

立木生物量线性联立模型研究

贺东北 骆期邦 曾伟生

(国家林业局中南林业调查规划设计院, 长沙 410014)

摘要 根据当今区域性森林生物量资源清查和监测的需要, 针对以往生物量模型存在总量与各维量互不兼容和与森林蓄积量清查不兼容等实际问题, 并以提高模型的估计精度和适用性为目标, 通过引进材积变量、线性联立方程组的建模原理和二段加权最小二乘拟合方法, 建立了马尾松超总体立木生物量线性联立方程模型, 解决了 2 个不兼容的实际问题, 并使模型的预估精度和适用性相对于以往生物量模型有明显提高。

关键词 马尾松; 生物量; 线性方程; 方程组; 超总体模型; 兼容性

中图分类号 S718. 556

资源和环境问题是人类社会面临的严峻挑战。森林作为一种再生资源和重要的环境资源, 不仅可持续提供社会发展所必不可少的林产品, 而且在全球环境保护方面占有独一无二的位置。因而, 在当今森林资源监测中, 基于从全树利用以节约资源和解决环境问题的需要, 将森林生物量作为监测内容已在全球受到普遍关注, 并已将它作为最主要的监测项目之一被国际林业研究机构联合会列入《国际森林资源监测大纲》^[1] 中。由于生物量测定的难度和工作量大, 对于区域性的生物量清查, 可以接受的首选方法是将其作为一个重要清查项目兼容于森林资源连续清查 (CFI) 体系中。就我国情况而言, 首先应建立能与立木材积模型兼容且生物总量和各维量也能兼容的超总体 (即通用或标准) 立木生物量模型, 然后将所建模型引入 CFI 抽样样本以估计样地、层和总体的生物量。本研究的主要目的, 在于为建立具有上述 2 个兼容性和广泛适用性的超总体立木生物量模型提供一套科学方法, 它对于推动我国森林生物量清查和监测有重要意义。由于树根生物量取样的难度大, 又受经费制约, 研究仅限于立木地上部分生物量。

1 材料和数据预处理

1.1 研究资料概况

收稿日期: 1998-04-06

第1作者简介: 贺东北, 男, 1962年生, 高级工程师

马尾松(*Pinus massoniana*)样本资料取自江西省德兴市, 按立地质量等级好、中2级和龄组幼、中、成双因素控制设置典型标准地6块, 在样地中分径阶选择生物量样本共50株。样本胸径(D)、树高(H)、冠幅(C_W)和冠长(C_L)的分布范围为: 4.8~47.6 cm, 7.23~24.70 m, 1.20~9.20 m 和 3.90~15.20 m。

样木材积以区分求积法测定。树高 <10 m 为 1 m 区分, >10 m 为 2 m 区分。采取全称量法野外实测干、枝、叶、果鲜质量, 树干分上、中、下截取烘干试样(圆盘), 枝、叶、果用标准取样法采集烘干样品, 将样品放入烘箱, 先在 105 °C 恒温下烘 2 h, 再在 85 °C 恒温下烘 5 h 进行第 1 次称量, 然后每隔 2 h 称量 1 次, 直至 2 次质量相对误差 $\leq 1.0\%$ 时, 将样品取出放入玻璃干燥器内冷却至室温再称其干量, 作为研究的基础数据。

1.2 数据预处理

以样木为单位, 根据样木各维量样品的鲜质量和干质量数据计算出相应的含水率, 进而推算出样木的木材、树皮、树枝和树叶(含果实, 因果实数量很少)的绝干质量, 各维量绝干质量之和为样木地上部分总生物量。

样品含水率的分析表明, 样木树干含水率存在由下而上逐渐增加的明显规律。因此, 树干含水率是将树干 3 等分后通过各段材积和相应的含水率加权求得的, 再据此推算出树干的绝干质量。

2 兼容性超总体立木生物量模型的定义和性质

兼容性是指所建立木生物量模型能与我国森林蓄积量清查的立木材积估计相兼容, 而且模型对立木生物量总量和各维量之间的预估值不产生矛盾; 所谓超总体模型, 是指可适用于一定区域范围内任一具体总体的模型, 即平常所说的通用性(或标准)模型(例如我国一直采用的标准二元立木材积模型即属于此), 而适用于任一总体的实质, 是指模型在任一具体待估总体中的应用都不能超过预定的误差限。由于不同总体立木构成的大小和年龄范围不同, 因此, 要保证其通用性, 必须控制预估变量的变动尽可能缩小, 并使所建模型随解释变量从小到大具有相同的预估精度。

