

周晏平, 雷泽勇, 赵国军, 等. 沙地樟子松不同树高-胸径模型比较分析 [J]. 华南农业大学学报, 2019, 40(3): 75-81.
ZHOU Yanping, LEI Zeyong, ZHAO Guojun, et al. Comparing different height-diameter models of *Pinus sylvestris* var. *mongolica* in sandy land[J]. Journal of South China Agricultural University, 2019, 40(3): 75-81.

沙地樟子松不同树高-胸径模型比较分析

周晏平¹, 雷泽勇¹, 赵国军², 韩艳刚¹

(1 辽宁工程技术大学 环境科学与工程学院, 辽宁 阜新 123000; 2 辽宁省固沙造林研究所, 辽宁 阜新 123000)

摘要:【目的】比较不同树高 (H)-胸径 (D) 模型精度, 确定适合章古台地区樟子松 *Pinus sylvestris* var. *mongolica* 的 H - D 模型。【方法】以 Sibbesen 模型为基础模型, 将优势木平均高 (H_T)、胸高断面积 (A_B) 和平方平均胸径 (D_{QM}) 3 个林分变量以不同组合加入基础模型中, 分别建立了 H - D 的基础模型 (1 个) 和广义模型 (3 个) 及对应的基础混合模型 (1 个) 和广义混合模型 (3 个)。对固定效应模型平均水平预测 (FPA)、混合模型的总体平均响应预测 (MPA) 和主体响应预测 (MPS) 的精度进行比较。对混合模型在使用随机抽取样本木和抽取平均木 (胸径接近平均值的样本) 2 种抽样方案计算随机参数时分析 MPS 精度与样本数量的关系。【结果】表征樟子松 H - D 关系的 4 种固定效应模型中, 含 H_T 和 A_B 的广义模型拟合精度最高, Akaike 信息量准则 (AIC)=2 167.7, Bayesian 信息量准则 (BIC)=2 196.3。相同预测变量的各模型预测精度均表现为: MPS>FPA>MPA, 仅含预测变量 D 的模型的 3 种预测精度差异最大。广义模型、广义混合模型、基础混合模型预测精度差异不大。使用验证数据检验模型精度时, 每块标准地中随机抽取 3 株样本木计算基础混合模型随机参数时, 该模型精度提升最为明显, MAE 和 RMSE 分别降低了 57.97% 和 57.63%; 而广义混合模型精度随抽取样本木数量的增多未出现大的变化。【结论】含有林分变量优势木平均高、胸高断面积的广义模型和基础混合模型均能较好地预测沙地樟子松人工林的单木树高。此外, 利用混合模型预测树高时, 推荐在标准地中随机抽取 3 株林木测量其树高, 并依此来计算随机参数。

关键词: 樟子松; 人工林; 树高-胸径模型; 林分变量; 抽样方案
中图分类号: S757 文献标志码: A 文章编号: 1001-411X(2019)03-0075-07

Comparing different height-diameter models of *Pinus sylvestris* var. *mongolica* in sandy land

ZHOU Yanping¹, LEI Zeyong¹, ZHAO Guojun², HAN Yangang¹

(1 College of Environmental Science and Engineering, Liaoning Technology University, Fuxin 123000, China;
2 Liaoning Province Sand-Fixation and Afforestation Research Institute, Fuxin 123000, China)

Abstract: 【Objective】To compare the accuracy of different height(H)-diameter (D) models to determine the optimal models for *Pinus sylvestris* var. *mongolica* in Zhanggutai area. 【Method】Sibbesen model was used as the basic model. Dominant height (H_T), stand basal area (A_B), and quadratic mean diameter (D_{QM}) with different combinations were added into Sibbesen model. We established one basic, three generalized, one basic mixed and three generalized mixed H - D models. The accuracies of population-averaged prediction (FPA) of fixed effects models, and mean response prediction (MPA) and specific-plot prediction (MPS) of mixed effects models were compared. For mixed models, two sampling designs, random sampling and medium-diameter tree sampling were used for random parameters estimation, and the relationship between MPS accuracy and sample size was analyzed. 【Result】In four fixed H - D models, the generalized model with H_T and A_B has the highest prediction precision. Akaike's information criterion (AIC) is 2 167.7. Bayesian information criterion (BIC) is

2 196.3. Models with the same predictor variables have precision in order of MPS> FPA>MPA, and models with D as the only variable have the largest variation among three types of prediction. There are little difference in prediction accuracy among generalized models, generalized mixed models and basic mixed model. Using three randomly selected sample trees per plot to estimate random parameters of basic mixed model results in the highest model precision based on the validation data, and MAE and RMSE decrease by 57.97% and 57.63% respectively. The accuracies of generalized mixed models do not change significantly with the increase of sample size. 【Conclusion】 Both generalized model including H_T and A_B and basic mixed model can well predict tree height for *P. sylvestris* var. *mongolica*. We recommend to randomly select three sample trees per plot measuring tree heights for parameters estimation of mixed models, and calculating random parameters.

