doi:10.6041/j.issn.1000-1298.2025.06.045

基于 CA/SPA－CARS 算法的小麦条锈病特征波段优选与监测模型构建

谷玲霄[1,2]　方　涛[1]　杜林丹[1]　吴喜芳[1,2]　李长春[1,2]　连增增[1]　岳　哲[1]
(1. 河南理工大学测绘与国土信息工程学院，焦作 454003；2. 河南理工大学定量遥感与智慧农业研究院，焦作 454003)

摘要：作物病害会严重制约作物产量和品质，传统的病害监测方法效率低且易受主观因素影响。高光谱遥感技术以其高光谱分辨率和客观真实性在作物病害监测中展现出重要潜力。本文利用多生育期冬小麦地面高光谱及田间病情指数(Disease index，DI)，基于相关性分析(Correlation analysis，CA)和连续投影法(Successive projections algorithm，SPA)分别对光谱数据进行光谱特征降维，通过构建最优参数的竞争性自适应重加权采样(Competitive adaptive reweighted sampling，CARS)算法优选小麦条锈病敏感波段，最后利用偏最小二乘回归(Partial least squares regression，PLSR)、反向传播神经网络(Back propagation neural network，BPNN)和极限学习机(Extreme learning machine，ELM)算法建立基于特征光谱的病情指数模型，比较不同建模方法的建模效果，实现小麦条锈病监测。研究结果表明，不同生育期均显示小麦条锈病敏感特征波段多集中于近红外和短波红外波段，其中挑旗期为 842、850、858 nm，灌浆期为 947、953、1 275、1 277、1 590、1 663、1 665 nm；对比不同建模算法，PLSR 模型表现最佳，满足小麦早期病虫害监测需求，且在病害中期显示更明显特征；挑旗期和灌浆期分别以 SPA－CARS－MCX 和 CA－CARS－MSC 数据构建 PLSR 模型预测效果最优，验证集 R^2 分别为 0.782 和 0.861，RMSE 分别为 0.022 和 0.094，RPD 分别为 2.140 和 2.687。本文构建算法能够为不同生育期小麦条锈病监测提供参考。

关键词：小麦条锈病；光谱变换；特征波段选择；相关性分析；连续投影法；竞争性自适应重加权采样
中图分类号：S435.121.4+2　　**文献标识码**：A　　**文章编号**：1000-1298(2025)06-0487-12　　**OSID**：

Feature Band Selection and Construction of Monitoring Model of Wheat Stripe Rust Based on CA/SPA－CARS Algorithm

GU Lingxiao[1,2]　FANG Tao[1]　DU Lindan[1]　WU Xifang[1,2]　LI Changchun[1,2]　LIAN Zengzeng[1]　YUE Zhe[1]
(1. *School of Surveying and Land Information Engineering*, *Henan Polytechnic University*, *Jiaozuo* 454003, *China*
2. *Institute of Quantitative Remote Sensing and Smart Agriculture*, *Henan Polytechnic University*, *Jiaozuo* 454003, *China*)

Abstract: Crop diseases can seriously restrict crop yield and quality. Traditional disease monitoring methods are inefficient and susceptible to subjective factors. Hyperspectral remote sensing technology has shown great potential in crop disease monitoring due to its high spectral resolution and objective authenticity. Ground hyperspectral and field disease index (DI) data of winter wheat with multiple growth stages were used, the spectral data were preprocessed using correlation analysis (CA) and successful projection algorithm (SPA) respectively, and the sensitive bands of wheat stripe rust through competitive adaptive reweighted sampling (CARS) algorithm that can construct optimal parameters were optimized. Finally, partial least squares regression (PLSR), back propagation neural network (BPNN) and extreme learning machine (ELM) were used to establish the disease index model based on the characteristic spectrum, and the modeling effects of different modeling methods were compared to realize the monitoring of wheat stripe rust. The research results indicated that the sensitive characteristic bands of wheat stripe rust in different growth stages were mainly concentrated in the near infrared and shortwave

收稿日期：2024－11－14　修回日期：2025－01－26
基金项目：河南省重点研发与推广专项(科技攻关)项目(232102210093、242102320198)、河南理工大学测绘科学与技术“双一流”学科创建项目(SYJX02)、河南理工大学博士基金项目(B2023－37)、无人机高光谱土壤环境污染评估在工业选址中的应用项目(H23－140)、河南省自然科学基金项目(242300420221)和河南理工大学国家级重大科研成果培育基金项目(NSFRF240101)
作者简介：谷玲霄(1994—)，女，讲师，博士，主要从事定量遥感研究，E-mail：gulx@hpu.edu.cn
通信作者：杜林丹(1994—)，女，讲师，博士，主要从事城市热岛、地理空间和景观生态风险研究，E-mail：du_lindan@163.com

