



基于高光谱植被指数的大豆地上部生物量估算模型研究

龚荣新¹, 鲁向晖^{1,2}, 张海娜^{1,2}, 王倩¹, 陈志琪¹, 杨宝城¹, 马露露¹

(1. 南昌工程学院 水利与生态工程学院, 江西 南昌 330099; 2. 南昌工程学院 江西省樟树繁育与开发利用工程研究中心, 江西 南昌 330099)

摘要:本研究以大豆为研究对象, 利用光谱仪测定大豆鼓粒期冠层高光谱数据并计算多种高光谱植被指数。分别采用一元线性回归(LR)、支持向量机(SVM)、反向传播神经网络(BPNN)和随机森林(RF)建立高光谱植被指数组合和大豆地上部生物量之间相互关系的数学模型。结果显示: 基于LR、SVM、BPNN和RF建立的估算AGB模型的決定系数(R^2)分别为0.59、0.71、0.73和0.76; 均方根误差(RMSE)分别为2 559.0、481.1、1 194.6和805.2 $\text{kg}\cdot\text{hm}^{-2}$; 相对分析误差(RPD)分别为1.22、1.55、1.87和1.92。基于RF建立模型的预测精准度比LR、SVM和BPNN模型更可靠, 因此运用RF模型可以更精确地估算大豆地上部生物量。

关键词: 大豆; 地上部生物量; 随机森林; 高光谱; 植被指数

Model Analysis of Estimating Soybean Above-ground Biomass by Hyperspectral Vegetation Index

GONG Rong-xin¹, LU Xiang-hui^{1,2}, ZHANG Hai-na^{1,2}, WANG Qian¹, CHEN Zhi-qi¹, YANG Bao-cheng¹, MA Lu-lu¹

(1. School of Water Conservancy and Ecological Engineering, Nanchang Institute of Technology, Nanchang 330099, China; 2. Jiangxi Camphor Tree Breeding and Development and Utilization Engineering Research Center, Nanchang Institute of Technology, Nanchang 330099, China)

Abstract: In this study, soybean was taken as the research object, and the hyper-spectral data of soybean canopy at the filling stage were measured by spectrograph and various hyper-spectral vegetation indexes were calculated. We used Linear Regression(LR), Support Vector Machine(SVM), Back Propagation Neural Network(BPNN) and Random Forest(RF) to established the mathematical model of the relationship between hyper-spectral vegetation index combination and soybean above-ground biomass. The results showed that the determination coefficients(R^2) of estimating AGB model based on LR, SVM, BPNN and RF were 0.59, 0.71, 0.73 and 0.76 respectively. The Root Mean Square Error(RMSE) was 2 559.0, 481.1, 1 194.6 and 805.2 $\text{kg}\cdot\text{ha}^{-1}$ respectively. The relative analysis error(RPD) was 1.22, 1.55, 1.87 and 1.92 respectively. The prediction accuracy of models based on RF was more reliable than LR, SVM and BPNN models. Therefore, the above ground biomass of soybean can be estimated more accurately by using RF models.

Keywords: soybean; above-ground biomass; random forest; hyper-spectral; vegetation index

大豆是中国重要粮油作物之一, 其种子含有丰富植物蛋白质^[1], 大豆产量的高低关乎国家粮食安全, 与人民生活质量之间存在紧密关联, 因此研究大豆产量的估算方法, 对产业发展具有重要意义。地上部生物量(AGB)是绿色植物光合作用的产物, 是衡量作物光能利用效率、估测作物生长状态的重要指标, 在作物田间监测和管理上有着广泛应用前景^[2-3]。

已有研究表明 AGB 的传统检测方法消耗时间较长且破坏性大, 检测成本高且结果不具备适时性, 同时无法适应现代农业生产管理的要求^[4-6]。而基于高光谱技术的检测方法凭借其快速高效且不损害作物的优势可以弥补传统监测方法的不足^[7-9]。然而, 前人基于高光谱反射率对植物 AGB

的预测主要应用于森林、草地等^[10-15], 近年来一些学者也开始尝试研究光谱数据与实测地上生物量的关系进而估算农作物的地上生物量^[16-19]。王秀珍等^[16]使用 ASD 地物光谱仪分析了不同氮处理的水稻光谱特征, 发现一些高光谱特征值及其组合构成的植被指数与水稻地上生物量密切相关, 用红边内一阶微分之和与蓝边内一阶微分之和的比值所构成组合建立的 AGB 线性估算模型精度最佳, 验证集決定系数(R^2)为 0.54。侯学会等^[17]使用 ASD 地物光谱仪获取了小麦的高光谱数据, 构建了不同光谱参数和光谱植被指数估算小麦 AGB 的指数模型, 光谱参数中红边位置(REP) 所构建模型最优, 验证集決定系数为 0.82, 植被指数中增强型植被指数(EVI)所构建模型精度最高, 验证集決定系数为

收稿日期: 2023-01-09

基金项目: 国家自然科学基金(52269013)。

第一作者: 龚荣新(1999—), 男, 硕士研究生, 主要从事农林水文遥感研究。E-mail: 2218804558@qq.com。

通讯作者: 鲁向晖(1976—), 男, 副教授, 博士, 主要从事农林水文遥感技术研究。E-mail: xianghuilu@nit.edu.cn。

0.74。童新等^[18]对牧草光谱反射率数据进行一至四阶的微分处理,发现基于二阶微分光谱的 DNVI 模型精度最好(R^2 为 0.7),模型精度随着微分阶数升高而下降。唐延林等^[19]分别测量了水稻、玉米和棉花的高光谱反射率,分析表明它们的高光谱反射率存在差异,其中棉花冠层光谱反射率最大值和红边参数最大值均高于水稻和玉米,表明对于不同植物,由于其生理信息存在着差异,导致光谱特征不同。吴琼等^[20]使用 ASD 地物光谱仪采集了大豆的高光谱数据,发现在大豆鼓粒期建立的光谱数据与大豆地上生物量估算模型精度较高,决定系数(R^2) 为 0.68,其他生育期较低。陆国正等^[21]利用无人机遥感获取光谱数据,采用最小二乘法建立基于光谱参数和植被指数的多元模型估算大豆 AGB,其验证集 R^2 为 0.71。以上研究说明利用高光谱技术监测大豆冠层生长来估算其 AGB 具有可行性,但鲜有报道比较基于不同分析模型进行估算的效果。