Key words: *Pinus sylvestris* var. *mongolica*; plantation; height-diameter model; stand variable; sampling design

树高和胸径是衡量林分结构和质量的最重要的 2 个指标^[1]。其中, 树高数据对估计树木材积、生物量, 描述林分生长变化不可或缺^[2-3]。但树高的测量耗时耗力、误差较大^[4], 而胸径的测量相对容易、精确、花费少。故实际操作中, 往往先测定少数树木的树高和胸径, 建立 $H-D$ 的函数关系, 再以胸径估计剩余树木的树高^[5-8]。由于立地、林龄和管理措施等因素影响, 预测树高时需考虑 $H-D$ 关系在不同林分间的差异^[9]。目前实现不同林分中树高预测的方法主要包括: 1) 采用足够数量的树高实测值构建的 $H-D$ 模型^[10]; 2) 针对具体样地分别建立适用于小面积相似林分的 $H-D$ 模型^[11]; 3) 建立含有表征林分特征的分变量或树木因子的 $H-D$ 模型^[1, 4, 8]; 4) 建立能够反映不同林分 (或样地) 间差异的混合模型^[5-7, 12]。前 2 种方法在森林规模较小时可以实现, 但是在大范围林分中所需工作量太大, 较难实现。后 2 种方法中, 广义模型具有较好的生物学解释和较高的精度^[11, 13-14], 混合模型在处理纵向数据和重复测量数据方面优势明显, 且精度较高^[15]。目前林业调查多使用后 2 种方法, 或结合 2 种方法建立广义混合模型预测树高^[6, 14]。

大量关于 $H-D$ 模型的研究表明: 不同树种的 $H-D$ 关系模型有明显差异^[1-2, 5-8]。多数是固定效应模型中的广义模型精度高于基础模型, 混合模型精度高于固定效应模型^[5, 14]。但基础混合模型和广义混合模型间差异却随树种不同而变化。例如, 对于欧洲云杉 *Picea abies*、欧洲赤松 *Pinus sylvestris* 和白桦 *Betula pubescens* 是广义混合模型精度高于基础混合模型^[16], 但对于颤杨 *Populus tremuloides* 却是基础混合模型预测精度高于广义混合模型^[7]。还有研究发现对于红松 *Pinus koraiensis*^[5]、杉木 *Cunninghamia lanceolata*^[6] 和落叶松 *Larix gmelinii*^[17], 2 种混合模型间差异不大。沙地樟子松 *Pinus sylvestris* var. *mongolica* 是我国北方地区主要造林树种之一, 在防沙治沙和改善环境方面具有重要生态价值。迄

今为止, 该树种已经推广到我国北方地区的 19 省、自治区、直辖市^[18]。随着造林面积的扩大, 如何合理经营沙地樟子松人工林? 不断扩大的沙地樟子松人工林资源的清查工作如何快速高效地进行? 解决这 2 个棘手的问题, 必须有精确的沙地樟子松 $H-D$ 关系模型, 而该方面的研究鲜见报道。鉴于此, 本文筛选合适的 $H-D$ 基础模型, 检验林分变量中优势木平均高 (H_T)、胸高断面面积 (A_B) 和平方平均胸径 (D_{QM}) 对 $H-D$ 模型的影响, 建立 $H-D$ 基础混合模型和广义混合模型, 并确定计算随机参数时的抽样方案, 以期对沙地樟子松林的生长预测和林分管理、资源清查提供科学依据。

1 材料与方法

1.1 研究区概况

研究区位于辽宁省阜新市章古台镇 (42°39'~42°43'N, 122°22'~122°33'E)。该地区位于科尔沁沙地东南缘, 属于半湿润气候, 近 40 年 (1973—2015) 年平均气温 6.99 °C; 年降水量平均 483.63 mm, 最大 713.2 mm, 最小 285.9 mm, 约 67.0% 集中在 6—8 月; 年均蒸发量 1 762 mm。最暖月份平均气温 -12.91 °C, 最冷月份平均气温 23.58 °C, 大于 5 °C 有效积温 2 570 °C, 全年无霜期 150~160 d。土壤主要为生草固定风沙土和流动风沙土。土壤养分含量比较低, 有机质质量分数为 0.6~1.6 g/kg, 全 N 质量分数为 0.17~0.38 g/kg, P₂O₅ 质量分数为 0.12~0.27 g/kg, pH 为 6.8~7.0^[18]。植物区系属于蒙古区系西辽河小区, 主要代表性植物有色木槭 *Acer mono*、山里红 *Crataegus pinnatifida* var. *major*、榆树 *Ulmus pumila*、大果榆 *U. macrocarpa*、山杏 *Armeniaca sibirica*、胡枝子 *Lespedeza bicolor*、中华隐子草 *Cleistogenes chinensis* 等^[19]。

1.2 数据收集

数据来源于 2016 年设置于章古台镇 35 块临时标准地, 标准地规模 20 m×20 m。所选标准地涵

盖了沙地樟子松不同的林龄 (A)、立地条件和林分密度 (N)。标准地的林分年龄为 13~62 年, 立地覆盖了平坦沙地 (11 块) 及各个坡向的坡上 (7 块)、坡中 (11 块) 和坡下 (6 块) 条件。测量标准地内沙地樟子松的树高 (H)、胸径 (D); 以每公顷 100 棵胸径最大沙地樟子松的平均树高为优势木高^[5, 7], 按标准地面积比例确定株数。从测量的数据中, 选择健康、