infrared bands, with 842 nm, 850 nm, and 858 nm in the flag leaf stage and 947 nm, 953 nm, 1 275 nm, 1 277 nm, 1 590 nm, 1 663 nm, and 1 665 nm in the filling stage. In the comparison of different modeling methods, PLSR model performed best, and the model met the needs of early monitoring of wheat diseases and pests, and showed more obvious characteristics in the middle of the disease. During the flag leaf stage and filling stages, the PLSR models constructed with SPA - CARS - MCX and CA - CARS - MSC respectively had the best prediction performance. The R^2 of the validation sets were 0.782 and 0.861, the RMSE were 0.022 and 0.094, and the RPD were 2.140 and 2.687, respectively. The algorithm constructed can provide ideas for monitoring wheat stripe rust at different growth stages.

Key words: wheat stripe rust; spectral transformation; feature band selection; correlation analysis; successive projections algorithm; competitive adaptive reweighted sampling

0 引言

作物病害会导致作物减产、品质下降,已经成为制约农业可持续发展的重要因素。小麦是全球最重要的粮食作物之一,在我国是仅次于水稻的第二大粮食作物,其产量稳定对保障国内粮食充足供应至关重要,是我国农业经济支柱之一[1]。条锈病作为小麦生育过程中的主要病害之一,具有传播范围广、爆发性强和危害程度大等特点,会显著降低小麦产量,影响小麦品质,严重危害着我国粮食安全[2]。通过快速、准确地监测和评估小麦条锈病病情,可以制定有效的防控措施,此外,了解条锈病发病机制有助于培育抗病小麦品种,从而提高农业生产稳定性和效益[3]。

传统的病害监测方法主要依赖于人工检测和目视评估,不仅耗时耗力,而且易受到主观因素的影响[4-5]。近年来,高光谱遥感技术由于其能够提供丰富的光谱信息而被广泛应用于作物病害监测[6-9]。高光谱数据能够捕捉植物叶片在不同波长下的反射特性,从而揭示病害对植物光谱特征的影响,通过分析这些光谱特征,可以实现对病害的早期检测和精准评估[10-13]。基于此,国内外众多学者借助高光谱遥感技术针对小麦条锈病展开了深入研究。HUANG 等[14]联合相关性分析和随机森林算法从与小麦条锈病相关的 21 个植被指数中提取最佳指数组合,使用极限梯度提升算法构建了条锈病严重度监测模型;DENG 等[15]基于无人机高光谱数据评估了植被指数及最佳波段组合,以实现小麦条锈病定量反演,研究表明,与单植被指数建模相比,适当的多植被指数建模可提高模型性能;GUO 等[16]通过组合无人机高光谱影像提取植被指数和纹理特征,建立了基于偏最小二乘回归的小麦条锈病染病早期监测模型,为利用无人机高光谱影像进行早期条锈病监测提供了参考;梅广源等[17]通过联合相关性分析与随机森林变量选择对小麦条锈病高光谱数据集进行研究,结果表明:在不同病情指数下,波长 821、846、537 nm 处的一阶微分有明显特征;YAO 等[18]使用主成分分析和连续投影法从小麦高光谱数据中提取条锈病敏感波段,并结合反向传播神经网络构建条锈病早期监测模型;姚志凤等[19]通过连续投影法、主成分分析和竞争性自适应重加权算法提取了小麦感染条锈病、白粉病以及健康叶片敏感波段,构建了病害分类模型;竞霞等[20]基于遗传算法和支持向量回归算法对小麦条锈病高光谱数据开展研究,确定了小麦条锈病严重度敏感因子,相比融合相关性分析法与支持向量回归算法,提高了条锈病监测精度。综上,众多学者在使用高光谱遥感技术监测小麦条锈病方面已经取得大量成果,但小麦条锈病光谱特征随生育期不同和发病程度的不同存在差异,而许多研究并未划分生育期,忽视了小麦条锈病高光谱特征在不同生育期的差异;学者们多基于已有植被指数或特征波段进行条锈病监测研究,而针对小麦条锈病特点对其敏感波段进行分析、选择和优化的较少,而如何优选反演条锈病的敏感波段,提高监测模型泛化能力和稳定性,还需要进一步探究。

本文以不同生育期下不同条锈病发病程度的小麦冠层非成像高光谱遥感数据为研究对象,结合地面病情指数调查资料,使用多元散射校正(Multiple scattering correction, MSC)、标准正态变量交化(Standard normal variate transformation, SNV)和数据中心化(Mean centering, MCX)等 3 种方法对原始光谱进行预处理,然后分别使用相关性分析法(Correlation analysis, CA)与连续投影法(Successive projections algorithm, SPA)进行特征波段初步提取,使用竞争自适应重加权采样(Competitive adaptive reweighted sampling, CARS)实现特征波段优选,最后利用偏最小二乘回归(Partial least squares regression, PLSR)、反向传播神经网络(Back propagation neural network, BPNN)和极限学习机(Extreme learning machine, ELM)算法建立小麦条锈病病情指数监测模型,探讨基于 CA/SPA - CARS

