

doi:10.6041/j.issn.1000-1298.2022.07.021

基于 XGBoost – Shapley 的玉米不同生育期 LAI 遥感估算

张宏鸣¹ 侯贵河¹ 孙志同¹ 杨欢瑜¹ 韩柯城¹ 韩文霆²

(1. 西北农林科技大学信息工程学院, 陕西杨凌 712100;

2. 西北农林科技大学机械与电子工程学院, 陕西杨凌 712100)

摘要: 针对当前快速准确获取叶面积指数 (Leaf area index, LAI) 时大部分遥感预测方法将光谱信息作为模型主要特征, 忽略时序变化特征的问题, 利用无人机搭载五通道多光谱相机获取研究区玉米不同生育期的影像数据, 基于该数据计算玉米相应生育期植被指数, 然后采用植被指数建立各生育期子模型, 采用 Shapley 理论计算子模型均方根误差对全生育期模型均方根误差的贡献度, 从而确定各子模型权重, 根据权重组合形成具有 LAI 时序变化特征的估算模型, 分别基于支持向量回归 (SVR)、多层感知机 (MLP)、随机森林 (RF) 和极限梯度提升树 (XGBoost) 算法构建组合估算模型。结果表明: 采用 Shapley 理论构建的组合 LAI 估算模型估算效果优于直接构建的全生育期 LAI 估算模型。相较于 SVR – Shapley、MLP – Shapley 以及 RF – Shapley 模型, XGBoost – Shapley 模型的估算效果最佳 (R^2 为 0.97, RMSE 为 0.021, RPD 为 6.9)。将最优模型 XGBoost – Shapley 应用于研究区 LAI 预测, 预测结果符合不同生育期玉米长势。本研究为大田玉米长势遥感监测提供了新的思路和方法。

关键词: 玉米; 叶面积指数; 极限梯度提升树; Shapley; 无人机遥感

中图分类号: TP391.4 文献标识码: A 文章编号: 1000-1298(2022)07-0208-09 OSID: 

Remote Sensing Estimation of Maize Leaf Area Index at Different Growth Periods Based on XGBoost – Shapley Algorithm

ZHANG Hongming¹ HOU Guihe¹ SUN Zhitong¹ YANG Huanyu¹ HAN Kecheng¹ HAN Wenting²

(1. College of Information Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China

2. College of Mechanical and Electronic Engineering, Northwest A&F University, Yangling, Shaanxi 712100, China)

Abstract: In view of the problem that most remote sensing prediction methods take spectral information as the main feature of the model and ignore the temporal variation characteristics when obtaining leaf area index (LAI) quickly and accurately, the UAV was equipped with a five channel multispectral camera to obtain the multispectral images of different growth periods of maize in the study area. The vegetation indices of maize in corresponding growth period were calculated based on the images. Then the sub models of each growth period were established by using the vegetation indices. The contribution of the root mean square error (RMSE) of each sub model to the RMSE of the whole growth period model was calculated based on Shapley theory. The weight of each sub model was given based on its contribution. The combination estimation model was built with LAI time-series variation characteristics according to the weight. And different combination models were built based on support vector regression (SVR) algorithm, multilayer perceptron (MLP), random forest (RF) algorithm and XGBoost algorithm for comparison. The results showed that the estimation effect of the combined LAI estimation model based on Shapley theory was better than that of the whole growth period LAI estimation model. Compared with other LAI estimation models (SVR – Shapley, MLP – Shapley and RF – Shapley), the XGBoost – Shapley model had the best estimation effect (R^2 was 0.97, RMSE was 0.021, RPD was 6.9). Thus the XGBoost – Shapley model was applied to LAI prediction in the study area. The research results showed that the LAI change rate in different growth periods were different, and the prediction results accorded with the growth trend of maize in different growth periods. The research result can provide a method for remote sensing monitoring of field maize growth.

Key words: maize; leaf area index; XGBoost; Shapley; UAV remote sensing

0 引言

叶面积指数 (Leaf area index, LAI) 量化了单位地面面积的叶面积,是控制与地表植被相关的生物和物理过程的一个重要参数^[1],和光合作用、蒸腾作用等生物物理过程有密切联系^[2]。快速、准确地获取作物 LAI 可以为现代农业生产提供科学指导。

传统 LAI 测量方法为直接测量法,该方法用时长、工作量大,会对农作物造成一定的破坏^[3]且不能实时和大面积获取。自 20 世纪 70 年代遥感技术在资源环境领域广泛应用,卫星遥感方法成为农业监测领域的一项有利手段^[4-5]。通过遥感影像可以快速获取大面积的地物信息,同时作物不同生长状况间的差异也可以体现在光谱影像的波段信息中^[6],但受空间分辨率限制,传统卫星遥感影像不能准确反映小尺度(田块、样地、植株)农作物长势情况^[7]。随着无人机(Unmanned aerial vehicle, UAV)相关技术的成熟,无人机遥感具有低空获取、搭载传感器空间分辨率较高的特点,无人机遥感手段既能有效弥补卫星遥感空间分辨率不高的缺点,又可在代价较小的野外工作条件下,即时、快速地获取较大面积无损可靠的农作物长势信息^[8]。

LAI 可以简单通过与冠层反射系数或植被指数(Vegetation index, VIs)的经验关系进行估算^[9],但传统 LAI 经验模型的估算精度不理想。随着计算机领域机器学习方法的出现,众多学者引入机器学习方法构建 LAI 估算模型,使得 LAI 估算模型的精度得到提高^[10-11]。近年来,许多学者针对 LAI 估算从模型算法角度进行了深入研究^[12-15],结果表明,选择合适的模型算法可以使 LAI 估算模型精度得到提高,但未深入探讨数据特征信息对模型的影响。OTHMAN 等^[16]研究发现从陆地卫星 7 号获得的多光谱数据可以估算生长在黏土和砂质土壤上的山核桃 LAI,李鑫川等^[17]研究了分段 LAI 的敏感植被指数,得出在反演不同植被 LAI 时选择最佳植被指数可以显著提高 LAI 模型反演精度,YANG 等^[18]通过利用不同的植被指数预测夏玉米 LAI 发现 EVI 作为参数的预测精度高于 NDVI 作为参数的预测模型,陶慧林等^[19]基于无人机高光谱影像进行冬小麦 LAI 估测,得出在冬小麦挑旗期、开花期、灌浆期和 LAI 相关性最高的 3 种光谱参数,均证明选择正确数据特征信息可以提高 LAI 估算模型的性能。植被指数等特征信息与 LAI 具有密切联系,同时作物 LAI 变化也具有明显的时序特征,且不同作物的时序变化各不相同^[19-20],已有学者利用历史 LAI 作为输入通过深度学习方法对未来 LAI 进行预测^[21],但

训练样本的稀缺性限制了深度学习在 LAI 预测方面的应用。时序特征信息很难像植被指数等参数被量化为数据信息,导致鲜有探讨将作物 LAI 时序特征信息引入 LAI 估算模型的影响。因此,设计一种策略将 LAI 时序变化特征引入模型训练是优化 LAI 估算模型的一个新思路。

