

DOI:10.12171/j.1000-1522.20190366

基于混合效应模型的人工红松枝下高模型研建

燕云飞 王君杰 姜立春

(东北林业大学林学院, 森林生态系统可持续经营教育部重点实验室, 黑龙江 哈尔滨 150040)

摘要:【目的】基于帽儿山红松人工林 63 块样地 2 972 株红松数据, 利用非线性混合模型构建红松枝下高模型, 为进一步研究生长与收获模型提供理论依据。【方法】本文首先使用 8 个常用的枝下高模型, 选出最优基础模型; 其次, 研究林分变量或单木变量对枝下高的影响, 建立含林分变量的枝下高模型; 最终在基础模型和含林分变量模型的基础上, 考虑样地效应对红松枝下高的影响, 构建红松枝下高基础混合效应模型和广义混合效应模型。模型用 4 种抽样方式(随机抽取、抽取最大树、抽取最小树、抽取平均树)和 8 种样本大小(1~8 株树)对基础混合效应模型和广义混合效应模型进行抽样检验。【结果】Logistic 模型拟合精度好, 符合生物学意义, 且模型形式简单, 选为最优基础模型。除树高、胸径以外, 大于对象木断面积之和、优势木高和冠幅与枝下高有显著相关性, 加入后明显提升模型的拟合精度。枝下高广义混合效应模型的拟合效果要优于其他模型。模型检验结果表明: 当应用基础混合效应模型预测时, 建议抽取胸径最小的 4 个样本; 当应用广义混合效应模型预测时, 建议随机抽取 4 个样本。【结论】枝下高广义混合效应模型在拟合效果和预测精度方面优于其他 3 种模型, 建议将此模型作为人工红松枝下高模型。当应用广义混合效应模型预测时, 建议随机抽取 4 个样本。

关键词: 红松; 非线性混合模型; 枝下高; 人工林; 混合模型校正

中图分类号: S758.5 **文献标志码:** A **文章编号:** 1000-1522(2020)09-0028-09

引文格式: 燕云飞, 王君杰, 姜立春. 基于混合效应模型的人工红松枝下高模型研建 [J]. 北京林业大学学报, 2020, 42(9): 28-36. Yan Yunfei, Wang Junjie, Jiang Lichun. Construction of the height to crown base mixed model for Korean pine[J]. Journal of Beijing Forestry University, 2020, 42(9): 28-36.

Construction of the height to crown base mixed model for Korean pine

Yan Yunfei Wang Junjie Jiang Lichun

(Key Laboratory of Sustainable Forest Ecosystem Management of Ministry of Education, School of Forestry,
Northeast Forestry University, Harbin 150040, Heilongjiang, China)

Abstract: [Objective] Based on the data of 2 972 Korean pine trees in 63 sample plots of Korean pine plantation in Maor Mountain of northeastern China, a nonlinear mixed model was used to construct the height to crown base model of Korean pine, which provided a theoretical basis for further research on growth and yield model. [Method] Firstly, the optimal basic model was selected from eight commonly used models of height to crown base. Secondly, the influence of stand variables or individual tree variables on height to crown base was studied, and a generalized model was established. Finally, on the basis of the basic model and the generalized model, considering the effects of sample plot on the height to crown base of Korean pine, the basic mixed effect model and generalized mixed effect model of height to crown base for Korean pine were constructed. Four alternatives of height to crown base (HCB) sampling designs (the randomly selected trees and selecting the maximum, medium-size and minimum trees) and eight sample sizes (1-8 trees) were studied for sampling correction of basic mixed effect model and generalized mixed effect model respectively. [Result] Logistic model had good fitting accuracy and biological significance,

收稿日期: 2019-09-20 修回日期: 2019-12-19

基金项目: 国家自然科学基金项目(31570624), 黑龙江省应用技术与开发计划项目(GA19C006), 中央高校基本科研业务费专项(2572019CP15)。

第一作者: 燕云飞。主要研究方向: 森林经理。Email: lwyanyunfei@163.com 地址: 150040 黑龙江省哈尔滨市香坊区和兴路 26 号东北林业大学林学院。

责任作者: 姜立春, 教授, 博士生导师。主要研究方向: 森林经理。Email: jlichun@nefu.edu.cn 地址: 同上。

本刊网址: <http://j.bjfu.edu.cn>; <http://journal.bjfu.edu.cn>

and its form was simple, so it was chosen as the the optimal basic model. In addition to tree height (H) and DBH, there was a significant correlation between the basal area sum larger than subject tree (BAL), dominant height (HD), crown width (CW) and HCB, and the fitting accuracy of the model was improved obviously. The fitting effect of height to crown base generalized mixed effect model was better than that of other models. The model validation showed that when the basic mixed effect model was used to predict, it was recommended to select four samples with the smallest DBH, and when the generalized mixed effect model was used to predict, it was recommended to randomly select four samples. [Conclusion] The generalized mixed effect model is superior to the other three models in fitting effect and prediction accuracy. It is recommended that this model can be used as the height to crown base model for Korean pine. When applying generalized mixed effect model prediction, it is recommended that four samples can be randomly selected.