算法优选特征波段在不同建模方法下对不同生育期小麦条锈病监测的可行性，以期为小麦条锈病快速监测提供方法。

1　研究区概况与研究数据

1.1　研究区概况

冬小麦冠层高光谱观测实验由北京农业信息技术研究中心遥感技术部于 2014 年 4—5 月在北京市昌平区小汤山国家精准农业示范研究基地(40°10.6′N，116°26.3′E)开展。实验区域分为对照区和病害区，4 月 5 日在病害区接种条锈菌，分别于接种后 17 d 的挑旗期和接种后 58 d 的灌浆期测量各样区的非成像高光谱冠层反射率数据以及生理生化参数，并同步记录各样区病情指数调查数据。挑旗期分别设置 9 个对照区和 9 个病害区，灌浆期分别设置 18 个对照区和 9 个病害区，共采集 45 个样本点。

1.2　冬小麦冠层高光谱观测与预处理

采用 ASD FieldSpec Pro 光谱仪测量小麦冠层高光谱数据，测定时间为 10:30—14:30，天气状况良好，晴朗无云。每个小区随机选取 1 个光谱调查点，每点测量 20 次，取其平均值作为该小区平均光谱值，各小区测量前后均立即使用标准参考板进行校正，采集波长为 350～2 500 nm。

在受到病害胁迫后，小麦光合色素含量的改变使得可见光波段冠层光谱反射率发生变化，同时近红外波段受细胞结构变化影响[21]，对植被病害表现敏感；病害胁迫也使得小麦叶片水分发生变化，进而影响到短波红外范围内冠层反射率特征[22]。由此选择波长 350～1 800 nm 的高光谱数据作为原始数据，记为 OR 数据，如图 1 所示。

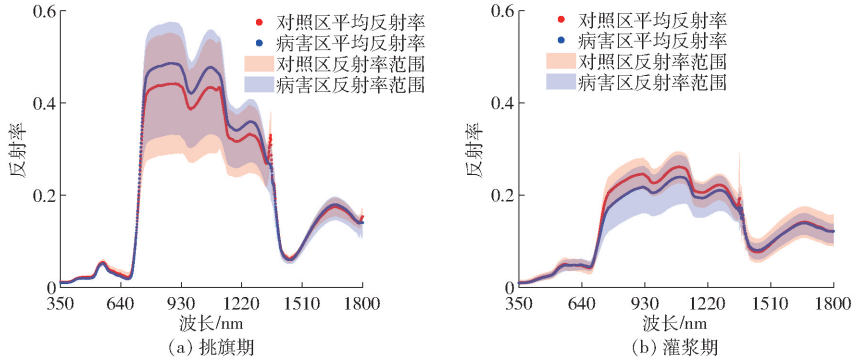

图 1　原始数据光谱反射率

Fig. 1　Spectral reflectance of OR

由于高光谱数据具有数据量大、相邻波段信息相关性高以及存在数据冗余等特点，既包含样本信息，也存在背景信息和噪声，通过光谱预处理降低噪声干扰、消除部分光散射，提高数据信噪比，从而增强数据可靠性和准确性[23]。为探究不同光谱预处理方法对构建小麦条锈病监测模型的影响，选取 MSC、SNV 和 MCX 等 3 种光谱变换方法对原始光谱数据进行预处理，减少噪声干扰以及散射影响，处理后光谱数据分别记为 MSC、SNV 和 MCX 数据[24]，如图 2 所示。FLS 表示挑旗期，FS 表示灌浆期。

小麦条锈病田间调查按照 GB/T 15795—1995《小麦条锈病测报调查规范》进行。分别于 4 月 22 日、5 月 29 日在测量小麦冠层光谱数据时同步进行田间病情指数(Disease index，DI)的调查。冠层病情调查采用五点取样法，在每个小区内选取对称的 5 点，每点约 2 m^2，每点随机选取 20 株小麦，分别调查其发病情况。将严重度分为 8 梯度(1%、5%、10%、20%、40%、60%、80% 和 100%)，分别记录各严重度小麦叶片数量，并计算病情指数，计算式为

$$D_I = \frac{\sum (tf)}{d\sum f}$$

式中　D_I——病情指数

t——各梯度极值

d——最高梯度，取 8

f——各梯度叶片数量

挑旗期时，对照区小麦无发病症状，病害区发病症状较轻；灌浆期时，对照区小麦有轻微发病症状，病害区发病症状严重。各生育期 DI 如表 1 所示。

2　研究方法

2.1　特征波段提取

利用高光谱数据进行小麦条锈病监测研究时，核心问题在于从大量高光谱波段中准确提取与病情相关的敏感波段。使用相关性分析和连续投影法分别对高光谱数据进行特征波段初步提取，再通过竞争自适应重加权采样算法优选特征波段，探究 CA/SPA－CARS 算法优选小麦条锈病敏感波段的可行性，技术路线如图 3 所示。

