林(占65%~90%)分别建模,也没有分起源按天然林和人工林分别建模。因此,用于预估更细的类型时模型的误差肯定会有所增加。若想进一步提高模型的预估精度,可以分别天然林、人工林和绝对纯林、相对纯林建模,或将起源、纯林类型等因子按哑变量对待,建立适应性更广的哑变量模型。另外,本研究所建材积源生物量模型是以立木生物量模型的估计结果为基础建立的,属于林分水平的模型;与单木水平的模型相比,其预估精度要略低^[40-41]。模型应用时,若具备单木水平模型的应用条件,应该首先采用单木模型;若只有样地、林分或小班水平的数据,不具备单木水平模型的应用条件,才考虑采用林分水平模型。

参考文献:

- [1] IUFRO. International Guidelines for Forest Monitoring[R]. IUFRO World Series, Volume 5, Vienna, 1994.
- [2] IPCC. IPCC Guidelines for National Greenhouse Gas Inventory[R/OL]. 2006, [2020-12-02]. <http://www.ipcc-nggip.iges.or.jp>.
- [3] FAO. Global Forest Resources Assessment 2020: Guidelines and Specifications[R]. FRA Working Paper 189, Rome, 2018.
- [4] Somogyi Z, Cienciala E, Mäkipää R, et al. Indirect methods of large-scale forest biomass estimation[J]. Eur J Forest Res, 2007, 126: 197-207.
- [5] Jagodziński A M, Dyderski M K, Gesikiewicz K, et al. Tree and stand level estimations of *Abies alba* Mill aboveground biomass[J]. Annals of Forest Science, 2019, 76: 56.
- [6] Luo Y J, Wang X K, Ouyang Z Y, et al. A review of biomass equations for China's tree species[J]. Earth SystSci Data, 2020, 12(1): 21-40.
- [7] 国家林业局. 立木生物量模型及碳计量参数—落叶松(LY/T 2654—2016)[S]. 北京: 中国标准出版社, 2017.
- [8] 国家林业局. 立木生物量模型及碳计量参数—云杉(LY/T 2655—2016)[S]. 北京: 中国标准出版社, 2017.
- [9] 国家林业局. 立木生物量模型及碳计量参数—冷杉(LY/T 2656—2016)[S]. 北京: 中国标准出版社, 2017.
- [10] 国家林业局. 立木生物量模型及碳计量参数—栎树(LY/T 2658—2016)[S]. 北京: 中国标准出版社, 2017.
- [11] 国家林业局. 立木生物量模型及碳计量参数—桦树(LY/T 2659—2016)[S]. 北京: 中国标准出版社, 2017.
- [12] Zeng W S. Developing tree biomass models for eight major tree species in China[M]//Tumuluru J S (ed.). Biomass Volume Estimation and Valorization for Energy. Croatia: InTech, 2017. Doi: 10.5772/65664.
- [13] Zeng W S. Developing one-variable individual tree biomass models based on wood density for 34 treespecies in China[J]. Forest Research: Open Access, 2018, 7(1): 217-221.
- [14] Shiver B D, Brister G H. Tree and stand volume functions for *Eucalyptus saligna*[J]. Forest Ecology and Management, 1992, 47(s1-4): 211-223.
- [15] Chamshama S A O, Mugasha A G, Zahabu E. Stand biomass and volume estimation for Miombo woodlands at Kitulungalo, Morogoro, Tanzania[J]. Southern African Forestry Journal, 2004, 200: 59-69.
- [16] Castedo-Dorado F, Gómez-García E, Diéguez-Aranda U, et al. Above-ground stand-level biomass estimation: A comparison of two meth-

