

网络出版时间:2017-05-10 13:36

DOI:10.13207/j.cnki.jnwafu.2017.06.008

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1390.S.20170510.1336.016.html>

# 基于线性混合效应模型的杉木树高-胸径模型

许崇华, 崔 珺, 黄兴召, 余 鑫, 徐小牛

(安徽农业大学 林学与园林学院, 安徽 合肥 230036)

**【摘要】**【目的】利用线性混合效应模型分析杉木树高与胸径的关系,为杉木树高测量提供支持。【方法】收集 688 组有效杉木研究数据,利用最小二乘法构建树高( $H$ )和胸径( $DBH$ )的线性基础模型,同时考虑林分密度效应和海拔效应,在 R 3.2.2 软件中拟合混合模型,比较基础模型和 2 种混合模型的赤池信息规则( $AIC$ )、贝叶斯信息规则( $BIC$ )和 $-2$  倍对数似然值( $-2\log lik$ ),在此基础上,引入误差效应方差协方差矩阵及指数函数、幂函数和恒等式函数,筛选较好的混合模型;基于独立验证数据对模型进行验证,选取  $R^2$ 、 $|\bar{E}|$ 、 $RMSE$  3 个评价指标对模型精度进行评价。【结果】固定模型的  $AIC=2\ 089.731$ ,  $BIC=2\ 102.151$ ,  $-2\log lik=2\ 083.732$ ,均大于混合模型,即混合效应模型拟合效果优于固定模型;考虑模型误差效应方差协方差矩阵,加入恒等式异方差函数能够显著提高模型的精度,且含有不同随机参数的混合模型精度不同,引入海拔随机效应的混合模型拟合精度( $R^2=0.804\ 4$ ,  $|\bar{E}|=1.553\ 9$ ,  $RMSE=2.143\ 0$ )高于含有林分密度效应的混合模型( $R^2=0.797\ 0$ ,  $|\bar{E}|=1.576\ 6$ ,  $RMSE=2.183\ 0$ )。【结论】考虑随机效应的混合模型既能反映杉木树高的总体变化趋势,还能体现不同组分间的差异,在估测精度和通用性上均优于固定模型。

**【关键词】** 杉木;树高-胸径方程;线性混合模型;随机效应

**【中图分类号】** S711

**【文献标志码】** A

**【文章编号】** 1671-9387(2017)06-0053-08

## Height-diameter model for Chinese fir based on linear mixed model

XU Chonghua, CUI Jun, HUANG Xingzhao, YU Xin, XU Xiaoniu

(School of Forestry & Landscape Architecture, Anhui Agricultural University, He fei, Anhui 230036, China)

**Abstract:** 【Objective】 Relationship between height ( $H$ ) and diameter at breast-height ( $DBH$ ) of Chinese fir (*Cunninghamia lanceolata*) was studied using linear mixed model to provide support for height measurement. 【Method】 Data of 688 Chinese fir plants were collected and least square method was used to established the  $H$ - $DBH$  linear model. Considering stand density and altitude effects, the Akaike information criterion ( $AIC$ ), Bayesian information criterion ( $BIC$ ) and the  $-2$  time of logarithm likelihood value ( $-2\log lik$ ) are compared between the basic model and the 2 mixed models. Based on this, the exponential function (Exp), power function (Power), and identity heteroscedasticity function (Ident) were added to select the better mixing models. The accuracy of the model was evaluated using of  $R^2$ ,  $|\bar{E}|$  and  $RMSE$  as indexes in R 3.2.2 software. 【Result】 The Akaike information criterion (2 089.731), Bayesian information criterion (2 102.151), and  $-2$  times of logarithm likelihood value (2 083.732) were higher than the mixed model, indicating that the mixed model was better. Considering the error variance-covariance matrix and

**【收稿日期】** 2016-04-22

**【基金项目】** 国家“973”计划项目(2012CB416905);国家自然科学基金项目(31370626)

**【作者简介】** 许崇华(1990—),男,安徽歙县人,在读硕士,主要从事森林生产力及其固碳研究。E-mail: xch422248787@live.com

**【通信作者】** 徐小牛(1961—),男,安徽庐江人,教授,博士,博士生导师,主要从事森林培育、森林生物地球化学研究。

E-mail: xn Xu2007@ahau.edu.cn

adding identity heteroscedasticity function (Ident) significantly improved the model accuracy. Compared with the density-level mixed model ( $R^2 = 0.7970$ ,  $|\bar{E}| = 1.5766$ ,  $RMSE = 2.1830$ ), the altitude-level mixed model had better precision ( $R^2 = 0.8044$ ,  $|\bar{E}| = 1.5539$ ,  $RMSE = 2.1430$ ). 【Conclusion】 The mixed model with the consideration of random effects was better than fixed model since it can reveal both overall trends and differences between different groups of Chinese fir plantations.

