

网络出版时间:2015-06-10 08:40

DOI:10.13207/j.cnki.jnwafu.2015.07.002

网络出版地址:http://www.cnki.net/kcms/detail/61.1390.S.20150610.0840.002.html

华山松种源对树高生长模型参数的影响

王亚南¹,王军辉¹,祁万宜²,高新章³,王庆灵⁴,张培⁴

(1 中国林业科学研究院 林业研究所,国家林业局林木培育重点实验室,北京 100091;

2 湖北省宜昌市林业科学研究所,湖北 宜昌 443111;3 湖北省宜昌市大老岭林场,湖北 宜昌 443132;

4 中国林业科学研究院 热带林业实验中心,广西 凭祥 532600)

[摘要] 【目的】研究华山松种源对树高生长模型参数的影响,为建立通用的华山松树高生长模型奠定基础。【方法】以湖北省宜昌市大老岭林场华山松种源试验林为研究对象,利用林龄 4~10 年、24 年和 32 年的树高数据,建立华山松 11 个种源的基础模型、哑变量模型和非线性混合模型,并对各模型的拟合精度进行比较。【结果】在 Richards、Logistic、Schumacher 和 Korf 4 个最常用的树高理论方程中,以 Richards 方程的拟合效果最好,故选其作为基础模型。哑变量模型和非线性混合模型的检验结果均显示,华山松种源对树高生长模型的渐进参数有显著影响,而对生长速率参数和形状参数无显著影响。对比 3 种模型的决定系数 R^2 、残差平方和 SSE、平均绝对残差 $|E|$ 和均方根误差 RMSE 4 个精度评价指标发现,哑变量模型和非线性混合模型的精度比基础模型明显提高,二者平均精度的决定系数 R^2 提高 1.37%,残差平方和 SSE 降低 18.27%,平均绝对残差 $|E|$ 降低 12.77%,均方根误差 RMSE 降低 9.6%;哑变量模型的拟合精度比非线性混合模型稍高,决定系数 R^2 提高 0.1%,残差平方和 SSE、平均绝对残差 $|E|$ 、均方根误差 RMSE 分别降低 1.8%、4% 和 0.9%。【结论】华山松种源对树高生长模型的渐进参数有显著影响,而对生长速率参数和形状参数无显著影响。在种源数量不多时,利用哑变量模型建立通用的华山松种源树高生长模型是一种较好的方法。

[关键词] 华山松;哑变量模型;非线性混合模型;种源;树高生长模型

[中图分类号] S758.2

[文献标志码] A

[文章编号] 1671-9387(2015)07-0074-08

Effects of *Pinus armandi* Franch provenances on parameters of height-age models

WANG Ya-nan¹, WANG Jun-hui¹, QI Wan-yi², GAO Xin-zhang³,

WANG Qing-ling⁴, ZHANG Pei⁴

(1 Research Institute of Forestry, Key Laboratory of Tree Breeding and Cultivation, State Forestry Administration, Chinese Academy of Forestry, Beijing 100091, China; 2 Yichang Institute of Forestry, Yichang, Hubei 443111, China;

3 Dalaoling Forest Farm of Yichang County of Hubei Province, Yichang, Hubei 443132, China; 4 Experiment Center of Tropical Forestry, Chinese Academy of Forestry, Pingxiang, Guangxi 532600, China)

Abstract: 【Objective】Parameters of height-age models for *Pinus armandi* Franch provenances were analyzed to improve the establishment of generalized height-age model for *Pinus armandi* Franch. 【Method】*Pinus armandi* Franch in provenance test forest in Dalaoling Forest Farm, Yichang, Hubei were chosen as experimental materials. Regression accuracies of basic model, nonlinear mixed model and dummy variable model describing height growth of 11 *Pinus armandi* Franch provenances with stand ages of 4—10, 24 and

[收稿日期] 2014-07-24

[基金项目] “十二五”国家科技支撑计划项目(2012BAD01B01)

[作者简介] 王亚南(1985—),男,河南汝南人,助理工程师,硕士,主要从事森林可持续经营研究。

E-mail: wangyanandiyici@163.com

[通信作者] 王军辉(1972—),男,河南郑县人,研究员,博士,主要从事云杉、楸树遗传育种研究。E-mail: wangjh808@sina.com

32 years were compared. 【Result】 Accuracy of Richards equation was higher than those of Logistic equation, Schumacher equation and Korf equation to describe the basic height-age model for *Pinus armandi* Franch provenances. Both nonlinear mixed model and dummy variable model showed that there was significant difference in asymptote parameter but not in growth-rate parameter or shape parameter for 11 *Pinus armandi* Franch provenances. Averaged determination coefficient (R^2), residual sum of squares (SSE), absolute residuals means ($|E|$) and minimum root mean square error (RMSE) of dummy variable model and nonlinear mixed model were 1.37% higher, 18.27% lower, 12.77% lower and 9.6% lower than those of the basic model. Dummy variable model was slightly better than nonlinear mixed model with increase of those four accuracies by 0.1%, 1.8%, 4%, and 0.9%, respectively. 【Conclusion】 *Pinus armandi* Franch provenances had significant effect on asymptote parameter but not in growth-rate parameter and shape parameter. The dummy variable model was suitable for the establishment of generalized height-age model for *Pinus armandi* Franch.

