

doi:10.6041/j.issn.1000-1298.2019.12.024

基于两种空间估算模型的乔木林地上碳密度估算

王海宾¹ 郑冬梅² 王少杰¹ 贾筱昕¹ 许等平¹

(1. 国家林业和草原局林产工业规划设计院, 北京 100010; 2. 国家林业和草原局调查规划设计院, 北京 100714)

摘要: 以乔木林地上碳密度为研究对象,基于调查获取的乔木林地上碳密度、Landsat 8 多光谱影像及 DEM 数据,提取植被指数、纹理特征、主成分变换因子、缨帽变换因子和地形因子作为建模变量,采用皮尔森相关系数法、结合平均残差平方和准则法对变量进行筛选,采用协同克里格插值和地理加权回归方法构建乔木林地上碳密度模型,分析对比两种方法的估算效果。结果表明:地理加权回归法构建的估算模型精度(R^2 为 0.74, RMSE 为 6.84 t/hm², MAE 为 5.13 t/hm², RE 为 0.74%) 优于协同克里格插值法(R^2 为 0.47, RMSE 为 9.72 t/hm², MAE 为 7.41 t/hm², RE 为 0.12%),并且较好地保留了估算变量的空间异质性,变异系数分别为 0.537 2、0.496 8,可获得较高的估算精度。本研究可为大尺度范围内的乔木林地上碳密度及其他森林参数的估算提供参考。

关键词: 乔木林; 地上碳密度; Landsat 8 影像; 协同克里格插值; 地理加权回归; 估算
中图分类号: S757; TP79 **文献标识码:** A **文章编号:** 1000-1298(2019)12-0212-09

Estimation of Above-ground Carbon Density Prediction of Arbor Forest Based on Two Spatial Estimation Models

WANG Haibin¹ ZHENG Dongmei² WANG Shaojie¹ JIA Xiaoxin¹ XU Dengping¹

(1. Planning and Design Institute of Forest Products Industry, National Forestry and Grassland Administration, Beijing 100010, China
2. Academy of Inventory and Planning, National Forestry and Grassland Administration, Beijing 100714, China)

Abstract: Based on Landsat 8 multi-spectral imagery and ground survey data, taking the above-ground carbon density of arbor forest as the research object, the field survey data of above-ground carbon density of arbor forest, Landsat 8 multi-spectral image and DEM data were used to extract vegetation indices, texture features, principal component transformation factors, cap transformation factors and topographic factors as modeling variables. Pearson correlation coefficient method combined with residual mean square criterion method was used to screen variables. Co-Kriging interpolation and geographic weighted regression method were used to construct above-ground carbon density of arbor forest. And the estimated effect of the two methods were compared and analyzed. The results showed that the accuracy of the estimated model constructed by the geographic weighted regression method (R^2 was 0.74, RMSE was 6.84 t/hm², MAE was 5.13 t/hm², RE was 0.74%), which was superior to the Co-Kriging interpolation method (R^2 was 0.47, RMSE was 9.72 t/hm², MAE was 7.41 t/hm², RE was 0.12%), and the spatial heterogeneity of the estimated variables was well preserved ($C_{V\text{GWR}} = 0.537\ 2$, $C_{V\text{COK}} = 0.496\ 8$), the geographic weighted regression method can obtain higher estimation accuracy. The research can provide a reference for estimating the aboveground carbon density of arbor forest and other forest parameters of forest at regional or large scale.

Key words: arbor forest; above-ground carbon density; Landsat 8 imagery; Co-Kriging interpolation; geographic weighted regression; estimation

0 引言

森林参数的准确估算是林学研究的热点之

—^[1-2],对森林经营管理、生物多样性保护、全球气候变化研究具有重要意义。掌握森林参数现状及分布情况,有利于森林经营管理者科学管理森林,对在

收稿日期: 2019-03-03 修回日期: 2019-07-12
基金项目: 国家重点林业工程监测技术示范推广项目([2015]02号)和国家林业局948项目(2015-4-32)
作者简介: 王海宾(1985—),男,助理工程师,博士,主要从事林业遥感与信息技术研究,E-mail: haibin-w@163.com
通信作者: 郑冬梅(1968—),女,教授级高级工程师,主要从事森林资源监测及遥感技术应用研究,E-mail: zhengdm2001@sina.com

不同层面制定森林经营管理政策、调整经营管理策略具有重要作用^[3]。

近年来,遥感技术得到了快速发展,因即时、快速和低成本等特点而使其在国土、农林、水利、矿业等领域得到了广泛应用,取得了较好的应用效果^[4-5]。尤其是在区域和大尺度范围的监测上,光学遥感数据覆盖范围广、获取周期短、数据成本低廉等,具有激光雷达、微波雷达等其他遥感数据无法比拟的优势^[6]。因此,应用光学遥感技术结合地面样地数据对森林参数进行估算成为必要的手段^[7]。

笔者曾参与全国森林植被碳储量遥感估算工作^[8-9],其中的乔木林碳储量估算主要基于 ZY-3 多光谱影像,结合地面调查数据,应用传统的线性模型,以省为单位来估算乔木林碳储量。由于覆盖全省的 ZY-3 多光谱影像数据量较大,存在时相不一致的问题,同时因拍摄角度、云雾等多因素干扰,增加了估算结果的不确定性。为改善估算效果,依据估算目标范围以及遥感影像空间分辨率的适用性等因素,选择 Landsat 8 多光谱影像,结合实地调查数据进行了森林植被碳储量估算的研究分析,以期为区域及全国尺度范围的森林碳储量估算提供更好的数据源及方法,减小估算结果的不确定性,达到增强监测成果时效性、提高监测效率、降低工作成本、改进和优化全国森林植被碳储量估算数据源和估算方法的目的。

