北京林业大学学报
JOURNAL OF BEIJING FORESTRY UNIVERSITY

基于混合效应和分位数回归的温带针阔混交林树高与胸径关系研究

程雯 武晓昱 叶尔江·拜克吐尔汉 王娟 赵秀海 张春雨

Research on the relationship between tree height and DBH of temperate coniferous and broadleaved mixed forests based on mixed effects and quantile regression

Cheng Wen, Wu Xiaoyu, Yeerjiang Baiketuerhan, Wang Juan, Zhao Xiuhai, Zhang Chunyu

引用本文:
程雯, 武晓昱, 叶尔江·拜克吐尔汉, 王娟, 赵秀海, 张春雨. 基于混合效应和分位数回归的温带针阔混交林树高与胸径关系研究[J]. 北京林业大学学报, 2024, 46(2):28-39. doi: 10.12171/j.1000-1522.20220428
Cheng Wen, Wu Xiaoyu, Yeerjiang Baiketuerhan, Wang Juan, Zhao Xiuhai, Zhang Chunyu. Research on the relationship between tree height and DBH of temperate coniferous and broadleaved mixed forests based on mixed effects and quantile regression[J]. Journal of Beijing Forestry University, 2024, 46(2):28-39. doi: 10.12171/j.1000-1522.20220428

在线阅读 View online: https://doi.org/10.12171/j.1000-1522.20220428

您可能感兴趣的其他文章

Articles you may be interested in

长白山云冷杉针阔混交林林木空间利用率混合模型

Mixed model of forest space utilization in spruce-fir coniferous and broadleaved mixed forest of Changbai Mountains, northeastern China

北京林业大学学报, 2020, 42(5): 71-79. https://doi.org/10.12171/j.1000-1522.20190112

天然栎类阔叶混交林林分平均高与平均胸径关系模型

Relationship model between stand mean height and mean DBH for natural *Quercus* spp. broadleaved mixed stands

北京林业大学学报, 2020, 42(9): 37-50. https://doi.org/10.12171/j.1000-1522.20190463

基于分位数回归法的杉木可变指数削度方程构建

Modeling variable exponential taper function for *Cunninghamia lanceolata* based on quantile regression

北京林业大学学报, 2021, 43(7): 70-78. https://doi.org/10.12171/j.1000-1522.20200253

基于混合效应模型的人工红松枝下高模型研建

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

北京林业大学学报, 2020, 42(9): 28-36. https://doi.org/10.12171/j.1000-1522.20190366

基于Possion回归混合效应模型的长白落叶松一级枝数量模拟

First-order branch number simulation for *Larix olgensis* plantation through Poisson regression mixed effect model

北京林业大学学报, 2017, 39(11): 45-55. https://doi.org/10.13332/j.1000-1522.20170204

利用分位数回归模拟人工樟子松树干干形

Modeling stem taper profile for *Pinus sylvestris* plantations using nonlinear quantile regression

北京林业大学学报, 2020, 42(2): 1-8. https://doi.org/10.12171/j.1000-1522.20190014

DOI:10.12171/j.1000−1522.20220428

基于混合效应和分位数回归的温带针阔混交林树高与胸径关系研究

程 雯[1] 武晓昱[2] 叶尔江·拜克吐尔汉[3] 王 娟[4] 赵秀海[1] 张春雨[1]
(1. 北京林业大学国家林业和草原局森林经营工程技术研究中心，北京 100083；2. 北京林业大学理学院，北京 100083；3. 新疆农业大学林学与风景园林学院，新疆干旱区林业生态与产业技术重点实验室，新疆 乌鲁木齐 830052；4. 北京林业大学生态与自然保护学院，北京 100083)

摘要:【目的】基于非线性回归和广义模型构建不同分位数回归和混合效应的树高预测方程，并对比分析非线性模型、不同分位点(τ = 0.1，0.2，0.3，0.4，0.5，0.6，0.7，0.8，0.9)模型、广义模型及非线性混合效应模型的拟合效果和预测精度，为研究林分生长和收获提供理论依据。**【方法】**本研究以吉林蛟河地区针阔混交林的主要树种(红松、色木槭、紫椴和水曲柳)为研究对象，基于 21.12 hm^2 样地数据，首先在 11 个广泛使用的树高方程基础模型中选定基础模型；其次探究林分变量对树高的影响并构建含林分变量的广义模型；最后在基础模型和广义模型的基础上，构建分位数模型，同时考虑样方效应对树高的影响，构建混合效应模型。**【结果】**(1)各树种均以 Richards 模型拟合精度更高，且具有生物学意义，选定为基础模型；考虑林分变量与树高的相关性以及模型收敛性，加入优势木高建立的广义模型能显著提高拟合效果。(2)各树种均为中位数 τ = 0.5 时模型拟合效果最佳，且与非线性回归预测精度相近，红松、色木槭、紫椴和水曲柳最高 R^2 值分别为 0.811、0.809、0.724 和 0.617，广义中位数回归预测能力得到进一步提高，R^2 值分别为 0.891、0.874、0.858 和 0.627。(3)混合效应模型相对其他模型能显著提高预测精度，其中基础混合模型略优于广义混合模型，4 个树种 R^2 值达到 0.937、0.919、0.906 和 0.643，表明包含样方效应的混合模型能得到更准确更稳定的预测结果。**【结论】**与传统方法建立的基础模型和广义模型以及两者的中位数回归模型相较，基于非线性混合效应构建的树高−胸径模型预测精度更高，其中基于基础混合效应构建的吉林蛟河地区混交林树高−胸径模型更具优越性和稳定性。

关键词：分位数回归；树高−胸径模型；混合效应模型；广义模型；针阔混交林
中图分类号：S757 **文献标志码：**A **文章编号：**1000−1522(2024)02−0028−12
引文格式：程雯，武晓昱，叶尔江·拜克吐尔汉，等. 基于混合效应和分位数回归的温带针阔混交林树高与胸径关系研究[J]. 北京林业大学学报，2024，46(2)：28−39. Cheng Wen, Wu Xiaoyu, Yeerjiang Baiketuerhan, et al. Research on the relationship between tree height and DBH of temperate coniferous and broadleaved mixed forests based on mixed effects and quantile regression[J]. Journal of Beijing Forestry University, 2024, 46(2): 28−39.

Research on the relationship between tree height and DBH of temperate coniferous and broadleaved mixed forests based on mixed effects and quantile regression

Cheng Wen[1] Wu Xiaoyu[2] Yeerjiang Baiketuerhan[3] Wang Juan[4] Zhao Xiuhai[1] Zhang Chunyu[1]
(1. Research Center of Forest Management Engineering of National Forestry and Grassland Administration, Beijing Forestry University, Beijing 100083, China;
2. School of Science, Beijing Forestry University, Beijing 100083, China;
3. College of Forestry and Landscape Architecture, Xinjiang Agricultural University, Key Laboratory of Xinjiang Uygur Autonomous Region for Forest Ecology and Industrial Technology in Arid Zone, Urumqi 830052, Xinjiang, China;
4. School of Ecology and Nature Conservation, Beijing Forestry University, Beijing 100083, China)

Abstract: [**Objective**] The aim of this study was to construct tree height equation for quantile regression

收稿日期: 2022−10−27 **修回日期:** 2022−12−05
基金项目: 国家重点研发计划重点专项(2022YFD2201004-4)。
第一作者: 程雯。主要研究方向：森林生态学。Email：Chengwen17@bjfu.edu.cn 地址：100083 北京市海淀区清华东路 35 号。
责任作者: 王娟，副教授。主要研究方向：森林生态学、繁殖生态学。Email：wangjuan@bjfu.edu.cn 地址：同上。
本刊网址: http://j.bjfu.edu.cn；http://journal.bjfu.edu.cn

and mixed-effects based on nonlinear regression and generalized models. And fitting effect and prediction accuracy of nonlinear models, different quntile models (τ = 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9), generalized models and nonlinear mixed-effects models were compared and analyzed, so as to provide theoretical basis for further study of stand growth and harvest. [**Method**] Based on 21.12 ha sample plot data, taking the main tree species (*Pinus koraiensis*, *Acer mono*, *Tilia amurensis* and *Fraxinus mandshurica*) from coniferous and broadleaved mixed forest of Jiaohe, Jilin Province of northeastern China as the research object. And the base model was first selected from 11 widely-used tree height equations, then we explored the influence of stand variables on tree height and constructed generalized model containing stand variables. Finally, on the basis of basic model and the generalized model, the quantile model was constructed, and the mixed-effect model was established considering the impact of sample effect on tree height. [**Result**] (1) Richards was selected as base model for all tree species because of its higher fitting accuracy and biological significance. And considering the correlation between stand variables and tree height and the convergence of models, the generalized model established by adding dominant tree height can significantly improve the fitting effect. (2) All models based on the median (τ = 0.5) performed best, and the prediction accuracy was close to the nonlinear regression. The highest R^2 values of *Pinus koraiensis*, *Acer mono*, *Tilia amurensis* and *Fraxinus mandshurica* were 0.811, 0.809, 0.724 and 0.617, respectively. The generalized median regression prediction ability was further improved, and R^2 values were 0.891, 0.874, 0.858 and 0.627, respectively. (3) Mixed-effect models can significantly improve the prediction accuracy compared with other models, among which base mixed model was slightly better than generalized mixed model, and the R^2 values of four tree species reached 0.937, 0.919, 0.906 and 0.643, respectively, indicating that mixed models including sample effect can improve the more accurate and stable prediction results. [**Conclusion**] Compared with base model, generalized model and median regression model established by traditional methods, the height-diameter model based on nonlinear mixed-effects has higher prediction accuracy, and base mixed-effects model has superiority and stability for height-DBH model construction of mixed forests in Jiaohe, Jilin Province of northeastern China.

