

基于 GF-2 的油松人工林地上生物量反演

苟睿坤^{1,2} 陈佳琦¹ 段高辉¹ 杨瑞¹ 卜元坤¹ 赵君³ 赵鹏祥^{1*}

(¹西北农林科技大学林学院, 陕西杨凌 712100; ²清华大学地球系统科学系地球系统数值模拟教育部重点实验室, 北京 100084; ³中国科学院生态环境研究中心城市与区域生态国家重点实验室, 北京 100085)

摘要 油松是黄土高原地区重要的造林树种.快速准确地估测其地上生物量,对开展该地区森林资源动态监测等具有重要作用.本研究选取陕西省黄龙山林区石堡林场的油松人工林为对象,结合国产卫星高分二号(GF-2)的多光谱遥感影像与野外同时段实测样地数据,对其地上生物量进行了估算.提取了5种植被指数和8种纹理信息,基于普通回归、逐步回归、岭回归、拉索回归与主成分回归5种方法在4种纹理窗口(3×3、5×5、7×7和9×9)下建模,使用留一法交叉验证测试了每个模型的估算精度.结果表明:提取的遥感因子之间存在着较为严重的多重共线性关系,大部分遥感因子与油松人工林地上生物量有较为显著的相关性;GF-2数据在石堡林场油松人工林地上生物量的反演中可以实现较高精度,其中估算效果最好的是使用了9×9纹理窗口的主成分回归模型,估算效果最差的是使用了3×3纹理窗口的普通回归模型.利用国产高分分辨率卫星影像对油松人工林地上生物量进行反演研究,可以为西北地区林业部门进行森林生物量监测、资源管理与可持续经营提供科学依据.

关键词 高分二号;油松人工林;地上生物量;留一法交叉验证;回归建模

Inversion of aboveground biomass of *Pinus tabuliformis* plantations based on GF-2 data. GOU Rui-kun^{1,2}, CHEN Jia-qi¹, DUAN Gao-hui¹, YANG Rui¹, BU Yuan-kun¹, ZHAO Jun³, ZHAO Peng-xiang^{1*} (¹College of Forestry, Northwest A&F University, Yangling 712100, Shaanxi, China; ²Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China; ³State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China).

Abstract: *Pinus tabuliformis* is an important afforestation species in the Loess Plateau. Quick and accurate estimation of aboveground biomass (AGB) of *P. tabuliformis* plantations plays an important role in monitoring regional forest resources. Here, we used multi-spectral remote sensing data of domestic satellite GF-2 and the field data to estimate the aboveground biomass of *P. tabuliformis* plantations in Shibao forest farm of Huanglong Mountain in Shaanxi Province. We calculated eight texture features and five vegetation indices, and then built models based four texture windows (3×3, 5×5, 7×7, 9×9) by using five regression methods including normal regression, stepwise regression, ridge regression, Lasso regression and principal component regression. We used the leave-one-out cross validation (LOOCV) to test the estimation accuracy of each model. We found serious multi-collinearity relationships between the extracted remote sensing factors. Most of the remote sensing factors had significant correlations with aboveground biomass of *P. tabuliformis* plantations. GF-2 data could achieve higher accuracy in the inversion of aboveground biomass of *P. tabuliformis* plantations in the Shibao forest farm. The best estimation result was the principal component regression model using 9×9 texture window, and the worst one was the normal regression model using 3×3 texture window. Inversion of aboveground biomass of *P. tabuliformis* plantation using domestic high-resolution satellite imagery could provide a scientific basis for forestry biomass monitoring,

本文由国家重点研发计划项目(2016YFD060020305)和国家自然科学基金项目(41801181)资助 This work was supported by the National Key R&D Program of China (2016YFD060020305) and the National Natural Science Foundation of China (41801181).

2019-07-26 Received, 2019-09-27 Accepted.

*通讯作者 Corresponding author. E-mail: zhaopengxiang@nwsuaf.edu.cn

resource management, and sustainable management in the forestry departments of northwest China.

Key words: GF-2; *Pinus tabuliformis* plantation; aboveground biomass; leave-one-out cross validation; regression modeling.

人类活动增加了大气中 CO₂、CH₄ 含量,从而导致全球温室效应日益加剧,而森林生态系统在减缓全球温室效应方面具有举足轻重的地位^[1-3].黄土高原地区是世界上水土流失最严重、生态环境最脆弱的地区之一.在国家实施退耕还林和天然林保护工程以来,该地区的森林面积不断增加,现有森林以人工林为主^[4-5].油松(*Pinus tabuliformis*)因其利用种子繁殖、萌发能力强、耐寒且耐旱等特性,具有较强应对气候变化的能力,被用作黄土高原地区主要造林树种之一,对该地区森林生态系统的构建和维护发挥着重要作用^[6].对黄土高原地区油松人工林开展实时准确的监测,不仅对评估该地区油松人工林碳储功能具有很好的促进作用,也对针对黄土高原这一生态脆弱地区开展森林健康评价和可持续发展潜力的研究具有重要意义^[7-8].

在此背景下,如何快速精确地估算油松人工林生物量是近些年来黄土高原地区森林资源动态监测的热点和难点.在估算森林生物量方面,除了采用传统的森林资源调查手段以外^[9-10],遥感估测法以其快速、高效等特性成为获取不同尺度生物量信息的主要方法^[11].目前,在进行区域森林生物量的反演时,采用的数据类型主要包括光学遥感、微波遥感和激光雷达等数据^[12].反演生物量的模型可大致归纳为过程模型、经验模型和非参数估计模型3类^[13-17].反演的过程一般是通过建立植被指数与地面实测数据等因子间的反演模型,推广到大范围的数据模拟,虽然使用的影像种类不同,但反演的精度误差均控制在40%以内^[18-24].目前,估算油松等针叶林的生物量研究主要集中在通过精度高但耗费巨大的传统测量方法^[25],利用较低分辨率遥感影像进行生物量反演^[26-27],或是仅使用单一建模方法进行生物量估算^[28].因此,油松的地上生物量反演需要基于更精细分辨率的光学遥感数据,结合多种建模方法进行比较,从而提升地上生物量的估算精度.

本研究以高分二号(GF-2)国产高分辨率遥感影像为主要数据源,结合同时段的外业样地调查数据,提取基于二阶概率统计的8种纹理特征,包括均值、方差、协同性、对比度、相异性、信息熵、二阶矩和相关性,并计算归一化植被指数、绿通道植被指数、增强植被指数、差值植被指数和土壤调节植被指数

5种植被指数.基于R软件,使用5种回归建模方法,建立油松人工林地上生物量与遥感因子间的模型,实现对陕西省黄龙山林区石堡林场的油松人工林地上生物量反演.并且,运用“留一法”评价不同纹理窗口下不同种回归模型的油松人工林地上生物量反演精度,定量分析回归模型的适用性,为黄土高原地区人工油松林的生物量监测、资源管理与可持续经营提供科学依据.

