

DOI: 10.13332/j.1000-1522.20150491

## 天然云冷杉针阔混交林单木胸径树高空间自回归模型研究

娄明华<sup>1,2</sup> 张会儒<sup>1</sup> 雷相东<sup>1</sup> 卢军<sup>1</sup>

(1 中国林业科学研究院资源信息研究所 2 宁波市农业科学研究院)

**摘要:** 林木间普遍存在着空间自相关,这直接关联着林木间的竞争与相互作用。单木胸径树高模型是森林生长、收获与预测的基础,忽略林木间的空间自相关将会导致胸径树高模型的普通最小二乘(OLS)回归违背残差独立分布假设,导致犯第一类错误的可能性变大,以及模型参数标准差的有偏估计和回归模型估计的有效性降低。因此,本文选择我国东北地区主要森林类型即天然云冷杉针阔混交林为研究对象,考虑林木间的空间自相关,选用合适的线性化单木胸径树高 OLS 模型为基准模型,利用3个同步自回归(SAR)模型即空间滞后模型(SLM)、空间误差模型(SEM)和空间 Durbin 模型(SDM),构建该混交林的单木胸径树高模型。与此同时,每个 SAR 模型分别采用5个不同的空间加权矩阵即 Delaunay 三角网(DT)矩阵、逆距离一次幂(ID1)、逆距离二次幂(ID2)、逆距离五次幂(ID5)和高斯变异函数(GV)矩阵,利用极大似然(maximum likelihood)估计3个 SAR 模型的参数。对 OLS 和3个 SAR 模型的回归参数进行 *t* 检验,对3个 SAR 模型的自回归参数进行似然比检验。选择 Moran's *I* (MI) 指数比较分析4个模型的残差空间自相关,选择决定系数( $R^2$ )、均方根误差(RMSE)和 Akaike 信息准则(AIC)3个拟合指标比较分析这4个模型的拟合效果,选择均方误差(MS)检验模型预测效果。结果表明:未考虑空间自相关的 OLS 模型残差存在正空间自相关;3个 SAR 模型拟合效果均优于 OLS,SDM 和 SEM 的拟合效果最好,SLM 最差;无论使用哪个空间加权矩阵,SLM 均不能消除模型残差空间自相关,但可降低空间自相关,在一定程度上提高了模型的拟合效果;5个空间矩阵应用于 SDM 和 SEM 时,均可以消除模型残差空间自相关,但空间加权矩阵 GV 只适用于 SEM;ID2 是5个空间加权矩阵中最好的空间加权矩阵,将 ID2 应用于4个模型进行预测时,SDM 和 SEM 的预测效果明显优于 SLM,但3个 SAR 模型的预测效果均优于 OLS。利用3个 SAR 模型提高了单木胸径树高模型拟合和预测的精度,为合理经营天然云冷杉针阔混交林提供了理论基础。

**关键词:** 空间自相关; 胸径; 树高; 自回归模型

中图分类号: S758.5<sup>+</sup>2 文献标志码: A 文章编号: 1000-1522(2016)08-0001-09

LOU Ming-hua<sup>1,2</sup>; ZHANG Hui-ru<sup>1</sup>; LEI Xiang-dong<sup>1</sup>; LU Jun<sup>1</sup>. **An individual height-diameter model constructed using spatial autoregressive models within natural spruce-fir and broadleaf mixed stands.** *Journal of Beijing Forestry University* (2016)38(8) 1-9 [Ch., 50 ref.]

1 Research Institute of Forest Resources Information Techniques, Chinese Academy of Forestry, Beijing, 100091, P. R. China;

2 Ningbo Academy of Agriculture Sciences, Ningbo, Zhejiang, 315040, P. R. China.

Spatial autocorrelation is a common phenomenon in forestry. It directly connects competition and interaction among individuals. Individual height-diameter models are fundamentally important for forest growth, yield modeling and forecasting. Violation of residual independent distribution assumption in ordinary least squares (OLS) will inflate type 1 errors, lead to biased estimates of the standard errors of model parameters, and decrease the efficiency of estimation in a regression model, if the spatial autocorrelation among the individuals is ignored. Therefore, three simultaneous autoregressive (SAR) models, including spatial lag model (SLM), spatial error model (SEM) and spatial Durbin model (SDM) within five spatial weight matrices, including Delaunay triangulation (DT), inverse distance raised to one power (ID1), inverse distance raised to two powers (ID2), inverse distance raised to five powers (ID5) and Gaussian variogram (GV), were applied to construct individual height-diameter

收稿日期: 2015-12-11 修回日期: 2016-03-29

基金项目: “十二五”国家科技支撑计划课题(2012BAD22B02)。

第一作者: 娄明华, 博士生。主要研究方向: 森林可持续经营。Email: mhlou1987@163.com 地址: 100091 北京市海淀区东小府1号中国林业科学研究院资源信息研究所。

责任作者: 张会儒, 研究员, 博士生导师。主要研究方向: 森林可持续经营。Email: huiru@ifrit.ac.cn 地址: 同上。

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

models of natural spruce-fir and broadleaf mixed stands which are the main forest type in northeast China, with linearization individual height-diameter OLS model as a benchmark model. Model parameters of three SAR models were estimated by maximum likelihood. Model coefficients of OLS and three SAR models were tested by *t*-test, the autoregressive parameters of three SAR models were all tested by likelihood ratio test. Moran's *I* (MI) was selected to compare autocorrelation of four model residuals. Three statistical indices, i. e. coefficient of determination ( $R^2$ ), root mean square error (RMSE) and Akaike information criterion (AIC), were regarded as the appropriate criteria to identify the model fitting among OLS, SLM, SDM and SEM. Mean square error (MS) was selected to identify the predictive validity among four models. Results show that residuals of OLS were positive spatial dependence for ignoring the spatial autocorrelation among individuals. The model fittings of three SAR models were better than that of OLS. Among the three SAR models, model fitting of SLM was worse than those of SDM and SEM. SLM do not remove but reduce the spatial autocorrelation of model residuals, and slightly improve the model fitting, no matter which spatial weight matrices are used in SLM. All of the spatial weight matrices used in SDM and SEM could remove the spatial autocorrelation of residuals; however, GV was only applicable to SEM. Among all spatial weight matrices, ID2 was the best spatial weight matrix. Using ID2 into four modes, the predictive validity of SDM and SEM was superior to that of SLM, while the predictive validity of three SAR was better than that of OLS. Using three SAR models, fitting and prediction of individual diameter at breast height and height models were improved, and it may provide a theoretical basis for reasonable management of natural spruce-fir and broadleaf mixed stands.