Key words: quantile regression; tree height-DBH model; mixed-effect model; generalized model; broadleaf-conifer mixed forest

树高和胸径是森林资源调查和经营效果评价的主要林分调查因子，也是林木生长发育和林分材积数量研究的重要变量。树高与胸径关系一直是林学研究的热点。通过量化特定树种的树高与胸径关系拟合树高–胸径预测模型，在计算树干材积、出材率表编制、生长和收获模型、森林资源调查、立地质量评价以及森林碳储量估算方面具有广泛应用。

分位数回归方法具有较长的统计学历史和理论基础，最早于 1978 年由 Koenker 等[1]根据中位数回归推导提出。由于对调查数据模型方差没有严格要求，利用自变量的多个分位数求得因变量条件分布的相应分位数方程，能够描述任意发散点处变量的变化趋势。对具有尖峰、厚尾、异方差显著的数据拟合效果更加稳健[2]，相对最小二乘回归具有更大优势。已广泛应用于林分自疏边界线、直径分布规律、林分密度指数、林木干形、树冠轮廓以及森林病虫害等领域。国外学者利用分位数回归建立了树高–胸径关系模型[3]。国内学者利用分位数回归拟合了树冠轮廓[4]、树冠表面积及树冠体积[5]，结合哑变量方法构建了大兴安岭落叶松（*Larix gmelinii*）的树高与胸径关系最优模型[6]。Cao 等[7]提出利用组合分位数法预测火炬松（*Pinus taeda*）干形曲线，并通过插值系数修正削度方程以提高预测精度。

树木生长受气候、立地条件和经营措施等多种因素的影响和制约，不同样地间的树高–胸径关系存在较大差异，仅含胸径一个变量的树高模型不能描述不同林况下树高–胸径关系的差异，因此很多引入林分变量或者单木变量构建的广义模型也被用于描述不同林分的模型差异。除此之外，近年来混合效应模型在树高模型中也得到了一定的应用，可以反映总体平均变化趋势及个体之间的变异，不需要假设观测值相互独立[8]，显著提高了模型的预测精度。Zhang 等[9]在以样地为随机效应的混合模型中引入林分密度和立地指数，结果表明该模型可以显著提高长白落叶松（*Larix olgensis*）树高–胸径模型的预测精度。燕云飞等[10]基于广义混合效应构建了拟合效

果最优的红松(*Pinus koraiensis*)人工林枝下高模型。

以往树高–胸径关系研究均选定某种方程进行非线性回归，且对多层次、多树种的混交林树高–胸径关系的描述相比纯林要困难和复杂很多[11]。同时随着方程通用程度的提高，用于局部地区的误差也会增大[12]。因此，本研究以吉林蛟河针阔混交林典型树种为研究对象，构建含有林分变量的广义树高–胸径模型，并基于基础模型和广义模型分布拟合不同树种的分位数回归和以样方效应作为分类变量的混合模型，对比分析各树种不同树高–胸径模型的预测效果和拟合差异，为温带针阔混交林资源调查、收获、经营与管理提供依据。

1 研究区概况与研究方法

1.1 研究区概况

吉林省林业实验区国有林保护中心位于吉林省东部、长白山西麓，隶属于长白山系张广才岭山脉。行政区域处于吉林省蛟河市前进乡境内，属温带大陆性季风气候，年平均气温 3.4 ℃，降水量 708.8 mm，无霜期为 120 ~ 130 d。地貌类型以低山丘陵区为主，海拔 459 ~ 517 m。森林资源和植物种类丰富，植被类型为次生针阔混交林，主要乔木树种包括红松、水曲柳(*Fraxinus mandshurica*)、蒙古栎(*Quercus mongolica*)、紫椴(*Tilia amurensis*)、色木槭(*Acer mono*)、白牛槭(*Acer mandshuricum*)和胡桃楸(*Juglans mandshurica*)等。

1.2 数据来源与统计

在研究区设置总面积大小为 21.12 hm^2 监测样地，经纬度为 43°57′35″ ~ 43°57′54″N，127°43′37″ ~ 127°44′09″E。2009 年对植被开展初次调查，2014 年、2019 年对样地进行复测。调查时将样地划分为 528 个 20 m × 20 m 的连续调查样方，记录所有样方中胸径大于 1 cm 的木本植物种类，测量胸径、树高、冠幅、坐标、林分郁闭程度等。基于样地内树种组成(表 1)和各树种树高–胸径模型的初步拟合结果，选择重要值最高的乔木树种作为研究样本，其中筛选出 1 531 株红松、3 066 株色木槭、1 822 株紫椴和 1 856 株水曲柳。建模时将各组数据随机分成两部分，其中 80% 作为训练数据，20% 作为测试数据(表 2)。根据调查数据，计算各样方的林分密度、胸高断面积和优势木高(表 3)。在每个 20 m × 20 m 样方内，分别选择 1 ~ 3 株最高的林木计算平均值，

表 1 优势树种主要指标统计

Tab. 1 Main indicator statistics of dominant tree species

树种 Tree species	重要值 Important value	多度 Abundance	频度 Frequency	胸高断面积/($m^2 \cdot hm^{-2}$) Basal area/($m^2 \cdot ha^{-1}$)
色木槭 *Acer mono*	11.42	3 071	505	87.58
水曲柳 *Fraxinus mandshurica*	11.13	1 861	420	119.81
紫椴 *Tilia amurensis*	9.46	1 825	472	84.91
红松 *Pinus koraiensis*	8.21	1 535	411	74.72
暴马丁香 *Syringa reticulata*	7.28	3 483	466	4.66
春榆 *Ulmus davidiana*	6.79	2 279	446	26.94
拧筋槭 *Acer triflorum*	6.51	1 474	444	41.76
蒙古栎 *Quercus mongolica*	5.62	646	337	55.80
白牛槭 *Acer mandshuricum*	4.67	1 492	358	15.56
白桦 *Betula platyphylla*	4.44	876	249	36.81

表 2 单木树高–胸径模型拟合及检验数据

Tab. 2 Fitting and testing data of individual tree height-DBH model

数据分类 Dataset classification	树种 Tree species	样本数 Number of sample	胸径 DBH/cm				树高 Tree height/m			
			平均值 Mean	最小值 Min. value	最大值 Max. value	标准差 SD	平均值 Mean	最小值 Min. value	最大值 Max. value	标准差 SD
拟合数据 Fitting data	红松 *Pinus koraiensis*	1 226	19.14	1.60	75.80	15.90	12.25	1.00	33.40	6.75
	色木槭 *Acer mono*	2 452	14.13	1.20	88.80	12.62	11.22	2.00	37.20	6.10
	紫椴 *Tilia amurensis*	1 458	19.86	1.20	70.40	13.76	14.42	1.60	32.70	6.36
	水曲柳 *Fraxinus mandshurica*	1 484	26.17	2.00	75.00	12.32	17.82	2.70	27.30	4.15
检验数据 Testing data	红松 *Pinus koraiensis*	305	19.28	2.00	65.30	16.12	12.11	1.90	34.40	6.98
	色木槭 *Acer mono*	614	14.32	1.40	62.00	12.55	11.49	1.20	29.40	6.42
	紫椴 *Tilia amurensis*	364	20.12	1.70	65.80	14.60	14.45	1.80	34.00	6.68
	水曲柳 *Fraxinus mandshurica*	372	24.84	1.80	62.40	11.76	17.75	4.50	27.00	4.21