未受损害且 $D \geq 5$ cm 的沙地樟子松样本数据用于模型建立。本研究共获得 1 070 组 H - D 数据, 分为 2 个数据集, 一个为由 28 块标准地中的 866 棵 (80%) 沙地樟子松构成的建模数据, 另一个为由其余 7 块标准地的 204 棵 (20%) 沙地樟子松组成的检验数据, 表 1 为数据基本概况。

表 1 沙地樟子松标准地的林分与树木指标统计
Table 1 Summary statistics of stand and tree indices of *Pinus sylvestris* var. *mongolica* sample plot

变量 ¹⁾ Variable	总数据 Total data			建模数据 Modeling data			检验数据 Validation data		
	$\bar{x} \pm SE$	最小值 Min.	最大值 Max.	$\bar{x} \pm SE$	最小值 Min.	最大值 Max.	$\bar{x} \pm SE$	最小值 Min.	最大值 Max.
H/m	8.5±0.1	2.1	16.1	8.3±0.1	2.1	16.1	9.2±0.2	2.2	14.3
D/cm	14.6±0.2	5.0	29.1	14.3±0.2	15.0	29.1	15.8±0.3	5.1	28.6
D_{range}/cm	10.9±0.4	5.7	16.5	10.9±0.4	5.7	16.5	11.0±0.9	7.3	14.2
D_{max}/cm	22.0±0.9	10.0	29.1	22.0±1.0	10.0	29.1	21.9±2.1	11.9	28.6
D_{min}/cm	11.0±0.8	1.9	18.0	11.1±0.9	1.9	18.0	10.9±1.9	3.3	18.0
A/a	36±2	13	62	36±3	13	62	36±6	13	56
$N/(trees \cdot hm^{-2})$	842±80	300	2 500	862±96	300	2 500	764±112	450	1 175
H_T/m	9.7±0.5	3.2	13.6	9.7±0.6	3.2	13.6	10.0±1.1	4.1	12.7
$A_B/(m^2 \cdot hm^{-2})$	15.26±1.04	2.92	33.34	15.16±1.19	4.28	33.34	15.65±2.31	2.92	21.97
D_{QM}/cm	16.7±0.9	7.0	24.0	16.6±1.0	7.0	23.9	16.8±1.9	7.7	21.9

1) D_{range} 、 D_{max} 和 D_{min} 分别为样地内胸径的分布范围值、最大值和最小值; A : 树龄; N : 林分密度; H_T : 优势木平均高; A_B : 胸高断面积; D_{QM} : 平方平均胸径

1) D_{range} , D_{max} and D_{min} are the distribution range value, maximum and minimum of DBH in the plot; A : Age of stand; N : Stand density; H_T : Dominant height; A_B : Stand basal area; D_{QM} : Quadratic mean diameter

1.3 基础模型

在以往报道的基础上选择了 30 个非线性 H - D 候选基础模型^[8, 20], 使用最小二乘法分别拟合总数据中沙地樟子松的树高-胸径关系, 并根据评价指标: 决定系数 (R^2)、平均绝对误差 (Mean absolute error, MAE) 和均方根误差 (Root mean square error, RMSE) 选出较好的 H - D 基础模型。 R^2 、MAE 和 RMSE 的计算公式为:

$$R^2 = 1 - \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (H_{ij} - \hat{H}_{ij})^2}{\sum_{i=1}^m \sum_{j=1}^{n_i} (H_{ij} - \bar{H}_{ij})^2}, \quad (1)$$

$$MAE = \frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} |H_{ij} - \hat{H}_{ij}|, \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^m \sum_{j=1}^{n_i} (H_{ij} - \hat{H}_{ij})^2}, \quad (3)$$

式中: H_{ij} 、 $\hat{H}_{ij}$ 和 $\bar{H}_{ij}$ 分别为树高观测值、预测值和平均观测值; m 为标准地数量; n 为样本木总量。

1.4 广义模型

基础模型是在总体平均水平上对 H - D 关系的描述, 不能反映 H - D 关系在不同林分中存在的实际差异^[1, 17]。基础模型中加入林分变量构建广义模型可以提升模型的生物学意义和预测精度。本文在确定基础模型后, 检验了优势木平均高 (H_T)、胸高断面积 (A_B) 和平方平均胸径 (D_{QM}) 对 H - D 模型的影响 (加入林分变量 1~3 个)。

1.5 混合模型

确定基础模型和广义模型后, 采用非线性混合模型描述不同标准地内的 H - D 关系, 模型的一般表达式为:

$$\begin{cases} y_i = f(\beta, u_i, x_i) + \varepsilon_i, \\ \mathbf{R}_i = \sigma^2 \mathbf{G}_i^{0.5} \mathbf{T}_i \mathbf{G}_i^{0.5}, \end{cases} \quad (4)$$