针对以上问题,本研究将分别基于玉米不同生育期构建 LAI 估算子模型,利用 Shapley 策略计算子模型均方根误差对全生育期模型均方根误差的贡献度来确定各子模型权重,然后组合得到具有 LAI 时序变化特征的组合 LAI 估算模型,同时基于不同机器学习算法构建组合模型进行对比,以期实现对不同生育期玉米 LAI 的较高精度估算,为玉米作物的长势监测提供技术参考。

1 材料与方法

1.1 数据获取与处理

夏玉米 LAI 采集实验在内蒙古自治区鄂尔多斯市达拉特旗昭君镇一实验农场进行,研究区中心经纬度坐标为(40°25'N,109°36'E)。研究区所处的昭君镇位于达拉特旗中西部,地貌呈丘陵、沙漠、滩涂状态,该区域属温带大陆性气候,冬寒夏热,降水量少且集中,年均日照时数 3 100 h,年均降水量 300 mm 左右,降雨主要集中在 7—9 月,适宜夏玉米种植。研究区面积约 1.13 hm²,如图 1 所示,研究区域形状为圆形,均匀划成 5 个大小一致的扇形区域,各区域按图 1 进行编号,其中在区域 1 和区域 4 进行水分胁迫实验,且区域 1 水分胁迫程度较弱。各扇形区域内取 3 个 4 m × 4 m 的实验样方用以采集夏玉米 LAI 实测值,研究区共有 15 个小样方区,根据各样方区的地理坐标信息制作样方区的 XML 像控点文件。

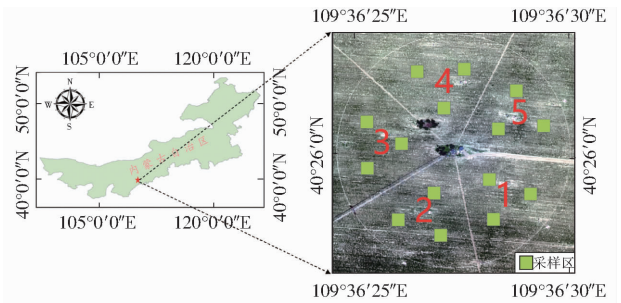


图 1 研究区位置和玉米田无人机影像

Fig. 1 Location of study area and UAV image of maize field

分别于 2018、2019 年 6—8 月进行地面数据获取工作,每年采集 9 次,结合研究区玉米作物生长特点,将研究区域玉米生长划分为 4 个生育期:苗期—拔节期、拔节—抽穗期、抽穗—乳熟期、乳熟—成熟期。

LAI 实测值采集时 4 次取冠下值,1 次取冠上值,每期每个样方重复采集 4 次取均值。夏玉米地面 LAI 实测值使用美国 LI-COR 公司 LAI-2200C 型植物冠层分析仪测量获取,LAI-2200C 型植物冠层分析仪利用“鱼眼”光学传感器(垂直视野 148°,水平视野 360°)测量树冠上、下 5 个角度的透射光线,利用植被树冠的辐射传输模型计算得到夏玉米实测 LAI。

多光谱影像数据通过无人机搭载多光谱相机低空拍摄获取。本次实验采用六翼无人机,无人机轴距 900 mm。多光谱相机为五波段多光谱相机 RedEdge,主要参数如表 1 所示,多光谱相机的 5 个光谱带分别为蓝(475 nm)、绿(560 nm)、红(668 nm)、近红外(840 nm)和红边(717 nm)。数据采集选取晴朗无云或少云无风的天气,时间选择太阳光辐射强度较为稳定的 11:00—13:00,无人机按照固定航线飞行,航向重叠度为 80%,旁向重叠度为 70%,同时采样区共设置 5 个地面控制点,用于影像拼接后的精度检测。已有研究工作表明空间分辨率在 1.00~5.69 cm 之间时,光谱信息对作物植株氮浓度、生物量等反演精度随空间分辨率提高而提升^[22],但影响较小^[23]。本文中数据采集实验从玉米叶片两指宽(约 3 cm)时开始,综合考虑玉米作物的形态特点,无人机采集效率以及多光谱影像空间分辨率较低导致植被指数饱和的因素,最终确定无人机飞行高度为 70 m,获取的多光谱影像空间分辨率为 4.71 cm。

表 1 RedEdge 多光谱相机主要参数
Tab.1 Main parameters of RedEdge multispectral camera

参数	数值/类型
尺寸(长×宽×高)/(cm×cm×cm)	12.1×6.6×4.6
质量/g	150
功耗/W	4(正常)、8(最大)
通道数	5
谱带	蓝、绿、红、近红外和红边
快捕获速度/(s·次 ⁻¹)	1(12 bit RAW)

无人机五波段多光谱影像数据处理过程分为光学辐射校正、影像拼接、几何校正、光谱提取 4 部分。光学辐射校正主要在多光谱影像拍摄阶段进行,光谱相机 RedEdge 配备了光学传感器和灰板两种校正设备,光学传感器主要校正影像拍摄过程中外界光线变化造成的影响,利用灰板的固定反射率可以校正影像的反射率。将采集的多幅多光谱原始影像按照日期索引导入 Pix4Dmapper 软件进行影像拼接,利用实时动态定位(Real-time kinematic, RTK)对影像进行精度检测,导入小图位置与姿态系统

(Position and orientation system, POS)数据,经过初始化、几何校正、三维模型构建等操作得到高清正射多光谱影像(分辨率为 4.71 cm)。基于拼接获取的多光谱影像进行光谱提取,利用 ENVI 软件将多光谱影像转换为五波段的灰度图,经过裁剪后得到实验区大小的影像数据,利用植被指数计算公式基于包含所有小样方区(图 1 中绿方块)的 XML 像控点文件,可以计算得到每个小样方区对应的各项植被指数^[24]。

将不同光谱波段按照一定的计算公式组合计算得到光谱参数,能有效降低背景环境信息的干扰^[25-26],这些光谱参数又称为植被指数。已有研究证明植被存在红光波段强吸收、近红外波段强反射的特性,且这两个光谱波段与植被 LAI 具有良好的相关性^[27]。本文重点选择红光、近红外和绿光 3 个光谱波段组合计算得到 8 种植被指数作为估算夏玉米 LAI 的估算因子^[15],8 种植被指数见表 2。

表 2 植被指数定义
Tab.2 Definition of vegetation index

植被指数	公式
归一化植被指数 (NDVI) ^[28]	$NDVI = \frac{R_{Nir} - R_{Red}}{R_{Nir} + R_{Red}}$
再归一化植被指数(RDVI) ^[29]	$RDVI = \frac{R_{Nir} - R_{Red}}{\sqrt{R_{Nir} + R_{Red}}}$
比值植被指数(RVI) ^[30]	$RVI = \frac{R_{Nir}}{R_{Red}}$
土壤调节植被指数(SAVI) ^[31]	$SAVI = 1.5 \frac{R_{Nir} - R_{Red}}{R_{Nir} + R_{Red} + 0.5}$
2 型增强植被指数(EVI2) ^[32]	$EVI2 = 2.5 \frac{R_{Nir} - R_{Red}}{R_{Nir} + 2.4R_{Red} + 1}$
修正型土壤调节植被指数(MSAVI) ^[33]	$MSAVI = \frac{2R_{Nir} + 1 - \sqrt{(2R_{Nir} + 1)^2 - 8(R_{Nir} - R_{Red})}}{2}$
三角植被指数(TVI) ^[34]	$TVI = \frac{120(R_{Nir} - R_{Green}) - 200(R_{Red} - R_{Green})}{2}$
优化土壤调节植被指数(OSAVI) ^[35]	$OSAVI = \frac{R_{Nir} - R_{Red}}{R_{Nir} + R_{Red} + 0.16}$