大豆鼓粒期,营养生长基本停止,生殖生长占主导地位,是大豆干物质积累最多的时期,做好对本时期大豆的田间检测和管理对提高大豆产量至关重要^[22]。因此本研究在大豆鼓粒期,使用 ASD 地物光谱仪测量大豆冠层高光谱数据,筛选与大豆 AGB 相关性较高的高光谱植被指数,分别采用一元线性回归(LR)、支持向量机(SVM)、反向传播神经网络(BPNN)和随机森林(RF)构建大豆冠层高光谱植被指数与地上生物量的估算模型,以期建立高效无损的大豆 AGB 估算方法,为基于高光谱技术估算大豆 AGB 提供理论依据,进而给更大规模农作物的长期监测研究提供参考。

1 材料与方法

1.1 材料

供试大豆品种为适合陕西关中地区夏播种植的高油型夏大豆品种山宁 17 号。

1.2 试验设计

试验区位于西北农林科技大学旱区节水灌溉试验站(108°43′27″E,34°17′1″N),属于暖温带,海拔约 521 m,大陆性季风气候,夏季湿度大、温度高,冬季温度低、湿度低,气温 -19.4 ~ 42.0 ℃。降雨季节分配不均,夏季多雨,多年平均降水量 635.1 mm。试验区土壤属于重壤土,黏粒较多,平均土壤干容积密度为 1.58 g·cm⁻³。在大豆鼓粒期于试验区采

集光谱数据及大豆 AGB 共 54 份有效样本,随机选择 2/3 样本作为建模集,剩余样本作为验证集,分别基于 LR、SVM、BPNN 和 RF 模型估算大豆 AGB,并验证和对比不同分析模型的估算精度。

1.3 方法

1.3.1 数据采集 光谱数据采集:采用美国 ASD 地物光谱仪测定大豆冠层光谱反射率。将仪器波长分为 300 ~ 1 000 nm 和 1 000 ~ 1 830 nm 两个波段,光谱分辨率分别为 3 和 10 nm,采样间隔分别设置为 1.4 和 2.0 nm。仪器输出间隔设置为 1.0 nm,预热 10 min 后设置光纤长度 1.5 m,采用 25°现场角,每次采样前后用反射率为 1 的标准白板校正。于干燥和无风尘的晴天测定光谱数据,时间为 2021 年 8 月 27 日,中午 11:45—12:15,测量时手持仪器,将探测针头竖直置于距大豆冠层顶部 0.75 m 左右的位置。每个试验区分别测定 3 组光谱样方数据,每个光谱样方分别采集其中 10 条光谱曲线,计算光谱加权的平均值,以其作为该光谱样方的光谱反射率^[23]。

大豆 AGB 采集:大豆 AGB 指单株大豆样本除去地下根部的干物质总量,于 2021 年 8 月 27 日在 54 个大豆小区中进行田间取样,每个小区取 3 株长势相近且具有代表性的大豆作为有效试验样本。使用烘箱将样本于 100 ℃烘烤 30 min 杀青,于 75 ℃烘烤至恒质量,并用电子天平测量大豆样本的干重,以大豆样本干重平均值为单株地上生物量,并计算单位面积地上生物量。

1.3.2 光谱数据预处理 光谱数据易受到如水环境和空气噪声环境等诸多因素带来的干扰。在 1 300 ~ 1 400 nm 处存在水分强吸收带,1 800 ~ 1 850 nm 处有较明显的边缘噪音。为减少无用信息对光谱数据质量的负面影响,本研究采用卷积平滑滤波法(Savitzky-Golay,SG)对这两处共 150 个光谱波段上获取的光谱反射率数据进行预先处理,采用二次多项式,平滑点数为 9,进行函数拟合和滤波去噪^[24]。

1.3.3 植被指数选择 植被指数是基于光谱数据得出的反映植物各种状态特征的指数^[25],本研究计算 18 种常见的高光谱植被指数,计算公式如表 1 所示, R_{800} 和 R_{670} 分别为波长 800 和 670 nm 处的光谱反射率,其余表达式同理。