1 研究地区与研究方法

1.1 研究区概况

固定样地布设于陕西省五大林区之一的黄龙山林区石堡林场(图1).石堡林场始建于1958年,位于陕西省延安市黄龙县中部,黄龙山主脉的西南端.东连虎沟门、大岭、圪台林场,南接界头庙林场,西与三岔林场毗邻,北和官庄、蔡家川林场接壤.地理坐标为35°33'—35°46' N, 109°47'—109°58' E,东西宽10 km,南北长18 km,总面积20111 hm².该地区乔木树种有22科29属100种,主要森林类型有油松林、辽东栎(*Quercus liaotungensis*)林和白桦(*Betula platyphylla*)林等,天然植被覆盖率高,素有黄河流域绿洲之称.

1.2 数据来源及预处理

使用的GF-2遥感数据获取时间为2017年8月8日.该数据由蓝(450~520 nm)、绿(520~590 nm)、红(630~690 nm)、近红外(760~890 nm)4个多光谱波段(4 m)和1个全色波段(1 m)组成,当日云量接近0,可以准确地提取各类植被指数及纹理信息.

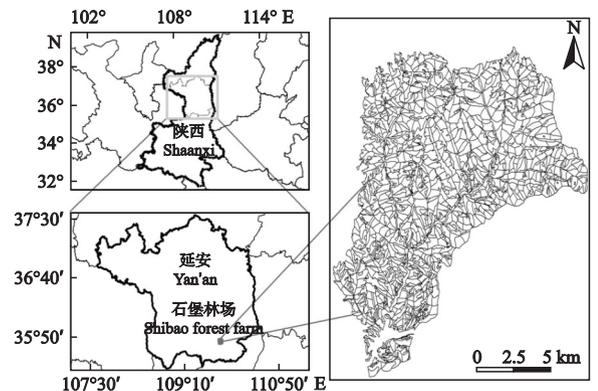


图1 石堡林场位置及示意图

Fig.1 Location and sketch of Shibao forest farm.

表 1 样地坐标及地上生物量

Table 1 Coordinate and aboveground biomass for sample site

样地 Sample site	经纬度 Latitude and longitude	地上生物量 Aboveground biomass (t · hm ⁻²)	样地 Sample site	经纬度 Latitude and longitude	地上生物量 Aboveground biomass (t · hm ⁻²)
1	35°41'15" N, 109°49'20" E	128.94	11	35°41'22" N, 109°49'18" E	156.51
2	35°42'44" N, 109°49'52" E	214.34	12	35°41'12" N, 109°49'34" E	178.79
3	35°42'59" N, 109°50'12" E	141.01	13	35°41'23" N, 109°49'33" E	170.04
4	35°43'06" N, 109°50'11" E	159.14	14	35°41'24" N, 109°49'32" E	178.87
5	35°43'05" N, 109°50'11" E	158.62	15	35°42'42" N, 109°49'38" E	116.39
6	35°42'44" N, 109°49'53" E	202.37	16	35°42'40" N, 109°49'38" E	122.08
7	35°42'42" N, 109°49'54" E	184.19	17	35°42'47" N, 109°49'39" E	189.69
8	35°43'12" N, 109°49'60" E	145.77	18	35°41'09" N, 109°49'08" E	160.40
9	35°43'11" N, 109°49'58" E	124.56	19	35°40'52" N, 109°49'28" E	152.58
10	35°43'14" N, 109°50'01" E	141.96	20	35°42'19" N, 109°49'24" E	168.88

对 GF-2 数据进行正射校正、辐射定标和大气校正,并在 1 : 10000 地形图上选取 50 个地面控制点对影像进行配准和几何精校正,使得位置误差小于 0.97 个像元,以上操作在 ENVI 5.3 软件中完成。

2017 年 7—8 月,在油松人工林分布区均匀选取了不同龄组具有相似立地条件的油松人工林,布设 20 块 20 m×30 m 固定样地(表 1),调查样地内乔木的胸径和冠幅等因子,并按胸径阶调查了部分树高。利用黄龙山二元材积表测算样地内每株油松的材积,汇总后得到各个样地的总蓄积量。再根据方精云等^[29]的油松蓄积量-生物量转换方程及地上-地下生物量之比,最终计算出各个人工油松林的地上生物量:

$$B = aV + b$$

式中: B 为单位面积的森林生物量(t · hm⁻²); V 为单位面积的森林蓄积量(m³ · hm⁻²)。针对油松林,参数 a 、 b 分别是 0.7554 和 5.0928,地上与地下生物量的比值为 4.12。

1.3 地上生物量反演模型构建

1.3.1 遥感因子提取 植被指数是指能够反映植被生长状况的数值,而纹理特征可以反映遥感影像中的细节与结构,并且可以有效地解决植被指数饱和的现象。选取各样地窗口内的像元平均值作为遥感提取的特征,得到 GF-2 影像的植被指数及纹理特征,计算公式见表 2 和表 3。在计算纹理时选择了 3×3、5×5、7×7 和 9×9 窗口来分别计算油松人工林纹理特征值。

1.3.2 因子相关性分析 为了分析不同种遥感因子与地上生物量之间的相互作用关系,进而对比在多重共线性影响下不同种回归模型的反演精度,基于 R 软件对 GF-2 影像提取出的 13 种遥感因子与油松

人工林地上生物量进行 Pearson 相关性分析,构建相关系数矩阵,分析 AGB 与多种遥感因子之间的相关关系。遥感因子及其对应变量见表 4。

1.3.3 基于普通回归的地上生物量反演 普通回归,也称全因子回归,利用的是普通最小二乘法(ordinary least square, OLS),OLS 的原理是通过将误差的平方和最小化,通过求导使其导数为零,求解得到最优的参数,找到数据匹配的最优函数,从而使回归模型中所有观察值的残差平方和最小^[30-31]。对多变量进行普通回归的建模为:

$$y = a_1x_1 + a_2x_2 + a_3x_3 + \dots + a_nx_n + b$$

式中: a 为系数; x 、 y 为变量; b 为偏置。基于前文提

表 2 植被指数计算公式

Table 2 Equations for vegetation index

植被指数 Vegetation index	计算公式 Calculation formula
归一化植被指数 NDVI	$NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$
增强植被指数 EVI	$EVI = G \frac{(\rho_{NIR} - \rho_{RED})}{(\rho_{NIR} + C_1\rho_{NIR} - C_2\rho_{BLUE} + C_3)}$
绿通道植被指数 GNDVI	$GNDVI = \frac{\rho_{NIR} - \rho_{GREEN}}{\rho_{NIR} + \rho_{GREEN}}$
差值植被指数 DVI	$DVI = \rho_{NIR} - \rho_{RED}$
土壤调节植被指数 SAVI	$SAVI = \frac{(1+L)(\rho_{NIR} - \rho_{RED})}{(\rho_{NIR} + \rho_{RED} + L)}$