CA简易直观、计算简单，可以高效快速地从大

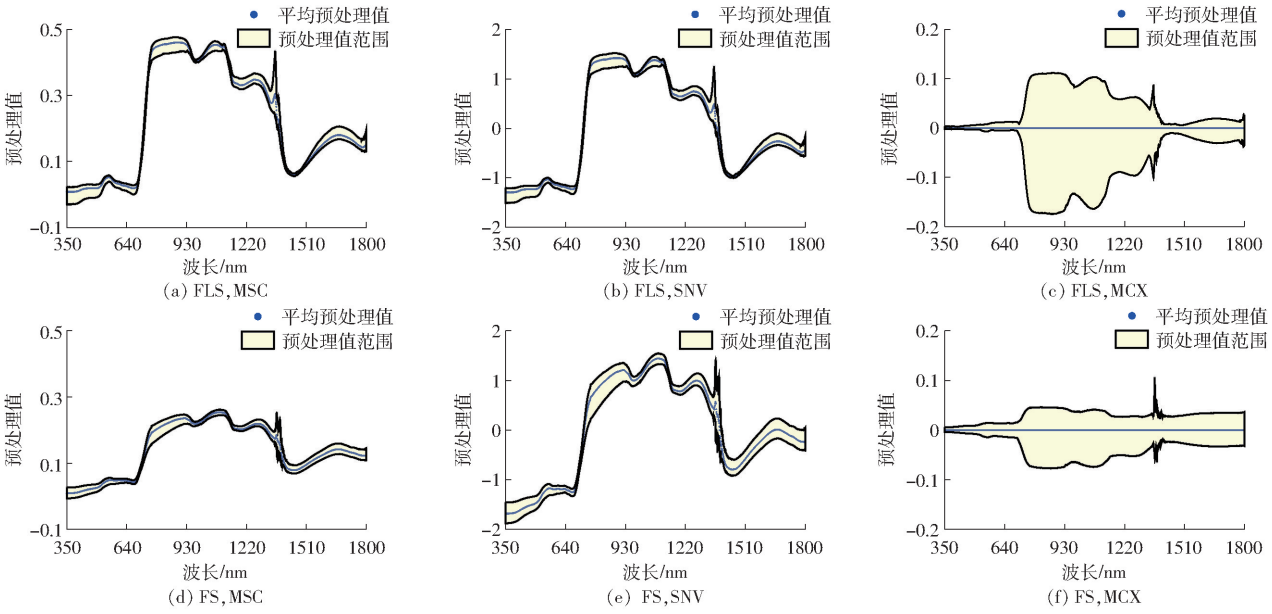

图 2 光谱预处理结果

Fig. 2 Results of spectra pretreatment

表 1 各生育期病情指数统计分析

Tab. 1 Statistical analysis of disease index in different reproductive stages

生育期	平均值	最大值	最小值	标准差	样本数量
挑旗期	0.060	0.133	0	0.060	18
灌浆期	0.404	0.939	0.111	0.276	27

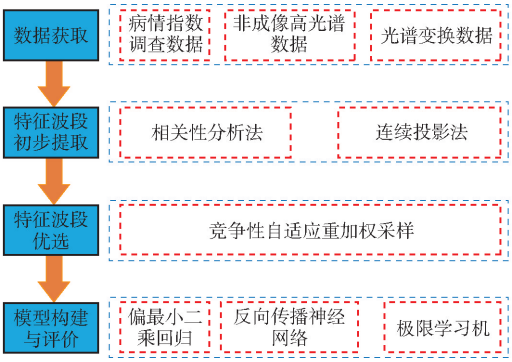

图 3 技术路线流程图

Fig. 3 Flowchart of technical route

量高光谱数据中筛选出与病虫害密切相关的波段，减少数据处理复杂性和计算量，适用于各种高光谱数据统计学分析；SPA 在处理高维数据领域应用广泛，从高光谱数据集中挑选一个波段，将其他波段投影在已选波段上，以投影向量最大的波段为待选波段，通过循环直至满足预设停止条件，筛选出含有最小冗余信息及共线性的波段组合[25]。