**Key words** spatial autocorrelation; diameter at breast height; height; autoregressive model

单木胸径树高模型是森林生长、收获与预测的基础<sup>[1-3]</sup>,广泛用于林木材积估计<sup>[4-5]</sup>、林分中林木地位确定<sup>[6]</sup>、优势高确定及立地生产力指数计算<sup>[5,7-8]</sup>、林分生长及其演替规律的描述<sup>[9-10]</sup>等。传统单木胸径树高模型均假定林木间是相互独立的,通常采用普通最小二乘(ordinary least squares, OLS)进行估计。然而,研究者们逐渐意识到林木间的关系不是相互独立的<sup>[11-13]</sup>,这是因为林木间存在空间自相关,林木间的空间自相关直接关联着林木间的竞争和相互作用<sup>[14-16]</sup>。根据著名的 Tobler 地理学第一定律(Tobler's first law of geography),可知相邻林木间的空间自相关程度强于相离林木间的空间自相关程度<sup>[17-20]</sup>。从林学角度上看,忽略林木间的空间自相关关系即林木间的竞争和相互作用关系,将使构建的单木胸径树高模型与实际的单木胸径树高生长规律存在着较大偏差,从而导致人们对森林生长、收获与预测的有偏估计<sup>[16]</sup>;从统计学角度上看,忽略林木间的空间自相关也会使 OLS 回归违背残差独立分布假设,导致犯第一类错误可能性变大<sup>[18,21-24]</sup>、模型参数标准差的有偏估计及回归模型估计有效性的降低<sup>[15,25-28]</sup>。因此,林木间的空间自相关不能忽视。

为了考虑林木间的空间自相关,国外已有学者将空间自相关模型加入到单木胸径树高模型中,而在国内尚无报道。目前,空间自相关模型有地理加

权回归(geographically weighted regression)模型<sup>[29-30]</sup>和同步自回归(simultaneous autoregressive, SAR)模型<sup>[16,31]</sup>等。其中 SAR 模型是基于空间自相关过程发生在模型应变量、模型解释变量及模型误差的可能性而构建的,直接将林木变量和林木间的空间自相关结构镶嵌于模型中,因此是最为常用的空间自相关模型<sup>[16,24,31]</sup>。具体来说,SAR 模型包括空间滞后模型(spatial lag model, SLM)、空间误差模型(spatial error model, SEM)和空间 Durbin 模型(spatial Durbin model, SDM),其中 SLM 假设空间自回归过程只发生在应变量,SEM 假设该过程只发生在模型误差项,SDM 假设该过程同时发生在应变量和解释变量<sup>[32-33]</sup>。

天然云冷杉林是我国寒温带分布最广的森林类型,也是我国东北地区主要森林类型之一,具有重要的生态、经济和社会效益<sup>[34-35]</sup>。其中,云冷杉针阔混交林是东北地区重要的用材林和风景林,在水土保持和水源涵养方面发挥着重要作用<sup>[36-37]</sup>。因此,本文以东北地区的天然云冷杉针阔混交林为研究对象,考虑林木间的空间自相关,利用 SAR 模型研究该混交林的单木胸径树高模型,为合理经营管理天然云冷杉针阔混交林提供理论基础。

## 1 材料与方法

### 1.1 研究区概况与数据

研究区位于吉林省东部的汪清林业局金沟岭林

场,其地理坐标为 130°05'53" ~ 130°19'55"E, 43°17'31" ~ 43°25'26"N。属长白山系老爷岭山脉雪岭支脉,地貌为低山丘陵,海拔 300 ~ 1 200 m,坡度一般在 5° ~ 25°。研究区属季风型气候,年均气温 3.9 °C 左右,年均降水量 600 ~ 700 mm,且集中在夏季,占全年总降水量的 80%。土壤以暗棕壤为主,平均厚度在 40 cm 左右。该区植被属长白山植物区系,群落结构复杂,植物种类较多。主要树种有云杉 (*Picea asperata*)、冷杉 (*Abies fabri*)、长白落叶松 (*Larix olgensis*)、白桦 (*Betula platyphylla*)、红松 (*Pinus koraiensis*)、山杨 (*Populus davidiana*)、榆树 (*Ulmus pumila*)、杨树 (*Populus spp.*)、椴树 (*Tilia tuan*)、黑桦 (*Betula dahurica*)、色木槭 (*Acer pictum*)、水曲柳 (*Fraxinus mandshurica*)、黄菠萝

(*Phelladendron amurense*) 等。

数据来源于 2013 年 7 月至 9 月中在金沟岭林场设置的 12 块 100 m × 100 m 天然云冷杉针阔混交林固定样地。利用罗盘仪将每块样地划分为 100 个 10 m × 10 m 的网格单元,在每个网格单元内,对胸径 5 cm 以上的树木进行每木调查,记录树种,测量胸径、树高、枝下高、东西南北冠幅、枯死原因等因子,并利用激光测距仪测定林木坐标 (X, Y)。本文选择坡度、坡位、坡向及海拔基本一致的 2 块样地,一块样地数据作为建模数据 (图 1A),另一块样地数据作为预测数据 (图 1B)。2 块样地的坡度均为 5°,坡位均为下坡,坡向均为东北坡,样地海拔 732 ~ 749 m。样地林分特征见表 1。