Key words: Korean pine; nonlinear mixed model; height to crown base; plantation; mixed model calibration

枝下高是指树冠的第一活枝到地面的高度^[1],它是单木树冠特征的一个重要指标,反映树木的生长活力和生产力^[2],以及林分内竞争水平等^[3-4]。枝下高也通常作为预测变量应用于冠幅模型^[5-7]、冠形模型^[8]以及生物量模型^[9]等,此外枝下高还可以用于计算冠长、冠长率等林分生长与收获模型的重要预测变量。在枝下高实际的测量中,尤其是在郁闭度较大的林分中^[10-12],测量精度较差,且浪费大量的人力和物力,因此构建树木枝下高预测模型具有重要的实际意义。

国内外学者对枝下高模型进行过一些研究。Ritchie等^[13]考虑树高、树冠竞争因子、林分断面积的自然对数、胸径和树高的比值等预测变量,利用Logistic模型对美国的14个树种建立了枝下高模型,模型的预测精度有较大的提高。Rijal等^[14]利用Logistic模型建立了13个树种的枝下高预估模型,模型包含林木大小变量(树高、胸径和高径比)以及竞争因子变量(冠幅竞争因子和大于对象木断面积和),使得各树种的预测精度有显著的提升。Yang等^[15]将林分密度和气象因子作为预测变量加入到白杨(*Populus tremuloides*)、扭叶松(*Pinus contorta*)和白云杉(*Picea glauca*)的枝下高模型中,得到较好的预测效果。段光爽等^[16]将竞争因子和林分密度作为预测变量加入到华北落叶松(*Larix principis-rupprechtii*)枝下高模型中,提高了模型的预测精度。

近年来,混合效应模型方法在枝下高模型中也得到了一定的应用^[10,16-17],主要是因为混合效应模型能提高模型的预测精度。目前利用混合效应模型建模需要解决的关键问题是当使用混合模型进行预测时,需要估计随机参数。估计随机参数需要二次抽样,因此确定合适的样本数是非常必要的,合适的样本数可以极大地减少消耗的人力和物力^[18-19]。当使用非线性混合模型建立树高曲线^[20-24]、断面积生长^[19,25]

以及胸径生长^[26-27]等模型时,合适的样本数估计随机参数已有较多讨论,但较少涉及混合效应枝下高模型,尤其是混合效应模型实际应用时不同抽样方案的比较,仅见Fu等^[10]和Sharma等^[17]的研究报道。

红松(*Pinus koraiensis*)主要分布在东北小兴安岭、长白山等林区。红松木材轻软、细致、纹理直、耐腐蚀性强,为建筑、桥梁、枕木、家具等优良用材,是东北地区的主要造林树种之一。根据查阅的国内外文献,关于红松枝下高模型的研究鲜有报道,因此本文以东北红松人工林为研究对象,构建含有林木大小,立地条件以及竞争变量等综合变量的广义枝下高模型,并以样地效应作为分类变量构建非线性混合效应枝下高模型,分析不同模型的预估效果,为红松人工林科学经营的精准预测提供理论依据。

1 数据与方法

1.1 研究区概况

帽儿山实验林场位于黑龙江省南部尚志市帽儿山镇,地理坐标为127°29'~127°44'E,45°14'~45°29'N,地处张广才岭西坡,位于阿什河上游,以山区丘陵为主,平均海拔300m;属中温带湿润性季风气候,年平均气温2.8℃,年平均降水量700~800mm,全年日照时数2150~2480h。林场总面积约264km²。由侏罗纪中酸性火册岩构成,其中帽儿山是哈尔滨附近最高峰,为东北林业大学下设的实验林场。地带性土壤为暗棕壤,有机质含量和各种化学元素、微量元素含量都较高。植被属于长白山植物区系,是由地带性顶级植被阔叶红松林经人为干扰破坏后形成的较典型的东北东部天然次生林。

1.2 实验数据

2018年在帽儿山实验林场不同林分设置了63块红松人工林样地,样地面积在0.02~0.06hm²之间。对样地内树木进行每木检尺,分别测量了树木

的胸径、树高、枝下高、冠幅以及每株树的相对位置坐标。通过实测数据计算林分密度(N)、高径比(RHD), 大于对象木断面积和(BAL)、冠幅(CW)和优势木高(HD)等变量指标。本研究实测红松样木

2 972 株, 按照 4:1 的比例随机抽样, 其中 51 块样地 2 417 株红松作为建模数据, 12 块样地 555 株红松作为检验数据。红松人工林样木及林分特征因子的统计见表 1。

表 1 红松人工林建模数据和检验数据基本统计量

Tab. 1 Statistics of fitting data and validation data of Korean pine plantation

数据类型 Data type	变量 Variable	均值 Mean	最小值 Minimum	最大值 Maximum	标准差 Standard deviation	变异系数 Coefficient of variation
建模数据 Fitting data	枝下高 Height to crown base (HCB)/m	5.27	1.10	13.20	2.49	47.24
	树高 Tree height (H)/m	10.86	3.30	19.60	3.16	29.12
	胸径 DBH (D)/cm	14.78	5.00	36.50	5.99	40.53
	高径比 Height-diameter ratio (RHD)	0.78	0.43	2.68	0.17	22.05
	大于对象木的断面积和/($\text{m}^2 \cdot \text{hm}^{-2}$) Basal area sum larger than subject tree (BAL)/($\text{m}^2 \cdot \text{ha}^{-1}$)	1.01	0.00	2.56	0.50	49.54
	冠幅 Crown width (CW)/m	1.66	0.33	4.30	0.44	26.46
	优势木高 Dominant height (HD)/m	17.88	14.90	20.62	1.23	6.86
检验数据 Validation data	枝下高 HCB/m	3.62	0.80	12.10	1.89	52.25
	树高 H /m	9.32	3.20	17.90	2.78	29.81
	胸径 DBH (D)/cm	11.72	5.00	28.20	4.57	38.98
	高径比 RHD	0.83	0.50	1.59	0.16	19.13
	大于对象木的断面积和/($\text{m}^2 \cdot \text{hm}^{-2}$) BAL/($\text{m}^2 \cdot \text{ha}^{-1}$)	0.87	0.00	1.40	0.31	35.45
	冠幅 CW/m	1.70	0.78	3.83	0.40	23.73
	优势木高 HD/m	17.93	16.86	19.47	0.81	4.54

1.3 非线性混合效应模型

1.3.1 基础模型的选择

本文选择了以下 8 个常用的枝下高模型作为候选基础模型(表 2)。利用建模数据中的树高(H)和胸径(D)作为基础变量拟合基础模型, 表中 $X = b_0 + b_1D$ 。根据均方根误差(RMSE)和决定系数(R^2)从以

