

唐国强, 刘梦云, 蒋丹垚, 等. 基于分数阶微分的猕猴桃叶片叶绿素含量估算[J]. 江苏农业学报, 2025, 41(2): 335-344.
doi:10.3969/j.issn.1000-4440.2025.02.014

基于分数阶微分的猕猴桃叶片叶绿素含量估算

唐国强, 刘梦云, 蒋丹垚, 宋正华, 常庆瑞
(西北农林科技大学资源环境学院, 陕西 杨凌 712100)

摘要: 叶片叶绿素含量是表征植被生长状态的重要生理生化参数, 传统的叶绿素含量测定方法操作复杂且会破坏叶片组织结构, 对植物造成不可逆的损伤。通过构建高精度叶绿素反演模型, 可以实现对猕猴桃叶片叶绿素含量的实时无损监测。本研究采集了猕猴桃冠层的高光谱数据, 并同步测定了叶片叶绿素相对含量(*SPAD* 值)。通过对原始光谱进行分数阶微分变换(阶数为0~2, 步长为0.2), 结合竞争自适应重加权采样算法(CARS)筛选得到敏感波段。分别基于原始波段和敏感波段训练随机森林模型(RF)、支持向量机模型(SVR)和极限学习机模型(ELM)。结果表明, 分数阶微分变换显著提高了光谱反射率与猕猴桃叶片叶绿素含量的相关性, CARS算法提升了模型精度。光谱反射率经过1.8阶微分处理后, 采用CARS算法筛选出敏感波段, 利用这些敏感波段训练随机森林模型, 取得了最佳效果。训练后的随机森林模型在验证集上决定系数(R^2)达到0.93, 均方根误差(RMSE)为2.56, 相对分析误差(RPD)为3.89。该研究结果可为猕猴桃叶片叶绿素含量的高精度估算提供理论依据和技术参考, 对猕猴桃生长监测和精准农业管理具有重要意义。

关键词: 猕猴桃; 叶绿素含量; 高光谱; 分数阶微分变换; 竞争自适应重加权采样

中图分类号: S663.4 **文献标识码:** A **文章编号:** 1000-4440(2025)02-0335-10

Estimation of kiwifruit leaf chlorophyll content based on fractional-order differential processing

TANG Guoqiang, LIU Mengyun, JIANG Danyao, SONG Zhenghua, CHANG Qingrui
(College of Natural Resources and Environment, Northwest A&F University, Yangling 712100, China)

Abstract: The leaf chlorophyll content is an important physiological and biochemical parameter for characterizing the growth status of vegetation. Traditional methods for measuring chlorophyll content are cumbersome and destructive to leaf tissue, causing irreversible damage to plants. By constructing high-precision chlorophyll inversion models, real-time, non-destructive monitoring of chlorophyll content in kiwifruit leaves can be achieved. In this study, high-spectral data of the kiwifruit canopy were collected, and the relative chlorophyll content (*SPAD*) of the leaves was measured synchronously. The original spectra were subjected to fractional-order differential transformation (with orders ranging from 0 to 2 at a step size of 0.2). Subsequently, the competitive adaptive reweighted sampling algorithm (CARS) was used to identify the sensitive bands. Random forest (RF), support vector machine (SVR), and extreme learning machine (ELM) models were

trained based on both the original bands and the sensitive bands. The results showed that the fractional-order differential transformation significantly enhanced the correlation between spectral reflectance and chlorophyll content in kiwifruit leaves, and the CARS algorithm improved model accuracy. After the spectral reflectance was processed by the 1.8th order differential, the sensitive bands

收稿日期: 2024-10-10

基金项目: 国家自然科学基金项目(41701398, 42071240)

作者简介: 唐国强(2000-), 男, 湖南衡阳人, 硕士研究生, 主要从事土地资源与空间信息技术研究。(E-mail) 1813377325@qq.com

通讯作者: 刘梦云, (E-mail) lmy471993@163.com

were screened out by the CARS algorithm, and the random forest model trained using these sensitive bands achieved the best performance. The trained RF model had a determination coefficient (R^2) of 0.93, a root mean square error ($RMSE$) of 2.56, and a relative percent deviation (RPD) of 3.89 on the validation set. The results of this study can provide a theoretical basis and technical reference for the high-precision estimation of chlorophyll content in kiwifruit leaves and are of great significance for kiwifruit growth monitoring and precision agricultural management.

Key words: kiwifruit; chlorophyll content; hyperspectral; fractional order differential transformations; competitive adaptive reweighted sampling

叶片叶绿素含量(LCC)是表征植被生长状态的重要生理生化参数,该参数可以反映植被生长阶段和营养状态^[1]。传统的叶绿素含量测定方法存在操作复杂、耗时、费力等问题,而且会破坏叶片组织结构,对植物造成不可逆的损伤^[2]。相比之下,使用叶绿素仪能够高效、无损、准确地表征叶绿素相对含量^[3]。中国猕猴桃种植面积在全球排第一,同时也是主要的猕猴桃出口国^[4]。陕西地区依托秦岭山脉独特的地理优势,所产猕猴桃果实大、糖分高、营养丰富,已成为中国猕猴桃的主要生产地。因此,建立快速、高效的猕猴桃叶片叶绿素含量监测方法具有重要意义。