3 立木生物量模型的建立

3.1 解释变量集的选择

经样本资料分析, D , H , C_W , C_L 和树干(带皮)材积 V 等因子均与生物量显著相关, 其中 V 的选入, 不仅可解决与立木材积估计的兼容问题, 而且可明显提高模型的精度。研究还表明, 在上述变量引入模型后, 年龄和立地质量对模型精度的提高已无明显作用。

3.2 基本模型的选择

为了解决立木生物总量与各维量之间的相互兼容问题, 根据总量与各维量之间的联立依存性, 我们选择了线性联立方程模型^[2] 为基本模型。

3.3 模型设计

3.3.1 联立方程模型的变量定义 ①外生变量: D , H , V , C_W , C_L ; ②内生变量: W_1 (立木总生物量)、 W_2 (树干去皮生物量)、 W_3 (树皮生物量)、 W_4 (树枝生物量)、 W_5 (树叶生物量); ③预定变量, 即方程右边的变量, 它可以是外生变量或内生变量。

3.3.2 联立模型的设计 根据超总体模型的性质, 为保证其通用性, 经充分研究, 总量、木材、树皮、树枝、树叶生物量模型的一般形式为:

$$W_i = C_{oi} + C_1 D^2 + C_2 DH + C_3 C_W^2 + C_4 V + u_i \quad (i = 1, 2, 3, 4, 5) \quad (1)$$

式中: W_i 为立木生物量总量和各维量, $C_{oi} \sim C_{4i}$ 为待估参数。

为解决总量和维量间的兼容问题, 经反复计算分析, 将用以描述立木生物量变量结构的完全方程组确定为:

$$\left\{ \begin{array}{l} W_1 = W_2 + W_3 + W_4 + W_5 \\ W_2 = a_0 + a_1 DH + a_2 V + a_3 W_1 + a_4 W_3 \end{array} \right. \quad (2) \quad (3)$$

$$\left\{ \begin{array}{l} W_3 = b_0 + b_1 D^2 + b_2 DH + b_3 V + b_4 W_2 \\ W_4 = c_0 + c_1 D^2 + c_2 C_W^2 + c_3 V + c_4 W_5 \end{array} \right. \quad (4) \quad (5)$$

$$\left\{ \begin{array}{l} W_5 = d_0 + d_1 C_W^2 + d_2 V + d_3 W_2 + d_4 W_4 \end{array} \right. \quad (6)$$

(2) ~ (6) 式称为兼容性立木生物量结构模型, 其中(2)式为定义方程, (3) ~ (6)式中右边的 $W_1 \sim W_5$ 由(1)式给出, $a_0 \sim a_4$, $b_0 \sim b_4$, $c_0 \sim c_4$, $d_0 \sim d_4$ 为待估参数。

3.4 联立模型的拟合

样本资料分析表明, 立木生物量的方差是随解释变量的增大而增大的。为了消除异方差和内生变量带来的偏差影响, 提高模型的通用性, 我们采用了二段加权最小二乘法^[2]对样本资料进行拟合。其最终结果如下:

$$W_1 = 1.882321 + 0.013875 D^2 - 0.097902 DH + 0.037108 C_W^2 + 671.7816 V \quad (7)$$

$$s = 6.0666 \quad R = 0.99778 \quad R^* = 0.99758$$

$$W_2 = 1.054721 - 0.022772 D^2 - 0.044532 DH - 0.100332 C_W^2 + 502.2074 V \quad (8)$$

$$s = 4.6731 \quad R = 0.99716 \quad R^* = 0.99690$$

$$W_3 = -0.244939 + 0.006119 D^2 + 0.010780 DH + 0.068186 C_W^2 + 9.8798 V \quad (9)$$

$$s = 1.4564 \quad R = 0.96176 \quad R^* = 0.95821$$

$$W_4 = 0.795481 + 0.029345 D^2 - 0.053013 DH + 0.032859 C_W^2 + 123.7982 V \quad (10)$$

$$s = 3.8549 \quad R = 0.98156 \quad R^* = 0.97986$$

$$W_5 = 0.277058 + 0.001183 D^2 - 0.011137 DH + 0.036395 C_W^2 + 35.8962 V \quad (11)$$

$$s = 1.2015 \quad R = 0.95985 \quad R^* = 0.95612$$

以上各式中 R^* 为修正复相关系数, 其定义为:

$$R^* = \sqrt{\frac{1 - \sum (W_i - \bar{W}_i)^2 / (n - T)}{\sum (W_i - \bar{W})^2 / (n - 1)}}$$

