#### DOI: 10.5846/stxb201603010352

秦立厚,张茂震,袁振花,杨海宾.基于人工神经网络与空间仿真模拟的区域森林碳估算比较——以龙泉市为例.生态学报,2017,37(10): - . Qin L H, Zhang M Z, Yuan Z H, Yang H B.Comparison of regional forest carbon estimation methods based on back-propagation neural network and spatial simulation: A case study in Longquan County.Acta Ecologica Sinica,2017,37(10): - .

# 基于人工神经网络与空间仿真模拟的区域森林碳估算 比较

——以龙泉市为例

秦立厚<sup>1,2</sup>,张茂震<sup>1,2,\*</sup>,袁振花<sup>1,2</sup>,杨海宾<sup>1,2</sup>

1 浙江农林大学 浙江省森林生态系统碳循环与固碳减排重点实验室,临安 311300
 2 浙江农林大学 环境与资源学院,临安 311300

**摘要:**森林是生态系统的重要组成部分,准确估算森林碳储量及其分布对于评价森林生态系统的功能具有重要意义。以龙泉市为研究区,利用 2009 年 99 个森林资源清查样地数据和同年度 Landsat TM 影像数据,采用高斯序列协同仿真(SGCS)与 BP 神经 网络方法(BPNN)分别模拟森林地上部分碳密度及其分布,并进行了对比分析。随机将样本数据分成 70 个建模样本和 29 个检验样本。通过模型检验,BP 神经网络预测值与实测值的相关性达到 0.67,相对均方根误差为 0.63,空间仿真方法预测值与实测值的相关性为 0.68,相对均方根误差为 0.63,空间仿真方法预测能力略高于神经网络方法。仿真结果表明,基于 BP 神经网络模拟的森林碳总量为 11042990 Mg,平均碳密度为 36.10 Mg/hm<sup>2</sup>,总体森林碳密度均值高于样地平均值 8.82%。基于空间仿真模拟的森林碳总量为 11388657 Mg,平均碳密度为 37.23 Mg/hm<sup>2</sup>,总体森林碳密度均值高于样地平均值 9.40%。对比分析可知:高斯协同仿真模拟和 BP 神经网络虽然在碳总量估算值上与抽样数据估计值相近,但两种方法在估测值的频率分布以及研究区碳分布上有较大的差异。与 BP 神经网络相比,序列高斯协同模拟结果更接近系统抽样样地实测值,全部样地预测值与实测值的相关性达到 0.75,在估计区域森林碳空间分布上有明显优势。在碳密度值域与频率分布方面,序列高斯协同模拟结果分布更合理。综上所述,序列高斯协同模拟在森林碳空间估计方面要优于 BP 神经网络。

关键词:森林碳储量;高斯协同仿真模拟;BP 神经网络;森林资源清查数据;TM 影像

# Comparison of regional forest carbon estimation methods based on backpropagation neural network and spatial simulation: A case study in Longquan County

QIN Lihou<sup>1,2</sup>, ZHANG Maozhen<sup>1,2,\*</sup>, YUAN Zhenhua<sup>1,2</sup>, YANG Haibin<sup>1,2</sup>

1 Zhejiang Provincial Key Laboratory of Carbon Cycling in Forest Ecosystems and Carbon Sequestration, Zhejiang A & F University, Lin'an 311300, China 2 School of Environmental & Resource Sciences, Zhejiang A & F University, Lin'an 311300, China

**Abstract**: Quantifying the carbon stocks of forest is critical for understanding the dynamics of carbon fluxes in terrestrial ecosystems and the atmosphere as well as monitoring ecosystem responses to environmental changes. However, due to the lack of methods and data, results of forest carbon estimation from different studies shown large difference, which presents a great uncertainty in the evaluation of forest carbon sink. Different methods can be used to estimate the carbon storage in the same study area, which can be compared with the advantages and disadvantages of each method and provides guidance for

基金项目:国家自然科学基金项目(30972360);国家自然科学基金项目(41201563);浙江农林大学农林碳汇与生态环境修复研究中心预研基金;浙江省林业碳汇与计量创新团队项目(2012R10030-01);浙江省林学一级重中之重学科学生创新计划项目资助(201515)

收稿日期:2016-00-00; 网络出版日期:2016-00-00

\* 通讯作者 Corresponding author.E-mail: zhangmaozhen@163.com

http://www.ecologica.cn

forest carbon estimation. On the basis of National Forest Inventory (NFI) data and the Land-sat TM image data collected in Longquan County, Zhejiang Province in 2009, we applied two methods, namely error back-propagation neural network (BPNN) and sequential Gaussian co-simulation (SGCS) to reproduce the distribution of above-ground forest carbon. We randomly divided plots into two sets, a 70-plot set for modeling and a 29-plot set for testing. For the model test, the correlation coefficient of predictive value and the plot data was 0.67 and 0.68 for BPNN and SGCS, respectively. Both of the two methods have the same RRMSE value (0.63). The predictive ability of SGCS was slightly higher than that of BPNN. The estimation results using BPNN showed that the sum of above-ground carbon is 11042990 Mg and the mean carbon density was 36.10 Mg/hm<sup>2</sup> which was higher than the average from the sample plots with a relative error of 8.82%. The SGCS showed that the sum of above-ground carbon was 11388657 Mg with a mean carbon density 37.23 Mg/hm<sup>2</sup> which was higher than the average from the sample plots with a relative error of 9.4%. Comparative analysis showed the carbon densities estimated using these two methods are both close to that calculated from the NFI data. However, there were some differences between the two methods with respect to the estimation of the frequency distribution and the carbon distribution in the study area. Predictive value of sample plot obtained using the SGCS method was closer to the plot data value than that obtained using the BPNN. And the correlation between predictive value and the plot data was 0.75, which proved that there were obvious advantages in estimating the spatial distribution of forest carbon. In addition, in terms of carbon density range and frequency distribution, SGCS was more reliable. This study further verifies the effectiveness of the SGSC which could provide effective methods for the estimation of regional forest carbon storage.