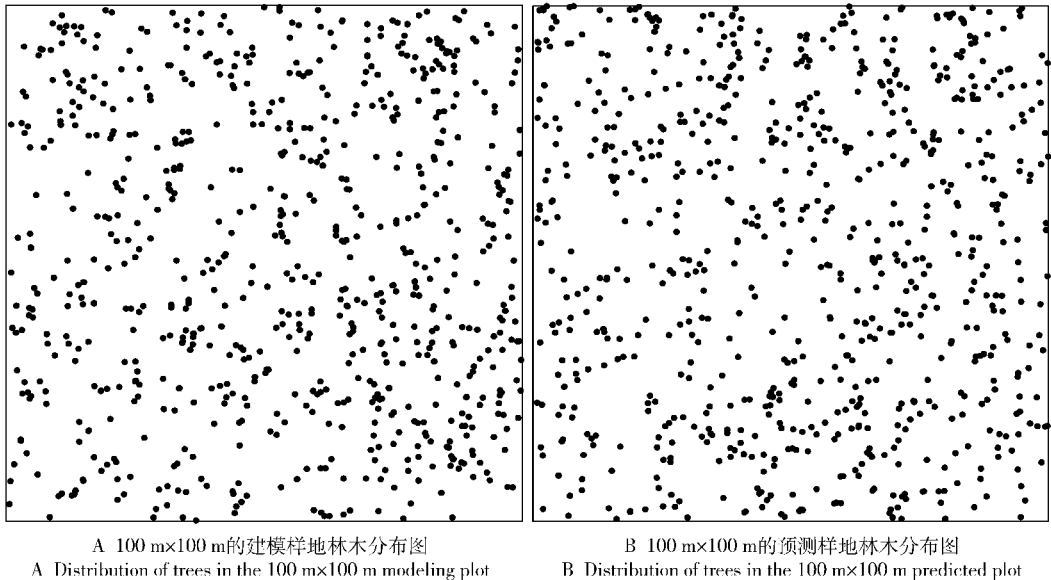


图 1 林木分布图

Fig. 1 Individual distribution diagram

表 1 2 块样地林分特征

Tab. 1 Stand characteristics of two plots

数据 Data	株数 Number of trees	平均胸径 ± 标准差 (最小值,最大值) Mean diameter at breast height ± standard deviation (min-value, max-value)/cm	断面积 Basal area/m <sup>2</sup>	平均高 ± 标准差 (最小值,最大值) Mean height ± standard deviation (min-value, max-value)/m	树种组成(100 分法) Species composition
建模数据 Modeling data	796	18.08 ± 8.87 (5.0, 55.5)	0.23	13.21 ± 5.74 (3.4, 34.2)	23 冷杉 <i>Abies fabri</i> 14 红松 <i>Pinus koraiensis</i> 14 白桦 <i>Betula platyphylla</i> 11 落叶松 <i>Larix olgensis</i> 10 云杉 <i>Picea asperata</i> 7 椴树 <i>Tilia tuan</i> 6 枫桦 <i>Betula costata</i> + 色木 <i>Acer pictum</i> - 杨树 <i>Populus spp.</i> - 榆树 <i>Ulmus pumila</i> - 水曲柳 <i>Fraxinus mandshurica</i> - 黄菠萝 <i>Phelladendron amurense</i> - 红豆杉 <i>Taxus chinensis</i>
预测数据 Validation data	765	19.06 ± 8.63 (5.0, 72.0)	0.26	13.81 ± 5.82 (3.6, 34.4)	19 冷杉 <i>A. fabri</i> 15 红松 <i>P. koraiensis</i> 13 椴树 <i>T. tuan</i> 11 枫桦 <i>B. costata</i> 10 落叶松 <i>L. olgensis</i> 10 云杉 <i>P. asperata</i> 6 杨树 <i>P. spp.</i> 5 色木 <i>A. pictum</i> + 白桦 <i>B. platyphylla</i> - 榆树 <i>U. pumila</i> - 水曲柳 <i>F. mandshurica</i> - 红豆杉 <i>T. chinensis</i>

## 1.2 普通最小二乘回归模型

由于当前同步自回归(SAR)模型均是线性的,而不是非线性的<sup>[16]</sup>,因此本文只考虑线性化的单木胸径树高普通最小二乘回归(OLS)模型。根据建模数据的胸径与树高散点图(图2),选用8个备选的线性化模型(见表2)。

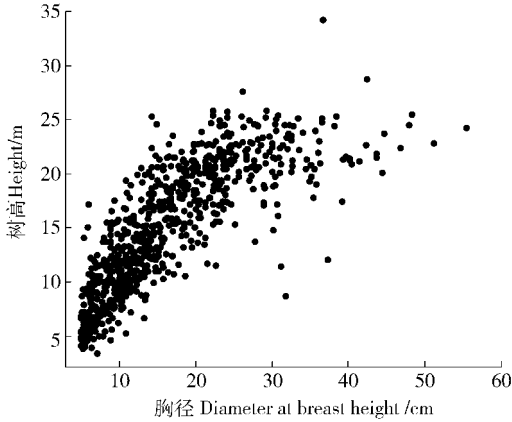


图2 胸径树高散点图

Fig.2 Scatterplot of diameter at breast height against height

表2 备选模型

Tab.2 Alternative models

模型序号 Model No.	模型表达式 Model expression	参考文献 References
A	$\ln H = \beta_0 + \beta_1 \ln D + \varepsilon$	[19]
B	$\ln H = \beta_0 + \beta_1 D + \varepsilon$	[38-39]
C	$\ln H = \beta_0 + \beta_1 D^{-1} + \varepsilon$	[9]
D	$H^{-1} = \beta_0 + \beta_1 D^{-1} + \varepsilon$	[39]
E	$H = \beta_0 + \beta_1 \ln D + \varepsilon$	[9]
F	$H = \beta_0 + \beta_1 D + \beta_2 D^2 + \varepsilon$	[9]
G	$H = \beta_0 + \beta_1 D^{-1} + \beta_2 D^2 + \varepsilon$	[9,38]
H	$H^{-1} = \beta_0 + \beta_1 D^{-1} + \beta_2 D^{-2} + \varepsilon$	[9,38]

注:  $H$  表示树高;  $D$  表示胸径;  $\beta_0$ 、 $\beta_1$  和  $\beta_2$  表示回归系数;  $\varepsilon$  表示模型残差。Note:  $H$  means tree height,  $D$  diameter of breast height,  $\beta_0$ ,  $\beta_1$  and  $\beta_2$  regression coefficient, and  $\varepsilon$  model residue.

