

基于混合效应的湖南栎类与杉木单木断面积生长模型研究^{*}

钟栋明 黄华蓉 李临兵 徐誉远

(广东省岭南综合勘察设计院, 广东 广州 510663)

摘要 构建栎类 (*Quercus* sp.) 杉木 (*Cunninghamia lanceolata*) 混交林单木断面积生长模型, 为该类型林分的生产经营提供参考。以湖南省 31 个样地共 1 646 株栎类与 1 154 株杉木为研究数据, 7 个常见模型作为栎类、杉木的单木基础模型, 基于样地随机效应构建栎类、杉木的单木断面积混合效应模型。结果表明, 栎类断面积的最优基础模型为 Gompertz, 杉木断面积的最优基础模型为 Logistic; 栎类断面积混合效应模型相比于基础模型, 决定系数 R^2 提高了 0.053, SSE 与 RRMSE 均有所降低; 杉木断面积混合效应模型相比于基础模型, 决定系数 R^2 增加了 0.035, SSE 与 RRMSE 均有所降低; 栎类断面积的成熟年龄在 50 年以后, 杉木断面积的生长速率在 30 年左右逐步放缓。文章在栎类、杉木单木模型的基础上进一步加入样地随机效应, 模型拟合精度明显提高, 说明栎类、杉木的断面积生长都受到年龄与样地的综合影响。

关键词 混合效应模型; 断面积; 年龄; 栎类

中图分类号: S757 文献标志码: A 文章编号: 2096-2053 (2020) 04-0018-06

Study on Growth Model of Individual Tree Basal Area of *Quercus* sp. and *Cunninghamia lanceolata* Forest Based on Mixed Effect in Hunan Province

ZHONG Dongming HUANG Huarong LI Linbin XU Yuyuan

(Lingnan Integrated Exploration and Design Institute of Guangdong Province, Guangzhou, Guangdong 510663, China)

Abstract The growth model of basal area of individual tree in mixed forest of *Quercus* sp. and *Cunninghamia lanceolata* was constructed to provide reference for the production and management of this type of forest. Based on the data of 1 646 *Quercus* sp. and 1 154 *C. lanceolata* in 31 plots in Hunan province, seven common models were used as the individual tree foundation models of *Quercus* sp. and *C. lanceolata*, and the mixed effect models of single wood basal area of *Quercus* sp. and *C. lanceolata* were constructed based on the sample random effect. The results showed that the optimal basic model of basal area of *Quercus* sp. was Gompertz, and the optimal basic model of basal area of *C. lanceolata* was Logistic. Compared with the basic model, the determination coefficient R^2 of basal area mixed effect model of *Quercus* sp. increased by 0.053, while SSE and RRMSE decreased. Compared with the basic model, the determination coefficient R^2 of basal area mixed effect model of *C. lanceolata* increased by 0.035, while SSE and RRMSE decreased. The mature age of *Quercus* sp. basal area was above 50 years, and growth rate of *C. lanceolata* slowed down gradually in about 30 years. After further considering the random effect of sample plots, the fitting accuracy of *Quercus* sp. and *C. lanceolata* individual models were significantly improved, it indicated that the basal area growth of *Quercus* sp. and *C. lanceolata* was influenced by age and plot.

Key words mixed effects model; basal area; age; *Quercus* sp.

^{*} 第一作者: 钟栋明 (1993—), 男, 助理工程师, 主要从事林业相关研究, E-mail: 125684179@qq.com。

通信作者: 黄华蓉 (1978—), 女, 高级工程师, 主要从事林业相关工作, E-mail: 48091984@qq.com。

把握林木生长发育规律,对于指导树木的生产经营与利用具有重要作用,其生长规律既有时间积累下来的定性经验,又可用生长模型等方法来定量测定。目前,林分的生长模型可分为全林分模型、径阶分布模型和单木生长模型^[1],前两种是反映林分或径阶的平均变化状态,着重于总体趋势的变化。而单木生长模型^[2-3]能直观反映年龄、立地和竞争等因子对单一树木的生长影响,预估精度高的单木模型能有效地预估树木的生长动态,为林分质量的精准提升提供依据,因此单木生长模型一直是林业科研工作者的研究重点。