为挖掘 Landsat 8 多光谱影像在全国森林植被碳储量遥感估算中的应用潜力,同时探索更好的估算方法,本研究以位于浙江省内的一景 Landsat 8 影像覆盖范围为研究区,以乔木林地上碳密度为研究对象,基于研究区内的地面样地调查数据,通过计算获得乔木林地上碳密度数据并将其作为因变量,采用协同克里格插值法和地理加权回归方法,构建乔木林地上碳密度模型,并对建模结果进行对比分析,以期为区域及全国范围的乔木林地上碳密度及森林参数遥感快速估算提供更好的数据源和估算方法,并提高遥感估算结果的精度和可靠性。

1 研究区及数据获取

1.1 研究区概况

研究区为浙江省内一景 Landsat 8 影像覆盖的范围(图 1),包括金华市、杭州市、绍兴市、衢州市、丽水市和温州市,地理坐标为东经 118°28′20″~120°49′11″,北纬 27°48′3″~29°55′26″。区内以丘陵和山地为主,属于亚热带季风气候,四季分明,雨热同季。区内土壤类型多样,主要有红壤、黄壤、粗骨土、水稻土、潮土、盐土等类型。研究区森林资源丰

富,树种较多,森林类型以针叶林为主,阔叶林、针阔混交林较多,主要包括马尾松(*Pinus massoniana* Lamb.)、杉木(*Cunninghamia lanceolata* (Lamb.) Hook.)、栎类(*Cryptomeria fortunei* Hooibrenk ex Otto et Dietr.)、木荷(*Schima superba* Gardn. et Champ.)以及其他软阔类等优势树种。

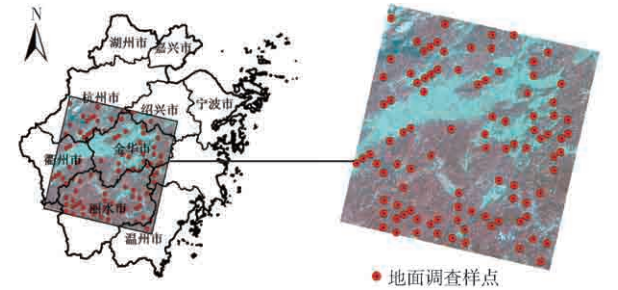


图 1 研究区内调查样地分布
Fig. 1 Distribution of survey samples in study area

1.2 数据获取

1.2.1 地面数据

地面数据获取时间为 2012 年。为使样地具有代表性,依据优势树种和林分年龄(幼龄林、中龄林、近熟林、成熟林、过熟林),设置和调查了 96 个样地(图 1),样地形状为正方形,面积为 0.067 hm²,涵盖针叶林、阔叶林、混交林等不同森林类型。样地属性因子包括树种、胸径(起测胸径为 5 cm)、平均树高、郁闭度、株数、年龄、坡向、坡度、海拔、乔木林蓄积量等因子。

将获取的样地乔木林地上蓄积量乘以生物量换算因子(Biomass expansion factor, BEF)得到乔木林地上生物量,并转换为每公顷生物量(t/hm²)。采用沈楚楚^[10]提出的浙江省四大树种的 BEF,得到每公顷生物量分别为杉木林 0.745 3 t/m³、马尾松林 0.883 9 t/m³、软阔类 1.657 2 t/m³、硬阔类 1.070 5 t/m³。地面样地乔木林地上生物量数据统计结果见表 1。

表 1 地面样地乔木林地上生物量统计结果
Tab.1 Statistic results of aboveground biomass of arbor forest for ground sample

		t/hm ²			
森林类型	样地数	最小值	最大值	平均值	标准差
针叶林	51	1.76	106.89	32.96	24.55
阔叶林	20	6.67	287.21	67.63	58.65
混交林	25	11.34	81.38	39.20	20.95

将每公顷生物量乘以树种含碳率即可得到单位面积乔木林地上碳密度。树种含碳率参考《土地利用、土地利用变化与林业碳汇计量监测技术指南》,指南中的树种含碳率可覆盖研究区内的各个优势树种。

在获得乔木林地上碳密度数据后,采用二倍标

准差法剔除异常值,结果剩余 95 个地面样地数据,将其作为因变量参与乔木林地上碳密度模型的构建。

1.2.2 遥感影像及辅助数据

本研究选用一景 Landsat 8 影像作为遥感数据源,覆盖整个研究区,影像轨道号为 119/040,空间分辨率为 30 m,采集时间为 2014 年 10 月 26 日。由于树木生长较慢,本研究假设地面数据采集时间与影像获取时间不存在差异。应用 ENVI 5.3 软件对 Landsat 8 多光谱影像进行预处理,包括辐射定标、大气校正。本研究所采用波段序号为 2~7,分别为蓝、绿、红、近红外、短波红外 1、短波红外 2 共 6 个波段。

本研究采用的辅助数据为覆盖研究区 1:50 000 数字高程模型 (Digital elevation model,DEM) 数据。

2 研究方法

2.1 空间自相关分析

采用协同克里格插值法对乔木林地上碳密度建模前,需要对使用的碳密度数据进行描述性统计分析,采用柯尔莫哥洛夫-斯蜜诺夫检验 (K-S 检验)^[9] 对数据进行分析,检验其是否符合正态分布。