表 1 植被指数及其计算式

Table 1 Vegetation indexes and their calculation formulas

植被指数 Vegetation index	公式 Formula	文献 Reference
归一化植被指数 NDVI	$\frac{R_{800} - R_{670}}{R_{800} + R_{670}}$	[26]
绿度归一化植被指数 GNDVI	$\frac{R_{801} - R_{550}}{R_{801} + R_{550}}$	[26]
土壤调节植被指数 SAVI	$\frac{1.5 \times (R_{800} - R_{670})}{R_{800} - R_{670} + 0.5}$	[26]
叶绿素吸收比值指数 CARI	$(R_{700} - R_{670}) - 0.2 \times (R_{700} + R_{670})$	[26]
MERIS 陆地叶绿素指数 MTCI	$\frac{R_{754} - R_{709}}{R_{709} - R_{681}}$	[26]
光化学植被指数 PRI	$\frac{R_{570} - R_{531}}{R_{570} + R_{531}}$	[26]
结构不敏感色素指数 SIPI	$\frac{R_{800} - R_{450}}{R_{800} + R_{450}}$	[26]
简单比值指数 SR	$\frac{R_{740}}{R_{720}}$	[26]
三角植被指数 TVI	$60 \times (R_{800} - R_{550}) - 100 \times (R_{670} - R_{550})$	[26]
绿红植被指数 GRVI	$\frac{R_{800}}{R_{550}}$	[26]
红外植被指数 IPVI	$R_{800} \times (R_{800} + R_{670})$	[26]
多时相植被指数 MTVI	$1.2 \times [1.2 \times (R_{800} - R_{550}) - 2.5 \times (R_{670} - R_{550})]$	[26]
增强植被指数 EVI	$2.5 \times (R_{800} - R_{670}) \times (R_{800} + 6 \times R_{760} - 7.5 \times R_{450} + 1)$	[26]
光化学植被指数 PRI	$\frac{R_{531} - R_{570}}{R_{570} + R_{531}}$	[26]
修正比值植被指数 MSR	$\frac{\frac{R_{800}}{R_{670}} - 1}{\sqrt{\frac{R_{800}}{R_{670}} + 1}}$	[27]
红边位置指数 REP	$40 \times \left(\frac{R_{670} + R_{700}}{2} - R_{700} \right)$ $700 + \frac{R_{740} - R_{700}}{R_{740} - R_{700}}$	[27]
改进红边归一化植被指数 mNDVI ₇₀₅	$\frac{R_{705} - R_{750}}{R_{750} + R_{705}^2 \times R_{445}}$	[27]
再归一化差植被指数 RDVI	$\frac{R_{800} - R_{670}}{\sqrt{R_{800} + R_{670}}}$	[27]

1.3.4 模型构建 本研究模型均在 MATLAB 软件中运行,采用 LR、SVM、BPNN 和 RF 构建大豆 AGB 估算模型。LR 可以反映单一自变量与因变量之间的线性关系,具有高可靠性和运用简单等优点^[28]。SVM 是一种监督式分类机器学习算法,可以很好地解决样本、非线性以及高维的机器学习问题^[29],本研究使用的 SVM 模型是基于 MATLAB 的 Libsvm 工具箱建立的,核函数类型设置为“poly”,并采用网格搜索方法优化 SVM 模型的参数,惩罚系数 C 和 γ 分别为 20 和 0.03。BPNN 是一种多层前馈神经网络,其最大的特点是信号前向传播和误差反向传播,这种传播机制使其输出数据与期望值的接近度

越来越高^[30],本研究中的 BPNN 估计模型是使用 MATLAB 的 Neural-Net-work 工具箱构建的,基于数值优化理论,将隐式层的传递函数设置为 Tansig 函数,网络训练采用 Levenberg-Marquardt (L-M) 算法,经过多次训练,中间层神经元数量设置为 17,最大迭代次数设置为 2 000,每次用来梯度下降的批处理数据大小为 9,神经网络训练后进行模拟,得到模拟值。RF 是由 Breiman 提出的一种基于分类树的算法,随机生成多个分类树,之后对它们进行汇总,是能较好取代传统机器学习方法的新的模型^[31]。本研究通过参数优化与训练,确定 RF 模型的决策树数量为 500。

1.3.5 模型精度评价 为选出最理想的模型,并对其可靠性和预测结果的精准性进行检验,选取 R^2 、 $RMSE$ 和 RPD 作为指标,其公式^[32]如下:

$$R^2 = \frac{\sum_{i=1}^n (\hat{a}_i - \bar{b})^2}{\sum_{i=1}^n (a_i - \bar{b})^2}; RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{a}_i - a_i)^2}{n}};$$
$$RPD = \frac{\sum_{i=1}^n (a_i - \bar{b})^2}{n - 1} \div \sqrt{\frac{\sum_{i=1}^n (a_i - \hat{a}_i)^2}{n}}.$$

式中, n 为样本个数,取 54, \hat{a}_i 为 54 个试验小区大豆 AGB 估测值, a_i 为 54 个试验小区大豆 AGB 实测值, \bar{b} 为大豆 AGB 实际测量值的均值。 R^2 越趋近 1,表明模型的拟合性越强; $RMSE$ 越接近于 0,模型的拟合性越强; RPD 为标准差与其均方根误差之比,当 RPD 大于 1.4 时可以预测模型^[33]。本研究对 4 种模型的精度参数进行统计和比较,以更直观地选择最优模型。

2 结果与分析

2.1 样本集数据分析

实际测量的大豆地上生物量样本数为 54,平均值为 11 449.1 kg·hm⁻²,标准差为 3 643.4 kg·hm⁻²,最大值 24 513.0 kg·hm⁻²,最小值 5 709.0 kg·hm⁻²,变异系数为 0.31。结果表明不同处理对冠层影响导致离散程度较大。

基于高光谱数据计算植被指数和大豆 AGB 实测值的相关系数如图 1 所示。相关系数取值范围为 [-1,1]。当两个变量同时增大时,它们之间为正相关,相关系数取值范围为 [0,1];当一个变量增大,另一个变量减小时,它们之间为负相关,相关系数取值范围为 [-1,0];当两个变量相关系数为 0 时,它们之间不存在线性相关关系。两个变量相关系数绝对值越大,则两者相关性越强。有 7 种植被指数与大豆 AGB 相关性较好,IPVI 与大豆 AGB 相关性最显著,其相关系数为 0.72,MTVI、EVI 和 TVI 均大于 0.70,SAVI 和 RDVI 为 0.69,mNDVI₇₀₅ 为 0.65,其余植被指数与大豆 AGB 相关性较差。