## 1.3 同步自回归模型

本文选用同步自回归(SAR)模型处理林木间的空间自相关。SAR模型包括空间滞后模型(SLM)、空间误差模型(SEM)和空间Durbin模型(SDM),其表达式<sup>[17,23,33,40-41]</sup>如下:

$$Y = \rho WY + X\beta + \varepsilon \quad (1)$$

$$Y = X\beta + \lambda W\xi + \varepsilon \quad (2)$$

$$Y = \rho WY + X\beta + WX\gamma + \varepsilon \quad (3)$$

式(1)、式(2)和式(3)分别表示SLM、SEM和SDM。其中, $Y$ 表示应变量矩阵, $X$ 表示解释变量矩阵, $W$ 表示空间加权矩阵, $\beta$ 表示回归系数, $\varepsilon$ 表示模型残差, $\rho$ 、 $\lambda$ 和 $\gamma$ 均表示空间自回归系数, $\xi$ 表示

空间相关残差。 $WY$ 、 $WY$ 和 $W\xi$ 为3个空间滞后变量。当 $\rho=0$ 时,式(1)则变为线性回归模型;当 $\lambda=0$ 时,式(2)则为线性回归模型;当 $\rho$ 和 $\gamma$ 同时为0时,式(3)则为线性回归模型。同时,采用极大似然(maximum likelihood)法估计3个SAR模型的参数,通过R统计语言<sup>[42]</sup>的spdep包的lagsarlm和errorsarlm函数<sup>[43]</sup>实现。

## 1.4 空间加权矩阵

空间加权矩阵(spatial weight matrix)是SAR模型重要的组成部分。本文选取常用的空间加权矩阵即相接邻近(contiguous neighbors)矩阵、逆距离幂(inverse distances raised to some power)矩阵和地统计(geostatistical)矩阵<sup>[31,44]</sup>。

相接邻近矩阵选用Delaunay三角网(Delaunay triangulation, DT)矩阵<sup>[45]</sup>等。DT矩阵中,林木间为相接邻近关系时,其空间加权值为1,否则为0。利用R统计语言的spdep包的trib2nb函数<sup>[43]</sup>计算空间加权值。

逆距离幂矩阵包括逆距离一次幂(inverse distances raised to one power, ID1)、逆距离二次幂(inverse distances raised to two power, ID2)和逆距离五次幂(inverse distances raised to five power, ID5)3种形式<sup>[44,46]</sup>。公式如下:

$$w_{ij} = \begin{cases} 1/d_{ij}^n & i \neq j \\ 0 & \text{其他} \end{cases} \quad (4)$$

式中: $d_{ij}$ 表示林木*i*与其相邻木*j*的距离, $w_{ij}$ 表示林木*i*与其相邻木*j*的空间加权值,当 $n=1$ 时,则为ID1;当 $n=2$ 时,则为ID2;当 $n=5$ 时,则为ID5。

Getis等<sup>[44]</sup>利用满足本征假设(intrinsic hypothesis)的变异函数构建了一个地统计矩阵,即高斯变异函数(Gaussian variogram, GV)矩阵。公式如下:

$$w_{ij} = \begin{cases} \exp(-d_{ij}^2/r^2) & d_{ij} < r \\ 0 & \text{其他} \end{cases} \quad (5)$$

式中: $r$ 表示变程。

利用R统计语言的nlme包的corGaus函数<sup>[47]</sup>计算GV矩阵的空间加权值。

本文采用DT、ID1、ID2、ID5和GV共5个空间加权矩阵,应用于3个SAR模型中。

## 1.5 空间自相关检验指标

选择常用的Moran's  $I$ (MI)检验模型残差的空间自相关程度。公式<sup>[48]</sup>如下:

$$MI = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (y_i - \bar{y})(y_j - \bar{y})}{\left( \sum_{i=1}^n (y_i - \bar{y})^2 \right) \left( \sum_{i=1}^n \sum_{j=1}^n w_{ij} \right)} \quad (6)$$

式中:MI 表示模型残差的空间自相关程度量化值, $n$  表示林木株数, $y_i$  和  $y_j$  分别表示第  $i$  株林木和第  $j$  株林木的模型残差值, $\bar{y}$  表示平均模型残差, $w_{ij}$  表示第  $i$  株林木和第  $j$  株林木的空间加权值。

其原假设为 MI 等于期望值  $MI_0 = -1/(n-1)$ , 表明不存在空间自相关。如果 MI 显著 ( $\alpha = 0.01$ ) 大于  $MI_0$ , 则模型残差为正自相关, 说明相邻林木间的相似程度高; 反之 MI 显著 ( $\alpha = 0.01$ ) 小于  $MI_0$ , 则模型残差为负自相关, 说明相邻林木间的不相似程度高。利用 R 统计语言的 spdep 包的 moran.test 函数<sup>[43]</sup> 进行检验。

### 1.6 模型拟合指标

为了方便比较模型之间的模型拟合程度, 本文选用决定系数 (coefficient of determination,  $R^2$ )、均方根误差 (root mean square error, RMSE) 和 Akaike 信息准则 (Akaike information criterion, AIC) 3 个拟合指标。

### 1.7 模型预测指标

预测误差 ( $e_i$ ) 为预测数据的观测值与预测值之差。其中,  $e_i$  的平均值 ( $e$ ) 表示模型预测准确度 (accuracy),  $e_i$  的标准差 ( $\sigma$ ) 作为模型预测精确度 (precision)。均方误差 (mean square error, MS) 是模型预测准确度和精确度的结合<sup>[31,49-50]</sup>, 公式如下:

$$MS = e^2 + \sigma^2 \quad (7)$$

本文将采用  $e$ 、 $\sigma$  和 MS 作为模型预测指标。

## 2 模型比较分析

选择最优的备选线性化模型作为本文的普通最小二乘回归模型, 然后与 SLM、SEM 和 SDM 3 个 SAR 模型进行比较, 分析其模型残差空间自相关、模型拟合、模型参数估计效果及模型预测效果。

### 2.1 普通最小二乘回归模型确定

利用  $R^2$  和 RMSE 2 个拟合指标, 比较 8 个备选模型。从表 3 可看出, 备选模型 C 可作为最合适单

木胸径树高 OLS 模型, 虽然其  $R^2$  略小于模型 F, RMSE 略大于模型 F, 但从模型形式上看, 模型 C 较模型 F 更为简单, 因为模型 F 增加了解释变量  $D^2$ 。