注:R_{Nir}为近红外波段平均反射率,R_{Red}为红光波段平均反射率,R_{Green}为绿光波段平均反射率。

1.2 研究方法

本研究的问题为回归类问题,近年来已有很多解决回归类问题的算法被提出,不同回归算法其实现原理不同导致它们在不同场景各有优劣。本研究中实验训练数据规模较小,初始特征维度(植被指数)较多,预测数据规模较大。针对数据规模较小的特点,选取在小规模数据时表现较优的支持向量回归(Support vector regression, SVR)算法作为对比

算法;针对初始数据特征维度较多的特点,选取基于决策树并行思想实现可以对特征进行重要性排序的随机森林(Random forest, RF)和同样基于决策树但选用串行思想实现的极限梯度提升树算法(XGBoost)作为对比算法;同时根据数据规模特点选择基础深度学习网络模型多层感知机(Multilayer perceptron, MLP)作为传统机器学习方法的对照组。

1.2.1 支持向量回归

支持向量回归是支持向量机算法用于处理回归问题的实现^[36-37],该算法在机器学习领域提出较早,其基本原理是在数据标签和特征之间寻找一个回归超平面,使得训练样本中所有数据尽可能地靠近该平面,SVR 可以通过更换合适核函数处理不同场景下的回归问题。

1.2.2 多层感知机

多层感知机是一种基础的深度学习网络模型^[38]。该网络由输入层、隐含层以及输出层组成,每一层都以神经元作为最小组成单位,把多个神经网络节点的信息映射到单一输出上。MLP 的隐含层可以包含多层网络,而输入、输出层则只有 1 层网络,且层与层之间均是全连接。针对本研究中的 LAI 预测问题,MLP 网络结构设计为:输入层包含 8 个神经元,4 层隐含层的神经元数分别为 512、128、64、32,输出层包含 1 个神经元。

1.2.3 随机森林

随机森林是 BREIMAN^[39]在 2001 年设计提出的一种基于决策树的机器学习方法,它可以处理分类问题和回归问题。RF 根据特征数量随机生成多棵相互独立的决策树,同时利用 bootstrap 重采样方法随机且有放回地从原始数据集中抽取训练样本作为每棵决策树的训练集,这些决策树进行平行训练,最终算法选取多棵树的平均值或权重最大的一类作为结果,这种投票策略称为袋装思想,RF 也因此具有较高的准确率,能处理具有高维特征的输入样本。

1.2.4 极限梯度提升树

极限梯度提升树算法是一种基于梯度提升决策树(Gradient boosting decision tree, GBDT)实现的集成学习方法^[40-41]。GBDT 的基学习器采用回归树,它的每一棵树拟合的是损失函数在前一棵树上的负梯度,最后将所有回归树的结果进行线性加权求和作为结果输出模型。XGBoost 是 GBDT 的高效实现,其基学习器包括分类树和回归树,本研究中的 LAI 估算问题属于回归问题,因此研究中采用的基学习器为回归树。相对于 GBDT 算法,XGBoost 在目标函数中显式地添加了正则化项,在更新基学习器时,GBDT 根据一阶导数迭代生成基学习器,而

XGBoost 不仅根据一阶导数还根据二阶导数更新基学习器,除此之外 XGBoost 算法在实现时还做了大量优化。一个含有 n 棵树的 XGBoost 模型可以表示为

$$\hat{y}_i = \sum_{n=1}^{N'} f_n(x_i) \tag{1}$$

式中 $\hat{y}_i$ ——第 i 个样本的 LAI 估算值
 x_i ——第 i 个样本中的输入变量
 f_n ——第 n 棵决策树的预测函数
 N' ——决策树数量

XGBoost 模型更新采用梯度提升(Gradient boosting)策略,通过更新损失函数的负梯度方向来更新模型,其优化模型可以表示为

$$O(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_n \Omega(f_n) \tag{2}$$

其中 $\Omega(f_n) = \gamma T + \frac{1}{2} \lambda \|\omega\|^2 \tag{3}$

式中 $l(\hat{y}_i, y_i)$ ——损失函数,即均方根误差
 $\Omega(f_n)$ ——正则化项
 γ ——模型复杂度
 T ——模型中树的叶子节点个数
 λ ——固定系数
 ω ——叶子节点量化权重向量

XGBoost 模型在保留训练好树模型的基础上,不断根据损失函数对预测函数求导,代入上一轮的预测函数中更新预测函数,通过迭代计算最终得到预测结果^[40]。

1.2.5 Shapley 值理论

Shapley 值是 SHAPLEY^[42]提出的合作博弈理论成果,其基本原理是贡献和所得相等,该理论被广泛应用到各领域中解决分配问题^[43-44]。LAI 估算组合模型中不同生育期单一算法模型的权重确定问题就是一个典型的分配问题,本研究通过计算单一模型估算误差的 Shapley 值和组合模型估算误差的 Shapley 值来确定各单一模型的权重,其过程主要包括 Shapley 值计算和权重计算两部分。其中 Shapley 值计算原理可以表示为

$$\phi_j(v) = \sum_{S \subseteq N-j} \frac{|S|! (|N| - |S| - 1)!}{|N|!} (v(S \cup \{j\}) - v(S)) \tag{4}$$

式中 j ——单一模型编号,即各生育期模型编号
 v ——模型估算均方根误差计算函数
 $\phi_j(v)$ ——第 j 个单一生育期模型 Shapley 值
 N ——RF 算法或 XGBoost 算法不同生育期子模型的集合
 S ——集合 N 中,除当前计算的第 j 个子模型

外所有子模型可能组合形成的集合
 $v(S)$ ——子模型集合 S 贡献的估算均方根误差
 $v(S \cup \{j\})$ ——子模型集合 S 加入子模型 j 后,贡献的估算均方根误差

组合模型中各子模型的权重确定依据的是各单一模型的估算误差,单一子模型的 Shapley 值越大,说明单一子模型对组合模型的误差贡献越大,其在组合模型中的权重则越小。单一子模型的权重可以定义为^[45]

$$w_j = \frac{1}{n' - 1} \frac{\phi - \phi_j}{\phi} \quad (j = 1, 2, \cdots, n') \quad (5)$$