式中: y_i 是第 i 块标准地的树高测量值, $i=1, 2, 3, \dots, m$; β 为固定效应参数; u_i 为随机效应参数, 假定 u_i 服从 Ψ 正态分布, 即 $u_i \sim N(0, \Psi)$, Ψ 为随机效应方差-协方差矩阵; x_i 为自变量测量值; f 为包含模

型参数的 x_i 与 y_i 之间的函数关系; ε_i 为误差项, 假定 $\varepsilon_i \sim N(0, \mathbf{R}_i)$; 假定 u_i 和 ε_i 之间相互独立; $\mathbf{R}_i$ 为样地内方差-协方差结构矩阵; σ^2 为误差扩散比例因子, 即残差方差; $\mathbf{G}_i$ 为描述方差异质性的对角矩阵; $\mathbf{\Gamma}_i$ 为自相关结构, 本文调查数据未涉及误差自相关性, 故用单位矩阵代替。

1.6 模型的拟合与检验

固定效应模型和混合模型分别使用 R 软件中 nls 函数和 nlme 程序包中极大似然估计方法实现。拟合优度指标为 Akaike 信息量准则 (Akaike's information criterion, AIC)、贝叶斯信息量准则 (Bayesian information criterion, BIC) 和似然函数值 (log-likelihood, LL)。

模型的预测类型包括: 固定效应模型的平均水平预测 (FPA), 混合模型的总体平均响应预测 (MPA) 与主体响应预测 (MPS)。其中 MPA 仅使用混合模型的固定效应部分, 预测时令随机参数为 0。对未包含在建模数据中的林木进行 MPS 时, 需使用经验线性最优无偏估计方法计算随机效应参数值^[11]。

$$\hat{u}_i = \hat{\Psi} \mathbf{Z}_i^T (\mathbf{Z}_i \hat{\Psi} \mathbf{Z}_i^T + \hat{\mathbf{R}}_i)^{-1} \mathbf{e}_i, \tag{5}$$

式中: $\hat{u}_i$ 为随机参数最优无偏估计值, $\hat{\Psi}$ 为随机效应方差-协方差矩阵, $\hat{\mathbf{R}}_i$ 为标准地内方差-协方差结构矩阵; $\mathbf{Z}_i$ 为随机参数的偏导数设计矩阵, $\mathbf{T}$ 表示转置矩阵; $\mathbf{e}_i$ 为 MPA 时所得的残差。

通常情况下, 用于计算随机参数的样本木数量越多, 预测效果越好, 但调查成本会提高。为了选出工作量较少且精度较高的样本木抽样方案, 本文对比了以下抽样方案: 1) 每块标准地随机抽取 1~8 棵树木, 测量其树高; 2) 参照平均胸径, 每块标准地抽取 1~8 棵平均木, 测量其树高。

评价模型精度时, 基于建模数据计算了不同模

型不同预测类型时的 R^2 、MAE 和 RMSE。使用检验数据计算 MAE 和 RMSE 来反映模型预测精度。

2 结果与分析

2.1 构建的基础模型和广义模型

拟合总数据时, 30 个基础模型的 R^2 、MAE 和 RMSE 差异很小。依据参数最少、效果最好的原则, 选择式 (6) 的 Sibbesen 模型^[17] 作为樟子松的 H - D 基础模型。在式 (6) 基础上, 将 H_T 、 A_B 和 D_{QM} 以不同的组合形式加入式 (6) 中, 式 (7)~(9) 为分别加入不同数量林分变量时的最优形式, 即广义模型。

$$H_{ij} = 1.3 + \beta_1 D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \tag{6}$$

$$H_{ij} = 1.3 + (\beta_1 + \beta_4 H_{Ti}) D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \tag{7}$$

$$H_{ij} = 1.3 + (\beta_1 + \beta_4 H_{Ti} + \beta_5 A_{Bi}) D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \tag{8}$$

$$H_{ij} = 1.3 + (\beta_1 + \beta_4 H_{Ti} + \beta_5 A_{Bi} + \beta_6 D_{QMi}) D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \tag{9}$$

式中: $\beta_1 \sim \beta_6$ 为模型参数; H_{ij} 和 D_{ij} 分别为第 i 块标准地中第 j 棵树的树高和胸径; H_{Ti} 、 A_{Bi} 和 D_{QMi} 分别为第 i 块标准地优势木平均高、胸高断面面积和林分平方平均胸径。

结果 (表 2) 表明: 加入林分变量后, 广义模型的拟合优度极显著高于基础模型 ($P < 0.01$), 且模型的精度明显提高, 模型的残差方差 (σ^2)、AIC 和 BIC 均减小, LL 值明显增大。只加入 1 个林分变量时, 加入 H_T 的式 (7) 的效果最好, 此时模型的统计指标变化较大。加入 2 个林分变量时, 含 H_T 和 A_B 的式 (8) 精度进一步提高, 且与式 (7) 差异极显著 ($P < 0.01$)。含 3 个林分变量的式 (9) 精度低于式 (8)。

表 2 基础和广义模型各参数估计结果及统计量

Table 2 Parameter estimates and evaluation statistics of basic and generalized models