Key words: *Pinus armandi* Franch; dummy variable model; nonlinear mixed model; provenances; height-age models

利用数学模型描述客观世界是科学研究和技术开发的一种重要手段,这种方法在力学、物理学和工程技术中早已获得了巨大的成功。随着数量遗传学、数量生态学等新兴学科的相继诞生以及计算机和 3S 技术的快速发展,数据库、数值模拟等技术已经逐渐成为许多行业不可缺少的研究基础^[1]。我国目前正在积极推进全国通用性立木生物量模型的研建工作。在建立通用性立木生物量模型时,如何在保证预估精度的前提下尽量简化所建模型,是值得研究的一个问题,而混合效应模型和哑变量模型的研建为解决这一问题提供了可能。从国内外混合效应模型和哑变量模型在林分生长模拟中的应用来看,其在国外已广泛应用于林业研究中,而国内对这 2 种建模技术的研究相对较少^[2-27]。鉴于哑变量和混合效应模型的优点以及森林生长和收获模型的发展方向,今后哑变量模型和混合效应模型必将成为我国林木生长和收获模型的主要类型,因此有必要对哑变量模型和混合效应模型做进一步研究。

国内有关华山松生长模型的研究较少,有关其树高生长模型的研究更少,仅见黄华梨^[28]对人工华山松 2 株样木 22 年的树高生长规律进行研究,得出华山松人工林树高生长规律符合 Logistic 模型;何爱喜^[29]对小陇山华山松人工林树高生长普遍规律进行研究,建立了实用的树高生长模型。因此,非常有必要对华山松树高生长模型开展研究。本研究以湖北省宜昌市大老岭林场华山松种源试验林为研究对象,利用哑变量模型和非线性混合模型分析华山松不同种源对树高生长方程参数的影响,并对哑变量模型和非线性混合模型 2 种建模方法进行分析,

以期建立通用的华山松树高生长模型奠定基础。

1 材料与方法

1.1 数据来源

华山松种源试验林的详细情况见马常耕的《华山松种源选择研究》^[30],在林龄为 4~10 年、24 年和 32 年时对树高进行每木调查,为了便于统计分析,本研究的种源编号与全国华山松种源协作组的编号保持一致。由于自然灾害和病虫害等其他原因造成部分种源保存率低或生长不适应,因此本研究选用 11 个生长较好的种源进行研究,种源号分别为 1, 2, 5, 6, 7, 11, 12, 20, 22, 23 和 28。

1.2 理论生长方程的选择

在选择数学模型时,一般选择能够正确描述树木的生长过程并具有生物学含义的生长方程。本研究选用其中最常用的树高理论方程,即 Richards 方程、Logistic 方程、Schumacher 方程和 Korf 方程^[15,20,31-34]。

Richards 方程: $y = a(1 - \exp(-kx))^c + \epsilon$;
 $a, k > 0$ 。

Logistic 方程: $y = a / (1 + k \cdot \exp(-cx)) + \epsilon$;
 $a, c > 0$ 。

Schumacher 方程: $y = a \cdot \exp(-k/x) + \epsilon$;
 $a, k > 0$ 。

Korf 方程: $y = a \cdot \exp(-k \cdot x^{-c}) + \epsilon$;
 $a, k, c > 0$ 。

式中: y 为树高, x 为林龄, a 为渐进参数, c 为形状参数, k 为生长速率参数, ϵ 为随机误差。

1.3 哑变量模型

通过对比 4 个理论方程的精度可知, Richards 方程的拟合效果最好, 故选其作为基础方程。哑变量方法是将种源效应引入 Richards 方程, 通过引入种源效应扩展模型和没有种源效应基础模型的残差平方和及自由度计算 F_{i-j} 的值, 来检验种源对树高-林龄曲线的 3 个参数是否存在显著性影响^[19]。

$$F_{i-j} = \left(\frac{SSE_i - SSE_j}{SSE_j} \right) \left(\frac{n_j}{n_i - n_j} \right)。$$

式中: SSE_i 、 SSE_j 表示基础方程 i 和扩展方程 j 拟合的残差平方和, n_i 、 n_j 表示基础方程和扩展方程的自由度。 $F_{i-j} \sim F_{a(n_i - n_j), n_j}$ 分布, 如果计算出 $F_{i-j} > F_{a(n_i - n_j), n_j}$, 则该事件发生的概率 $P < \alpha$, 即方程 i 和方程 j 差异显著。

基础方程(1)Richards: $y = a(1 - \exp(-kx))^c + \varepsilon$; $a, k > 0$ 。

扩展方程(2): $y = (a_1 P_1 + a_2 P_2 + a_3 P_3 + a_4 P_4 + a_5 P_5 + a_6 P_6 + a_7 P_7 + a_8 P_8 + a_9 P_9 + a_{10} P_{10} + a_{11} P_{11})(1 - \exp(-kx))^c + \varepsilon$ 。