Key Words: forest carbon storage; sequential Gaussian co-simulation; back-propagation neural network; National Forest Inventory; TM image

森林是全球陆地生态系统中的最大有机碳库<sup>[1]</sup>,森林生态系统贮存了全球陆地植被碳库的77%,全球土 壤碳库的39%<sup>[2]</sup>,在维护区域生态环境和全球碳平衡中起着巨大的作用。准确估算区域森林碳储量及其分 布,对于评价森林生态系统的功能具有重要意义。但是,由于方法和数据的缺乏,不同学者估测的森林碳储量 相差较大,导致森林碳汇功能评价具有较大的不确定性<sup>[3]</sup>。

日益发展的遥感技术具有快速、准确、对森林无破坏性并能进行宏观监测的优势,使得遥感成为获取森林 地上生物量的主要途径。目前基于遥感数据计算森林碳储量的方法主要有:回归估计法<sup>[4-5]</sup>,神经网络 法<sup>[6-7]</sup>,遥感数据与过程模型融合的方法<sup>[8]</sup>和空间仿真模拟<sup>[9-10]</sup>等方法。与其余方法相比,神经网络与空间 仿真方法具有较高的估算精度<sup>[11-12]</sup>。人工神经网络(ANN)是通过模拟人脑神经系统建立起来的一类模型, 在建模时不需要给出具体的数学函数,可以一次性引入多个解释变量,并同时输出多个估测量,适合复杂的非 线性模型的模拟,被广泛地应用于工学、天文学、生物学等领域。近年来,神经网络也常被用来估算森林生物 量。与其他模型相比神经网络方法虽然可以提高森林的预测精度<sup>[13-14]</sup>,但是在模拟预测过程中易陷入局部 最优,从而导致了部分区域模拟结果与实际值不相符<sup>[15]</sup>。而且在模拟过程中,由于隐含层个数和单元数的确 定缺乏理论指导,学习与记忆具有不确定性<sup>[16]</sup>,虽然使用相同方法、相同数据但是却得出不同的结果,使研究 结果不能重复。

空间仿真模拟采用与地统计学方法相结合的随机算法模拟森林碳分布,通过对局部森林碳分布特征量的 分析,得到局部森林碳分布函数,再用蒙特卡洛方法实现对局部的估计<sup>[10]</sup>。它不像克里格方法,追求的是特 定点位某个属性的局部最优估值<sup>[17]</sup>,也不像回归方法那样只考虑保证总体平均数的估计精度。空间仿真模 拟追求的是森林碳储量模拟的真实性,尽可能地接近真实的空间分布<sup>[9]</sup>。由于不同的方法对于同一地区碳 储量的估算结果是不同的,因此对同一地区使用不同方法进行碳储量估算,可以对比各方法的优劣,为森林碳 估算提供指导。本文以龙泉市为研究对象,采用 BP 神经网络与空间仿真模拟方法对其地上部分碳储量和碳 分布进行仿真,并对两种方法的估算能力进行对比分析。

#### 1 研究区与数据

#### 1.1 研究区概况

龙泉市(118°43′—119°26′E,27°42′—28°21′N)地处浙江省西南部,隶属丽水市,东西长 68.9 km,南北宽 70.8 km,总面积 3059 km<sup>2</sup>。

龙泉市在地貌上属于浙南山地,地形复杂、海拔高低悬殊,因此气候基本呈垂直变化分布,光、温、水地域 差异明显。该市地处亚热带季风气候区,温暖湿润,四季分明,雨水充沛,光热较优,适宜各种林木生长,植物 资源丰富。全市共有高等植物1800余种,其中木本植物1105种(含种下分类群),占全省3/4以上。

在全国森林资源经营管理分区方案中,龙泉市属于南方山地丘陵区中的南方低山丘陵亚区,是重要的集体林区。全市林业用地面积 265633 hm<sup>2</sup>,占总面积 87.17%,森林蓄积量 1 455.9 万 m<sup>3</sup>,森林覆盖率 84.2%,乔木林年总生长量为 101.9 万 m<sup>3</sup>,生长率 8.53%(2008 年)。森林类型主要有常绿落叶阔叶混交林、针叶阔叶混交林、常绿阔叶林、黄山松林、马尾松林、杉木林、毛竹林以及山地矮林、灌丛等类型。