公式 Formula	β_1	β_2	β_3	β_4	β_5	β_6	σ^2	AIC	BIC	LL
(6)	14.242 3	-27.039 5	1.784 9				3.305 1	3 498.3	3 517.4	-1 745.2
(7)	-0.105 1	-5.428 3	1.442 1	1.073 6			0.724 7	2 184.8	2 208.6	-1 087.4
(8)	0.190 2	-4.472 6	1.217 7	1.112 6	0.055 7		0.709 8	2 167.7	2 196.3	-1 077.9
(9)	0.404 3	-4.646	1.222 3	1.177 7	0.045 9	-0.034 5	0.709 3	2 168.2	2 201.5	-1 077.1

2.2 构建的混合模型

拟合基础混合模型时, 只含 1 个随机参数的模型拟合优度最高。拟合广义混合模型时, 模型含有 2 个随机参数时的拟合优度略高于含 1 个随机参数

的模型, 但无显著差异 ($P > 0.05$)。因此, 4 个混合模型中均只含 1 个随机参数, 具体见式 (10)~(13)。

$$H_{ij} = 1.3 + (\beta_1 + u_i) D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \tag{10}$$

$$H_{ij} = 1.3 + [\beta_1 + (\beta_4 + u_i)H_{Ti}]D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \quad (11)$$

$$H_{ij} = 1.3 + [\beta_1 + (\beta_4 + u_i)H_{Ti} + \beta_5 A_{Bi}]D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \quad (12)$$

$$H_{ij} = 1.3 + [\beta_1 + (\beta_4 + u_i)H_{Ti} + \beta_5 A_{Bi} + \beta_6 D_{QMi}]D_{ij}^{\beta_2 D_{ij}^{-\beta_3}} + \varepsilon_{ij}, \quad (13)$$

式中： u_i 为随机效应参数；其余同上。图 1 为式 (10) 的残差分布图 [式 (11)~(13) 与此相似而未列出]，从图中未发现模型存在明显的方差异质性，因此不考虑对混合模型进行异方差处理。式 (10)~(13) 的 AIC、BIC 及 LL 的变化规律与 4 个固定模型相一致。式 (10) 的随机参数方差 (σ_i^2) 在 4 个混合模型中最大，说明式 (10) 中参数 β_1 最不稳定，在不同标准地间变化最大。

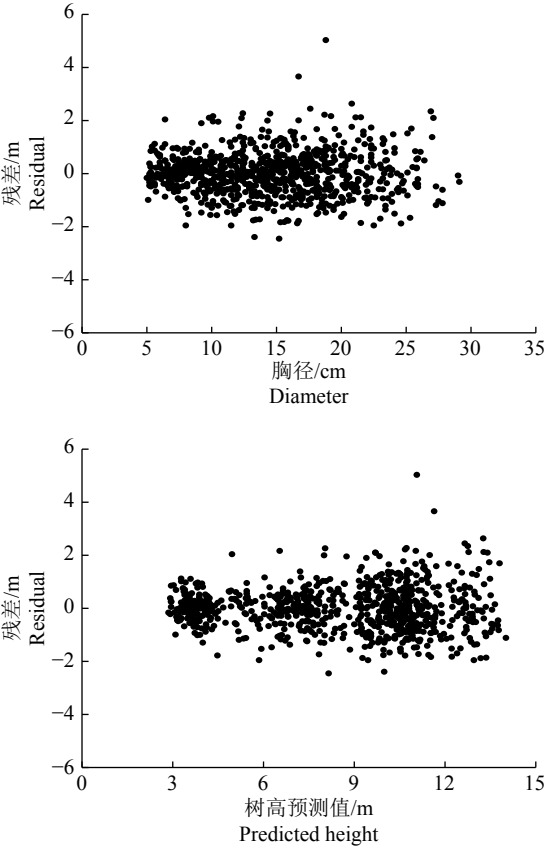


图 1 混合模型残差分布图
Fig. 1 Distribution of residuals for mixed model

由表 2 和表 3 可知，混合模型的拟合精度均高于相应的固定效应模型。加入随机参数对 H - D 基础模型改良效果最明显，混合模型式 (10) 的 AIC 和 BIC 较基础模型式 (6) 分别降低了 34.70% 和 34.38%，LL 值升高了 34.84%。说明混合模型可以提高 H - D 模型的预测精度，随机参数可以描述不同立地条件、样地间树高和胸径关系的差异。广义混合模型与对应的广义模型间都存在极显著差异 ($P<0.01$)，但精度提高的幅度较小 (AIC、BIC 和 LL 变化幅度低于 1.50%)。

2.3 模型评价与对比

2.3.1 模型拟合与预测精度评价 拟合结果 (图 2) 表明，式 (10) 的 MPS 在所有模型中精度最高，但与广义模型的 FPA、MPA 和 MPS 差异均不明显。除了式 (12)MPA 时的 MAE 略小于对应 FPA 外，其余指标表明，4 对模型精度均为：MPS> FPA> MPA。模型只含预测变量 D 时，FPA 和 MPS 精度明显较其 MPA 提高， R^2 分别升高了 39.51% 和 93.00%，MAE 分别降低了 27.68% 和 68.88%，RMSE 分别降低了 20.69% 和 64.38%。含林分变量时，各模型的拟合精度较高但差异较小。相较于对应 MPA，各模型 FPA 和 MPS 时精度变化不明显，FPA 时 R^2 、MAE 和 RMSE 改良范围不足 1%，MPS 时精度变化稍微增大但低于 5%。