扩展方程(3): $y = (a_1 P_1 + a_2 P_2 + a_3 P_3 + a_4 P_4 + a_5 P_5 + a_6 P_6 + a_7 P_7 + a_8 P_8 + a_9 P_9 + a_{10} P_{10} + a_{11} P_{11})(1 - \exp(-(k_1 P_1 + k_2 P_2 + k_3 P_3 + k_4 P_4 + k_5 P_5 + k_6 P_6 + k_7 P_7 + k_8 P_8 + k_9 P_9 + k_{10} P_{10} + k_{11} P_{11})))^c + \varepsilon$ 。

扩展方程(4): $y = (a_1 P_1 + a_2 P_2 + a_3 P_3 + a_4 P_4 + a_5 P_5 + a_6 P_6 + a_7 P_7 + a_8 P_8 + a_9 P_9 + a_{10} P_{10} + a_{11} P_{11})(1 - \exp(-(c_1 P_1 + c_2 P_2 + c_3 P_3 + c_4 P_4 + c_5 P_5 + c_6 P_6 + c_7 P_7 + c_8 P_8 + c_9 P_9 + c_{10} P_{10} + c_{11} P_{11}))) + \varepsilon$ 。

式中: a_i 为第 i 种源的渐进参数, k_i 为第 i 种源的生长速率参数, c_i 为第 i 种源的形状参数, P_i 为第 i 种源的哑变量, 其赋值方法^[13]如下。

$$P_i = \begin{cases} 1, & \text{当种源为第 } i \text{ 个种源时;} \\ 0, & \text{当种源为其他种源时;} \end{cases} \quad i=1, 2, \dots, 11。$$

1.4 非线性混合模型

非线性混合模型通过考虑回归函数依赖于固定和随机效应的非线性关系而建立, 其主要特征是参数向量随样地的不同而变化, 参数被分成固定效应和随机效应 2 部分。本研究将渐进参数 a 、生长速率参数 k 和形状参数 c 的影响作为固定参数, 种源对固定参数 a 、 k 、 c 的影响为随机效应 m_i 、 n_i 、 t_i ; 利用方程 i 和方程 j 的对数似然比 LRT_{i-j} 差值来检验假设: 种源对树高-林龄曲线固定参数 a 、 k 、 c 的随机效应显著为 0, LRT_{i-j} 服从 $\chi^2_{(n_i - n_j, a)}$ 的 χ^2 分布, 如果

LRT_{i-j} 差值的出现概率 $P > 0.05$, 则接受原假设, 即种源对固定参数 a 、 k 、 c 的随机效应显著为 0, 种源对树高-林龄曲线的影响只有固定效应, 随机效应为 0; 若 LRT_{i-j} 差值的出现概率 $P < 0.05$, 则拒绝原假设, 即种源对固定参数 a 、 k 、 c 的随机效应不为 0, 种源对树高-林龄曲线的影响由固定效应和随机效应 2 部分组成^[10]。

$$LRT_{i-j} = 2 \log \left(\frac{L_i}{L_j} \right) = 2 [\log(L_i) - \log(L_j)]。$$

式中: L_i 和 L_j 分别为基础方程 i 和扩展方程 j 的似然值。 LRT_{i-j} 服从 $\chi^2_{(n_i - n_j, a)}$ 的 χ^2 分布, 如果计算出 $LRT_{i-j} > \chi^2_{(n_i - n_j, a)}$, 则该事件发生的概率 $P < \alpha$, 即基础方程 i 和扩展方程 j 差异显著。

混合模型方程(5): $y = (a + m_i) \cdot (1 - \exp(-kx))^c$ 。

混合模型方程(6): $y = (a + m_i) \cdot (1 - \exp(-(k + n_i)x))^c$ 。

混合模型方程(7): $y = (a + m_i) \cdot (1 - \exp(-kx))^{c+t_i}$ 。

式中: a 、 k 、 c 为种源的固定效应, m_i 、 n_i 、 t_i 为种源的随机效应。

1.5 模型评价指标

以决定系数 R^2 、残差平方和 SSE 、均方根误差 $RMSE$ 和平均绝对残差 $|E|$ 4 个精度评价指标^[20,22]对模型的模拟结果进行精度对比, 模型的 SSE 、 $RMSE$ 和 $|E|$ 越小, R^2 越大, 说明该模型拟合精度越高。

$$R^2 = 1 - \left[\frac{\sum_{i=0}^m \sum_{j=0}^{n_j} (y_{ij} - \hat{y}_{ij})^2}{\sum_{i=0}^m \sum_{j=0}^{n_j} (y_{ij} - \bar{y}_{ij})^2} \right];$$

$$SSE = \sum_{i=0}^m \sum_{j=0}^{n_j} (y_{ij} - \hat{y}_{ij})^2;$$

$$RMSE = \sqrt{\frac{\sum_{i=0}^m \sum_{j=0}^{n_j} (y_{ij} - \hat{y}_{ij})^2}{n-1}};$$

$$|E| = \frac{\sum_{i=0}^m \sum_{j=0}^{n_j} |y_{ij} - \hat{y}_{ij}|}{n}。$$

式中: i 为种源, j 为区组, m 为种源数量, n_j 为区组数量, n 为总样本数, y_{ij} 为观测值, $\hat{y}_{ij}$ 为预测值, $\bar{y}_{ij}$ 为观测值的平均值。