式中 n' ——子模型数量

确定单一子模型的权重以后,则组合模型可以定义为

$$P = \sum_j^{n'} w_j P_j \quad (6)$$

式中 P ——组合模型估算的 LAI

P_j ——子模型估算的 LAI

1.3 评价方法

将样本数据按照玉米的 4 个生育期分为出苗—

拔节期、拔节—抽穗期、抽穗—乳熟期和乳熟—成熟期 4 个样本集,采用交叉验证的方式将 4 个样本集按照 7:3 的比例随机划分为训练集(315 条)和验证集(135 条)。采用决定系数 R^2 、均方根误差(Root mean square error, RMSE)和相对预测偏差(Relative predictive deviation, RPD)3 个指标对 LAI 估算模型进行评价^[46]。RPD 是评价模型预测能力的一个指标,它将模型的预测能力分为了 4 个等级^[47]:RPD 小于 1.5 时,模型不可用;RPD $\in [1.5, 2.0)$ 时,模型可以区分预测的高值和低值;RPD $\in [2.0, 2.5)$ 时,模型可以进行定量预测;RPD 大于等于 2.5 时,模型具有很好的定量预测能力。

2 结果与分析

2.1 植被指数相关性分析

利用光谱波段组合提取得到 8 种植被指数,再按照夏玉米的 4 个生育期对 8 种植被指数和实测 LAI 进行皮尔逊相关性分析。如表 3 所示,不同生育期下训练集和验证集中 LAI 和植被指数均在 $P < 0.01$ 水平上呈现极显著相关,相关系数均大于 0.58。

表 3 植被指数与 LAI 相关系数
Tab.3 Correlation coefficient between vegetation index and LAI

植被指数	出苗-拔节期		拔节-抽穗期		抽穗-乳熟期		乳熟-成熟期		全生育期	
	训练集	验证集	训练集	验证集	训练集	验证集	训练集	验证集	训练集	验证集
NDVI	0.853 **	0.852 **	0.750 **	0.750 **	0.723 **	0.721 **	0.840 **	0.843 **	0.722 **	0.849 **
RDVI	0.834 **	0.833 **	0.757 **	0.756 **	0.742 **	0.751 **	0.762 **	0.767 **	0.699 **	0.821 **
RVI	0.817 **	0.817 **	0.726 **	0.713 **	0.713 **	0.719 **	0.718 **	0.708 **	0.615 **	0.648 **
SAVI	0.837 **	0.835 **	0.756 **	0.772 **	0.742 **	0.744 **	0.739 **	0.741 **	0.695 **	0.817 **
EVI2	0.835 **	0.834 **	0.752 **	0.752 **	0.743 **	0.742 **	0.723 **	0.727 **	0.691 **	0.814 **
MSAVI	0.832 **	0.831 **	0.751 **	0.761 **	0.750 **	0.750 **	0.722 **	0.729 **	0.692 **	0.815 **
TVI	0.829 **	0.829 **	0.728 **	0.729 **	0.701 **	0.708 **	0.584 **	0.581 **	0.660 **	0.788 **
OSAVI	0.847 **	0.846 **	0.762 **	0.761 **	0.739 **	0.738 **	0.817 **	0.821 **	0.712 **	0.835 **

注: ** 表示在 $P < 0.01$ 水平上极显著相关。

由表 3 可知,夏玉米 8 种植被指数均和 LAI 呈极显著相关,且在不同生育期植被指数和 LAI 的相关性发生明显变化。通过皮尔逊相关性分析,将相关性最低的 RVI、TVI 两种植被指数排除,结合文献和 XGBoost 模型给出的特征得分,最终确定 NDVI、RDVI、EVI2 和 SAVI 共 4 种植被指数作为模型的特征输入。

2.2 LAI 时序变化

玉米生长过程中 LAI 变化规律为:缓慢增长-快速增长-缓慢下降,这是因为玉米的抽苗-拔节期和拔节-抽穗期是茎叶生长的营养生长阶段,在该阶段玉米的叶片逐渐变大;进入抽穗-乳熟期以后,玉米主要是果穗生长,此时玉米茎叶基本不再发育^[45]。

研究区各扇形区域夏玉米各生育期的 LAI 变化趋势均不相同,将研究区采集的玉米 LAI 数据集绘制玉米 LAI 时序变化图。

图 2 中区域 4 的 LAI 峰值明显小于其他 4 个区域的峰值,这是因为区域 4 在实验过程中进行了水分胁迫实验,玉米的长势受到影响。由图 2 可以看出,研究区玉米 LAI 曲线随时间整体呈现先增长达到峰值稳定,然后小幅度下降的变化趋势,LAI 曲线的变化基本符合玉米的生长规律。玉米 LAI 的变化规律随生育期的不同而不同,因此研究中分别基于玉米 4 个不同的生育期 LAI 数据训练建立 LAI 估算模型。

基于玉米不同生育期建立的 LAI 估算模型在特定的玉米生育期具有较高的 LAI 估算精度,但不同

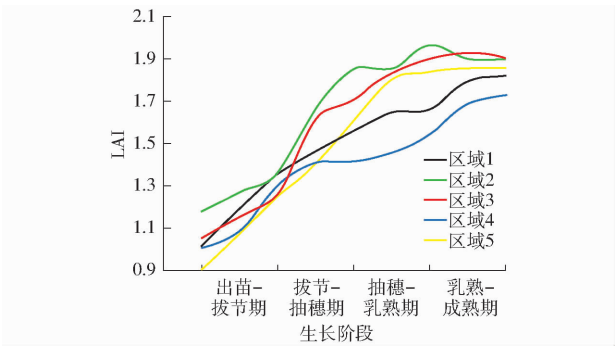


图 2 玉米 LAI 时序变化曲线

Fig.2 Temporal variation of maize LAI

生育期的子模型在玉米的全生长过程中并不具备良好的 LAI 估算精度,因此,将不同生育期建立的 LAI 估算子模型组合得到全生育期 LAI 估算模型(图 3),以使得该模型在玉米的全生育期均能取得较好的 LAI 估算效果。

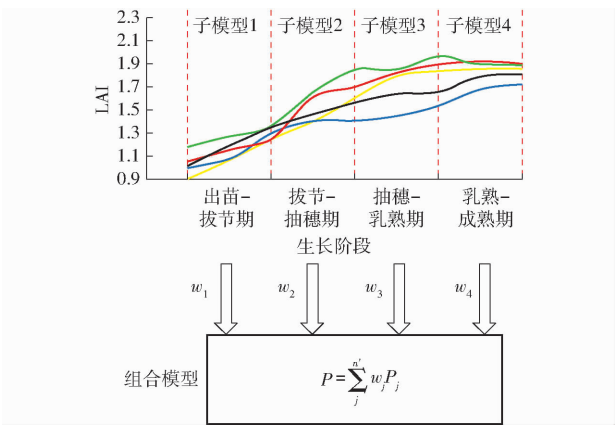


图 3 组合模型结构示意图

Fig.3 Structural diagram of combined model

基于 SVR、MLP、RF 和 XGBoost 算法分别以玉米各生育期子模型的估算均方根误差进行 Shapley 值计算,得到组合模型各生育期子模型的权重如表 4 所示。