表 3　林分调查因子统计表

Tab. 3　Statistical table of stand description factors

数据分类 Dataset classification	树种 Tree species	胸高断面积/($m^2 \cdot hm^{-2}$) Basal area/($m^2 \cdot ha^{-1}$)				林分密度/(株·hm^{-2}) Stand density/(tree·ha^{-1})				优势木高 Dominant tree height/m			
		平均值 Mean	最小值 Min. value	最大值 Max. value	标准差 SD	平均值 Mean	最小值 Min. value	最大值 Max. value	标准差 SD	平均值 Mean	最小值 Min. value	最大值 Max. value	标准差 SD
拟合数据 Fitting data	红松 *Pinus koraiensis*	32.91	15.10	257.89	13.62	1 223.63	350	2 900	384.76	16.30	1.80	33.40	6.97
	色木槭 *Acer mono*	32.66	15.02	257.89	17.34	1 244.02	350	3 100	400.18	16.98	4.00	37.20	5.56
	紫椴 *Tilia amurensis*	34.00	15.02	257.89	14.38	1 284.93	350	2 900	387.05	19.62	3.40	34.00	4.79
	水曲柳 *Fraxinus mandshurica*	31.44	15.02	257.89	15.41	1 420.70	350	3 100	447.84	24.21	2.80	35.30	4.49
检验数据 Testing data	红松 *Pinus koraiensis*	33.49	19.21	58.18	7.35	1 215.16	450	2 575	346.34	16.88	2.00	33.40	7.14
	色木槭 *Acer mono*	33.46	16.19	257.89	16.51	1 258.15	450	2 900	374.19	17.23	3.20	37.20	5.31
	紫椴 *Tilia amurensis*	33.82	15.02	257.89	16.92	1 310.16	600	2 900	412.81	19.67	6.20	34.00	5.15
	水曲柳 *Fraxinus mandshurica*	31.91	15.10	257.89	21.16	1 488.58	600	3 100	443.16	24.95	15.20	34.80	3.97

作为该树种在该样方的优势木高。利用 R 软件的 quantreg 包和 nlme 包进行非线性分位数回归和混合效应模型分析。

1.3　研究方法

1.3.1　基础模型与广义模型

以往大多选定某种非线性方程描述树种的树高−胸径关系，而混交林相对于纯林或年龄结构和林层单一的林分来说，其树高与胸径的关系更为复杂[13]。根据以往的研究，基于树高−胸径散点图和树种的分布特征，选择以下 11 个常用的树高模型作为候选基础模型(表 4)，对树高−胸径分布数据进行非线性拟合，比较各模型的拟合精度，筛选预测精度最高的作为研究区树高−胸径的最优基础模型。

但在结构更加复杂的针阔混交林中，仅仅将胸径作为预测变量，无法反映其他林分变量对树高的影响。本研究采用再参数化的方法建立树高−胸径模型，依据林分变量与树高之间的关系，适当融合林分密度、胸高断面积和优势木高度等变量。具体做法如下：基于所选择的最优基础模型，对林分变量与树高进行相关性分析；比较不同变量引入对模型的影响，筛选模型中的其他变量，确定与树高相关性较大且能显著提高模型预测能力的变量作为协变量；基于最优基础模型构建各树种预测精度最高的广义模型。

1.3.2　分位数回归

分位数回归是利用解释变量估计响应变量的条件分位数分布，全面刻画整个条件分布的综合回归方法[25−26]。不仅可以度量响应变量对分布中心的影响，而且可以度量对分布上尾和下尾的影响。最小二乘法假定误差服从正态分布，而分位数回归的响

表 4　候选树高−胸径关系模型

Tab. 4　Candidate models of tree height-DBH relation

模型编号 Model No.	表达式 Equation	参考文献 Reference
M1	$H = 1.3 + \beta_1 D^{\beta_2}$	[14]
M2	$H = 1.3 + \beta_1 \left[\frac{D}{1+D}\right]^{\beta_2}$	[15]
M3	$H = 1.3 + \exp\left(\beta_1 + \frac{\beta_2}{D+1}\right)$	[16]
M4	$H = 1.3 + \beta_1 [1 - \exp(-\beta_2 D)]$	[17]
M5	$H = 1.3 + \beta_1 \exp\left(\frac{\beta_2}{D}\right)$	[18]
M6	$H = 1.3 + \beta_1 [1 - \exp(-\beta_2 D)]^{\beta_3}$	[19]
M7	$H = 1.3 + \beta_1 D^{\left[\beta_2 D^{(-\beta_3)}\right]}$	[20]
M8	$H = 1.3 + \exp\left(\beta_1 + \frac{\beta_2}{D+\beta_3}\right)$	[21]
M9	$H = 1.3 + \beta_1 \exp[-\beta_2 \exp(-\beta_3 D)]$	[22]
M10	$H = 1.3 + \frac{D^2}{\beta_1 + \beta_2 D + \beta_3 D^2}$	[23]
M11	$H = 1.3 + \frac{\beta_1}{1 + \beta_2 \exp(-\beta_3 D)}$	[24]

注：H为树高预测值；D为胸径；β_1、β_2、β_3为模型参数。Notes: H is the predicted tree height; D is the DBH; β_1, β_2 and β_3 are the model parameters.

应变量的平均预测值对残差分布没有特殊要求[27]，从而确保参数估计的稳健性和模型拟合的灵活性。最小二乘回归的目标是最小化误差平方和，分位数回归的核心思想是从均值推广到分位数，也是通过最小化残差绝对值的非对称损失函数来估计分为点处参数值。

$$\min\left[\sum_{y_\tau \geqslant \hat{y}_\tau} \tau(y_\tau - \hat{y}_\tau) + \sum_{y_\tau < \hat{y}_\tau} (1-\tau)(\hat{y}_\tau - y_\tau)\right] \quad (1)$$

式中：τ 表示要估计的分位数，y_τ 和 $\hat{y}_\tau$ 分别为响应变量在分位数 τ 处的观测值和估计值。

分位数组合法利用抽样数据拟合不同组合的分位数曲线，基于相邻两条分位数回归曲线插值得到修正的分位数曲线。假设第 i 个样方第 j 个树木的胸径值为 x_{ij}，树高观测值为 y_{ij} 分布在第 m 和 $(m+1)$ 分位数估计值之间，即 $\hat{y}_m(x_{ij}) \leqslant y_{ij} \leqslant \hat{y}_{m+1}(x_{ij})$。使用插值法计算通过该点的修正曲线。

$$\hat{y}_{ij} = \alpha \hat{y}_m(x_{ij}) + (1-\alpha)\hat{y}_{m+1}(x_{ij}) \quad (2)$$

式中：$\alpha = \dfrac{\hat{y}_{m+1}(x_{ij}) - y_{ij}}{\hat{y}_{m+1}(x_{ij}) - \hat{y}_m(x_{ij})}$ 表示插值系数，$\hat{y}_{ij}$ 为树高估计值，$\hat{y}_m(x_{ij})$ 和 $\hat{y}_{m+1}(x_{ij})$ 分别为胸径为 x_{ij} 的树木在第 m 和 $(m+1)$ 分位数处的树高估计值。

当抽样数量大于 1 时，在各分位数组合中选择使平均误差值符号发生改变的相邻曲线。当胸径值大于最大分位数回归的预测值时，将 $\hat{y}_m$ 和 $\hat{y}_{m+1}$ 分别定义为 $\hat{y}_{m-1}$ 和 $\hat{y}_m$，之后利用公式(2)得到树高曲线。

由于基础模型只有一个预测变量缺乏其他变量信息，很难完整描述响应变量在不同分位点的分布。本研究选择 9 个分位点（τ = 0.1，0.2，0.3，0.4，0.5，0.6，0.7，0.8，0.9），基于所选的最优基础模型拟合树高−胸径曲线，同时使用三分位数、五分位数和九分位数组合法得到树高修正曲线，通过模型评价和比较得到基于最佳分位数的基础模型，进而在该模型中融入林分变量构建广义分位数模型。

1.3.3 非线性混合效应模型

在筛选出的最优基础模型和广义模型加入随机效应构建非线性混合模型，本研究中选择样方为随机效应，混合效应模型一般形式可表示为

$$\hat{\boldsymbol{H}}_{ij} = f(\hat{\boldsymbol{\phi}}_i, \hat{\boldsymbol{D}}_{ij}) + \boldsymbol{\varepsilon}_{ij},\ \boldsymbol{\varepsilon}_{ij} \sim \mathrm{N}(0,\ \sigma^2) \quad (3)$$