1.6 统计分析软件

方差分析利用 SAS 9.2 的 GLM 模块 type 3; 4 个理论生长方程以及哑变量模型和非线性混合模型, 利用中国林业科学研究院资源信息所研发的林

业统计分析软件 Forstat 2.2 标准版进行统计,模型的评价指标利用 SAS 9.2、Forstat 2.2 和 Excel 2007 配合计算。

2 结果与分析

2.1 华山松不同种源 9 个林龄树高的方差分析

华山松 11 个种源 9 个林龄树高的方差分析结

表 1 华山松 11 个种源 9 个林龄树高的方差分析

Table 1 Variance analysis for 11 *Pinus armandi* Franch provenances on height at 9 stand ages

林龄/年 Age	种源 Provenance				区组 Block			
	残差平方和 Square sum of residual	均方 Mean square	F 值 F value	Pr>F	残差平方和 Square sum of residual	均方 Mean square	F 值 F value	Pr>F
4	0.298	0.017	11.88	<0.000 1	0.020	0.004	2.86	0.019 1
5	1.095	0.061	10.40	<0.000 1	0.028	0.006	0.95	0.452 9
6	1.927	0.107	9.88	<0.000 1	0.035	0.007	0.64	0.667 8
7	4.117	0.229	9.38	<0.000 1	0.281	0.056	2.30	0.051 2
8	6.550	0.364	11.49	<0.000 1	0.498	0.100	3.14	0.011 6
9	9.439	0.524	10.51	<0.000 1	0.805	0.161	3.23	0.010 0
10	9.723	0.540	5.99	<0.000 1	1.053	0.211	2.33	0.048 5
24	254.391	14.133	4.45	<0.000 1	64.657	12.931	4.07	0.002 3
32	209.254	11.625	2.00	0.022 6	50.395	10.079	1.73	0.140 0

表 2 华山松 11 个种源 9 个林龄树高的方差分量和方差分量百分比

Table 2 Variance component and variance component percent for 11 *Pinus armandi* Franch provenances on height at 9 stand ages

林龄/年 Age	种源方差分量 Variance component of provenance	区组方差分量 Variance component of block	随机误差方差分量 Variance component of random error	种源方差分量百分比 Variance component percent of provenance	区组方差分量百分比 Variance component percent of provenance
4	0.002 5	0.000 1	0.001 4	62.5	2.5
5	0.010 5	0.000 2	0.006 2	62.1	1.2
6	0.019 6	0.000 1	0.010 1	65.8	0.3
7	0.037 1	0.002 2	0.022 6	59.9	3.6
8	0.059 9	0.004 8	0.031 5	62.3	5.0
9	0.085 6	0.007 8	0.052 5	58.7	5.3
10	0.095 3	0.010 1	0.087 2	49.5	5.2
24	1.732 4	0.324 5	3.748 1	29.8	5.6
32	0.558 7	0.466 9	7.747 5	6.4	5.3

2.2 华山松不同种源树高生长基础模型的选择

用 Richards 方程、Logistic 方程、Schumacher 方程和 Korf 方程对华山松 11 个种源 9 个林龄的 576 个树高值进行拟合,通过对比 4 个基础模型对华山松不同种源树高生长的拟合结果(表 3)可以得

果见表 1。由表 1 可以看出,华山松 11 个种源的树高在林龄 32 年时差异显著,在林龄 4~10 年、24 年时差异达到极显著。华山松 11 个种源 9 个林龄树高的方差分量和方差分量百分比结果见表 2。由表 2 可以看出,树高生长的差异主要是由种源引起的。

出,Richards 方程拟合的决定系数 R^2 最大,残差平方和 SSE、均方根误差 RMSE 和平均绝对残差 $|E|$ 最小,说明该模型拟合的效果最好,因此选用 Richards 方程作为基础模型方程。

表 3 4 个基础模型对华山松不同种源树高生长的拟合结果

Table 3 Regression results of four basic equations for height growth of 11 *Pinus armandi* Franch provenances

方程 Equation	a	k	c	决定系数 R^2	残差平方和 SSE	均方根误差 RMSE	平均绝对残差 $ E $
Richards	13.95±0.57	0.100 3±0.008 1	4.18±0.35	0.930 6	641.952 5	1.055 7	0.525 9
Logistic	12.14±0.22	0.224 0±0.007 6	57.95±5.15	0.929 8	648.970 1	1.061 5	0.561 0
Korf	23.99±3.16	1.052 0±0.100 9	27.03±4.76	0.928 7	659.227 5	1.069 8	0.567 6
Schumacher	25.71±0.59		24.69±0.54	0.928 7	659.524 3	1.070 1	0.561 8

注:a. 渐进参数;k. 生长速率参数;c. 形状参数。
Note:a. Asymptote parameter;k. Growth-rate parater;c. Shape parameter.