高光谱遥感技术凭借其快速、高效和无损检测的优势,被广泛用于植物叶片叶绿素含量的估测^[5]。研究结果表明,作物光谱反射率与叶片叶绿素含量之间存在显著相关性。在玉米的 4 个关键生长阶段,基于机器学习的 $LCC-MR$ 模型能够准确反映叶绿素含量与光谱反射率之间的关系,表现出较高的整体准确性^[6]。王宇等^[7]发现, $MCC-GAPLS-PLSR$ 模型在利用光谱数据反演叶绿素含量方面表现出了较强的能力。Song 等^[8]发现,结合高光谱图像中提取到的光谱信息和纹理信息,能够提升模型对感染花叶病苹果叶片叶绿素含量的检测精度;李紫琴等^[9]利用高光谱技术实现对不同施氮水平下棉花叶绿素含量的快速无损检测;Niu 等^[10]采用分数阶微分处理高光谱数据,通过分析叶绿素敏感波段,构建了典型植被指数,并建立了高精度模型以寻找叶绿素含量与光谱反射率之间的关联性。阿热孜古力·肉孜等^[11]采用了一系列光谱处理方法(对数变换、平方根变换、包络线去除、分数阶微分变换、小波变换)和多植被指数相结合的方法,构建了基于鲸鱼优化算法的随机森林模型,用于定量反演叶绿素含量。现有叶绿素含量无损检测方法多采用单一模型,容易出现反演精度不足和模型泛化能力差等

问题^[12]。传统机器学习模型在参数设置不当的情况下,不仅会导致拟合精度显著降低,还可能产生训练集表现良好而测试集效果欠佳的过拟合现象^[13]。

目前将分数阶微分与 CARS 算法相结合用于筛选最优变量子集,进而反演猕猴桃叶片叶绿素含量的研究较少。本研究拟以杨凌区猕猴桃作为研究对象,在叶片高光谱数据和实测叶绿素相对含量($SPAD$)值的基础上,对原始光谱进行分数阶微分变换,并采用 CARS 算法剔除冗余数据,筛选敏感波段构建优选变量集,再用优选变量集训练随机森林模型、支持向量机模型和极限学习机模型,对冠层叶片叶绿素含量进行预测和分析,以期对猕猴桃生长发育期叶绿素含量的检测提供技术支持。

1 材料与方法

1.1 试验区与试验材料

试验区(图 1)位于陕西省杨凌农业高新技术产业示范区五泉镇蒋家寨村($108^{\circ}01'20''E$, $34^{\circ}18'11''N$),该地区地势平坦,海拔 516.4~540.1 m,土壤肥沃,水资源丰富、水利条件优越,适宜多种农作物生长,年降水量 635.1~663.9 mm,年平均气温 $12.9^{\circ}C$,属暖温带季风半湿润气候,适宜种植猕猴桃。

试验所用猕猴桃品种为徐香,采样时间为 2023 年 6 月 29 日,正值猕猴桃的膨果期。膨果期是猕猴桃生长的关键阶段,该时期猕猴桃果实快速发育。选取健康且树龄较大的猕猴桃树,采用隔行采样法,选取 8 行,每一行选择 10 个点,在每个样点都摘取长势差异大的健康无虫害的新老叶片各 2 片,80 个采样点共采集 320 片叶片,叶片采摘后立即装入密封袋送回实验室。

1.2 试验方法

1.2.1 叶片叶绿素含量测定 采用叶绿素测量仪(型号 $SPAD-502Plus$, 日本 Konica Minolta 公司产品)测定膨果期猕猴桃叶片的 $SPAD$ 值, $SPAD$ 值与



图1 试验区位置

Fig.1 Location of the test area

叶绿素的实际含量之间存在直接的正相关关系,通过一定的转换关系,可以将 SPAD 值换算为叶绿素含量^[14]。测定时避开叶脉,从叶尖端至叶柄端,在叶片的左右侧各选择 3 个区域,每个区域测量 2 次,共测 6 次,最终取平均值。

1.2.2 叶片高光谱反射率测定 利用便携式非成像地物光谱仪测定高光谱反射率,光谱波段为 350~2 500 nm。为保证试验的科学性和可靠性,在测量前需进行多次白板校正,直至结果稳定。测量时,使用纸巾清洁叶片表面,确保无灰尘和泥土等杂物,并且每隔 30 min 需进行 1 次白板校正。按编号顺序用叶片夹夹持叶片,避开叶脉,从叶尖端至叶柄端,在叶片的左右侧各选择 3 个区域,每个区域测量 2 次。每片叶片测得 12 条光谱曲线,取平均值。

1.2.3 光谱数据处理 利用便携式非成像地物光谱仪对原始数据进行重采样,采样的间隔为 1 nm,区间为 400~1 000 nm,随后利用 SG 滤波法 (Savitzky golay filter) 对数据进行平滑处理,以有效去噪并保留植被生长过程中的局部突变信息^[15]。尽管便携式非成像地物光谱仪仅能采集部分光谱数据,但其涵盖叶绿素敏感的关键波段。与使用全光谱数据相比,使用部分关键光谱数据可提高模型处理效率和分析速度,在实际应用中可以更快地进行叶绿素反演。此外,便携式非成像地物

光谱仪的光谱分辨率和灵敏度较高,可以精确捕捉目标的光谱特征。

1.2.4 分数阶微分处理 本研究采用分数阶微分 (Fractional order differentiation, FOD) 算法对原始光谱进行处理,对光谱反射率实施 0~2 阶 (步长 0.2) 的分数阶微分变换。分数阶微分是一种用于计算非整数阶导数的方法,能够突出光谱曲线中斜率较高的区域,忽略平缓部分,从而提高数据分析的精度和效果,尤其在处理复杂光谱数据时具有显著优势^[16]。与整数阶微分相比,分数阶微分能够更全面地捕捉光谱细节信息。研究结果表明,整数阶模型的精度显著低于分数阶模型^[17]。

1.2.5 竞争性自适应重加权采样 竞争性自适应重加权采样算法 (Competitive adaptive reweighted sampling, CARS) 是一种基于达尔文生物进化理论的变量选择算法,通过迭代统计信息筛选出对生化参数敏感的波段,解决变量组合中的相关问题^[18]。采样过程包括蒙特卡罗采样、指数衰减波长选择、自适应重加权采样和循环迭代 4 个步骤。