NDVI: Normalized difference vegetation index; EVI: Enhanced vegetation index; GNDVI: Green normalized difference vegetation index; DVI: Difference vegetation index; SAVI: Soil-adjusted vegetation index. 下同 The same below. ρ_{NIR} 、 ρ_{RED} 、 ρ_{GREEN} 和 ρ_{BLUE} 分别为 GF-2 影像近红外、红、绿、蓝波段的反射率 ρ_{NIR} 、 ρ_{RED} 、 ρ_{GREEN} and ρ_{BLUE} were the reflectances of the near-infrared, red, green and blue bands of the GF-2 image, respectively; EVI 计算中系数 $C_1 = 6.0$, $C_2 = 7.5$, $C_3 = 1.0$, $G = 2.5$ The coefficients in the EVI calculation were $C_1 = 6.0$, $C_2 = 7.5$, $C_3 = 1.0$, $G = 2.5$; SAVI 计算中 L 为土壤调节系数,一般取 0.5 In the SAVI calculation, L was the soil adjustment coefficient, generally taken as 0.5.

表3 基于二阶概率统计的纹理特征

Table 3 Texture feature based on co-occurrence measures

纹理特征 Texture feature	计算公式 Calculation formula
均值 Mean	$\sum_{ij=0}^{N-1} iP_{ij}$
方差 Variance	$\sum_{ij=0}^{N-1} P_{ij} \times (i - \text{mean})^2$
熵 Entropy	$\sum_{ij=0}^{N-1} P_{ij} \times (-\ln P_{ij})$
二阶矩 Second moment	$\sum_{ij=0}^{N-1} P_{ij}^2$
相关性 Correlation	$\sum_{ij=0}^{N-1} P_{ij} \left[\frac{(i - \text{mean}) - (j - \text{mean})}{\sqrt{\text{variance}_i \times \text{variance}_j}} \right]$
均匀性 Homogeneity	$\sum_{ij=0}^{N-1} i \frac{P_{ij}}{1 + (i - j)^2}$
对比度 Contrast	$\sum_{ij=0}^{N-1} iP_{ij} \times (i - j)^2$
异质性 Dissimilarity	$\sum_{ij=0}^{N-1} iP_{ij} \times i - j $

表4 遥感因子及对应变量

Table 4 Remote sensing factors and corresponding variables

遥感因子 Remote sensing factor	对应变量 Corresponding variable
均值 Mean	x_1
二阶矩 Second moment	x_2
对比度 Contrast	x_3
相关性 Correlation	x_4
异质性 Dissimilarity	x_5
熵 Entropy	x_6
均匀性 Homogeneity	x_7
方差 Variance	x_8
土壤调节植被指数 SAVI	x_9
归一化植被指数 NDVI	x_{10}
绿通道植被指数 GNDVI	x_{11}
增强植被指数 EVI	x_{12}
差值植被指数 DVI	x_{13}

取的5种植被信息与分别从 3×3 、 5×5 、 7×7 和 9×9 窗口下提取的8种纹理特征作为自变量,油松人工林地上实测生物量作为因变量,运用OLS构建普通回归模型。

1.3.4 基于逐步回归的地上生物量反演 逐步回归的原理和方法是将因子依次代入模型,在代入解释因子后,执行 F 检验并对所选因子逐个完成 t 检验。如果之前引入的解释因子由于之后引入的因子而变得不再重要,那么就将它剔除,以保证在代入每个新的因子之前,只有重要的因子包含在回归方程

中^[32-33]。选取向后逐步筛选变量的回归模型来确定最准确预测油松人工林地上生物量的遥感因子组合方程。输入的自变量同样包括前文提取的13类遥感因子,因变量为油松人工林地上实测生物量,输出内容包括一系列不同遥感因子组成的多元线性回归方程、相关的判定系数及 F 检验值等。建模方程可表示为:

$$Y = b_1X_1 + b_2X_2 + \dots + b_nX_n + c$$

式中: Y 为估算得到的油松人工林地上生物量; X_i 为第 i 个预测的遥感因子变量; b_i 为第 i 个变量的回归系数; c 为回归方程中的常数项。

1.3.5 基于岭回归的地上生物量反演 岭回归属于一类有偏估计,是补充改进后的最小二乘法。它虽然在无偏性方面有损失,但其数值获得了较高的稳定性和精确度。存在一个矩阵 X ,若其中某一个数值的改动,能够引发最终计算结果的巨大误差,则该矩阵 X 就被称作“病态矩阵”。对于一个线性回归模型,对 β 这一回归参数的岭估计公式可表示为^[34]:

$$\beta(k) = (X'X + kl)^{-1}X'y$$

首先对矩阵 X 和 y 进行标准化, k 为岭参数($k > 0$), $X'X$ 为自变量样本的相关阵,表示回归参数 β 标准化的岭回归估计,应该比普通最小二乘法估计更准确、稳定。但是,因为 k 值并不是唯一确定值,所以得到的标准化岭回归估计也仅是回归参数 β 存在的一个估计值。因此,可以取不同的 k 值进行运算试验,以分析回归参数 β 众多估计值的稳定性。

在构建岭回归模型之前,首先构建一个惩罚函数来解决遥感因子间可能存在的共线性而导致的参数估计误差过大的问题;在模型中输入的自变量为提取出的13种遥感因子,因变量为油松人工林地上实测生物量;针对回归参数 β 岭估计公式中不同的 k 值,得到岭迹图;最终根据岭迹图,选择合适的岭回归模型中应该删除的遥感因子变量。

1.3.6 基于拉索回归的地上生物量反演 拉索回归的方法和原理与岭回归较为相似,同样也是设置了惩罚函数来压缩回归模型中的系数值,即强制使一部分回归系数缩小,甚至部分绝对值比较小的系数会缩小为0,惩罚函数通常用回归模型的绝对值函数来代表。在拉索回归模型中输入的自变量 X_i 包括前文提取的13种遥感因子,因变量 Y_i 为第 i 个样地的油松人工林地上实测生物量。拉索回归的函数表达式为^[35]:

$$Y_i = \partial_i + \sum_{j=1}^n \beta_j X_i + \varepsilon_i \quad [\varepsilon_i \sim N(0, \sigma^2)]$$

