

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

孙钰森¹ 王维芳^{1*} 李国春²

(¹东北林业大学林学院, 哈尔滨 150049; ²黑龙江省林业监测规划院, 哈尔滨 150070)

摘要 森林碳储量对于全球气候变化具有重要影响, 以往的模型估算未考虑到模型残差的空间相关性和碳储量数据的非平稳性, 影响模型的预测精度. 本研究基于东北林业大学帽儿山实验林场的 ETM+遥感影像数据和 193 块固定样地, 利用地理加权克里格回归 (GWRK) 建立森林碳储量与遥感和地形因子的回归模型, 同时对比最小二乘模型 (OLS)、地理加权回归模型 (GWR) 的预测精度. 结果表明: 对于帽儿山地区的森林碳储量估算, GWRK 的平均绝对误差 (MAE)、均方根误差 (RMSE) 低于 OLS 模型和 GWR 模型, GWRK 模型的平均误差 (ME) 低于 GWR 模型, 与 OLS 模型相近. GWRK 模型的预测精度为 83.2%, 较 OLS 模型 (73.7%) 和 GWR 模型 (77.3%) 分别提高 6% 和 10%, 拟合精度明显提高, 说明 GWRK 模型是森林碳储量估算的有效方法. 利用 GWRK 模型预测的研究区森林碳储量平均值为 $70.31 \text{ t} \cdot \text{hm}^{-2}$, 在海拔较高的地区, 森林碳储量值相对较高, 说明海拔对其有较大影响.

关键词 森林碳储量; 最小二乘回归; 地理加权回归; 地理加权回归克里格; 空间异质性

Spatial distribution of forest carbon storage in Maershan region, Northeast China based on geographically weighted regression kriging model. SUN Yu-sen¹, WANG Wei-fang^{1*}, LI Guo-chun² (¹*School of Forestry, Northeast Forestry University, Harbin 150040, China*; ²*Heilongjiang Institute of Forestry Monitoring and Planning, Harbin 150070, China*).

Abstract: Forest carbon storage has important impacts on climate change. The previous models do not take into consideration of the inherent spatial correlation structure of residual and non-stationary of forest carbon storage which limits the prediction accuracy. Based on ETM+ remotely sensed imagery and 193 fixed plots of Maershan Experimental Forest Farm of Northeast Forestry University, we established the geographically weighted regression kriging (GWRK) model between forest carbon storage and extracted factors from remotely sensed imagery and topographic factors. The prediction accuracy of GWRK, ordinary least square (OLS) model and geographically weighted regression (GWR) were compared. The results showed that the mean absolute error (MAE) and root mean square error (RMSE) of GWRK were lower than those of OLS and GWR models, while the mean error (ME) of GWRK model was lower than that of GWR model and was close to that of OLS model. The prediction accuracy of GWRK model was 83.2%, which was 6% and 10% higher than that of OLS model (73.7%) and GWR model (77.3%). Therefore, the GWRK model was more effective in estimating forest carbon storage than the others. The mean value of forest carbon storage predicted by GWRK model was $70.31 \text{ t} \cdot \text{hm}^{-2}$. The relatively high values presented in high altitude area, indicating that altitude had a great impact on forest carbon storage.

Key words: forest carbon storage; ordinary least square; geographically weighted regression; geographically weighted regression Kriging; spatial heterogeneity.

森林是最大的陆地生态系统,在全球碳循环中占有重要地位^[1].作为碳收支以及碳汇交易的前提,精准估算森林碳储量已成为现如今全球气候变化与碳循环研究的重要内容^[2].

目前,森林碳储量估算方法主要有3种:样地清查法、涡度相关法和模型估算法^[3].传统的样地清查方法对于森林碳储量估算具有一定的局限性,随着科技的发展,基于遥感技术的森林碳储量估计方法是目前的主要估计方法之一.如邢素丽等^[4]利用多光谱数据对落叶松(*Larix gmelinii*)的生物量模型进行拟合研究,推动了碳储量估算的发展;李明泽^[5]采用光学遥感、多角度遥感、多源遥感等技术,对东北地区的森林生物量进行估算,并对森林生物量进行时空分析,探讨影响森林生物量变换的驱动机制;于颖等^[6]根据东北林区的标准地调查数据探究东北林区不同尺度森林的含碳率,结果表明东北林区的含碳率相对稳定,总含碳率约为0.44.然而,森林碳储量数据存在空间异质性,这类非平稳的数据导致基于传统统计方法的全局模型所得结果必然不够精确.Fotheringham^[7]设计出基于非参数建模思想的地理加权回归(geographically weighted regression, GWR)模型是一种空间异相关局部模型,通过空间权函数将空间位置参与到模型的拟合,用于处理具有非平稳性的数据.相对于全局的统计模型, GWR模型具有更小的残差项,同时 GWR模型可将研究区域内模型的回归诊断以及模型参数的空间变化进行可视化分析.2004年纽约州立大学的Zhang等^[8-9]首次将GWR模型应用于林业研究,探究树木生长与环境影响之间的关系以及树木之间的竞争变化.刘畅^[10-11]利用GWR模型根据黑龙江省固定样地数据以林分平均胸径(DBH)、每公顷株数(TPH)等林分因子对黑龙江省森林碳储量空间分布进行研究,结果表明GWR模型的拟合以及预估效果均好于全局模型.然而, GWR模型也存在一定局限性,它未考虑到其残差固有的空间相关结构^[12-13].Wasko等^[14]对模型的不同结合方式进行比较研究,利用对局部模型的全局残差进行插值,对于克服模型的残差空间自相关结果具有良好效果.地理加权回归克里格(geographically weighted regression kriging, GWRK)是GWR模型的一种拓展,将GWR与克里格法结合,利用克里格法对GWR模型的残差部分进行插值,消除了残差的空间相关对模型拟合的影响,同时该方法可以展示出被空间非平稳所掩盖的局部变异.Kumar等^[15-16]利用GWRK模型对美国中

西部地区的土壤有机质(SOC)进行估算,发现GWRK模型具有较高的拟合和预测精度,明确解决了SOC数据的空间非平稳性与空间相依的问题;Kumari等^[17]利用GWRK模型对新西兰地区的降水情况进行估算,这些研究均证实对于具有空间异质性以及空间依赖性的数据, GWRK模型具有很好的预测和拟合效果.

