

黔中马尾松木荷混交林树高-胸径模型

冉佳璇, 戚玉娇

(贵州大学 林学院, 贵州 贵阳 550025)

摘要: 【目的】建立马尾松 *Pinus massoniana*-木荷 *Schima superba* 混交林树高-胸径模型, 将树种作为哑变量引入模型, 考虑模型残差空间自相关和异质性, 为混交林树高-胸径模型构建和科学经营提供理论依据。【方法】基于贵州省开阳县马尾松-木荷混交林 727 组树高-胸径调查数据, 构建普通最小二乘法模型 (OLS)、广义可加模型 (GAM)、线性混合模型 (LMM)、地理加权回归模型 (GWR) 和地理加权回归克里格模型 (GWRK) 的树高-胸径全林木模型, 在此基础上, 将树种作为哑变量引入, 选择全局莫兰指数 (Moran' I)、局域 Moran' I 和组内方差分析 5 种模型残差空间自相关与空间异质性, 并采用决定系数 (R^2)、均方误差 (MSE) 和赤池信息准则 (AIC) 对模型进行评价。【结果】①马尾松-木荷混交林全林木基础模型的拟合精度从低到高依次为 OLS、GAM、LMM、GWR、GWRK。②将树种作为哑变量引入模型后, 各模型拟合精度均高于全林木基础模型。③OLS 和 GAM 模型残差的全局 Moran' I 在 $\alpha=0.05$ 水平下显著 ($Z>1.96$), 局域 Moran' I 分布图中存在较多热点, 表现出强烈的空间自相关。而 LMM、GWR 和 GWRK 模型残差全局 Moran' I 在 $\alpha=0.05$ 水平下不显著 ($-1.96\leq Z\leq 1.96$), 且在局域 Moran' I 分布图中存在较多冷点, 说明模型残差空间自相关已被消除。④5 种模型残差的组内方差均表现随着滞后距离增大而增大的趋势, 但 GWR 和 GWRK 模型具有更小的组内方差, 能较好地降低模型残差空间的异质性。【结论】OLS 和 GAM 模型拟合精度不高, 并且不能消除模型残差空间自相关和异质性, 因此不是用来建立树高-胸径模型的最佳选择。LMM、GWR 和 GWRK 模型在提高模型拟合精度和降低空间自相关性方面表现良好, 但 GWR 和 GWRK 模型在降低空间异质性方面显著, 是最适合的树高-胸径模型。图 2 表 3 参 38

关键词: 马尾松; 木荷; 混交林; 树高-胸径模型; 模型残差; 空间自相关; 空间异质性

中图分类号: S758.5 文献标志码: A 文章编号: 2095-0756(2024)02-0343-10

Height-diameter model of *Pinus massoniana* and *Schima superba* mixed forest in central Guizhou Province

RAN Jiakuan, QI Yujiao

(College of Forestry, Guizhou University, Guiyang 550025, Guizhou, China)

Abstract: [Objective] To establish a tree height-diameter model for mixed forests of *Pinus massoniana* and *Schima superba*, introduce tree species as dummy variables into the model, and consider the spatial autocorrelation and heterogeneity of residuals of the model, in order to provide theoretical basis for the construction of the tree height-diameter model of mixed forests and the scientific management of mixed forests. [Method] Based on the survey data of 727 groups of tree height-diameter in mixed forests of *P. massoniana* and *S. superba* in Kaiyang County, Guizhou Province, we constructed ordinary least squares (OLS), generalized additive model (GAM), linear mixed model (LMM), geographically weighted regression model (GWR), and geographically weighted regression kriging (GWRK) models for tree height-diameter-whole-forest model, on

收稿日期: 2023-06-12; 修回日期: 2023-10-18

基金项目: 贵州省科技计划项目后补助计划项目 (黔科合平台人才 [2018] 5261); 国家重点研发计划项目 (2017YFD0600302)

作者简介: 冉佳璇 (ORCID: 0009-0007-6288-279X), 从事森林经理学研究。E-mail: 1007935488@qq.com。通信作者: 戚玉娇 (ORCID: 0000-0002-5769-2234), 教授, 博士, 从事森林经理学和森林生态学研究。E-mail: yjq@gzu.edu.cn

the basis of which tree species was introduced as a dummy variable, and five model residuals spatial autocorrelation and spatial heterogeneity were selected for global Moran' I , local Moran' I , and intra-block variance analyses with the coefficients of determination (R^2), mean squared error (MSE), and the Akaike information criterion (AIC) to evaluate the models. [Result] (1) The fitting accuracies of the whole-forest base models of *P. massoniana*-*S. superba* mixed forest were OLS < GAM < LMM < GWR < GWRK in descending order. (2) The fitting accuracies of the models were higher than those of the whole-forest base models after introducing tree species as a dummy variable into the models. (3) The global Moran' I of the residuals of the OLS and GAM models was significant at the $\alpha=0.05$ level ($Z > 1.96$), and there were more hot spots in the local Moran' I distribution maps, which showed strong spatial autocorrelation. In contrast, the global Moran' I of the residuals of the LMM, GWR and GWRK models is insignificant at the $\alpha=0.05$ level ($-1.96 \leq Z \leq 1.96$) and there are more cold spots in the local Moran' I distribution plot, indicating that spatial autocorrelation of the model residuals has been eliminated. (4) The intra-block variance of the residuals of the five models show a tendency to increase with the lag distance, but the GWR and GWRK models have smaller intra-block variance, which can better reduce the heterogeneity of the model residual space. [Conclusion] The OLS and GAM models do not have high fitting accuracy and cannot eliminate spatial autocorrelation and heterogeneity of model residuals, so they are not the best choices for modeling tree height-diameter. The LMM, GWR, and GWRK models perform well in improving the model fitting accuracy and decreasing the spatial autocorrelation, but the GWR and GWRK models are more significant in decreasing the spatial heterogeneity, and they are the most appropriate of the tree height-breast diameter models. [Ch, 2 fig. 3 tab. 38 ref.]