表 3 备选模型的拟合指标

Tab. 3 Fitting index of eight alternative models

模型序号 Model No.	$R^2$	RMSE
A	0.715	0.237
B	0.571	0.290
C	0.736	0.228
D	0.692	0.233
E	0.733	0.228
F	0.740	0.226
G	0.710	0.251
H	0.696	0.226

### 2.2 模型残差空间自相关

由表 4 可知, 5 个空间加权矩阵的 OLS 残差 MI 值均显著 ( $P < 0.01$ ) 大于期望值  $MI_0 = -1/(796-1) = -0.0013$ , 说明 OLS 残差存在正空间自相关。5 个空间加权矩阵的 SLM 残差 MI 值均显著 ( $P < 0.01$ ) 大于  $MI_0$ , 但其 MI 值小于相同空间加权矩阵的 OLS 残差 MI 值, 说明 SLM 没有消除模型残差的空间自相关, 但是降低了空间自相关。除了 GV, 其他 4 个空间加权矩阵的 SDM 残差 MI 值均与  $MI_0$  差异不显著 ( $P > 0.1$ ), 说明选择合适的空间加权矩阵, SDM 可以消除模型残差的空间自相关, 同时也表明产生模型残差的空间自相关可能是由滞后变量  $WY$  (即相邻加权树高) 和滞后变量  $WX$  (即相邻加权胸径) 共同作用产生的。5 个空间加权矩阵的 SEM 残差 MI 值均与  $MI_0$  差异不显著 ( $P > 0.1$ ), 表明 SEM 可以消除模型残差的空间自相关, 但空间自相关可能由滞后变量  $W\xi$  (即相邻加权未观测的变量) 产生的, 因为 SEM 的前提假设是空间自相关过程来源于模型误差; 同时也表明 GV 空间加权矩阵适用于 SEM, 而不适用于 SDM。

表 4 4 个模型残差的空间自相关

Tab. 4 Test for spatial autocorrelation of four model residuals

空间加权矩阵 Spatial weight matrix	OLS		SLM		SDM		SEM	
	MI	$P$	MI	$P$	MI	$P$	MI	$P$
DT	0.120	0.000 0	0.097	0.000 0	-0.006	0.820 6	-0.006	0.831 2
ID1	0.082	0.000 0	0.069	0.000 0	-0.014	0.345 2	-0.010	0.505 9
ID2	0.167	0.000 0	0.115	0.000 0	-0.014	0.616 8	-0.013	0.637 2
ID5	0.204	0.000 0	0.129	0.000 9	-0.005	0.920 8	-0.006	0.910 6
GV	0.568	0.000 0	0.385	0.004 6	0.342	0.011 9	-0.066	0.635 5

注: DT 是 Delaunay 三角网; ID1 是逆距离一次幂; ID2 是逆距离二次幂; ID5 是逆距离五次幂; GV 是高斯变异函数。下同。Notes: DT, Delaunay triangulation; ID1, inverse distances raised to one power; ID2, inverse distances raised to two power; ID5, inverse distances raised to five power; GV, Gaussian variogram. The same below.

### 2.3 模型拟合

利用  $R^2$ 、RMSE 和 AIC 3 个拟合指标,比较分析 OLS 与 3 个 SAR 模型即 SLM、SDM 和 SEM 的模型拟合效果。由表 5 和表 3 可知:3 个 SAR 模型的  $R^2$  均大于 OLS, RMSE 和 AIC 均小于 OLS,说明 3 个 SAR 模型的拟合效果均优于 OLS。3 个 SAR 模型

中,SDM 和 SEM 拟合效果明显优于 SLM,这是因为 SLM 没有消除模型残差的空间自相关,SDM 与 SEM 的拟合效果基本相同。由表 5 还可以看出,依据  $R^2$ 、RMSE 和 AIC 3 个拟合指标,5 个空间加权矩阵在 SDM 和 SEM 中的拟合效果比较结果是:ID2 最好,DT 排第二,ID1 和 ID5 并列第三,GV 最差。

表 5 3 个 SAR 模型的拟合  
Tab. 5 Three SAR models fitting

空间加权矩阵 Spatial weight matrix	SLM			SDM			SEM		
	$R^2$	RMSE	AIC	$R^2$	RMSE	AIC	$R^2$	RMSE	AIC
DT	0.737	0.227	-93.188	0.750	0.221	-121.113	0.749	0.222	-119.047
ID1	0.737	0.227	-92.765	0.748	0.222	-117.564	0.748	0.222	-116.797
ID2	0.739	0.226	-99.366	0.753	0.220	-127.342	0.753	0.220	-128.421
ID5	0.739	0.226	-98.456	0.748	0.222	-115.410	0.748	0.222	-117.162
GV	0.741	0.225	-106.228	0.742	0.225	-105.994	0.742	0.225	-104.509

### 2.4 模型参数

对 3 个 SAR 模型及 OLS 的回归参数进行  $t$  检验,对 3 个 SAR 模型的自回归参数进行似然比检验(likelihood ratio test)。由表 6 可知:无论用哪个空间加权矩阵,3 个 SAR 模型的回归参数均显著不等于 0 ( $P < 2.2 \times 10^{-16}$ )。3 个 SAR 模型的回归参数  $\beta_1$  均与 OLS 的  $\beta_1$  比较相似,而回归参数  $\beta_0$  则不然。SEM 的  $\beta_0$  与 OLS 的  $\beta_0$  相似,而 SDM 和 SLM 的  $\beta_0$  与 OLS 的  $\beta_0$  不相似, $\beta_0$  随着空间加权矩阵的变化而发生变化,这种不相似性在 SDM 中表现尤为突出。事实上,这种不相似性主要是因为 SDM 和 SLM 的模型形式中增加了空间滞后变量  $WY$  和  $WX$  所致。

由表 7 可见,GV 应用于 SDM 中时,其自回归参数  $\rho$  和  $\gamma$  均不显著 ( $P > 0.1$ ),说明 GV 不适用于 SDM。DI1 应用于 SLM 中时,其自回归参数  $\rho$  不显著 ( $P > 0.1$ ),说明 ID1 不适用于 SLM。自回归参数  $\rho$  在 SDM 中的值均大于 SLM 的值(GV 除外)。在 SDM 中,所有空间加权矩阵(GV 除外)的自回归参数  $\rho$  和  $\gamma$  均显著不为 0,说明模型残差的空间自相