本研究利用Landsat-TM数据对东北林业大学帽儿山实验林场进行碳储量估算,分别建立最小二乘模型(OLS)、GWR模型、GWRK模型,对预测结果进行比较,优选出研究区域最适合的预测模型,并对森林的碳储量分布格局进行讨论.

1 研究地区与研究方法

1.1 研究区域概况

东北林业大学帽儿山实验林场(45°15'—45°29' N, 127°23'—127°43' E)位于黑龙江省尚志市,总面积为26496 hm².该区域位于长白山支脉张广才岭西北部的余脉,属于大陆性季风气候,年平均降水629 mm,年平均气温3.1℃.植被是由顶级植被阔叶红松林经人为干扰破坏后形成的较典型的东北东部天然次生林、人工林^[18].次生林类型多样且具有代表性,群落类型有硬阔叶林、软阔叶林、针叶林、针阔混交林.主要阔叶树种包括白桦(*Betula papyrifera*)、杨树(*Populus davidiana*)、柞树(*Quercus mongolica*),次要树种包括枫桦(*Betula costata*)、椴树(*Tilia amurensis*)、榆树(*Ulmus pumila*)、五角槭(*Acer mono*)、水曲柳(*Fraxinus mandshurica*)、胡桃楸(*Juglans mandshurica*)、山杨(*Populus davidiana*),还有少量的花曲柳(*Fraxinus rhynchophylla*)、黄波罗(*Phellodendron amurense*)和兴安落叶松(*Larix gmelinii*)、红松(*Pinus koraiensis*)、樟子松(*Pinus sylvestris* var. *mongolica*)等.

1.2 野外数据的收集与处理

本研究数据采用东北林业大学帽儿山实验林场2016年调查的263块森林资源连续清查固定样地数据,其中,有林地193块.样地为菱形样地,分布在1 km×1 km的网格上,面积为0.06 hm²,对样地内所有的乔木进行每木检尺,起测径阶为5 cm,记录每株树的胸径(DBH)、树种以及状态.获取样地中心点的GPS坐标,精度在5 m内.本研究采取随机抽样的方法,将193块有林地固定样地按照3:1分为两部分,其中,147块作为建模样本,46块作为检验样本.使用Dong等^[19-21]提出的东北林区主要树种林分类

型生物量模型计算样地内林木生物量,其模型的实验样本来自于帽儿山实验林场,与本研究区域吻合,且该生物量方程为可加性生物量方程,利用联立方程组对不同器官的生物量进行累加,计算精度显著提高.对于超出生物量方程的 DBH 适用范围的大树,采用材积方程与木材密度计算树干生物量方程校正误差,再利用生物量扩展因子计算单木整体生物量^[22].将每株乔木的生物量乘以相应的生物量/碳储量转换系数得到单木的碳储量,其中,生物量/碳储量转换系数来源于文献[23];将样地内每株乔木的碳储量进行累加求和,除以样地面积 0.06 hm²,得到各样地单位面积森林碳储量.

1.3 遥感数据收集与处理

遥感数据来源于地理空间数据云 (<http://www.gscloud.cn/>). 美国 Landsat-ETM+ 数据的影像为 WGS-84 椭球、UTM 投影,时间为 2016 年 7 月 21 日.本景遥感影像由于 Landsat 数据自身原因存在一定的条带,会对实验结果造成一定影响,因此,本研究利用 ESRI 公司提供的 Landsat_GapFill 插件对数据进行修复 (http://landsat7.usgs.gov/slc_enhancements/gapfilled1.php),消除条带的影响.为消除大气、地形和传感器自身成像对遥感数据造成的影响,利用 ENVI 对 TM 影像的 6 个波段(不包含第 6 波段;远红外波段)进行预处理;辐射定标将遥感数据的图像亮度值通过增益和偏移转化为辐射亮度值.通过 ENVI 的 FLAASH 模块对遥感影像进行大气校正,消除水汽对遥感影像的干扰,利用地形图对遥感影像进行几何精校正,消除地形对遥感数据造成的干扰,几何校正后误差在 1 个像元内,处理后的遥感影像如图 1 所示.地形地貌因子采用 AW3D-30 DEM 数据,该数据由日本宇宙航空研研制,并于 2014 年向其他地区提供 30 m 分辨率的新一代数字高程模型数据(<http://www.eorc.jaxa.jp/>),在此基础上提取高程、坡度、坡向等基本地形属性.

1.4 模型变量

初步筛选出 65 个自变量因子,主要包括各原始波段的灰度值及其非线性组合,包括归一化植被指数(NDVI)、垂直植被指数(PVI)、土壤调整植被指数(SAVI)等植被指数、灰度共生矩阵的平均值、对比度、熵等纹理特征(3×3)以及高程、坡度、坡向等地形因子^[24].通过相关性分析,筛选出 13 个显著性水平 $\alpha=0.01$ 的变量因子.由表 1 可知,研究区域内样地森林碳储量与 EI、NDVI、SAVI、TM452 和 TM42 呈显著正相关,即森林碳储量随着高程等因子的增



图 1 固定样地分布
Fig.1 Distribution of fixed plots.