Key words: *Pinus massoniana*; *Schima superba*; mixed forest; tree height-diameter model; model residuals; spatial autocorrelation; spatial heterogeneity

树高和胸径是森林连续清查和管理中的2个重要因子,相较于胸径的测量,树高的测量具有耗时、成本高且误差大等缺点^[1]。因此,研究者将树高-胸径模型作为一种常见、快速、便捷的方法预测林木高度^[2],估算森林的生长与产量^[3]、生物量和碳储量^[4-5]、林木材积^[6]和立地指数^[7]等。普通最小二乘法(ordinary least squares, OLS)常被学者作为基础模型预测树高,但该模型在拟合具有空间分布特征的变量时,消除空间自相关的能力较弱^[8-10]。广义可加模型(generalized additive model, GAM)属于数据驱动模型,采用链接函数建立响应变量和解释变量的平滑函数之间的关系,优势在于能够处理响应变量和解释变量之间高度非线性和非单调的关系^[11]。线性混合模型(linear mixed model, LMM)由固定效应和随机效应两部分组成,可以同时反映整体变化规律和不同个体在整体中的变化,方差协方差结构还能反映数据之间的自相关和异质性^[12]。因此,与传统模型OLS相比,GAM和LMM模型能提高模型拟合精度^[9,13]。地理加权回归模型(geographically weighted regression, GWR)作为能够显著消除模型空间自相关性的局域模型,研究对象位置不同导致权函数不同,进而使每个位置上参数随着空间的变化而变化,产生的结果更加精确,该模型已被运用于很多研究领域。近年来,学者们从各方面对GWR模型进行拓展,如李泽坤等^[14]将时间维度引入GWR模型回归参数部分拓展出时空地理加权回归模型(geographically and temporally weighted regression, GTWR)分析了浙江省各影响因素对碳排放总量影响的时空异质性,陈科屹等^[15]对模型残差进行普通克里格插值,拓展出地理加权回归克里格模型(geographically weighted regression kriging, GWRK)对吉林省汪清林业局浪溪林场天然次生林碳储量空间分布进行了研究,吴子豪等^[16]从解决自变量差异化方面拓展出多尺度地理加权回归模型(multi-scale geographic weighted regression model, MGWR),探明了湖北省武汉市黄陂区土壤镉在不同空间位置的影响因子和污染来源等。

马尾松 *Pinus massoniana* 作为中国南方主要先锋造林树种,具有速生、丰产、适应性强等特点,但马尾松纯林存在结构简单、生态系统不稳定,容易造成土壤退化,引发水土流失等问题^[17]。因此,常把马尾松与其他树种进行混交造林,提高林分稳定性。木荷 *Schima superba* 因为其材质优良,适应能力

强，能在酸性土壤和贫瘠山地等不利的条件下生长，被列入国家珍贵用材发展名录^[18]。有研究发现：马尾松-木荷混交林在提高林地生产力，改善林地生态环境，促进林木生长^[19]等方面优于马尾松纯林，但与纯林相比，混交林的复杂性和多样性使得建立树高-胸径模型更为复杂。目前，许多学者做了关于混交林树高-胸径模型的研究^[20-21]。但对于马尾松-木荷混交林研究主要在不同混交比例对林分生长的影响^[22]、林下植被物种组成及多样性^[23]、凋落物和水土保持性能^[24]等方面，对树高-胸径模型鲜有研究。因此，本研究以贵州省开阳县马尾松-木荷混交林为研究对象，对比分析 OLS、GAM、LMM、GWR 和 GWRK 树高-胸径回归哑变量模型拟合效果及降低模型残差空间自相关和异质性的程度，为开阳县马尾松-木荷混交林科学经营管理提供理论依据。

1 材料与方法

1.1 研究区概况与样地调查

研究区位于贵州省贵阳市开阳县，该县西连安顺市与毕节市，东与南连接黔南布依族苗族自治州，北接遵义。位置为 26°11'~27°22'N，106°07'~107°16'E，地势南北高、中间低、平均海拔为 1 200 m，地形以山地、丘陵为主、喀斯特地貌分布广泛，占全区面积的 71.8%。气候类型为亚热带高原季风湿润性气候，年均气温为 15.3 ℃，年均降水量为 1 196.9 mm。所选研究区森林类型为天然马尾松-木荷混交林，林内树种有马尾松、木荷、水青冈 *Fagus longipetiolata*、光叶海桐 *Pittosporum glabratum*、云贵鹅耳枥 *Carpinus pubescens* 和杉木 *Cunninghamia lanceolata* 等。

2017 年 8 月，在研究区马尾松-木荷混交林内，根据 CTFS(Centre for Tropical Forest Science) 样地建设标准方法设置 40 m×50 m 的固定样地，并将样地划分为 20 个 10 m×10 m 的小样方，对每个样方内所有胸径≥1 cm 的树木进行每木检尺，即记录树种名称，测量胸径、树高、冠幅、坐标等因子。样地内林木调查因子描述性统计见表 1。