2.3.2 不同混合模型抽样对比 图 3 为各模型 FPA、MPA 预测及不同抽样方案下 MPS 预测时 MAE 和 RMSE 的变化。进行 FPA 和 MPA 预测时，广义模型精度均高于基础模型，且 FPA 中式 (8) 的精度最高。进行 MPS 时，混合模型的 MAE 和 RMSE 变化趋势相同。其中，基础模型精度一直为：MPS>FPA>MPA。使用随机选择样本木计算模型随机参数时的模型预测效果好于使用平均木计算。每块标准地内随机测量 3 棵树的树高时，式 (10) 的 MAE 和 RMSE 的降低速率最大，此时模型 MAE 和 RMSE 较其 MPA 分别降低了 57.97% 和 57.63%。广义混合模型 [式 (11)~(13)] 的预测精度同样随样本数量增多而降低 (1~8 棵)，随机选择

表 3 混合模型参数估计及统计量

Table 3 Parameter estimates and evaluation statistics of mixed model

公式 Formula	β_1	β_2	β_3	β_4	β_5	β_6	σ_i^2	σ^2	AIC	BIC	LL
(10)	10.258 1	-7.630 1	1.565 7				9.617 6	0.682	2 284.3	2 308.2	-1 137.2
(11)	0.211 3	-6.414 4	1.491 4	1.046 9			7.98×10^{-3}	0.681	2 163.3	2 191.9	-1 075.7
(12)	0.248 3	-5.571 8	1.378 5	1.030 0	0.043 0		5.81×10^{-3}	0.680	2 158.2	2 191.6	-1 072.1
(13)	0.375 3	-5.693 8	1.380 3	1.085 9	0.034 9	-0.029 1	5.65×10^{-3}	0.680	2 159.3	2 197.4	-1 071.6

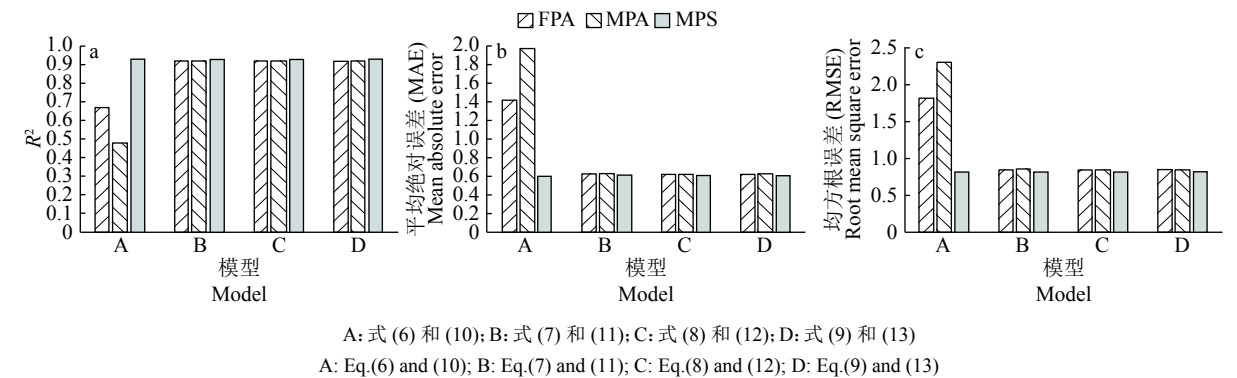


图 2 基于建模数据计算所得不同模型的不同预测类型指标统计

Fig. 2 Indices calculated based on model fitting data for different types of predictions in different models