表 1 样地碳储量与自变量的相关系数
Table 1 Correlation coefficients between the plot carbon storage and independent variables

变量 Variance		相关系数 Correlation coefficient
原始波段 Original band	b1	0.402 **
	b2	0.360 **
	b3	0.362 **
植被指数 Vegetation index	b4	0.293 **
	NDVI	0.429 **
	SAVI	0.427 **
波段组合 Band combination	PVI	0.267 *
	TM452	0.499 **
	TM42	0.481 **
高程 Elevation	EI	0.556 **
灰度共生矩阵的熵	Ent2	-0.428 **
Entropy of gray level cooccurrence matrix	Ent3	-0.271 *
灰度共生矩阵的方差	Var2	-0.261 *
Variance of gray level co-occurrence matrix	Var3	-0.321 **

TM452=TM(4+5-2)/TM(4+5+2); TM42=TM4/TM2; Ent2、Ent3 分别为第二波段、第三波段灰度共生矩阵的熵 Ent2 and Ent3 were the entropy of gray level co-occurrence matrix in the second and third bands, respectively; Var2、Var3 分别为第二波段、第三波段灰度共生矩阵的方差 Var 2 and Var 3 were variances of gray level co-occurrence matrices in the second and third bands, respectively. NDVI: 归一化植被指数 Normalized differential vegetation index; SAVI: 土壤调整植被指数 Soil-adjusted vegetation index; PVI: 垂直植被指数 Perpendicular vegetation index.

大而增大;森林碳储量与 Ent2 呈显著负相关,即随着 Ent2 的降低,森林碳储量增加.

各植被指数的公式如下:

$$NDVI = \frac{TM4 - TM3}{TM4 + TM3} \tag{1}$$

$$PVI = \sqrt{(0.335TM4 - 0.149TM3)^2 + (0.335TM3 - 0.852TM4)^2} \tag{2}$$

$$SAVI = \frac{(1 + L)(TM4 - TM3)}{TM4 + TM3 + L} \tag{3}$$

式中： L 为冠层背景调整因子，当 L 值为 0.5 时，可将土壤亮度减到最小；TM4、TM3 分别为近红外波段和红光波段。

1.5 研究方法

1.5.1 OLS 模型 OLS 是一种利用 n 组观察值求得 p 个独立变量与因变量之间残差平方和最小的拟合方法.其公式如下：

$$Y = \sum_{i=1}^p X_i \beta_i + \varepsilon \tag{4}$$

式中： Y 是样地碳含量的预测值； β_i 为未知的模型拟合参数； p 是用于建模的样地个数； ε 为模型的残差项，服从正态分布 $N(0, \sigma^2)$.通过矩阵方程(5)进行估计：

$$\hat{\beta} = (X'X)^{-1}X'Y \tag{5}$$

1.5.2 GWR 模型 GWR 模型是 OLS 模型的拓展，将地理信息考虑到建模中，是一种局部参数的估计方法.其实质是将空间距离权重加入到模型的拟合中，构建空间权重矩阵来描述局部区域的空间非平稳性，通过对每一个未知点带宽内所有的观测值进行局部线性回归获取该点的预测值，克服空间异质性对模型拟合的干扰，回归点的参数随着空间位置的变化而变化.其表达式如下：

$$Y_{GWR} = \beta_k(\mu_i, v_i) + \sum_{k=1}^p \beta_k(\mu_i, v_i) x_{ik} + \varepsilon_i \tag{6}$$

$(i = 1, 2, \cdots, n)$

式中： (μ_i, v_i) 为 i 点的坐标； $\beta_k(\mu_i, v_i)$ 为 i 点的回归参数； $\beta_0(\mu_i, v_i)$ 为 i 点的截距项； ε_i 为残差项，遵循正态分布 $N(0, \sigma^2)$.

模型参数通过下式矩阵方程进行估计：

$$\hat{\beta}_i = (X'W_iX)^{-1}X'W_iY \tag{7}$$

式中： W_i 是 i 点处的空间权重对角矩阵，表示周围参与建模的观测值与 i 点的空间权重.本研究选择 Gaussian 函数计算各点权重，表达式如下：

$$W_{ij} = \left[1 - \left(\frac{d_{ij}}{b} \right)^2 \right]^2 \tag{8}$$

式中： d_{ij} 为 i 点与其他观测值的距离，其值越大，所对应的权重越小； b 为带宽，带宽的选择是 GWR 模型的关键，本研究中选择以 AICc 作为准则，利用黄金分割搜索法，选择 AICc 值最小时的 b 值作为最佳带宽^[25]，AICc 计算公式如下：

$$AICc = 2n \ln \hat{\sigma}^2 + n \ln 2\pi + n \left\{ \frac{n + \text{tr}(S)}{n - 2 - \text{tr}(S)} \right\} \tag{9}$$

式中： n 为样本数； $\hat{\sigma}^2$ 为预测值误差项的标准误； $\text{tr}(S)$ 为带宽函数帽子矩阵的迹。

1.5.3 GWRK 模型 GWRK 是 GWR 模型的一种拓展，是对 GWR 模型拟合的残差部分进行普通克里格插值(OK)，然后将残差部分与 GWR 模型拟合的趋势项进行相加，表达式如下：

$$Y_{GWRK} = \beta_0(\mu_i, v_i) + \sum_{k=1}^p \beta_k(\mu_i, v_i) x_{ik} + \hat{\varepsilon}_{OK}(\mu_i, v_i) \tag{10}$$