利用 Lasso 回归, 在 OLS 方法的基础上代入惩罚函数, 来压缩上述回归系数 α 和 β , c 为调整参数, 可根据 Robert Tibshirani 的广义验证法进行选择^[35], 公式如下:

$$(\hat{\alpha}, \hat{\beta}) = \operatorname{argmin} \left\{ \sum_i (Y_i - \alpha_i - \sum_j \beta_j X_{ij})^2 \right\}$$

$$\text{s.t. } \sum_j |\beta_j| \leq c$$

1.3.7 基于主成分回归的地上生物量反演 主成分回归是结合从遥感因子提取出的主成分自变量与因变量油松人工林地上生物量来构建回归模型, 从而使建立的模型更加简洁, 避免自变量之间出现多重共线性关系^[36-38].

首先计算输入自变量数据的协方差与相关系数矩阵, 然后得到相关系数矩阵的特征向量与特征值, 其中 PC1 的系数是最大特征值所对应的特征向量, PC2 的系数是第二大特征值所对应的特征向量, 输入模型的主成分因子数量取决于每个主成分对因变量的贡献率和累计贡献率。一般认为, 若前 i 个主成分的累计贡献率达到 80%, 那么前 i 个主成分就能够代替所有输入的自变量, 最终利用前 i 个累计贡献率达到 80% 的主成分建立多元线性回归模型, 公式如下:

$$Y' = c_1 X_1' + c_2 X_2' + \dots + c_n X_n' + d$$

式中: Y' 为油松人工林地上生物量; X_i' 为第 i 个主成分值。

选取标准分数法对全部样本数据进行标准化, 以避免不同的输入数据具有不同的量纲 (数量单位), 影响最终模型预测的结果。

2 结果与分析

2.1 因子相关性分析结果

将相关系数矩阵通过热力图的方式直观展示出来。通过颜色种类反映二维矩阵中不同因子间相关性正负, 颜色深浅表示相关性大小。由图 2 可以看出, 9 个因子与油松人工林地上生物量 (AGB) 呈显著相关 ($P < 0.01$)。在 13 个遥感因子中, 有 3 个因子与 AGB 呈极低相关 ($|r| < 0.19$), 2 个因子与 AGB 呈低度相关 ($0.20 < |r| < 0.39$), 4 个因子与 AGB 呈中度相关 ($0.40 < |r| < 0.69$), 4 个因子与 AGB 呈高度相关 ($0.70 < |r| < 0.89$)。

在 5 种植被指数因子中, 归一化植被指数与 AGB 呈显著正相关, 相关性最高; 在 8 种纹理特征因子中, 异质性与 AGB 呈显著正相关, 相关性最高。针对石堡林场提取 GF-2 影像的植被指数因子与

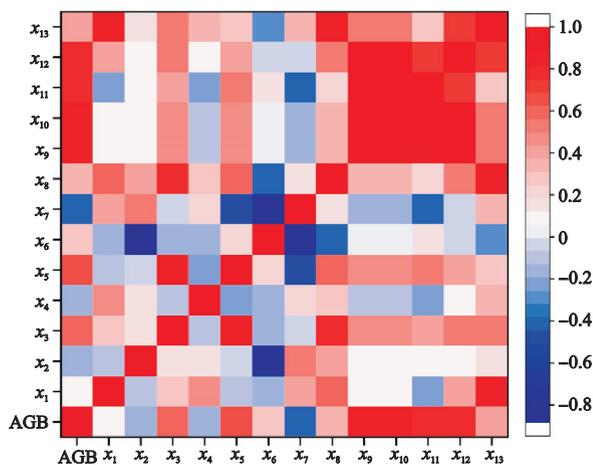


图 2 相关系数矩阵热图

Fig. 2 Heat map of correlation coefficient matrix.

AGB: 地上生物量 Aboveground biomass. 下同 The same below.

AGB 的相关性要强于纹理特征因子与 AGB 的相关性。此外, 从图中颜色较深的部分可以看出, 遥感因子之间存在着较严重的多重共线性问题。

2.2 地上生物量反演回归建模结果

利用 R 软件, 将提取的 13 类遥感因子作为自变量, 油松人工林地上生物量作为因变量, 分别构建不同纹理窗口下的普通回归模型、逐步回归模型、岭回归模型、拉索回归模型和主成分回归模型, 模型预测结果见表 5 和图 3~7。

由图 3~7 可以看出, 对于普通回归模型, 预测地上生物量与实测地上生物量最接近的纹理设置为 7×7 窗口, 预测效果最差的纹理设置为 3×3 窗口; 对于逐步回归模型, 预测地上生物量与实测地上生物量最接近的纹理设置为 9×9 窗口, 预测效果最差的纹理设置为 3×3 窗口; 对于岭回归模型, 预测明显比普通回归模型和逐步回归模型的效果更好, 其中, 预测地上生物量与实测地上生物量最接近的纹理设置为 9×9 窗口, 预测效果最差的纹理设置为 3×3 窗口; 对于拉索回归模型, 预测地上生物量与实测地上

表 5 不同回归模型下不同纹理窗口的 RMSE

Table 5 RMSE of different regression models with different texture windows

不同纹理窗口	普通回归模型	逐步回归模型	岭回归模型	拉索回归模型	主成分回归模型
Different texture window	Normal regression model	Stepwise regression model	Ridge regression model	Lasso regression model	Principal component regression model
3×3	0.95	0.86	0.53	0.81	0.49
5×5	0.80	0.74	0.48	0.46	0.45
7×7	0.72	0.69	0.47	0.49	0.44
9×9	0.76	0.65	0.45	0.46	0.43