CARS 算法结合蒙特卡洛采样与偏最小二乘（Partial least squares，PLS）模型回归系数实现特征变量选择，通过指数衰减函数和自适应重加权采样方法去除 PLS 模型中回归系数绝对值较小的变量，基于剩余变量建立新的 PLS 模型继续筛选，经过多次运算，最终通过交叉验证选取交叉验证均方根误差（Root mean square error of cross-validation，RMSECV）最低的子集作为特征变量[26]。

CARS 算法涉及多次特征采样和权重调整，计算复杂度较高，尤其在特征维度非常大的情况下，可能需要较长时间来完成特征选择过程。为结合不同特征提取算法优势，分别采用 CA 和 SPA 算法对高光谱数据进行初步特征波段提取，去除大量无信息光谱，缩减数据量，再结合 CARS 算法进行特征波段优选，完成整个特征波段提取过程。

2.2 模型构建与评价

偏最小二乘回归（Partial least squares regression，PLSR）是一种线性回归分析方法，结合多元线性回归分析、主成分分析和典型相关分析的方法，适用于数量大且相关性强的变量间的回归方程建立，广泛地应用在各研究领域[27]；反向传播神经网络（Back propagation neural network，BPNN）是一种按误差反向传播算法训练的多层前馈神经网络，具有学习能力强，适应性广泛等优点，具有处理复杂非线性问题能力[28]；极限学习机是由 HUANG 等[29]提出的一种基于单层前馈神经网络理论改进的机器学习方法，具有训练速度快，泛化性能良好等优点，在回归分析领域发挥着重要作用。

为保证所建模型随机性，对两个生育期样本均按照比例 3∶1 随机划分训练集和验证集。模型精度评价利用决定系数（Coefficient of determination，R^2）、均方根误差（Root mean square error，RMSE）和相对分析误差（Relative percent deviation，RPD）等评价指标。R^2 越高，越趋向于 1，该模型预测能力越

强，性能越稳定；RMSE 越小，模型实测值与预测值之间偏差越小，模型估算能力越强[30]；RPD 大于等于 2 时，表明模型可对样本进行有效预测，模型稳定性良好；RPD 为 1.4～2 时，表明模型对样本进行粗略预测，模型稳定性一般；RPD 小于 1.4 时，表明模型不能对样本进行预测，模型稳定性很差[31]。

3　结果及分析

3.1　特征波段初步提取

对不同生育期下小麦条锈病原始光谱以及预处理数据集分别进行 CA 与 SPA 处理，完成对各高光谱数据集初步特征波段提取工作。

3.1.1　CA 法提取特征波段

对各数据集与小麦条锈病 DI 进行相关性分析，提取出相关系数绝对值 $|r|$ 大于等于 0.5 且达到极显著性水平波段，各数据集与小麦条锈病 DI 的相关系数如图 4 所示。由图 4 可知，挑旗期时，OR 数据中紫外线（350～400 nm）、绿光及部分黄光区域（507～604 nm）、红边（696～703 nm）和短波红外（1 355～1 470 nm、1 799 nm、1 800 nm）范围内共 155 个波段满足提取条件，记为 CA－OR 数据，MCX 数据满足条件的波段范围与 OR 数据相同，共 155 个波段，记为 CA－MCX 数据；MSC 数据中主要为红边（737～740 nm）、近红外（967～1 120 nm）和短波红外（1 340～1 382 nm、1 784～1 800 nm）共 182 个波段满足提取条件，记为 CA－MSC 数据，SNV 数据满足条件的波段范围与 MSC 数据类似，共 169 个波段，记为 CA－SNV 数据；灌浆期时，CA－OR、CA－MCX 数据位于红边及近红外（755～843 nm）内，共 89 个波段；CA－MSC 数据位于红边及近红外（727～960 nm、1 145～1 300 nm）和短波红外（1 300～1 409 nm、1 561～1 682 nm）内，共 565 个波段，CA－SNV 数据满足条件的波段范围与 CA－MSC 数据相同，共 581 个波段。