2.3 哑变量模型检验华山松不同种源对树高生长模型参数的影响

许多研究表明,渐进参数是最不稳定的参数^[10,19,23-24],故本研究首先对渐进参数进行种源差异性检验。利用扩展方程(2)和基础方程(1)对华山松不同种源树高生长拟合的残差平方和与自由度计算的 F_{1-2} 来检验假设: $H_1:a_1=a_2=\cdots=a_{11}$ 。

由哑变量模型对华山松不同种源树高生长拟合的残差平方和及 F 检验结果(表 4)可知, $F_{1-2}=12.08>F_{0.05(10,563)}=1.83,P<0.05$,拒绝原假设 H_1 ,即华山松不同种源的树高-林龄曲线的渐进参

数 a 不相同,在此结果的基础上检验假设 $H_2:k_1=k_2=\cdots=k_{11}$ 。由表 4 可知, $F_{2-3}=1.16<F_{0.05(10,553)}=1.83,P>0.05$,接受假设 H_2 ,即华山松不同种源的树高-林龄曲线的渐进参数 a 不相同,生长速率参数 k 相同,在此结果的基础上检验假设 $H_3:c_1=c_2=\cdots=c_{11}$ 。由表 4 可知, $F_{2-4}=0.91<F_{0.05(10,553)}=1.83,P>0.05$,接受假设 H_3 ,即华山松不同种源的树高-林龄曲线的渐进参数 a 不相同,生长速率参数 k 和形状参数 c 相同。哑变量模型检验的结果为:华山松不同种源树高生长方程的渐进参数 a 不相同,而生长速率参数 k 和形状参数 c 相同。

表 4 哑变量模型对华山松不同种源树高生长的拟合及 F 检验结果

Table 4 Regression results and F test of the dummy variable model for height growth of 11 <i>Pinus armandi</i> Franch provenances			
方程序号 Equation	残差平方和 SSE	F	P
(4)	519.935	$F_{2-4}=0.91$	>0.05
(3)	517.679	$F_{2-3}=1.16$	>0.05
(2)	528.534	$F_{1-2}=12.08$	<0.05
(1)	641.952		

2.4 非线性混合模型检验华山松不同种源对树高生长模型参数的影响

利用非线性混合模型检验华山松种源对树高模型渐进参数 a 的随机效应 m_i 是否显著为 0。利用方程(5)和基础方程(1)的似然值与自由度计算 LRT_{1-5} 来检验假设 $H_4:m_i=0$ 。由非线性混合模型对华山松不同种源树高生长的拟合及 LRT 检验结果(表 5)可知, $LRT_{1-5}=76.64,P<0.05$,拒绝假设

H_4 ,即种源对参数 a 的随机效应 m_i 不为 0。利用方程(6)、方程(7)和方程(5)的似然值与自由度计算的 LRT_{5-6} 、 LRT_{5-7} 来检验假设 $H_5:m_i\neq 0,n_i=0$ 和假设 $H_6:m_i\neq 0,t_i=0$ 。由表 5 可知, $LRT_{5-6}=0.14$, $LRT_{5-7}=0.09$,二者的 $P>0.05$,接受假设 H_5 和 H_6 ,即非线性混合模型检验的结果为:华山松不同种源对树高-林龄曲线渐进参数 a 的随机效应不为 0,对生长速率参数 k 和形状参数 c 的随机效应为 0。

表 5 非线性混合模型对华山松不同种源树高生长的拟合及 LRT 检验结果

Table 5 Regression results and likelihood ratio test (LRT) of nonlinear mixed model for height growth of 11 <i>Pinus armandi</i> Franch provenances						
方程序号 Equation	混合效应参数 Mixed effect	固定效应参数 Fixed effect	-2 倍似然值的对数 $-2\times\log\text{-likelihood}$	参数个数 Number of parameters	似然值的比 LRT	P
(7)	a,k	c	1 633.36	7	$LRT_{5-7}=0.09$	>0.05
(6)	a,c	k	1 633.31	7	$LRT_{5-6}=0.14$	>0.05
(5)	a	k,c	1 633.45	5	$LRT_{1-5}=76.64$	<0.05
(1)	无 No	a,k,c	1 710.09	4		

2.5 华山松不同种源对树高生长模型的影响

根据拟合华山松 11 个种源树高生长的哑变量模型(2)和非线性混合模型(5)的参数(表 6)可知,哑变量模型中生长速率参数 k 和形状参数 c 分别为 0.098 6和 4.12,而各种源的渐进参数 a 不同,变幅为 11.74~15.90,标准差为 1.34;非线性混合模型中生长速率参数 k 和形状参数 c 分别为 0.101 和 4.21,而各种源的渐进参数 a 不同,变幅为 11.78~15.56,标准差为 1.21。哑变量模型的渐进参数平均值比非线性混合模型的高 0.33,误差占平均值的 1.26%;哑变量模型的生长速率参数比非线性混合