式中： (μ_i, v_i) 为 i 点的坐标； $\beta_k(\mu_i, v_i)$ 为 i 点的回归参数； $\hat{\varepsilon}_{OK}(\mu_i, v_i)$ 为模型在 i 点的残差值.OK 插值属于一种无偏线性插值方法，表达式为：

$$\hat{\varepsilon}_{OK}(\mu_i, v_i) = \sum_{i=1}^n \hat{\lambda}_i \hat{\varepsilon}_i \mu_i v_i \tag{11}$$

式中： $\hat{\lambda}_i$ 是邻域观察值对 i 点观察值的贡献权重，可通过变异函数确定，表达式为：

$$y(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [z(x_i + h) - z(x_i)]^2 \tag{12}$$

式中： $y(h)$ 为变异函数； h 为步长； $N(h)$ 为 h 范围内的观测数； $z(x_i + h)$ 为 $(x_i + h)$ 处的观测值； $z(x_i)$ 为 x_i 处的观测值.图 2 是理论变异函数示意图，块金常数 (C_0) 表示为实验误差或尺度引起的变异，基台值 ($C_0 + C_1$) 表示为随机变异部分与固定变异部分之和，块基比 $[C_0 / (C_0 + C_1)]$ 是反映空间自相关程度的重要指标，其值越小，空间自相关程度越好.变程表示存在空间相关性的最大距离，当观察值之间的距离大于变程时，观察值的自相关程度消失^[26-27]。

1.6 模型评价

应用平均误差(mean error, ME)、平均绝对误差(mean absolute error, MAE)、均方根误差(root mean square error, RMSE)和预测精度(accuracy, Acc)对模型进行精度评价，计算公式如下：

$$ME = \frac{1}{n} \sum_{i=1}^n [Z^*(x_i) - Z(x_i)] \tag{13}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |Z^*(x_i) - Z(x_i)| \tag{14}$$

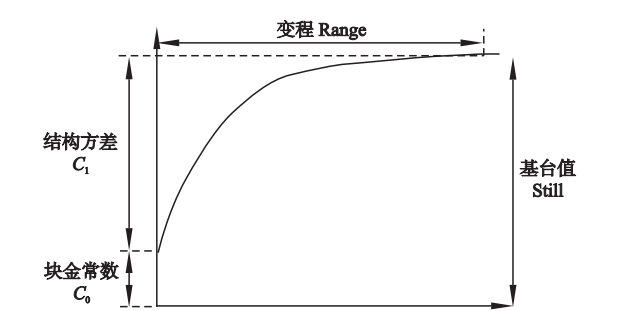


图 2 理论变异模型示意图
Fig.2 Sketch of theoretical variogram model.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n [Z^*(x_i) - Z(x_i)]^2} \tag{15}$$

$$Acc = \left[1 - \frac{1}{n} \sum_{i=1}^n \left| \frac{Z^*(x_i) - Z(x_i)}{Z(x_i)} \right| \right] \tag{16}$$

式中: $Z^*(x_i)$ 为 x_i 处的预测值; $Z(x_i)$ 为 x_i 处的观测值;ME 是衡量模型预测偏移程度的指标,该值越接近于 0,偏移程度越小;MAE 和 RMSE 是衡量模型估算精度的指标,该值越小,模型精度越高;Acc 是描述模型预测精度更直观的指标,该值越接近 1,模型预测能力越好.

2 结果与分析

2.1 森林碳储量的统计描述

由表 2 可以看出,研究区森林碳储量在 2.98 ~ 132.76 t · hm⁻²,全部数据(建模+检验)的森林碳储量平均值为 75.53 t · hm⁻²,与 2015 年黑龙江省碳储量分布数据基本符合.样地碳储量变异系数(coefficient of variation,CV)值为 32.0%,属于中度变异,偏度(Skewness)值 0.13,说明数据基本符合正态分布,其频数直方图如图 3 所示.

2.2 OLS 模型

通过逐步回归(stepwise regression)的方法进行变量选择,既能保证与森林碳储量相关性显著的自变量因子进入回归模型,又能检查模型的多重共线性.变量进入模型的显著性水平 SLE=0.1,变量保留在模型中的显著水平 SLS=0.05,最终筛选出高程值(EI)、土壤调整植被指数(SAVI)、第 2 波段灰度共生矩阵的熵(Ent2)为基础模型的变量,其统计特征如表 3 所示.对筛选出的变量进行 OLS 模型拟合,由表 4 可以看出,随着高程的增加,人类活动减小,森

林立地质量越高;SAVI 在 NDVI 的基础上,通过引入冠层背景调节因子减小土壤背景对 NDVI 的影响,其值越大,说明森林的生长程度越好;Ent2 是绿光波段灰度共生矩阵的熵,表明林木的分化程度越大,随着碳储量的增加,树木的竞争越大,森林分化程度增大.所有变量系数预测值的 P 值均小于 0.05,同时 VIF 值小于 10,不存在多重共线性,模型决定系数(R²)达到 0.42,在一定程度上可以满足统计模型的基本要求.

2.3 GWR 模型

采用高程值(EI)、土壤调整植被指数(SAVI)、第 2 波段灰度共生矩阵的熵(Ent2)作为 GWR 模型的解释变量进行参数回归.采用高斯函数时,通过 AICc 值最小的准则选取最佳带宽为 9239.87 m,每个样点平均有 53 个样地参与建模,GWR 模型中各变量的回归系数的统计如表 5 所示.GWR 模型 R²为 0.43,Ent2 的回归系数均为负值,与森林碳储量数据呈负相关关系,SAVI 和 EI 的回归系数均为正值,与森林碳储量数据呈正相关关系.同时数据的分布较为集中,这与全局模型的相关性检验结果基本一致.