**Key words:** Chinese fir;  $H$ -DBH equation; linear mixed model; random-effects

树高是森林资源调查中不可缺少的调查因子,是描述森林垂直结构、树木生长状况、立地指数以及计算林木蓄积与生物量的重要参数<sup>[1-4]</sup>,但研究中使用的树高数据多是在林地中基于一定的水平距离和角度测量获得的,故其精确度难以得到保障。超声和激光技术在林业上的应用,一定程度上提高了树高测量的准确性,但存在耗时费力的不足。基于树高( $H$ )与胸径(DBH)间存在的相关性,Wiedemann拟合了第一个树高-胸径( $H$ -DBH)方程,自此用  $H$ -DBH 模型预测林木树高的方法逐渐得到推广<sup>[5]</sup>。当前,描述  $H$ -DBH 关系的生长模型有很多,常用于林业研究的有线性模型和非线性模型 2 种。

随着  $H$ -DBH 模型研究的不断深入,发现环境因子对  $H$  与 DBH 的关系有一定的影响。Grubb<sup>[6]</sup>研究热带雨林树木生长时指出,随着海拔梯度的不断变化,树木的  $H$ -DBH 关系也会随之发生改变。林分密度同样也会影响  $H$ -DBH 模型预测树高的准确性,Hummel<sup>[7]</sup>发现,相同树高的同一树种在密度小的区域胸径远大于密度大的区域。此外,温度<sup>[8]</sup>、年均降水<sup>[9-10]</sup>、林龄变化<sup>[11-12]</sup>等也会影响  $H$ -DBH 的关系。传统的固定模型只考虑了样本总体的平均水平却不能体现不同林地间的差异性,而混合模型既可以体现样本总体的平均变化趋势,又能以随机效应参数的方差、协方差等信息来反映样本个体间的差异<sup>[13]</sup>。相对于只拥有固定效应参数的模型,混合模型具有更好的灵活性,因而成为了当前研究的热点。

线性混合效应模型(LME)包含固定效应参数和随机效应参数,且 2 种参数都以线性形式出现在模型中。随机效应可分为截距随机效应与系数随机效应 2 种,许多学者将线性混合效应模型应用于林业研究中。Lynch 等<sup>[14]</sup>将 Korf 函数线性化,利用 561 株黑樱桃树高-胸径数据构建了线性混合效应模型,结果表明加入随机效应的混合效应模型拟合精度远大于只含固定参数的模型;Uzo 等<sup>[15]</sup>在拟合美国黄松树高生长模型时加入了位置与样地的随机效应参数,预测结果显示线性混合效应模型有效提

高了预测精度;李春明<sup>[16]</sup>基于两层次的线性混合效应模型模拟了杉木林的胸径生长模型,提高了胸径生长量的预测精度。

杉木(*Cunninghamia lanceolata*)主要分布于我国南方地区,具有速生、丰产、材质优的特性,是重要的用材树种<sup>[17-18]</sup>。由于缺乏精确定量杉木树高的生长模型,在经营过程中常根据国家森林抚育标准和森林经营经验进行林木的管理,因此亟需研究建立考虑了随机效应的适合于杉木树高测量的混合模型。本研究对文献中各地区杉木树高、胸径等数据进行整理,以线性方程作为基础模型,基于线性混合效应模型构建考虑林分密度与海拔的  $H$ -DBH 模型,以期对杉木森林资源的可持续经营与优化配置提供依据。

## 1 材料与方法

### 1.1 试验数据的收集与整理

从中国知网、万方、维普和 ScienceDirect 等网络数据库中收集以往有关杉木的研究文献,共获得 142 篇 720 组杉木相关研究数据,整个数据集包括了杉木树高(Height,  $H$ )和胸径(Diameter at breast height, DBH)以及杉木林分密度(Density)、海拔(Altitude)、样地区域(Region)等的基本信息。首先,将获取的数据进行筛选,删除重复值、异常值以及胸径小于 5 cm 的数据,并对收集的每组杉木数据进行来源编码,最终获得 688 组杉木数据。整个杉木数据集涵盖了安徽省、湖南省、湖北省、福建省、江西省、广西省、广东省、四川省、江苏省、贵州省等我国南方 10 个省份,其中福建省、湖南省、江西省、广西省、广东省是杉木研究数据来源较多的地区(图 1)。将这些数据集分为 2 部分,选取其中 468 组数据用于拟合  $H$ -DBH 模型,平均胸径、树高分别为 14.55 cm 和 12.73 m,林分密度和海拔分别是 300~10 000 株/hm<sup>2</sup> 和 30~1 450 m;其余 220 组数据作为模型的检验数据,平均胸径、树高分别为 14.68 cm 和 12.36 m,其林分密度和海拔分别是 400~10 000 株/hm<sup>2</sup> 和 30~1 450 m(表 1)。