# 1.2 研究数据预处理

# 1.2.1 地面样地数据

研究区地面调查数据为 2009 年龙泉市森林资源连续清查样地调查数据,全市共有按系统抽样方法布设 的固定样地 102 个,样地间距 6 km×4 km,样地形状为正方形,样地面积 0.08 hm<sup>2</sup>。本研究采用的有效样地个 数为 99 个,样地数据特征值见表 1。由于现有的森林生物量模型有限,本研究将树种分为杉木、马尾松、硬阔 和软阔 4 个树种组,根据已发表的生物量模型<sup>[18-20]</sup>进行样地内单株地上部分生物量计算。如果样地含有毛 竹,毛竹生物量根据文献[21]的毛竹单株生物量模型计算。单木碳储量由单木生物量乘以碳储量转换系数 0.5<sup>[22]</sup>得到,最后累加求得样地碳储量。结合样地碳储量与遥感因子利用神经网络和空间仿真方法来估计研 究区森林地上部分碳储量及其碳密度。

Table 1 statistics for plot data							
最大值/(Mg/hm <sup>2</sup> ) Maximum	最小值/(Mg/hm <sup>2</sup> )	平均值/(Mg/hm <sup>2</sup> ) Mean	标准差/(Mg/hm <sup>2</sup> ) SD	95%置信区间 95% Confidence interval of the difference			
	Minimum			上限 Upper	下限 Lower		
119.88	0	33.30	28.11	27.77	38.84		

表 1 样地数据统计

# 1.2.2 遥感数据

遥感数据选用 2009 年龙泉市全境 Landsat TM 影像数据,由 119/40 和 119/41 两景组成。并对其进行了 几何校正和辐射校正,总误差小于一个像元。对于用于建模的遥感因子,通过 ArcGIS 提取样地所对应的遥感 图像 6 个波段灰度值、相关植被指数以及其它波段组合等遥感变量。通过分析比较各样地碳密度与对应遥感 变量之间的相关性,选取其中与碳密度相关性较大的遥感变量进行碳储量的估算。主要遥感变量与森林碳密 度之间的关系见表 2。

れる 工会通常文里 可較低度之间的個人は							
Table 2         Correlation between remote sensing factors and carbon density							
TM1         TM2         TM3         TM4         TM5							
-0.405 **	-0.453 **	-0.376 **	-0.151	-0.465 **	-0.425 **		
NDVI	MSAVI	RSR	TM5+TM7	TM4/TM5			
0.29 **	-0.451 **	0.440 **	-0.465 **	0.34 **			

表 2 主要遥感变量与碳密度之间的相关性

表中 TM1-TM7 为 TM 影像 1—7 波段, NDVI 为归一化植被指数, MSAVI 为修正土壤调整植被指数, RSR 为简单比率指数, TM5+TM7 为第 5 波段与第 7 波段的和, TM4/TM5 为第 4 波段与第 5 波段的比值。\*\* 表示在 0.01 水平上显著

# 2 研究方法

4

#### 2.1 BP 神经网络仿真建模

BP 神经网络是1种多层前馈神经网络,该网络的主要特点是信号前向传递,误差反向传递。主要由3部 分组成:输入层、隐含层、输出层。隐含层可以分为一层或多层,一个包含两层隐含层的 BP 神经网络的拓扑 结构如图1所示[23]。

BP 网络模型处理信息的基本原理是:输入信号通 过隐层节点作用于输出层节点,经过非线形变换,产生 输出信号,网络训练的每个样本包括输入向量 P 和期 望输出量 Y,网络输出值 A 与期望输出值 Y 之间的偏 差,通过调整输入节点与隐层节点的联接强度取值和隐 层节点与输出节点之间的联接强度以及阈值,使误差沿 梯度方向下降,经过反复学习训练,确定与最小误差相 对应的网络参数(权值和阈值),训练即告停止。此时 经过训练的神经网络即能对类似样本的输入信息,自行 处理输出误差最小的经过非线形转换的信息。



Fig.1 Schematic diagram of BPNN

设一个三层 BP 神经网络,输入节点  $x_i$ ;隐含节点  $y_i$ ,输出节点  $z_k$ ,输入节点与隐含层节点的网络权值为  $w_{ii}$ ,隐含层节点与输出节点的网络权值为 $w_{ii}$ ,当输出节点的期望输出为 $t_l$ 时,BP模型的计算公式如下:

隐含层节点的输出:

$$y_j = f(\sum w_{ij}x_i - \theta_j) \tag{1}$$

输出节点的计算输出为:

$$z_k = f(\sum_i T_{li} y_i - \theta_i)$$
<sup>(2)</sup>

输出节点的误差计算公式为:

$$E = \frac{1}{2} \sum_{l} \left( t_{l} - z_{k} \right)^{2} = \frac{1}{2} \sum_{l} \left( t_{l} - f\left(\sum_{j} T_{li}y_{j} - \theta_{j}\right) \right)^{2} = \frac{1}{2} \sum_{l} \left( t_{l} - f\left(\sum_{j} T_{li}f\left(\sum_{j} w_{ij}x_{i} - \theta_{j}\right) - \theta_{j}\right) \right)^{2}$$
(3)