2.4 GWRK 模型

利用 GS+对 GWR 模型的参数进行地统计分析,拟合其残差的理论变异函数模型为高斯模型.其决定系数达到 0.8 以上,拟合效果较好,块基比小于 25%,有效变程为 2460 m(表 6),说明在该范围内 GWR 模型的残差存在一定的空间自相关,可以进行普通克里格插值.

2.5 模型精度评价

由表 7 可以看出,GWRK 模型的平均误差(MAE)、均方根误差(RMSE)均小于 OLS 模型、

表 2 森林碳储量的描述统计
Table 2 Descriptive statistic of forest carbon storage

组别 Group	样本数 No. of samples	平均值 Mean (t · hm ⁻²)	标准差 SD (t · hm ⁻²)	变异系数 CV (%)	最小值 Minimum (t · hm ⁻²)	最大值 Maximum (t · hm ⁻²)	偏度 Skewness	峰度 Kurtosis
建模组 Modeling dataset	147	74.46	22.22	31.2	2.98	129.70	-0.19	0.33
检验组 Validation dataset	46	76.13	21.08	33.6	7.71	132.76	0.22	0.06
总数据 Total dataset	193	75.53	24.17	32.0	2.98	132.76	0.13	0.07

表 3 检验数据变量的描述统计
Table 3 Descriptive statistic of variables of validation dataset

变量 Variable	平均值 Mean	标准差 SD	变异系数 CV(%)	最小值 Minimum	最大值 Maximum	偏度 Skewness	峰度 Skewness
高程 Elevation(m)	447.91	98.08	21.2	263	727	0.50	-0.17
土壤调整植被指数 SAVI	1.19	0.04	3.4	0.90	1.23	-4.60	28.83
灰度共生矩阵的熵 Ent2	1.26	0.28	22.2	0.49	2.04	0.02	0.09

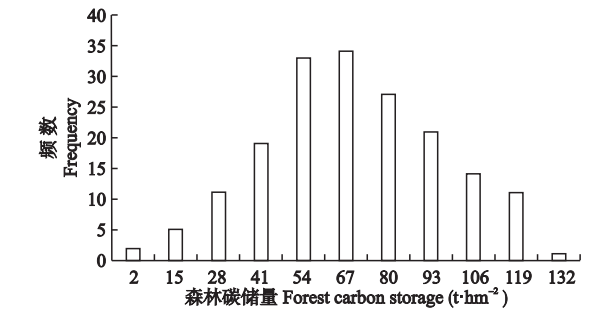


图3 碳储量频数直方图

Fig.3 Histogram of the forest carbon storage.

表4 回归模型系数与显著性及共线性检验

Table 4 Regression model coefficients, signifiacne and the collinearity test

变量 Variable	系数 Coefficient	方差膨胀 因子 VIF	P 值 P value
常数 Constant	-150.75	-	<0.01
高程 Elevation	0.10	1.14	<0.01
植被指数 SAVI	167.14	1.2	<0.01
灰度共生矩阵的熵 Ent2	-15.10	1.17	0.01

表5 GWR 模型系数估算的统计特征

Table 5 Statistics characteristics of model coefficient estimates of GWR model

变量 Variable	最小值 Minimum	第一四分 位数 Q1	中值 Median	第三分 位数 Q3	最大值 Maximum
高程 Elevation (m)	0.083	0.091	0.097	0.108	0.124
土壤调整植被指数 SAVI	174.849	190.990	211.312	244.200	309.300
灰度共生矩阵 的熵 Ent2	-19.647	-15.751	-12.164	-9.015	-4.022

表6 GWR 模型残差的变异函数模型参数

Table 6 Variogram model parameters of regression residuals of GWR

拟合模型 Fitted model	块金值 C_0	基台值 C_0+C_1	块基比 $C_0/(C_0+C_1)$ (%)	变程 Range (m)	决定系数 R^2
高斯函数 Gaussian function	0.59	5.62	10.5	2460	0.948

GWR 模型;且预测精度 (Acc) 高于其他方法,达到 83.2%.GWRK 模型的 ME 值为 0.71,低于 GWR 模型,与 OLS 模型相近,说明 GWRK 模型估算的偏移程度不弱于其他两种方法.在 3 种方法中,GWRK 模型的 MAE 和 RMSE 值最小,GWR 模型其次,OLS 模型相对较差,GWRK 的 Acc 值最大,表明 GWRK 方法的预测结果更接近真实值,预测效果最高.在本研究区域,GWRK 模型为预测效果最好的方法,3 种模型拟合残差分布图如图 4 所示.

表7 3 个模型预测能力的精度评价指标比较

Table 7 Comparison of precision evaluation of three models

模型 Model	平均误差 ME ($t \cdot hm^{-2}$)	平均绝对 误差 MAE ($t \cdot hm^{-2}$)	均方根误差 RMSE ($t \cdot hm^{-2}$)	预测精度 Acc (%)
OLS	-0.132	16.721	19.490	73.7
GWR	6.843	16.691	19.957	77.3
GWRK	1.514	15.174	16.225	83.2

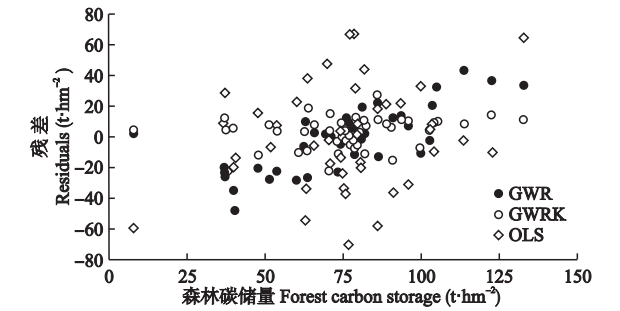


图4 3 种模型拟合帽儿山地区森林碳储量的残差分布

Fig.4 Residual distribution of forest carbon storage of Maoershan region fitted by three models.