模型的低 0.002 4,误差占平均值的 2.4%;哑变量模型的形状参数比非线性混合模型的低 0.09,误差占平均值的 2.2%。这说明华山松种源对哑变量模型和非线性混合模型参数的影响结果一致,即华山松种源对树高生长模型的渐进参数有显著影响,对生长速率参数和形状参数无显著影响。

2.6 基础模型、哑变量模型和非线性混合模型对华山松树高生长的拟合效果

由表 7 可以看出,哑变量模型和非线性混合模型明显要比基础模型好,二者的平均精度与基础模型相比决定系数 R^2 提高 1.37%,残差平方和 SSE

降低 18.27%,平均绝对残差|E|降低 12.77%,均方根误差 RMSE 降低 9.6%;哑变量模型的拟合精度比非线性混合模型稍高,决定系数 R^2 提高0.1%,

残差平方和 SSE、平均绝对残差|E|、均方根误差 RMSE 分别降低 1.8%,4%和 0.9%。

表 6 拟合华山松 11 个种源树高生长的哑变量模型(2)和非线性混合模型(5)的参数

Table 6 Regression parameters of the dummy variable model (2) and the nonlinear mixed model (5) for height growth of 11 *Pinus armandi* Franch provenances

种源 Provenance	哑变量模型(2) Dummy variable model(2)			非线性混合模型(5) Nonlinear mixed model(5)		
	<i>a</i>	<i>k</i>	<i>c</i>	<i>a</i>	<i>k</i>	<i>c</i>
1	15.32			15.03		
2	12.68			12.63		
5	14.69			14.45		
6	14.02			13.86		
7	14.42			14.22		
11	14.21	0.098 6	4.12	14.03	0.101	4.21
12	15.90			15.56		
20	14.03			13.86		
22	11.74			11.78		
23	12.38			12.36		
28	15.56			15.23		

表 7 基础模型(1)、哑变量模型(2)和非线性混合模型(5)对华山松树高生长的拟合精度

Table 7 Regression accuracies of basic model(1),dummy variable model (2) and nonlinear mixed model (5) for height growth of 11 *Pinus armandi* Franch provenances

方程序号 Equation	决定系数 R^2	残差平方和 SSE	平均绝对残差 E	均方根误差 RMSE
(1)	0.930 6	641.952 5	0.478 1	1.055 7
(2)	0.943 8	519.934 6	0.408 4	0.950 1
(5)	0.942 8	529.426 3	0.425 6	0.958 7

3 结论与讨论

3.1 种源对华山松树高生长模型的影响

根据哑变量模型和非线性混合模型检验华山松种源对树高生长模型参数的影响,显示 2 种模型的检验结果相同;华山松种源对树高生长模型的渐进参数有显著影响,而对生长速率参数和形状参数无显著影响。这一结论与许多针对种源和家系的研究结果相同。Nance 等^[25]运用 Richards 方程研究火炬松种源对树高-林龄模型中 3 个参数的影响,结果发现,相同立地条件下,不同种源的树高-林龄曲线的形状参数和生长速率参数相同,而树高的渐进值随种源的不同而不同。Buford 等^[19]运用 Schumacher 方程检验火炬松种源和自由授粉家系对模型参数的影响,发现 17 个地点的种源试验结果中,16 个地点的不同种源树高曲线斜率相同,60 个样地中 38 个样地的种源截距不相同;同时,也对火炬松子代测定林进行检验,结果与种源试验的结果相同:不同家系树高-林龄曲线的斜率相同,截距不同。李希菲等^[23]和王明亮等^[24]应用线性统计推断理论检验了杉木不同种源对立地指数曲线的影响,认为同一地

区不同种源的立地指数不同,相同种源不同地区的立地指数不同,说明立地指数的差异是遗传和环境差异的综合体现;而不同种源立地指数曲线的斜率相同,即各种源优势高生长可以用同一个模型来描述。但也有不同结论,如孙晓梅^[26]研究了 23 个日本落叶松自由授粉家系对树高生长模型参数的影响,认为家系效应对 Richards 方程的渐进参数和生长速率参数具有显著影响,而对形状参数影响不显著,相同立地条件下不同家系间树高-林龄曲线的渐进值(对数线性方程的截距参数)存在显著差异,而且曲线的形状(对数线性方程的斜率参数)也发生了明显改变,因此不同日本落叶松家系的优势高生长规律不同,应采用不同的生长模型来描述。

综上所述,相同立地条件下,不同种源或家系对树高模型参数的渐进值有显著影响,对形状参数和生长速率参数一般无显著影响,但家系水平有可能对形状参数有显著影响。

3.2 基础模型、哑变量模型和非线性混合模型对华山松树高生长拟合效果的比较

哑变量模型和非线性混合模型对树高生长的拟合效果比基础模型要好,这已被许多研究者所证实。

对比本研究中哑变量模型和非线性混合模型的精度发现,二者精度相似,决定系数和均方根误差很接近,但哑变量模型残差平方和、平均绝对残差比非线性混合模型的降低了 1.8%和 4%,这表明哑变量模型比非线性混合模型精度稍高。在模型结构上,哑变量模型要构造 11 个哑变量,在进行模拟时要给出 11 个种源的初始值,拟合编程相对比较繁琐;而非线性混合模型只需构造一个随机效应参数,模型结构简单,易于理解,编程方便。如果在模拟分类数目较大时,非线性混合模型的优势将非常明显。