2.2 序列高斯协同仿真

为了得到研究区碳密度分布图,采用基于地统计学的序列高斯协同仿真模拟来进行仿真。该方法通过将 研究区划分成块的方法进行仿真并假设每个单元的估计值是一个随机函数在该位置的随机变量 Z(u) 的实 现,其概率分布假定为正态分布,并由该点周围的样地数据确定。地统计学中,半方差函数用来描述随机函数 空间关系,相交的半方差函数可以度量两个随机函数相互的空间相关关系。设变量 Z 为森林碳储量,则 Z(u)为定义在二维空间 u 处的随机函数,其半方差函数  $\gamma_{zz}(h)$ 、空间协方差  $C_{zz}(h)$  通过关于距离 h 的方程 式进行计算:

$$\gamma_{ZZ}(h) = \frac{1}{2N} \sum_{\alpha=1}^{N} \left[ Z(u_{\alpha}) - Z(u_{\alpha} + h) \right]^{2}$$
(4)

$$C_{ZZ}(h) = \frac{1}{N} \sum_{\alpha=1}^{N} Z(u_{\alpha}) \times Z(u_{\alpha} + h) - m_{-h} \times m_{+h}$$
(5)

方程式中  $\alpha$  是变程范围内第  $\alpha$  个样本;相距 h 的两个样本称为一个样本对; N 为变程范围内样本对的数 量;为了区分样本对中的两个样本数据,分别称为头和尾, m-h和 m+h 分别为若干个样本对的尾和头数据的平 均值。

采用序列高斯协同仿真模拟时,需要由一个统计平均数和方差来确定特定的概率密度函数,而这个统计 平均数和方差可以用同位协同简单克里格估计来获得。

$$z^{sck}(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{sck}(u) [z(u_{\alpha}) - m_{z}] + \lambda_{y}^{sck}(u) [y(u) - m_{y}] + m_{z}$$
(6)

$$\sigma^{2(sck)}(u) = C_{zz}(0) - \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}^{sck}(u) C_{zz}(u_{\alpha} - u) - \lambda_{y}^{sck}(u) C_{zy}(0)$$
(7)

式中, u 为一个被估计的位置;  $z^{sct}(u)$  为位置 u 的森林碳估计值;  $z(u_{\alpha})$  为位置 u 局部范围内第  $\alpha$  个样地的森 林碳实测值; n(u) 为在搜索范围内所获得的样地数; y(u) 为在像元 u 处的遥感影像数据;  $m_z \ m_y$  分别为地 面样地数据和遥感影像数据的均值;  $\lambda_{\alpha}^{sct}(u)$  和  $\lambda_{y}^{sct}(u)$  分别为样地数据和影像数据的权重;  $C_{zt}(0)$  为森林碳 地面样地数据的方差;  $C_{zy}(0)$  为森林碳和遥感影像数据的协方差。当  $h = (u_{\alpha} - u)$  时,  $C_{zy}(h)$  为森林碳与遥 感影像数据的交叉协方差函数。

由平均数和方差确定的密度函数 f[z(u)] 可以用下式表示:

$$f[z(u)] = \frac{1}{\sigma^{(sck)} \sqrt{2\pi}} e^{-\frac{[Z(u)-\mu^{sck}]^2}{2\sigma^{2(sck)}}}$$
(8)

概率密度函数积分得到其条件累积分布,并假定这个分布符合正态分布,该分布中随机抽取一个数作为 待估位置 u 处的模拟实现。本文选取与碳储量相关性最高的 TM5 参与仿真模拟。在仿真过程中,样本半方 差可以用下式模拟:

$$\mathbf{r}^{\rm sph}(h) = \mathbf{c}_0 + c_1 \left[ 1.5 \times \frac{h}{a} - 0.5 \times \left(\frac{h}{a}\right)^3 \right]$$
(9)

其中,  $c_0$  为块金值,  $c_1$  为结构参数, a 为变程, h 为距离。

2.3 模型精度验证

为了检验两种模型的精度,本研究选取 70 个样地数据(建模样本)用来建模,剩余 29 个样地数据(检验 样本)用来检验模型精度,两种方法两组数据保持一致。两组数据的统计见表 3。

	Table 3 I	Descriptive statistics for d	ata modeling and data	testing		
数据集	样地个数	最大值/(Mg/hm <sup>2</sup> )	最小值/(Mg/hm <sup>2</sup> )	平均值/(Mg/hm <sup>2</sup> )	标准差/(Mg/hm <sup>2</sup> )	
Dataset	Simple number	Maximum	Minimum	Mean	SD	
建模样本 Data Modeling	70	119.88	0	36.37	28.48	
检验样本 Data Testing	29	98.36	0	25.80	20.04	

#### 表 3 建模数据和检验数据的基本统计

模型确定后,以决定系数(*R*<sup>2</sup>)、均方根误差(RMSE)以及相对均方根误差(RRMSE)对模型拟合精度进行 评价。

# 3 结果与分析

# 3.1 模型的确定

# 3.1.1 BP 神经网络建模

Funahashi<sup>[24]</sup>指出单隐含层 BP 神经网络模型能够以任意精度逼近任意函数。鉴于本研究输入输出因子 相对简单,选取单隐含层进行模型构建。利用 matlab2010b 神经网络工具箱及相关程序,以样地所在位置对 应的 6 个波段信息及相关性较高的 MSAVI、RSR、TM5+TM7 经归一化处理后作为输入层,通过多次训练并调 整神经网络目标误差以及隐含层神经元个数,从而选取较为理想的模型来估算研究区森林碳储量,最后对神 经网络输出结果进行反归一化得到森林碳储量的预测值。用前面提到的检验样本数据集检验样地位置的估 计值与实测值的吻合程度。计算可知,检验样本预测值与实测值的相关性为 0.67,决定系数(*R*<sup>2</sup>)为 0.45,均 方根误差(RMSE)为 20.42 Mg/hm<sup>2</sup>,相对均方根误差(RRMSE)为 0.63。其中建模数据与检验数据的预测值 基本统计见表 4。