式中：$\hat{\boldsymbol{H}}_{ij}$ 和 $\hat{\boldsymbol{D}}_{ij}$ 分别为第 i 个样方第 j 株树的树高和胸径预测向量，f 为包含树高和胸径变量的函数表达式，$\hat{\boldsymbol{\phi}}_i$ 为混合效应参数向量，$\boldsymbol{\varepsilon}_{ij}$ 为误差项。

参数向量 $\hat{\boldsymbol{\phi}}_i$ 的固定效应参数和随机效应参数可表示为

$$\boldsymbol{\phi}_i = \hat{\boldsymbol{X}}_i\hat{\boldsymbol{\beta}} + \hat{\boldsymbol{Z}}_i\hat{\boldsymbol{u}}_i,\ \hat{\boldsymbol{u}}_i \sim \mathrm{N}(0,\ \boldsymbol{\varphi}) \quad (4)$$

式中：$\hat{\boldsymbol{\beta}}$ 为固定效应参数向量，$\hat{\boldsymbol{u}}_i$ 为第 i 个样方的随机效应参数向量，遵循均值为 0、协方差矩阵为 $\boldsymbol{\varphi}$ 的正态分布，$\hat{\boldsymbol{X}}_i$ 和 $\hat{\boldsymbol{Z}}_i$ 分别是 $\hat{\boldsymbol{\beta}}$ 和 $\hat{\boldsymbol{u}}_i$ 的设计矩阵。

非线性混合效应模型中，方差−协方差结构由包括相关因子和加权因子的矩阵 $\boldsymbol{R}_i$ 来表示。

$$\boldsymbol{R}_i = \sigma^2 \boldsymbol{G}_i^{0.5} \boldsymbol{\Gamma}_i \boldsymbol{G}_i^{0.5} \quad (5)$$

式中：σ^2 指模型的残差方差值，$\boldsymbol{\Gamma}_i$ 为组内误差相关性结构，$\boldsymbol{G}_i$ 为描述样方内方差异质性的对角矩阵，其对角元素由方差函数提供[28]。

随机效应的方差−协方差结构（$\boldsymbol{D}$）反映了样方间的差异，矩阵因随机参数个数的不同也会有不同的形式，本研究中选用广义正定矩阵结构，当混合模型中具有两个随机参数时表达式如下

$$\boldsymbol{D} = \begin{bmatrix} \sigma_{u_1}^2 & \sigma_{u_1u_2} \\ \sigma_{u_1u_2} & \sigma_{u_2}^2 \end{bmatrix} \quad (6)$$

式中：u_1、u_2 为随机参数，$\sigma_{u_1}^2$、$\sigma_{u_2}^2$ 分别为 u_1、u_2 的方差，$\sigma_{u_1u_2}$ 为 u_1 和 u_2 的协方差。

混合效应模型中固定效应部分的检验方法与传统模型相同，对于随机效应的检验则需要计算随机参数值，本研究需计算在拟合数据中不涉及的样方的随机参数，利用式 7[29]进行计算。

$$\hat{u}_i \approx \hat{D}\hat{Z}_i^{\mathrm{T}}\left(\hat{Z}_i\hat{D}\hat{Z}_i^{\mathrm{T}} + \hat{R}_i\right)^{-1}\hat{e}_i \quad (7)$$

式中：$\hat{\boldsymbol{D}}$ 为随机效应参数的方差协方差矩阵，$\hat{\boldsymbol{R}}_i$ 为样方内方差协方差矩阵，$\hat{\boldsymbol{Z}}_i$ 为设计矩阵，$\hat{\boldsymbol{Z}}_i^{\mathrm{T}}$ 为转置矩阵，$\hat{e}_i$ 为实际值减去用固定效应参数计算的预测值。

1.3.4 模型评价与检验

研究中通过决定系数（R^2）、平均绝对误差（MAE）和均方根误差（RMSE）评价和比较模型拟合结果，并利用未参与建模的 20% 独立样本检验 11 个候选非线性模型，上述评价指标的计算公式。

$$R^2 = 1 - \sum_{i,j=1}^{n}\left(H_{ij} - \hat{H}_{ij}\right)^2 / \sum_{i,j=1}^{n}\left(H_{ij} - \overline{H}_{ij}\right)^2 \quad (8)$$

$$\mathrm{MAE} = \sum_{i,j=1}^{n}\left|H_{ij} - \hat{H}_{ij}\right| / n \quad (9)$$

$$\mathrm{RMSE} = \sqrt{\sum_{i,j=1}^{n}\left(H_{ij} - \hat{H}_{ij}\right)^2 / n} \quad (10)$$

式中：H_{ij} 和 $\hat{H}_{ij}$ 为第 i 个样方第 j 个样本的树高实际值和预测值，$\overline{H}_{ij}$ 为预测值的均值；n 为样本数。

除考虑以上指标外，还选择赤池信息准则（AIC）、贝叶斯信息准则（BIC）和似然比检验（LRT）对混合效应模型进行评价，AIC 和 BIC 数值越小则表明模型拟合精度越高，LRT 得到的 P 值若 < 0.000 1，则认为两个模型差异显著。

$$\mathrm{AIC} = -2L + 2\lambda \quad (11)$$

$$\mathrm{BIC} = -2L + \lambda \ln n \tag{12}$$

式中：L 为模型的对数似然值，λ 为固定效应参数的个数。

2　结果与分析

2.1　基础模型拟合

基于红松、色木槭、紫椴和水曲柳的建模数据，对表 4 中的 11 个基础模型进行拟合，比较其 MAE、RMSE 和 R^2 值（表 5）。结果表明，对于 4 个树种的树高–胸径关系拟合，各模型均具有良好表现。MAE 范围为 1.887 ~ 2.818，RMSE 范围为 2.557 ~ 3.582，R^2 则为 0.552 ~ 0.811，其中红松、紫椴水曲柳以 M6 即 Richards 模型拟合结果最优，色木槭的 Richards 模型拟合结果与其最优模型相差极小。使用检验数据对拟合模型进行评价比较，得到相同结果。且 Richards 方程生物学意义明确、灵活性较高易收敛，被认为很适合描述树高与胸径之间的关系[30–31]，因此选择 Richards 作为 4 个树种的基础模型。

2.2　广义模型构建

对表 3 的林分特征因子进行相关性分析，结果如表 6 所示。4 个树种的树高和胸径与优势木高呈显著正相关关系（$P < 0.01$），色木槭和紫椴树高、胸

表 5　候选模型拟合与评价

Tab. 5　Fitting and evaluation of candidate models

模型 Model	红松 *Pinus koraiensis*			色木槭 *Acer mono*			紫椴 *Tilia amurensis*			水曲柳 *Fraxinus mandshurica*		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
M1	2.381	3.153	0.780	2.118	2.846	0.781	2.818	3.582	0.684	2.248	2.774	0.552
M2	2.361	3.128	0.784	2.080	2.806	0.787	2.521	3.357	0.722	2.079	2.571	0.615
M3	2.211	2.982	0.803	1.926	2.697	0.803	2.503	3.344	0.724	2.083	2.575	0.614
M4	2.179	2.945	0.808	1.895	2.654	0.809	2.530	3.343	0.724	2.088	2.575	0.614
M5	2.279	3.032	0.797	2.030	2.772	0.793	2.547	3.374	0.719	2.076	2.568	0.616
M6	2.140	2.925	0.811	1.890	2.654	0.809	2.491	3.332	0.726	2.063	2.557	0.619
M7	2.149	2.936	0.809	1.887	2.653	0.809	2.495	3.337	0.725	2.069	2.562	0.617
M8	2.144	2.932	0.810	1.888	2.652	0.810	2.496	3.335	0.725	2.074	2.565	0.617
M9	2.174	2.945	0.808	1.931	2.691	0.804	2.513	3.348	0.723	2.077	2.570	0.615
M10	2.146	2.935	0.809	1.887	2.652	0.810	2.493	3.336	0.725	2.068	2.560	0.618
M11	2.231	2.991	0.802	1.986	2.742	0.796	2.553	3.381	0.718	2.091	2.587	0.610

注：MAE.平均绝对误差；RMSE.均方根误差；R^2.决定系数。下同。Notes: MAE, mean absolute error; RMSE, root mean squared error; R^2, coefficient of determination. The same below.