表 1 样地林分调查因子描述性统计

Table 1 Descriptive statistics of forest survey factors in sample plots

树种	胸径/cm	树高/m	密度/(株·hm ⁻²)	胸高断面积/(m ² ·hm ⁻²)
全林木	7.3±7.9	7.2±5.8	3 635	6.580 0
马尾松 <i>Pinus massoniana</i>	27.9±12.8	22.2±6.9	225	3.320 0
木荷 <i>Schima superba</i>	6.6±4.2	6.3±2.8	1 545	1.510 0
水青冈 <i>Fagus longipetiolata</i>	8.3±7.6	9.5±6.2	290	0.580 0
杉木 <i>Cunninghamia lanceolata</i>	8.1±6.4	8.0±5.6	235	0.390 0
枫香 <i>Liquidambar formosana</i>	12.0±10.7	12.3±8.5	45	0.180 0
云贵鹅耳枥 <i>Carpinus pubescens</i>	4.8±3.3	5.8±2.8	290	0.150 0
山樱花 <i>Cerasus serrulata</i>	18.2±10.5	19.5±6.1	15	0.100 0
光叶海桐 <i>Pittosporum glabratum</i>	2.8±1.5	2.4±0.3	330	0.050 0
光皮桦 <i>Betula luminifera</i>	10.7±4.3	14.6±1.5	25	0.050 0
白栎 <i>Quercus fabri</i>	3.6±4.5	4.9±3.9	95	0.050 0
老鼠矢 <i>Symplocos stellaris</i>	3.6±2.7	4.2±2.5	145	0.040 0
响叶杨 <i>Populus adenopoda</i>	14.4±8.1	15.4±4.1	10	0.040 0
虎皮楠 <i>Daphniphyllum oldhamii</i>	7.7±1.7	6.9±2.1	55	0.020 0
栲 <i>Castanopsis fargesii</i>	3.1±1.6	4.9±2.7	110	0.020 0
山胡椒 <i>Lindera glauca</i>	2.4±0.9	5.2±2.6	65	0.006 0
穗序鹅掌柴 <i>Heptapleurum delavayi</i>	3.3±1.1	3.1±1.7	30	0.005 0
栓皮栎 <i>Quercus variabilis</i>	5.2±0.3	8.7±0.0	10	0.004 0
川榛 <i>Corylus heterophylla</i> var. <i>sutchuenensis</i>	2.4±0.9	5.8±2.6	35	0.003 0
山矾 <i>Symplocos sumuntia</i>	2.3±1.2	2.7±0.8	25	0.002 0
象鼻藤 <i>Dalbergia mimosoides</i>	2.7±0.6	8.4±1.1	10	0.001 0
灰毛大青 <i>Clerodendrum canescens</i>	1.6±0.3	4.6±0.6	15	0.000 6

说明：胸径和树高数值为平均值±标准差。

1.2 模型

根据调查数据(表1)可知林内主要树种为马尾松和木荷。因此,将树种作为哑变量引入OLS、GAM、LMM、GWR和GWRK模型中,分别构建马尾松、木荷和其他树种树高-胸径模型。树种分类用定性代码0和1表示,第*i*个树种表示为 S_i ,定性数据 S_i 转换为(0,1)形式; S_i 为0或1,当 x 是第*i*个树种时为1,否者为0。其中*i*=1、2、3; S_1 、 S_2 、 S_3 分别是马尾松、木荷和其他树种的定性代码。将没有引入哑变量的模型作为全林木基础模型。

1.2.1 普通最小二乘法(OLS) 采用幂函数($Y=aX^b$)的对数变换形式($\ln Y=a+b\ln X$)作为本研究的基础模型。其模型表达式如下:

$$\ln H_t = a + b \ln D_{BH} + \varepsilon. \quad (1)$$

式(1)中, H_t 为树高(m), D_{BH} 为胸径(cm), a 与 b 为模型参数, ε 为模型误差。

1.2.2 线性混合模型(LMM) 增加样方水平的单水平随机效应,在模型参数 a 和 b 的位置上分别增加随机参数 μ_1 和 μ_2 ,如式(2)所示:

$$\ln H_t = (a + \mu_1) + (b + \mu_2) \ln D_{BH} + \varepsilon. \quad (2)$$

式(2)中, a 和 b 为模型参数, μ_1 和 μ_2 为随机参数, ε 为模型随机误差。

1.2.3 广义可加模型(GAM) GAM是广义线性模型的半参数拓展,唯一的基本假设是平滑函数是可加的^[25]。将 $\ln D_{BH}$ 作为解释变量, $\ln H_t$ 作为响应变量,平滑回归项采用平滑样条函数构建GAM,其表达式:

$$\ln H_t = \alpha + s(\ln D_{BH}) + \varepsilon. \quad (3)$$

式(3)中, s 为平滑样条函数, ε 为模型误差, α 为截距。

1.2.4 地理加权回归模型(GWR) GWR模型作为传统回归方法的拓展,将样本点的地理位置信息加入建模,使得模型既能描述响应变量与解释变量间的关系,又能消除模型残差的空间自相关^[26]。使用Gauss函数作为权函数,构建开阳县马尾松-木荷混交林树高-胸径模型。

$$\ln H_t = \beta_0(u_i, v_i) + \sum_{k=1}^p \beta_k(u_i, v_i) X_{ki}(\ln D_{BH}) + \varepsilon. \quad (4)$$

式(4)中, $\ln H_t$ 为响应变量, (u_i, v_i) 为第*i*个样本点的坐标, β_0 为第*i*个样本点的常数估计值, X_{ki} 为第*k*个自变量在第*i*个样本点的值, $\beta_k(u_i, v_i)$ 为第*i*个样本点上的第*k*个回归参数,是关于地理位置的函数, p 为用于建模的林木总数, ε 为模型误差。

1.2.5 地理加权回归克里格模型(GWRK) GWRK模型是GWR模型的延伸与拓展,将GWR与克里格插值相结合的模型,通过GWR回归残差进行空间插值,然后将插值结果和GWR回归估计值相加,从而获得GWRK估计值^[27]。

1.3 模型残差评价

1.3.1 模型残差的空间自相关性 分别采用全局与局域莫兰指数(Moran's I)评价模型残差的空间分布^[9]。全局Moran's I 是用来描述所有的空间单元在整个区域上与周边地区的平均关联程度,取值为-1~1,多数统计检验在开始时都先确定零假设,在计算全局Moran's I 时通常会给出 Z 值,如果 $Z > 1.96$ 或者 $Z < -1.96$,拒绝零假设,表示空间分布可能是聚集模式或分散模式;若 $-1.96 \leq Z \leq 1.96$,接受零假设,表示空间分布可能是随机过程产生的结果, Z 能够判断*n*个样点内的观测值之间是否存在空间相关性^[28]。而局域Moran's I 可以进一步研究局域空间聚集现象,评价模型残差是否存在局域聚集情况,分析模型残差局域空间自相关性。

1.3.2 模型残差的空间异质性 空间异质性作为生态系统的主要属性之一,是产生空间格局的主要原因^[29]。使用模型残差组内方差评价模型残差空间异质性^[28,30]。

1.4 模型检验

模型评价指标选用决定系数(R^2)、均方误差(MSE)和赤池信息准则(AIC)进行模型比较,除了决定系数值越大越好外,其他指标均是越小越好。综合以上评价指标选择最优模型。