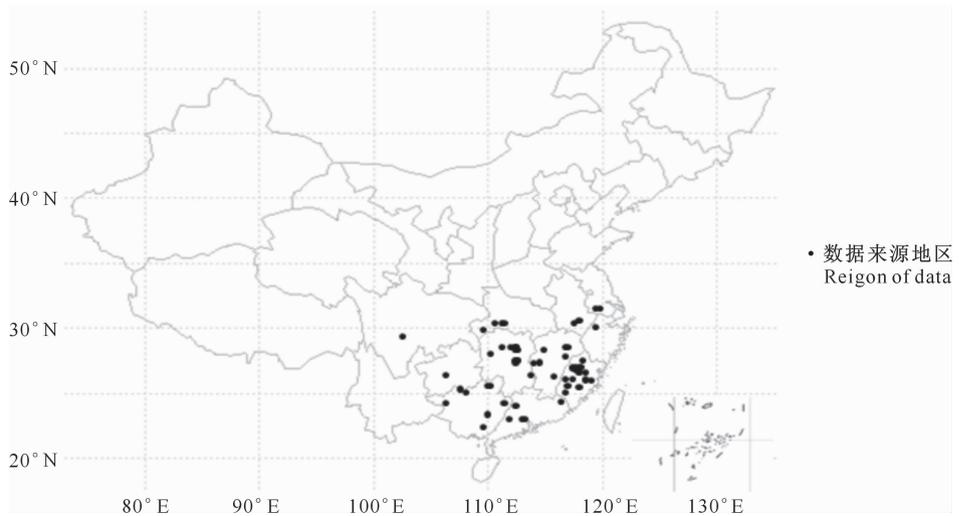


图 1 本研究中数据采集样地的空间分布

Fig. 1 Spatial distribution of sampling sites in this study

表 1 本研究中杉木数据集中的树高与胸径因子统计

Table 1 Statistical results of height and diameter at breast-height

研究数据 Research data	样本数 Number	测量因子 Measure factor	最小值 Min	最大值 Max	平均值 Mean	标准差 SD
建模数据 Fitting data	468	DBH/cm	5.10	36.41	14.55	5.41
		H/m	2.64	29.79	12.73	4.85
检验数据 Validation data	220	DBH/cm	5.50	33.59	14.68	6.04
		H/m	4.00	32.20	12.36	5.66

## 1.2 研究方法

1.2.1 基础模型的确定 根据收集数据绘制的  $H-DBH$  关系散点图见图 2。由图 2 可以看出,  $H-DBH$  呈线性分布, 因而选用固定线性模型作为拟

合  $H-DBH$  关系的基础模型, 其模型具体形式如下:

$$H_i = \beta_1 + \beta_2 DBH_i + \epsilon_i. \quad (1)$$

式中:  $H_i$  为第  $i$  株样木的树高(m),  $DBH_i$  为第  $i$  株样木的胸径(cm),  $\beta_1$  和  $\beta_2$  为待估参数,  $\epsilon$  为误差项。

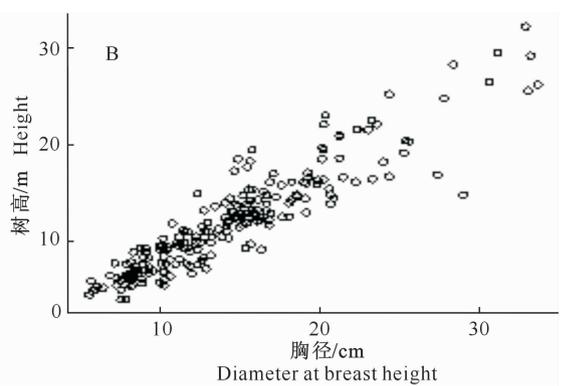
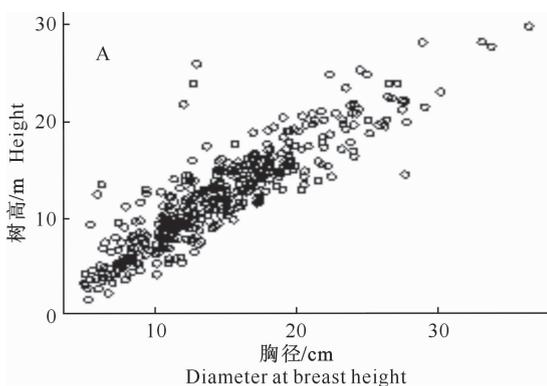


图 2 利用建模数据(A)和验证数据(B)得到的树高-胸径关系

Fig. 2  $H-DBH$  plot of calibration (A) and validation (B) subsets

1.2.2 线性混合模型 根据模型中引入随机效应因子的个数, 将线性混合效应模型分为单水平和多水平 2 种基本形式。其中单水平线性混合效应模型只包含 1 个随机效应因子, 本研究即是基于单水平线性模型构建树高-胸径关系模型, 其一般表达式为:

$$\begin{cases} H_j = DBH_j \beta + Z_j \mu_j + \epsilon_j, \\ j = 1, 2, \dots, M, \\ \mu_j \sim N(0, \mathbf{D}), \\ \epsilon_j \sim N(0, \mathbf{R}_j). \end{cases} \quad (2)$$

式中:  $H_j$ 、 $DBH_j$  分别表示海拔为  $j$  或林分密度为  $j$  时所对应的树高(m)、胸径(cm);  $\beta$  为固定效应参数

向量;  $\mu_j$  为第  $j$  海拔水平或第  $j$  林分密度水平的随机效应向量;  $Z_j$  为与  $\mu_j$  对应的随机向量设计矩阵;  $\epsilon_j$  为误差项, 且与参数  $\mu_j$  相互独立;  $M$  为海拔或林分密度等级;  $N$  表示  $\mu_j$  或  $\epsilon_j$  服从正态分布;  $D$  为海拔或林分密度对应的随机参数的方差协方差结构矩阵;  $R_j$  为第  $j$  海拔水平或第  $j$  林分密度水平误差效应的方差协方差结构矩阵。

1.2.3 模型参数与结构的确定 1) 模型中固定参数与随机参数的确定。两种参数的确定一般依赖于具体的试验数据, 对于拟合模型满足收敛条件但缺乏随机效应方差、协方差结构先验知识的情况, Pinheiro 等<sup>[19]</sup> 提出先将模型中所有参数都看作随机效应参数, 再分别拟合含有不同随机参数的模型, 比较 AIC(赤池信息准则)、BIC(贝叶斯信息准则)、 $-2\log lik$ (对数似然值)<sup>[1-5]</sup> 参数指标, AIC、BIC、 $-2\log lik$  值越小, 则模型效果越好。为避免参数过多, 含有不同参数个数的模型还需进行似然比检验(LRT), 若  $P < 0.05$  则认为差异显著, 选择收敛且精度高的模型进行效果评价。