2.6 森林碳储量空间分布格局

用 EI、SAVI、Ent2 为自变量,通过 GWRK 模型对帽儿山实验林场范围内的森林碳储量进行预测.由于仍有一些栅格数据会受到道路、农田等其他地类的影响,使碳储量值的范围不合理,本研究根据杨国亭等^[23]研究中黑龙江省最小森林碳储量值($2.52 t \cdot hm^{-2}$)对栅格数据进行数据选择,裁切受道路等其他地类影响的数据.从图 5 可以看出,研究区森林碳储量范围为 $2.25 \sim 126.25 t \cdot hm^{-2}$,平均值为 $70.31 t \cdot hm^{-2}$,高值主要集中于东部、北部和中部地区,南部地区的碳储量值相对较低,这是由于人类活动的影响造成.在南部人口密集区域,人类干扰大,森林碳储量值较低,同时随着海拔的增高,人类活动减小,森林的林分密度增加,致使森林碳储量逐渐增大,这与以往的研究结果一致^[28].

3 讨 论

本研究分别采用 OLS、GWR 和 GWRK 3 种回归模型对研究区域内森林碳储量的空间分布进行预测.OLS 模型的前提假设是变量间关系在不同位置上不变,属于全局模型,它没有考虑到空间异质性对模型拟合的影响,存在一定局限性;GWR 模型同时考虑了不同采样点的地理位置和采样点周围的影响,依靠权函数的变化,能够捕捉到局部空间信息,属于局部模型,利用该模型可以得到不同空间位置

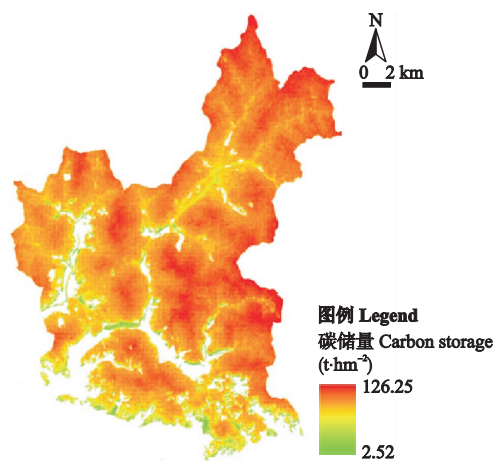


图5 GWRK 模型估算的帽儿山地区森林碳储量的空间分布
Fig.5 Spatial distribution of forest carbon storage of Maoershan region estimated by GWRK model.

上森林碳储量与影响因子间的关系,消除了空间异质性对模型拟合的影响,使模型的估算和预测能力得到提高;GWRK 模型对 GWR 模型进行了拓展,它既考虑到空间异质性对模型估算的影响,还考虑到 GWR 模型的残差固有的空间相关结构,对 GWR 模型拟合的残差部分进行普通克里格插值,然后将残差部分与 GWR 模型拟合的趋势项进行相加,提高

了精度.杨顺华等^[32] 比较了普通克里格 (OK) 与 GWRK 模型在土壤有机质空间插值精度以及制图效果方面的差异,结果表明,GWRK 的成图效果更加精细,同时土壤有机质的局部模拟效果更加精确.Kumar 等^[15-16] 运用 GWRK 模型对美国中西部地区土壤有机质进行估算和制图,同样证明了 GWRK 模型对非平稳性数据估算的优越性.本研究利用 GWRK 模型对森林碳储量进行估算,其预测精度与上述学者的研究结果一致,模型的预测精度 (Acc) 为 83.2%,优于其他 2 种方法.图 6 分别为 3 种模型估算的森林碳储量空间分布,可以看出,与其他 2 种方法相比,GWRK 模型在各个区域的值变化更加丰富,其高低值之间的过渡更加光滑,不存在明显的斑块现象,其对应的估算与地形变换更加吻合.

森林碳储量受林分立地质量、树种组成、林龄、林分密度和人为因素等多种固定效应、随机效应的影响.研究区域为国家级森林公园,近年来受人为干扰较少,森林质量得到很大改善,使得森林碳储量值较高,研究区域内森林碳储量平均值为 $70.31 t \cdot hm^{-2}$.空间分布上,在人类居住的低海拔区域,受人为干扰的影响,森林质量降低,碳储量值相应较低;随着海拔的逐渐升高,森林碳储量值逐渐增大,这是因为随着海拔的升高,人为干扰程度降低,森林