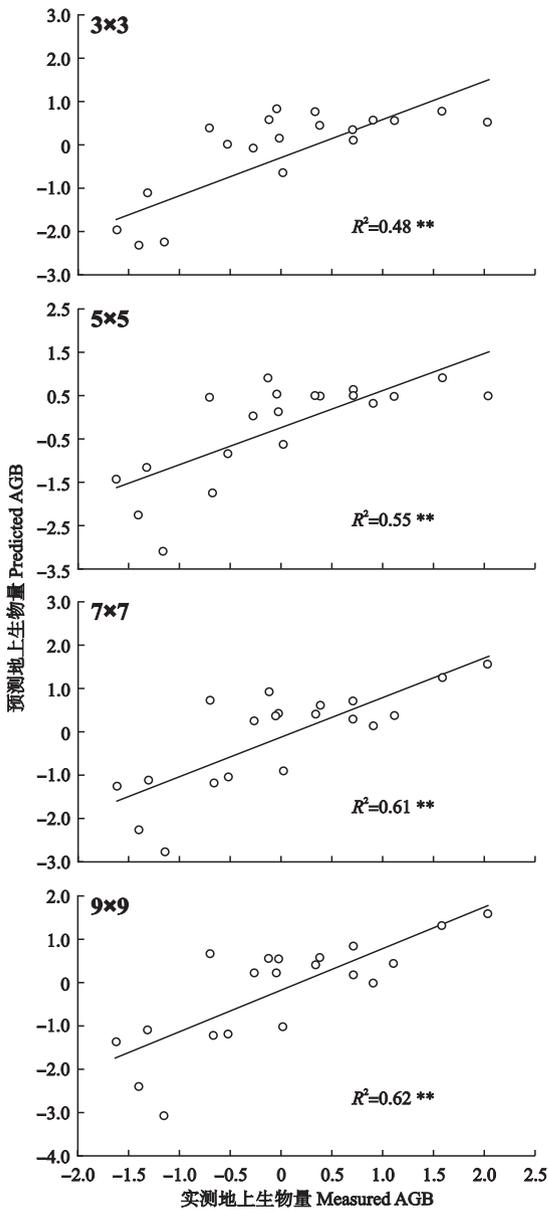


图3 普通回归模型预测与实测生物量(标准化值)散点图
Fig.3 Scatter diagram of measured AGB and predicted AGB (nomalization) of ordinary regression model.

* $P < 0.05$; ** $P < 0.01$. 下同 The same below.

生物量最接近的纹理设置为 9×9 和 5×5 窗口, 预测效果最差的纹理设置为 3×3 窗口; 对于主成分回归模型, 预测地上生物量与实测地上生物量最接近的纹理设置为 9×9 窗口, 预测效果最差的纹理设置为 3×3 窗口。

由表5和表6可以看出, 不管是对 3×3 、 5×5 、 7×7 , 还是 9×9 窗口, 主成分回归模型在5类不同的回归模型中, 预测油松人工林地上生物量的能力最好, 反演精度最高; 在4种不同的窗口中, 纹理设置为 9×9 窗口的预测精度最高。

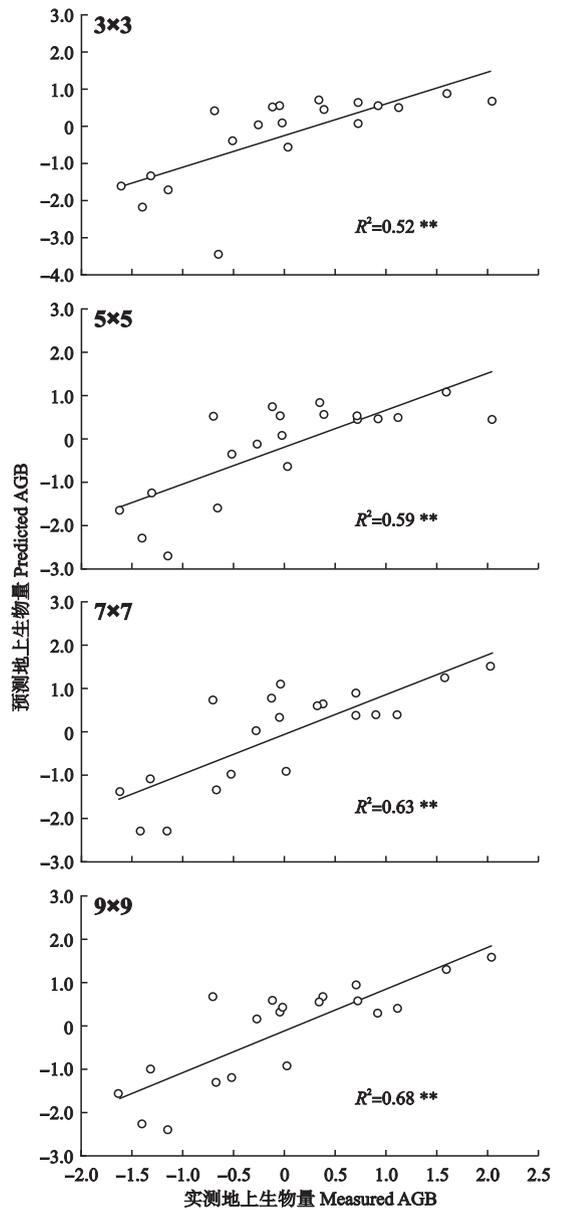


图4 逐步回归模型预测与实测生物量(标准化值)散点图
Fig.4 Scatter diagram of measured AGB and predicted AGB (nomalization) of stepwise regression model.

2.3 研究区油松人工林地上生物量反演

经过综合比较分析, 在陕西省黄龙山石堡林场

表6 不同回归模型下不同窗口的 R^2

Table 6 R^2 of different regression models with different texture windows

不同纹理窗口 Different texture window	普通回归模型 Normal regression model	逐步回归模型 Stepwise regression model	岭回归模型 Ridge regression model	拉索回归模型 Lasso regression model	主成分回归模型 Principal component regression model
3×3	0.48	0.52	0.71	0.59	0.75
5×5	0.55	0.59	0.76	0.78	0.79
7×7	0.61	0.63	0.77	0.76	0.80
9×9	0.62	0.68	0.79	0.79	0.81