此外,采用地理加权回归法建模前,需要对乔木林地上碳密度进行空间自相关分析,检验研究数据是否随机分布。采用莫兰指数 (Moran's I) 进行检验^[11],计算公式如下

$$I = \frac{n \sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x}) (x_j - \bar{x})}{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})^2} \quad (1)$$

式中 n ——样本总数
 I ——莫兰指数
 $\bar{x}$ ——变量平均值
 x_i, x_j ——空间位置 i 和 j 的变量取值
 w_{ij} ——空间位置 i 和 j 的权重

Moran's I 的取值范围在 $[-1, 1]$ 之间,小于 0 表示负相关,等于 0 表示不相关,大于 0 则表示正相关。

2.2 自变量提取及优选

2.2.1 自变量提取

自变量一般选择光学遥感数据、地形数据以及经计算衍生的一些因子,常用的包括植被指数^[12]、纹理特征^[13]、主成分变换因子^[3]、缨帽变换因子^[3]、地形因子^[14]等。本研究基于 Landsat 8 多光谱波段及 DEM 数据,计算了植被指数 (9 个)、纹理特征 (48 个)、主成分变换因子 (3 个)、缨帽变换因子 (3 个)

和地形因子 (3 个) 共 66 个特征作为自变量参与模型构建。其中,选用的植被指数见表 2,其中 B_{LUE} 、 R_{ED} 、 N_{IR} 、 S_{WIR1} 、 S_{WIR2} 分别为蓝、红、近红、短波红外 1、短波红外 2 波长;SAVI 中, L 为 0.5;ARVI 中, r 为校正系数,取 1.0;EVI 中,增益系数 G 、土壤调整因子 L 、修正系数 C_1 和 C_2 分别为 2.5、0.10、6.0 和 7.5。纹理特征采用灰度共生矩阵方法中的 8 个纹理特征 (平均值、方差、均一性、对比度、相异性、熵、角二阶矩、相关性);主成分变换和缨帽变换因子分别选择主成分分析获取的前 3 个波段和缨帽变换的前 3 个波段 (亮度、绿度和湿度);地形因子采用 DEM 数据提取的海拔、坡度、坡向 3 个因子。按照地面实测样地位置,提取各个样地对应的自变量因子。

表 2 植被指数
Tab.2 Vegetation indices

变量	计算公式
归一化植被指数 (NDVI)	$NDVI = \frac{N_{IR} - R_{ED}}{N_{IR} + R_{ED}}$
比值植被指数 (RVI)	$R_{VI} = N_{IR} / R_{ED}$
差值植被指数 (DVI)	$D_{VI} = N_{IR} - R_{ED}$
比值植被指数 1 (RVI54)	$R_{VI54} = S_{WIR1} / N_{IR}$
比值植被指数 2 (RVI64)	$R_{VI64} = S_{WIR2} / N_{IR}$
土壤植被指数 (SAVI)	$S_{AVI} = \frac{(1 + L) (N_{IR} - R_{ED})}{N_{IR} + R_{ED} + L}$
优化土壤调整指数 (NLI)	$N_{LI} = \frac{N_{IR}^2 - R_{ED}}{N_{IR}^2 + R_{ED}}$
大气抗阻植被指数 (ARVI)	$A_{RVI} = \frac{N_{IR} - R_{ED} + r (B_{LUE} - R_{ED})}{N_{IR} + R_{ED} - r (B_{LUE} - R_{ED})}$
增强植被指数 (EVI)	$EVI = \frac{G (N_{IR} - R_{ED})}{N_{IR} + C_1 R_{ED} - C_2 B_{LUE} + L}$

2.2.2 自变量优选

在构建模型时,自变量对所构建模型的稳定性有很大影响,有些因子对碳密度估算的影响很小,甚至没有影响,则将其从自变量因子集中剔除,将剩余的因子组成最优自变量集。在本研究中,分两步对自变量进行优选:首先利用相关性分析提取自变量因子;其次选用平均残差平方和 (Residual mean square, RMSQ) 准则^[15],对自变量做进一步优选。

2.3 建模方法

2.3.1 协同克里格插值法

协同克里格插值法 (Co-Kriging, COK) 在变量估测上具有较好的预测精度^[8,16-18],它是利用两个或两个以上的变量,将其中一个作为主变量,剩余变量作为辅助变量,将主变量的自相关性和主辅变量的交互相关性结合起来进行无偏最优估计^[17-18],协同克里格插值公式如下

$$Z^* (x_0) = \sum_{i=1}^n \lambda_{1i} Z_1 (x_i) + \sum_{j=1}^m \lambda_{2j} Z_2 (x_j) \quad (2)$$

其中
$$\sum_{i=1}^n \lambda_{1i} = 1 \quad \sum_{i=1}^n \lambda_{2i} = 0$$

式中 $Z^*(x_0)$ ——待估点处的蓄积量估测值
 λ_{1i} 、 λ_{2i} ——主变量 Z_1 和辅变量 Z_2 实测值的权重
 n 、 m ——参与估测 x_0 点的蓄积量和辅助变量 Z_2 的实测值数目