10 期

农业 建铁奶油和过滤奶油以肉但盔伞儿们						
Table 4 Descriptive statistics for predicted value of data modeling and data testing						
数据集 Dataset	Table 4 Descriptive statistics for predicted value of data modeling and data testing         最大值/(Mg/hm <sup>2</sup> )       最小值/(Mg/hm <sup>2</sup> )       平均值/(Mg/hm <sup>2</sup> )       标准差/(Mg/hm <sup>2</sup> )         Maximum       Minimum       Mean       SD         值 Predicted value of data modeling       87.89       3.66       36.03       28.48         值 Predicted value of data testing       101.65       5.80       32.28       21.20					
建模样本预测值 Predicted value of data modeling	87.89	3.66	36.03	28.48		
检验样本预测值 Predicted value of data testing	101.65	5.80	32.28	21.20		

# 表 4 建模数据和检验数据预测值基本统计

3.1.2 半方差函数拟合

在进行半方差函数拟合时,通过反复调节块金值、基台值以及变程3个参数来选择较好的组合。当块金值为0.35,基台值为0.65,变程为7560时达到最优值。此时半方差函数为:

$$r^{\text{sph}}(h) = 0.35 + 0.65 \left[ 1.5 \times \frac{h}{7560} - 0.5 \times \left(\frac{h}{7560}\right)^3 \right]$$

式中,*h*为距离, r<sup>sph</sup>(*h*)为标准半方差。仿真完成后,提取检验样本所对应的像元值对仿真结果进行精度验证。计算可知,预测值与实测值的相关性为 0.68,决定系数(*R*<sup>2</sup>)为 0.47,均方根误差(RMSE)为 20.04 Mg/hm<sup>2</sup>,相对均方根误差(RRMSE)0.63。建模数据与检验数据的预测值基本统计见表 5。

表 5 建模数据和检验数据预测值基本统计

Table 5 Descriptive statistics for data modeling and data testing

数据集 Dataset	最大值/(Mg/hm <sup>2</sup> ) Maximum	最小值/(Mg/hm <sup>2</sup> ) Minimum	平均值/(Mg/hm <sup>2</sup> ) Mean	标准差/(Mg/hm <sup>2</sup> ) SD
建模样本预测值 Predicted value of data modeling	97.45	5.83	36.46	21.41
检验样本预测值 Predicted value of data testing	88.27	5.14	31.77	21.59

## 3.2 模拟结果

从统计特征来看,基于 BP 神经网络模拟的森林平均碳密度为 36.10 Mg/hm<sup>2</sup>,碳密度最大值为 130.92 Mg/hm<sup>2</sup>,最小值为-107.33 Mg/hm<sup>2</sup>,碳总量为 11042990 Mg,标准差为 20.09 Mg/hm<sup>2</sup>。基于空间仿真模拟的森 林平均碳密度为 37.23 Mg/hm<sup>2</sup>,碳密度最大值为 116.78 Mg/hm<sup>2</sup>,最小值为 0.18 Mg/hm<sup>2</sup>,碳总量为 11388657 Mg,标准差为 19.35 Mg/hm<sup>2</sup>。

在森林碳分布方面,由于神经网络方法估测结果含有部分负值,与实际值不相符,而且由于负值的存在使 得森林碳分布表现不明显。因此在制作森林碳分布图时,本研究将小于0的值赋值为0。图2为两种方法估 测的碳密度分布图。从碳分布来看,基于空间仿真模拟的森林碳储量较高的区域主要分布在研究区南部,西 北部以及北部。基于神经网络模拟的森林碳储量较高的区域主要分布在研究区东南部和西北部,整体上呈现 出南北高中部低的趋势。与神经网络方法相比,基于空间仿真方法估算的森林碳储量在分布上相对分散,未 表现出明显的南北多中间少的趋势。两者相差较大的地区在东南部,基于神经网络估测的碳密度要高于空间 仿真方法。

#### 3.3 结果分析

为了对比两种方法的估测精度,可将两种方法的估测结果与抽样统计估计结果对比,结果见表 6。表 6显示,两种方法估测的碳总量要高于系统抽样估计结果,但神经网络方法估测的碳总量更接近抽样估计结果。与抽样估计结果相比,基于空间仿真方法估计的碳总量要高 9.40%;基于 BP 神经网络估计的碳总量要高 8.82%。两种方法相比,BP 神经网络方法估算的平均碳密度要低 1.13 Mg/hm<sup>2</sup>,最大碳密度高 14.41 Mg/hm<sup>2</sup>,最小碳密度低 107.51 Mg/hm<sup>2</sup>。两种方法最小值相差较大是因为神经网络方法在建设用地、水体等区域碳密度估算存在过度拟合的现象,使得估测值为负值。