1.5 数据分析

使用R语言的minpack.lm包、mgcv包、lme4包和GWmodel包分别对OLS、GAM、LMM和GWR

模型进行建模分析，随机选择 581 株 (80%) 林木作为建模数据，剩余的 146 株 (20%) 林木作为检验数据。用 ArcMap 10.6 对模型残差进行普通克里格插值，ROOKCASE (用 VBA 编写的 Excel 宏文件) 计算全局和局域 Moran's I ，作图在 Excel 中完成。

2 结果与分析

2.1 林木树高-胸径模型拟合效果

全林木基础模型拟合结果见表 2，5 种模型的 R^2 为 0.63~0.80，MSE 为 0.08~0.16，AIC 为 462.77~759.23，都能很好地反映树高与胸径之间的关系。将树种作为哑变量引入后，模型拟合精度整体高于全林木基础模型，5 种回归模型的 R^2 提高了 0.01~0.06，MSE 和 AIC 也有不同程度的降低，全林木基础模型和哑变量模型的 3 种拟合指标从劣到优均依次为 OLS、GAM、LMM、GWR、GWRK。OLS 和 LMM 哑变量模型参数估计值均在 0.001 水平上显著，GWR 模型的参数估计值为一个范围，能很好地表现林木间的差异 (表 3)。

表 2 5 种回归模型拟合统计和全局 Moran's I (Z)

Table 2 Model fitting statistics and Global Moran's I (Z-value) for five regression models

模型类型	模型	R^2	MSE	AIC	全局 Moran's I (Z)
哑变量模型	普通最小二乘法(OLS)	0.65	0.16	723.97	0.2494(26.521 0)
	线性混合模型(LMM)	0.78	0.09	506.29	-0.0065(-0.5448)
	广义可加模型(GAM)	0.66	0.15	714.54	0.2047(22.0847)
	地理加权回归模型(GWR)	0.85	0.07	271.83	-0.0125(-1.1784)
	地理加权克里格回归模型(GWRK)	0.86	0.06	260.66	-0.0095(-1.0208)
全林木基础模型	普通最小二乘法(OLS)	0.63	0.16	759.23	0.2878(30.1583)
	线性混合模型(LMM)	0.77	0.10	522.88	-0.0071(-0.6050)
	广义可加模型(GAM)	0.65	0.15	723.95	0.2797(29.6680)
	地理加权回归模型(GWR)	0.79	0.09	488.08	-0.0144(-1.3773)
	地理加权克里格回归模型(GWRK)	0.80	0.08	462.77	-0.0122(-0.4704)

说明： R^2 为决定系数，MSE 为均方误差，AIC 为赤池信息准则，全局 Moran's I (Z) 为空间自相关性评价指标。

表 3 OLS、LMM 和 GWR 哑变量模型参数估计

Table 3 OLS, LMM and GWR dummy variable model parameter estimation

模型类型	拟合参数	估计值	标准误	P
普通最小二乘法 (OLS)	a	1.200 7	0.087 6	<0.001
	b	0.581 5	0.020 5	<0.001
	木荷	-0.440 0	0.069 7	<0.001
	其他	-0.398 2	0.073 0	<0.001
线性混合模型 (LMM)	a	0.969 6	0.093 1	<0.001
	b	0.637 6	0.017 8	<0.001
	木荷	-0.282 8	0.060 2	<0.001
	其他	-0.249 7	0.062 2	<0.001
地理加权回归模型 (GWR)	a	-1.003 4~2.401 6		
	b	0.285 9~1.314 6		
	木荷	-1.241 3~-0.842 6		
	其他	-1.245 5~1.368 1		

说明： a 和 b 为模型拟合参数； P 表示拟合参数的显著性，其中 $P < 0.001$ 表示在 0.001 水平上显著。GAM 模型为非参数模型；GWRK 模型是将基础模型残差进行普通克里格插值，也没有参数估计值。

2.2 模型残差分析

2.2.1 模型残差空间自相关性 由表 2 可知：OLS 和 GAM 模型残差全局 Moran's I 在 $\alpha=0.05$ 的显著水平下， $Z > 1.96$ ，存在显著的空间自相关性。而 LMM、GWR 和 GWRK 模型残差具有更低的全局 Moran's I ，

且在 $\alpha=0.05$ 条件下，模型残差全局 Moran's I 的 Z 在 $-1.96\sim 1.96$ 范围，模型残差空间自相关性不显著。以 5 m 为间隔，计算 5~20 m 滞后距离内模型残差全局 Moran's I (图 1)。OLS 和 GAM 模型残差全局 Moran's I 随着滞后距离增加不断降低，而 LMM、GWR 和 GWRK 模型残差全局 Moran's I 在各滞后距离处在 0 上下小幅度波动，且值相接近。

OLS 和 GAM 模型残差的局域 Moran's I 整体分布不均匀且存在大而多的黑色气泡 (正局域 Moran's I) (图 1B~F)，说明这 2 种模型对热点的预测不足或过度。相比之下，LMM、GWR 和 GWRK 模型残差产生更少的热点，局域 Moran's I 多为负值，相邻点之间表现为相反特征。