究竟应该选择哑变量模型还是混合模型,在生物计量学和统计学领域一直还存在着争论。曾伟生等^[17]利用哑变量模型和线性混合模型对 2 个树种(杉木和马尾松)及其 2 个分布区(中心区和一般区)进行对比研究发现,2 种模型反映了同样的规律,而且在程度上差异非常小,哑变量模型和线性混合模型均可以显著减少估计误差,而且 2 种模型的估计效果几乎相同。Wang 等^[27]对 2 种模型进行了深入的对比研究,认为从实用性角度出发,一般可以根据划分类型的数量及每一类型的样本量来决定:如果类型数量少(如 10 个以下),则选择哑变量模型可能较好;如果类型数量多,而且每个类型包含的样本量又少,则一般推荐混合模型;如果每个类型的样本量都较大,则选择哪个模型都无关紧要。

[参考文献]

[1] 唐守正,李 勇.生物数学模型的统计学基础 [M].北京:科学出版社,2002.
Tang S Z,Li Y. The basic of statistic for biomathematical model [M]. Beijing:Science Press,2002. (in Chinese)

[2] Fox J C,Bi H Q,Ades P K. Spatial dependence and individual-tree growth models [J]. Forest Ecology and Management, 2007,245(1/2/3):20-30.

[3] Chi E M,Reinsel G C. Models for longitudinal data with random effects and ar(1) errors [J]. Journal of American Statistics Association,1989,84(406):452-459.

[4] Little R C,Milliken G A,Stroup W W. SAS system for mixed models [M]. North Carolina:SAS Press,1996.

[5] Nanos N,Rafael C,Gregorio M. Geostatistical prediction of height-diameter models [J]. Forest Ecology and Management, 2004,195:221-235.

[6] Rafael C,Gregorio M. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain [J]. Canadian Journal of Forest Research,2004,34(1):150-163.

[7] Lappi J,Bailey R L. Random parameter height-age models when stand parameters and stand age are correlated [J]. Forest Science,1994,40(4):715-731.

[8] Hall D B,Bailey R L. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models [J]. Forest Science,2001,47:311-321.

[9] 李永慈,唐守正.用 Mixed 和 Nlmixed 过程建立混合生长模型 [J]. 林业科学研究,2004(3):279-283.
Li Y C,Tang S Z. Using Mixed and Nlmixed to construct mixed growth models [J]. Forest Research,2004(3):279-283. (in Chinese)

[10] Fang Z,Bailey R L. Nonlinear mixed effects modeling for slash pine dominant height growth following intensive silvicultural treatments [J]. Forest Science,2001,47:287-300.

[11] Fang Z,Bailey R L,Shiver B D. A multivariate simultaneous prediction system for stand growth and yield with fixed and random effects [J]. Forest Science,2001,47:550-562.

[12] Zhao D,Wilson M,Borders B E. Modeling response curves and testing treatment effects in repeated measures experiments:A multilevel nonlinear mixed-effects model approach [J]. Canadian Journal of Forest Research,2005,35:122-132.

[13] 李丽霞,郜艳晖,张 瑛. 哑变量在统计分析中的应用 [J]. 数理医药学杂志,2006(1):51-53.
Li L X,Gao Y H,Zhang Y. The application of dummy variable in statistics analysis [J]. Journal of Mathematical Medicine, 2006(1):51-53. (in Chinese)

[14] 李希菲,洪玲霞.用哑变量法求算立地指数曲线族的研究 [J]. 林业科学研究,1997(2):108-112.
Li X F,Hong L X. Research on the use of dummy variables method to calculate the family of site index curves [J]. Forest Research,1997(2):108-112. (in Chinese)

[15] 孙晓梅,张守攻,孔凡斌,等.日本落叶松家系对树高生长模型参数的影响 [J]. 林业科学,2005,41(1):78-84.
Sun X M,Zhang S G,Sun F B,et al. Analyzing parameters of height-age models for open-pollinated Japanese Larch families [J]. Scientia Silvae Sinicae,2005,41(1):78-84. (in Chinese)

[16] 李忠国,孙晓梅,陈东升,等.基于哑变量的日本落叶松生长模型研究 [J]. 西北农林科技大学学报:自然科学版,2011,39(8):69-74.
Li Z G,Sun X M,Chen D S,et al. The dummy variables model of increment of *Larix kaempferi* [J]. Journal of Northwest A&F University:Nat Sci Ed,2011,39(8):69-74. (in Chinese)