关是滞后变量  $WY$ (即相邻加权树高)和滞后变量  $WX$ (即相邻加权胸径)共同作用产生的。自回归参数  $\gamma$  远大于  $\rho$ ,这意味着应变量(如树高)的空间自相关性显著强于解释变量(如胸径)的空间自相关性。在 SEM 中,所有空间加权矩阵(GV 除外)的自回归参数  $\lambda$  显著不为 0,说明空间自相关是由滞后变量  $W\xi$ (即相邻加权未观测的变量)产生的。

### 2.5 模型预测

由 OLS、SLM、SDM 和 SEM 4 个模型的模型残差空间自相关、模型拟合及参数分析比较结果可知,ID2 是最好的空间加权矩阵。因此,基于预测数据,分别计算 ID2 构建的 4 个模型的  $e$ 、 $\sigma$  和 MS 模型预测指标,以及预测残差的 MI 值及其  $P$  值(见表 8)。由表 8 可知,3 个 SAR 模型的  $\sigma$  和 MS 均小于 OLS 相应的预测指标值,3 个 SAR 模型的预测误差 MI 值均小于 OLS 的 MI 值,说明 3 个 SAR 模型预测效果均优于 OLS。在 3 个 SAR 模型中,SDM 和 SEM 的  $\sigma$  和 MS 值基本相等,均明显小于 SLM 的预测指标值,同时 SDM 和 SEM 的预测误差 MI 值也相近,均比

表 6 3 个 SAR 模型的回归参数

Tab. 6 Regressive parameters of three SAR models

空间加权矩阵 Spatial weight matrix	SLM				SDM				SEM			
	$\beta_0$	$P$	$\beta_1$	$P$	$\beta_0$	$P$	$\beta_1$	$P$	$\beta_0$	$P$	$\beta_1$	$P$
DT	3.152	$<2.2 \times 10^{-16}$	-8.622	$<2.2 \times 10^{-16}$	2.333	$<2.2 \times 10^{-16}$	-8.758	$<2.2 \times 10^{-16}$	3.327	$<2.2 \times 10^{-16}$	-8.749	$<2.2 \times 10^{-16}$
ID1	3.065	$<2.2 \times 10^{-16}$	-8.619	$<2.2 \times 10^{-16}$	2.052	0.0000	-8.776	$<2.2 \times 10^{-16}$	3.323	$<2.2 \times 10^{-16}$	-8.753	$<2.2 \times 10^{-16}$
ID2	3.069	$<2.2 \times 10^{-16}$	-8.585	$<2.2 \times 10^{-16}$	2.397	$<2.2 \times 10^{-16}$	-8.762	$<2.2 \times 10^{-16}$	3.325	$<2.2 \times 10^{-16}$	-8.754	$<2.2 \times 10^{-16}$
ID5	3.169	$<2.2 \times 10^{-16}$	-8.598	$<2.2 \times 10^{-16}$	2.777	$<2.2 \times 10^{-16}$	-8.721	$<2.2 \times 10^{-16}$	3.323	$<2.2 \times 10^{-16}$	-8.717	$<2.2 \times 10^{-16}$
GV	3.313	$<2.2 \times 10^{-16}$	-8.733	$<2.2 \times 10^{-16}$	3.315	$<2.2 \times 10^{-16}$	-8.760	$<2.2 \times 10^{-16}$	3.318	$<2.2 \times 10^{-16}$	-8.717	$<2.2 \times 10^{-16}$

注:OLS:  $\beta_0 = 3.318$ ,  $P < 2.2 \times 10^{-16}$ ;  $\beta_1 = -8.650$ ,  $P < 2.2 \times 10^{-16}$ 。Note: OLS:  $\beta_0 = 3.318$ ,  $P < 2.2 \times 10^{-16}$ ;  $\beta_1 = -8.650$ ,  $P < 2.2 \times 10^{-16}$ 。

表7 3个SAR模型的自回归参数

Tab.7 Autoregressive parameters of three SAR models

空间加权矩阵 Spatial weight matrix	SLM		SDM				SEM	
	$\rho$	$P$	$\rho$	$P$	$\gamma$	$P$	$\lambda$	$P$
DT	0.062	0.094 3	0.280	0.000 0	3.242	0.000 0	0.289	0.000 0
ID1	0.096	0.123 2	0.357	0.000 0	4.181	0.000 0	0.390	0.000 0
ID2	0.094	0.002 7	0.271	0.000 0	2.694	0.000 0	0.276	0.000 0
ID5	0.056	0.004 5	0.162	0.000 0	1.517	0.000 0	0.162	0.000 0
GV	0.039	0.000 1	0.018	0.349 1	0.609	0.164 6	0.265	0.000 2

SLM的MI值小,说明SDM和SEM的预测效果相近,SLM的预测效果最差。SLM的预测效果明显比SDM和SEM差,其主要原因是SLM的预测误差MI值与MI<sub>0</sub>具有显著差异( $P < 0.01$ ),而SDM和SEM的预测误差MI值与MI<sub>0</sub>差异不显著( $P > 0.1$ ),这说明了模型预测误差的空间自相关可以有效提高预测效果。

表8 模型预测指标及其预测误差的MI值

Tab.8 Model prediction index and MI of prediction residuals

模型 Model	$e$	$\sigma$	MS	MI	$P$
OLS	0.000 0	0.227 6	0.051 8	0.167 2	0.000 0
SLM	0.000 0	0.226 1	0.051 1	0.115 2	0.000 0
SDM	0.000 0	0.220 0	0.048 4	-0.013 9	0.616 8
SEM	0.000 0	0.220 1	0.048 4	-0.013 2	0.637 2