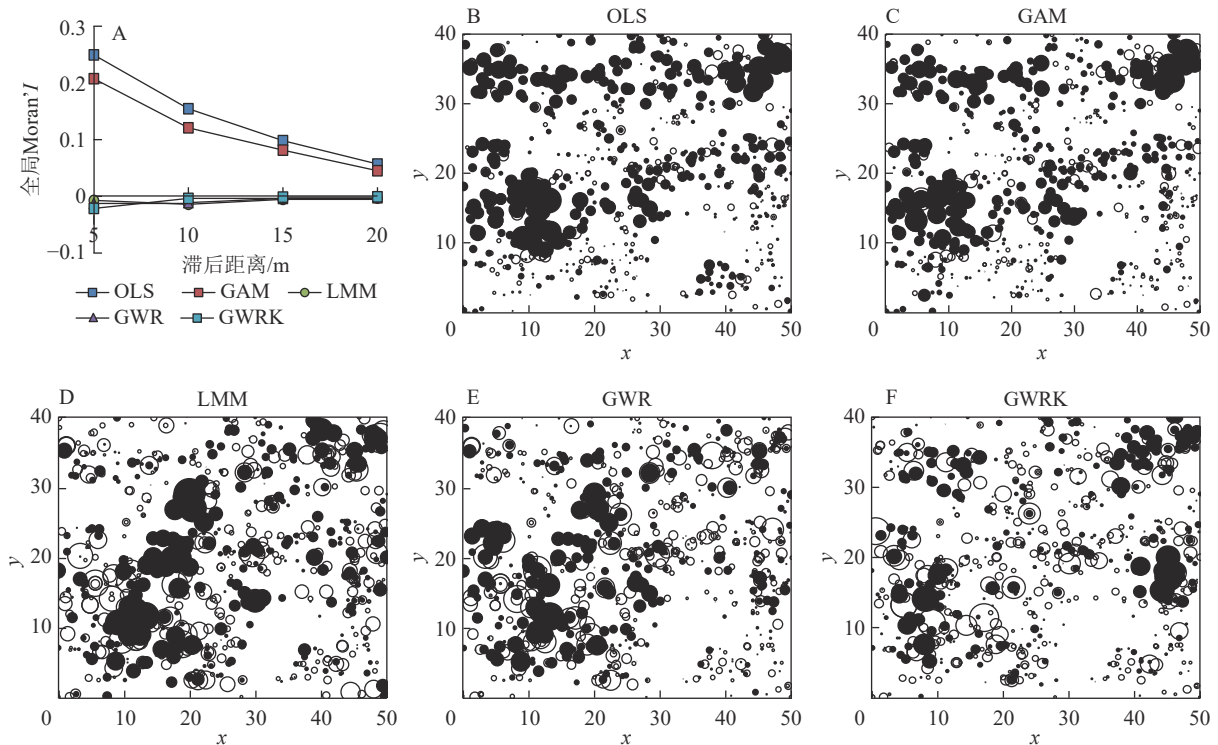


图 1 5 种模型残差空间相关性及其局域 Moran's I 分布
Figure 1 Spatial correlation and local distribution of Moran's I values of the five models' residuals

2.2.2 模型残差空间异质性 图 2 为 5~20 m 滞后距离内各模型残差的组内方差。在该范围内，5 种模型残差组内方差在 5 m 处最小，并随着滞后距离的增加而增加。GWR 和 GWRK 模型残差组内方差在各滞后距离处均显著小于 OLS、GAM 和 LMM 模型残差组内方差。5 种模型残差在各滞后距离处的组内方差从小到大排序为 GWRK、GWR、LMM、GAM、OLS。

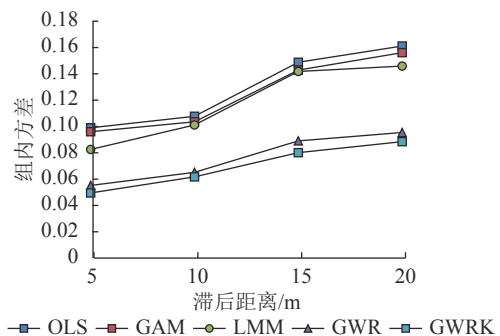


图 2 5 种模型残差组内方差
Figure 2 Intra-block variance of residuals of five models

3 讨论

3.1 林木树高-胸径模型拟合

在全林木基础模型中，模型拟合精度由高到低排序为 GWRK、GWR、LMM、GAM、OLS。OLS 模型是对数据进行全局拟合，不考虑样本点之间空间自相关性的全局模型，即所得到的模型参数估计与数据的地理位置无关，估计结果是该点的最优无偏估计的同时也是研究区内所有样本点的最优无偏估计^[28]。GAM 模型作为数据驱动模型，使用链接函数建立响应变量和解释变量的平滑函数之间的关系，为响应变量提供比 OLS 模型更好的预测^[31]。LMM 模型拟合效果同样优于 OLS 模型，本研究通过增加样方水平的随机效应，将参数 a 和 b 同时作为随机效应参数，发现将样方作为随机效应可以消除样方差异

对模型的影响^[32], 在拟合过程中展现出更好的效果。但在实际研究中, 不同地理位置之间具有空间非平稳性, OLS、GAM 和 LMM 模型属于全局模型, 没有考虑样本点位置不同会造成不同参数预测值的问题, 导致模型拟合效果不佳。GWR 和 GWRK 模型作为一种加入空间权重函数的局域空间回归模型, 考虑了不同地理位置之间的差异性, 反映了局域参数的变化, 比 OLS 模型拟合精度更高。GWRK 模型对 GWR 模型残差进行普通克里格插值后, 消除了残差空间相关对模型拟合的影响^[33], 模型拟合精度比 GWR 显著提高。将树种作为哑变量引入模型后, 模型拟合精度比全林木基础模型均有提升, 原因在于哑变量模型通常能更加精确地反映树木生长受到本身遗传性质的影响, 即将不同树种存在的差异考虑进模型。娄明华等^[8]使用树种作为哑变量构建混交林林木树高-胸径模型时, 同样发现哑变量模型拟合精度高于基础模型。有研究发现: 不同树种所拟合的树高-胸径模型曲线不同^[20], 本研究构建的马尾松、木荷以及其他树种树高-胸径模型具有不同模型参数预测值, 同样也说明了遗传特征是混交林树木生长的决定因素之一。整体而言, 局域模型拟合效果优于全局模型, 哑变量模型拟合精度高于全林木基础模型。

3.2 模型残差空间自相关性和异质性

考虑模型残差空间自相关时, LMM、GWR 和 GWRK 模型比 OLS 和 GAM 模型得到显著改进。5 种模型残差全局 Moran's I 均表现在滞后距离为 20 m 时趋近于 0, 这表明树高之间的空间自相关性随着距离的增加逐渐减小, 且当距离超过 20 m 时, 空间相关性微弱且不显著。