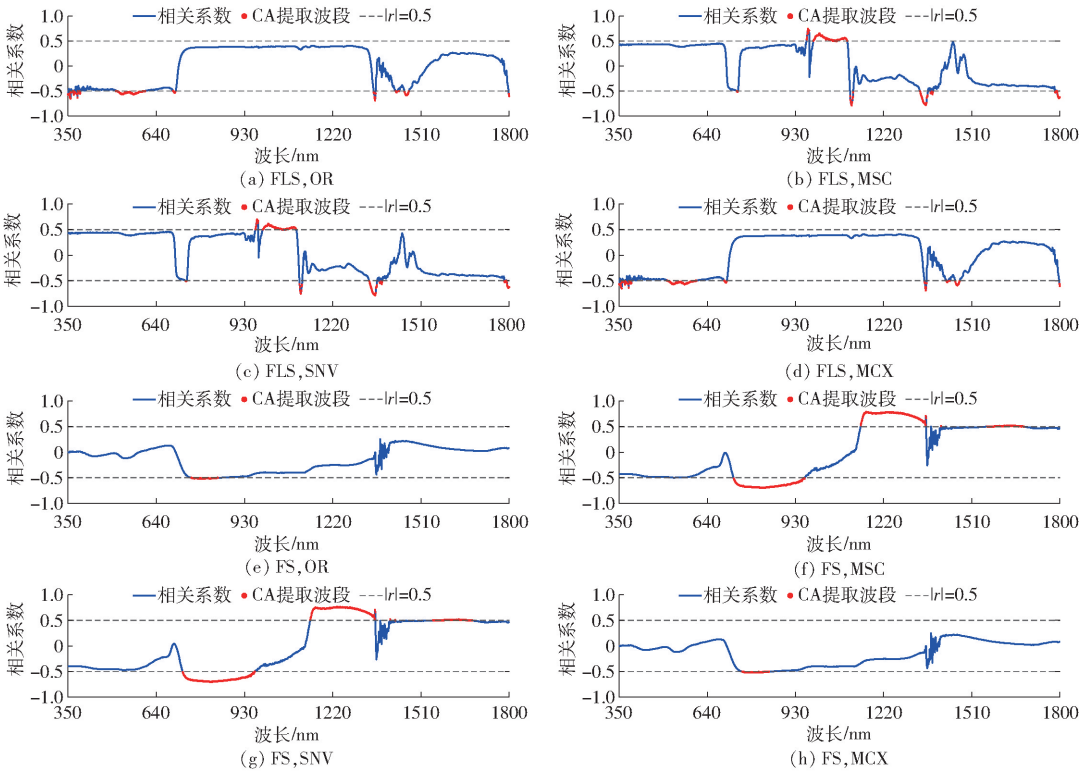

图 4　各数据集与小麦条锈病冠层 DI 的相关系数

Fig. 4　Correlation coefficient between each dataset and canopy DI of wheat stripe rust

3.1.2　SPA 法提取特征波段

使用 SPA 对各数据集进行特征波段提取，设置提取最大变量数为 100，在每组光谱数据中，将每个波段均作为初始变量运行一次 SPA，共得到 1 451 组数据，将结果组成一个 100×1 451 的矩阵，统计矩阵中出现频率最高的前 100 个元素作为 SPA 特征波段提取结果，如图 5 所示。

由图 5 可知，挑旗期时，OR、MCX 数据经过 SPA 处理后，提取到的波段位于近红外（800～920 nm）和短波红外（1 359～1 365 nm），分别记为 SPA－OR 和 SPA－MCX 数据；MSC 数据经 SPA 处理后，提取到的波段位于近红外（789～884 nm）和短波红外（1 347～1 365 nm），记为 SPA－MSC 数据；SNV 数据经过 SPA 处理后，提取到的波段位于紫外线（350～402 nm）以及近红外（813～869 nm）和短波红外（1 349～1 365 nm），记为 SPA－SNV 数据。