2.3.2 地理加权回归法

地 理 加 权 回 归 法 (Gergraphic weighted regression, GWR)是用回归原理研究具有空间(或区域)分布特征的两个或多个变量之间数量关系的方法,它是基于普通全局回归模型的扩展,将空间影响到模型构建中,以距离权重的形式加入到模型之中,故地理加权回归模型^[3]的基本形式如下

$$\hat{y} = \hat{\beta}_0(x,y) + \hat{\beta}_1(x,y)X_1 + \hat{\beta}_2(x,y)X_2 + \cdots + \hat{\beta}_N(x,y)X_N + \varepsilon \tag{3}$$

式中 x,y ——空间样本的坐标值
 N ——样本数 $\hat{\beta}_0$ ——截距
 $\hat{\beta}_n$ ——第 N 个变量的回归参数
 ε ——模型残差,服从 $N(0,\sigma^2)$ 分布
 X_n ——因子变量 $\hat{y}$ ——因变量

采用地理加权回归法构建模型时,模型带宽的选择至关重要,在很大程度上决定模型的预测准确度^[3,14]。本研究采用最小化信息准则 (Akaike information criterion, AIC)^[9] 来确定最佳带宽:当 AIC 值取最小值时,即可获得最优带宽及地理加权回归模型。AIC 值计算公式为

$$A_{IC} = -2\ln H + 2k \tag{4}$$

式中 k ——未知参数的数量
 H ——似然函数 A_{IC} ——AIC 值
最优模型带宽在 AIC 值最小时获得。

此外,在完成 GWR 模型构建后,需进一步对 GWR 模型进行参数的非平稳性检验,检验模型效果。地理加权回归理论和 GWR 模型参数的非平稳性检验详见文献[3,14]。

2.4 模型精度评价

本文采用留一交叉验证法对建模结果进行评价。采用的评价指标包括:决定系数 (R^2)、均方根误差 (Root mean square error, RMSE)、平均绝对误差 (Mean absolute error, MAE) 和总预报偏差的相对误差 (Relative error, RE)。

3 研究结果

3.1 描述性统计及空间自相关性分析

3.1.1 描述性统计分析

使用 SPSS 20.0 统计软件,采用 K-S 检验对数据进行检验。经分析,乔木林地上碳密度数据服从正态分布,K-S 检验值为 0.374,其平均值为 19.61 t/hm²,标准差 13.29 t/hm²,中位数 17.91 t/hm²,最小值为 0.88 t/hm²,最大值为 56.02 t/hm²。

3.1.2 空间自相关分析

本研究对乔木林地上碳密度进行空间自相关分析,结果显示研究区内乔木林地上碳密度的莫兰指数为 0.083 7,呈空间正相关,并且在 10% 的显著性水平 ($p = 0.098$) 下 Z 为 1.654 (大于 1.65),高度显著。依据莫兰指数^[11]的定义,当其取值范围在 (0,1] 之间时,乔木林地上碳密度在空间分布上呈现聚集状态,乔木林地上碳密度具有整体空间分布特性,说明研究区内的乔木林地上碳密度数据分布是非随机的。

3.2 自变量优选结果

应用 SPSS 20.0 软件对提取的乔木林地上碳密度和 66 个自变量因子进行皮尔森 (Pearson) 相关性分析,筛选出与碳密度极显著相关 ($p < 0.01$) 的自变量因子,提取的自变量因子共计 22 个,结果详见表 3。B1、B2、B3、B6 分别表示 Landsat 8 影像的第 2 (蓝)、3 (绿)、4 (红)、7 (短波红外 2) 波段; Mean 表示平均值纹理, Entropy 表示熵纹理, Second 表示角二阶矩纹理, Correlation 表示相关性纹理, Variance 表示方差纹理, Contrast 表示对比度纹理, Diss 表示相异性纹理。

表 3 自变量因子与乔木林地上碳密度的 Pearson 相关系数

Tab.3 Pearson correlation coefficients between independent variables and above-ground carbon density of arbor forest					
变量	相关系数	变量	相关系数	变量	相关系数
NDVI	0.644 **	Greenness	0.306 **	B6_Mean	-0.320 **
RVI	0.678 **	海拔	0.321 **	B6_Variance	-0.345 **
RVI54	-0.400 **	B1_Mean	-0.336 **	B6_Contrast	-0.295 **
RVI64	-0.398 **	B2_Mean	-0.342 **	B6_Diss	-0.277 **
SAVI	0.644 **	B3_Mean	-0.407 **	B6_Entropy	-0.318 **
NLI	0.387 **	B3_Entropy	-0.321 **	B6_Second	0.285 **
ARVI	0.646 **	B3_Second	0.316 **		
PCA2	-0.500 **	B3_Correlation	0.282 **		

注: **表示在 0.01 水平 (双侧) 上显著相关,下同。

在相关性分析的基础上,本研究采用 RMSQ 法对 22 个自变量因子进一步优选,结果见表 4。经筛选后,RVI、RVI54、RVI64、Greenness、B3_Correlation、B6_Mean 共计 6 个因子被优选出来。

表 4 基于 RMSQ 筛选的自变量因子
Tab.4 Independent factors filter based on RMSQ

变量	相关系数
RVI	0.678 **
RVI54	-0.400 **
RVI64	-0.398 **
Greenness	0.306 **
B3_Correlation	0.282 **
B6_Mean	-0.320 **

表 5 基于 RMSQ 不同变量拟合的最优方差函数模型及参数

Tab.5 Optimal variance function model and parameters fitted by different variables extracted by RMSQ