表 4 玉米 LAI 估算子模型权重

Tab.4 Weights of sub maize LAI estimation models

模型	子模型 1	子模型 2	子模型 3	子模型 4
SVR - Shapley	0. 23	0. 26	0. 27	0. 24
MLP - Shapley	0. 22	0. 30	0. 32	0. 16
RF - Shapley	0. 12	0. 34	0. 36	0. 18
XGBoost - Shapley	0. 24	0. 29	0. 30	0. 17

根据组合模型中各个子模型的权重,SVR 组合模型、MLP 组合模型、RF 组合模型和 XGBoost 组合模型分别为

$$P_{\text{SVR-Shapley}} = 0. 23P_1 + 0. 26P_2 + 0. 27P_3 + 0. 24P_4 \tag{7}$$

$$P_{\text{MLP-Shapley}} = 0. 22P_1 + 0. 30P_2 + 0. 32P_3 + 0. 16P_4 \tag{8}$$

$$P_{\text{RF-Shapley}} = 0. 12P_1 + 0. 34P_2 + 0. 36P_3 + 0. 18P_4 \tag{9}$$

$$P_{\text{XGBoost-Shapley}} = 0. 24P_1 + 0. 29P_2 + 0. 30P_3 + 0. 17P_4 \tag{10}$$

式中 $P_1 \sim P_4$ 分别表示玉米出苗-拔节期、拔节-抽穗期、抽穗-乳熟期、乳熟-成熟期 4 个生育时期 LAI 估算子模型的估算值。

2.3 LAI 估算模型精度评价

分别利用 SVR、MLP、RF 和 XGBoost 算法在玉米的全生育期数据集上训练构建 LAI 估算模型,将这 4 个模型作为组合 LAI 估算模型的对照组进行精度评价。

表 5 中,SVR - Shapley、MLP - Shapley、RF - Shapley 以及 XGBoost - Shapley 为组合模型。其中 XGBoost - Shapley 模型的决定系数最高,为 0.97。相较于直接在玉米全生育期数据集上建立的 LAI 估算模型,基于各单生育期利用 Shapley 值建立的组合 LAI 估算模型决定系数均有提高:SVR - Shapley 模型相较于 SVR 模型, R^2 提高 0.1, MLP - Shapley 模型相较于 MLP 模型, R^2 提高 0.21, RF - Shapley 模型相较于 RF 模型, R^2 提高 0.08, XGBoost - Shapley 模型相较于 XGBoost 模型, R^2 提高 0.06。这一结果说明基于各单生育期利用 Shapley 值建立组合模型的策略可以使得模型拟合度提高,其中 XGBoost - Shapley 模型的拟合度最好。

表 5 玉米 LAI 估算模型精度对比

Tab.5 Comparison of LAI estimation models' precision of maize

模型	R^2	RMSE	RPD
SVR	0. 72	0. 187 1	1. 9
MLR	0. 68	0. 198 5	1. 8
RF	0. 86	0. 090 1	3. 8
XGBoost	0. 91	0. 083 2	4. 6
SVR - Shapley	0. 82	0. 123 3	2. 4
MLP - Shapley	0. 89	0. 093 1	4. 1
RF - Shapley	0. 94	0. 053 0	5. 1
XGBoost - Shapley	0. 97	0. 021 0	6. 9

XGBoost - Shapley 模型的 RMSE 为 0.021,其他模型的 RMSE 分别为 0.187 1 (SVR)、0.198 5 (MLP)、0.090 1 (RF)、0.083 2 (XGBoost)、0.123 3 (SVR - Shapley)、0.093 1 (MLP - Shapley)、0.053 (RF - Shapley),XGBoost - Shapley 模型的 RMSE 最小,同时,采用了 Shapley 值组合策略的模型 RMSE 均有所降低:SVR - Shapley 模型相较于 SVR 模型 RMSE 降低 0.053 8,MLP - Shapley 模型相较于 MLP 模型 RMSE 降低 0.105 4,RF - Shapley 模型相较于

RF 模型 RMSE 降低 0.037 1,XGBoost-Shapley 模型相较于 XGBoost 模型 RMSE 降低 0.062 2。结果表明,XGBoost-Shapley 模型在 4 个模型中 LAI 估算精度最高,Shapley 值组合模型 LAI 估算精度高于其他模型。

采用相对预测偏差评价 LAI 估算模型的综合性能。由表 5 发现,8 个模型中 XGBoost-Shapley 模型的 RPD 最高,为 6.9,说明 XGBoost-Shapley 模型的综合预测性能最佳。RF、XGBoost、MLP-Shapley、RF-Shapley 4 个模型的 RPD 均高于 2.5,说明这 4 个 LAI 估算模型均有较好的定量预测

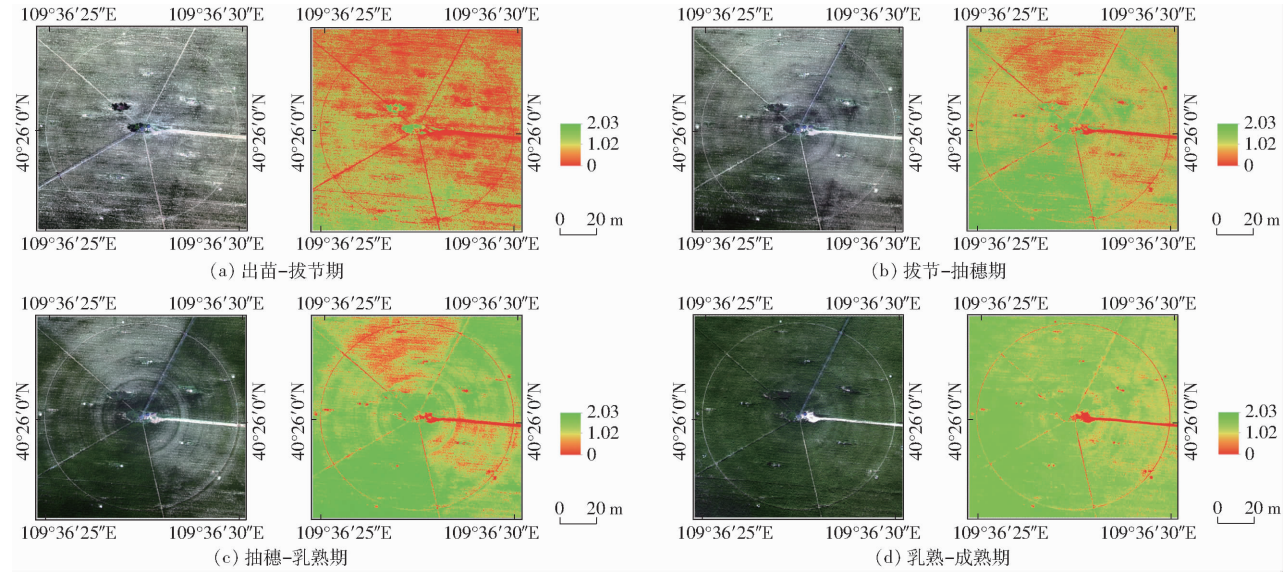


图 4 研究区玉米 LAI 估算结果

Fig.4 LAI estimation results of maize in study area

从 LAI 空间分布来看,LAI 估算结果可以反映出研究区域中不同地块的玉米生长情况,实验前期对区域 1 和区域 4 进行了不同程度的水分胁迫,这使得不同区域的玉米生长状况不一致,在 LAI 估算图上可以明显看出这一差异。