OLS 模型前提假设认为林木之间是相互独立的, 然而大量研究发现林木间的关系并不是相互独立的^[34], 因此忽略空间自相关会使 OLS 违背独立性假设导致对模型参数的标准误差的偏见估计^[9], 导致犯第一类错误(原假设为真, 假设检验拒绝了原假设)的可能性变大^[8]。与 OLS 模型相比, GAM 模型残差全局 Moran's I 降低, 但空间自相关性仍表现显著($Z > 1.96$, $\alpha = 0.05$), 表明 GAM 并没有显著降低残差空间自相关, 因为 GAM 并没有将空间自相关纳入建模过程, 尽管 GAM 模型由于其稳健性和灵活性而改进了模型拟合并产生了更好的预测, 但是该模型本质上还是非空间的^[31]。同时, GAM 和 OLS 模型残差局域 Moran's I 存在较多聚集的热点, 残差之间存在显著空间正相关性, 即相邻点之间的表现为相同特征, 造成对热点有较多过度预测或预测不足的值。然而, 在 $\alpha = 0.05$ 条件下, LMM、GWR 和 GWRK 模型残差全局 Moran's I 随着滞后距离的增加越来越接近于 0, 且 $-1.96 \leq Z \leq 1.96$, 残差空间自相关性不显著。LMM、GWR 和 GWRK 模型残差局域 Moran's I 在研究区内整体表现为较多冷点, 研究区相邻点残差呈现相反的特征。LMM 模型中协方差结构可以反映个体间的差异, 类似 GWR 模型中的权函数, 因此 LMM 模型在提高模型拟合效果的同时还能消除模型残差空间自相关性的影响^[35]。GWR 和 GWRK 模型对空间数据具有较强的分析能力, 在建模过程中解释变量的回归系数随着位置的变化而变化, 因此能够有效地探测空间数据的空间非平稳性^[36]。张凌宇等^[37]在研究大兴安岭中部天然次生林更新分布时, 比较全局模型和局域模型残差空间自相关时研究发现: 局域模型残差自相关明显小于全局模型, 这与本研究结果相同。陈科屹等^[38]利用 GWR 模型研究进界木空间分布, 同样得到类似的结论。

5 种回归模型在 5~20 m 滞后距离范围内, 模型残差组内方差随着距离的增加逐渐增加。与 OLS 模型相比, LMM 和 GAM 模型的组内方差在各滞后距离内变化不大, 而 GWR 和 GWRK 模型却存在很大程度的降低, 说明全局模型不能降低模型残差空间异质性, 而局域模型不仅能消除模型残差空间自相关, 还在降低空间异质性方面表现出明显的优势。ZHANG 等^[31]对加拿大 3 种不同林分建立 OLS、LMM 和 GWR 树高-胸径模型, 研究模型残差空间异质性时, 得出结果与本研究结论相同。综上, OLS 和 GAM 模型在模型拟合精度、降低空间自相关和空间异质性方面表现不佳, GWR 和 GWRK 模型不仅能提高模型拟合精度, 还能降低模型残差空间自相关和异质性, 建议用于林木树高-胸径模型的建立。

4 结论

本研究以开阳县马尾松-木荷混交林为研究对象, 以 OLS、GAM、LMM、GWR 和 GWRK 模型为全林木树高-胸径基础模型, 将树种作为哑变量引入基础模型, 并比较全林木基础模型和哑变量模型的拟合效果, 同时分析哑变量模型降低空间自相关和异质性程度, 发现哑变量模型拟合效果整体优于基础模

型, 局域模型较全局模型不仅能显著提高模型拟合效果, 还能降低模型残差空间自相关性和异质性。今后研究马尾松-木荷混交林树高-胸径模型时, 建议应用局域模型得到精度更高的拟合结果, 若考虑区域尺度问题, 可以将混交比、立地质量等因子作为哑变量引入模型, 进而构建更大尺度范围的混交林树高-胸径模型。