自变量组	模型	R^2	残差平方和 $R_{ss}/$ ($t^2 \cdot \text{hm}^{-4}$)	块金值 C_0	偏基台值 C	$C/(C+C_0)$	变程/m
RVI	孔穴效应模型	0.43	10 189.80	0	174.33	1	15 557.59
RVI、RVI54	K-贝塞耳模型	0.47	8 995.34	0	195.10	1	62 507.09
RVI、RVI54、RVI64	K-贝塞耳模型	0.47	8 967.56	0	194.87	1	61 702.32

为 0,块金值与基台值的比值小于 25%,表明系统具有很强的相关性。依据变异函数的评价标准可知,在使用 RVI 进行协同克里格插值时,孔穴效应模型可获得最优的变异函数拟合效果,在使用 RVI、RVI54 插值时,K-贝塞耳模型可获得最优的拟合结果,在使用 RVI、RVI54、RVI64 时,K-贝塞耳模型可获得最优的拟合结果。

(2)精度评价

基于 3 组自变量,应用协同克里格插值法对乔木林地上碳密度进行建模,模型精度评价指标中,RVI、RVI54、RVI64 组成的自变量组所构建模型具有最高预测精度(R^2 为 0.47, RMSE 为 9.72 t/hm^2 , MAE 为 7.41 t/hm^2 , RE 为 0.12%),优于 RVI 构建的模型精度(R^2 为 0.43, RMSE 为 10.36 t/hm^2 , MAE 为 8.22 t/hm^2 , RE 为 0.19%)和 RVI、RVI54 所构建的模型精度(R^2 为 0.47, RMSE 为 9.73 t/hm^2 , MAE 为 7.44 t/hm^2 , RE 为 0.09%)。由以上分析可知,应用协同克里格插值估算乔木林地上碳密度时,选用相关性较高的自变量因子参与插值,可获得较好的预测效果。

3.3.2 地理加权回归法

本研究使用地理加权回归法构建乔木林地上碳密度模型时,采用与协同克里格插值相同的自变量组构建 GWR 模型,并对构建的 GWR 模型进行对比分析。

(1)带宽选择

在模型带宽的选择上,分别采用固定距离法和

3.3 模型构建

3.3.1 协同克里格插值法

(1)模型拟合

基于优选的自变量,提取出与乔木林地上碳密度相关性最高的 3 个自变量(RVI、RVI54、RVI64),分别建立 3 个自变量组,见表 5。

采用协同克里格插值法对乔木林地上碳密度进行建模。本文对 ArcGIS 10.1 地统计分析模块中的 11 种变异模型进行了计算和分析,根据变异函数的理论和评价方法,得到最优变异函数的拟合结果(表 5)。由表 5 可知,在协同克里格插值中,采用基于 RMSQ 提取的 3 组自变量参与插值后,块金值 C_0

自适应方法,确定最优带宽。本研究基于植被指数 RVI 因子的模型带宽的计算结果见表 6 和图 2。当采用固定距离法时,模型带宽为 $2\,514\,047 \text{ m}$,模型决定系数 R^2 为 0.46,调整后 R^2 为 0.45;采用自适应方法选择带宽时,根据 AIC 确定的最佳带宽可知,随着带宽(临近点个数)的增加,AIC 值在最开始呈现大幅度减小趋势,随后减小的趋势逐渐趋于平缓,在临近点个数为 30 个时,AIC 值的变化斜率明显降低,变化幅度减小(图 2)。因此,可确定临近点个数 30 为最优带宽。对比两种方法(表 6)可知,基于自适应法选择的带宽具有较高的调整决定系数(0.48)和较低残差平方和 R_{ss} ($28\,276.96 \text{ t}^2/\text{hm}^4$),优于固定距离法(调整决定系数为 0.45, R_s 为 $36\,238.38 \text{ t}^2/\text{hm}^4$)。因此,本研究选择自适应法校准权重函数。由以上分析可知,基于植被指数 RVI 构建的 GWR 模型的最优临近点个数为 30 个,即表示选择回归点周边的 30 个点作为核局部带宽中最为

表 6 GWR 模型参数计算结果

Tab.6 Parameter estimation results of GWR model

模型参数	固定距离法	自适应法
带宽	$2\,514\,047 \text{ m}$	30 个
残差平方和 $R_{ss}/(t^2 \cdot \text{hm}^{-4})$	36 238.38	28 276.96
有效个数/个	2	18.72
残差估计标准差/ $(t \cdot \text{hm}^{-2})$	19.75	19.25
AIC 值	840.67	847.75
决定系数	0.46	0.58
调整决定系数	0.45	0.48

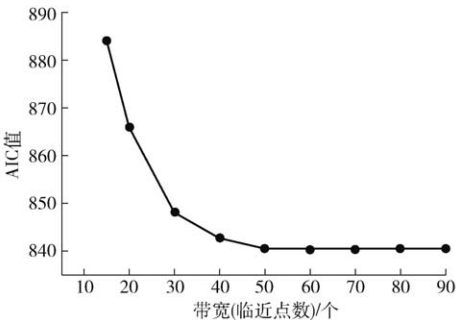


图2 GWR模型最优带宽

Fig.2 Selection of optimal bandwidth for GWR model

临近要素的点,参与建模。

(2)模型拟合

在确定最优带宽后,本研究借助 ArcGIS 10.1 软件中的地理加权回归模块,分别基于协同克里格法中采用的2种自变量组构建 GWR 模型。对 95 个样地碳密度实测值进行估算和精度评价。自变量组为 RVI、RVI54、RVI64 构建的模型具有最高的预测精度 (R^2 为 0.74, RMSE 为 6.84 t/hm², MAE 为 5.13 t/hm², RE 为 0.74%), 优于 RVI、RVI54 构建的模型预测精度 (R^2 为 0.71, RMSE 为 7.18 t/hm², MAE 为 5.37 t/hm², RE 为 0.65%) 和 RVI 构建的模型预测精度 (R^2 为 0.58, RMSE 为 8.63 t/hm², MAE 为 6.58 t/hm², RE 为 0.21%), 可以获得较为满意的估算效果。由此可知,在应用 GWR 构建模型时,随着自变量个数的增加,GWR 模型的拟合精度逐渐提高,可以获得较为满意的估算效果。因此,在构建 GWR 模型时,选择合适的解释变量对提高 GWR 模型预测精度具有重要作用。