按 LAI 时间变化来看,研究区玉米 4 个生育期 LAI 均值分别为 1.09、1.34、1.49、1.59,4 个生育期的 LAI 估算结果按时间顺序逐渐递增,且增速逐渐变慢,这一特征符合玉米真实生长过程,在最后一个生育期研究区不同地块玉米 LAI 差值缩小,逐渐到达极值,因为前期水分胁迫实验主要影响玉米前期和中期的叶片生长速度,随着时间变长这一影响逐渐变小。综合比较可见光影像和 LAI 估算结果可得:XGBoost-Shapley 模型的 LAI 估算结果可以较好地反映不同生育期玉米的真实长势。

3 讨论

3.1 各 LAI 估算模型的差异

对不同玉米 LAI 估算模型进行对比分析,结

能力。

总体而言,基于 Shapley 值建立的多生育期组合 LAI 估算模型相较于直接建立的 LAI 估算模型在决定系数、均方根误差以及相对预测偏差 3 项指标上均有一定的优化,其中 XGBoost-Shapley 模型相较于其他模型表现最优。

2.4 研究区 LAI 估算

利用 LAI 估算表现最优的 XGBoost-Shapley 模型对研究区的玉米按不同生育期进行 LAI 估算(图 4),其中左图为可见光图像,右图为 LAI 估算结果图像。

合玉米 LAI 曲线时序变化特点可以发现,玉米生长过程中 LAI 变化随生育期的不同呈现不同的变化速率,且不同 LAI 估算模型所关注的特征以及采用的子模型权重策略各有差异:SVR-Shapley 模型各子模型的权重几乎相等,这一现象可能因为 SVR 算法寻找最优超平面使得所有数据特征离这个平面距离最近,实验中各个生育期数据样本数量相等,使得 SVR-Shapley 模型将每个生育期子模型贡献度视为相同;MLP-Shapley 模型和 XGBoost-Shapley 模型的子模型权重分配方式几乎相同,但 XGBoost-Shapley 模型的表现却明显优于 MLP-Shapley 模型,这可能因为 MLP 作为一种简单神经网络,参数量和所需训练数据量都较大,在本研究数据规模偏小的场景下表现较差;RF-Shapley 和 XGBoost-Shapley 子模型权重分布相似,但是 RF 拔节-抽穗期、抽穗-乳熟期子模型贡献度明显高于其他两个生育期,这可能导致 RF 组合模型对其他两个生育期数据拟合度欠佳,同时由于 XGBoost 算法比 RF 算法存在一定回归优

势^[15],最终使得 XGBoost - Shapley 模型的 LAI 估算效果更好。

3.2 Shapley 组合模型的时序特征

对研究区不同生育期玉米 LAI 估算结果对比分析,图 4 拔节-抽穗期、抽穗-乳熟期两个生育期的 LAI 估算结果中,受到水分胁迫的区域 1、4 的 LAI 估值明显低于其他 3 个区域,结合玉米 LAI 变化曲线(图 2)发现,拔节-抽穗期、抽穗-乳熟期区域 1、4 的 LAI 曲线增长速度明显小于其他 3 个区域,这一现象说明区域 1、4 的玉米叶片生长受水分胁迫影响,生长变慢,图 4 中拔节-抽穗期、抽穗-乳熟期区域 1 的 LAI 估值大于区域 4,这是因为区域 1 比区域 4 受到的水分胁迫程度轻。图 4 中出苗-拔节期、乳熟-成熟期两个生育期,区域 1、4 的 LAI 估值和其他 3 个区域的估值差异较小,这一现象的原因是:在出苗-拔节期玉米幼苗整体 LAI 偏低,且水分胁迫实验处于初始阶段,各区域间 LAI 差异较小,而在乳熟-成熟期,区域 2、3、5 玉米完成营养生长转入生殖生长阶段,叶片生长变为果实生长,区域 1、4 玉米由于水分胁迫尚未完成营养生长,叶片继续生长,导致各区域间 LAI 差值逐渐缩小。表 4 中基于 Shapley 策略的组合模型在拔节-抽穗期、抽穗-乳熟期的权重均大于其他两个生育期,表明 Shapley 组合模型拔节-抽穗期、抽穗-乳熟期的 LAI 估算子模型相较于其他两个生育期更能体现玉米 LAI 生长变化差异,权重分配结果符合研究区玉米各生育期的生长特点。

3.3 存在的不足

传统基于光谱波段的 LAI 估算模型重点关注可以反映作物叶片信息的波段或植被指数,忽略 LAI 在不同时间的变化特征。本文在传统模型的基础上,分生育期构建模型,再利用 Shapley 策略进行融合,从而将 LAI 的时序变化特征引入模型,使得 LAI 估算模型的估算精度进一步提高。同时,研究中虽然利用 Shapley 策略对玉米的不同生育期子模型进行了权重计算使得 LAI 估算模型得到优化,但实验过程缺乏对作物生长知识的考虑,今后的研究可以根据不同作物的生长特点选择合适的权重计算方法,进一步提高 LAI 估算模型的泛化能力。此外,本文中通过参考前人研究工作^[22-23]结合玉米作物的形态特点确定无人机的飞行高度,但前人实验对象多为小麦、马铃薯等作物,由于不同作物间的植株形态、光谱信息存在一定的差异,前人的研究结论在玉米作物上的适用性有待进一步论证,后续研究可以从无人机高度的角度研究不同空间分辨率多光谱数据对作物 LAI 反演精度的影响。

4 结论

- (1)基于 Shapley 理论引入 LAI 时序变化特征的组合玉米 LAI 估算模型估算效果优于直接建立的全生育期 LAI 估算模型。
- (2)相较于 SVR - Shapley、MLP - Shapley 以及 RF - Shapley 模型,XGBoost - Shapley 模型表现最佳, R^2 为 0.97, RMSE 为 0.021, RPD 为 6.9。

参 考 文 献

[1] CHEN J M, CIHLAR J. Retrieving leaf area index of boreal conifer forests using Landsat TM images[J]. Remote Sensing of Environment, 1996, 55(2): 153 - 162.

[2] VEROUSTRAETE F, PATYN M R B. Estimating net ecosystem exchange of carbon using the normalized difference vegetation index and an ecosystem model[J]. Remote Sensing of Environment, 1996, 58(1): 115 - 130.

[3] 高林, 李长春, 王宝山, 等. 基于多源遥感数据的大豆叶面积指数估测精度对比[J]. 应用生态学报, 2016, 27(1): 191 - 200.

GAO Lin, LI Changchun, WANG Baoshan, et al. Comparison of precision in retrieving soybean leaf area index based on multi-source remote sensing data[J]. Chinese Journal of Applied Ecology, 2016, 27(1): 191 - 200. (in Chinese)

[4] BASTIAANSEN W, MOLDEN D J, MAKIN I W. Remote sensing for irrigated agriculture: examples from research and possible applications[J]. Agricultural Water Management, 2000, 46(2): 137 - 155.