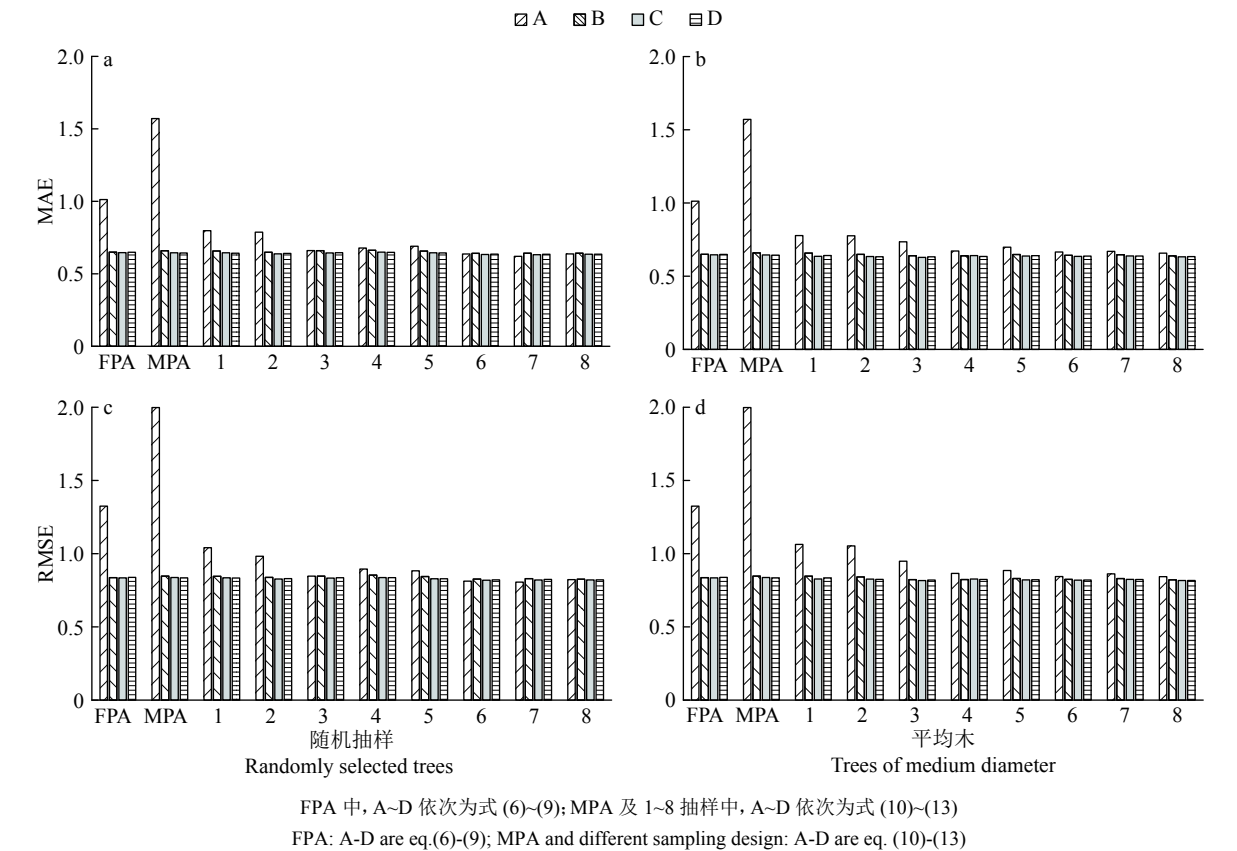


图 3 不同抽样方案下各模型平均绝对误差 (MAE) 和均方根误差 (RMSE) 随样本木数量变化情况

Fig. 3 Relationship between mean absolute error (MAE), root mean square error (RMSE) and number of sample trees based on different sampling designs

2 个样本木和选择 3 个平均木时模型 MAE 和 RMSE 降低速率最大, 但整体变化缓慢。

3 讨论与结论

本文以辽宁章古台地区樟子松人工纯林调查数据为例, 从 30 个常用模型中选出了 Sibbesen 模型为适合樟子松 ($D>5\text{ cm}$) 的 H - D 基础模型。以此为基础, 建立不同的 H - D 广义模型和混合模型时发现, 广义模型和混合模型都能提高模型精度, 且广义模型和基础混合模型精度均较高。

研究表明, 立地条件和林分密度是影响 H - D 关

系最为明显的因素^[1, 7]。 H_T 可以代表林分的立地质量和发展阶段^[21], 王冬至等^[13] 在研究华北落叶松 *Larix principis-rupprechtii* 和白桦 H - D 模型时的结果表明, 包含 H_T 的模型拟合效果较好。本文的研究也表明, H_T 对沙地樟子松 H - D 模型的影响最大, 含 H_T 的式 (7) 模型精度明显高于式 (6) 的基础模型, 且模型 R^2 、MAE 和 RMSE 的变化幅度高于 30%。 A_B 可以表示林分密度和林分竞争, 它可以反映其他因素 (如林龄) 对 H - D 关系的影响^[9]。当模型中含有 H_T 和 A_B 时, 加入 D_{QM} 反而降低了模型精度。此外, H - D 关系是一个动态关系, 即林龄不同时期 H -

D 关系模型的参数也不同, 由于混合模型中随机参数综合反映了包含林龄和立地条件在内的多种差异^[7], 且广义模型中 A_B 和 H_T 均与林龄关系密切, 故而本文未将树木按照林分阶段划分后分别建模。

模型中无随机参数时 (MPA 和 FPA), 模型的预测精度主要受预测变量的影响^[22], 此时 FPA 与对应 MPA 随林分变量个数变化产生的趋势相同。还可能因为混合模型的 MPA 不能够充分代表其总体水平预测^[23], 而出现 FPA 优于 MPA (尤其基础模型)。林分变量和随机参数都能反映不同林分间差异^[22], 4 个混合模型间 MPS 差异不明显, 说明林分变量对混合模型的作用不大^[17]。但并不意味广义模型无构建价值, 因为广义模型具有较强的生物学意义且数据量较小时具有明显优势^[22]。

模型预测时, 如何抽取样本木使得模型精度较高且工作量较小十分重要。使用经验线性最优无偏估计法计算随机参数时, 式 (10) 在随机选择 3 棵样本木时, 模型的 MAE 和 RMSE 降低速率最大, 之后降低速率变慢。基础混合模型的 MPS 精度与广义模型各预测形式间差异不大。由于广义混合模型预测时, 同时需要林分变量和先验信息, 调查成本较高, 但模型预测精度变化较小, 不太实用。因此, 本文建议使用基础混合模型, 并通过随机测量每块样地内 3 棵树的树高, 以此计算随机参数来预测树高, 当无先验信息时, 使用含有林分变量优势木平均高、胸高断面面积的广义模型预测树高。

参考文献:

[1] SHARMA M. Comparing height-diameter relationships of boreal tree species grown in plantations and natural stands[J]. *For Sci*, 2016, 62(1): 70-77.