式中: W_i 为实际值, \bar{W}_i 为理论值, \bar{W} 为实际值的平均值, n 为样本数, T 为回归参数个数。

3.5 联立模型的检验

3.5.1 模型整体检验 采用总相对误差(RS)、平均系统误差(E)、平均相对误差绝对值(RMA)和预估精度(P)等4个统计指标进行检验。各指标的定义为:

$$RS = \frac{\sum W_i - \bar{W}_i}{\sum W_i} \times 100\% \quad (12)$$

$$E = (\sum \frac{W_i - \bar{W}_i}{W_i}) / n \times 100\% \quad (13)$$

$$RMA = \sum \left| \frac{W_i - \bar{W}_i}{W_i} \right| / n \times 100\% \quad (14)$$

$$P = (1 - \frac{t_a \sqrt{\sum (W_i - \bar{W}_i)^2}}{\bar{W} \cdot \sqrt{n(n-1)}}) \times 100\% \quad (15)$$

(15) 式中 t_a 和 \bar{W} 分别为置信水平(取 95%) 和平均预估值(由模型 $W_i = f(\bar{D}, \bar{H}, \bar{V}, \bar{C}_W)$ 给出)。

模型的检验结果列于表 1。

表 1 模型整体检验结果

Table 1 Test results of the models for total sample

模型号	RS	E	RMA	P
7	0.74	-0.01	5.15	97.96
8	0.54	0.08	5.63	97.75
9	1.27	0.70	14.55	93.50
10	2.00	0.45	16.23	92.90
11	-2.28	0	23.07	91.45

表 1 数据表明, 各模型对样本生物量估计的总相对误差均在 $\pm 3\%$ 以内, 预估精度均在 91% 以上, 其中立木总生物量和树干去皮生物量的估计精度达 97% 以上, 从整体上讲精度是高的。

3.5.2 模型残差的随机性检验 以上 4 个整体检验指标尚不足以识别所建模型是否有偏, 还需要通过模型残差分布是否随机来加以检验, 这对于保证超总体模型的

通用性而言是至关重要的, 因为当待估总体的立木大小范围正处于模型存在较大系统误差的区段时, 其误差必然传递到估计结果中去。为此, 我们采用残差随 D, H, V, C_W 和 C_L 而变化的残差分布图的方法³ 对其进行了检验(图略)。图示表明, 所建总量、木材、树皮、树枝和树叶 5 个生物量模型的残差均为随机分布, 不存在有明显偏差的区段。

3.5.3 模型分段检验 为了进一步用数量指标来验证所建模型对解释变量变化范围的全面切合性能, 将建模样本资料按胸径从小到大排序, 等分成 5 个区段, 每段 10 株样木, 进行全范围的分段精度检验, 结果如表 2 所示。

从表 2 可以看出, 各段立木总生物量的相对误差绝大多数在 $\pm 1\%$ 左右, 最大值(第 2 段)也在 $\pm 3\%$ 以内; 各段立木总生物量的预估精度, 绝大多数在 91% 以上, 平均精度为 92.2%, 其变化范围为 88.36% (第 2 段) ~ 94.28% (第 4 段)。树干生物量的分段检验结果与立木总生物量相近似。其他各维量也得到了有效控制, 树皮、树枝和树叶的生物量虽然本身因生物量少而变动较大, 但其分段估计的总相对误差, 绝大多数在 $\pm 5\%$ 以内, 最大也只有 9.45%。树皮分段平均估计精度和变化范围分别为: 76.96% 和 72.07% ~ 85.23%; 树枝为 76.99% 和 71.59% ~ 84.97%; 树叶为 70.59% 和 60.59% ~ 79.93%。这一检验结果表明, 模型具有较好的全面切合性能。

表2 分段检验结果

Table 2 Test results of the models for sub-samples

分段	模型号	RS	E	RMA	P
第1段	7	0.48	0.03	5.88	91.28
	8	-0.18	0.75	7.20	89.58
	9	-2.16	-1.37	11.67	74.40
	10	4.99	3.44	21.56	72.42
	11	2.26	0.20	21.76	67.31
	7	-2.90	-1.89	6.75	88.36
	8	-4.48	-3.47	5.94	88.69
	9	-0.75	-1.98	17.87	72.92
	10	-1.39	1.53	19.47	71.59
	11	9.45	7.15	38.90	60.59
	7	0.87	0.86	4.11	94.14
第2段	8	1.17	0.74	6.11	90.23
	9	1.86	3.70	15.92	80.20
	10	-0.57	-0.34	11.89	84.97
	11	-0.03	1.34	21.52	74.99
第3段	7	0.31	-0.19	4.71	94.28
	8	2.02	2.05	4.26	95.15
	9	1.73	1.38	11.36	85.23
	10	-5.18	-7.46	15.56	77.56
第4段	11	-3.70	-6.15	14.91	79.93
	7	1.43	1.25	4.22	92.95
	8	-0.02	0.34	4.53	91.85
	9	1.48	1.88	16.09	72.07
第5段	10	7.54	5.57	12.26	78.40
	11	-3.60	-2.80	17.73	70.15