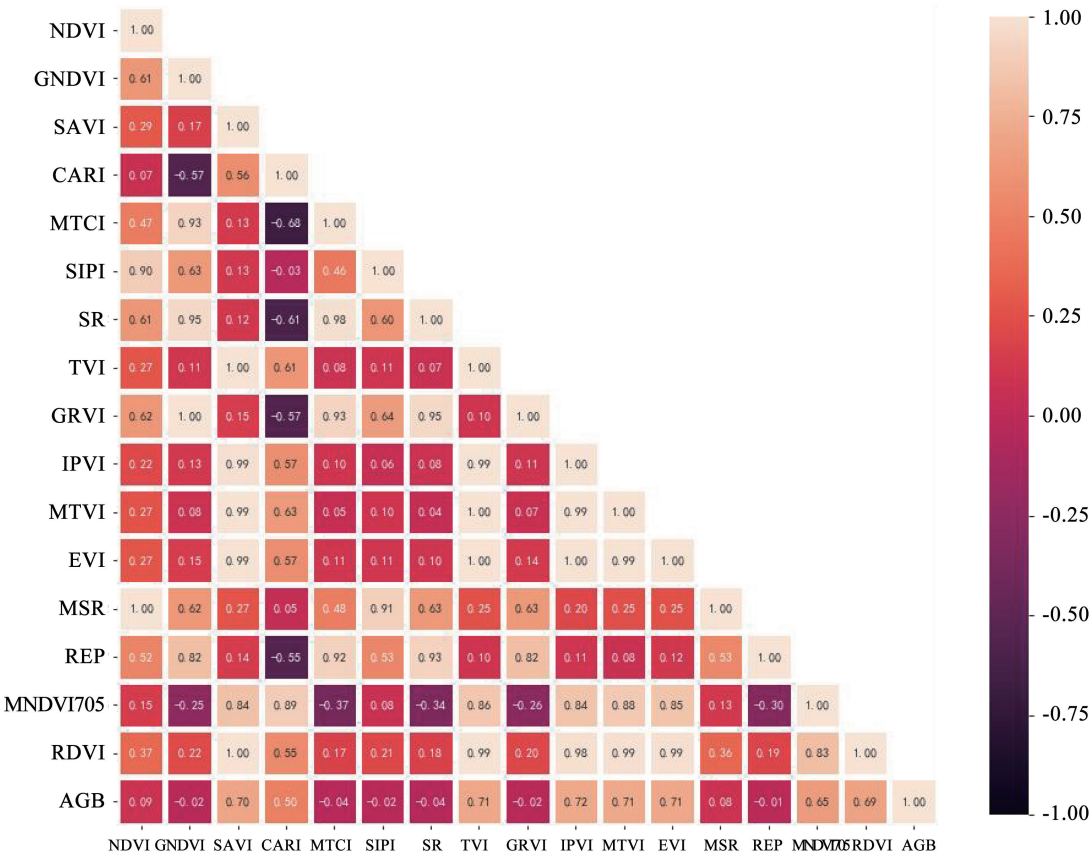


图1 植被指数和大豆 AGB 的 pearson 相关性矩阵

Fig.1 Pearson correlation matrix of vegetation index and soybean AGB

2.2 基于一元线性回归(LR)模型估算

由图 1 可知,SAVI、TVI、IPVI、MTVI、mNDVI705、RDVI 和 EVI 这 7 种植被指数与大豆 AGB 相关系数较高,相关系数均大于 0.65,因此本研究以这 7 种植被指数为自变量,分别与大豆 AGB 实测

值建立 LR 模型。如表 2 所示;对比基于 7 种植被指数建立的模型,其中基于 IPVI 建立的模型优于其他植被指数,其建模集 R^2 为 0.44, RMSE 为 $2\,571.3\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 0.91;验证集 R^2 为 0.59, RMSE 为 $2\,259.2\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 1.22。

表 2 植被指数与生物量一元线性模型及验证

Table 2 Uni-variate linear model of vegetation index and biomass and its verification

植被指数 Vegetation index	表达式 Formula	建模集 Modeling set			验证集 Validation set		
		R^2	RMSE/($\text{kg}\cdot\text{hm}^{-2}$)	RPD	R^2	RMSE/($\text{kg}\cdot\text{hm}^{-2}$)	RPD
SAVI	$y=81543x-48465$	0.43	2606.4	0.88	0.54	2679.7	1.03
TVI	$y=1002x-19096$	0.44	2583.6	0.90	0.57	2609.2	1.10
IPVI	$y=59306x-4315$	0.44	2571.3	0.91	0.59	2559.2	1.22
MTVI	$y=40377x-19041$	0.44	2606.4	0.90	0.58	2612.6	1.09
mNDVI ₇₀₅	$y=3320x-5799$	0.43	2599.1	0.89	0.58	2593.6	0.89
RDVI	$y=134298x-2899$	0.38	2719.9	0.79	0.58	3014.8	0.81
EVI	$y=79057x-41415$	0.40	2678.5	0.82	0.55	2660.7	1.02

2.3 基于 SVM 估算

基于 SVM 的大豆地上生物量模型的拟合结果如图 2 所示;建模集的 R^2 为 0.64, RMSE 为 $1\,194.7\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 1.65;验证集的 R^2 为 0.73, RMSE 为 $1\,194.7\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 1.88。

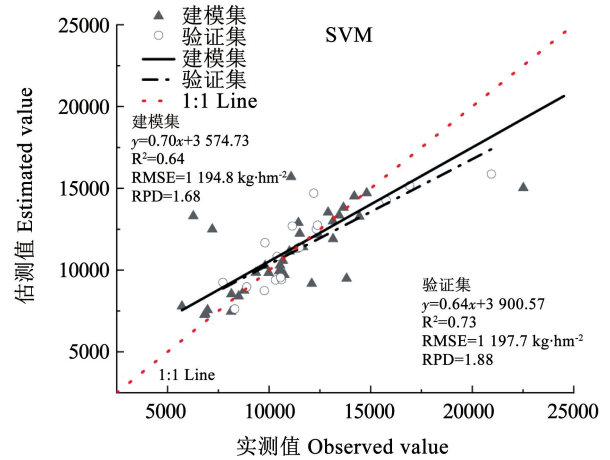


图 2 基于支持向量机的大豆地上生物量估算模型
Fig. 2 Support vector machine model for estimating above ground biomass of soybean