2) 误差效应方差协方差结构矩阵( $R_j$ )的确定。 $R_j$  的确定必须解决误差方差问题, 林业研究中常用于描述误差效应方差协方差结构的表达式为:

$$R_j = \sigma^2 \phi_j^{0.5} \Gamma_j \phi_j^{0.5}. \quad (3)$$

式中:  $\sigma^2$  为模型的误差方差值,  $\phi_j^{0.5}$  为描述误差方差异质性的对角矩阵,  $\Gamma_j$  为描述误差自相关的结构矩阵。

考虑误差的异方差问题, 本研究以指数函数(Exp)、幂函数(Power)以及恒等式(Ident)3种异方差结构来消除数据间的异方差性。虽然本研究采用的数据集时间跨度较大, 但并非同一样木或样地在时间序列上的复测数据, 因而无须考虑时间序列自相关对混合模型误差相关性的影响。

3) 随机效应方差协方差结构矩阵( $D$ )。该矩阵反映了随机效应间的差异性及  $D$  随随机效应参数个数的变化情况, 以包含 2 个随机参数的随机效应方差协方差结构为例, 其结构可表达为:

$$D = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \sigma_{\mu_1}^2 & \sigma_{\mu_1 \mu_2} \\ \sigma_{\mu_1 \mu_2} & \sigma_{\mu_2}^2 \end{bmatrix}. \quad (4)$$

式中:  $\sigma_{\mu_1}^2$ 、 $\sigma_{\mu_2}^2$  分别为随机效应参数  $\mu_1$ 、 $\mu_2$  的方差;  $\sigma_{\mu_1 \mu_2} = \sigma_{\mu_2 \mu_1}$ , 分别为  $\mu_1$ 、 $\mu_2$  的协方差。

1.2.4 模型检验 拟合模型的检验主要通过树高预测并比较其精度来完成, 需要对随机效应参数值进行估算, 通常采用公式(5)来计算随机参数  $\hat{\mu}_k$ , 即:

$$\hat{\mu}_k = \hat{D} \hat{Z}_k^T (\hat{Z}_k \hat{D} \hat{Z}_k^T + \hat{R}_k)^{-1} \hat{\epsilon}_k. \quad (5)$$

式中:  $\hat{\mu}_k$  为随机参数预估值,  $\hat{D}$  为随机效应方差协方差预估值矩阵,  $\hat{Z}_k$  为随机效应预估值矩阵,  $\hat{Z}_k$  与  $\hat{Z}_k^T$  互为转置矩阵,  $\hat{R}_k$  为误差效应方差协方差预估值矩阵,  $\hat{\epsilon}_k$  为误差向量。

1.2.5 模型评价及指标检验 通过决定系数( $R^2$ )、平均绝对残差( $|\bar{E}|$ )和均方根误差(RMSE)等评价指标来比较拟合和检验结果, 若  $|\bar{E}|$ 、RMSE 值越小,  $R^2$  值越大, 则模型越优。  $|\bar{E}|$ 、RMSE 和  $R^2$  的表达式为:

$$R^2 = 1 - \left[ \frac{\sum_{j=1}^n (y_j - \hat{y}_j)^2}{\sum_{j=1}^n (y_j - \bar{y})^2} \right]. \quad (6)$$

$$|\bar{E}| = \frac{|\sum_{j=1}^n (y_j - \hat{y}_j)|}{n}. \quad (7)$$

$$RMSE = \sqrt{\frac{\sum_{j=1}^n (y_j - \hat{y}_j)^2}{n-1}}. \quad (8)$$

式中:  $y_j$  为实测值,  $\hat{y}_j$  为模型预测值,  $\bar{y}$  为实测值的平均值,  $n$  为实测值个数。

## 2 结果与分析

### 2.1 基础模型的拟合

在建立线性混合效应模型的  $H$ -DBH 方程前, 通过 R 3.2.2 软件对式(1)模型采用最小二乘法进行基础模型拟合, 利用  $t$  检验和  $F$  检验对模型及模型参数进行检验。结果显示, 式(1)模型所有参数的估计值均达到极显著水平( $P < 0.005$ ), 说明参数  $\beta_1$ 、 $\beta_2$  对模型变化影响显著。对方程进行  $F$  检验( $F = 1788.1$ ,  $P < 0.0001$ ), 结果表明, 方程式(1)对树高、胸径关系的描述具有显著意义。因而, 本研究选取式(1)模型( $H = 1.1836 + 0.7939DBH + \epsilon$ )为构建  $H$ -DBH 关系混合模型的基础模型。

### 2.2 混合效应模型的拟合

2.2.1 随机效应的确定 本研究数据来自不同的省份, 其中样本数超过 30 个的省份有 5 个, 以不同省份区域作为随机参数拟合混合模型, 似然比检验结果显示混合模型与基本模型差异不显著, 因而考虑选择其他因子作为模型的随机参数。由于数据集包含的海拔和林分密度范围较大且较为分散, 不可直接用于拟合线性混合方程, 因而需要将林分密度和海拔划分等级后方可用于混合模型的拟合。本研究将林分密度按区间划分为  $\leq 1000$  株/hm<sup>2</sup>,  $> 1000 \sim \leq 1300$  株/hm<sup>2</sup>,  $> 1300 \sim \leq 1600$  株/hm<sup>2</sup>,  $> 1600 \sim \leq 1900$  株/hm<sup>2</sup>,  $> 1900 \sim \leq 2200$  株/hm<sup>2</sup>,  $> 2200 \sim \leq 2500$  株/hm<sup>2</sup>,