(3)参数的非平稳性检验

本研究应用 ArcGIS 10.1 软件计算了普通最小二乘 (Ordinary least square, OLS) 模型和 GWR 模型的回归系数,结果见表 7。根据文献[3]的非平稳性检验标准,GWR 模型各回归系数的第 1 四分位(25%) 和第 3 四分位(75%) 值的变化范围都大于 OLS 模

型的二倍标准误值,因此,认为乔木林地上碳密度的空间关系是非平稳性的。通过计算 OLS 和 GWR 模型的 AIC 值,结果显示两种模型的 AIC 值之差大于 3,进一步证明了乔木林地上碳密度存在明显的空间非平稳性。

表 7 空间平稳性检验

Tab.7 Stationary test of relationship

自变量	OLS 模型		GWR 模型		变化幅度
	标准误	二倍标准误	第 1 四分位值	第 3 四分位值	
截距	7.610 4	15.220 8	-0.010 7	20.903 1	20.913 8
RVI	0.343 2	0.686 4	2.412 1	3.239 3	0.827 2
RVI54	25.454 2	50.908 4	-92.687 3	1.426 7	94.114 0
RVI64	34.811 2	69.622 4	-28.980 8	80.901 1	109.881 9

3.3.3 两种建模方法比较

为对比两种建模方法的估算效果,本研究计算了两种方法预测的乔木林地上碳密度的变异系数,并对两种模型的预测结果进行对比分析,采用 R^2 、残差平方和 (Residual sum of squares, RSS) 以及均方根误差 (RMSE)、变异系数 (Coefficient of variation, CV) 等指标对其进行验证 (表 8)。由表 8 可知, GWR 模型可获得最好的预测效果 ($R^2 = 0.74, R_{ss} = 2\,500.11\text{ t}^2/\text{hm}^4$, RMSE 为 5.13 t/hm²), 其次为协同克里格插值 ($R^2 = 0.47, R_{ss} = 8\,975.45\text{ t}^2/\text{hm}^4$, RMSE 为 9.72 t/hm²), 说明 GWR 模型的预测精度较好。经统计,实测碳密度的变异系数为 0.677 7。由表 8 可知,GWR 模型的变异系数最大,为 0.537 2, 优于协同克里格插值法的 0.496 8, GWR 模型预测的乔木林地上碳密度保留了 0.537 2 的空间异质性特征。由以上分析可知 GWR 模型采用了局部回归方式,在一定程度上提高了模型的预测精度,较好地保留了乔木林地上碳密度的空间异质性,与协同克里格插值法相比,具有较高的空间模拟预测精度。

表 8 两种模型估算精度比较

Tab.8 Comparison of estimation accuracy of two models

模型	变量组	R^2	$R_{ss}/(\text{t}^2 \cdot \text{hm}^{-4})$	$\text{RMSE}/(\text{t} \cdot \text{hm}^{-2})$	CV
协同克里格插值法	RVI、RVI54、RVI64	0.47	8 975.45	9.72	0.496 8
GWR	RVI、RVI54、RVI64	0.74	2 500.11	5.13	0.537 2

两种模型的乔木林地上碳密度预测偏差分布如图 3 所示,协同克里格插值法的预测偏差为 -19.54 ~ 27.79 t/hm², GWR 模型的预测偏差为 -14.52 ~ 18.30 t/hm², 由此可知 GWR 模型的残差分布区间小于协同克里格插值法,说明 GWR 模型具有更好的预测效果。

4 讨论

笔者分别采用了协同克里格插值和地理加权回归对蓄积量和生物量进行了估算^[8-9], 结果显示地理加权回归估算效果优于协同克里格插值法,但其研究仅仅采用了测树因子作为自变量进行了估算,

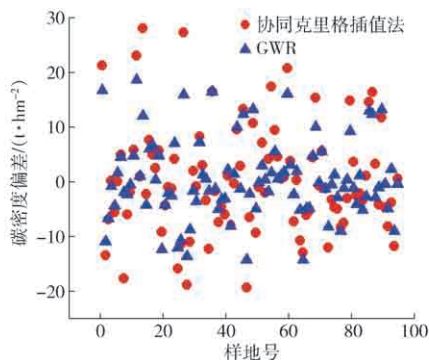


图 3 两种模型的残差分布

Fig. 3 Distribution of residuals of two models

而这些自变量往往缺少空间连续性特征。本研究将 Landsat 8 多光谱影像及 DEM 数据提取的因子作为自变量参与建模,这些衍生的自变量因子具有空间连续性特征,可以有效避免之前研究中自变量因子不具有空间连续性特征的缺点,并且可以减少外业调查的工作量。由本文的研究结果可知,地理加权回归模型适用于地级市及区域尺度的乔木林地上碳密度估算,因为乔木林地上碳密度存在空间非平稳性。本研究采用 Landsat 8 多光谱影像结合地理加权回归法对乔木林地上碳密度处理可以获得较好的估算效果,与前期应用 ZY-3 多光谱影像进行测试估算的效果相比, Landsat 8 多光谱影像在空间分辨率以及多光谱波段方面具有更适合估算区域尺度乔木林地上碳密度的优势,为区域森林植被碳储量和森林参数的估算提供了较好的数据源和估算方法。