[17] 曾伟生,唐守正,夏忠胜,等.利用线性混合模型和哑变量模型方法建立贵州省通用性生物量方程 [J]. 林业科学研究,2011(3):285-291.
Zeng W S,Tang S Z,Xia Z S,et al. Using linear mixed model and dummy variable model approaches to construct generalized single-tree biomass equations in Guizhou [J]. Forest Research,2011(3):285-291. (in Chinese)

[18] 孙晓梅,张守攻,李凤日,等.遗传改良林分生长和收获预估模型的研究进展 [J]. 林业科学研究,2004(4):525-532.
Sun X M,Zhang S G,Li F R,et al. An overview on growth and yield projection models for genetically improved stands [J]. Forest Research,2004(4):525-532. (in Chinese)

- [19] Buford M A, Burkhart H. Genetic improvement effects on growth and yield of loblolly pine plantations [J]. Forest Science, 1987, 33(3): 707-724.
- [20] 李春明. 混合效应模型在森林生长模拟研究中的应用 [D]. 北京: 中国林业科学研究院, 2010.
Li C M. Application of mixed effects models in forest growth models [D]. Beijing: Chinese Academy of Forestry, 2010. (in Chinese)
- [21] 李春明, 张会儒. 利用非线性混合模型模拟杉木林优势木平均高 [J]. 林业科学, 2010(3): 89-95.
Li C M, Zhang H R. Modeling dominant height for Chinese fir plantation using a nonlinear mixed-effects modeling approach [J]. Scientia Silvae Sinicae, 2010(3): 89-95. (in Chinese)
- [22] 姜立春, 杜书立. 基于非线性混合模型的东北兴安落叶松树高和直径生长模拟 [J]. 林业科学研究, 2012(1): 11-16.
Jiang L C, Du S L. Height and diameter growth modeling of *Dahurian larch* based on nonlinear mixed model in northeastern China [J]. Forest Research, 2012(1): 11-16. (in Chinese)
- [23] 李希菲, 王明亮. 杉木种源对立地指数模型的影响 [J]. 林业科学研究, 2000(6): 641-645.
Li X F, Wang M L. The impact of Chinese fir provenances on site index curves [J]. Forest Research, 2000(6): 641-645. (in Chinese)
- [24] 王明亮, 李希菲. 杉木种源对断面积模型影响的检验和比较 [J]. 林业科学研究, 1999(6): 585-590.
Wang M L, Li X F. Testing and comparing basal-area models for Chinese fir provenances [J]. Forest Research, 1999(6): 585-590. (in Chinese)
- [25] Nance W L, Wells O O. Site index models for height growth of planted loblolly pine seeds sources [C]//Proc 16th south for-tree improv conf. [S. l.]: [s. n.], 1981: 140-148.
- [26] 孙晓梅. 日本落叶松纸浆材优良家系选择及家系生长模型的研究 [D]. 北京: 中国林业科学研究院, 2003.
Sun X M. Selection of superior *Larix kaempferi* families for pulpwood purpose and family-based growth simulation [D]. Beijing: Chinese Academy of Forestry, 2003. (in Chinese)
- [27] Wang M L, Borders B E, Zhao D H. An empirical comparison of two subject-specific approaches to dominant heights modeling: The dummy variable method and the mixed model method [J]. Forest Ecology and Management, 2008, 255(7): 2659-2669.
- [28] 黄华梨. 人工华山松高生长规律初步研究 [J]. 甘肃林业科技, 2007, 32(2): 52-53.
Huang H L. Study on the height growth of *Pinus armandi* Franch plantation [J]. Journal of Gansu Forestry Science and Technology, 2007, 32(2): 52-53. (in Chinese)
- [29] 何爱喜. 秦岭西端小陇山林区华山松树高生长模型的研究 [J]. 甘肃科技, 2012, 28(13): 140-142.
He A X. Study on the height growth models of *Pinus armandi* Franch forest in Xiaolong Mountain in the west of Qinling Mountains [J]. Gansu Science and Technology, 2012, 28(13): 140-142. (in Chinese)
- [30] 马常耕. 华山松种源选择研究 [M]. 北京: 北京农业大学出版社, 1991: 13-21.
Ma C G. *Pinus armandi* Franch provenances selection [M]. Beijing: China Agricultural University Press, 1991: 13-21. (in Chinese)
- [31] Liu Z G, Li F R. The generalized Chapman-Richards function and applications to tree and stand growth [J]. Journal of Forestry Research, 2003, 14(1): 19-26.
- [32] 张建国, 段爱国. 理论生长方程与直径结构模型的研究 [M]. 北京: 科学出版社, 2004.
Zhang J G, Duan A G. Approach to theoretical growth equations and diameter structure models [M]. Beijing: Science Press, 2004. (in Chinese)
- [33] 李凤日, 吴俊民, 鲁胜利. Richards 函数与 Schnute 生长模型的比较 [J]. 东北林业大学学报, 1993(4): 15-24.
Li F Z, Wu J M, Lu S L. Comparison of Richards growth and Schnute model [J]. Journal of Northeast Forestry University, 1993(4): 15-24. (in Chinese)
- [34] Zeide B. Accuracy of equations describing diameter growth [J]. Can J For Res, 1989, 19: 1283-1286.