在传统建模中,往往只考虑单一因素对树木生长的影响,如胸径对树高^[4]、胸径对冠幅^[5]的影响,这样的单木预估模型拟合精度一般,对树木的生长反映不够灵敏。树木的生长受自身遗传因素的影响较大,但林木竞争、立地条件、气候等因子的影响也不可忽略。因此,一些学者尝试引入混合效应来提高生长模型预估精度,同时综合反映不同因子对树木生长的影响。如符利勇等^[6]考虑林场与样地效应构建的蒙古栎(*Quercus mongolica*)单木断面积模型拟合精度明显优于传统模型;李春明等^[7]基于样地效应构建落叶松(*Larix gmelinii*)林分断面积模型,其拟合精度明显高于传统的非线性回归方法。目前,混合效应模型已应用于杉木、落叶松、马尾松(*Pinus massoniana*)等速生针叶树种,而关于栎类次生林的研究较少。

栎类是我国常绿阔叶林的主要建群树种之一,是我国分布最广、面积最大的重要森林资源。但目前很多栎类林都是经砍伐后萌生的次生林,普遍存在林分过密、萌生、细长杆材等质量低下问题^[8]。湖南的栎类资源丰富,总面积占全省阔叶林面积的13%,栎类树种达到6属77种,主要树种为青冈栎(*Cyclobalanopsis glauca*)、石栎(*Lithocarpus glaber*)、甜槠(*Castanopsis eyrei*)等,其中栎类的主要混交类型包括有栎类-杉木、栎类-马尾松、栎类-阔叶树混交林,栎类-杉木混交林类型广泛分布于全省岳阳、怀化、郴州等14个地级市和自治州,总面积占据栎类资源的18%以上,是湖南省一种主要的森林类型^[9]。考虑到混交树种的复杂性以及栎类杉木混交林的重要性,本研究以栎类和杉木占主要优势的栎类杉木混交次生林为主要研究对象,基于样地随机效应分别构建栎类与杉木的断面积混合效应模型,

为进一步把握栎类与杉木的生长规律,提高栎类资源的林分质量提供参考。

1 研究区概况

湖南省位于长江中下流,洞庭湖以南,地理坐标为108°47'~114°15'E, 24°38'~30°08'N。境内海拔分布范围24~2 122 m,地形地貌复杂,包括有丘陵、平原、山地等,全年春夏湿润多雨,秋冬少雨干旱,为典型的大陆性亚热带季风湿润气候,年平均气温在15~18℃,年均降雨量1 200~1 700 mm。境内的土壤以红壤、黄壤为主,少量分布有石灰土、潮土等。湖南省位于中亚热带常绿阔叶林地带,境内植被资源丰富,主要分布树种有杉木、马尾松、枫香(*Liquidambar formosana*)、樟树(*Cinnamomum camphora*)、青冈栎等。

2 数据来源与方法

2.1 数据来源

研究数据来源于湖南省第八次森林资源连续清查数据库,从中筛选以栎类和杉木为主要组成树种的样地共31块,其中岳阳市7块、怀化市6块、永州市6块、郴州市5块、益阳市4块、湘西自治州3块。为确保研究样本的一致性与代表性,样地筛选时按照以下标准进行:样地郁闭度在0.7以上;每公顷株树在1 000株以上;样地内栎类与杉木的总株数占样地总数的70%以上;样地受人为干扰程度较少。研究样地为25.8 m×25.8 m的正方形,主要调查计算因子包括每木胸径、年龄、断面积和树种等林木因子,海拔、坡度、坡向、坡位、土壤、郁闭度等林地因子。林木年龄由龙时胜等^[10]方法计算而来,其中,在6块样地中分别选取20株树木,利用生长锥钻取树木木芯样本,通过计算年轮的方法确定了树木的实际年龄,结果表明实测年龄与该方法的误差控制在20%以内。

经筛选后确定的样本数为栎类1 646株,杉木1 154株,样地、栎类与杉木的基本概况见表1。为确保所建立的模型科学合理,研究选取总样本的2/3(1 867株)作为建模数据,1/3(933株)作为验证数据。