准确估算森林参数的关键环节包括两部分:高质量的遥感数据和地面数据以及估算方法。在地面数据确定的基础上,遥感数据源和估算方法成为决定估算效果的关键因素。首先,在遥感数据上,光学遥感数据在森林参数的估算中得到了广泛的研究和应用,但只反映森林的水平信息,而目前激光雷达数据能够提供森林垂直信息(如树高),并且在很多研究中显示了很好的适用性^[19-20]。因此,在以后的研究和推广应用中可以引入激光雷达数据,将其与光学遥感数据进行协同,估算森林参数,获取更高的估算精度。其次,在估算方法上,应用较为广泛的方法包括参数法、非参数法和机理模型 3 种主要方法^[21],每种方法都有自己的优势和劣势,参数法大多是通过回归模型来进行估算,非参数法虽然可提高估算精度,但其无法反映估算变量和自变量间的内在关系,机理模型虽然考虑了估算变量和自变量间的内在机理关系,但其模型机理往往较为复杂,在中大尺度范围内进行应用往往需要大量的数据支撑。由此可知,应用较为简单的估算方法获得精度更高的估算效果成为中大尺度森林参数估算的关键。

在自变量的提取上,本文采用了 Landsat 8 多光谱影像的植被指数及纹理特征,文献[22]表明, Landsat 系列影像的全色波段在遥感定量估算中具有较好的适用性,在后续研究中,建议进一步测试分析全色波段纹理特征在模型构建中的适用性。此外,地形因子与乔木林地上碳密度显示了较好的相关性,由此推测与其相关的立地条件、气象因子及生态因子等因素也可能成为碳密度估算的影响因素,这部分因子对估算结果的影响有待进一步挖掘分析和研究。在自变量的优选上,本研究分两步完成变量优选:一是通过相关性分析提取自变量,目的是减少共线性因子的影响;二是通过平均残差平方和准则进一步优选自变量,目的是去除对模型构建影响较小的因子,使模型的残差平方和更小,获得拟合效果更优的模型。本研究结果显示,应用相关性分析与平均残差平方和准则相结合的方法来优选自变量可获得较好的估算结果。

在本研究中还存在以下不足:①受限于数据使用权限,本研究未对多光谱影像特征对估算结果的影响进行分析。②本研究假设影像数据获取时间和地面数据获取时间不存在时间差,同时受限于地面调查数据量及调查时间的限制,只采用了留一交叉验证法进行验证。③相关性分析结合平均残差平方和准则优选变量进行建模的效果是否优于单一变量筛选方法建模的结果,有待于进一步研究分析。在后续研究中应对存在的不足进行深入分析,减小估算结果的不确定性,并在全国或更大尺度范围内,结合地形、环境、气候等诸多因素,做进一步的研究和推广应用,为区域及全国森林植被碳储量及森林参数遥感估算提供更优的方法,提高估算结果的精度和可靠性。

5 结束语

基于 Landsat 8 多光谱影像和 DEM 数据,提取了遥感及地形因子共计 66 个,结合地面样地调查数据,采用皮尔森相关系数法和 RMSQ 法对自变量进行优选,分别采用协同克里格插值和地理加权回归法构建乔木林地上碳密度模型,并对结果进行了分析对比。结果显示,地理加权回归法构建的估算模型精度(R^2 为 0.74, RMSE 为 6.84 t/hm², MAE 为 5.13 t/hm², RE 为 0.74%) 优于协同克里格插值法(R^2 为 0.47, RMSE 为 9.72 t/hm², MAE 为 7.41 t/hm², RE 为 0.12%),并且较好地保留了估算变量的空间异质性,变异系数分别为 0.537 2、0.496 8,说明应用 Landsat 8 多光谱影像结合地理加权回归模型可获得较好的估算效果。