- ods for major forest species in north west Spain[J]. *Annals of Forest Science*, 2012, 69: 735-746.
- [17] Usoltsev V A, Shobairi S O R, Chasovskikh V P. Triple harmonization of transcontinental allometric of *Picea* spp. and *Abies* spp. forest stand biomass[J]. *Eco Env& Cons*, 2018, 24(4): 1966-1972.
- [18] Jagodziński A M, Dyderski M K, Gesikiewicz K, et al. How do tree stand parameters affect young Scots pine biomass?-Allometric equations and biomass conversion and expansion factors[J]. *Forest Ecology and Management*, 2018, 409: 74-83.
- [19] Jagodziński A M, Dyderski M K, Gesikiewicz K, et al. Tree- and stand-level biomass estimation in a *Larix decidua* Mill. chronosequence[J]. *Forests*, 2018, 9: 587.
- [20] Jagodziński A M, Dyderski M K, Gęsikiewicz K, et al. Effects of stand features of aboveground biomass and biomass conversion and expansion factors based on a *Pinus sylvestris* L. chronosequence in western Poland[J]. *European Journal of Forest Research*, 2019, 138: 673-683.
- [21] 方精云, 刘国华, 徐嵩龄. 我国森林植被的生物量和净生产量[J]. *生态学报*, 1996, 16(5): 497-508.
- [22] Fang J Y, Chen A P, Peng C H, et al. Changes in forest biomass carbon storage in China between 1949 and 1998[J]. *Science*, 2001, 292: 2320-2322.
- [23] 侯振宏, 张小全, 徐德应, 等. 杉木人工林生物量和生产力研究[J]. *中国农学通报*, 2009, 25(5): 97-103.
- [24] 王 斌, 刘某承, 张 彪. 基于森林资源清查资料的森林植被净生产量及其动态变化研究[J]. *林业资源管理*, 2009, (1): 35-42.
- [25] Hou Y N, Wu H L, Zeng W X, et al. Conversion parameters for stand biomass estimation of four subtropical forests in southern China[C]. *DEStech Transactions on Environment Energy and Earth Science*, DEStech Publications, 2017, doi: [10.12783/dteees/eccsd2016/5846](https://doi.org/10.12783/dteees/eccsd2016/5846).
- [26] Mei G Y, Sun Y J, Saeed S. Models for predicting the biomass of *Cunninghamia lanceolata* trees and stands in southeastern China[J]. *PLoS ONE*, 2017, 12(1): 1-14.
- [27] Zhao M M, Yang J L, Zhao N, et al. Estimation of China's forest stand biomass carbon sequestration based on the continuous biomass expansion factor model and seven forest inventories from 1977 to 2013[J]. *Forest Ecology and Management*, 2019, 448: 528-534.
- [28] Dong L H, Zhang L J, Li F R. Evaluation of stand biomass estimation methods for major forest types in the eastern Da Xing'an Mountain, northeast China[J]. *Forests*, 2019, 10: 715.
- [29] 曾伟生. 云南省森林生物量与生产力研究[J]. *中南林业调查规划*, 2005, 24(4): 1-3, 13.
- [30] 欧阳钦. 长沙望城区森林植被生物量及碳储量研究[D]. 长沙: 中南林业科技大学, 2014.
- [31] 梁兴军. 济南市森林植被生物量和碳储量调查研究[D]. 济南: 山东师范大学, 2015.
- [32] 陈小林. 湖南安仁县森林生态系统生物量和碳储量研究[D]. 长沙: 中南林业科技大学, 2016.
- [33] 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.
- [34] 张连金, 曾伟生, 唐守正. 用带截距的非线性方程和分段建模方法对立木生物量估计的比较[J]. *林业科学研究*, 2011, 24(4): 453-457.
- [35] 贺东北, 曾伟生, 骆期邦. 通用性二元立木材积模型的建模样本研究[J]. *中南林业调查规划*, 2001, 20(1): 1-9.
- [36] 国家林业局. 立木生物量建模样本采集技术规程(LY/T 2259—2014)[S]. 北京: 中国标准出版社, 2015.
- [37] 曾伟生. 加权回归估计中不同权函数的对比分析[J]. *林业资源管理*, 2013, (5): 55-61.
- [38] 曾伟生, 唐守正. 立木生物量模型的优度评价和精度分析[J]. *林业科学*, 2011, 47(11): 106-113.
- [39] Zeng W S, Zhang L J, Chen X Y, et al. Construction of compatible and additive individual-tree biomass models for *Pinus tabulaeformis* in China[J]. *Canadian Journal of Forest Research*, 2017, 47: 467-475.
- [40] 曾伟生, 杨学云, 陈新云. 单木和林分水平一元与二元材积模型的预估精度对比[J]. *中南林业调查规划*, 2017, 36(4): 1-6.
- [41] 李海奎, 赵鹏祥, 雷渊才, 等. 基于森林清查资料的乔木林生物量估算方法的比较[J]. *林业科学*, 2012, 48(5): 44-52.

Development of Volume-derived Biomass Models for Three Coniferous Forest Types in China

ZENG Wei-sheng

(Academy of Forest Inventory and Planning, National Forestry and Grassland Administration, Beijing 100714, China)

Abstract: [Objective] Stand-level biomass models/tables are important quantitative tools for implementing forest resources inventory and monitoring. Developing volume-derived biomass models for three coniferous forest types in China is not only an exploration of methodology, but also provides reference for practice. [Method] Based on field measurement data of 3000 sample plots from three coniferous forest types (*Larix* spp., *Pinus tabulaeformis* and *Cunninghamia lanceolata*) in China, the volume-derived biomass models were developed through ordinary regression (OR), weighted regression (WR), and segmented modeling (SM) approaches; and the relevant published models were compared. [Result] The coefficients of determination (R^2) of the volume-derived biomass models for the three coniferous forest types based on WR approach were between 0.912~0.937, the mean prediction errors (MPEs) were between 0.93%~1.58%, the total relative errors (TREs) were within $\pm 2.0\%$, and the TREs for validation were within $\pm 2.6\%$. The R^2 values of the models based on SM approach were between 0.915~0.953, the MPEs were between 0.81%~1.55%, the TREs were within $\pm 0.3\%$, and the TREs for validation were within $\pm 1.3\%$. Using the data of this study to test the applicability of the relevant published biomass models for the three coniferous forest types, the TREs were 11.62%, -25.19% and -6.26%, respectively, and the errors for different biomass classes were quite higher, even systematic deviations appeared, and seriously exceeded the allowable error. [Conclusion] The stand-level biomass per hectare is linearly related to volume stock. The WR approach should be used preferentially when developing volume-derived biomass models, and the sample plots should be much enough and evenly distributed. When one model is not enough to obtain accurate estimates for different biomass classes perfectly, the SM approach can be used. The volume-derived biomass models developed in this study have low MPEs, indicating that they can be applied in practice.

Keywords: biomass; weighted regression; segmented modeling; *Larix* spp.; *Pinus tabulaeformis*; *Cunninghamia lanceolata*

(责任编辑：彭南轩)