表 2 候选基础枝下高模型

Tab. 2 Model of HCB candidate

模型 Model	模型表达式 Model expression	参考文献 Reference
(1)	$HCB = H / (1 + \exp(X))$	[31]
(2)	$HCB = H / \sqrt{1 + \exp(X)}$	[14]
(3)	$HCB = H / \sqrt[6]{1 + \exp(X)}$	[14]
(4)	$HCB = H(1 - \exp(X))$	[29]
(5)	$HCB = H(a + \exp(X))$	[30]
(6)	$HCB = H(1 - a \exp(X^2))$	[31]
(7)	$HCB = H / \sqrt{1 + a \exp(X)}$	[14]
(8)	$HCB = H(1 - a \exp(X^c))$	[14]

注: a 、 c 为模型参数; X 为关于林木大小、竞争因子、立地条件的函数, 在表中特指 $X = b_0 + b_1D$, b_0 、 b_1 为模型参数。下同。Notes: a and c refer to model parameters; X is the function about tree size competition index and site condition, here in the table, $X = b_0 + b_1D$, b_0 , b_1 are model parameters. The same below.

下模型中筛选最优基础模型[10,14,16,28-31]。

1.3.2 广义枝下高模型

在最优基础模型基础上考虑林分变量对枝下高的影响, 构建广义模型。根据以往的研究结果以及红松林分变量对枝下高影响的分析, 主要采用林木大小变量、立地质量以及竞争变量等林分变量, 即 X 函数的表达式为:

$$X = f(\text{SIZE}, \text{SITE}, \text{COMP})$$

式中: SIZE 表示林木大小变量, SITE 表示立地质量, COMP 表示竞争变量。

(1)林木大小变量: 含有胸径(D)和树高(H)变量的模型是基础形式, 高径比是另外一个很重要的因子。高径比和枝下高都受林分密度的影响, 高径比的值越大, 枝下高越高。SIZE 可表示为:

$$\text{SIZE} = b_1D + b_2H + b_3\text{RHD}$$

(2)立地质量: 优势木高(HD)表示林分立地条件的优劣。SITE 可表示为:

$$\text{SITE} = c_1\text{HD}$$

(3)竞争变量: 林分密度(N)是影响林分生长和林分稳定性的重要因子。大于对象木断面积和(BAL)是密度的体现形式之一, 表明了林分内一株树木的

群落等级和相对的优势程度;冠幅(CW)的竞争影响到林分中树木的自然整枝,间接影响枝下高的大小。因此 COMP 可以表示为:

$$\text{COMP} = d_1 \text{BAL} + d_2 \text{CW} + d_3 N$$

通过决定系数(R^2)、平均绝对误差(MAE)和均方根误差(RMSE)等对添加林分变量构建的广义枝下高模型进行评价和选择。

1.3.3 非线性混合模型的构建

将效果最好的基础模型和广义模型加入随机效应,构建混合模型。首先,将不同随机参数组合的模型进行拟合,利用 AIC、BIC 等指标进行筛选,且为避免过度参数化,对所选模型进行 LRT 检验。其次,根据以往研究结果,本文选定广义正定结构为方差协方差结构。为了确定样地内的方差协方差结构,必须解决异方差的问题。目前林业上基本都采用下面的公式来描述:

$$R_i = \sigma^2 G_i^{0.5} \Gamma_i G_i^{0.5} \quad (9)$$

式中: σ^2 为模型的误差方差值; Γ_i 为组内误差相关性结构; $G_i^{0.5}$ 为描述方差异质性的对角矩阵。

常用的指数函数和幂函数以及常数加幂函数的方法来描述混合模型产生的异方差现象,并求得 AIC 和 BIC 来求得最合适的模型。3 个模型的形式^[32]如下:

指数函数:

$$g(u_{ij}, \alpha) = \exp(u_{ij}\alpha) \quad (10)$$

幂函数:

$$g(u_{ij}, \beta) = |u_{ij}|^\beta \quad (11)$$

常数加幂函数:

$$g(u_{ij}, \delta) = |u_{ij}|^{\delta_1} + \delta \quad (12)$$

式中: u_{ij} 代表第 i 个样地第 j 棵树木的胸径、枝下高观测值和枝下高拟合值; α 、 β 、 δ 、 δ_1 均为待估参数。

1.3.4 随机效应模型的抽样设计

混合效应模型的检验不同于传统模型的检验,包含固定效应参数以及随机效应参数。固定参数的检验与传统最小二乘模型的检验方法相同,而随机效应参数的检验需要计算随机参数值,在每个样地抽取 1~8 株样本,通过提取样本的数据,利用公式计算随机参数估计值:

$$\hat{b}_k \approx \hat{D} \hat{Z}_k^T (\hat{Z}_k \hat{D} \hat{Z}_k^T + \hat{R}_k)^{-1} \hat{e}_k \quad (13)$$

式中: \hat{b}_k 为随机参数估计值, \hat{D} 为随机效应参数的方差协方差矩阵, \hat{Z}_k 为设计矩阵, \hat{R}_k 为样地内方差协方差矩阵, \hat{e}_k 为实际值减去固定效应参数计算的预测值。

不同的抽样方案会有不同的校正参数值,为了探索模型预测精度和抽样方案的关系,本研究采用

常用的 4 种抽样方案分别对基础混合效应模型和广义混合效应模型校正并进行选择,其中随机抽取的方案模拟计算 100 次,4 种抽样方案分别为:

(1)每个样地随机抽取 1~8 株树,测量各样木胸径、树高、冠幅以及枝下高等样本数据。

(2)每个样地抽取胸径最大的 1~8 株树,测量各样木胸径、树高、冠幅以及枝下高等样本数据。

(3)每个样地抽取胸径最小的 1~8 株树,测量各样木胸径、树高、冠幅以及枝下高等样本数据。

(4)根据样地的平均胸径,抽取 1~8 株与平均胸径相近的树,测量各样木胸径、树高、冠幅以及枝下高等样本数据。

1.4 模型评价指标

本研究所使用的评价指标为决定系数(R^2)、平均绝对误差(MAE)以及均方根误差(RMSE)来进行模型的筛选以及模型的预测能力的评估。

$$R^2 = 1 - \frac{\sum (HCB - \hat{HCB})^2}{\sum (HCB - \bar{HCB})^2} \quad (14)$$

$$\text{MAE} = \sum |HCB - \hat{HCB}| / n \quad (15)$$

$$\text{RMSE} = \sqrt{\sum (HCB - \hat{HCB})^2 / (n - p)} \quad (16)$$

式中: n 为观测的样本数;HCB 为枝下高的观测值; \hat{HCB} 为枝下高预测值; \bar{HCB} 为枝下高的平均值; p 为模型参数数量。

在评价和比较混合模型方面,除考虑以上评价指标,还参考赤池信息量(AIC)和贝叶斯信息量(BIC)等评价指标。为避免混合模型中参数过多导致模型过参数化问题,还对参数不同的模型采取似然比检验(LRT),若 $P < 0.0001$,则认为模型差异显著,选择出收敛且精度高的模型^[16,24]。

$$\text{AIC} = -2\text{LOGLIK} + 2\lambda \quad (17)$$

$$\text{BIC} = -2\text{LOGLIK} + \lambda \log(n) \quad (18)$$

$$\text{LRT} = 2(\text{LOGLIK1} - \text{LOGLIK2}) \quad (19)$$

式中:LOGLIK 为拟合模型的对数似然值, λ 为模型中固定效应参数的个数, n 为观测样本数,LOGLIK1、LOGLIK2 分别为需要比较的两个不同随机参数模型的对数似然值。

2 结果与分析

2.1 枝下高基础模型

基于建模数据利用 R 软件的 NLS 拟合候选基础模型,得到决定系数(R^2)、均方根误差(RMSE)以及模型的参数(表 3)。在拟合过程中发现模型(5)、模型(6)、模型(7)和模型(8)不收敛。其他 4 个模型的参数估计值都显著($P < 0.05$)且都能解释大于

表3 候选模型评价指标和拟合参数

Tab. 3 Evaluation indices and fitting parameters of candidate model

模型 Model	参数 Parameter	参数估计值 Parameter estimate	拟合优度 Goodness-of-fit statistics	
			R^2	RMSE
(1)	b_0	0.396 4	0.579 5	1.614 8
	b_1	-0.020 0		
(2)	b_0	1.633 6	0.577 9	1.617 7
	b_1	-0.026 9		
(3)	b_0	5.085 8	0.577 7	1.618 3
	b_1	-0.050 0		
	a	0.561 9		
(4)	b_0	0.039 0	0.578 5	1.616 7
	b_1	-0.008 4		
	a			
(5)	b_0	不收敛 No convergence		
	b_1			
	a			
(6)	b_0	不收敛 No convergence		
	b_1			
	a			
(7)	b_0	不收敛 No convergence		
	b_1			
	c			
(8)	b_0	不收敛 No convergence		
	b_1			
	c			

57%的枝下高的变化,其中模型(1)为 Logistic 模型,显示了较小的均方根误差 RMSE 以及最大决定

$$HCB_{ij} = \frac{H_{ij}}{1 + \exp((b_0 + u_0) + (b_1 + u_1)D_{ij} + c_1HD_{ij} + d_1BAL_{ij} + (d_2 + u_4)CW_{ij})} + \varepsilon_{ij} \quad (22)$$

式中: u_4 为随机效应参数。

本研究也比较了混合模型中常用的3种方差协方差结构,根据AIC和BIC可得到广义正定矩阵结构为最佳方差协方差结构。由于异方差的存在,本文对于3种方差函数分别进行了研究,指数函数和常数项加幂函数在异方差的校正过程不收敛,幂函数能够消除混合效应模型的异方差,校正广义混合效应模型的幂函数形式为

$$G_i = |u_{ij}|^{0.4995} \quad (23)$$

式中: u_{ij} 代表第*i*个样地第*j*棵树木的枝下高拟合值。

系数 R^2 ,且模型(1)可以将枝下高限制在(0, H]之间,因此选择模型(1)作为拟合枝下高模型的最优基础模型。

2.2 广义枝下高模型

使用随机效应模型的方法将不同预测变量的组合引入到基础模型中,依据避免多重共线性和过度参数化的原则,将拟合效果较好的林分变量组合列入表4。当选择大于对象木断面积和(BAL)、优势木高(HD)和冠幅(CW)等3个林分变量组合时,相较于其他的林分变量组合, R^2 值提升了0.1~0.13,MAE下降了0.17~0.23, RMSE下降了0.21~0.27,因此选择该变量组合。广义枝下高模型的形式为:

$$HCB = \frac{H}{1 + \exp(b_0 + b_1D + c_1HD + d_1BAL + d_2CW)} + \varepsilon \quad (20)$$

式中: b_0 、 b_1 、 c_1 、 d_1 、 d_2 为模型参数。

2.3 混合效应模型

在基础模型和广义模型基础上考虑样地效应对枝下高的影响。首先,分析基础模型,2个参数均添加随机效应模型收敛且拟合效果最好,得到以下模型:

$$HCB_{ij} = \frac{H_{ij}}{1 + \exp((b_0 + u_0) + (b_1 + u_1)D_{ij})} + \varepsilon_{ij} \quad (21)$$