1.2.5.1 蒙特卡罗采样 采用随机法划分数据集,使用偏最小二乘法 (PLS) 构建回归模型,通过计算回归系数的绝对值百分比,估算每个变量对目标变量的重要性。

1.2.5.2 指数衰减波长选择 CARS 算法初始包含所有变量,并在迭代中逐步减少变量数量。第 i 次采样中,变量保留比例的计算公式为:

$$r_i = ae^{-ki} \quad (1)$$

上式的约束条件为 $r_1 = P, r_N = 2/P$, 可得函数参数为:

$$a = (P/2)^{1/(N-1)} \quad (2)$$

$$k = \ln(p/2)/(N-1) \quad (3)$$

式中, r_i 为第 i 次采样时保留的变量比例; a 为初始变量比例; k 为衰减常数; N 为迭代次数。

CARS 算法中变量选择包括 2 个阶段: 快选阶段 ($i < n$), 在这一阶段, 算法会快速地剔除大量变量; 精选阶段 ($n < i < N$), 在这一阶段, 算法会更精确地分析剩余的变量, 每次迭代仅删除少量变量。

1.2.5.3 自适应重加权采样 根据指数衰减选择波长对应的变量, 剔除无用变量, 构建预测模型并计算均方根误差 (RMSE), 采用交叉验证均方根误差 (RMSECV) 评估模型性能。

1.2.5.4 循环迭代 通过多次循环迭代, 对比偏最

小二乘回归模型的 $RMSECV$ 值,确定最优变量子集。CARS 算法通过 Matlab2023b 实现^[18]。

1.2.6 建模方法 利用光谱反射率训练随机森林模型 (RF)、支持向量机模型 (SVM) 与极限学习机模型 (ELM)。随机森林模型具有高准确率、处理大型数据集能力强和抗过拟合的优点,但模型可解释性较差,训练和预测速度较慢^[19]。本研究中随机森林模型的树数量设为 5,每棵树深度为 5。支持向量机模型适用于高维空间且去噪能力强,但不适用于大规模数据集^[20]。本研究中,支持向量机模型的核心函数选择 fitrsvm 函数,最大迭代次数设为 30。极限学习机模型具有良好的泛化能力、多样性和灵活性以及一定的鲁棒性^[21],本研究所用激活函数为 sigmoid 函数,隐藏层节点个数设置为 50。

通过决定系数 (R^2)、均方根误差 ($RMSE$) 和相对分析误差 (RPD) 比较模型精度。 R^2 越接近 1,表明模型对目标变量的解释和预测能力越强,拟合效果越好; $RMSE$ 越小,表明模型的预测精度越高,误差越小。 RPD 越大,表明模型对数据集的适用性越好。

$$RPD = SD / RMSEP \quad (4)$$

式中, SD 为实测值的标准偏差, $RMSEP$ 为预测均方根误差。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2} \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

式中, \hat{y}_i 、 y_i 、 \bar{y} 分别为估测值、实测值、实测值平均值, n 为样本数。

2 结果与分析

2.1 分数阶微分处理的光谱反射率与叶片叶绿素含量的相关性

利用 0~2 阶 (步长为 0.2) 的分数阶微分对光谱反射率进行预处理,以分析光谱反射率与叶绿素含量之间的相关性。如图 2 所示,波长 400~865 nm,光谱反射率与叶绿素含量呈负相关;波长 550 nm 处,光谱反射率与叶绿素含量的相关系数最小 (-0.86);波长 866~1 000 nm,光谱反射率与叶绿素含量呈正相关。

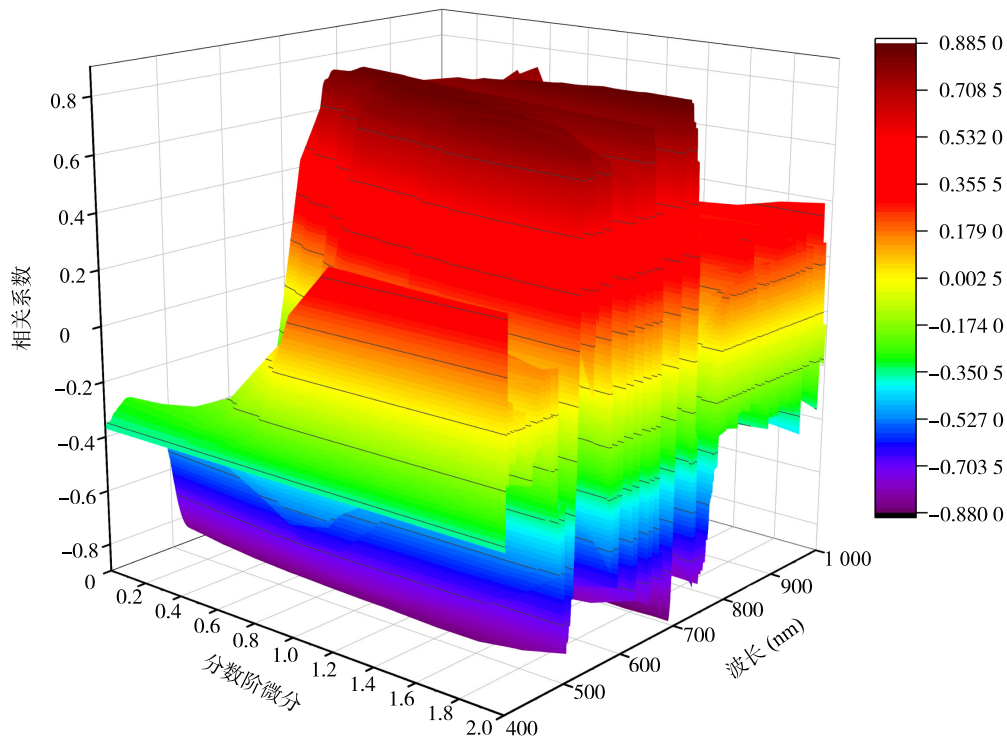
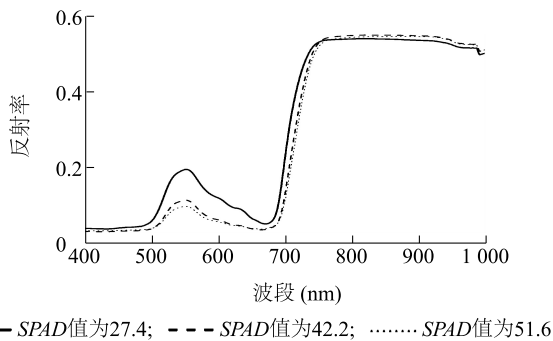