### 3 结论与讨论

以线性化单木胸径树高OLS模型为基准模型,利用3个SAR模型和5个空间加权矩阵构建基于空间自相关的天然云冷杉针阔混交林单木胸径树高模型,可得出以下结论:1)未考虑空间自相关的OLS模型残差存在正空间自相关;2)3个SAR模型中,SDM和SEM的拟合效果最好,SLM最差,但三者均比OLS拟合效果好,这一结论与Kissling等<sup>[24]</sup>、Meng等<sup>[16]</sup>和Lu等<sup>[31]</sup>的结论类似;3)无论使用哪个空间加权矩阵,SLM均不能消除模型残差空间自相关,但可降低空间自相关,在一定程度上提高了模型的拟合效果;4)5个空间矩阵应用于SDM和SEM时,均可以消除模型残差空间自相关,但空间加权矩阵GV只适用于SEM;5)5个空间加权矩阵中,ID2是最好的空间加权矩阵;6)将ID2应用于4个模型进行预测时,SDM和SEM的预测效果明显优于SLM,但3个SAR模型的预测效果均优于OLS。

从林分类型方面,本文只针对天然云冷杉针阔混交林,将空间自回归模型应用于其他林分类型如阔叶混交林和针叶混交林时,是否可以得出类似结

论,还有待进一步研究。对于分树种问题,本文只研究了不分树种的单木胸径树高模型,这是由于分树种构建空间自相关模型时,会改变林木间原有的空间自相关结构,而林木间的空间自相关结构直接关联着林木间的竞争和相互作用,如果原有的空间自相关结构改变了,意味着林木间的竞争和相互作用也改变了。将发生变化的林木间竞争和相互作用考虑到胸径树高模型中,这样建立的模型将与实际不符。因此,在不改变原有林木间空间自相关结构即不改变林木间的竞争和相互作用条件下如何分树种构建模型,是今后主要的研究工作。针对线性或非线性模型问题,本文只用了线性化的空间自相关模型,如何研究非线性的空间自相关模型,就目前来说还是一个难题,也是今后的重点研究方向。

### 参 考 文 献

- [1] 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.
- [2] 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]. The Forestry Chronicle, 2009, 85(1): 43-56.
- [3] GÓMEZ-GARCÍA E, FONSECA T F, CRECENTE-CAMPO F, et al. Height-diameter models for maritime pine in Portugal: a comparison of basic, generalized and mixed-effects models [J/OL]. iForest, 2015[2015-08-01]. <http://www.sisef.it/forest/pdf/?id=ifor1520-1008>.
- [4] LASERN D R, HANN D W. Height-diameter equations for seventeen tree species in Southwest Oregon [C]. Corvallis: Oregon State University, 1987.
- [5] JAYARAMAN K, LAPPI J. Estimation of height-diameter curves through multilevel models with special reference to even-aged teak stands [J]. Forest Ecology and Management, 2001, 142: 155-162.
- [6] COLBERT K, LARSEN D, LOOTENS J. Height-diameter equations for thirteen midwestern bottomland hardwood species [J]. Northern Journal of Applied Forestry, 2002, 19(4): 171-176.