表 6　树高与林分变量相关性分析

Tab. 6　Correlation analysis between tree height and stand variables

树种 Tree species	因子 Factor	树高 Tree height	胸径 DBH	胸高断面积 Basal area	林分密度 Stand density	优势木高 Dominate tree height
红松 *Pinus koraiensis*	树高 Tree height	1.000				
	胸径 DBH	0.855**	1.000			
	胸高断面积 Basal area	0.020	0.040	1.000		
	林分密度 Stand density	0.010	0.006	0.092**	1.000	
	优势木高 Dominate tree height	0.587**	0.440**	0.069*	0.018	1.000
色木槭 *Acer mono*	树高Tree height	1.000				
	胸径 DBH	0.845**	1.000			
	胸高断面积 Basal area	0.031	0.030	1.000		
	林分密度 Stand density	−0.284**	−0.270**	0.089**	1.000	
	优势木高 Dominate tree height	0.484**	0.353**	0.139**	−0.261**	1.000
紫椴 *Tilia amurensis*	树高 Tree height	1.000				
	胸径 DBH	0.776**	1.000			
	胸高断面积 Basal area	0.068**	0.064**	1.000		
	林分密度 Stand density	−0.207**	−0.200**	0.105**	1.000	
	优势木高 Dominate tree height	0.537**	0.294**	0.178**	−0.175**	1.000
水曲柳 *Fraxinus mandshurica*	树高Tree height	1.000				
	胸径 DBH	0.684**	1.000			
	胸高断面积 Basal area	0.067**	0.133**	1.000		
	林分密度 Stand density	−0.037	−0.081**	0.084**	1.000	
	优势木高 Dominate tree height	0.256**	0.264**	−0.021**	0.318**	1.000

注：** 表示两个变量在 $P < 0.01$ 水平上显著相关，* 表示两个变量在 $P < 0.05$ 水平上显著相关。Notes: ** means significant correlation at $P < 0.01$ level between two variables, * means significant correlation at $P < 0.05$ level between two variables.

径与林分密度呈显著负相关关系($P < 0.01$)。因此将优势木高和林分密度两个变量分别引入基础模型构建广义模型，但包含林分密度因子的色木槭和紫椴树高方程不收敛，而引入胸高断面积的水曲柳树高方程也不收敛，这表明添加优势木高度可显著提高模型精度。因此本研究最优广义模型的表达式为

$$H = 1.3 + (\beta_1 + a_1 H_{D,i})[1 - \exp(-\beta_2 D)]^{\beta_3} \quad (13)$$

式中：H_D 为某个树种的优势木高，$H_{D,i}$ 为第 i 个样方的优势木高，β_1、β_2、β_3、a_1 为模型参数。

2.3 分位数回归

在不同分位点处($\tau = 0.1$，0.2，0.3，0.4，0.5，0.6，0.7，0.8，0.9)，基于 Richards 方程利用分位数回归分别拟合树高–胸径关系模型(表 5)。利用回归评价指标检验不同分位数模型对观测值的拟合程度(表 7)，基础模型及分位数回归模型的决定系数 R^2 值均在 0.6 以上，表明各分位点的树高方程均具有较好的拟合效果。其中 $\tau = 0.5$ 的中位数回归与非线性回归的拟合效果基本一致，反映了分位数回归模型的灵活性。

利用未参与建模的 20% 测试数据检验模型的预测能力(表 8)。分位数组合法分别采用三分位数组合($\tau = 0.1$，0.5，0.9)、五分位数组合($\tau = 0.1$，0.3，0.5，0.7，0.9)和九分位数组合($\tau = 0.1$，0.2，0.3，0.4，0.5，0.6，0.7，0.8，0.9)。随机抽样或抽取平均木时幼树会带来较大预测误差，基于王君杰等[32]的研究结果，选取 10 株最大树木计算插值系数进行预测。结果表明：对于分位数组合法，三分位数和五分位数组合的预测精度高于九分位数组合；其中红松树高的拟合效果略优于中位数回归，但其他树种模型预测精度均低于中位数回归。因此，本研究选择中位数为 4 个树种的最优分位数树高曲线式(13)构建广义分位数模型。

2.4 非线性混合效应模型构建

在基础模型和广义模型基础上考虑样方效应对树高的影响，在 4 个树种的基础模型上增加样方为随机效应，共 8 种情况收敛(表 9)，红松、紫椴和水曲柳增加 1 个随机效应参数时拟合效果最好，而色木槭引入 2 个随机参数的混合模型预测精度更高，模型表达式分别为

$$H = 1.3 + (\beta_1 + u_{1i})[1 - \exp(-\beta_2 D)]^{\beta_3} \quad (14)$$

$$H = 1.3 + (\beta_1 + u_{1i})[1 - \exp(-\beta_2 D)]^{(\beta_3 + u_{2i})} \quad (15)$$

式中：u_{1i}、u_{2i} 为样方水平的随机参数。

在广义模型基础上，引入不同组合形式的随机效应作为自变量构建广义混合效应模型，4 个树种共有 13 种情况收敛(表 9)，当引入 3 个以上随机参

表 7 分位数回归模型拟合精度

Tab. 7 Goodness-of-fit tests for quantile regression models

分位数 Quantile	红松 *Pinus koraiensis*			色木槭 *Acer mono*			紫椴 *Tilia amurensis*			水曲柳 *Fraxinus mandshurica*		
	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2	MAE	RMSE	R^2
Richards	2.140	2.925	0.811	1.890	2.654	0.809	2.491	3.332	0.726	2.063	2.557	0.619
0.1	3.415	4.497	0.780	3.103	4.105	0.770	4.029	5.166	0.696	3.426	4.156	0.614
0.2	2.779	3.826	0.793	2.460	3.406	0.793	3.140	4.238	0.715	2.680	3.387	0.615
0.3	2.368	3.322	0.805	2.081	2.963	0.804	2.742	3.791	0.720	2.323	2.976	0.615
0.4	2.179	3.063	0.809	1.924	2.764	0.807	2.520	3.493	0.723	2.116	2.693	0.615
0.5	2.122	2.942	0.811	1.873	2.666	0.809	2.461	3.374	0.724	2.056	2.583	0.617
0.6	2.191	2.953	0.810	1.918	2.675	0.808	2.542	3.363	0.724	2.119	2.596	0.616
0.7	2.440	3.181	0.807	2.161	2.897	0.803	2.830	3.630	0.719	2.372	2.846	0.614
0.8	2.912	3.690	0.797	2.577	3.354	0.791	3.388	4.268	0.697	2.887	3.422	0.615
0.9	4.081	5.027	0.753	3.451	4.293	0.762	4.629	5.519	0.679	3.965	4.582	0.602

表 8 分位数模型验证评价结果

Tab. 8 Quantile model validation evaluation results

方法 Method	红松 *Pinus koraiensis*		色木槭 *Acer mono*		紫椴 *Tilia amurensis*		水曲柳 *Fraxinus mandshurica*	
	RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
中位数回归 Median regression	2.963	2.094	2.898	2.020	3.540	2.549	2.560	2.050
九分位数组合 Nine quantile combination	2.956	2.118	2.997	2.185	3.510	2.761	2.601	2.150
五分位数组合 Five quantile combination	2.927	2.095	2.971	2.155	3.507	2.758	2.602	2.152
三分位数组合 Triquartile combination	2.926	2.092	2.983	2.162	3.596	2.851	2.592	2.139

表 9　混合效应模型验证评价结果

Tab. 9　Verification and evaluation results of mixed effect model

树种 Tree species	基础混合效应模型 Base mixed-effect model						广义混合效应模型 Generalized mixed-effect model					
	随机参数 Random parameter	AIC	BIC	L	LRT	P	随机参数 Random parameter	AIC	BIC	L	LRT	P
红松 *Pinus koraiensis*	u_1	5 619.3	5 644.8	−2 804.6			u_1	5 346.9	5 377.5	−2 667.4		
	u_1, u_2	5 643.2	5 678.9	−2 814.6	19.878 2	< 0.001	u_3	5 354.2	5 384.8	−2 671.1		
							u_1, u_3	5 350.9	5 391.8	−2 667.5	7.238 0	0.026 8
色木槭 *Acer mono*	u_2	10 966.2	10 995.2	−5 478.1			u_1	10 597.5	10 632.3	−5 292.7		
	u_1, u_2	10 962.2	11 002.9	−5 474.1	7.964 9	0.018 6	u_3	10 602.8	10 637.6	−5 295.4		
							u_1, u_3	10 594.2	10 640.6	−5 289.1	12.571 9	0.001 9
							u_1, u_2	10 601.4	10 647.8	−5 292.8		
紫椴 *Tilia amurensis*	u_1	7 109.6	7 136.0	−3 549.8			u_1	6 662.5	6 694.2	−3 325.2		
							u_3	6 661.8	6 693.5	−3 324.9		
							u_1,u_3	6 665.8	6 708.0	−3 324.9	4.415 0	0.998 7
水曲柳 *Fraxinus mandshurica*	u_1	7 007.8	7 034.4	−3 498.9			u_1	6 966.0	6 997.8	−3 477.0		
	u_2	7 009.0	7 035.5	−3 499.5			u_3	6 964.3	6 996.1	−3 476.1		
	u_1, u_2	7 011.8	7 049.0	−3 498.9	1.126 9	0.569 2	u_1,u_3	6 964.9	7 007.4	−3 474.5	3.322 9	0.189 9

注：AIC.赤池信息准则；BIC.贝叶斯信息准则；L.对数似然值；LRT.似然比检验。Notes: AIC, Akaike information criterion; BIC, Bayesian information criterion; L, log-likelihood value; LRT, likelihood ratio test.