图 2 分数阶微分处理的光谱反射率与叶片叶绿素含量的相关性

Fig.2 The correlation between spectral reflectance processed by fractional order differential and chlorophyll content in leaves

2.2 叶片高光谱特征

2.2.1 原始光谱反射率 根据 SPAD 值将 160 个猕猴桃叶片的光谱反射率数据分为 3 组,每组的最小值、最大值与平均值分别为 (16.2, 35.5, 27.4)、(35.5, 46.6, 42.2) 和 (46.6, 62.1, 51.6)。如图 3 所示,3 条光谱反射率曲线的走势基本一致,在可见光波段 (380~750 nm),550 nm 附近的绿光波段出现反射峰,680 nm 附近的红光波段出现吸收谷,这是由于猕猴桃叶片对红光具有强吸收性。在红边波段 (680~750 nm),光谱反射率急剧上升,形成典型的红边效应。在近红外光波段,光谱曲线呈现高反射平台,光谱反射率随着 SPAD 值的增加而升高,表明叶片的内部结构更加稳定^[22]。在可见光波段,光谱反射率随着 SPAD 值的增大而降低,同时红边向长波方向移动,表明猕猴桃叶片对可见光的吸收能力增强。



SPAD:叶绿素相对含量。

图3 叶片原始光谱反射率曲线

Fig.3 The original spectral reflectance curve of the leaves

2.2.2 分数阶微分处理的光谱反射率 如图 4 所示,随着微分阶数的增加,曲线变化幅度越小,光谱反射率逐渐趋近于 0,特征峰更加明显。在 0.4 阶~2.0 阶微分处理下,光谱反射率出现负值。

2.3 利用竞争性自适应重加权采样算法 (CARS) 筛选敏感波段

在 CARS 计算过程中,设置因子数为 5,交叉验证折数为 10,蒙特卡罗采样次数为 50,旨在通过 50 次迭代寻找偏最小二乘法 (PLS) 模型最小的交叉验证均方根误差 ($RMSECV$)。如图 5 所示,在原 CARS 迭代过程中,剩余光谱波段数量呈指数下降,而 PLS 模型的 $RMSECV$ 总体呈现先急剧下降后波动上升的趋势。在前 23 次迭代中,PLS 模型的 $RMSECV$ 整

体呈下降趋势,表明与叶绿素含量无关的噪声信息被逐步剔除。在第 23 次迭代时,PLS 模型的 $RMSECV$ 最小,表明模型的预测精度达到最高。此后,PLS 模型的 $RMSECV$ 总体呈上升趋势,主要是因为与叶绿素含量密切相关的关键波段被剔除。因此,选择迭代 23 次时的波段组合作为子集。

如图 6 所示,对光谱反射率进行 1.0 阶微分预处理,使用 CARS 算法筛选得到的敏感波段最少 (7 个);对光谱反射率进行 0.4 阶微分预处理,使用 CARS 算法筛选得到的敏感波段最多 (149 个)。对光谱反射率进行 0~2.0 阶微分预处理,使用 CARS 算法筛选得到的敏感波段为 7~149 个,且大部分波段属于可见光波段。如表 1 所示,基于 0~2.0 阶微分处理光谱反射率,PLS 模型 $RMSECV$ 最小值为 4.99~5.54,模型的预测误差水平较为稳定,表明 CARS 算法对各阶微分处理的光谱反射率均有较好的筛选效果。

2.4 模型构建

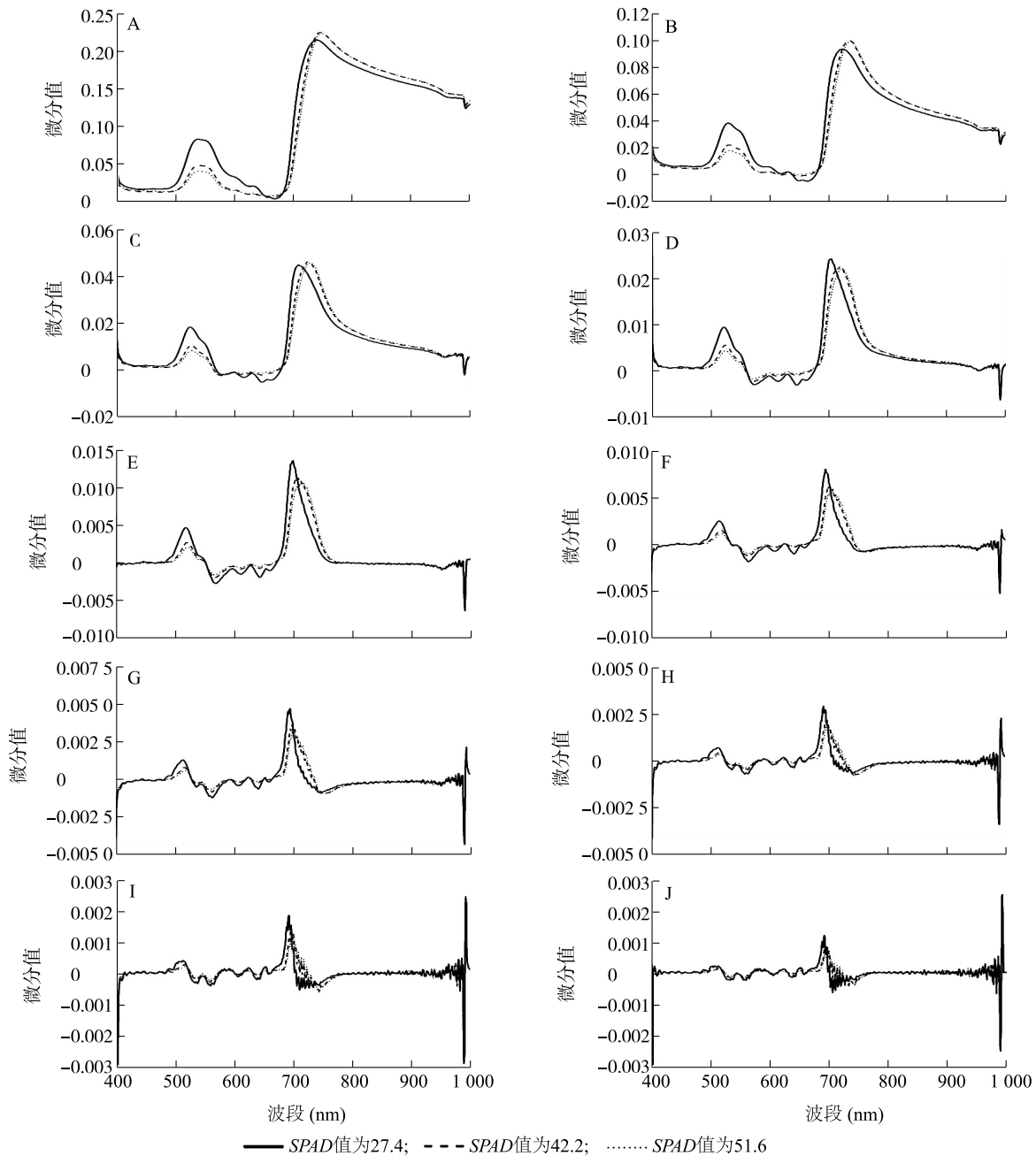
本研究分别基于各阶分数阶微分变换的光谱数据及其经 CARS 筛选后的最优变量集构建模型,训练集与测试集按 4:1 的数量比例分配。