2.4 基于 BPNN 估算

基于 BPNN 的大豆地上生物量估算模型的拟合结果如图 3 所示;建模集的 R^2 为 0.65, RMSE 为 $544.6\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 1.35;验证集的 R^2 为 0.71, RMSE 为 $481.1\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 1.55。

2.5 基于 RF 估算

基于 RF 的大豆地上生物量估算模型的拟合结果如图 4 所示;建模集的 R^2 为 0.83, RMSE 为 $149.5\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 2.37;验证集的 R^2 为 0.76, RMSE 为 $805.1\text{ kg}\cdot\text{hm}^{-2}$, RPD 为 1.92。

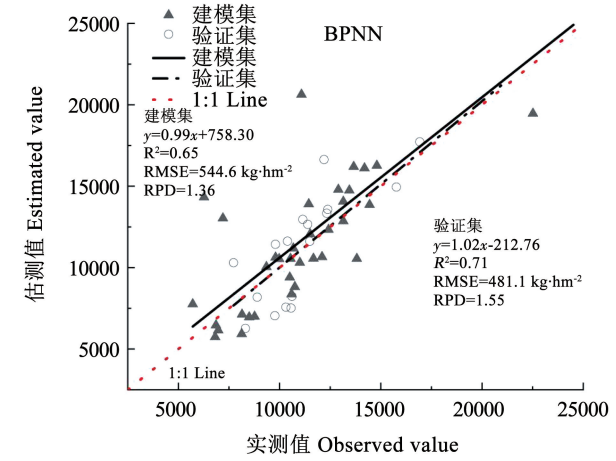


图 3 基于反向传播神经网络的大豆地上生物量估算模型
Fig. 3 Soybean above ground biomass estimation model based on back propagation neural network

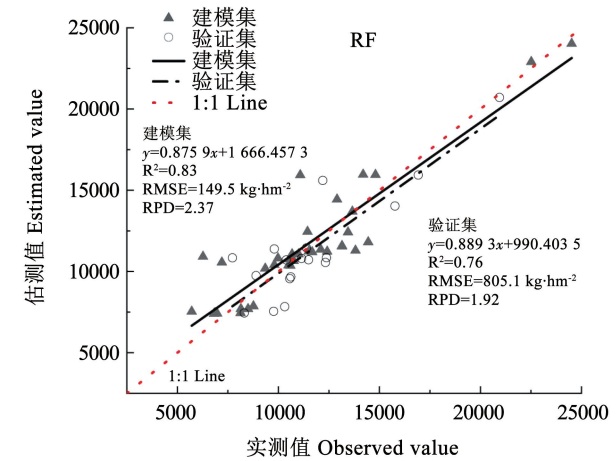


图 4 基于随机森林的大豆地上生物量估算模型
Fig. 4 Above ground biomass estimation model of soybean based on random forest

2.6 模型评价

精度参数统计如表 4 所示:RF 模型建模集和验证集 R^2 均高于其余模型,分别为 0.83 和 0.76;建模集和验证集 RMSE 均最小,建模集为 $149.5\text{ kg}\cdot\text{hm}^{-2}$,验证集为 $805.1\text{ kg}\cdot\text{hm}^{-2}$;建模集 RPD 为 2.37,验

证集 RPD 为 1.92。结果表明:RF 模型可以很好地估算样本,为最优模型;其次是 BPNN 和 SVM 模型,两者估算精度相似,建模集决定系数均大于 0.6,验证集决定系数分别为 0.71 和 0.73,估算样本能力较差;LR 模型精度最差,无法很好估算样本。

表 4 模型精度对比
Table 4 Comparison of model accuracy

模型类型 Model type	建模集 Modeling set			验证集 Validation set		
	R^2	$RMSE/(\text{kg}\cdot\text{hm}^{-2})$	RPD	R^2	$RMSE/(\text{kg}\cdot\text{hm}^{-2})$	RPD
LR	0.44	2571.3	0.91	0.59	2559.0	1.22
BPNN	0.65	544.6	1.35	0.71	481.1	1.55
SVM	0.63	1194.7	1.65	0.73	1194.7	1.87
RF	0.83	149.5	2.37	0.76	805.1	1.92

3 讨论

3.1 相关性分析

当两个变量的相关系数大于 0.50 时,表示两者相关性较好^[33],本研究中大豆 AGB 实测值与所选 7 种植被指数的相关系数均大于 0.65,说明大豆 AGB 与这 7 种植被指数均具有一定相关性。其中,IPVI 与大豆 AGB 的相关系数最高,为 0.72,表明近红外光构成植被指数对大豆地上生物量较为敏感,与赵翊含等^[31]和曹盛中等^[34]的研究结果类似。

3.2 模型精度讨论

通过比较 R^2 、 $RMSE$ 和 RPD 发现,不同建模方式对模型精度评价影响较大,以 SAVI、TVI、IPVI、MTVI、 $mNDVI_{705}$ 、RDVI 和 EVI 这 7 种植被指数分别作为自变量,依次建立估算大豆 AGB 的 LR 模型,精度整体较低,说明单一的植被指数无法很好地估算大豆 AGB,即大豆 AGB 与植被指数之间并非简单的线性关系。