判断估测精度的另一种方法,是将估计结果与样地实测值进行比较。将全部样地实测数据与碳分布图叠 加得到森林碳密度对比图(图3)。从图3可以看出,两种方法均可以在一定程度体现森林碳密度的空间分布

6

10 期



格局。但神经网络方法在东南部森林碳密度估测值要高于样地实测值,与实测样地差距较大。这可能是由神 经网络算法本身的固有缺陷造成,本研究采用 BP 算法来训练神经网络,该算法在预测过程中易陷入局部最 优.从而导致了部分区域模拟结果偏高<sup>[15]</sup>。

1 ab 6 Comparison between the two estimates and plot data						
项目 Item	空间仿真 SGCS	神经网络 BPNN	抽样估计 Sampling estimation	仿真模拟与 抽样估计之差 Difference between SGCS and Sampling estimation	神经网络与 抽样估计之差 Difference between BPNN and Sampling estimation	
碳总量 Total/(Mg)	11388657	11042990	10409777	978880	633213	
均值 Mean/( Mg/hm <sup>2</sup> )	37.23	36.10	33.30	3.20	2.07	
最大值 Maximum/(Mg/hm <sup>2</sup> )	116.78	130.92	119.88	-3.10	11.04	
最小值 Minimum/(Mg/hm <sup>2</sup> )	0.18	-107.33	0	0.18	-107.33	

表 6 2 种估计结果与样地数据统计结果比较

SCCS:高斯序列协同仿真 Sequential Gaussian co-simulation; BPNN: BP 神经网络 Error back-propagation neural network

为了更准确地体现两种方法与样地实测值的一致性,提取全部样地所在像元的预测值,通过对比分析从 数值上来体现两种方法与实测值的差异。从两种方法估计结果中分别提取与地面样地位置对应像元的碳密 度估计结果,以地面样地数据为真值,分别就两种方法的估计结果进行比较。图4为两种方法碳密度估测值 与实测值的关系图。图中显示,基于空间仿真估测的结果与样地实测数据的拟合程度较好,相关性达到0.75, R<sup>2</sup>= 0.56。而基于神经网络估测的结果与样地实测数据的相关性为 0.63, R<sup>2</sup>= 0.39。图 5 为两种方法森林碳 密度预测误差曲线图,显示空间仿真方法的估算结果与实际值更相符,最小绝对差值为0.07 Mg/hm<sup>2</sup>,最大绝 对差值为 52.78 Mg/hm<sup>2</sup>,均方根误差为 14.23 Mg/hm<sup>2</sup>。基于神经网络方法的最小绝对差值为 0.29 Mg/hm<sup>2</sup>, 最大绝对差值为 63.01 Mg/hm<sup>2</sup>,均方根误差为 22.59 Mg/hm<sup>2</sup>。由此可见,空间仿真方法对于样地位置的估算 结果与实测值更接近,更能表现森林碳密度的分布。

以上分析仅对样地实测数据与预测数据进行了对比分析,不能从全部像元值上体现两种方法估测结果的

7



图 3 基于神经网络方法(左)和空间仿真方法(右)模拟森林碳密度与样地实测值比较 Fig.3 Forest carbon density from SGCS and BPNN compared with the plot data





差异。表6显示,BP 神经网络含有部分负值,但碳总量却与空间仿真方法相差不多。这可能与两种方法估测 值的频率分布有关,频率分布曲线和累积频率分布曲线在揭示各值频率分布上具有明显优势。图6和图7分 别为两种方法预测值的频率分布曲线和累积频率分布曲线。由图6和图7可见,两种方法频率较高的值都集 中在20—40 Mg/hm<sup>2</sup>的值域范围内,累积频率都为40%。在此值域范围内神经网络平均值为29.43 Mg/hm<sup>2</sup>, 空间仿真方平均值为29.73 Mg/hm<sup>2</sup>。与空间仿真方法相比,BP 神经网络在40—70 Mg/hm<sup>2</sup>之间有一个明显 的突起,在此值域范围内累积频率百分比为31.5%,均值为53.13 Mg/hm<sup>2</sup>,而空间仿真方法频率百分比为 34.5%,均值为51.09 Mg/hm<sup>2</sup>,BP 神经网络在此值域的频率分布低于空间仿真方法但均值高于空间仿真方 法。两种方法在大于70 Mg/hm<sup>2</sup>的值相对较少,累积频率百分比分别为5.88%和7.2%,神经网络方法也低于 空间仿真方法。由此可见,神经网络方法在估算森林碳密度时,低值的估计要多于空间仿真方法,再加上神经 网络方法含有部分负值(所占比例为2.23%),使得神经网络方法值算得的碳总量要低于空间仿真方法,也使 得神经网络估计的碳总量更接近抽样估计值。但空间仿真方法值域分布更为合理,极值与样地数据更接近,

9



图 6 神经网络方法估测值频率分布图(左)与累积频率图分布图(右) Fig.6 Frequency and cumulative frequency of BPNN