数时，模型不收敛；仅引入参数u_1的红松树高模型和仅引入参数u_3的紫椴及水曲柳树高模型 AIC、BIC 值最小，且 L 值达到最大，而色木槭以随机效应作用于u_1和u_3两个参数时模型预测精度最高，一般表达式分别为

$$H = 1.3 + (\beta_1 + u_{1i} + a_1 H_{\mathrm{D},i})[1 - \exp(-\beta_2 D)]^{\beta_3} \quad (16)$$

$$H = 1.3 + [\beta_1 + u_{1i} + (a_1 + u_{3i}) H_{\mathrm{D},i}][1 - \exp(-\beta_2 D)]^{\beta_3} \quad (17)$$

$$H = 1.3 + [\beta_1 + (a_1 + u_{3i}) H_{\mathrm{D},i}][1 - \exp(-\beta_2 D)]^{\beta_3} \quad (18)$$

式中：u_{1i}、u_{3i}为样方水平的随机参数。

选择各树种拟合效果最好的混合效应模型，参数估计值见表 10。

2.5　模型评价与检验

4 个树种的各模型拟合统计量见表 11。结果表明：基于混合效应的模型精度明显高于不含随机参数的模型；在不含随机效应的模型中，中位数回归形式和一般形式的广义模型精度差异较小；广义中位

表 10　混合效应模型参数估计

Tab. 10　Parameter estimates of mixed-effect models

参数 Parameter	基础混合效应模型 Base mixed-effect model				广义混合效应模型 Generalized mixed-effect model			
	红松 *Pinus koraiensis*	色木槭 *Acer mono*	紫椴 *Tilia amurensis*	水曲柳 *Fraxinus mandshurica*	红松 *Pinus koraiensis*	色木槭 *Acer mono*	紫椴 *Tilia amurensis*	水曲柳 *Fraxinus mandshurica*
公式编号 Equation No.	（14）	（15）	（14）	（14）	（16）	（17）	（18）	（18）
a	21.192(0.387***)	18.626(0.251***)	19.366(0.263***)	20.021(0.190***)	9.186(0.535***)	7.776(0.467***)	4.522(0.468***)	16.513(0.523***)
b	0.065(0.004***)	0.085(0.004***)	0.100(0.005***)	0.101(0.007***)	0.071(0.004***)	0.092(0.004*)	0.104(0.005***)	0.110(0.007***)
c	1.366(0.056***)	1.142(0.033***)	1.367(0.066***)	1.445(0.126***)	1.305(0.057***)	1.142(0.034***)	1.322(0.066***)	1.567(0.142***)
a_1					0.577(0.024***)	0.554(0.023***)	0.715(0.022***)	0.132(0.019***)
$\sigma^2_{u_1}$	14.927 2	10.920 5	11.410 7	0.201 9	2.877 0	14.850 5		
$\sigma^2_{u_2}$		0.011 9						
$\sigma^2{}_{u_3}$						0.038 0	0.001 7	0.000 4
$\sigma_{u_1 u_2}$		0.162 1						
$\sigma_{u_1 u_3}$						−0.718 2		
σ^2	3.637 0	3.638 8	4.908 0	6.422 3	3.811 3	3.858 9	5.275 9	6.201 5

注：***表示在 $P < 0.001$ 水平上显著。Note: *** means significant at $P < 0.001$ level.

表 11 树高–胸径模型拟合评价统计量

Tab. 11 Fitting and evaluation statistics of tree height-DBH models

树种 Tree species	指标 Index	基础模型 Base model	广义模型 Generalized model	中位数回归 Median quantile regression		非线性混合效应 Nonlinear mixed-effects	
				简单模型 Simple model	广义模型 Generalized model	简单模型 Simple model	广义模型 Generalized model
红松 *Pinus koraiensis*	MAE	2.140	1.669	2.122	1.666	1.259	1.358
	RMSE	2.925	2.225	2.942	2.228	1.687	1.826
	R^2	0.811	0.891	0.811	0.891	0.937	0.927
色木槭 *Acer mono*	MAE	1.890	1.575	1.873	1.568	1.258	1.359
	RMSE	2.654	2.159	2.666	2.166	1.740	1.875
	R^2	0.809	0.875	0.809	0.874	0.919	0.906
紫椴 *Tilia amurensis*	MAE	2.491	1.722	2.461	1.706	1.399	1.611
	RMSE	3.332	3.379	3.374	2.401	1.955	2.235
	R^2	0.726	0.860	0.724	0.858	0.906	0.877
水曲柳 *Fraxinus mandshurica*	MAE	2.063	2.041	2.056	2.031	2.004	2.527
	RMSE	2.557	2.524	2.583	2.536	2.478	2.038
	R^2	0.619	0.629	0.617	0.627	0.643	0.629

数模型在预测精度上高于基础模型的中位数回归，红松、色木槭、紫椴和水曲柳的模型 R^2 分别提高了 9.9%、8.0%、18.5% 和 1.6%，MAE 值分别降低了 27.4%、19.5%、44.3% 和 1.2%，RMSE 值分别下降了 32.0%、23.1%、40.5% 和 1.9%。在增加随机效应的模型中，基础混合模型的拟合结果略优于广义混合模型，4 个树种 R^2 仅分别提高了 1.1%、1.4%、3.3% 和 2.2%，但两者并未有显著差异。红松、色木槭、紫椴和水曲柳的基础混合效应和广义混合效应模型的残差分布图如图 1 所示，残差分布较为均匀，没有出现明显的不规则形状，这说明各树种的混合模型均没有明显异方差现象。

混合效应模型中随机效应部分的检验需要估算随机参数值，本研究以 20% 的样木作为检验数据，利用公式(7)分别计算未包含在拟合数据的样方的随机参数；同时与基础模型、广义模型和分位数模型的预测精度进行比较，检验结果如表 12。可以看出，混合效应模型的 MAE 和 RMSE 值均小于其他模型，拟合效果相较分位数回归得到了显著提高，其中基础混合模型相比于广义混合模型预测精度更高。

3 讨 论

3.1 基于基础模型的分位数回归

分位数回归既能描述均值条件下树高的分布趋势，也能描述特定分位点树高–胸径曲线的形状变化[33]。传统的最小二乘法易受残差分布限制，与之

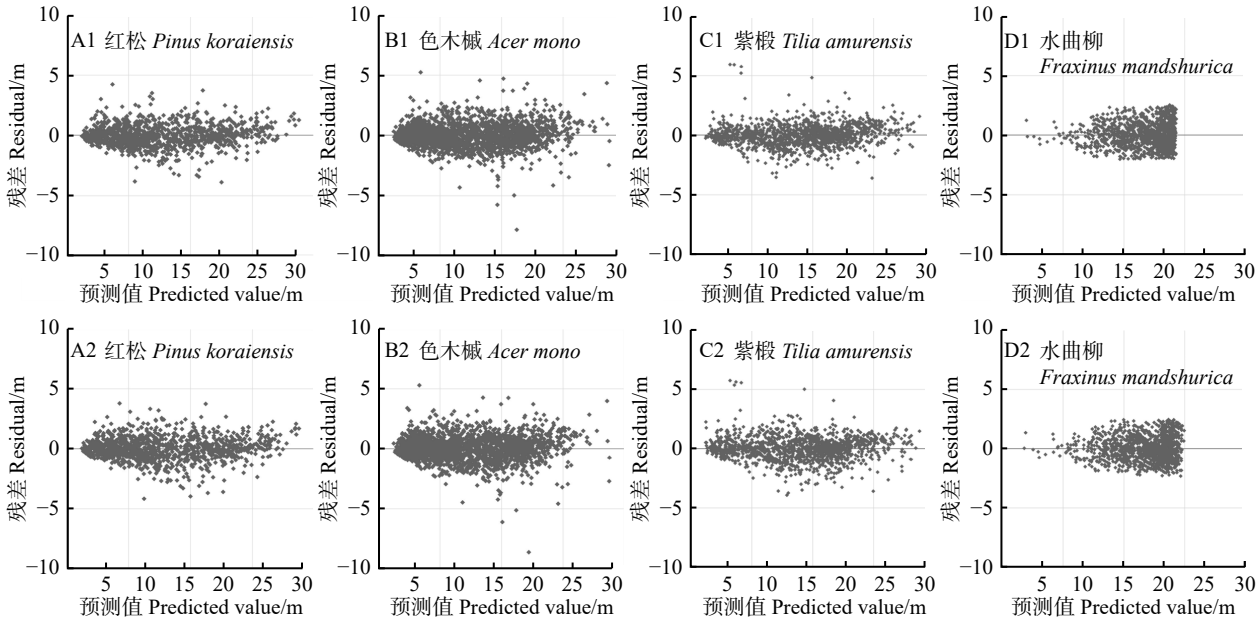


图 1 基础混合效应模型(A1、B1、C1、D1)和广义混合效应模型(A2、B2、C2、D2)的残差分布图

Fig. 1 Residual plots of base mixed-effect models (A1, B1, C1, D1) and generalized mixed-effect models (A2, B2, C2, D2)