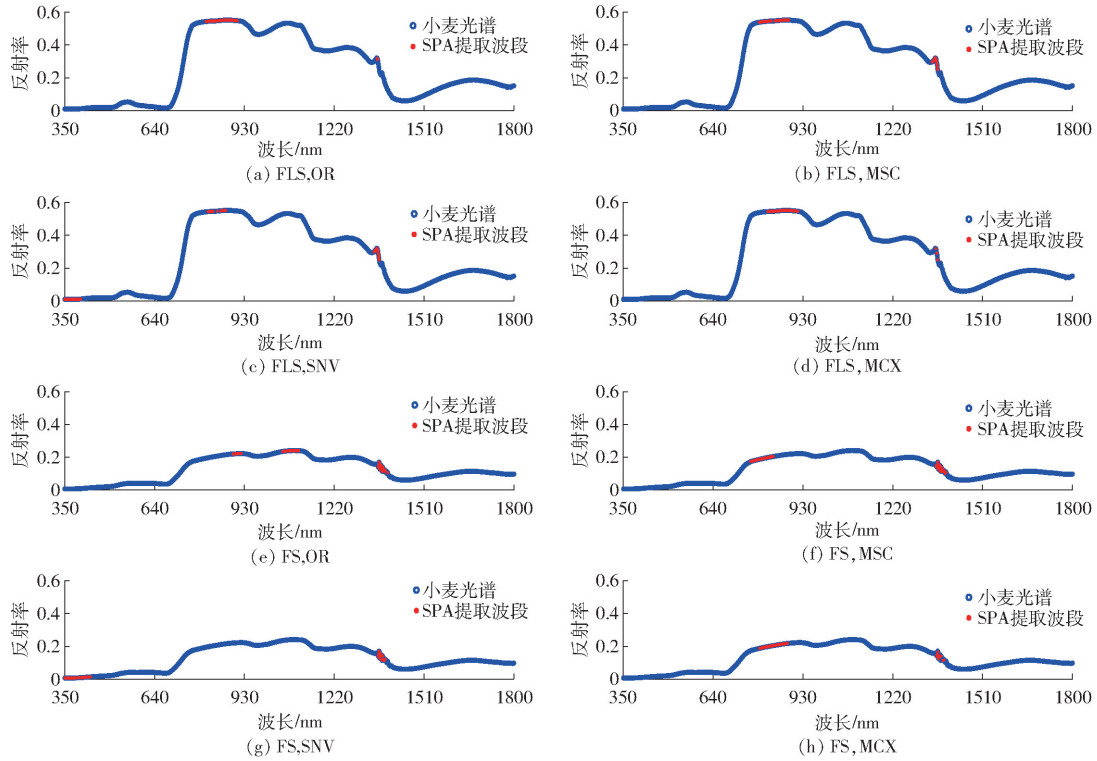

图 5 SPA 特征波段提取结果

Fig. 5 Extraction results of SPA characteristic bands

灌浆期时,SPA – OR 数据位于近红外(893 ~ 920 nm、1 054 ~ 1 105 nm)和短波红外(1 362 ~ 1 392 nm);SPA – MSC 数据位于近红外(761 ~ 834 nm)和短波红外(1 360 ~ 1 391 nm);SPA – SNV 数据位于紫外线以及到蓝光过渡区域(350 ~ 433 nm)和短波红外(1 362 ~ 1 384 nm);SPA – MCX 数据位于近红外(790 ~ 880 nm)和短波红外(1 362 ~ 1 385 nm)。

3.2 CARS 特征波段优选

CARS 通过 PLS 模型波段绝对回归系数判断选择波段重要性,建立 PLS 模型需预设最大主成分数量,再通过交叉验证确定最佳主成分数量,以此构建合适的 PLS 模型。CARS 还需设置采样次数,为研究最大主成分数量、采样次数设置对 CARS 算法运行的影响,将采样次数分别设置为 100、200 和 300,最大主成分数量分别设置为 10、20 和 30,对各特征波段初步提取的数据集进行 CARS 处理,对比分析上述参数不同输入下 CARS 输出结果,以此确定最大主成分数量和采样次数的最优输入,以挑旗期 CA – OR数据为例,CARS 输出结果对应的采样次数以及最佳主成分数量统计结果如表 2 所示。

由表 2 可知,通过 CARS 算法对挑旗期 CA – OR 数据进行特征波段优选,不同采样次数设置下 CARS 输出结果对应采样次数均小于 100;设置不同最大主成分数量,CARS输出结果对应的PLS模型

表 2 挑旗期 CA – OR 数据的 CARS 输出结果对应采样次数以及最佳主成分数量统计

Tab. 2 Number of sampling times and optimal number of principal components for CARS output results of CA – OR data during flag leaf stage

采样次数	最大主成分数量	最佳主成分数量	输出结果对应采样次数
100	10	4	8
100	20	2	89
100	30	2	95
200	10	4	22
200	20	5	17
200	30	2	32
300	10	7	18
300	20	3	63
300	30	3	39

的最佳主成分数量均小于 10。其余初步提取的特征波段数据均满足以上情况,且采样次数和最大主成分数量设置越小,CARS 算法运行时间越短,有利于提高特征波段优选效率,因此在 CARS 处理中,将采样次数设置为 100,最大主成分数量设为 10。

CARS 优选波段过程如图 6 所示,以灌浆期 CA – SNV 数据为例:设置采样次数为 100,由图 6 可知,当采样次数为第 60 次时,为 CARS 输出结果(图 6c 蓝色星号竖线所在采样次数位置):由图6a 可知,当采样次数为第 60 次时,波段数减少至 15 个;图 6b 显示随着采样次数增加至 60 次,RMSECV 取最小

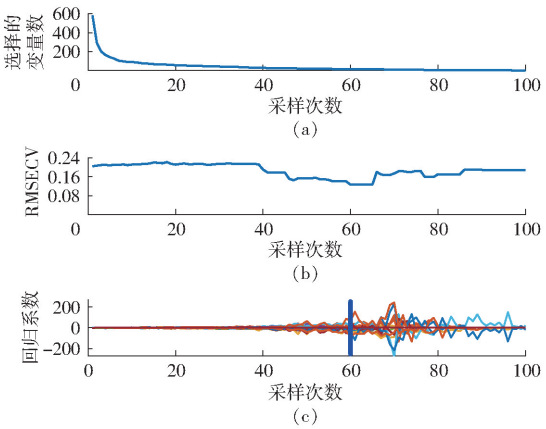

图 6　CARS 提取特征波段流程

Fig. 6　Process of extracting characteristic bands by CARS

值，为 0.126 7；图 6c 显示所有波段的回归系数随采样次数增加，CARS 保留了对模型贡献更大的波段，增加其回归系数，同时逐渐去除贡献率小的波段，实现对特征波段的优选，当第 60 次采样时，模型 RMSECV 最低，此时保留的 15 个波段即为特征波段，记为 CA－CARS－SNV 数据。