$> 2\ 500 \sim \leq 2\ 800$  株/hm<sup>2</sup>,  $> 2\ 800 \sim \leq 3\ 100$  株/hm<sup>2</sup>,  $> 3\ 100 \sim \leq 3\ 600$  株/hm<sup>2</sup> 以及  $> 3\ 600$  株/hm<sup>2</sup> 等 10 个等级;海拔则按区间划分为  $\leq 250$  m,  $> 250 \sim \leq 500$  m,  $> 500 \sim \leq 750$  m 和  $> 750$  m 4 个梯度。

2.2.2 林分密度层次线性混合效应模型的拟合  
不同随机参数组合对模型拟合结果的影响不同,本研究先以林分密度作为随机参数拟合混合模型,选出收敛的模型比较 AIC、BIC、 $-2\log lik$  等 3 个评价指标,3 个指标的数值越小则说明方程拟合的精度越高,拟合效果越好。引入林分密度层次随机因子的线性混合模型拟合结果见表 2。由表 2 可见,

表 2 线性混合效应模型的拟合结果

Table 2 Fitting results of mixed models

模型 Model	随机参数 Random parameter	自由度 df	AIC	BIC	$-2\log lik$	LRT	P
固定模型 Fixed model	无 without	3	2 089.731	2 102.151	2 083.732		
密度效应 Density effect	$\mu_{11}$	4	2 085.704	2 102.264	2 077.704	6.027	0.014 1
	$\mu_{12}$	4	2 083.828	2 100.387	2 075.828	7.903	0.004 9
	$\mu_{11} \cdot \mu_{12}$	6	2 083.411	2 108.250	2 071.410	12.321	0.006 4
海拔效应 Altitude effect	$\mu_{21}$	4	2 080.782	2 097.341	2 072.782	10.949	0.000 9
	$\mu_{22}$	4	2 090.289	2 106.849	2 082.290	1.442	0.229 8
	$\mu_{21} \cdot \mu_{22}$	6	2 063.671	2 088.510	2 051.670	32.060	<0.000 1

考虑误差效应的方差协方差结构,本研究将指数函数(Exp)、幂函数(Power)以及恒等式(Ident)作为异方差函数代入模型中,结果(表 3)表明,含恒等式异方差结构的混合模型的拟合优度(AIC = 2 045.705, BIC = 2 099.523,  $-2\log lik = 2 019.704$ )较不考虑异方差模型有很大提高,且差异极显著( $P < 0.000 1$ );而以指数函数和幂函数作为异方差函数的模型的拟合优度均较差(即 AIC、BIC 和  $-2\log lik$  值均较大)。因此以含异方差函

模型收敛的情况共有 3 种(分别为含随机参数  $\mu_{11}$ ,  $\mu_{12}$  和  $\mu_{11} \cdot \mu_{12}$  的混合模型),其 AIC、 $-2\log lik$  值均小于固定模型,似然比检验(LRT)结果表明:混合模型与固定模型差异显著( $P < 0.05$ ),说明加入林分密度随机参数的混合模型的拟合效果优于最小二乘法拟合的固定模型。比较混合模型拟合结果中 AIC、BIC、 $-2\log lik$  值的大小,结果表明,3 个混合模型的拟合效果差异不大(即 AIC、BIC 和  $-2\log lik$  值相近),而似然比检验(LRT)结果表明:含随机参数  $\mu_{12}$  的模型与固定模型差异极显著( $P < 0.000 5$ ),因此将含随机参数  $\mu_{12}$  的模型作为基于林分密度层次的最优模型。

数 Ident( $\alpha x_j$ )的模型为最优模型,其基本形式为:

$$\begin{cases} H_j = \beta_1 + (\beta_2 + \mu_{12j})DBH_j + \epsilon_j, \\ \mu_j \sim (\mu_{12j}) \sim N(0, \mathbf{D}), \\ \epsilon_j \sim N(0, \mathbf{R}_j), \mathbf{R}_j = \sigma^2 \psi_j^{0.5} \Gamma_j \psi_j^{0.5}, \\ \psi_j = Ident(-4.063\ 9x_j). \end{cases} \quad (9)$$

式中:  $\mu_{12j}$  为第 j 林分密度对应的随机参数, Ident( $\alpha x_j$ )为恒等式函数, DHB<sub>j</sub> 为第 j 林分密度对应的胸径。其余参数定义如前所述。

表 3 考虑异方差结构的混合效应模型的拟合结果

Table 3 Fitting results of mixed models based on considering heteroscedasticity

随机效应 Random effect	异方差结构 Heteroscedasticity structure	自由度 df	AIC	BIC	$-2\log lik$	LRT	P
密度效应 Density effect	Ident	13	2 045.705	2 099.523	2 019.704	64.026	<0.000 1
	Power	5	2 085.825	2 106.524	2 075.824	7.906	0.019 2
	Exp	5	2 085.822	2 106.521	2 075.822	7.909	0.019 2
海拔效应 Altitude effect	Ident	9	2 032.762	2 070.021	2 014.762	68.969	<0.000 1
	Power	10	2 039.998	2 081.397	2 019.998	63.733	<0.000 1
	Exp	10	2 051.882	2 093.280	2 031.882	51.850	<0.000 1

林分密度层次的线性混合模型的拟合结果表明,林分密度因子对树高生长具有显著影响,加入林分密度随机参数可以提高模型的拟合精度,分别考虑随机参数效应和随机参数效应与异方差的组合,模型的拟合精度都得到进一步提高。因此,混合模