[5] DORAISWAMY P C, HATFIELD J L, JACKSON T J, et al. Crop condition and yield simulations using Landsat and MODIS [J]. Remote Sensing of Environment, 2004, 92(4): 548 - 559.

[6] 邹文涛, 吴炳方, 张森, 等. 农作物长势综合监测——以印度为例[J]. 遥感学报, 2015, 19(4): 539 - 549.

ZOU Wentao, WU Bingfang, ZHANG Miao, et al. Synthetic method for crop condition analysis: a case study in India[J]. Journal of Remote Sensing, 2015, 19(4): 539 - 549. (in Chinese)

[7] DUVEILLER G, DEFOURNY P. A conceptual framework to define the spatial resolution requirements for agricultural monitoring using remote sensing[J]. Remote Sensing of Environment, 2010, 114(11): 2637 - 2650.

[8] 刘忠, 万伟, 黄晋宇, 等. 基于无人机遥感的农作物长势关键参数反演研究进展[J]. 农业工程学报, 2018, 34(24): 60 - 71.

LIU Zhong, WAN Wei, HUANG Jinyu, et al. Progress on key parameters inversion of crop growth based on unmanned aerial vehicle remote sensing[J]. Transactions of the CSAE, 2018, 34(24): 60 - 71. (in Chinese)

[9] KIMURA R, OKADA S, MIURA H, et al. Relationships among the leaf area index, moisture availability, and spectral reflectance in an upland rice field[J]. Agricultural Water Management, 2004, 69(2): 83 - 100.

[10] 陶惠林, 冯海宽, 杨贵军, 等. 基于无人机成像高光谱影像的冬小麦 LAI 估测[J]. 农业机械学报, 2020, 51(1): 176 - 187.

TAO Huilin, FENG Haikuan, YANG Guijun, et al. Leaf area index estimation of winter wheat based on UAV imaging

- hyperspectral imagery[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(1): 176–187. (in Chinese)
- [11] 马怡茹, 吕新, 易翔, 等. 基于机器学习的棉花叶面积指数监测[J]. 农业工程学报, 2021, 37(13): 152–162.
MA Yiru, LÜ Xin, YI Xiang, et al. Monitoring of cotton leaf area index using machine learning[J]. Transactions of the CSAE, 2021, 37(13): 152–162. (in Chinese)
- [12] JIANG F, SMITH A R, KUTIA M, et al. A modified KNN method for mapping the leaf area index in arid and semi-arid areas of China[J]. Remote Sensing, 2020, 12(11): 1884.
- [13] 张春兰, 杨贵军, 李贺丽, 等. 基于随机森林算法的冬小麦叶面积指数遥感反演研究[J]. 中国农业科学, 2018, 51(5): 855–867.
ZHANG Chunlan, YANG Guijun, LI Heli, et al. Remote sensing inversion of leaf area index of winter wheat based on random forest algorithm[J]. Scientia Agricultura Sinica, 2018, 51(5): 855–867. (in Chinese)
- [14] LÜ Xin. Estimation of cotton leaf area index(LAI) based on spectral transformation and vegetation index[J]. Remote Sensing, 2021, 14(1): 136.
- [15] 张宏鸣, 刘雯, 韩文霆, 等. 基于梯度提升树算法的夏玉米叶面积指数反演[J]. 农业机械学报, 2019, 50(5): 251–259.
ZHANG Hongming, LIU Wen, HAN Wenting, et al. Inversion of summer maize leaf area index based on gradient boosting decision tree algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 251–259. (in Chinese)
- [16] OTHMAN Y A, HILAIRE R S. Using multispectral data from Landsat ETM+ to estimate leaf area index of pecan orchards[J]. Fresenius Environmental Bulletin, 2021, 30(3): 2613–2618.
- [17] 李鑫川, 徐新刚, 鲍艳松, 等. 基于分段方式选择敏感植被指数的冬小麦叶面积指数遥感反演[J]. 中国农业科学, 2012, 45(17): 3486–3496.
LI Xinchuan, XU Xin'gang, BAO Yansong, et al. Retrieving LAI of winter wheat based on sensitive vegetation index by the segmentation method[J]. Scientia Agricultura Sinica, 2012, 45(17): 3486–3496. (in Chinese)
- [18] YANG Fei, SUN Jiulin, FANG Hongliang, et al. Comparison of different methods for corn LAI estimation over northeastern China[J]. International Journal of Applied Earth Observation Geoinformation, 2012, 18: 462–471.
- [19] 陶惠林, 徐良骥, 冯海宽, 等. 基于无人机高光谱长势指标的冬小麦长势监测[J]. 农业机械学报, 2020, 51(2): 180–191.
TAO Huilin, XU Liangji, FENG Haikuan, et al. Monitoring of winter wheat growth based on UAV hyperspectral growth index[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 180–191. (in Chinese)
- [20] 王鹏新, 陈弛, 张树誉, 等. 基于 LAI 和 VTCI 及 Copula 函数的冬小麦单产估测[J]. 农业机械学报, 2021, 52(10): 255–263.
WANG Pengxin, CHEN Chi, ZHANG Shuyu, et al. Winter wheat yield estimation based on Copula function and remotely sensed LAI and VTCI[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(10): 255–263. (in Chinese)
- [21] 龙泽昊, 秦其明, 张添源, 等. 基于长短期记忆网络的冬小麦连续时序叶面积指数预测[J]. 光谱学与光谱分析, 2020, 40(3): 898–904.
LONG Zehao, QIN Qiming, ZHANG Tianyuan, et al. Prediction of continuous time series leaf area index based on long short-term memory network: a case study of winter wheat[J]. Spectroscopy and Spectral Analysis, 2020, 40(3): 898–904. (in Chinese)
- [22] 刘杨, 冯海宽, 孙乾, 等. 不同分辨率无人机数码影像的马铃薯地上生物量估算研究[J]. 光谱学与光谱分析, 2021, 41(5): 1470–1476.
LIU Yang, FENG Haikuan, SUN Qian, et al. Estimation study of above ground biomass in potato based on UAV digital images with different resolutions[J]. Spectroscopy and Spectral Analysis, 2021, 41(5): 1470–1476. (in Chinese)
- [23] 贾丹, 陈鹏飞. 低空无人机影像分辨率对冬小麦氮浓度反演的影响[J]. 农业机械学报, 2020, 51(7): 164–169.
JIA Dan, CHEN Pengfei. Effect of low-altitude UAV image resolution on inversion of winter wheat nitrogen concentration[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(7): 164–169. (in Chinese)
- [24] HUETE A, JUSTICE C, LEEUWEN W V. MODIS vegetation index(MOD13)[J]. Algorithm Theoretical Basis Document, 1991, 3: 131.
- [25] QI J G, CHEHBOUNI A R, HUETE A R, et al. A modified soil adjusted vegetation index[J]. Remote Sensing of Environment, 1994, 48(2): 119–126.
- [26] BROGE N H, LEBLANC E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density[J]. Remote Sensing of Environment, 2001, 76(2): 156–172.
- [27] RONDEAUX G, STEVEN M, BARET F. Optimization of soil-adjusted vegetation indices[J]. Remote Sensing of Environment, 1996, 55(2): 95–107.
- [28] 张宏鸣, 谭紫薇, 韩文霆, 等. 基于无人机遥感的玉米株高提取方法[J]. 农业机械学报, 2019, 50(5): 241–250.
ZHANG Hongming, TAN Ziwei, HAN Wenting, et al. Extraction method of maize height based on UAV remote sensing[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(5): 241–250. (in Chinese)
- [29] 张东彦, 刘谔源, 宋晓宇, 等. 应用近地成像高光谱估算玉米叶绿素含量[J]. 光谱学与光谱分析, 2011, 31(3): 771–775.
ZHANG Dongyan, LIU Rongyuan, SONG Xiaoyu, et al. A field-based pushbroom imaging spectrometer for estimation chlorophyll content of maize[J]. Spectroscopy and Spectral Analysis, 2011, 31(3): 771–775. (in Chinese)
- [30] LIANG Liang, DI Liping, ZHANG Lianpeng, et al. Estimation of crop LAI using hyperspectral vegetation indices and a hybrid inversion method[J]. Remote Sensing of Environment, 2015, 165: 123–134.
- [31] 刘伟东, 项月琴, 郑兰芬, 等. 高光谱数据与水稻叶面积指数及叶绿素密度的相关分析[J]. 遥感学报, 2000, 21(4): 279–283.
LIU Weidong, XIANG Yueqin, ZHENG Lanfen, et al. Relationships between rice LAI, CHL D and hyperspectral data[J]. National Remote Sensing Bulletin, 2000, 21(4): 279–283. (in Chinese)