式中: HCB_{ij} 、 H_{ij} 、 D_{ij} 分别为第*i*块样地第*j*株树的枝下高、树高、胸径; u_0 、 u_1 为随机效应参数。

对于广义枝下高模型,随机效应共有31种不同的组合方式,当分别引入4个和5个随机参数时,模型不收敛,因此表5中仅列出了随机参数相同时效果最优的组合。当随机参数组合为 b_0 、 b_1 、 d_2 时,AIC、BIC值最小,且LOGLIK达到最大值,且通过似然比检验,各随机参数组合之间有明显的不同($P < 0.0001$),因此选择 b_0 、 b_1 、 d_2 随机参数组合,得到模型:

2.4 模型比较

4个模型的固定参数、方差估计以及拟合统计量如表6所示,在不含随机效应模型方面,广义模型在拟合精度上高于基础模型, R^2 提升21.0%,MAE和RMSE分别降低18.1%和15.6%;在含随机效应模型方面,广义混合效应模型在拟合精度上略高于基础混合效应模型, R^2 提升1.1%,MAE和RMSE分别降低2.2%和1.8%。含随机效应模型的拟合精度与相应不含随机效应模型相比均有很大提高,含随机效应的基础混合效应模型较于不含随机效应的基础模型, R^2 提升32.62%,RMSE和MAE分别下降25.65%和27.87%,含随机效应的广义混合效应模型

表 4 不同预测变量组合模型拟合结果

Tab. 4 Model fitting results based on different prediction variables

变量组合 Variable combination	R^2	MAE	RMSE
—	0.567 0	1.284 0	1.638 6
大于对象木的断面积和-优势木高 BAL-HD	0.683 9	1.087 8	1.400 1
大于对象木的断面积和-林分密度 BAL-N	0.671 1	1.113 2	1.428 1
大于对象木的断面积和-冠幅 BAL-CW	0.687 4	1.080 3	1.392 3
大于对象木的断面积和-优势木高-冠幅 BAL-HD-CW	0.700 0	1.050 2	1.363 8
大于对象木的断面积和-高径比-优势木高 BAL-RHD-HD	0.690 0	1.070 5	1.386 5
大于对象木的断面积和-优势木高-林分密度 BAL-HD-N	0.690 0	1.088 7	1.386 5

注: —表示除胸径外, 无林分变量添加。Notes: — indicates the models without other stand variables except D .

表 5 广义非线性混合模型评价指标

Tab. 5 Evaluation indices of generalized nonlinear mixed model

随机效应参数 Random effect parameter	AIC	BIC	LOGLIK	LRT	P值 P value
—	8 167.35	8 207.88	-4 076.67		
b_0	7 987.10	8 027.63	-3 985.55	182.24	< 0.000 1
b_0, d_2	7 969.87	8 021.98	-3 975.93	19.24	< 0.000 1
b_0, b_1, d_2	7 815.16	7 890.44	-3 894.58	162.7	< 0.000 1

注: —表示无随机参数。Note: — indicates the models without random parameters.

表 6 模型参数估计和模型拟合统计量

Tab. 6 Parameter estimates and fitting statistics for each model

项目 Item	参数 Parameter	不含随机效应模型 Model without random effect		含随机效应模型 Model with random effect	
		基础模型 Base model	广义模型 Generalized model	基础混合效应模型 Base mixed effect model	广义混合效应模型 Generalized mixed effect model
固定参数 Fixed parameter	b_0	0.396 4	-1.201 6	0.020 2	-1.739 8
	b_1	0.020 0	-0.043 9	0.010 5	-0.024 7
	c_1		0.093 9		0.112 8
	d_1		-0.363 3		-0.276 2
	d_2		0.383 7		0.285 9
方差组成 Composition of variance	$\sigma_{b_0}^2$			0.416 6	0.135 4
	$\sigma_{b_1}^2$			0.000 2	0.000 3
	$\sigma_{d_2}^2$				0.054 6
	$\sigma_{b_0 b_1}$			-0.008 9	-0.000 7
	$\sigma_{b_0 d_2}$				-0.031 9
	$\sigma_{b_1 d_2}$				-0.003 0
拟合统计量 Fitting statistics	R^2	0.578 1	0.699 5	0.766 7	0.775 2
	MAE	1.284 0	1.051 8	0.926 2	0.905 5
	RMSE	1.617 5	1.364 9	1.202 6	1.180 5

注: $\sigma_{b_0}^2, \sigma_{b_1}^2, \sigma_{d_2}^2$ 分别为 b_0, b_1, d_2 的方差; $\sigma_{b_0 b_1}, \sigma_{b_0 d_2}, \sigma_{b_1 d_2}$ 分别为 b_0 与 b_1, b_0 与 d_2, b_1 与 d_2 的协方差; MAE 为平均绝对误差, RMSE 为均方根误差。
Notes: $\sigma_{b_0}^2, \sigma_{b_1}^2, \sigma_{d_2}^2$ are the variance of b_0, b_1, d_2 , respectively; $\sigma_{b_0 b_1}, \sigma_{b_0 d_2}, \sigma_{b_1 d_2}$ are the covariance of b_0 and b_1, b_0 and d_2, b_1 and d_2 , respectively; MAE is mean absolute error, RMSE is root mean square error.