型在实际运用中要想取得较为准确的拟合效果,需要综合考虑误差效应的方差协方差结构,消除误差异方差。

2.2.3 海拔层次线性混合效应模型的拟合 表 2 显示,基于海拔效应,线性混合模型的 3 种拟合结果

均收敛,拟合优度较固定模型有所提高。其中含有随机参数  $\mu_{21}$  和  $\mu_{22}$  混合模型的 AIC、BIC、 $-2\log \text{lik}$  值均最小,拟合效果最好。似然比检验(LRT)表明:包含 2 个随机参数( $\mu_{21}$ 、 $\mu_{22}$ )的模型与固定模型差异极显著( $P < 0.0001$ ),同时与另外 2 个收敛模型差异显著(表 2),说明增加模型随机参数的个数能提高模型的精度,因而将同时含  $\mu_{21}$  和  $\mu_{22}$  的模型作为基于海拔层次的最优混合模型。

将指数函数、幂函数和恒等式作为异方差函数加入模型中,结果(表 3)显示,含 3 种不同异方差函数的模型均收敛,且与不考虑异方差结构的模型差异极显著( $P < 0.0001$ )。比较 3 个评价指标,最终选定常数形式异方差结构的模型作为最优模型,其表达式为:

$$\begin{cases} H_j = (\beta_1 + \mu_{21j}) + (\beta_2 + \mu_{22j})DBH_j + \epsilon_j, \\ \mu_j \sim \begin{pmatrix} \mu_{21j} \\ \mu_{22j} \end{pmatrix} \sim N(0, \mathbf{D}), \\ \epsilon_j \sim N(0, \mathbf{R}_j), \mathbf{R}_j = \sigma^2 \psi_j^{0.5} \Gamma_j \psi_j^{0.5}, \\ \psi_j = \text{Ident}(0.217 \ 3x_j). \end{cases} \quad (10)$$

式中: $\mu_{21j}$ 、 $\mu_{22j}$  为第  $j$  海拔效应对应的随机参数,

表 4 固定模型与混合效应模型的参数及其拟合效果

Table 4 Parameters and fitting results of fixed models and mixed models

变量 Variable	参数 Parameter	固定模型 Fixed model		密度层次 Density level		海拔层次 Altitude level	
		估值 Estimate value	标准误 SE	估值 Estimate value	标准误 SE	估值 Estimate value	标准误 SE
固定效应 Fixed effect	$\beta_1$	1.183 6**	0.298 6	0.645 1*	0.298 6	-0.219 4	1.115 1
	$\beta_2$	0.793 9**	0.019 2	0.829 2**	0.023 0	0.876 1**	0.061 5
随机效应 Random effect	$D$			[0.001 3]		[4.572 6 -0.248 0]	[-0.248 0 0.013 5]
误差方差 Error variance	$\sigma^2$		5.049 0		4.283 3		3.545 1
评价指标 Evaluation index	$R^2$		0.785 9		0.797 0		0.804 4
	$ \bar{E} $		1.633 7		1.576 6		1.553 9
	RMSE		2.242 1		2.183 0		2.143 0

注(Notes): \*\* .  $P < 0.0001$ ; \* .  $P < 0.01$ ; .  $P < 0.5$ .

2.2.5 模型检验 建立  $H$ - $DBH$  关系的线性混合效应模型后,利用 220 组独立检验数据,根据公式(5)以及表 4 中各参数的模拟结果求出检验数据的随机参数  $\mu_k$  预测值,进而计算各树高的估计值,利用检验指标均方根误差(RMSE)、平均绝对残差( $|\bar{E}|$ )与固定模型进行精度比较,具体计算基于 Excel 和 R 3.2.2 软件实现,结果见表 5。从表 5 可以看出,林分密度层次的线性混合效应模型平均绝对残差  $|\bar{E}| = 1.5813$ 、均方根误差  $RMSE = 2.0321$ ;海拔层次混合模型平均绝对残差  $|\bar{E}| = 1.5802$ 、均方根误差  $RMSE = 1.9638$ ,均小于固定模型的平均绝对误差和均方根误差值( $|\bar{E}| = 1.6086$ ,  $RMSE = 2.0349$ ),证明本研究中加入随机效应参数的混合

$DBH_j$  为第  $j$  海拔效应对应的胸径。其余参数定义如前所述。

海拔层次线性混合模型的拟合结果表明,海拔因子对树高生长影响极为显著,加入海拔随机参数可较大程度地提高模型的拟合精度,同时考虑方差协方差结构,引入异方差函数(Ident)消除模型的异方差,模型的拟合精度可以进一步提高。

2.2.4 模型评价 利用 R 3.2.2 软件对上述混合效应模型进行参数估计。表 4 分别列出了固定模型与混合效应模型的固定参数值、误差方差以及模型评价指标值。表 4 结果显示,固定模型的拟合精度均低于混合效应模型,海拔层次混合效应模型拟合精度高于林分密度层次混合模型,说明混合效应模型的拟合效果优于固定模型,且不同随机参数对模型拟合效果的影响也不同。本研究结果表明,海拔对杉木树高生长的影响大于林分密度对其生长的影响,同时也说明随机参数体现了样本个体间的差异性。因此,在实际运用中要选择适当的随机参数以提高模型的精度。

效应模型的预测精度高于固定模型。林分密度层次混合模型的  $|\bar{E}|$  与  $RMSE$  值均大于海拔层次混合模型,因此加入海拔随机参数的线性混合模型在预测精度方面优于含有林分密度随机参数的混合模型,同时也证明在大环境中,海拔因子对胸径-树高关系的影响大于林分密度。

表 5 固定模型与线性混合效应模型的验证结果比较

Table 5 Comparison results of fixed model and linear mixed model

评价指标 Evaluation index	固定模型 Fixed model	混合模型 Mixed model	
		密度层次 Density level	海拔层次 Altitude level
$ \bar{E} $	1.608 6	1.581 3	1.580 2
RMSE	2.034 9	2.032 1	1.963 8