- [ 7 ] HUANG S, TITUS S J. An index of site productivity for uneven-aged or mixed-species stands [ J ]. Canadian Journal of Forest Research, 1993, 23(3) : 558-562.
- [ 8 ] VANCLAY J K. Modelling forest growth and yield: applications to mixed tropical forests [ M ]. Wallingford: CAB International, 1994.
- [ 9 ] CURTIS R O. Height-diameter and height-diameter-age equations for second-growth Douglas-Fir [ J ]. Forest Science, 1967, 13: 365-375.
- [ 10 ] PENG C, ZHANG L, LIU J. Developing and validating nonlinear height-diameter models for major tree species of Ontario's boreal forests [ J ]. Northern Journal of Applied Forestry, 2001, 18(3) : 87-94.
- [ 11 ] FORD E D, DIGGLE P J. Competition for light in a plant monoculture modelled as a spatial stochastic process [ J ]. Annals of Botany, 1981, 48: 481-500.
- [ 12 ] TOMPPA E. Models and methods for analysing spatial patterns of trees [ M ]. Helsinki: The Finnish Forest Research Institute, 1986.
- [ 13 ] MATEU J, USO J, MONTES F. The spatial pattern of a forest ecosystem [ J ]. Ecological Modelling, 1998, 108(1) : 163-174.
- [ 14 ] MAGNUSSEN S. Application and comparison of spatial models in analyzing tree-genetics field trials [ J ]. Canadian Journal of Forest Research, 1990, 20: 536-546.
- [ 15 ] FOX J C, ADES P K, BI H. Stochastic structure and individual-tree growth models [ J ]. Forest Ecology and Management, 2001, 154(1) : 261-276.
- [ 16 ] MENG Q, CIESZEWSKI C J, STRUB M R, et al. Spatial regression modeling of tree height-diameter relationships [ J ]. Canadian Journal of Forest Research, 2009, 39(12) : 2283-2293.
- [ 17 ] ANSELIN L. Spatial econometrics: methods and models [ M ]. Dordrecht: Kluwer Academic Publishers, 1988.
- [ 18 ] LEGENDRE P. Spatial autocorrelation trouble or new paradigm? [ J ]. Ecology, 1993, 74: 1659-1673.
- [ 19 ] LEGENDRE P, LEGENDRE L. Numerical ecology [ M ]. Amsterdam: Elsevier, 1998.
- [ 20 ] MILLER H J. Tobler's First Law and Spatial Analysis [ J ]. Annals of The Association of American Geographers, 2004, 94(2) : 284-289.
- [ 21 ] LENNON J J. Red-shifts and red herrings in geographical ecology [ J ]. Ecography, 2000, 23(1) : 101-113.
- [ 22 ] LEGENDRE P, DALE M R T, FORTIN M J, et al. The consequences of spatial structure for the design and analysis of ecological field surveys [ J ]. Ecography, 2002, 25(5) : 601-615.
- [ 23 ] DORMANN C F, MCPHERSON J M, ARAUJO M B, et al. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review [ J ]. Ecography, 2007, 30(5) : 609-628.
- [ 24 ] KISSLING W D, CARL G. Spatial autocorrelation and the selection of simultaneous autoregressive models [ J ]. Global Ecology and Biogeography, 2008, 17: 59-71.
- [ 25 ] KRMER W. Finite sample efficiency of ordinary least squares in linear regression model with autocorrelated errors [ J ]. Journal of the American Statistical Association, 1980, 75: 1005-1099.
- [ 26 ] WEST P W, RATKOWSKY D A, DAVIS A W. Problems of hypothesis testing of regressions with multiple measurements from individual sampling units [ J ]. Forest Ecology and Management, 1984, 7(3) : 207-224.
- [ 27 ] GREGOIRE T G. Generalized error structure for forestry yield models [ J ]. Forest Science, 1987, 33(2) : 423-444.
- [ 28 ] ZHANG L, SHI H J. Local modeling of tree growth by geographically weighted regression [ J ]. Forest Science, 2004, 50: 225-244.
- [ 29 ] ZHANG L, BI H, CHENG P, et al. Modeling spatial variation in tree diameter-height relationships [ J ]. Forest Ecology and Management, 2004, 189(1-3) : 317-329.
- [ 30 ] ZHANG L, GOVE J H, HEATH L S. Spatial residual analysis of six modeling techniques [ J ]. Ecological Modelling, 2005, 186(2) : 154-177.
- [ 31 ] LU J F, ZHANG L J. Modeling and prediction of tree height-diameter relationship using spatial autoregressive models [ J ]. Forest Science, 2011, 57(3) : 252-264.
- [ 32 ] CRESSIE N A C. Statistics for spatial data [ M ]. Wiley Series in Probability and Mathematical Statistics. New York: Wiley, 1993.
- [ 33 ] HAINING R. Spatial data analysis: theory and practice [ M ]. Cambridge: Cambridge University Press, 2003.
- [ 34 ] 李金良, 郑小贤, 王昕. 东北过伐林区林业局级森林生物多样性指标体系研究 [ J ]. 北京林业大学学报, 2003, 25(1) : 48-52.
- LI J L, ZHENG X X, WANG X. Study on forest biodiversity index system of northeast over cutting forest region based on forestry bureau level [ J ]. Journal of Beijing Forestry University, 2003, 25(1) : 48-52.
- [ 35 ] 杨华, 李艳丽, 沈林, 等. 长白山云冷杉针阔混交林主要树种空间分布及其关联性 [ J ]. 生态学报, 2014, 34(16) : 4698-4706.
- YANG H, LI Y L, SHEN L, et al. Spatial distributions and associations of main tree species in a spruce-fir forest in the Changbai Mountains area in northeastern China [ J ]. Acta Ecologica Sinica, 2014, 34(16) : 4698-4706.
- [ 36 ] 李冰, 樊金拴, 车小强. 我国天然云冷杉针阔混交林结构特征、更新特点及经营管理 [ J ]. 世界林业研究, 2012, 25(3) : 43-49.
- LI B, FAN J S, CHE X Q. A review of studies on structural features, regeneration features and management of natural spruce-fir mixed stand of coniferous and broadleaved trees in China [ J ]. World Forestry Research, 2012, 25(3) : 43-49.
- [ 37 ] 李艳丽, 杨华, 亢新刚, 等. 长白山云冷杉针阔混交林天然更新空间分布格局及其异质性 [ J ]. 应用生态学报, 2014, 25(2) : 311-317.
- LI Y L, YANG H, KANG X G, et al. Spatial heterogeneity of natural regeneration in a spruce-fir mixed broadleaf-conifer forest in Changbai Mountains [ J ]. Chinese Journal of Applied Ecology, 2014, 25(2) : 311-317.
- [ 38 ] HUANG S, TITUS S, WIENS D P. Comparison of nonlinear height-diameter functions for major Alberta tree species [ J ].



- Canadian Journal of Forest Research, 1992, 22: 1297-1304.
- [39] 李希菲,唐守正,袁国仁,等.自动调控树高曲线和一元立木材积模型[J].林业科学研究,1994,7(5):512-518.  
LI X F, TANG S Z, YUAN G R, et al. Self-adjusted height-diameter curves and one entry volume model [J]. Forest Research, 1994, 7(5): 512-518.
- [40] CLIFF A D, ORD J K. Spatial processes-models and applications [M]. Economic Geography. London: Pion Ltd, 1981.
- [41] FORTIN M J, DALE M R T. Spatial analysis: a guide for ecologists [M]. Cambridge: Cambridge University Press, 2005.
- [42] R Development Core Team. R: a language and environment for statistical computing [R/OL]. Vienna: R Foundation for Statistical Computing, 2010 [2015-03-03]. <http://www.R-project.org>.
- [43] BIVAND R. SPDEP: spatial dependence: weighting schemes, statistics and models: R package version 0.5-88 [R/OL]. [2015-03-03]. <http://www.cran.r-project.org/web/packages/spdep/spdep.pdf>.
- [44] GETIS A, ALDSTADT J. Constructing the spatial weights matrix using a local statistic [J]. Geographical Analysis, 2004, 36(2): 90-104.
- [45] SMIRNOV O, ANSELIN L. Fast maximum likelihood estimation of very large spatial autoregressive models: a characteristic polynomial approach [J]. Computational Statistics and Data Analysis, 2001, 35(3): 301-319.
- [46] LESAGE J P. Spatial econometrics [M/OL]. [2015-07-03]. [https://www.researchgate.net/publication/277298256\\_Spatial\\_Econometrics](https://www.researchgate.net/publication/277298256_Spatial_Econometrics).
- [47] PINHEIRO J C, BATES D M, DEBROY S, et al. Nlme: Linear and nonlinear mixed effects models: R package version 3.1-120 [R/OL]. [2015-03-03]. <http://www.cran.r-project.org/web/packages/nlme/nlme.pdf>.
- [48] PARADIS E. Moran's autocorrelation coefficient in comparative methods [M/OL]. [2015-03-14]. <http://www.cran.r-project.org/web/packages/ape/vignettes/MoranI.pdf>.
- [49] ARABATZIS A A, BURKHART H E. An evaluation of sampling methods and model forms for estimating height-diameter relationships in loblolly pine plantations [J]. Forest Science, 1992, 38(1): 192-198.
- [50] ZHANG L J. Cross-validation of nonlinear growth functions for modeling tree height-diameter distributions [J]. Annals of Botany, 1997, 79(3): 251-257.

(责任编辑 冯秀兰  
责任编辑 黄选瑞)