[2] HUANG S M, TITUS S J. An age-independent individual tree height prediction model for boreal spruce-aspen stands in Alberta[J]. *Can J For Res*, 1994, 24(7): 1295-1301.

[3] 陈义刚, 谢正生, 张祥生, 等. 粤北低山丘陵地区小红栲生长过程分析[J]. *华南农业大学学报*, 1994, 15(2): 124-128.

[4] CORRAL-RIVAS S, ÁLVAREZ-GONZÁLEZ J G, CRECENTE-CAMPO F, et al. Local and generalized height-diameter models with random parameters for mixed, uneven-aged forests in Northwestern Durango, Mexico[J]. *For Ecosyst*, 2014, 1(1): 6.

[5] 臧颢, 雷相东, 张会儒, 等. 红松树高-胸径的非线性混合效应模型研究[J]. *北京林业大学学报*, 2016, 38(6): 8-16.

[6] 樊伟, 许崇华, 崔珺, 等. 基于混合效应的大别山地区杉木树高-胸径模型比较[J]. *应用生态学报*, 2017, 28(9): 2831-2839.

[7] HUANG S M, WIENS D P, YANG Y Q, et al. Assessing the impacts of species composition, top height and density on individual tree height prediction of quaking

aspen in boreal mixed woods[J]. *For Ecol Manage*, 2009, 258(7): 1235-1247.

[8] LEI X D, PENG C H, WANG H Y, et al. Individual height-diameter models for young black spruce (*Picea mariana*) and jack pine (*Pinus banksiana*) plantations in New Brunswick, Canada[J]. *For Chron*, 2009, 85(1): 43-56.

[9] ZEIDE B, VANDERSCHAAF C. The effect of density on the height-diameter relationship[C]// OUTCALT K W. Proceedings of the 11th biennial southern silvicultural research conference. Knoxville: General Technical Report SRS-48, 2002: 463-466.

[10] VAN LAAR A, ACKA A. Forest Mensuration[M]. Heidelberg: Springer, 2007: 383.

[11] SHARMA M, ZHANG S Y. Height-diameter models using stand characteristics for *Pinus banksiana* and *Picea mariana*[J]. *Scand J For Res*, 2004, 19(5): 442-451.

[12] 符利勇, 张会儒, 唐守正. 基于非线性混合模型的杉木优势木平均高[J]. *林业科学*, 2012, 48(7): 68-71.

[13] 王冬至, 张冬燕, 张志东, 等. 基于非线性混合模型的针阔混交林树高与胸径关系[J]. *林业科学*, 2016, 52(1): 30-36.

[14] CALAMA R, MONTERO G. Interregional nonlinear height-diameter model with random coefficients for stone pine in Spain[J]. *Can J For Res*, 2004, 34(1): 150-163.

[15] 符利勇, 孙华, 张会儒, 等. 不同郁闭度下胸高直径对杉木冠幅特征因子的影响[J]. *生态学报*, 2013, 33(8): 2434-2443.

[16] SHARMA R P, BREIDENBACH J. Modeling height-diameter relationships for Norway spruce, Scots pine, and downy birch using Norwegian national forest inventory data[J]. *For Sci Technol*, 2015, 11(1): 44-53.

[17] ZANG H, LEI X, ZENG W. Height-diameter equations for larch plantations in northern and northeastern China: A comparison of the mixed-effects, quantile regression and generalized additive models[J]. *Forestry*, 2016, 89(4): 434-445.

[18] 雷泽勇, 周晏平, 赵国军, 等. 竞争对辽宁西北部樟子松人工固沙林树高生长的影响[J]. *干旱区研究*, 2018(1): 144-149.

[19] 韩辉, 张学利, 党宏忠, 等. 基于树干液流通量的沙地樟子松合理林分密度的确定[J]. *林业科学研究*, 2015, 28(6): 797-803.

[20] HUANG S, TITUS S J, WIENS D P. Comparison of nonlinear height-diameter functions for major Alberta tree species[J]. *Can J For Res*, 1992, 22(9): 1297-1304.

[21] SIBBESSEN E. Some new equations to describe phosphate sorption by soils[J]. *J Soil Sci*, 1981, 32(1): 67-74.

[22] YANG Y, HUANG S. Allometric modelling of crown width for white spruce by fixed- and mixed-effects models[J]. *For Chron*, 2017, 93(2): 138-147.

[23] MENG S X, HUANG S M, YANG Y Q, et al. Evaluation of population-averaged and subject-specific approaches for modeling the dominant or codominant height of lodgepole pine trees[J]. *Can J For Res*, 2009, 39(6): 1148-1158.