### 3 结论与讨论

利用从142篇杉木文献中收集的688组有效数据,基于R 3.2.2软件nlme包的lme功能,建立了杉木树高预测的*H-DBH*线性混合效应模型,并选用评价指标*AIC*、*BIC*、 $-2\log lik$ 和似然比检验(*LRT*)对模型的拟合优度进行评价。结果表明,混合模型的评价值均小于固定模型( $AIC = 2\ 089.731$ ,  $BIC = 2\ 102.151$ ,  $-2\log lik = 2\ 083.732$ ),基于独立检验数据对模型进行检验,混合模型的 $|\bar{E}|$ 、*RMSE*值均小于固定模型( $|\bar{E}| = 1.608\ 6$ ,  $RMSE = 2.034\ 9$ ),说明引入林分密度和海拔随机参数的混合效应模型的估测效果均优于固定模型,这也验证了以往学者的研究结果<sup>[20-23]</sup>,即海拔梯度与林分密度都会影响树高预测的准确性。

拟合林分密度层次的混合效应模型时,利用*AIC*、*BIC*、 $-2\log lik$ 3个评价指标对混合模型进行评价,结果表明加入随机参数 $\mu_{12}$ 的混合模型拟合效果最好。考虑误差效应的方差协方差结构,为解决数据间的异方差问题,在混合模型中添加指数函数、幂函数和恒等式等3种异方差结构,结果显示引入随机效应以及随机效应和恒等式异方差函数(*Ident*)的组合,混合模型的精度得到进一步提高,与仅考虑随机效应的模型差异极显著( $P < 0.000\ 1$ )。拟合海拔层次的混合效应模型时,引入随机参数 $\mu_{21}$ 和 $\mu_{22}$ 的混合模型的拟合效果最优,*AIC*、*BIC*、 $-2\log lik$ 3个评价指标值最小。相对于不考虑方差协方差结构矩阵的模型,加入异方差函数(*Ident*)的混合模型拟合效果更优( $AIC = 2\ 032.762$ ,  $BIC = 2\ 070.021$ ,  $-2\log lik = 2\ 014.762$ ),且差异极显著( $P < 0.000\ 1$ )。

对于含有不同随机效应参数的混合模型,利用独立数据检验的结果表明:引入海拔随机效应、恒等式形式的异方差函数的树高-胸径线性混合模型的平均绝对误差 $|\bar{E}| = 1.580\ 2$ 、均方根误差 $RMSE = 1.963\ 8$ ,在预测精度上高于含林分密度随机效应并考虑误差效应方差协方差结构的混合模型( $|\bar{E}| = 1.581\ 3$ ,  $RMSE = 2.032\ 1$ ),说明海拔效应对树高生长的影响大于林分密度效应,即来源于海拔间树高的误差差异性大于不同林分密度间的树高误差差异。这是因为不同海拔地区的年均温度、昼夜温差、年均降水、年活动积温等均不同,随着海拔升高物候期推迟,影响植被的生长<sup>[24-25]</sup>;而林分密度梯度的不同仅会导致林内小环境的改变,因此前者对杉木树

高生长的影响大于后者。

杉木树高的生长受海拔、林分密度、气候条件、立地条件、林种起源、遗传作用、经营措施等多种因素的共同影响与制约<sup>[26-28]</sup>。混合效应模型中的各参数均有其生物学意义,可以反映出杉木树高受各随机效应的影响程度。本研究探讨了林分密度效应、海拔效应对杉木树高的影响,为科学的杉木经营管理方案的编制和大尺度区域杉木树高生长的通用模型的研建提供了依据。然而,本研究拟合的*H-DBH*线性混合模型相对于固定模型在估测精度上提高幅度有限,且与混合模型中随机因子的选取有关,由于仅考虑了单一林分密度效应或海拔效应对*H-DBH*关系的影响,而对于大尺度区域模型的拟合需要考虑更多的影响因子(包括因子间的嵌套交互作用),因此提高大尺度模型的估测精度还需要进一步研究。

### [参考文献]

- [1] Peng C H, Zhang L J, Liu J X, et al. Developing and validating nonlinear height-diameter models for major tree species of Ontario's boreal forests [J]. Northern Journal of Applied Forestry, 2001, 18(3): 87-94.
- [2] Huang S, Titus S J, Wiens D P, et al. Comparison of nonlinear height-diameter functions for major Alberta tree species [J]. Canadian Journal of Forest Research, 2011, 22(9): 1297-1304.
- [3] Rupšys P. Height-diameter models with stochastic differential equations and mixed-effects parameters [J]. Journal of Forest Research, 2014, 20(1): 9-17.
- [4] Gomezgarcia E, Dieguezaranda U, Castedodorado F, et al. A comparison of model forms for the development of height-diameter relationships in even-aged stands [J]. Forest Science, 2014, 60(3): 560-568.
- [5] Vargaslarreta B, Castedodorado F, Alvarezgonzalez J G, et al. A generalized height-diameter model with random coefficients for uneven-aged stands in El Salto, Durango (Mexico) [J]. Forestry, 2009, 82(4): 445-462.
- [6] Grubb P J. Control of forest growth and distribution on wet tropical mountains, with special reference to mineral nutrition [J]. Annu Rev Ecol Syst, 1977, 8(1): 83-107.
- [7] Hummel S. Height, diameter and crown dimensions of *Cordia alliodora* associated with tree density [J]. Forest Ecology and Management, 2000, 127(1/2/3): 31-40.
- [8] Wang X, Fang J, Tang Z, et al. Climatic control of primary forest structure and DBH-height allometry in Northeast China [J]. Forest Ecology and Management, 2006, 234(1): 264-274.
- [9] Ryan M G, Yoder B J. Hydraulic limits to tree height and tree growth [J]. Bioscience, 1997, 47(4): 235-242.
- [10] Choat B, Jansen S, Brodribb T J, et al. Global convergence in the vulnerability of forests to drought [J]. Nature, 2012, 491