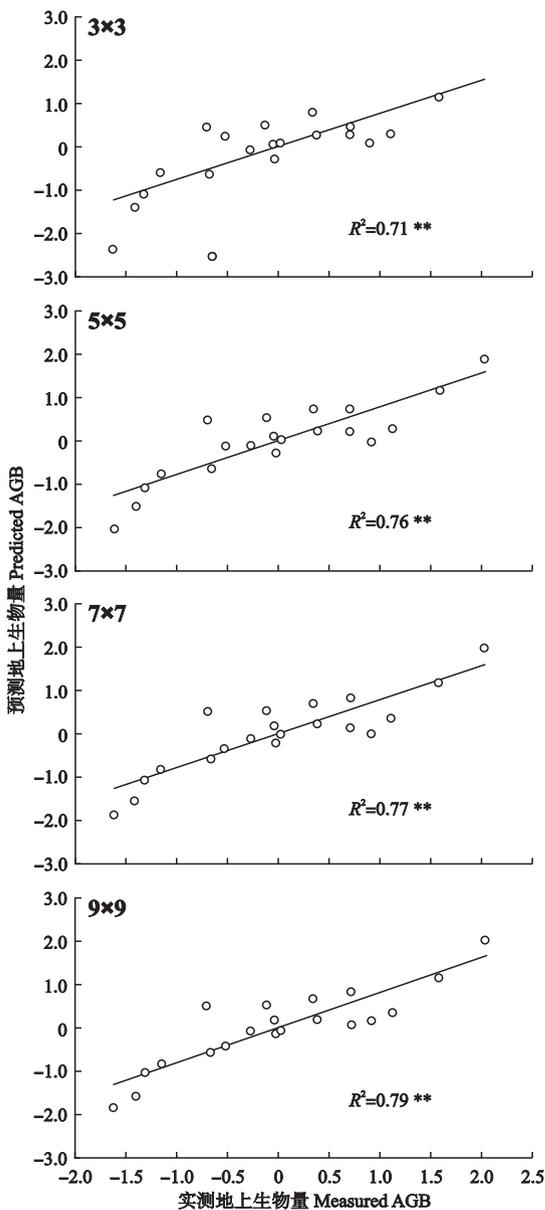


图5 岭回归模型预测与实测生物量(标准化值)散点图
Fig.5 Scatter diagram of measured AGB and predicted AGB (normalization) of ridge regression model.

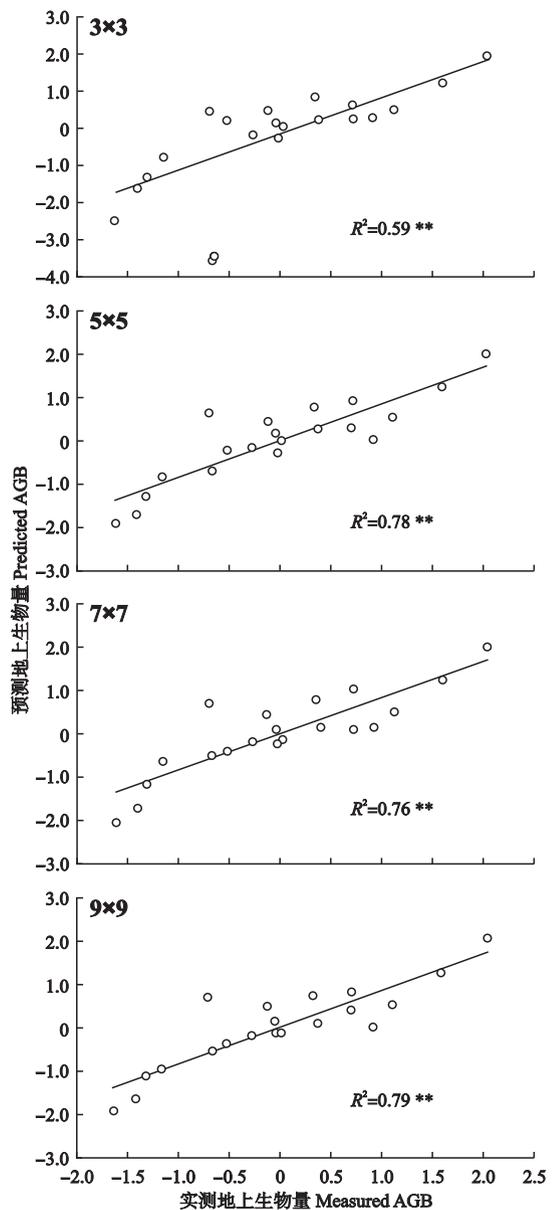


图6 拉索回归模型预测与实测生物量(标准化值)散点图
Fig.6 Scatter diagram of measured AGB and predicted AGB (normalization) of Lasso regression model.

油松人工林地上生物量的估测中,运用主成分回归建模并将纹理设置为 9×9 窗口时的精度最高,反演效果最好.因此,首先根据陕西省森林资源二类调查数据提取出石堡林场范围内油松人工林主要分布区;然后基于 9×9 纹理窗口的主成分回归模型,提取每个林分区域对应的植被指数和纹理特征因子作为该模型的输入因子,估测出石堡林场范围内的油松人工林地上生物量;最终利用 ArcGIS 10.3 平台制作了该研究区油松人工林地上生物量专题图(图8).

$$Y' = 0.37X_1' - 0.18X_2' + 0.01X_3' + 0.02X_4'$$

式中: X_1' 、 X_2' 、 X_3' 和 X_4' 分别为入选主成分回归建模的主成分 PC1、PC2、PC3 和 PC4.

3 讨论

本研究选取陕西省黄龙山林区石堡林场作为研究区,结合国产卫星 GF-2 的多光谱遥感影像与野外同时段实测样地数据,对油松人工林地上生物量反演进行了研究.结论如下: 1) 对于 GF-2 影像衍生出的遥感因子,9 个因子与油松人工林地上生物量 (AGB) 呈显著相关 ($P < 0.01$), 在 5 种植被指数因子中,归一化植被指数与 AGB 呈显著正相关,相关性最高;在 8 种纹理特征因子中,异质性与 AGB 呈显著正相关,相关性最高.同时,针对石堡林场这一研究区提取 GF-2 影像的植被指数因子与 AGB 的相关性

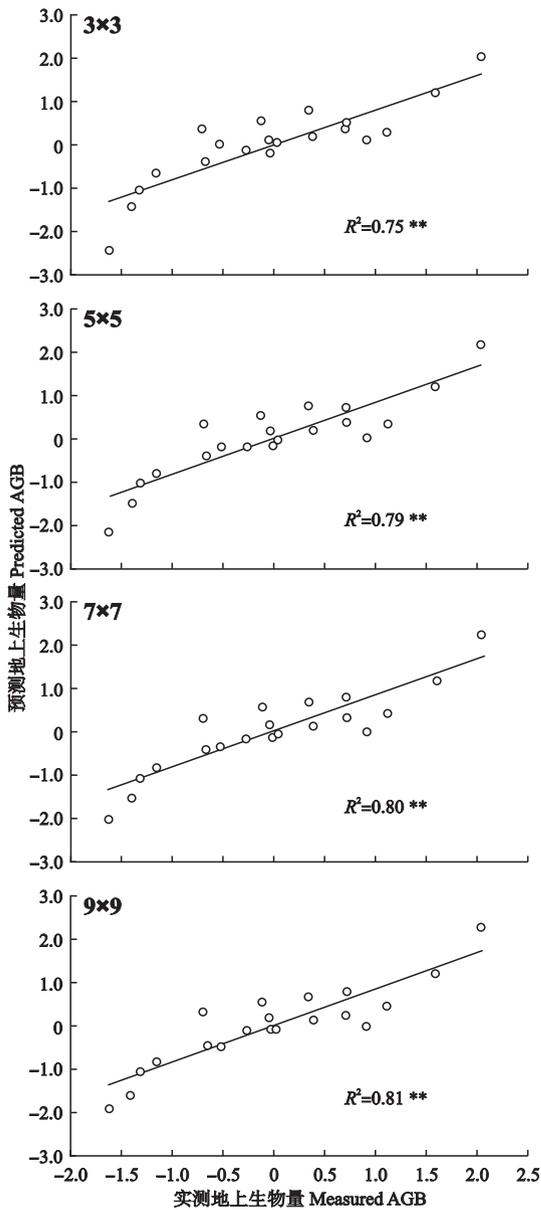


图7 主成分回归模型预测与实测生物量(标准化值)散点图
Fig.7 Scatter diagram of measured AGB and predicted AGB (normalization) of principal component regression model.