2.4.1 随机森林模型 分别基于各阶分数阶微分处理后的光谱反射率数据及经 CARS 算法筛选后的敏感波段训练随机森林模型 (RF)。如表 2 和表 3 所示,光谱反射率经过 1.8 阶微分处理后,采用 CARS 算法筛选出敏感波段,利用这些敏感波段训练随机森林模型,取得了最佳效果。训练后的随机森林模型在验证集上 R^2 达到 0.93, $RMSE$ 为 2.56, RPD 为 3.89。0~2.0 阶微分预处理光谱反射率后,利用经过 CARS 算法筛选得到的敏感波段训练的模型,其在验证集上的决定系数 (R^2) 和相对分析误差 (RPD) 均高于未经过 CARS 算法筛选的原始波段训练的模型,其在验证集上的均方根误差 ($RMSE$) 均低于利用未经过 CARS 算法筛选的原始波段训练的模型。表明利用经过 CARS 算法筛选得到的敏感波段训练的随机森林模型性能更好,精度更高。

2.4.2 支持向量机模型 分别基于各阶分数阶微分处理后的光谱反射率数据及经 CARS 算法筛选后的敏感波段训练支持向量机模型 (SVM)。如表 4 和表 5 所示,光谱反射率经过 1.6 阶微分处理后,采用 CARS 算法筛选出敏感波段,利用这些敏感波段

训练支持向量机模型,取得了最佳效果。训练后的支持向量机模型在验证集上 R^2 达到 0.88, $RMSE$ 为 3.98, RPD 为 2.96。除 0.2 阶和 1.4 阶微分处理外,其他分数阶微分预处理光谱反射率后,利用经过 CARS 算法筛选得到的敏感波段训练的模型,其在验证集上的决定系数 (R^2) 和相对分析误差 (RPD)

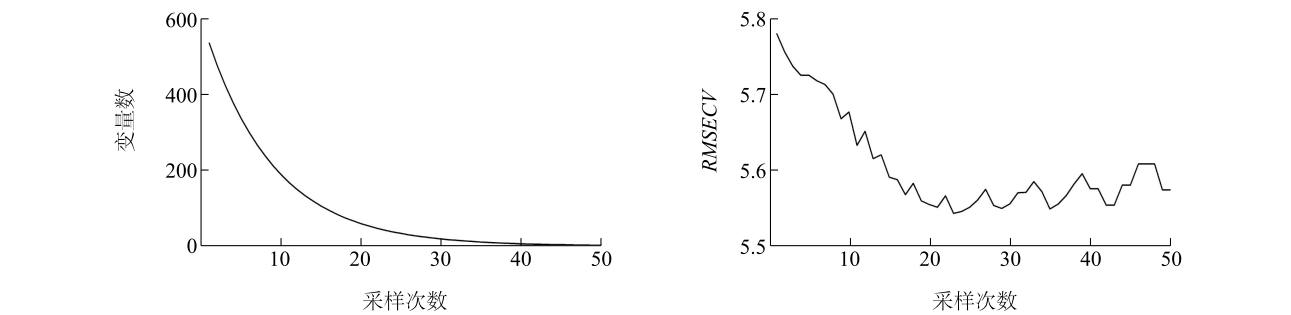
均高于未经过 CARS 算法筛选的原始波段训练的模型,其在验证集上的均方根误差 ($RMSE$) 均低于利用未经过 CARS 算法筛选的原始波段训练的模型。表明利用经过 CARS 算法筛选得到的敏感波段训练的支持向量机模型性能更好,精度更高。



SPAD: 叶绿素相对含量。A、B、C、D、E、F、G、H、I、J 分别为 0.2、0.4、0.6、0.8、1.0、1.2、1.4、1.6、1.8 和 2.0 阶微分处理的光谱反射率。

图 4 分数阶微分处理的叶片光谱反射率曲线

Fig.4 The spectral reflectance curves of leaves processed by fractional order differential



RMSECV;交叉验证均方根误差。

图 5 竞争性自适应重加权采样算法 (CARS) 的敏感波段筛选过程

Fig.5 The screening process of sensitive bands using the competitive adaptive reweighted sampling algorithm (CARS)