5 参考文献

- [1] LEI Xiangdong, PENG Changhui, WANG Haiyan, *et al.* Individual height-diameter models for young black spruce (*Picea mariana*) and jack pine (*Pinus banksiana*) plantations in New Brunswick, Canada [J]. *The Forestry Chronicle*, 2009, **85**(1): 43 – 56.
- [2] PATRICIO M S, DIAS C R G, NUNES L. Mixed-effects generalized height-diameter model: a tool for forestry management of young sweet chestnut stands [J/OL]. *Forest Ecology and Management*, 2022, **514**: 120209[2023-06-05]. doi: 10.1016/j.foreco.2022.120209.
- [3] CHAI Zongzheng, TAN Wei, LI Yuanyuan, *et al.* Generalized nonlinear height-diameter models for a *Cryptomeria fortunei* plantation in the Pingba Region of Guizhou Province, China [J]. *Web Ecology*, 2018, **18**(1): 29 – 35.
- [4] EERIKAINEN K. Predicting the height-diameter pattern of planted *Pinus kesiya* stands in Zambia and Zimbabwe [J]. *Forest Ecology and Management*, 2003, **175**(1/3): 355 – 366.
- [5] 贾炜玮, 孙赫明, 李凤日. 包含哑变量的黑龙江省落叶松人工林碳储量预测模型系统[J]. *应用生态学报*, 2019, **30**(3): 814 – 822.
JIA Weiwei, SUN Heming, LI Fengri. Prediction model system with dummy variables for carbon storage of larch plantation in Heilongjiang Province, China [J]. *Chinese Journal of Applied Ecology*, 2019, **30**(3): 814 – 822.
- [6] 粟军. 利用一元线性回归方程计算盗伐林木材积试验[J]. *林业调查规划*, 2017, **42**(5): 14 – 19.
SU Jun. Timber volume calculation of illegal logging with unary linear regression equation [J]. *Forestry Investigation and Planning*, 2017, **42**(5): 14 – 19.
- [7] 何潇, 周超凡, 雷相东, 等. 长白落叶松人工林林分碳储量生长模型系研究[J]. *北京林业大学学报*, 2021, **43**(11): 1 – 10.
HE Xiao, ZHOU Chaofan, LEI Xiangdong, *et al.* Study on carbon storage growth model of *Larch larix* plantation [J]. *Journal of Beijing Forestry University*, 2021, **43**(11): 1 – 10.
- [8] 娄明华, 张会儒, 雷相东, 等. 基于空间自相关的天然蒙古栎阔叶混交林林木胸径-树高模型[J]. *林业科学*, 2017, **53**(6): 67 – 76.
LOU Minghua, ZHANG Huiru, LEI Xiangdong, *et al.* Individual diameter-height models for mixed *Quercus mongolica* broadleaved natural stands based on spatial autocorrelation [J]. *Scientia Forestry Sinica*, 2017, **53**(6): 67 – 76.
- [9] ZHANG Lianjun, MA Zhihai, GUO Luo. Spatially assessing model errors of four regression techniques for three types of forest stands [J]. *Forestry*, 2008, **81**(2): 209 – 225.
- [10] LU Junfeng, ZHANG Lianjuan. Modeling and prediction of tree height-diameter relationships using spatial autoregressive models [J]. *Forest Science*, 2011, **57**(3): 252 – 264.
- [11] GUIGAN A, EDWARDS T C, HASTIE T. Generalized linear and generalized additive models in studies of species distributions: setting the scene [J]. *Ecological Modelling*, 2002, **157**(2/3): 89 – 100.
- [12] 姜立春, 张锐, 李凤日. 基于线性混合模型的落叶松枝条长度和角度模型[J]. *林业科学*, 2012, **48**(5): 53 – 60.
JIANG Lichun, ZHANG Rui, LI Fengri. Modeling branch length and branch angle with linear mixed effects for Dahurian larch [J]. *Scientia Silvae Sinicae*, 2012, **48**(5): 53 – 60.
- [13] 许崇华, 崔璐, 黄兴召, 等. 基于线性混合效应模型的杉木树高-胸径模型[J]. *西北农林科技大学学报(自然科学版)*, 2017, **45**(6): 53 – 60.
XU Chonghua, CUI Jun, HUANG Xingzhao, *et al.* Height-DBH model of Chinese fir based on linear mixed effects model [J]. *Journal of Northwest A&F University (Natural Science Edition)*, 2017, **45**(6): 53 – 60.
- [14] 李泽坤, 任丽燕, 马仁锋, 等. 基于时空地理加权回归模型的浙江省碳排放时空格局及驱动因素分析[J]. *宁波大学学报(理工版)*, 2021, **34**(6): 105 – 113.
LI Zekun, REN Liyan, MA Renfeng, *et al.* Analysis of spatial-temporal pattern and driving factors of carbon emissions in Zhejiang Province based on spatial-temporal geographic weighted regression model [J]. *Journal of Ningbo University (Science and Technology Edition)*, 2021, **34**(6): 105 – 113.
- [15] 陈科屹, 张会儒, 张博, 等. 基于地理加权回归拓展模型的自然次生林碳储量空间分布[J]. *应用生态学报*, 2021, **32**(4):

- 1175 – 1183.
- CHEN Keyi, ZHANG Huiru, ZHANG Bo, *et al.* Spatial distribution of carbon storage in natural secondary forests based on geographically weighted regression expansion model [J]. *Chinese Journal of Applied Ecology*, 2021, **32**(4): 1175 – 1183.
- [16] 吴子豪, 刘耀林, 冯向阳, 等. 基于多尺度地理加权回归的土壤镉污染局部影响因子分析[J]. *地球信息科学学报*, 2023, **25**(3): 573 – 587.
- WU Zihao, LIU Yaolin, FENG Xiangyang, *et al.* Analysis of local influencing factors of cadmium pollution in soil by using multi-scale geographically weighted regression [J]. *Journal of Geo-Information Science*, 2023, **25**(3): 573 – 587.
- [17] 李鹏, 杨章旗, 颜培栋, 等. 不同混交比例马尾松和红锥人工混交林水土流失特征和土壤理化性质[J]. *中南林业科技大学学报*, 2022, **42**(4): 104 – 116.
- LI Peng, YANG Zhangqi, YAN Peidong, *et al.* Quality evaluation of mixed plantations of *Pinus massoniana* and *Castanopsis hystrix* based on the soil erosion characteristics and soil physical and chemical properties [J]. *Journal of Central South University of Forestry and Technology*, 2022, **42**(4): 104 – 116.
- [18] 李志辉, 李柏海, 祁承经, 等. 我国南方珍贵用材树种资源的重要性及其发展策略[J]. *中南林业科技大学学报*, 2012, **32**(11): 1 – 8.
- LI Zhihui, LI Bohai, QI Chengjing, *et al.* Studies on importance of valuable wood species resources and its development strategy [J]. *Journal of Central South University of Forestry & Technology*, 2012, **32**(11): 1 – 8.
- [19] 陈黑虎. 马尾松、木荷混交林生长效果分析[J]. *安徽农学通报*, 2014, **20**(17): 105 – 107, 121.
- CHEN Heihu. Analysis on the growth effect of masson pine and *Schima superba* mixed forest [J]. *Anhui Agricultural Science Bulletin*, 2014, **20**(17): 105 – 107, 121.
- [20] TIAN Dongyuan, JIANG Lichun, SHAHZAD M K, *et al.* Climate-sensitive tree height-diameter models for mixed forests in Northeastern China [J/OL]. *Agricultural and Forest Meteorology*, 2022, **326**: 109182[2023-06-10]. doi: 10.1016/j.agrformet.2022.109182.
- [21] 孙拥康, 汤景明, 王怡. 基于分位数回归的马尾松青冈栎混交林树高-胸径模型[J]. *中南林业科技大学学报*, 2021, **41**(12): 18 – 25.
- SUN Yongkang, TANG Jingming, WANG Yi. Height-diameter model of *Pinus massoniana* and *Cyclobalanopsis glauca* mixed forest based on quantile regression [J]. *Journal of Central South University of Forestry and Technology*, 2021, **41**(12): 18 – 25.
- [22] 李方兴, 张意苗, 易伟东, 等. 马尾松、木荷纯林及混交林的生长差异分析[J]. *南方林业科学*, 2016, **44**(5): 17 – 20.
- LI Fangxing, ZHANG Yimiao, YI Weidong, *et al.* Analysis on growth differences between mixed plantations of *Pinus massoniana* and *Schima superba* and their pure plantation [J]. *Scientia Forestry Sinica of Southern China*, 2016, **44**(5): 17 – 20.
- [23] 黎芳, 潘萍, 宁金魁, 等. 马尾松-木荷不同比例混交林林下植被物种组成及其多样性分析[J]. *西北林学院学报*, 2016, **31**(6): 34 – 40.
- LI Fang, PAN Ping, NING Jinkui, *et al.* Undergrowth species composition and diversity of *Pinus massoniana* and *Schima superba* mixed stands with different proportions [J]. *Journal of Northwest Forestry College*, 2016, **31**(6): 34 – 40.
- [24] 殷沙, 赵芳, 欧阳勋志. 马尾松木荷不同比例混交林枯落物和土壤持水性能比较分析[J]. *江西农业大学学报*, 2015, **37**(3): 454 – 460.
- YIN Sha, ZHAO Fang, OUYANG Xunzhi. A comparison on water-holding capacity of forest litter and soil of mixed forests of *Pinus massoniana* and *Schima superba* in different proportions [J]. *Journal of Jiangxi Agricultural University*, 2015, **37**(3): 454 – 460.
- [25] 刘晓彤, 黄金金, 张逸如, 等. 基于广义可加模型的广东省森林土壤有机质影响因子[J]. *生态学杂志*, 2022, **41**(11): 2278 – 2288.
- LIU Xiaotong, HUANG Jinjin, ZHANG Yiru, *et al.* Analysis of influencing factors on forest soil organic matter in Guangdong Province based on GAM model [J]. *Chinese Journal of Ecology*, 2022, **41**(11): 2278 – 2288.
- [26] 杜一尘, 李明泽, 范文义, 等. 基于地理加权回归模型与林火遥感数据估算森林年龄[J]. *林业科学*, 2019, **55**(6): 184 – 194.
- DU Yichen, LI Mingze, FAN Wenyi, *et al.* Estimation of forest age based on geographically weighted regression model and forest fire remote sensing data [J]. *Scientia Silvae Sinicae*, 2019, **55**(6): 184 – 194.
- [27] 李运龙, 熊立华, 闫磊. 基于地理加权回归克里金的降水数据融合及其在水文预报中的应用[J]. *长江流域资源与环境*, 2017, **26**(9): 1359 – 1368.