为了体现两种方法在碳密度分布上的差异,将两种方法估测结果进行差值运算,图 8 为空间仿真结果减 去 BP 神经网络模拟结果得到的碳密度差值图,该图由两种方法对应像元位置的森林碳密度估计值相减得 到。由图 8 可知,空间仿真方法在森林碳密度较高的区域要高于神经网络方法,在研究区东南部却低于神经 网络方法。通过前面分析可知,空间仿真方法与样地实测值更接近,更能表示研究区真实的碳分布。若以空 间仿真方法为标准,说明 BP 神经网络方法低估了森林碳密度较高区域的值,高估了部分森林碳密度较低的 值。在像元尺度上对差值图进行统计分析可得,两种方法估测结果最大差值为 89.68 Mg/hm<sup>2</sup>,最小差值为-74.39 Mg/hm<sup>2</sup>,均值为1.25 Mg/hm<sup>2</sup>,标准差 18.09。说明两种方法总体上差距不大,但正负方向上的极值差距 较大。由差值频率分布图和累积频率分布图(图 9)可知,两种方法估算结果的差值基本以 0 为对称轴呈左右 分布,值域在-40—40 Mg/hm<sup>2</sup>的范围内,而那些差值较大的点,只是个别像元的影响。

# 4 结论与讨论

本研究基于 BP 神经网络方法和空间仿真模拟法对龙泉市森林碳储量及碳分布进行了估算,并对两种方法进行了对比分析。通过对比分析可知空间仿真模拟和神经网络方法在碳总量上与抽样数据估计结果相差





不多,BP 神经网络方法估算的森林碳总量更接近抽样 估计结果,但在值域分布上空间仿真方法更合理。两种 方法在森林碳分布上相差较大,空间仿真方法与实测值 更接近,更能反映真实的碳分布。这主要是因为空间仿 真模拟是基于图像的地统计条件模拟技术,以区域化变 量理论为基础,通过量化随机函数的空间关系,得到其 条件累积分布函数进而取得区域碳密度。它追求的是 森林碳储量分布的真实性,尽可能的接近碳分布的真实 情况。而 BP 神经网络在进行网络训练时追求的是误 差最小,并没有考虑到变量间的空间关系,在反应局部 特征上可能会有一定的差异。

利用神经网络模型模拟森林碳储量分布已有相关 学者进行了研究,但是,目前的研究内容主要是生物量 建模<sup>[6,25]</sup>,关于碳分布的研究区域主要为林场、湿地等 小区域<sup>[26-28]</sup>,对于县级以上碳分布还少有研究<sup>[15]</sup>。小 区域森林碳估算,由于研究区用地类型单一,植被类型 简单,测结果较为精确,如翟晓江等<sup>[29]</sup>利用神经网络方





法对黄龙山林区森林生物量进行可估算预测精度达到 87.49%,相关性达到 0.738;李丹丹等<sup>[26]</sup>利用 BP 神经 网络建立了旺业甸林场森林生物量模型仿真检验结果的平均相对误差为 15.7%,相对系数达 0.8022。大区域 森林碳估算时,土地覆盖与土地利用类型的多样性以及植被的复杂性会对估算精度有一定的影响,没有小区 域森林碳估算精度高,汪少华等<sup>[15]</sup>利用 BP 神经网络方法估算了临安市森林碳储量,检验样本与实测值得相 关性为 0.61,决定系数为 0.37;陈蜀蓉等<sup>[30]</sup>利用 Erf-BP 神经网络对缙云县公益林森林生物量进行了估算,预 测数据的决定系数为 0.513。此外,Cutler 等<sup>[31]</sup>利用多源遥感数据估算了 3 个不同研究区的森林地上部分生 物量。研究表明,只使用单一研究区数据进行生物量估算,预测值与实测值的相关性在 0.79 以上。若将 3 个研究区数据结合使用,相关性仅为 0.55,因此不同研究区间的差异也会对生物量估测精度产生一定的影响。

与空间仿真方法相比,神经网络方法仿真得出的森林碳密度较高的区域较多,部分地区碳密度预测值小于0。这可能由2方面的原因造成:第一,神经网络本身的缺陷,使用误差反向传播算法来训练神经网络,容





易使训练结果陷入局部最优,从而导致了部分区域模拟结果与实际值不相符。第二,本文所使用的有效样地 数仅为 99 个,研究指出如果森林生物量建模野外实测数据的分布和代表性不足,对应用神经网络估算森林碳 密度的估算精度有一定影响<sup>[32]</sup>,再加上神经网络外延性较差<sup>[33]</sup>,因此对于缺少样地信息的河流等区域的预 测能力较差,使预测值与实际值差别较大。