要强于纹理特征因子与 AGB 的相关性。此外,本研究选取的 13 种遥感因子之间存在着较为严重的多重共线性问题;2) 留一法交叉验证结果显示,4 种纹理窗口(3×3、5×5、7×7 和 9×9)下绝对系数 R^2 基本都大于 0.50,说明 GF-2 数据在石堡林场油松人工林地上生物量的反演中可以实现较高精度。其中,AGB 估算效果最好的是使用了 9×9 纹理窗口的主成分回归模型($R^2=0.81$, RMSE=0.43),估算效果最差的是使用了 3×3 窗口的普通回归模型($R^2=0.48$, RMSE=0.95)。综合来看,主成分回归模型在 5 类不同的回归模型中,预测的 AGB 值与油松人工林实测 AGB 值都最接近,反演的精度最高;纹理设置为 9×9

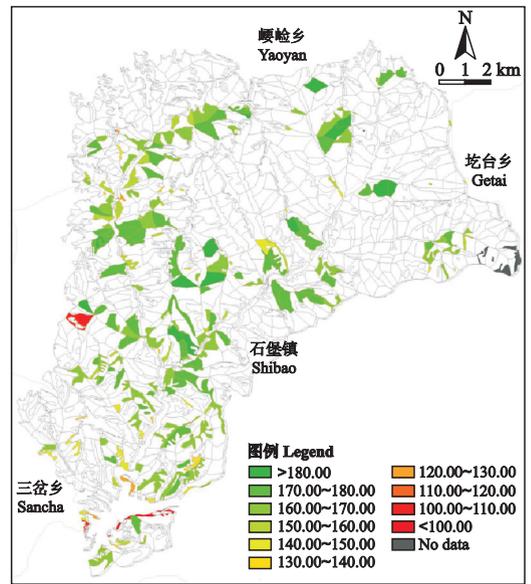


图8 石堡林场油松人工林地上生物量分布
Fig.8 Aboveground biomass distribution of *Pinus tabulaeformis* plantation in Shibao forest farm.

窗口在 4 种不同的窗口中,预测的精度表现得最好。

在研究过程中也存在一些不足,主要包含以下 5 个方面:1) 在研究区进行样地调查时未采伐足够的标准木,所以实测地上生物量是基于油松蓄积量-生物量转换方程计算的,准确性不如标准木法高;2) 调查的样地数量偏少,对研究区的代表性不足;3) 本研究仅使用 GF-2 影像针对石堡林场的油松人工林地上生物量进行了遥感反演研究,尚需要通过黄土高原其他不同地区的不同树种进一步比较与完善;4) 本研究仅从 GF-2 数据中提取了遥感因子结合实测 AGB 进行建模,而未讨论地形、海拔以及卫星数据各单波段等因素对模型可能造成的影响;5) 在 5 类回归模型中,参数如何进行优化设置,惩罚因子如何影响油松人工林地上生物量的估算精度,还有待进一步研究。

参考文献

- [1] Oliver JE. Intergovernmental panel in climate change (IPCC). *Encyclopedia of Energy Natural Resource & Environmental Economics*, 2013, **26**: 48-56
- [2] Parmesan C. Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution & Systematics*, 2006, **37**: 637-669
- [3] Wang S-W (王绍武), Ye J-L (叶瑾琳). An analysis of global warming during the last one hundred years. *Scientia Atmospherica Sinica* (大气科学), 1995, **19** (5): 545-553 (in Chinese)
- [4] Huang Y-L (黄艳丽), Li Z-B (李占斌), Su H (苏辉), et al. Effect of man-made forest on soil moisture of