较于不含随机效应的广义模型, R^2 增长了 10.82%, RMSE 和 MAE 分别减少了 13.51% 和 13.91%。综合评价指标, 含随机效应的广义混合效应模型拟合效

果优于其他 3 种模型。基础模型和广义模型以及含随机效应的基础混合效应模型和广义混合效应模型的残差图分别对应图 1 中的 A、B、C、D, 异方差校正

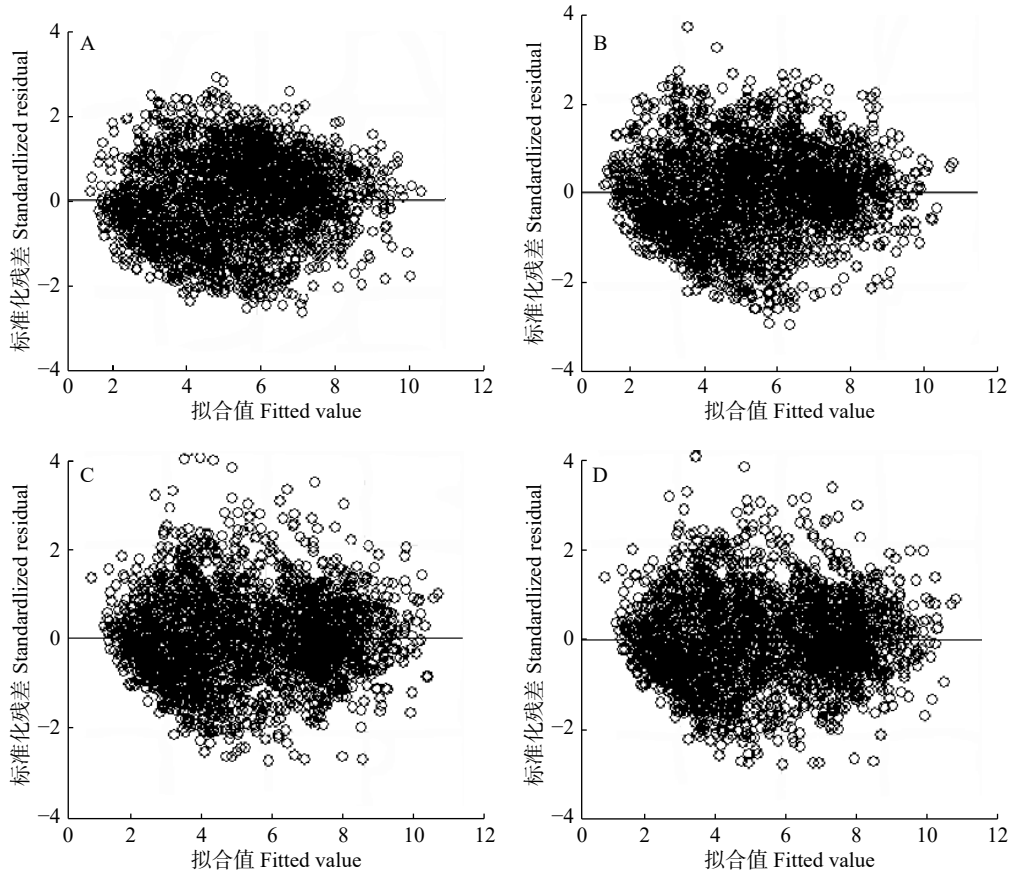


图1 基本模型(A)、广义模型(B)、基础混合效应模型(C)和广义混合效应模型(D)的残差分布图

Fig. 1 Residual plots of basic model (A), generalized model (B), basic mixed effect model (C), and generalized mixed effect model (D)

后都没有显示极不规则的形状。

2.5 混合效应模型应用抽样设计比较

本研究对随机效应的基础混合效应模型和广义混合效应模型分别校正,利用检验数据进行4种抽样,即随机抽取样本、抽取样本最大值、抽取样本最小值和抽取样本均值。每个样地抽取1~8个样本,即实测抽取样木的枝下高,基于公式(13)和表6的参数估计值计算随机参数估计值,具体采用R语言的矩阵编程来实现。为了避免得到随机参数估计值的偶然性,重复100次抽样校正过程。通过分析不同抽样设计下样本量与预测精度(RMSE)的关系(图2和表7)可以发现:含随机效应的基础混合效应模型和广义混合效应模型的4种抽样方式均呈现RMSE总体下降的趋势,且下降趋势相似。当抽取1~4株样木时,下降速度较快,然后下降缓慢,因此当应用含随机效应的基础混合效应模型和广义混合效应模型进行预测时,根据抽样成本和精度的要求,建议抽取4个样本。当比较抽取4个样本的4种抽样方案时,对于基础混合效应模型,抽取最小值的样本精度最高,而对于广义混合效应模型,随机抽取样本的精度最高。

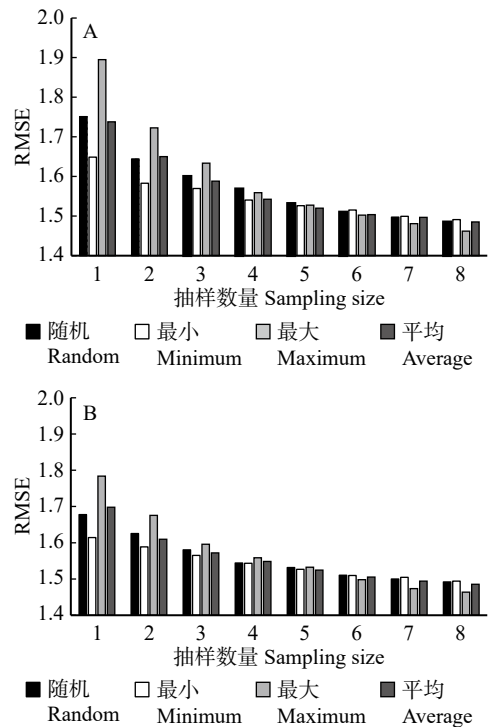


图2 不同抽样设计下基础混合效应模型(A)和广义混合效应模型(B)的预测精度

Fig. 2 Prediction accuracy of basic mixed effect model (A) and generalized mixed effect model (B) based on different sampling designs