#### 参考文献(References):

- [1] Dixon R K, Solomon A M, Brown S, Houghton R A, Trexler M C, Wisniewski J. Carbon pools and flux of global forest ecosystems. Science, 1994, 263(5144): 185-190.
- Bousquet P, Peylin P, Ciais P, Le Quéré C, Friedlingstein P, Tans P P. Regional changes in carbon dioxide fluxes of land and oceans since 1980.
   Science, 2000, 290(5495): 1342-1347.
- [3] Tans P P, Fung I Y, Takahashi T. Observational constraints on the global atmospheric CO<sub>2</sub> budget. Science, 1990, 247(4949): 1431-1438.
- [4] 蒋蕊竹, 李秀启, 朱永安, 张治国. 基于 MODIS 黄河三角洲湿地 NPP 与 NDVI 相关性的时空变化特征. 生态学报, 2011, 31(22): 6708-6716.
- [5] 王长委, 胡月明, 沈德才, 黄胜利, 朱剑云, 王璐. 基于 CBERS 数据的亚热带森林地上碳储量估算. 林业科学, 2014, 50(1): 88-96.
- [6] 王淑君, 管东生. 神经网络模型森林生物量遥感估测方法的研究. 生态环境, 2007, 16(1): 108-111.
- [7] Foody G M, Cutler M E, McMorrow J, Pelz D, Tangki H, Boyd D S, Douglas I. Mapping the biomass of Bornean tropical rain forest from remotely sensed data. Global Ecology and Biogeography, 2001, 10(4): 379-387.
- [8] 耿君, 阮宏华, 涂丽丽, 吴国训. 基于 CASA 模型的瓦屋山林场植被净初级生产力估算. 林业科技开发, 2012, 26(3): 90-96.
- [9] Wang G X, Oyana T, Zhang M Z, Adu-Prah S, Zeng S Q, Lin H, Se J Y. Mapping and spatial uncertainty analysis of forest vegetation carbon by combining national forest inventory data and satellite images. Forest Ecology and Management, 2009, 258(7): 1275-1283.
- [10] 张茂震, 王广兴, 葛宏立, 徐丽华. 基于空间仿真的仙居县森林碳分布估算. 林业科学, 2014, 50(11): 13-22.
- [11] 张超, 彭道黎. 基于 PCA-RBF 神经网络的森林碳储量遥感反演模型研究. 中国农业大学学报, 2012, 17(4): 148-153.
- [12] 沈希,张茂震,祁祥斌.基于回归与随机模拟的区域森林碳分布估计方法比较.林业科学,2011,47(6):1-8.
- [13] Vahedi A A. Artificial neural network application in comparison with modeling allometric equations for predicting above-ground biomass in the Hyrcanian mixed-beech forests of Iran. Biomass and Bioenergy, 2016, 88: 66-76.
- [14] 范文义,张海玉,于颖,毛学刚,杨金明.三种森林生物量估测模型的比较分析.植物生态学报,2011,35(4):402-410.
- [15] 汪少华,张茂震,赵平安,陈金星.基于 TM 影像、森林资源清查数据和人工神经网络的森林碳空间分布模拟.生态学报,2011,31(4): 998-1008.
- [16] 宰松梅,郭冬冬,韩启彪,温季.基于人工神经网络理论的土壤水分预测研究.中国农学通报,2011,27(8):280-283.
- [17] 史舟,李艳,程街亮.水稻土重金属空间分布的随机模拟和不确定评价.环境科学,2007,28(1):209-214.

- [18] 国家林业局森林资源管理司. LY/T 2263—2014 立木生物量模型及碳计量参数—马尾松. 北京:中国标准出版社, 2014.
- [19] 国家林业局森林资源管理司. LY/T 2264—2014 立木生物量模型及碳计量参数—杉木. 北京:中国标准出版社, 2014.
- [20] 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: 3-3.
- [21] 陈辉, 洪伟, 兰斌, 郑郁善, 何东进. 闽北毛竹生物量与生产力的研究. 林业科学, 1998, 34(S1): 60-64.
- [22] 王效科, 冯宗炜, 欧阳志云. 中国森林生态系统的植物碳储量和碳密度研究. 应用生态学报, 2001, 12(1): 13-16.
- [23] 闻新,李新,张兴旺.应用 MATLAB 实现神经网络.北京:国防工业出版社,2015:95-105.
- [24] Funahash K I. On the approximate realization of continuous mappings by neural networks. Neural Networks, 1989, 2(3): 183-192.
- [25] 王轶夫, 孙玉军, 郭孝玉. 基于 BP 神经网络的马尾松立木生物量模型研究. 北京林业大学学报, 2013, 35(2): 17-21.
- [26] 李丹丹, 冯仲科, 汪笑安, 张凝, 张巍巍. BP 神经网络反演森林生物量模型研究. 林业调查规划, 2013, 38(1): 5-8.
- [27] 王立海,邢艳秋.基于人工神经网络的天然林生物量遥感估测.应用生态学报,2008,19(2):261-266.
- [28] Ingram J C, Dawson T P, Whittaker R J. Mapping tropical forest structure in southeastern Madagascar using remote sensing and artificial neural networks. Remote Sensing of Environment, 2005, 94(4): 491-507.
- [29] 翟晓江,郝红科,麻坤,李鹏,杨延征.基于TM的陕北黄龙山森林生物量模型.西北林学院学报,2014,29(1):41-45.
- [30]陈蜀蓉,张超,郑超超,张伟,伊力塔,余树全.公益林生物量估算方法研究——以浙江省缙云县公益林为例.浙江林业科技,2015,35 (5):20-28.
- [31] Cutler M E J, Boyd D S, Foody G M, Vetrivel A. Estimating tropical forest biomass with a combination of SAR image texture and Landsat TM data: an assessment of predictions between regions. ISPRS Journal of Photogrammetry and Remote Sensing, 2012, 70: 66-77.
- [32] 许等平,李晖,智长贵,韩爱惠.基于 CEBERS-WFI 遥感数据的森林生物量估测方法研究.林业资源管理,2010,(3):104-109.
- [33] 陈文烯. 基于遥感数据的森林生物量测定理论与方法. 亚热带水土保持, 2013, 25(2): 41-43, 55-55.