参 考 文 献

- [1] 刘茜, 杨乐, 柳钦火, 等. 森林地上生物量遥感反演方法综述[J]. 遥感学报, 2015, 19(1): 62–74.
LIU Qian, YANG Le, LIU Qinhuo, et al. Review of forest above ground biomass inversion methods based on remote sensing technology[J]. Journal of Remote Sensing, 2015, 19(1): 62–74. (in Chinese)
- [2] LU D S, CHEN Q, WANG G, et al. A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems[J]. International Journal of Digital Earth, 2014, 9(1): 63–105.
- [3] 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 of forest above ground carbon storage using remote sensing in Daxing'anling Mountains[D]. Harbin: Northeast Forestry University, 2014. (in Chinese)
- [4] 李德仁. 我国第一颗民用三线阵立体测图卫星——资源三号测绘卫星[J]. 测绘学报, 2012, 41(3): 317–322.
LI Deren. China's first civilian three-line-array stereo mapping satellite[J]. Acta Geodaetica et Cartographica Sinica, 2012, 41(3): 317–322. (in Chinese)
- [5] JIN X L, MA J H, WEN Z D, et al. Estimation of maize residue cover using Landsat-8 OLI image spectral information and textural features [J]. Remote Sensing, 2015, 7(11): 14559–14575.
- [6] MENG J H, LI S M, WANG W, et al. Estimation of forest structural diversity using the spectral and textural information derived from SPOT-5 satellite images[J]. Remote Sensing, 2016, 8(2): 125.
- [7] SARKER L R, NICHOL J E. Improved forest biomass estimates using ALOS AVNIR-2 texture indices[J]. Remote Sensing of Environment, 2011, 115(4): 968–977.
- [8] 王海宾, 彭道黎, 范应龙, 等. 基于辅助信息的森林蓄积量空间模拟[J/OL]. 农业机械学报, 2016, 47(6): 283–289.
WANG Haibin, PENG Daoli, FAN Yinglong, et al. Spatial modeling of forest stock volume based on auxiliary information[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2016, 47(6): 283–289. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20160637&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2016.06.037. (in Chinese)
- [9] 王海宾, 侯瑞萍, 郑冬梅, 等. 基于地理加权回归模型的亚热带地区乔木林生物量估算[J/OL]. 农业机械学报, 2018, 49(6): 184–190.
WANG Haibin, HOU Ruiping, ZHENG Dongmei, et al. Biomass estimation of arbor forest in subtropical region based on geographically weighted regression model[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(6): 184–190. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?flag=1&file_no=20180621&journal_id=jcsam. DOI:10.6041/j.issn.1000-1298.2018.06.021. (in Chinese)
- [10] 沈楚楚. 浙江省主要树种(组)生物量转换因子研究[D]. 杭州: 浙江农林大学, 2013.
SHEN Chuchu. The research on biomass expansion factors of the dominant tree species in Zhejiang Province[D]. Hangzhou: Zhejiang A&F University, 2013. (in Chinese)
- [11] 毛学刚, 王静文, 范文义. 基于遥感与地统计的森林生物量时空变异分析[J]. 北京林业大学学报, 2016, 38(2): 10–19.
MAO Xuegang, WANG Jingwen, FAN Wenyi. Spatial and temporal variation of forest biomass based on remote sensing and geostatistics [J]. Journal of Beijing Forestry University, 2016, 38(2): 10–19. (in Chinese)
- [12] 张煜星, 王祝雄. 遥感技术在森林资源清查中的应用研究[M]. 北京: 中国林业出版社, 2007.
- [13] BEGUET B, GUYON D, BOUKIR S, et al. Automated retrieval of forest structure variables based on multi-scale texture analysis of VHR satellite imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 96: 164–178.
- [14] 郭含茹. 基于地理加权回归的区域森林碳储量估计[D]. 杭州: 浙江农林大学, 2015.
GUO Hanru. Geographically weighted regression based estimation of regional forest carbon storage[D]. Hangzhou: Zhejiang A&F University, 2015. (in Chinese)
- [15] 向安民, 刘凤铃, 于宝义, 等. 基于 k-NN 方法和 GF 遥感影像的森林蓄积量估测[J]. 浙江农林大学学报, 2017, 34(3): 406–412.
XIANG Anmin, LIU Fengling, YU Baoyi, et al. Forest stock volume estimation based on the k-NN method and GF remote sensing data [J]. Journal of Zhejiang A&F University, 2017, 34(3): 406–412. (in Chinese)
- [16] 王政权. 地统计学及在生态学中的应用[M]. 北京: 科学出版社, 1999: 35–141.
- [17] 范磊, 程永政, 郑国清, 等. 基于 Cokriging 插值修正冬小麦面积遥感监测[J]. 农业工程学报, 2010, 26(10): 206–211.
FAN Lei, CHENG Yongzheng, ZHENG Guoqing, et al. Monitoring of winter wheat area by remote sensing based on Cokriging [J]. Transactions of the CSAE, 2010, 26(10): 206–211. (in Chinese)
- [18] 贺鹏, 张会儒, 雷相东, 等. 基于地统计学的森林地上生物量估计[J]. 林业科学, 2013, 49(5): 101–109.
HE Peng, ZHANG Huiru, LEI Xiangdong, et al. Estimation of forest above-ground biomass based on geostatistics [J]. Scientia Silvae Sinicae, 2013, 49(5): 101–109. (in Chinese)
- [19] VASTARANTA M, HOLOPAINEN M, KARJALAINEN M, et al. TerraSAR-X stereo radargrammetry and airborne scanning LiDAR height metrics in imputation of forest aboveground biomass and stem volume[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(2): 1197–1204.
- [20] GREGOIRE T G, NAESSET E, MCROBERTS R E, et al. Statistical rigor in LiDAR-assisted estimation of aboveground forest biomass[J]. Remote Sensing of Environment, 2016, 173: 98–108.
- [21] 韩爱惠. 森林生物量及碳储量遥感监测方法研究[D]. 北京: 北京林业大学, 2009.
HAN Aihui. Study on monitoring method of forest biomass and carbon storage based on remote sensing[D]. Beijing: Beijing Forestry University, 2009. (in Chinese)
- [22] 王月婷, 张晓丽, 杨慧乔, 等. 基于 Landsat8 卫星光谱与纹理信息的森林蓄积量估算[J]. 浙江农林大学学报, 2015, 32(3): 384–391.
WANG Yueting, ZHANG Xiaoli, YANG Huiqiao, et al. Forest volume estimation based on spectral and textural information from the Landsat 8 satellite [J]. Journal of Zhejiang A&F University, 2015, 32(3): 384–391. (in Chinese)