2.2 研究方法

研究拟构建以样地为随机效应的栎类、杉木断面积生长混合效应模型,其中断面积为因变量,

表 1 样地与样木基本概况
Table 1 The basic situation of sample and wood

因子类型 Factor type	因子变量 Variable	样本量 Sample size	最小值 Min	最大值 Max	平均值 Mean	标准差 Standard error
样地	海拔 /m	31	60	1 170	495	315
	坡度	31	12°	46°	31°	8°
	年龄 / a	1 646	11	159	26	12
栎类	直径 /cm	1 646	5.0	59.0	9.4	2.9
	断面积 /m ²	1 646	0.002 0	0.273 3	0.008 4	0.003 1
	年龄 / a	1 154	5	32	10	3
杉木	直径 /cm	1 154	5	39.8	10.9	3.1
	断面积 /m ²	1 154	0.002 0	0.124 3	0.011 0	0.004 2

表 2 基础模型表达式
Table 2 The expression of base model

序号 Serial number	模型名称 Name of model	模型表达式 Expression of model
1	线性	$Y = \varphi_1 \cdot A + \varphi_2$
2	指数	$Y = \varphi_1 \cdot A^{\varphi_2}$
3	二次函数	$Y = \varphi_1 \cdot A^2 + \varphi_2 \cdot A + \varphi_3$
4	Schumacher	$Y = \varphi_1 \cdot e^{\frac{-\varphi_2}{A}}$
5	Richards	$Y = \varphi_1 \cdot (1 - e^{-\varphi_2 \cdot A})^{\varphi_3}$
6	Gompertz	$Y = \varphi_1 \cdot e^{-\varphi_2 \cdot e^{-\varphi_3 \cdot A}}$
7	Logistic	$Y = \frac{\varphi_1}{1 + \varphi_2 \cdot e^{-\varphi_3 \cdot A}}$

注：Y 为断面积；A 为林木年龄； φ_1 、 φ_2 、 φ_3 为模型参数。
Note: Y is the basal area; A is the age of trees; φ_1 、 φ_2 、 φ_3 are the model parameters.

林木年龄为自变量。混合效应模型的具体构建与评价步骤如下。

2.2.1 基础模型的筛选 相关研究表明，理论生长模型在描述林木的生长过程时往往优于经验模型，为进一步验证该理论，研究考虑选取 3 个经验模型（线性、指数和二次函数）和 4 个生长模型（Schumacher、Richards、Gompertz、Logistic）作为模拟断面积生长的基础模型，基础模型的表达式如表 2。

2.2.2 混合效应模型的构建 混合效应模型是依据回归函数依赖于固定效应参数和随机效应参数的线性或非线性关系而建立的^[11]。考虑到不同样地之间的树木生长可能存在差异，本研究以林木年龄为固定效应，样地为随机效应，拟构建栎类与杉木的单木生长混合效应模型，其表达式为：

$$\begin{cases} Y_{ij} = f(\phi_{ij}, A_{ij}) + \varepsilon_{ij} \\ i = 1, \dots, m; j = 1, \dots, n_i \\ \phi_{ij} = a_{ij}\beta + b_{ij}u_i \\ u_i \sim N(0, \psi) \end{cases}$$

式中， Y_{ij} 为第 i 个样地第 j 株林木的断面积； A_{ij} 为第 i 个样地第 j 株林木的年龄； m 为样地个数； n_i 为第 i 个样地的林木株数； f 是关于参数向量 ϕ_{ij} 的函数； β 为 $p \times 1$ 维固定效应参数； u_i 为 $q_1 \times 1$ 维的样地随机效应参数，假定服从期望为 0，方差为 ψ 的正态分布； ϕ_{ij} 为形式参数，它与 β 、 u_i 呈线性函数关系； a_{ij} 、 b_{ij} 分别为 β 、 u_i 的设计矩阵； ε_{ij} 为随机误差项，假定 u_i 、 ε_{ij} 之间相互独立。