对地上生物量的估算模型精度进行对比时,RF 模型明显优于 SVM、BPNN 和 LR 模型,表明 RF 在反演大豆地上生物量时较其他模型效果更好。栗方亮等^[35]建立并对比蜜柚叶片氮素含量估算模型,发现 RF 模型比 SVM 和 BPNN 模型在反演 AGB 时更具优势,与本研究结果类似;蔡鑫垚^[36]基于 RF 模型估测天然次生林地上生物量,也得到了相似结论。产生以上结果的原因可能是 BPNN 虽然有较高的容错性,且具有学习复杂非线性关系以应对干扰的能力,但是它反向学习过程较为繁琐,梯度下降的学习方法导致其容易局部收敛并找不到最优解,可能导致结果为局部最优解而不是全局优解^[37],此外 BPNN 模型是模拟大脑神经的网络状模型,学习过多由于随机抽取产生的特殊样本可能造成过拟合,从而降低模型精度。而 SVM 主要用于解决二分

类问题,虽然它的解是唯一和全局的,具有较好的学习与估测能力,但是 SVM 对大规模样本训练速度慢,精度低,且 SVM 的核函数参数和误差惩罚因子参数的组合不够完善^[38]。相比于 SVM 模型和 BPNN 模型,RF 模型解决了以往模型可靠性较差的问题,有效防止模型的单向预测,其在预测过程中随机生成分类树和节点,进而降低了发生过拟合的概率,因而可以获得更高的精度^[39]。本研究样本数量较少,选用植被指数间相关性较高,可能降低模型精度,且得到的模型限于鼓粒期单一大豆品种,对于不同生育期大豆以及不同品种大豆的地上部生物量估算还有待于进一步研究。因此,后续模型优化可以扩大样本数量,进一步筛选对大豆地上部生物量更敏感的光谱波段和植被指数,并根据大豆不同品种和生育期分别讨论和建立模型,以期得到更准确和全面的大豆地上部生物量估算模型。

4 结论

本研究筛选与大豆 AGB 相关的高光谱植被指数并建立利用高光谱植被指数估算大豆 AGB 模型。SAVI、TVI、IPVI、MTVI、 $mNDVI_{705}$ 、RDVI 和 EVI 这 7 种高光谱植被指数与大豆地上部都有显著相关性,相关系数均大于 0.65,说明基于这 7 种植被指数估算大豆 AGB 具有可行性。采用 LR 建立的大豆 AGB 估算模型与单一植被指数的估算模型精度整体较低,说明大豆 AGB 与植被指数之间不是简单的线性关系。不同建模方式对模型精度评价影响较大,RF 模型的精度高于 LR、SVM 和 BPNN 模型。其建模集 R^2 大于 0.80, RMSE 为 $149.5\text{ kg}\cdot\text{hm}^{-2}$, RPD 高于 2.00;验证集的 R^2 为 0.76, RMSE 为 $805.1\text{ kg}\cdot\text{hm}^{-2}$, RPD 高于 1.90。因此 RF 可以作为监测大豆地上生物量的优选模型。

参考文献

[1] ZHOU X, ZHAO H J, ZHAO X, et al. Oil bodies extracted from high-oil soybeans (*Glycine max*) exhibited higher oxidative and physical stability than oil bodies from high-protein soybeans[J]. Food & Function, 2022, 13(6): 3271-3282.

[2] 杨鹏万, 付刚, 李云龙, 等. 多光谱相机估算藏北高寒草甸地上生物量[J]. 草业科学, 2014, 31(7): 1211-1217. (YANG P W, FU G, LI Y L, et al. Estimation of above-ground biomass of alpine meadow in northern Tibet by multi-spectral camera[J]. Grassland Science, 2014, 31(7): 1211-1217.)

[3] 张晋玉, 徐新娟, 晁毛妮, 等. 大豆 GmZAT12 基因的克隆和表达分析[J]. 华北农学报, 2022, 37(3): 1-7. (ZHANG J Y, XU X J, CHAO M N, et al. Cloning and expression analysis of soybean GmZAT12 gene[J]. North China Journal of Agriculture, 2022, 37(3): 1-7.)

[4] TODA Y, KAGA A, KAJIYAKANE GAE H, et al. Genomic prediction modeling of soybean biomass using UAV-based remote sensing and longitudinal model parameters[J]. The Plant Genome, 2021, 14(3): E20157.

[5] HUYNH T, LEWIS T, APPELEGATE G, et al. Allometric equations to estimate above-ground biomass in spotted gum (*Corymbia citriodora* Subspecies *variegata*) plantations in Queensland[J]. Forests, 2022, 13(3): 486.

[6] 宋开山, 张柏, 李方, 等. 高光谱反射率与大豆叶面积及地上鲜生物量的相关分析[J]. 农业工程学报, 2005, 21(1): 36-40. (SONG K S, ZHANG B, LI F, et al. Correlation analysis between hyper-spectral reflectance and soybean leaf area and above-ground fresh biomass[J]. Journal of Agricultural Engineering, 2005, 21(1): 36-40.)

[7] 郭志勋. 棉花冠层叶片叶绿素相对含量的高光谱反演研究[D]. 阿拉尔: 塔里木大学, 2022. (GUO Z X. Hyper-spectral retrieval of relative chlorophyll content in cotton canopy leaves[D]. Alar: Tarim University, 2022.)

[8] 郭建彪. 基于高光谱遥感的小麦生长及氮素利用效率估测研究[D]. 郑州: 河南农业大学, 2021. (GUO J B. Research on estimation of wheat growth and nitrogen use efficiency based on hyper-spectral remote sensing [D]. Zhengzhou: Henan Agricultural University, 2021.)

[9] 白丽. 基于高光谱数据的玉米 LAI 和叶绿素含量遥感反演方法的研究[D]. 石河子: 石河子大学, 2020. (BAI L. Study on remote sensing retrieval method of maize lai and chlorophyll content based on hyperspectral data [D]. Shihezi: Shihezi University, 2020.)

[10] LUO Z K, TAN K Y. Hyper-spectral remote sensing image classification based on random forest[J]. Academic Journal of Computing & Information Science, 2021, 4(6): 67-71.