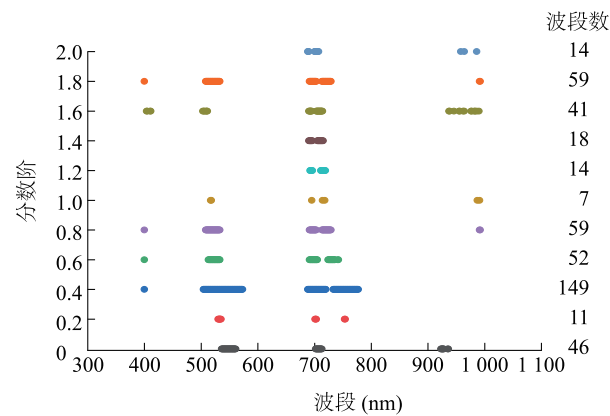


图 6 竞争性自适应重加权采样算法 (CRAS) 筛选的敏感波段分布

Fig.6 Distribution of sensitive bands screened by the competitive adaptive reweighted sampling algorithm (CARS)

表 1 光谱反射率数据基于不同分数阶微分处理的最小二乘 (PLS) 模型均方根误差最小值

Table 1 The minimum root mean square error of the partial least squares (PLS) model based on different fractional-order differential processing of spectral reflectance data

阶数	RMSECV
0	5.54
0.2	5.37
0.4	5.37
0.6	5.30
0.8	5.30
1.0	5.31
1.2	5.22
1.4	5.22
1.6	5.08
1.8	5.26
2.0	4.99

RMSECV;交叉验证均方根误差。

表 2 光谱反射率数据基于不同分数阶微分处理的随机森林模型

Table 2 Random forest models based on different fractional-order differential processing of spectral reflectance data

阶数	训练集			验证集		
	R^2	RMSE	RPD	R^2	RMSE	RPD
0	0.91	3.12	3.32	0.78	5.72	2.18
0.2	0.93	2.85	3.70	0.80	5.41	2.25
0.4	0.94	2.76	3.96	0.79	4.82	2.34
0.6	0.94	2.61	4.04	0.75	5.94	2.05
0.8	0.92	3.18	3.47	0.77	4.87	2.11
1.0	0.93	2.87	3.85	0.82	4.15	2.41
1.2	0.93	2.80	3.71	0.83	4.81	2.44
1.4	0.92	3.13	3.46	0.73	5.63	1.97
1.6	0.92	3.03	3.51	0.74	5.90	1.99
1.8	0.94	2.62	4.11	0.76	5.44	2.07
2.0	0.93	2.80	3.85	0.73	5.64	1.97

R^2 : 决定系数;RMSE;均方根误差;RPD;相对分析误差。

表 3 光谱反射率数据基于分数阶微分处理和竞争自适应算法的随机森林模型

Table 3 Random forest models based on fractional-order differential processing of spectral reflectance data and the competitive adaptive reweighted sampling algorithm

阶数	训练集			验证集		
	R^2	RMSE	RPD	R^2	RMSE	RPD
0	0.92	2.99	3.60	0.82	4.64	2.38
0.2	0.92	3.09	3.56	0.83	4.13	2.50
0.4	0.93	2.74	3.78	0.84	4.91	2.55
0.6	0.93	2.87	3.70	0.85	4.44	2.66
0.8	0.93	2.85	3.30	0.85	4.42	2.64
1.0	0.92	2.92	3.61	0.89	4.02	3.03
1.2	0.96	1.72	5.35	0.90	2.30	3.27
1.4	0.94	2.17	4.13	0.85	3.19	2.74
1.6	0.93	2.43	3.70	0.87	2.94	2.89
1.8	0.98	1.29	6.70	0.93	2.56	3.89
2.0	0.96	2.94	5.17	0.86	1.75	2.84

R^2 : 决定系数;RMSE;均方根误差;RPD;相对分析误差。

表 4 光谱反射率数据基于不同分数阶微分处理的支持向量机模型
Table 4 Support vector machine models based on different fractional-order differential processing of spectral reflectance data

阶数	训练集			验证集		
	R^2	$RMSE$	RPD	R^2	$RMSE$	RPD
0	0.76	5.38	2.06	0.77	4.77	2.11
0.2	0.78	5.25	2.12	0.82	4.00	2.40
0.4	0.78	5.10	2.16	0.80	4.48	2.25
0.6	0.81	4.78	2.33	0.77	4.65	2.12
0.8	0.83	4.51	2.44	0.80	4.54	2.29
1.0	0.84	4.25	2.49	0.78	4.52	2.18
1.2	0.83	4.45	2.46	0.79	4.57	2.24
1.4	0.83	4.48	2.43	0.78	5.08	2.15
1.6	0.83	4.57	2.41	0.81	4.38	2.35
1.8	0.83	4.29	2.41	0.82	5.27	2.37
2.0	0.81	4.66	2.30	0.82	4.75	2.41

R^2 : 决定系数; $RMSE$: 均方根误差; RPD : 相对分析误差。

表 5 光谱反射率数据基于分数阶微分和竞争自适应算法处理的支持向量机模型
Table 5 Support vector machine models based on fractional-order differential processing of spectral reflectance data and competitive adaptive reweighted sampling algorithm

阶数	训练集			验证集		
	R^2	$RMSE$	RPD	R^2	$RMSE$	RPD
0	0.83	4.24	2.45	0.80	5.64	2.25
0.2	0.80	4.86	2.23	0.84	4.43	2.52
0.4	0.88	3.71	2.97	0.82	4.23	2.38
0.6	0.81	4.63	2.33	0.86	4.00	2.76
0.8	0.85	4.09	2.61	0.86	4.22	2.77
1.0	0.80	4.71	2.25	0.88	4.12	2.89
1.2	0.83	4.47	2.44	0.83	4.41	2.46
1.4	0.89	3.48	2.97	0.83	5.11	2.45
1.6	0.83	4.38	2.43	0.88	3.98	2.96
1.8	0.82	4.54	2.35	0.88	4.06	2.87
2.0	0.83	4.54	2.44	0.85	3.93	2.58