混合效应模型的构建关键在于形式参数的构建，形式参数如何构造争议较大，目前还没有一

个统一的标准。本研究选用一种常用的方法,即在保证模型收敛的情况下,考虑基础模型中所有形式参数都为混合参数,利用评价指标-赤池信息准则(Akaike Information Criterion, AIC)和贝叶斯信息准则(Bayesian Information Criterion, BIC)对所有形式参数组合类型进行比较。AIC与BIC两者的原理有所差异,AIC是从预测角度选择一个好的模型用来预测,BIC是从拟合角度选择一个对现有数据拟合最好的模型,但两者都能有效避免模型精度过高而造成模型过度复杂的现象,模型评价过程中AIC和BIC越小越好^[12],两个评价指标的表达式分别为:

$$AIC = -2\ln(L) + 2k \quad \dots\dots\dots(1)$$

$$BIC = -2\ln(L) + \ln(n) \times k \quad \dots\dots\dots(2)$$

式中, k 为参数个数; n 为样本数; L 为模型的似然函数。

2.2.3 模型评价与检验 模型的评价指标包括决定系数(R^2)、相对均方根误差(RRMSE)、残差平方和(SSE)共3个指标, R^2 越大,RRMSE和SSE越小,模型拟合效果越佳。根据建模数据求出7个模型的参数,依据3个评价指标进行对比,就可筛选出最优的基础模型,然后利用检验数据评价基础模型与混合效应模型的拟合效果。其中,3个评价指标的表达式为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad \dots\dots\dots(3)$$

$$SSE = \sum (y_i - \hat{y}_i)^2 \quad \dots\dots\dots(4)$$

$$RRMSE = 100\% * \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - k}} / \bar{y} \quad \dots\dots\dots(5)$$

式中, n 为样本数; y 为测量值; $\hat{y}$ 为预估值; $\bar{y}$ 为测量值的平均值; k 为参数个数。

3 结果与分析

3.1 基础模型拟合结果

利用7个基础模型分别拟合栎类与杉木的断面积生长模型,结果见表3。除线性模型外,栎类单木断面积生长模型的決定系数均在0.711以上,非线性模型的拟合效果较好,其中Gompertz

模型的拟合效果最佳($R^2=0.845$, $SSE=0.037$, $RRMSE=5.64\%$),可作为栎类单木断面积的最优基础模型;除线性模型外,杉木单木断面积生长模型的決定系数均在0.733以上,非线性模型的拟合效果较好,其中Logistic模型的拟合效果最佳($R^2=0.746$, $SSE=0.032$, $RRMSE=4.79\%$),可作为杉木单木断面积的最优基础模型。

3.2 混合效应模型拟合结果

基于样地水平,构建栎类的断面积混合效应生长模型,针对不同形式参数的组合,利用AIC、BIC两个统计指标对模型的拟合优度进行比较,AIC与BIC值越小,模型拟合效果越好。由表4的结果可知,栎类的断面积生长混合效应模型随机参数位于 φ_1 时,模型的AIC、BIC指标最小。其中,栎类的混合效应生长模型AIC、BIC均明显小于基础模型的AIC、BIC,说明基于样地水平的栎类混合效应模型拟合效果优于基础模型。

基于样地水平,构建杉木的断面积混合效应生长模型。由表5可知,杉木的断面积生长混合效应模型随机参数位于 φ_1 时,模型的AIC、BIC指标最小。其中,杉木的混合效应生长模型AIC、BIC均明显小于基础模型的AIC、BIC,说明基于样地水平的杉木混合效应模型拟合效果优于基础模型。

3.3 栎类与杉木的生长规律

对比分析栎类与杉木的断面积生长规律,结果如图1。栎类的断面积生长过程较为缓慢,且随着年龄的增大呈增加的趋势,说明栎类的成熟年龄在50年以后;杉木的断面积生长速率明显高于栎类,且在30年左右生长速率逐步放缓,之后生长趋于稳定。

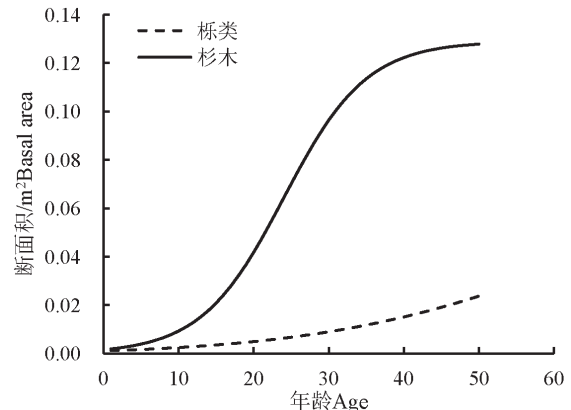


图1 栎类与杉木的生长规律

Fig.1 The growth law of *Quercus* sp. and *Cunninghamia lanceolata*

表 3 栎类与杉木断面积生长模型拟合结果
Table 3 The fitting results of basal area model of *Quercus* sp. and *Cunninghamia lanceolata*