[11] 许婷婷, 董智, 郭建英, 等. 放牧对内蒙古典型草原植物群落特征与土壤有机碳的影响[J]. 草地学报, 2022, 30(9): 2273-2279. (XU T T, DONG Z, GUO J Y, et al. Effects of grazing on characteristics of typical grassland plant communities and soil organic carbon in Inner Mongolia[J]. Journal of Grassland Science, 2022, 30(9): 2273-2279.)

[12] 温雨笑, 吕杰, 马庆勋, 等. 高光谱和 Li-DAR 联合反演森林

生物量研究[J]. 测绘通报, 2022(7): 38-42. (WEN Y X, LY U J, MA Q X, et al. Study on joint retrieval of forest biomass by hyper spectral and Li-DAR [J]. Bulletin of Surveying and Mapping, 2022(7): 38-42.)

[13] 周蔚, 吕延杰, 林起楠. 联合光学和合成孔径雷达数据的太平湖森林地上生物量反演研究[J]. 西北林学院学报, 2023, 38(2): 193-200. (ZHOU W, LYU Y J, LIN Q N. Inversion of above-ground biomass of Taiping Lake forest using optical and synthetic aperture radar data[J]. Journal of Northwest Forestry University, 2023, 38(2): 193-200.)

[14] 刘岳含, 黎松松, 郑伟, 等. 旅游干扰下山地草甸光谱特征及其与冠层特征的关系[J]. 中国草地学报, 2020, 42(2): 83-94. (LIU Y H, LI S S, ZHENG W, et al. Spectral characteristics of mountain meadow under tourism interference and their relationship with canopy characteristics[J]. Chinese Journal of Grassland Science, 2020, 42(2): 83-94.)

[15] KANG X Y, ZHANG A W, PANG H Y. Estimation of grassland above-ground biomass from UAV-mounted hyper-spectral image by optimized spectral reconstruction[J]. Spectroscopy and Spectral Analysis, 2021, 41(1): 250-256.

[16] 王秀珍, 黄敬峰, 李云梅, 等. 水稻地上鲜生物量的高光谱遥感估算模型研究[J]. 作物学报, 2003, 29(6): 815-821. (WANG X Z, HUANG J F, LI Y M, et al. Study on hyper-spectral remote sensing estimation model of rice above-ground fresh biomass[J]. Journal of Crops, 2003, 29(6): 815-821.)

[17] 侯学会, 牛铮, 黄妮, 等. 小麦生物量和真实叶面积指数的高光谱遥感估算模型[J]. 国土资源遥感, 2012, 24(4): 30-35. (HOU X X, NIU Z, HUANG N, et al. Hyper-spectral remote sensing estimation model of wheat biomass and real leaf area index [J]. Remote Sensing of Land and Resources, 2012, 24(4): 30-35.)

[18] 童新, 杨震雷, 张亦然, 等. 基于不同阶微分高光谱植被指数的牧区草地上生物量估算[J]. 草地学报, 2022, 30(9): 2438-2448. (TONG X, YANG Z L, ZHANG Y R, et al. Estimation of pasture aboveground biomass using different orders of differential hyperspectral vegetation indices [J]. Journal of Grassland Science, 2022, 30(9): 2438-2448.)

[19] 唐延林, 黄敬峰, 王秀珍, 等. 水稻、玉米、棉花的高光谱及其红边特征比较[J]. 中国农业科学, 2004, 37(1): 29-35. (TANG Y L, HUANG J F, WANG X Z, et al. Comparison of hyper-spectral and red edge characteristics of rice, corn and cotton [J]. China Agricultural Sciences, 2004, 37(1): 29-35.)

[20] 吴琼, 齐波, 赵团结, 等. 高光谱遥感估测大豆冠层生长和籽粒产量的探讨[J]. 作物学报, 2013, 39(2): 309-318. (WU Q, QI B, ZHAO T J, et al. Estimation of soybean canopy growth and seed yield by hyper-spectral remote sensing [J]. Journal of Crops, 2013, 39(2): 309-318.)

[21] 陆国政, 杨贵军, 赵晓庆, 等. 基于多载荷无人机遥感的大豆地上鲜生物量反演[J]. 大豆科学, 2017, 36(1): 41-50. (LU G Z, YANG G J, ZHAO X Q, et al. Retrieval of above-ground fresh biomass of soybean based on multi load UAV remote sensing[J]. Soybean Science, 2017, 36(1): 41-50.)

[22] 孟田. 不同年代育成大豆品种物质形成和养分吸收规律的研究[D]. 沈阳: 沈阳农业大学, 2016. (MENG T. Studies on the law of material formation and nutrient absorption of soybean cultivars raised in different ages [D]. Shenyang: Shenyang

Agricultural University, 2016.)

[23] 吴建付,陈功,杨红丽,等. 利用高光谱技术进行草地上部生物量估测[J]. 草业与畜牧, 2009(4): 1-3. (WU J F, CHEN G, YANG H L, et al. Estimation of above-ground biomass of grassland using hyperspectral technology [J]. Grassland and Animal Husbandry, 2009(4): 1-3.)

[24] 李燕丽, 吴士文, 刘娅, 等. 基于 FastICA 盲源分离法去除土壤干扰的小麦生物量高光谱估算[J]. 生态学杂志, 2017, 36(4): 7. (LI Y L, WU S W, LIU Y, et al Hyper-spectral estimation of wheat biomass based on FastICA blind source separation method to remove soil disturbance [J]. Journal of Ecology, 2017, 36(4): 7.)

[25] 刘峥, 殷勇. 基于高光谱技术的香肠亚硝酸盐快速检测方法[J]. 食品与机械, 2019, 35(5): 78-82. (LIU Z, YIN Y. A rapid detection method for nitrite in sausage based on hyper-spectral technology [J]. Food and Machinery, 2019, 35(5): 78-82.)