- (7426);752-755.
- [11] 赵俊卉, 亢新刚, 张慧东, 等. 长白山主要针叶树种胸径和树高变异系数与竞争因子的关系 [J]. 应用生态学报, 2009, 20(8):1832-1837.
- Zhao J H, Kang X G, Zhang H D, et al. Relationships between coefficient of variation of diameter and height and competition index of main coniferous trees in Changbai mountains [J]. Chinese Journal of Applied Ecology, 2009, 20(8):1832-1837.
- [12] Paulo J A, Tome J, Tome M, et al. Nonlinear fixed and random generalized height-diameter models for *Portuguese cork oak* stands [J]. Annals of Forest Science, 2011, 68(2):295-309.
- [13] 李春明. 混合效应模型在森林生长模型中的应用 [J]. 林业科学, 2009, 45(4):131-138.
- Li C M. Application of mixed effects models in forest growth model [J]. Scientia Silvae Sinicae, 2009, 45(4):131-138.
- [14] Lynch T B, Holley A G, Stevenson D J. A random-parameter height-DBH model for *Cherrybark oak* [J]. Southern Journal of Applied Forestry, 2005, 29(1):22-26.
- [15] Uzoh F C, Oliver W W. Individual tree height increment model for managed even-aged stands of ponderosa pine throughout the western United States using linear mixed effects models [J]. Forest Ecology and Management, 2006, 221(1):147-154.
- [16] 李春明. 基于两层线性混合效应模型的杉木林单木胸径生长量模型 [J]. 林业科学, 2012, 48(3):66-73.
- Li C M. Individual tree diameter increment model for Chinese fir plantation based on two-level linear mixed effects models [J]. Scientia Silvae Sinicae, 2012, 48(3):66-73.
- [17] 袁晓红, 李际平. 杉木人工林南北坡向树高-胸径生长曲线研究 [J]. 西北林学院学报, 2012, 27(2):180-183.
- Yuan X H, Li J P. Height, DBH growth models of fir artificial forest on northern and southern slopes [J]. Journal of Northwest Forestry University, 2012, 27(2):180-183.
- [18] 李文毅, 朱开明. 论杉木林的目标经营 [J]. 湖南林业科技, 2004, 31(6):84-85, 89.
- Li W Y, Zhu K M. Target management of *Cunninghamia lanceolata* forest [J]. Hunan Forestry Science & Technology, 2004, 31(6):84-85, 89.
- [19] Pinheiro J C, Bates D M. Mixed-effects models in S and S-Plus [M]. New York: Springer Verlag, 2000.
- [20] Huang S, Price D, Titus S J, et al. Development of ecoregion-based height-diameter models for white spruce in boreal forests [J]. Forest Ecology and Management, 2000, 129(1):125-141.
- [21] Feldpausch T R, Banin L, Phillips O L, et al. Height-diameter allometry of tropical forest trees [J]. Biogeosciences, 2011, 8(5):1081-1106.
- [22] 王怀芳. 林分密度对 30 年生杉木人工林生长的影响 [J]. 亚热带农业研究, 2012, 8(3):156-159.
- Wang H F. Impact of different density on the growth of 30-year-old Chinese fir plantation [J]. Subtropical Agriculture Research, 2012, 8(3):156-159.
- [23] Henry H L, Aarssen L W. The interpretation of stem diameter-height allometry in trees: biomechanical constraints, neighbour effects, or biased regressions? [J]. Ecology Letters, 1999, 2(2):89-97.
- [24] 王力, 李凤霞, 周万福, 等. 气候变化对不同海拔高山嵩草物候期的影响 [J]. 草业科学, 2012, 29(8):1256-1261.
- Wang L, Li F X, Zhou W F, et al. Impacts of climate change on *Kobresia pygmaea* phenophase at different altitude regions in Qinghai Plateau [J]. Pratacultural Science, 2012, 29(8):1256-1261.
- [25] 夏浩铭, 李爱农, 赵伟, 等. 2001—2010 年秦岭森林物候时空变化遥感监测 [J]. 地理科学进展, 2015, 34(10):1297-1305.
- Xia H M, Li A N, Zhao W, et al. Spatiotemporal variations of forest phenology in the Qinling zone based on remote sensing monitoring, 2001—2010 [J]. Progress in Geography, 2015, 34(10):1297-1305.
- [26] Scaranello M A, Alves L F, Vieira S A, et al. Height-diameter relationships of tropical Atlantic moist forest trees in south-eastern Brazil [J]. Scientia Agricola, 2012, 69(1):26-37.
- [27] Thomas S C, Martin A R, Mycroft E E, et al. Tropical trees in a wind-exposed island ecosystem: height-diameter allometry and size at onset of maturity [J]. Journal of Ecology, 2015, 103(3):594-605.
- [28] Banin L, Feldpausch T R, Phillips O L, et al. What controls tropical forest architecture? Testing environmental, structural and floristic drivers [J]. Global Ecology & Biogeography, 2012, 21(12):1179-1190.