R^2 : 决定系数; $RMSE$: 均方根误差; RPD : 相对分析误差。

2.4.3 极限学习机模型 分别基于各阶分数阶微分处理后的光谱反射率数据及经 CARS 算法筛选后的敏感波段训练极限学习机模型 (ELM)。如表 6 和表 7 所示, 光谱反射率经过 0.6 阶微分处理后, 采用 CARS 算法筛选出敏感波段, 利用这些敏感波段训练极限学习机模型, 取得了最佳效果。训练后的极限学习机模型在验证集上 R^2 达到 0.86, $RMSE$ 为 4.03, RPD 为 3.29。0~2.0 阶微分预处理光谱反射

率后, 利用经过 CARS 算法筛选得到的敏感波段训练的模型, 其在验证集上的决定系数 (R^2) 和相对分析误差 (RPD) 均高于未经过 CARS 算法筛选的原始波段训练的模型, 其在验证集上的均方根误差 ($RMSE$) 均低于利用未经过 CARS 算法筛选的原始波段训练的模型。表明利用经过 CARS 算法筛选得到的敏感波段训练的极限学习机模型性能更好, 精度更高。

表 6 光谱反射率数据基于不同分数阶微分处理的极限学习机模型
Table 6 Extreme learning machine models based on different fractional-order differential processing of spectral reflectance data

阶数	训练集			验证集		
	R^2	$RMSE$	RPD	R^2	$RMSE$	RPD
0	0.80	4.60	2.93	0.83	5.10	3.08
0.2	0.82	4.35	3.03	0.81	5.42	2.94
0.4	0.81	4.58	2.98	0.82	4.90	2.88
0.6	0.81	4.72	3.12	0.81	4.53	2.95
0.8	0.81	4.55	3.05	0.81	5.48	2.82
1.0	0.77	5.18	2.60	0.74	5.65	2.64
1.2	0.75	5.39	2.65	0.76	5.28	2.54
1.4	0.71	5.67	2.33	0.73	6.08	2.53
1.6	0.75	5.37	2.48	0.77	5.18	2.82
1.8	0.76	5.09	2.64	0.65	7.30	2.26
2.0	0.72	5.59	2.38	0.70	6.42	2.37

R^2 : 决定系数; $RMSE$: 均方根误差; RPD : 相对分析误差。

表 7 光谱反射率数据基于分数阶微分和竞争自适应算法处理的极限学习机模型
Table 7 Extreme learning machine models based on fractional-order differential processing of spectral reflectance data and competitive adaptive reweighted sampling algorithm

阶数	训练集			验证集		
	R^2	$RMSE$	RPD	R^2	$RMSE$	RPD
0	0.82	4.66	3.14	0.85	3.69	3.22
0.2	0.81	4.67	3.10	0.83	4.59	3.48
0.4	0.82	4.56	3.21	0.83	4.34	3.61
0.6	0.81	4.69	3.20	0.86	4.03	3.29
0.8	0.81	4.62	3.07	0.83	4.98	3.13
1.0	0.81	4.75	3.04	0.84	4.16	3.46
1.2	0.80	5.00	2.97	0.83	4.04	3.10
1.4	0.80	4.84	3.00	0.82	4.63	3.07
1.6	0.83	4.28	3.10	0.82	5.11	3.07
1.8	0.80	4.67	3.02	0.81	5.22	3.04
2.0	0.83	4.24	3.14	0.83	5.30	3.12

R^2 : 决定系数; $RMSE$: 均方根误差; RPD : 相对分析误差。

3 讨论

叶绿素是植物进行光合作用的重要色素,其含量与光合作用强度密切相关,直接影响植物的生长发育。因此,快速、准确地获取植物叶绿素含量对检测植物生长状态具有重要意义。本研究通过分析不同 SPAD 值下的光谱特征发现,原始光谱的整体走势基本一致。随着 SPAD 值的增加,可见光区域的反射率总体下降,而近红外区域的反射率则呈上升趋势,这表明 SPAD 值高的猕猴桃叶片细胞栅栏组织稳定性提高^[23]。研究还发现,随着分数阶微分阶数的增加,光谱反射峰更加明显,但可见光和部分近红外区域的反射率一致性降低,噪声现象更严重,导致建模精度并未随阶数增加而持续提高。

针对相邻波段光谱反射率相关性较高的问题,本研究采用分数阶微分方法显著提升了光谱数据与猕猴桃叶片叶绿素含量之间的相关性。并进一步结合竞争自适应重加权采样方法(CARS)对分数阶微分后的光谱数据进行处理,结果表明,CARS 算法能够有效剔除与叶绿素含量无关的波段,筛选出最优变量子集,显著提高建模精度。

分别基于原始波段和分数阶微分变换+CARS 算法筛选得到的敏感波段训练随机森林模型(RF)、支持向量机模型(SVM)与极限学习机模型(ELM),所有模型在验证集和训练集上的精度 R^2 均达到 0.65 及以上。其中,光谱反射率经过 1.8 阶微分处理后,采用 CARS 算法筛选出敏感波段,利用这些敏感波段训练随机森林模型,取得了最佳效果。训练后的随机森林模型在验证集上 R^2 达到 0.93, RMSE 为 2.56, RPD 为 3.89。RF 模型因其优异的集成学习机制和对特征的鲁棒性,在处理本研究数据时表现出最佳的预测能力,而支持向量机模型在处理高维数据和边界定义方面表现良好^[24]。在未来的研究中可进一步优化模型参数,同时可以尝试集成各模型的优势以获得最佳预测结果^[25]。