表 7 4 种抽样方式 RMSE 对比

Tab. 7 Comparison of RMSE of four sampling methods

样本数 Sample size	含随机效应基础模型 Basic model with random effect				含随机效应广义模型 Generalized model with random effect			
	随机抽取 Random sampling	最小值 Minimum value	最大值 Maximum value	平均值 Average value	随机抽取 Random sampling	最小值 Minimum value	最大值 Maximum value	平均值 Average value
1	1.751 2	1.648 7	1.894 7	1.737 9	1.635 1	1.691 9	1.704 4	1.684 3
2	1.644 4	1.583 0	1.722 5	1.650 1	1.588 8	1.624 7	1.636 6	1.636 4
3	1.600 6	1.569 5	1.633 5	1.588 3	1.550 6	1.580 8	1.569 2	1.572 5
4	1.569 0	1.540 6	1.558 9	1.542 5	1.524 3	1.533 7	1.537 5	1.545 0
5	1.532 2	1.526 4	1.527 6	1.520 1	1.513 2	1.514 9	1.514 3	1.511 9
6	1.510 4	1.515 4	1.502 3	1.503 7	1.493 3	1.482 3	1.490 7	1.485 2
7	1.497 5	1.499 5	1.481 0	1.496 9	1.479 3	1.466 0	1.471 8	1.472 7
8	1.485 5	1.491 0	1.462 0	1.485 3	1.468 1	1.448 2	1.450 7	1.446 5

3 讨 论

将优势木高作为立地质量的代表引入到模型中,表现了对枝下高显著的相关性。优势木高反映了林分生长和收获能力,优势木高与树木生长和林分发育之间存在着很强的关系。从表 6 中可以看出优势木高的参数估计值是负号,这可能因为红松林竞争压力大,树木由于自然整枝的原因导致下层枝条自然死亡,随着优势木高的增加,枝下高增加。

在竞争变量中,选择了大于对象木断面积和(BAL)和冠幅(CW),这 2 个变量与枝下高存在显著的相关性,其他的研究也可以证明,如 Fu 等^[10]在研究蒙古栎(*Quercus mongolica*)枝下高模型时得到相同结论。大于对象木的断面积和(BAL)表述样地内树木的竞争指标,在冠长率模型中经常出现,与相邻木相比,BAL 越小,树木的直径越大,对象木的竞争能力越强,就越能获得光照、水和土壤养分。冠幅变量可表示样地内的密度竞争指标,在表 6 中可以看出大于对象木的断面积和的参数估计是负值,表示随着大于对象木断面积之和的增加,树木的竞争压力越大,枝下高变大;冠幅的参数估计是正值,表示随着冠幅的增加,对象木的竞争降低,枝下高变小。

混合效应模型是分析分组或分层数据的一种方法,在本研究中混合模型方法能够解释样地之间枝下高的变化,并能提高模型的预测精度。在模型应用计算随机参数过程中,随着二次抽样样本数量的增加,预测精度也提高,所耗费抽样成本也增加,因此在应用混合效应模型时,确定合适的抽样数量计算随机参数是非常必要的。

目前关于枝下高模型的研究较少,尤其是混合效应枝下高模型应用时随机参数的不同抽样方案比较仅见 2 篇。Sharma 等^[17]在研究挪威云杉(*Picea*

abie)和欧洲山毛榉(*Fagus sylvatica*)枝下高模型时,随机抽取 1~10 株树进行预估随机参数,考虑预测精度与成本,建议随机抽取 4 株树进行抽样校正;Fu 等^[10]在研究蒙古栎枝下高模型抽样数量与预测精度问题时,也建议随机抽取 4 株树进行校正,这与本研究的结果基本一致。而 4 种抽样方案对比来说略有不同,对于基础混合效应模型,抽取最小值的样本最好,对于广义基础混合效应模型,随机抽取样本最好。

4 结 论

本文从 8 个候选基础模型中得到最优基础模型 Logistic 模型,加入立地质量因子(HD)和竞争因子(BAL 和 CW)后,显著提高了基础模型的拟合精度。随后将基础模型和广义模型分别引入随机效应,模型的拟合精度进一步提高,可以表明林分变量以及样地效应对枝下高模型的影响是显著的。含随机效应的广义混合效应模型在拟合效果与预测精度方面优于其他 3 种模型,建议此模型作为人工红松枝下高预测模型。

参 考 文 献

- [1] Hasenauer H, Monserud R. A crown ratio model for Austrian forests[J]. *Forest Ecology and Management*, 1996, 84(1/3): 49-60.
- [2] Sharma R P, Bilek L, Vacek, Z, et al. Modelling crown width-diameter relationship for Scots pine in the central Europe[J]. *Trees*, 2017, 31(6): 1875-1889.
- [3] Kuprevicius A, Auty D, Achim A, et al. Quantifying the influence of live crown ratio on the mechanical properties of clear wood[J]. *Forestry: An International Journal of Forest Research*, 2013, 86(3): 361-369.
- [4] Monserud R A, Sterba H. A basal area increment model for individual trees growing in even- and uneven-aged forest stands in Austria[J]. *Forest Ecology and Management*, 1996, 80(1/3): 57-80.
- [5] 吕乐,董利虎,李凤日. 黑龙江省东部地区天然椴树单木冠幅预测模型[J]. *东北林业大学学报*, 2019, 47(7): 37-42.