YUAN Wei, ZHOU Tian, XI Zongshun, et al. A multi-scale adaptive deep learning network for remote sensing semantic segmentation[J]. Journal of Geomatics Science and Technology, 2020, 37(6): 581–588. (in Chinese)

[18] LI H, LI Y, ZHANG G, et al. Global and local contrastive self-supervised learning for semantic segmentation of HR remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1–14.

[19] 刘文萍, 赵磊, 周焱, 等. 基于深度学习的无人机土地覆盖图像分割方法[J]. 农业机械学报, 2020, 51(2): 221–229. LIU Wenping, ZHAO Lei, ZHOU Yan, et al. Deep learning based unmanned aerial vehicle landcover image segmentation method[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(2): 221–229. (in Chinese)

[20] WEI X, GUO Y, XIN G, et al. A new semantic segmentation model for remote sensing images[C] // IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2017: 1776–1779.

[21] LASHGARI E, LIANG D, MAOZ U. Data augmentation for deep-learning-based electroencephalography[J]. Journal of Neuroscience Methods, 2020, 346: 108885.

[22] MA N, ZHANG X, SUN J. Funnel activation for visual recognition[J]. arXiv preprint arXiv, 2020: 2007.11824.

[23] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision[C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 2818–2826.

[24] DING X, GUO Y, DING G, et al. ACNet: strengthening the kernel skeletons for powerful CNN via asymmetric convolution blocks[C] // Proceeding of the IEEE/CVF International Conference on Computer Vision (ICCV), 2019: 1911–1920.

[25] 柏宇阳, 朱福珍. 改进的残差卷积神经网络遥感图像超分辨率重建[J]. 黑龙江大学自然科学学报, 2020, 37(3): 372–378. BAI Yuyang, ZHU Fuzhen. An improved remote sensing images super-resolution method based on residual convolution neural network[J]. Journal of Natural Science of Heilongjiang University, 2020, 37(3): 372–378. (in Chinese)

[26] HE K, ZHANG X, REN S, et al. Delving deep into rectifiers: surpassing human-level performance on imageNet classification [C] // Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2015: 1026–1034.

(上接第 216 页)

[32] ROUSEL J, HAAS R, SCHELL J. Monitoring vegetation systems in the great plains with ERTS[J]. NASA Special Publication, 1973, 351: 309–317.

[33] GILABERT M A, GONZALEZ-PIPUERAS J, GARCE-HARO F J, et al. A generalized soil-adjusted vegetation index[J]. Remote Sensing of Environment, 2002, 82(2): 303–310.

[34] BIRTH G S, MCVEY G R. Measuring the color of growing turf with a reflectance spectrophotometer[J]. Agronomy Journal, 1968, 60(6): 640–643.

[35] HUETE A R. A soil-adjusted vegetation index(SAVI)[J]. Remote Sensing of Environment, 1988, 25(3): 295–309.

[36] SAUNDERS C, STITSON M O, WESTON J, et al. Support vector machine[J]. Computer Science, 2002, 1(4): 1–28.

[37] BASAK D, SRIMANTA P, PATRANBIS D C, et al. Support vector regression[J]. Neural Information Processing—Letters and Reviews, 2007, 11: 67–80.

[38] GARDNER M W, DORLING S R. Artificial neural networks(the multilayer perceptron)—a review of applications in the atmospheric sciences[J]. Atmospheric Environment, 1998, 32(14): 2627–2636.

[39] BREIMAN L. Random forest[J]. Machine Learning, 2001, 45: 5–32.

[40] FRIEDMAN J H. Greedy function approximation: a gradient boosting machine[J]. Annals of Statistics, 2001, 29(5): 1189–1232.

[41] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C] // The 22nd ACM SIGKDD International Conference, 2016.

[42] SHAPLEY L S. A value for n-person games[J]. Contributions to the Theory of Games, 1953, 28: 307–319.

[43] AN Q, WEN Y, DING T, et al. Resource sharing and payoff allocation in a three-stage system: integrating network DEA with the Shapley value method[J]. Omega, 2019, 85: 16–25.

[44] 陈志鼎, 程丛. 基于改进 Shapley 值的 PPP 项目利益分配研究[J]. 建筑经济, 2020, 41(5): 40–44. CHEN Zhiding, CHENG Cong. Research on profit distribution of PPP project based on revised Shapley value[J]. Construction Economy, 2020, 41(5): 40–44. (in Chinese)

[45] 王鹏新, 乔琛, 李俐, 等. 基于 Shapley 值组合预测的玉米单产估测[J]. 农业机械学报, 2021, 52(9): 221–229. WANG Pengxin, QIAO Chen, LI Li, et al. Estimation of maize yield based on Shapley value combination forecasting[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(9): 221–229. (in Chinese)

[46] 张宏鸣, 陈丽君, 刘雯, 等. 基于 Stacking 集成学习的夏玉米覆盖度估测[J]. 农业机械学报, 2021, 52(7): 1–13. ZHANG Hongming, CHEN Lijun, LIU Wen, et al. Estimation of summer maize fractional vegetation coverage based on Stacking ensemble learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(7): 1–13. (in Chinese)

[47] SAEYS W, MOUAZEN A M, RAMON H. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy[J]. Biosystems Engineering, 2005, 91(4): 393–402.