4 分析与结论

4.1 根据立木总生物量和各维量间的相互依存性, 通过线性联立模型的结构设计, 解决了以往总量和各维量均采用独立建模所导致的不兼容问题。

4.2 以立木树干材积作为变量引入所建模型, 解决了生物量估计与立木材积估计的兼容问题。因为立木材积是我国现行 CFI 体系和二类调查所必测的主要因子, 只要将材积估计值代入生物量模型即可得出相应的生物量估计值, 这就为样地(林分)和区域性森林生物量作为一个估计项目列入现行各类森林蓄积量清查体系中提供了可靠依据。

4.3 所建模型中的解释变量除 D , H 及其组合变量外, 增加了 C_W , C_L , V , 使模型的预估精度相对于传统的建立在以 D , H 及其组合为解释变量的二元相对生长方程立木生物量模型有明显提高[按常规模型 $W_1 = C_0(D^2H)^{c_1}$ 的拟合结果为 $s = 16.6255$, $R^* = 0.98166$, 其精度明显低于模型(7)]。

4.4 研究表明, 立木生物量随解释变量的变化存在异方差特性。在模型的拟合中, 经采用二段加权最小二乘法和最佳权函数的选定, 使异方差的影响得到了有效消除, 从而保证了模型的无偏性和适用性。值得着重指出的是, 林业数表模型所研究的问题, 绝大多数存在异方

差性而不满足普遍最小二乘法拟合模型对方差齐性的基本要求, 若不采用消除异方差的有效方法而直接采用普通最小二乘法进行拟合, 必然导致模型有偏, 应予以高度重视。

4.5 由于 C_W 和 C_L 在现行森林蓄积量清查体系中不是必测因子, 研究中对含与不含 C_W 和 C_L 因子分 2 种方案进行了对比研究。结果表明, 不带 C_W 和 C_L 建立的模型其预估精度略低, 其优点是应用时无需野外实测 C_W 和 C_L 。考虑到森林作为环境资源的重要性越来越突出, C_W 和 C_L 必将成为未来森林资源清查和监测的必测因子。从国际发展趋势看, 欧洲一些林业发达国家早已将树冠作为重要的测定因子。因此, 我们建议在建模时可同时建立 2 套模型以适应我国目前和未来发展的需要。对于含树冠因子的模型, 为了尽可能减少野外实测树冠因子的工作量, 可同时分别建立立木相对冠长和相对冠幅模型, 只需实测平均木的冠长和冠幅, 便可估计出各立木的相应冠幅与冠长估计值。这一方法我们已同时进行了研究, 效果很好。

4.6 本研究所提供的立木生物量模型的建模原理和方法具有通用性, 可适用于其他树种, 但结构模型的具体设计, 应根据各自样本资料反映的具体规律进行。

参 考 文 献

- 1 IUFRO S4. 02—05. *International guidelines for forest monitoring*. Vienna: [s. n.], 1994. 15~16
- 2 曹焕勋, 娄彦博. 经济计量学. 长沙: 湖南大学出版社, 1986. 222~227
- 3 骆期邦, 宁辉, 贺东北等. 二元立木材积动态模型研究. 林业科学研究, 1992, 5 (3): 263~270

He Dongbei (South-central Forest Inventory and Planning Institute, Changsha 410014, Hunan, PRC), Luo Qibang, and Zeng Weisheng. **Study on linear simultaneous model of tree biomass.** *Journal of Zhejiang Forestry College*, 1998, 15 (3): 298~303

Abstract: Biomass is one of the important items for today's regional forest resources inventory and monitoring, but the current biomass models have two big problems, that is, biomass estimator is not compatible with volume estimator, and total biomass estimators is not compatible with stem, branch, leaf biomass estimators. A new super-population tree biomass model for *Pinus massoniana* was presented based on linear simultaneous equation system and 2-stage weighting least squares. The model could solve the two problems above and have higher precision and better applicability than the currently-used biomass models.

Key words: *Pinus massoniana*; biomass; linear equation; simultaneous equations; compatibility; super-population model