- Lü L, Dong L H, Li F R. Individual tree crown width prediction models for natural *Tilia tuan* in eastern Heilongjiang Province[J]. Journal of Northeast Forestry University, 2019, 47(7): 37–42.
- [6] Sharma R P, Vacek Z, Vacek S. Individual tree crown width models for Norway spruce and European beech in Czech Republic[J]. *Forest Ecology and Management*, 2016, 366: 208–220.
- [7] Fu L Y, Sharma R P, Hao K J, et al. A generalized interregional nonlinear mixed-effects crown width model for Prince Rupprecht larch in northern China[J]. *Forest Ecology and Management*, 2017, 389: 364–373.
- [8] Crecente-Campo F, Álvarez-González J G, Castedo-Dorado F, et al. Development of crown profile models for *Pinus pinaster* Ait. and *Pinus sylvestris* L. in northwestern Spain[J]. *Forestry*, 2013, 86(4): 481–491.
- [9] Ritson P, Sochacki S. Measurement and prediction of biomass and carbon content of *Pinus pinaster* trees in farm forestry plantations, southwestern Australia[J]. *Forest Ecology and Management*, 2003, 175(1/3): 103–117.
- [10] Fu L Y, Zhang H R, Sharma R P, et al. A generalized nonlinear mixed-effects height to crown base model for Mongolian oak in northeast China[J]. *Forest Ecology and Management*, 2017, 384: 34–43.
- [11] McRoberts R, Hahn J T, Hefty G J, et al. Variation in forest inventory field measurements[J]. *Canadian Journal of Forest Research*, 1994, 24(9): 1766–1770.
- [12] Temesgen H, Lemay V, Mitchell S J. Tree crown ratio models for multi-species and multi-layered stands of southeastern British Columbia[J]. *The Forestry Chronicle*, 2005, 81(1): 133–141.
- [13] Ritchie M W, Hann D W. Equations for predicting height to crown base for fourteen tree species in southwest Oregon[R]. Corvallis: Oregon State University, 1987.
- [14] Rijal B, Weiskittel A R, Kershaw J A. Development of regional height to diameter equations for 15 tree species in the North American Acadian Region[J]. *Forestry: An International Journal of Forest Research*, 2012, 85(3): 379–390.
- [15] Yang Y Q, Huang S M. Effects of competition and climate variables on modelling height to live crown for three boreal tree species in Alberta, Canada[J]. *European Journal of Forest Research*, 2018, 137(2): 153–167.
- [16] 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.
Duan G S, Li X D, Feng Y, et al. Generalized nonlinear mixed-effects crown base height model of *Larix principis-rupprechtii* natural secondary forests[J]. Journal of Nanjing Forestry University (Natural Science Edition), 2018, 42(2): 170–176.
- [17] Sharma R P, Vacek Z, Vacek S, et al. Modelling individual tree height to crown base of Norway spruce (*Picea abies* (L.) Karst.) and European beech (*Fagus sylvatica* L.)[J]. *PLoS ONE*, 2017, 12(10): e0186394.
- [18] Calama R, Montero G. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain[J]. *Canadian Journal of Forest Research*, 2004, 34(1): 150–163.
- [19] Yang Y Q, Huang S M, Meng S X, et al. A multilevel individual tree basal area increment model for aspen in boreal mixedwood stands[J]. *Canadian Journal of Forest Research*, 2009, 39(11): 2203–2214.
- [20] Temesgen H, Monleon V J, Hann D W. Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests[J]. *Canadian Journal of Forest Research*, 2008, 38(3): 553–565.
- [21] Crecente C F, Tomé M, Soares P, et al. A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain[J]. *Forest Ecology and Management*, 2010, 259(5): 943–952.
- [22] Sharma R P, Breidenbach J. Modeling height-diameter relationships for Norway spruce, Scots pine, and downy birch using Norwegian national forest inventory data[J]. *Forest Science and Technology*, 2015, 11(1): 44–53.
- [23] 段光爽, 李学东, 冯岩, 等. 华北落叶松天然次生林树高曲线的混合效应模型[J]. 南京林业大学学报(自然科学版), 2018, 42(2): 163–169.
Duan G S, Li X D, Feng Y, et al. Developing a height-diameter relationship model with mixed random effects for *Larix principis-rupprechtii* natural secondary forests[J]. Journal of Nanjing Forestry University (Natural Science Edition), 2018, 42(2): 163–169.
- [24] 李婉婷, 姜立春, 万道印. 基于混合效应的兴安落叶松树高与胸径关系模拟[J]. 植物研究, 2014, 34(3): 343–348.
Li W T, Jiang L C, Wan D Y. Simulation of height-diameter relationships for *Larix gmelinii* based on mixed effects[J]. Bulletin of Botanical Research, 2014, 34(3): 343–348.
- [25] 雷相东, 李永慈, 向玮. 基于混合模型的单木断面积生长模型[J]. 林业科学, 2009, 45(1): 74–80.
Lei X D, Li Y C, Xiang W. Individual basal area growth model using multi-level linear mixed model with repeated measures[J]. Scientia Silvae Sinicae, 2009, 45(1): 74–80.
- [26] 李春明. 基于两层次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.
- [27] Calama R, Montero G. Multilevel linear mixed model for tree diameter increment in stone pine (*Pinus pinea*): a calibrating approach[J]. *Sliva Fennica*, 2005, 39(1): 394.
- [28] 李想, 董利虎, 李凤日. 基于联立方程组的人工樟子松枝下高模型构建[J]. 北京林业大学学报, 2018, 40(6): 9–18.
Li X, Dong L H, Li F R. Building height to crown base models for Mongolian pine plantation based on simultaneous equations in Heilongjiang Province of northeastern China[J]. Journal of Beijing Forestry University, 2018, 40(6): 9–18.
- [29] Wykoff W R, Crookston N L, Stage A R. User's guide to the stand prognosis model[R]. Ogden: Forest Service, United States Department of Agriculture, 1982.
- [30] Popoola F S, Adesoye P O. Crown ratio models for *Tectona grandis* (Linn. f) stands in Osho Forest Reserve, Oyo State, Nigeria[J]. *Journal of Forest & Environmental Science*, 2012, 28(2): 63–67.
- [31] Walters D K, Hann D W. Taper equations for six conifer species in southwest Oregon[M]. Corvallis: Oregon State University, 1986..
- [32] 韩斐斐, 姜立春. 基于树干不同高度直径的落叶松立木材积方程[J]. 东北林业大学学报, 2017, 45(4): 65–69.
Han F F, Jiang L C. Tree volume function based on diameter at different relative heights of Dahurian larch[J]. Journal of Northeast Forestry University, 2017, 45(4): 65–69.

(责任编辑 冯秀兰
责任编委 张会儒)