树种 Species	模型 序号 Model number	模型参数 Model parameter			评价指标 Evaluation indicator		
		φ_1	φ_2	φ_3	R^2	SSE	RRMSE/%
栎类	1	0.001	-0.012		0.625	0.089	8.75
	2	9.3×10^{-6}	2.006		0.838	0.038	5.72
	3	1.1×10^{-5}	-0.0001	0.003	0.841	0.038	5.72
	4	0.424	136.149		0.711	0.068	7.65
	5	0.336	0.011	2.997	0.802	0.047	6.36
	6	0.478	6.044	0.014	0.845	0.037	5.64
	7	0.027	0.039	0.000	0.823	0.042	6.01
杉木	1	0.003	-0.015		0.686	0.04	5.35
	2	7.9×10^{-5}	2.076		0.742	0.033	4.86
	3	0.000 1	-0.000 2	0.001	0.741	0.033	4.86
	4	0.175	29.568		0.692	0.039	5.28
	5	0.286	0.034	2.725	0.733	0.034	4.93
	6	0.285	5.999	0.056	0.736	0.034	4.93
	7	0.129	79.405	0.182	0.746	0.032	4.79

表 4 栎类的混合效应模型拟合结果
Table 4 The fitting results of mixed effect model of *Quercus* sp.

模型 Model	随机参数 Random parameter	评价指标 Evaluation indicator		
		AIC	BIC	LL
基础模型	None	-11 682.7	-11 648.2	5 848.233
混合效应模型	φ_1	-15 812.9	-15 778.5	7 913.362

表 5 杉木的混合效应模型拟合结果
Table 5 The fitting results of mixed effect model of *Cunninghamia lanceolata*

模型 Model	随机参数 Random parameter	评价指标 Evaluation indicator		
		AIC	BIC	LL
基础模型	None	-5 797.26	-5 767.81	2 905.531
混合效应模型	φ_1	-6 720.09	-6 690.63	3 366.943

表 6 基础模型与混合效应模型的拟合效果对比
Table 6 Comparison of the fitting results between basic and mixed effect model

模型 Model	栎类 <i>Quercus</i> sp.			杉木 <i>Cunninghamia lanceolata</i>		
	R^2	SSE	RRMSE/%	R^2	SSE	RRMSE/%
基础模型	0.842	5.58	0.747	0.011	4.82	5.58
混合效应模型	0.895	4.72	0.813	0.009	3.94	4.72

3.4 模型检验与评价

基于检验样本对栎类与杉木的最优基础模型与混合效应模型拟合效果进行检验。由表 6 可知，栎类断面积生长混合效应模型的决定系数 R^2 比基础模型增加 0.053，残差平方和 SSE 与相对均方根

误差 RRMSE 均有所降低，检验数据的结果进一步表明栎类的混合效应模型拟合效果优于基础模型；杉木断面积生长混合效应模型的决定系数 R^2 比基础模型增加 0.035，残差平方和 SSE 与相对均方根误差 RRMSE 均有所降低，检验数据的结果

进一步表明杉木的混合效应模型拟合效果优于基础模型。

4 结论与讨论

莫辉等^[13]在构建木荷的直径、树高与材积生长模型时发现,与经验模型相比,理论生长模型在模型拟合精度及生物学解释上均具有一定优势;吕勇^[14]在研究杉木与马尾松树高生长模型时也具有同样的结论。本研究在构建栎类与杉木的断面积生长模型时发现,栎类的最优基础模型为 Gompertz 理论生长模型,杉木的最优基础模型为 Logistic 模型,理论生长模型的拟合精度总体高于经验模型,说明理论生长模型在拟合林木的生长过程时能比较准确的反映出树木随年龄的累积变化规律,且在生物学解释具有一定的优势。