- LI Yunlong, XIONG Lihua, YAN Lei. A geographically weighted regression kriging approach for Trmm-rain gauge data merging and its application in hydrological forecasting [J]. *Resources and Environment in the Yangtze Basin*, 2017, **26**(9): 1359 – 1368.
- [28] 刘畅. 黑龙江省森林碳储量空间分布研究[D]. 哈尔滨: 东北林业大学, 2014.
LIU Chang. *Spatial Distribution of Forest Carbon Storage in Heilongjiang Province* [D]. Harbin: Northeast Forestry University, 2014.
- [29] 李海滨, 王政权, 王庆成. 空间异质性定量研究理论与方法[J]. *应用生态学报*, 1998, **9**(6): 651 – 657.
LI Habin, WANG Zhengquan, WANG Qingcheng. Theory and method ology of spatial heterogeneity quantification [J]. *Chinese Journal of Applied Ecology*, 1998, **9**(6): 651 – 657.
- [30] 戚玉娇. 大兴安岭森林地上碳储量遥感估算与分析[D]. 哈尔滨: 东北林业大学, 2014.
QI Yujiao. *Estimation and Analysis of Forest above-ground Carbon Storage by Remote Sensing in Greater Hingan Mountains*[D]. Harbin: Northeast Forestry University, 2014.
- [31] ZHANG Lianjun, GOVE J H. Spatial assessment of model errors from four regression techniques [J]. *Forest Science*, 2005, **51**(4): 334 – 346.
- [32] 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 Hao, LUO Yang. Construction of nonlinear mixed effect height-diameter model for *Pinus massoniana* [J]. *Journal of Forestry and Environment*, 2021, **41**(4): 439 – 448.
- [33] 孙钰森, 王维芳, 李国春. 基于地理加权回归克里格模型的帽儿山地区森林碳储量空间分布[J]. *应用生态学报*, 2019, **30**(5): 1642 – 1650.
SUN Yusen, WANG Weifang, LI Guochun. Spatial distribution of forest carbon storage in Maershan region, Northeast China based on geographically weighted regression kriging model [J]. *Chinese Journal of Applied Ecology*, 2019, **30**(5): 1642 – 1650.
- [34] MATEU J, USO J L, MONTES F. The spatial pattern of a forest ecosystem [J]. *Ecological Modelling*, 1998, **108**(1/3): 163 – 174.
- [35] QI Y J, ZHANG Y C, WANG K, *et al.* Application of spatial regression models for forest biomass estimation in Guizhou Province, Southwest China [J]. *Applied Ecology and Environmental Research*, 2020, **18**(5): 7215 – 7232.
- [36] FOTHERINGHAM A S, BRUNSDON C, CHARLTON M. *Geographically Weighted Regression: the Analysis of Spatially Varying Relationships* [M]. New York: John Wiley & Sons, 2003.
- [37] 张凌宇, 刘兆刚. 基于地理加权回归模型的大兴安岭中部天然次生林更新分布[J]. *林业科学*, 2019, **55**(11): 105 – 116.
ZHANG Lingyu, LIU Zhaogang. Regeneration and distribution of natural secondary forests in the central part of Daxing'an Mountains based on geographically weighted regression model [J]. *Scientia Forestry Sinica*, 2019, **55**(11): 105 – 116.
- [38] 陈科屹, 张会儒, 张博, 等. 基于地理加权回归的天然次生林进界木空间分布模拟[J]. *北京林业大学学报*, 2021, **43**(2): 1 – 9.
CHEN Keyi, ZHANG Huiru, ZHANG Bo, *et al.* Spatial distribution simulation of recruitment trees of natural secondary forest based on geographically weighted regression [J]. *Journal of Beijing Forestry University*, 2021, **43**(2): 1 – 9.