本研究基于膨果期猕猴桃开展叶绿素监测试验,该结论是否适用于其他生育期仍需进一步验证。此外,探索更多阶数的分数阶微分变换对光谱信息的处理效果也具有重要的研究意义。

4 结论

本研究利用便携式非成像地物光谱仪测得

400~1 000 nm 高光谱反射率,通过 0~2.0 阶微分变换光谱反射率,并采用 CARS 算法筛选敏感波段,分别将原始波段和筛选得到的敏感波段用于训练随机森林模型(RF)、支持向量机模型(SVM)和极限学习机模型(ELM)。光谱反射率经过 1.8 阶微分处理后,采用 CARS 算法筛选出敏感波段,利用这些敏感波段训练随机森林模型,取得了最佳效果。训练后的随机森林模型在验证集上 R^2 达到 0.93, RMSE 为 2.56, RPD 为 3.89。

参考文献:

- [1] 王 烁,常庆瑞,刘梦云,等.基于高光谱遥感的棉花叶片叶绿素含量估算[J]. 中国农业大学学报,2017,22(4):16-27.
- [2] 陈 澜,常庆瑞,高一帆,等.猕猴桃叶片叶绿素含量高光谱估算模型研究[J]. 西北农林科技大学学报(自然科学版),2020,48(6):79-89,98.
- [3] 李粉玲,王 力,刘 京,等.基于高分一号卫星数据的冬小麦叶片 SPAD 值遥感估算[J]. 农业机械学报,2015,46(9):273-281.
- [4] 张计育,莫正海,黄胜男,等.21 世纪以来世界猕猴桃产业发展以及中国猕猴桃贸易与国际竞争力分析[J]. 中国农学通报,2014,30(23):48-55.
- [5] ANSAR A, MUHAMMAD M. Evaluating the potential of red edge position (REP) of hyperspectral remote sensing data for real time estimation of LAI & chlorophyll content of kinnow mandarin (*Citrus reticulata*) fruit orchards[J]. Scientia Horticulturae, 2020, 267:109326.
- [6] GUO Y M, JIANG S Y, MIAO H L, et al. Ground-based hyperspectral estimation of maize leaf chlorophyll content considering phenological characteristics[J]. Remote Sensing, 2024, 16(12). DOI:10.3390/rs16122133.
- [7] 王 宇,汪 泓,肖玖军,等.基于 MCC-GAPLS-PLSR 的辣椒叶绿素含量高光谱定量反演[J]. 江苏农业学报,2024,40(5):865-873.
- [8] SONG Z H, LIU Y F, YU J R, et al. Estimation of chlorophyll content in apple leaves infected with mosaic disease by combining spectral and textural information using hyperspectral images[J]. Remote Sensing, 2024, 16:2190.
- [9] 李紫琴,王家强,李 贞,等.基于光谱指数的棉花叶片叶绿素密度估算研究[J]. 中国农业科技导报,2024,26(8):103-111.
- [10] NIU L Y, GAO C Y, SUN J B, et al. Study on hyperspectral estimation model of chlorophyll content in grape leaves[J]. Agricultural Biotechnology, 2018(4):55-58,61.
- [11] 阿热孜古力·肉孜,买买提·沙吾提,何旭刚,等.基于多植被指数组合的棉花叶片叶绿素含量估算[J]. 干旱区研究,2023,40(11):1865-1874.
- [12] 符欣彤,常庆瑞,张佑铭,等.基于 Stacking 集成学习的猕猴桃叶片叶绿素含量估算[J]. 干旱地区农业研究,2023,41(4):

- 247-256.
- [13] 林少喆,彭致功,王春堂,等. 基于“三边”参数的冬小麦冠层 SPAD 值监测模型[J]. 排灌机械工程学报,2021,39(1):102-108.
- [14] 刘楠,杨海波,高飞,等. 基于查找表法和优化光谱指数的马铃薯叶绿素反演[J]. 中国马铃薯,2022,36(6):495-507.
- [15] 姜海玲,姚奕旭,洪绣超,等. 基于 sg 滤波去噪的时间序列谐波分析重建算法研究[J]. 吉林师范大学学报(自然科学版),2021,42(3):133-140.
- [16] 向友珍,王辛,安嘉琪,等. 基于分数阶微分和最优光谱指数的大豆叶面积指数估算[J]. 农业机械学报,2023,54(9):329-342.
- [17] 蒋宇恒,晏博,庄清源,等. 基于分数阶微分的土壤重金属锌和镍的定量反演模型研究[J]. 光谱学与光谱分析,2024,44(10):2850-2857.
- [18] 李铠,常庆瑞,陈倩,等. 基于连续小波变换耦合 cars 算法的冬小麦冠层叶片含水量估算[J]. 麦类作物学报,2023,43(2):251-258.
- [19] 张瑞杰,周春艳,陈辉,等. 基于 GF-5B 卫星的随机森林模型反演京津冀地区 $PM_{2.5}$ [J]. 中国环境科学,2024,44(11):5961-5970.
- [20] 池涛,曹广溥,李丙春,等. 基于高光谱数据和 SVM 方法的土壤盐渍度反演[J]. 山东农业大学学报(自然科学版),2018,49(4):585-590.
- [21] 艾小童. 基于 elm 模型的极化 sar 土壤水分降尺度研究[J]. 地理空间信息,2023,21(7):11-15.
- [22] YUAN Z R, YE Y, WEI L F, et al. Study on the optimization of hyperspectral characteristic bands combined with monitoring and visualization of pepper leaf SPAD value [J]. Sensors, 2022, 22(1):183.
- [23] 王婷婷. 基于高光谱和高分一号卫星影像的冬小麦叶绿素遥感反演[D]. 咸阳:西北农林科技大学,2019.
- [24] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报,2000,26(1):36-46.
- [25] HAN X H, MA S F, SHI Z W, et al. A novel power transformer fault diagnosis model based on Harris-Hawks-optimization algorithm optimized kernel extreme learning machine [J]. Journal of Electrical Engineering & Technology, 2022, 17(3):1993-2001.

(责任编辑:成纾寒)