表 12　树高−胸径模型统计检验

Tab. 12　Goodness-of-fit statistics of tree height-DBH models

树种 Tree species	指标 Index	基础模型 Base model	广义模型 Generalized model	中位数回归 Median quantile regression		非线性混合效应 Nonlinear mixed-effects	
				简单模型 Simple model	广义模型 Generalized model	简单模型 Simple model	广义模型 Generalized model
红松 *Pinus koraiensis*	MAE	2.128	2.119	2.094	2.163	1.600	1.856
	RMSE	2.963	3.024	2.963	3.063	2.216	2.635
色木槭 *Acer mono*	MAE	2.018	1.906	2.020	2.052	1.620	1.712
	RMSE	2.867	2.782	2.898	2.944	2.341	2.529
紫椴 *Tilia amurensis*	MAE	2.571	1.685	2.549	2.697	1.609	1.635
	RMSE	3.471	2.349	3.540	3.754	2.058	2.206
水曲柳 *Fraxinus mandshurica*	MAE	2.041	2.037	2.050	2.051	2.039	2.049
	RMSE	2.526	2.517	2.560	2.561	2.528	2.534

相比分位数回归具有更加稳健的参数估计值，模型稳定性和灵活性也更强，能够避免选取数据的人为影响。在不同分位点拟合树高−胸径模型，同时基于参数估计和建模样本计算三分位数、五分位数和九分位数组合的插值系数，王君杰等[34]最近发现中位数回归优于非线性回归；本文与之不同发现中位数回归与非线性回归的拟合效果基本一致。经过对 11 个非线性模型的评价比较，选择 Richards 方程作为基础模型进行分位数回归，其中中位数回归拟合效果最好，4 个树种的 R^2 值均大于 0.6，对树高的预测能力不逊于非线性回归，这表明 Richards 方程能广泛适用树高−胸径模型的拟合。

三分位数组合的预测效果均优于五分位数和九分位数组合，尤其在红松树高的预测中，两类分位数组合的预测精度相较中位数和非线性回归均有提高；以往研究显示五分位数组合对树高曲线和削度方程的拟合要优于三分位数组合，但都没有考虑抽取方案对分位数组合的影响[35]。因此抽取方案的不同会导致分位数组合之间的差异不同，总体上三分位数组合（$\tau = 0.1, 0.5, 0.9$）的预测能力较好且效果稳定，本研究中选取多个分位数组合也并未使预测结果有明显提高[32]。

3.2　引入林分变量和随机效应对树高模型的影响

在林分生长过程中，有学者利用含有优势木胸径、树高和林分密度等因子的模型来描述树高−胸径关系，模型拟合结果得到提升[36]。为了比较不同样方间树高−胸径的变化规律，本研究对胸高断面积、林分密度和优势木高 3 个林分变量进行 Pearson 相关性分析，发现引入变量的广义模型的拟合效果均优于基础模型，当引入 1 个变量时，红松、色木槭、紫椴和水曲柳均仅增加优势木高变量的模型可以收敛，部分其他林分变量与树高有显著关系，但含优势木高的模型相较含林分密度或胸高断面积的广义模型对树高−胸径的关系影响更大；4 个树种中水曲柳的广义模型相较基础模型的拟合效果提升幅度最小，R^2 仅提高 0.01，这可能与水曲柳的直径分布有关，幼树数量较少，成熟林木较多，在林分中多位于林冠层，内部资源和空间竞争相对稳定，处于优势地位，这类大个体的树高不易受林分变量较大影响。这说明即使是在具有相似地形和气候条件的林分中，不同树种的树高−胸径关系可能不同。森林群落中特定的环境条件和树木之间的复杂联系可能会导致树高−胸径生长关系不同[37]。优势木高反映了立地质量以及林分生长和收获，优势木高的参数估计值为正值，这表明由于立地质量的不同，优势木高与树高之间呈正相关关系，引入优势木高的广义模型更适用于树高−胸径模拟的拟合。

基于基础模型和各树种的最优广义模型，对应构建以样方水平为随机效应的混合模型，比较不同随机参数下各树种树高模型的拟合效果。研究发现，红松、紫椴和水曲柳基础混合效应模型均以增加随机参数 u_1，广义混合效应模型中红松增加参数 u_1，紫椴和水曲柳将 u_3 同时作为随机参数，拟合效果最优；色木槭基础混合效应模型增加随机参数 u_1、u_2，广义混合效应模型则是将 u_1、u_3 同时作为随机参数，预测精度最高。通过对各模型进行检验评价，结果表明混合效应模型对树高的预测效果优于固定效应模型，这与王冬至等[38]的研究结果一致，表明考虑样方水平的随机影响能更好地描述树高生长规律。此外，本研究的模型评价和检验结果显示基础混合效应模型的预测能力略优于广义混合效应模型，但未有显著差异，同样的情况也出现在陈浩等[39]对马尾松和周晏平等[40]对樟子松的研究结果中，说明在增加样方效应后，林分变量对树高−胸径关系的影响被样方水平所涵盖，对混合效应模型的影响较弱。在 Zang 等[41]对中国北部和东北落叶松人工林树

高–胸径关系的研究中，基础混合效应模型比其他模型提供了更好的拟合和更稳定的估计，表明不需要额外的树木和林分变量。

4 结　　论

本研究基于分位数回归和混合效应模型对阔叶红松林中优势树种红松、色木槭、紫椴和水曲柳的树高–胸径关系进行模型拟合，在 11 个非线性候选模型中确定 Richards 方程为基础模型，并引入优势木高构建广义模型，模型拟合结果得到较大提高；4 个树种基于中位数回归的基础模型和广义模型拟合精度与基础模型的非线性回归相近。

加入样方水平为随机效应后，混合效应模型相较固定效应模型更具优越性，模型预测能力得到显著提高，而基础混合模型和广义混合模型的拟合优度差异较小。这表明林分变量和随机效应对树高模型的影响是显著的，在不考虑样方效应的情况下，可选择广义模型作为东北地区针阔混交林树高模型；若考虑样方效应，基础混合效应模型能提供更稳定的预测趋势和拟合结果。

参 考 文 献

[1] Koenker R, Bassett G. Regression quantiles[J]. Econometrica: Journal of the Econometric Society, 1978, 46(1): 33−50.

[2] 田德超, 李凤日, 董利虎. 依据分位数回归建立的长白落叶松潜在最大冠幅预测模型 [J]. 东北林业大学学报, 2019, 47(8): 41−46.
Tian D C, Li F R, Dong L H. Potential maximum crown width prediction model of *Larix olgensis* by quantile regression[J]. Journal of Northeast Forestry University, 2019, 47(8): 41−46.

[3] Özçelik R, Cao Q V, Trincado G, et al. Predicting tree height from tree diameter and dominant height using mixed-effects and quantile regression models for two species in Turkey[J]. Forest Ecology and Management, 2018, 419−420: 240−248.

[4] 高慧淋. 东北林区针叶树树冠轮廓及特征因子模拟 [D]. 哈尔滨: 东北林业大学, 2017.
Gao H L. Modelling crown profile and characteristic attributes of coniferous tree species in northeast China[D]. Harbin: Northeast Forestry University, 2017.

[5] 赵梦草, 顾海燕. 用分位数回归建立的混交林树高生长模型 [J]. 东北林业大学学报, 2020, 48(2): 24−28, 53.
Zhao M C, Gu H Y. Tree height growth model of mixed forest using quantile regression[J]. Journal of Northeast Forestry University, 2020, 48(2): 24−28, 53.