- different slopes in upper and lower reaches of small watershed of Loess Plateau. *Transactions of the Chinese Society of Agricultural Engineering* (农业工程学报), 2018, **34**(15): 108–116 (in Chinese)
- [5] Liu Y-L (刘玉林), Zhu G-Y (朱广宇), Deng L (邓蕾), *et al.* Effects of natural vegetation restoration and afforestation on soil carbon and nitrogen storage in the Loess Plateau, China. *Chinese Journal of Applied Ecology* (应用生态学报), 2018, **29**(7): 2163–2172 (in Chinese)
- [6] Niu S (牛 硕), Wang G-R (王广儒), Li W-Z (李卫忠), *et al.* Simulated tending and predicting the *Pinus tabulaeformis* plantation in Huanglong Mountain. *Journal of Northeast Forestry University* (东北林业大学学报), 2016, **44**(3): 16–20 (in Chinese)
- [7] Sun WY, Mu XM, Song XY, *et al.* Changes in extreme temperature and precipitation events in the Loess Plateau (China) during 1960–2013 under global warming. *Atmospheric Research*, 2016, **168**: 33–48
- [8] Xin ZB, Xu JX, Zheng W. Spatiotemporal variations of vegetation cover on the Chinese Loess Plateau (1981–2006): Impacts of climate changes and human activities. *Science in China Series D: Earth Sciences*, 2008, **51**: 67–78
- [9] Zhao M (赵 敏), Zhou G-S (周广胜). Forest Inventory Data (FID)-based biomass models and their prospects. *Chinese Journal of Applied Ecology* (应用生态学报), 2004, **15**(8): 1468–1472 (in Chinese)
- [10] Fang X-W (方向文), Jiang Z-R (蒋志荣). The review of remote sensing application to investigate forest resources. *Journal of Gansu Agriculture University* (甘肃农业大学学报), 2003, **38**(3): 267–273 (in Chinese)
- [11] Wang W-F (王维枫), Lei Y-C (雷渊才), Wang X-F (王雪峰), *et al.* A review of forest biomass models. *Journal of Northwest Forestry University* (西北林学院学报), 2008, **23**(2): 58–63 (in Chinese)
- [12] Zhi Z, Xin T, Chen EX. Review of methods on estimating forest aboveground biomass. *Journal of Beijing Forestry University*, 2011, **33**: 144–150
- [13] Zhu XL, Liu DS. Improving forest aboveground biomass estimation using seasonal landsat NDVI time-series. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, **102**: 222–231
- [14] Dube T, Mutanga O. Evaluating the utility of the medium-spatial resolution Landsat 8 multispectral sensor in quantifying aboveground biomass in uMgeni catchment, South Africa. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, **101**: 36–46
- [15] Yin GD, Zhang Y, Sun Y, *et al.* MODIS based estimation of forest aboveground biomass in China. *PLoS One*, 2015, **10**(6): 130–143
- [16] Li Y (李 燕), Zhang J-G (张建国), Duan A-G (段爱国), *et al.* Selection of biomass estimation models for Chinese fir plantation. *Chinese Journal of Applied Ecology* (应用生态学报), 2010, **21**(12): 3036–3046 (in Chinese)
- [17] Zhou JJ, Zao Z, Zhao QX, *et al.* Quantification of aboveground forest biomass using quickbird imagery, topographic variables, and field data. *Journal of Applied Remote Sensing*, 2013, **7**: 73–84
- [18] Wang B (王 彬), Wei T-X (魏天兴), Liu Z (刘钊). Characteristic of nutrient cycling and ecosystem structure of *Larix principis-rupprechtii* in hilly Loess Plateau. *Research of Soil and Water Conservation* (水土保持研究), 2017, **24**(6): 45–51 (in Chinese)
- [19] Meng S-L (蒙诗栋), Pang Y (庞 勇), Zhang Z-J (张钟军), *et al.* Estimation of aboveground biomass in a temperate forest using texture information from WorldView-2. *Journal of Remote Sensing* (遥感学报), 2017, **21**(5): 812–824 (in Chinese)
- [20] Wu P (吴 平), Huang Q-F (黄庆丰), Tang X-H (唐雪海), *et al.* Estimation of aboveground tree biomass based on rapideye imagery. *Journal of Anhui Agricultural University* (安徽农业大学学报), 2016, **43**(6): 891–898 (in Chinese)
- [21] Wang J-R (王璟睿), Shen W-J (沈文娟), Li W-Z (李卫正), *et al.* Performances comparison of multiple non-linear models for estimation plantations' biomass based on rapideye imagery. *Journal of Northwest Forestry University* (西北林学院学报), 2015, **30**(6): 196–202 (in Chinese)
- [22] Propastin P. Large-scale mapping of aboveground biomass of tropical rainforest in Sulawesi, Indonesia, using landsat ETM⁺ and MODIS data. *GIScience & Remote Sensing*, 2013, **50**: 633–651
- [23] Ji L, Wylie BK, Nossor DR, *et al.* Estimating aboveground biomass in interior Alaska with landsat data and field measurements. *International Journal of Applied Earth Observation and Geoinformation*, 2012, **18**: 451–461
- [24] Ni WJ, Zhang ZY, Sun GQ, *et al.* The penetration depth derived from the synthesis of ALOS/PALSAR InSAR data and ASTER GDEM for the mapping of forest biomass. *Remote Sensing*, 2014, **6**: 7303–7319
- [25] Cheng X-Q (程小琴), Han H-R (韩海荣), Kang F-F (康峰峰). Biomass, carbon accumulation and its partitioning of a *Pinus tabulaeformis* plantation ecosystem in Shanxi Province, China. *Chinese Journal of Ecology* (生态学杂志), 2012, **31**(10): 2455–2460 (in Chinese)
- [26] Liu Q (刘 茜), Yang L (杨 乐), Liu Q-H (柳钦火), *et al.* Review of forest aboveground biomass inversion methods based on remote sensing technology. *Journal of Remote Sensing* (遥感学报), 2015, **19**(1): 62–74 (in Chinese)
- [27] Dube T, Mutanga O, Shoko C, *et al.* Remote sensing of aboveground forest biomass: A review. *Tropical Ecology*, 2016, **57**: 125–132
- [28] Huang HB, Liu CX, Wang XY, *et al.* Integration of multi-resource remotely sensed data and allometric models for forest aboveground biomass estimation in China. *Remote Sensing of Environment*, 2019, **221**: 225–234
- [29] Fang J-Y (方精云), Liu G-H (刘国华), Xu S-L (徐嵩龄). Biomass and net production of forest vegetation

- in China. *Acta Ecologica Sinica* (生态学报), 1996, **16** (5): 497–508 (in Chinese)
- [30] Liu M (刘明), Huang H-J (黄恒君). The least squares of spatial regression models. *Statistics & Information Forum* (统计与信息论坛), 2014, **29**(10): 9–14 (in Chinese)
- [31] Feng S-P (冯守平), Shi Z (石泽), Zou J (邹瑾). Comparison of several methods for parameter estimation in unitary linear regression model. *Statistics & Decision* (统计与决策), 2008 (24): 152–153 (in Chinese)
- [32] Wang X-Z (王秀珍), Huang J-F (黄敬峰), Li Y-M (李云梅), *et al.* Study on hyperspectral remote sensing estimation models for the ground fresh biomass of rice. *Acta Agronomica Sinica* (作物学报), 2003, **29**(6): 815–821 (in Chinese)
- [33] Pu R-L (浦瑞良), Gong P (宫鹏). Relationships between forest biochemical concentrations and CASI data along the Oregon transect. *Journal of Remote Sensing* (遥感学报), 1997, **1**(2): 115–123 (in Chinese)
- [34] Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 1970, **12**: 55–67
- [35] Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1996, **58**: 267–288
- [36] Guo C-Q (郭呈全), Chen X-Z (陈希镇). SPSS implementation of principal component regression. *Statistics & Decision* (统计与决策), 2011(5): 157–159 (in Chinese)
- [37] Qu L-P (渠丽萍), Zhang L-Q (张丽琴), Hu W-Y (胡伟艳). Factors influencing change in urban land intensive use: A case study of Wuhan City. *Resources Science* (资源科学), 2010, **32**(5): 970–975 (in Chinese)
- [38] Shu X-H (舒晓惠), Liu J-P (刘建平). Several problems in dealing with multicollinearity by principal component regression. *Statistics & Decision* (统计与决策), 2004(10): 25–26 (in Chinese)

作者简介 苟睿坤, 男, 1994年生, 博士研究生. 主要从事生态遥感和森林经理学研究. E-mail: grk_123@163.com

责任编辑 孙菊