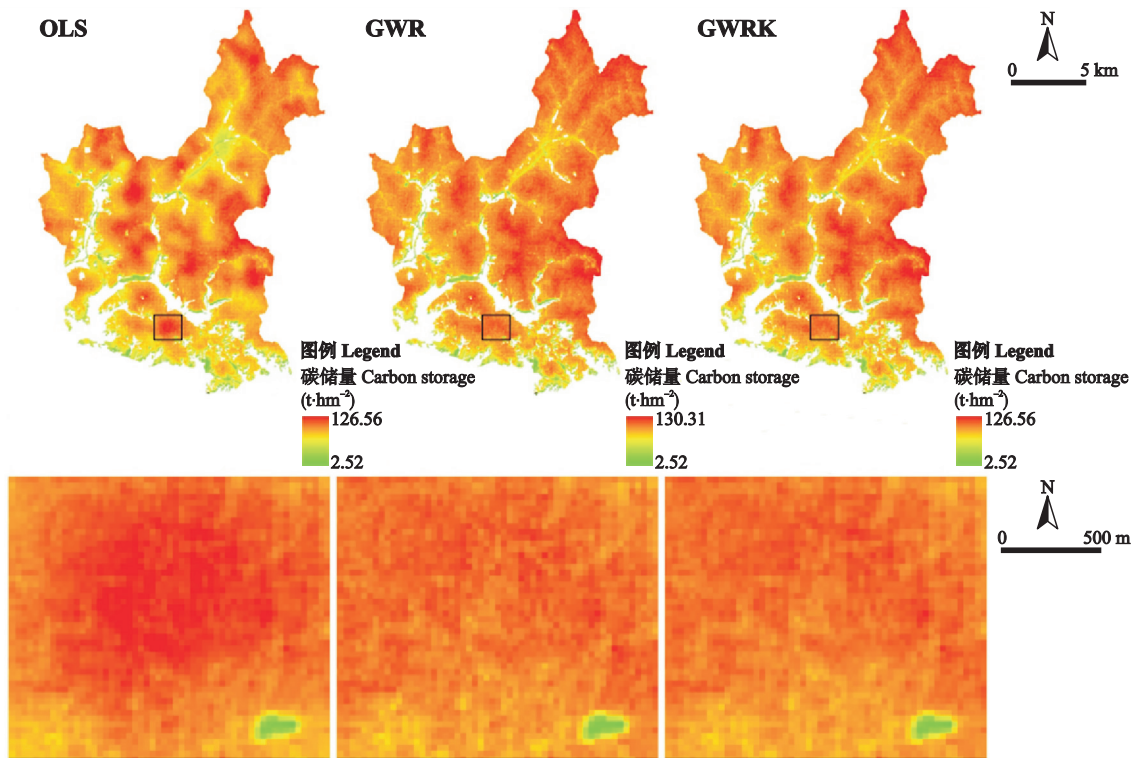


图6 3 种模型估算的帽儿山地区森林碳储量局部效果
Fig.6 Local view of forest carbon storage of Maoershan region estimated by three models.

群落结构更加完整,反映在遥感影像上,植被指数逐渐增大,这与模型的研究结果一致。随着林木竞争加剧,导致林木分化程度增大,使遥感影像纹理更加复杂,灰度共生矩阵熵值逐渐增大,这在一定程度上抑制了碳储量的增长。

本研究尝试将 GWRK 模型应用于林业方面,利用 GWRK 模型对帽儿山森林碳储量进行估算,结果表明 GWRK 模型对于森林碳储量此类非平稳性数据具有较好的估算能力,这是一次有益尝试。本研究尚未考虑到灌木层、草本层、枯落物等其他类森林碳储量,模型估算精度在一定程度上受到影响。同时光学遥感自身存在饱和性局限,可通过多源遥感数据进行精度提高。

4 结 论

本研究基于遥感数据,采用 OLS、GWR、GWRK 3 种模型对帽儿山实验林场的森林碳储量进行空间分布预测,并对比模型的预测精度和制图效果。通过相关性分析和共线性检验,选取高程、植被指数和灰度共生矩阵的熵为自变量,三者都与森林碳储量呈极显著相关关系。与其他 2 种方法对比,GWRK 模型可以对森林碳储量进行有效预测,精度比 OLS 和 GWR 模型分别提高 6% 和 10%,有效地克服了森林碳储量数据的空间非平稳性。GWRK 模型的制图效果相对其他 2 种方法的制图效果有所提高,图像的信息更加丰富,GWRK 模型的估算结果与实际情况更加吻合,效果更佳。

参考文献

- [1] Fang J, Guo Z, Hu H, *et al.* Forest biomass carbon sinks in East Asia, with special reference to the relative contributions of forest expansion and forest growth. *Global Change Biology*, 2014, **20**: 2019–2030
- [2] Li H-K (李海奎), Lei Y-C (雷渊才), Zeng W-S (曾伟生). Forest carbon storage in China estimated using forestry inventory data. *Scientia Silvae Sinicae* (林业科学), 2011, **47**(7): 7–12 (in Chinese)
- [3] Cao J-X (曹吉鑫), Tian Y (田 赞), Wang X-P (王小平), *et al.* Estimation methods of forest sequestration and their prospects. *Ecology and Environmental Sciences* (生态环境学报), 2009, **18**(5): 2001–2005 (in Chinese)
- [4] Xing S-L (邢素丽), Zhang G-L (张广录), Liu H-T (刘慧涛), *et al.* The estimating model of *Larix gmelinii* forests biomass using Landsat ETM data. *Journal of Fujian College of Forestry* (福建林学院学报), 2004(2): 153–156 (in Chinese)
- [5] Li M-Z (李明泽). Estimation and Analysis of Forest