[6] 王君杰, 夏宛琦, 姜立春. 基于分位数回归和哑变量模型的大兴安岭兴安落叶松树高–胸径模型 [J]. 中南林业科技大学学报, 2020, 40(9): 24−32, 40.
Wang J J, Xia W Q, Jiang L C. Height-diameter models for *Larix gmelinii* in Daxing' anling based on quantile regression and dummy variable model[J]. Journal of Central South University of Forestry & Technology, 2020, 40(9): 24−32, 40.

[7] Cao Q V, Wang J. Evaluation of methods for calibrating a tree taper equation[J]. Forest Science, 2015, 61(2): 213−219.

[8] 佟艺玟, 陈东升, 冯健, 等. 基于线性分位数混合效应的辽东山区红松冠幅模型 [J]. 应用生态学报, 2022, 33(9): 2321−2330.
Tong Y W, Chen D S, Feng J, et al. Crown width model for planted Korean pine in eastern Liaoning mountains based on mixed effect linear quantile[J]. Chinese Journal of Applied Ecology, 2022, 33(9): 2321−2330.

[9] Zhang X F, Fu L Y, Sharma R P, et al. A nonlinear mixed-effects height-diameter model with interaction effects of stand density and site index for *Larix olgensis* in Northeast China[J]. Forests, 2021, 12(11): 1460.

[10] 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 Y F, Wang J J, Jiang L C. Construction of the height to crown base mixed model for Korean pine[J]. Journal of Beijing Forestry University, 2020, 42(9): 28−36.

[11] Schmidt M, Kiviste A, Gadow K. A spatially explicit height-diameter model for Scots pine in Estonia[J]. European Journal of Forest Research, 2011, 130(2): 303−315.

[12] 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, 24(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]. Forestry Research, 2011, 24(3): 285−291.

[13] Temesgen H, Zhang C H, Zhao X H. Modelling tree height-diameter relationships in multi-species and multi-layered forests: A large observational study from Northeast China[J]. Forest Ecology and Management, 2014, 316: 78−89.

[14] Stage A R. Prediction of height increment for models of forest growth[M/OL]. Ogden: Intermountain Forest and Range Experiment Station, Forest Service, U.S. Dept. of Agriculture, 1975 [2022−05−19]. https://doi.org/10.5962/bhl.title.69034.

[15] Curtis R. Height-diameter and height-diameter-age equations for second-growth Douglas-fir[J]. Forest Science, 1967, 13(4): 365−375.

[16] Wykoff W R, Crookston N L, Stage A R. User's guide to the stand prognosis model[R/OL]. Gen. Tech. Rep. INT-GTR-133. Ogden, UT: U.S. Department of Agriculture, Forest Service, Intermountain Forest and Range Experiment Station, 1982: 112. https://doi.org/10.2737/INT-GTR-133.

[17] Meyer H A. A mathematical expression for height curves[J]. Journal of Forestry, 1940, 38(5): 415−420.

[18] Schumacher F X. New growth curve and its application to timber-yield studies[J]. Journal of Forestry, 1939, 37: 819−820.

[19] Richards F J. Aflexible growth function for empirical use[J]. Journal of Experimental Botany, 1959, 10(2): 290−300.

[20] Sibbesen E. Some new equations to describe phosphate sorption

by soils[J]. Journal of Soil Science, 1981, 32(1): 67−74.

[21] Ratkowsky D A. Handbook of non-linear regression[M]. New York: Marcel and Dekker, 1990: 120.

[22] Winsor C P. The gompertz curve as a growth curve[J]. PNAS, 1932, 18(1): 1−8.

[23] Prodan M. Growth functions in: forest biometrics[M]. Oxford: Pergamon Press, 1968.

[24] Pearl R, Reed L J. On the rate of growth of the population of United States since 1790 and its mathematical representation[M]// Smith D P, Keyfitz N. Mathematical demography: selected papers. Berlin: Springer, 1977: 341−347.

[25] Zhang L, Bi H, Gove J H, et al. A comparison of alternative methods for estimating the self-thinning boundary line[J]. Canadian Journal of Forest Research, 2005, 35(6): 1507−1514.

[26] Mehtätalo L, Gregoire T G, Burkhart H E. Comparing strategies for modeling tree diameter percentiles from remeasured plots[J]. Environmetrics: the Official Journal of the International Environmetrics Society, 2008, 19(5): 529−548.

[27] 辛士冬, 姜立春. 利用分位数回归模拟人工樟子松树干干形 [J]. 北京林业大学学报, 2020, 42(2): 1−8.
Xin S D, Jiang L C. Modeling stem taper profile for *Pinus sylvestris* plantations using nonlinear quantile regression[J]. Journal of Beijing Forsetry University, 2020, 42(2): 1−8.

[28] Ma A Y, Miao Z, Xie L F, et al. Crown width prediction for *Larix olgensis* plantations in Northeast China based on nonlinear mixed-effects model and quantile regression[J]. Trees, 2022, 36(6): 1761−1776.

[29] Ciceu A, Garcia-Duro J, Seceleanu I, et al. A generalized nonlinear mixed-effects height–diameter model for Norway spruce in mixed-uneven aged stands[J/OL]. Forest Ecology and Management, 2020, 477 [2020−08−20]. https://doi.org/10.1016/j.foreco.2020.118507.

[30] Huang S, Price D, Titus S J. Development of ecoregion-based height-diameter models for white spruce in boreal forests[J]. Forest Ecology and Management, 2000, 129(3): 139−141.

[31] 臧 颢, 雷相东, 张会儒. 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.
Zang H, Lei X D, Zhang H R. Nonlinear mixed-effects height-diameter model of *Pinus koraiensis*[J]. Journal of Beijing Forestry University, 2016, 38(6): 8−16.

[32] 王君杰, 姜立春. 基于线性分位数组合的兴安落叶松冠幅预测 [J]. 南京林业大学学报 (自然科学版), 2021, 45(5): 161−170.
Wang J J, Jiang L C. Predicting crown width for *Larix gmelinii* based on linear quantiles groups[J]. Journal of Nanjing Forestry University (Natural Sciences Edition), 2021, 45(5): 161−170.

[33] 高慧淋, 董利虎, 李凤日. 基于分位数回归的长白落叶松人工林最大密度线 [J]. 应用生态学报, 2016, 27(11): 3420−3426.
Gao H L, Dong L H, Li F R. Maximum density-size line for *Larix olgensis* plantations based on quantile regression[J]. Chinese Journal of Applied Ecology, 2016, 27(11): 3420−3426.

[34] 王君杰, 姜立春. 利用分位数组合预测兴安落叶松枝下高 [J]. 北京林业大学学报, 2021, 43(3): 9−17.
Wang J J, Jiang L C. Predicting height to crown base for *Larix gmelinii* using quantile groups[J]. Journal of Beijing Forestry University, 2021, 43(3): 9−17.

[35] Özçelik R, Diamantopoulou M J, Trincado G, et al. Evaluation of potential modeling approaches for Scots pine stem diameter prediction in north-eastern Turkey[J]. Computers and Electronics in Agriculture, 2019, 162: 773−782.

[36] Xu H, Sun Y J, Wang X J, et al. Height-diameter models of Chinese fir (*Cunninghamia lanceolata*) based on nonlinear mixed effects models in Southeast China[J]. Advance Journal of Food Science and Technology, 2014, 6(4): 445−452.

[37] Cui K D, Wu X Y, Zhang C Y, et al. Estimating height-diameter relations for structure groups in the natural forests of Northeastern China[J/OL]. Forest Ecology and Management, 2022, 519 [2022−06−09]. https://doi.org/10.1016/j.foreco.2022.120298.

[38] 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.
Wang D Z, Zhang D Y, Zhang Z D, et al. Height-diameter relationship for conifer mixed forest based on nonlinear mixed-effects model[J]. Scientia Silvae Sinicae, 2016, 52(1): 30−36.

[39] 陈 浩, 罗 扬. 马尾松树高−胸径非线性混合效应模型构建 [J]. 森林与环境学报, 2021, 41(4): 439−448.
Chen H, Luo Y. Construction of nonlinear mixed effect height-diameter model for *Pinus massoniana*[J]. Journal of Forest and Environment, 2021, 41(4): 439−448.

[40] 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): 75−91.
Zhou Y P, Lei Z Y, Zhao G J, 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): 75−91.

[41] Zang H, Lei X D, Zeng W S. 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.

（责任编辑　范　娟
责任编委　臧润国）