树木的生长变化是林木遗传特性、林木竞争以及立地条件综合影响的结果,单一方面的指标无法很全面的反映树木的生长过程。混合效应模型在考虑林木的平均变化趋势时,还能体现出个体或立地对林木生长的影响,在构建树木的生长模型方面具有优势。部分研究表明,模型中加入样地随机效应能有效的提高模型的拟合精度,如李春明^[15]在构建杉木蓄积联立方程的基础上引入样地混合效应,结果显示,混合效应模型的拟合效果明显优于基础模型;刘洵等^[16]在栎林生长收获模型中加入样地层次的混合效应后,模型精度得到很大改善。样地随机效应是对立地因子差异性的反映,不同样地的海拔、坡度、坡向、土壤等往往存在差异,而这种差异性就会导致树木的生长规律表现不一。本研究在栎类与杉木的断面积生长模型中加入样地随机效应后,模型拟合精度得到提高,说明栎类与杉木的生长不仅受固定效应年龄的影响,在不同样地之间也会表现出差异性。

异龄混交林中单株树木年龄的测定一直是林业工作中的重点与难点,目前还没有一种准确性高、操作成本低廉的有效测定方法。龙时胜等^[10]采用林木多期直径数据估计异龄林的年龄,能确保单木的平均绝对百分误差控制在 14.8% 以内,对于树木年龄的估计具有一定的参考价值。本研究参考此方法,利用生长锥选取了部分林木进行了年龄的实测,结果表明实测年龄与该方法的相对误差控制在 20% 以内。本文以年龄为自变量,

拟合栎类与杉木的断面积平均生长曲线,反映的是树木平均年龄与断面积的生长过程,而估计年龄是在实测年龄的上下波动,20% 的误差对平均年龄的影响不大,可以认为估计年龄对本文的研究结果影响较小,可以应用于断面积的生长建模。

参考文献

- [1] 李安明, 郭小龙, 张岗岗, 等. 小陇山林区日本落叶松人工林全林整体模型的建立与验证[J]. 西北林学院学报, 2016, 31(6): 26-29; 40.
- [2] 萨如拉, 徐加睿, 王智慧, 等. 基于树干解析的兴安落叶松人工林单木生长模型研究[J]. 林业资源管理, 2019(2): 88-92.
- [3] 林丽平, 徐期瑚, 罗勇, 等. 广东主要乡土阔叶树种单木生长模型构建[J]. 林业与环境科学, 2018, 34(3): 14-22.
- [4] 胥辉, 全宏波, 王斌. 思茅松标准树高曲线的研究[J]. 西南林学院学报, 2000(2): 74-77.
- [5] 张冬燕, 王冬至, 范冬冬, 等. 不同立地类型华北落叶松人工林冠幅与胸径关系研究[J]. 林业资源管理, 2019(4): 69-73.
- [6] 符利勇, 唐守正, 张会儒, 等. 基于多水平非线性混合效应蒙古栎林单木断面积模型[J]. 林业科学研究, 2015, 28(1): 23-31.
- [7] 李春明, 唐守正. 基于非线性混合模型的落叶松云冷杉林分断面积模型[J]. 林业科学, 2010, 46(7): 106-113.
- [8] 刘洵. 湖南栎类次生林生长收获模型研究[D]. 长沙: 中南林业科技大学, 2019.
- [9] 罗小浪. 湖南栎类天然林立地质量评价研究[D]. 长沙: 中南林业科技大学, 2018.
- [10] 龙时胜, 曾思齐, 甘世书, 等. 基于林木多期直径测定数据的异龄林年龄估计方法[J]. 中南林业科技大学学报, 2018, 38(9): 1-8.
- [11] 符利勇, 孙华. 基于混合效应模型的杉木单木冠幅预测模型[J]. 林业科学, 2013, 49(8): 65-74.
- [12] 段光爽, 李学东, 冯岩, 等. 基于广义非线性混合效应的华北落叶松天然次生林枝下高模型[J]. 南京林业大学学报(自然科学版), 2018, 42(2): 170-176.
- [13] 莫辉, 赵建强. 木荷单木生长模型构建及生长特征分析[J]. 林业与环境学, 2019, 35(2): 55-61.
- [14] 吕勇. 林木树高曲线模型研究[J]. 中南林学院学报, 1997(4): 86-89.
- [15] 李春明. 基于混合效应模型的杉木人工林蓄积联立方程系统[J]. 林业科学, 2012, 48(6): 80-88.
- [16] 刘洵, 曾思齐, 贺东北, 等. 基于混合效应的湖南省栎林生长收获模型[J]. 森林与环境学报, 2019, 39(5): 475-482.