[26] 田明璐. 西北地区冬小麦生长状况高光谱遥感监测研究[D]. 杨凌: 西北农林科技大学, 2017. (TIAN M L. Hyper-spectral remote sensing monitoring of winter wheat growth in northwest China [D]. Yangling: Northwest University of Agriculture and Forestry Science and Technology, 2017.)

[27] 安德帅,徐丹丹,刘月,等. 高光谱与拟合多光谱植被指数反演武夷山亚高山草甸 LAI 的对比研究[J]. 生态科学, 2022, 41(5): 187-196. (AN D S, XU D D, LIU Y, et al. Comparative study on LAI retrieval of sub-alpine meadow in Wuyi Mountain by hyper-spectral and fitting multi-spectral vegetation index [J]. Ecological Science, 2022,41(5): 187-196.)

[28] 张鹏举. 结合昌马水库渗压计观测数据探讨一元线性回归法在坝体渗压分析中的应用[J]. 中国水能及电气化, 2015(4): 49-52. (ZHANG P J. Combined with the observation data of Changma Reservoir's osmometer, this paper discusses the application of one-dimensional linear regression method in the analysis of dam seepage pressure [J]. China Hydro-power and Electrification, 2015 (4): 49-52.)

[29] 杨建坤,熊坚,徐东霞. 基于支持向量机的车辆换道行为识别仿真研究[J]. 贵州大学学报(自然科学版), 2020, 37(5): 120-124. (YANG J K, XIONG J, XU D X. Simulation study on recognition of vehicle lane changing behavior based on support vector machine [J]. Journal of Guizhou University (Natural Science Edition), 2020,37(5): 120-124.)

[30] 侯启真, 李泽, 姬雨初, 等. 基于 CS-BPNN 算法的飞机客舱 PMV 指标预测[J]. 计算机仿真, 2020, 37(3): 52-55, 99. (HOU Q Z, LI Z, JI Y C, et al. Prediction of aircraft cabin PMV index based on CS-BPNN algorithm [J]. Computer Simulation, 2020,37(3): 52-55, 99.)

[31] 赵翊含, 侯蒙京, 冯琦胜, 等. 基于 Landsat 8 和随机森林的青海门源天然草地上部生物量遥感估算[J]. 草业学报, 2022, 31(7): 1-14. (ZHAN Y H, HOU M J, FENG Q S, et al. Remote sensing estimation of above-ground biomass of natural grassland in Menyuan, Qinghai Province based on Landsat 8 and random forest [J]. Journal of Grass Industry, 2022, 31(7): 1-14.)

[32] 龚辉, 夏乔浪, 黄媛媛. 近红外漫反射技术快速检测黄酒酒醅中酒精度[J]. 酿酒科技, 2021(7): 125-129. (GONG H, XIA Q L, HUANG Y Y. Rapid detection of alcohol content in fermented grains of yellow rice wine by near-infrared diffuse reflectance technology [J]. Brewing Technology, 2021(7): 125-129.)

[33] 张丽强, 丛伟, 董罡, 等. 基于多元线性回归的单相电表相别判断方法[J]. 电力自动化设备, 2020, 40(5): 144-156, 187. (ZHANG L Q, CONG W, DONG G, et al. Phase discrimination method for single-phase ammeters based on multiple linear regression [J]. Power Automation Equipment, 2020, 40(5): 144-156, 187.)

[34] 曹中盛, 李艳大, 黄俊宝, 等. 监测花生叶面积指数和地上部生物量的最优植被指数及适宜波段带宽[J]. 中国油料作物学报, 2022, 44(6): 1320-1328. (CAO Z S, LI Y D, HUANG J B, et al. Monitoring peanut leaf area index and AGB's optimal vegetation index and suitable bandwidth [J]. Chinese Journal of Oil Crops, 2022, 44(6): 1320-1328.)

[35] 栗方亮, 孔庆波, 张青, 等. 瑯溪蜜柚叶片氮素含量多种高光谱估算模型对比研究[J]. 果树学报, 2022, 39(5): 882-891. (LI F L, KONG Q B, ZHANG Q, et al. Comparison of multiple hyper-spectral estimation models for nitrogen content in Guanxi honey pomelo leaves [J]. Journal of Fruit Trees, 2022, 39(5): 882-891.)

[36] 蔡鑫垚. 基于多源遥感数据的天然次生林地上生物量估测及随机森林偏差校正[D]. 哈尔滨: 东北林业大学, 2021. (CAI X Y. Estimation of biomass on natural secondary forest land and correction of random forest deviation based on multi-source remote sensing data [D]. Harbin: Northeast Forestry University, 2021.)

[37] 侯启真, 李泽, 姬雨初, 等. 基于 CS-BPNN 算法的飞机客舱 PMV 指标预测[J]. 计算机仿真, 2020, 37(3): 52-55, 99. (HOU Q Z, LI Z, JI Y C, et al. Prediction of aircraft cabin PMV index based on CS-BPNN algorithm [J]. Computer Simulation, 2020, 37(3): 52-55, 99.)

[38] LIU X P, LU M Z, CHAI Y Z, et al. A comprehensive framework for HSPF hydrological parameter sensitivity, optimization and uncertainty evaluation based on SVM surrogate model-A case study in Qinglong River watershed, China [J]. Environmental Modelling & Software, 2021, 143: 105126.

[39] 左雪燕, 崔丽娟, 李伟, 等. 基于高光谱数据的互花米草叶片功能性状反演[J]. 生态学报, 2021, 41(15): 6159-6169. (ZUO X Y, CUI L J, LI W, et al. Retrieval of functional traits of *Spartina alterniflora* leaves based on hyper-spectral data [J]. Journal of Ecology, 2021, 41(15): 6159-6169.)