- Biomass in Northeast Forest Region using Remote Sensing Technology. PhD Thesis. Harbin: Northeast Forestry University, 2010 (in Chinese)
- [6] Yu Y (于 颖), Fan W-Y (范文义), Li M-Z (李明泽). Forest carbon rates at different scales in Northeast China forest area. *Chinese Journal of Applied Ecology* (应用生态学报), 2012, **23**(2): 341–346 (in Chinese)
- [7] Fotheringham AS. Spatial variations in school performance: A local analysis using geographically weighted regression. *Geographical and Environmental Modelling*, 2001, **5**: 43–66
- [8] Zhang L, Bi H, Cheng P, *et al.* Modeling spatial variation in tree diameter-height relationships. *Forest Ecology and Management*, 2004, **189**: 317–329
- [9] Zhang L, Shi H. Local modeling of tree growth by geographically weighted regression. *Forest Science*, 2004, **50**: 225–244
- [10] Liu C (刘 畅). Spatial Distribution of Forest Carbon Storage in Heilongjiang Province. PhD Thesis. Harbin: Northeast Forestry University, 2014 (in Chinese)
- [11] Liu C. Spatial modeling of the carbon stock of forest trees in Heilongjiang Province, China. *Journal of Forestry Research*, 2014, **25**: 269–280
- [12] Sun W, Zhu Y, Huang S, *et al.* Mapping the mean annual precipitation of China using local interpolation techniques. *Theoretical and Applied Climatology*, 2015, **119**: 171–180
- [13] Harris P, Fotheringham AS, Crespo R, *et al.* The use of geographically weighted regression for spatial prediction: An evaluation of models using simulated data sets. *Mathematical Geosciences*, 2011, **43**: 399–399
- [14] Wasko C, Sharma A, Rasmussen P. Improved spatial prediction: A combinatorial approach. *Water Resources Research*, 2013, **49**: 3927–3935
- [15] Kumar, S. Estimating spatial distribution of soil organic carbon for the Midwestern United States using historical database. *Chemosphere*, 2015, **127**: 49–57
- [16] Kumar S, Lal R, Liu D. A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma*, 2012, **189–190**: 627–634
- [17] Kumari M, Singh C K, Basistha A, *et al.* Non-stationary modelling framework for rainfall interpolation in complex terrain. *International Journal of Climatology*, 2017, **37**: 4171–4185
- [18] Wang, C. Biomass allometric equations for 10 co-occurring tree species in Chinese temperate forests. *Forest Ecology and Management*, 2006, **222**: 9–16
- [19] Dong L, Zhang L, Li F. A three-step proportional weighting system of nonlinear biomass equations. *Forest Science*, 2015, **61**: 35–45
- [20] Dong L, Zhang L, Li F. A compatible system of biomass equations for three conifer species in Northeast, China. *Forest Ecology and Management*, 2014, **329**: 306–317
- [21] Dong L, Zhang L, Li F. Developing additive systems of biomass equations for nine hardwood species in Northeast China. *Trees*, 2015, **29**: 1–15

[22] Liu F (刘 帆), Wang C-K (王传宽), Wang X-C (王兴昌), *et al.* Spatial patterns of biomass in the temperate broadleaved deciduous forest within the fetch of the Maoershan flux tower. *Acta Ecologica Sinica* (生态学报), 2016, **36**(20): 1-14 (in Chinese)

[23] Yang G-T (杨国亭), Li F-R (李凤日), Yin T (殷彤), *et al.* Estimation and Analysis of Spatial Distribution of Forest Carbon Storage in Heilongjiang Province. Harbin: Northeast Forestry University Press, 2017 (in Chinese)

[24] Li M-Z (李明泽), Mao X-G (毛学刚), Fan W-Y (范文义). Forest biomass estimation using remote sensing based on canopy density simultaneous equations model. *Scientia Silvae Sinicae* (林业科学), 2014, **50**(2): 85-91 (in Chinese)

[25] Liu Z-X (刘正显). Research on the Spatial Distribution of Liangshui Nature Reserve Forest Biomass Based on Geographically Weighted Regression. Master Thesis. Harbin: Northeast Forestry University, 2015 (in Chinese)

[26] Christakos G. Random Field Models in Earth Sciences. New York: Academic Press, 2012

[27] Zhen Z (甄 贞), Guo Z-Y (郭志英), Zhao Y-H (赵颖慧), *et al.* Spatial distribution of soil total nitrogen in Liangshui National Nature Reserve based on local model. *Chinese Journal of Applied Ecology* (应用生态学报), 2016, **27**(2): 549-558 (in Chinese)

[28] Wang W-F (王维芳), Song L-N (宋丽楠), Sui X (隋 欣). Estimation of forest biomass and its temporal and spatial distribution patterns in Maoershan Forest Farm. *Journal of Northeast Forestry University* (东北林业大学学报), 2010, **38**(1): 47-49 (in Chinese)

[29] Réjouméchain M, Mullerlandau HC, Detto M, *et al.* Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. *Biogeosciences*, 2014, **11**: 6827-6840

[30] Ciotoli G, Voltaggio M, Tuccimei P, *et al.* Geographically weighted regression and geostatistical techniques to construct the geogenic radon potential map of the Lazio region: A methodological proposal for the European Atlas of Natural Radiation. *Journal of Environmental Radioactivity*, 2016, **166**: 355-375

[31] Song L-N (宋丽楠). Estimation of Forest Biomass and Its Temporal and Spatial Distribution Patterns Analysis in Maoershan Forest Farm. Master Thesis. Harbin: Northeast Forestry University, 2010 (in Chinese)

[32] Yang S-H (杨顺华), Zhang H-T (张海涛), Guo L (郭 龙), *et al.* Spatial interpolation of soil organic matter using regression Kriging and geographically weighted regression Kriging. *Chinese Journal of Applied Ecology* (应用生态学报), 2015, **26**(6): 1649-1656 (in Chinese)

作者简介 孙钰森,男,1996 年生,硕士研究生.主要从事林业遥感与地理信息系统研究. E-mail: 304589278@qq.com

责任编辑 杨 弘