3.2.1　小麦挑旗期 CARS 特征波段优选

使用 CARS 对 CA 和 SPA 处理后的挑旗期数据集进行特征波段优选，特征波段信息及分布特征如图 7 和表 3 所示。由图 7、表 3 可知，CARS 算法对挑旗期不同数据集提取的特征波段数目不同，分布也有所差异。对于 CA 处理后的数据集，经 CARS 处理，CA－CARS－OR 数据主要位于紫外线、可见光以及红边范围，少量位于短波红外范围；CA－CARS－MSC 和 CA－CARS－SNV 数据大多数分布于近红外范围，少量位于短波红外范围，CA－CARS－MCX 数据位于紫外线及可见光范围；联合 SPA 和 CARS 处理后的数据集整体相较于上述数据，包含波段数更少，其中 SPA－CARS－OR 和 SPA－CARS－MCX 数据所含波段数最少，分布于近红外波段；其次为 SPA－CARS－MSC 数据，大多数分布于近红外范围，少量位于短波红外范围；SPA－CARS－SNV 数据包含波段数最多，分布于紫外线、近红外及短波红外范围。

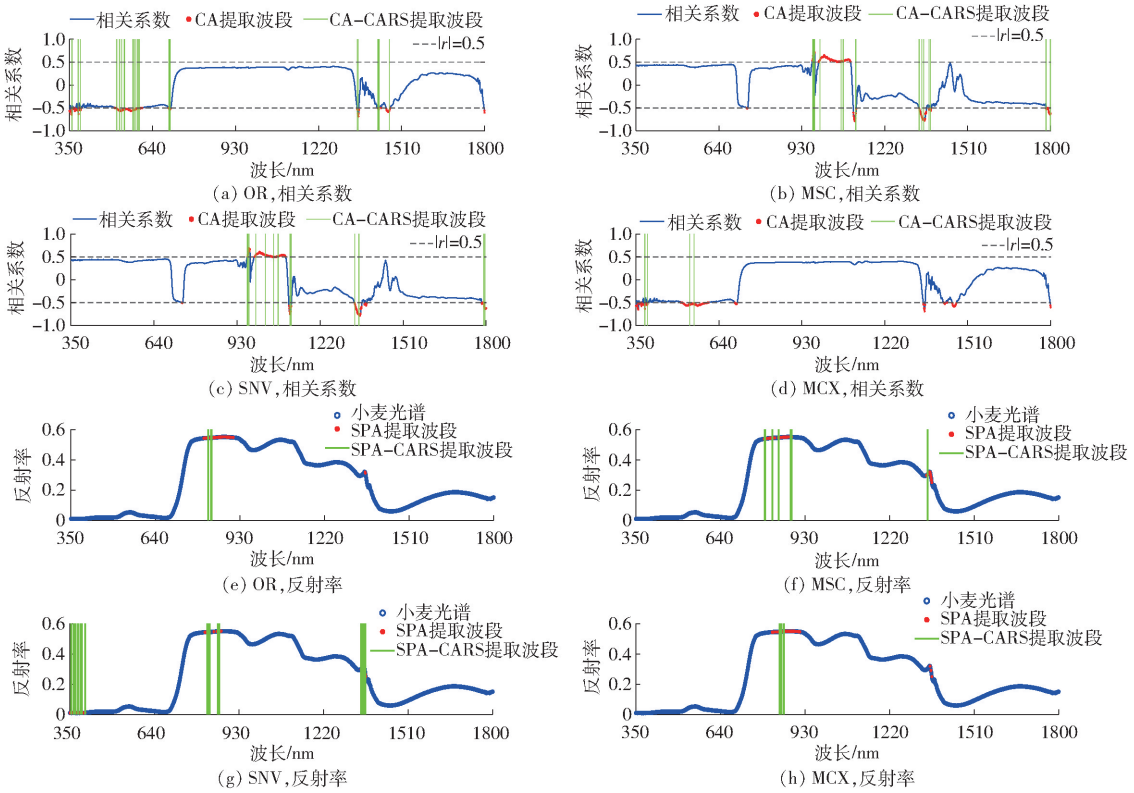

图 7　挑旗期提取特征波段分布

Fig. 7　Distributions of extracted characteristic bands in flag leaf stage

3.2.2　小麦灌浆期 CARS 特征波段优选

小麦灌浆期下各数据集经 CARS 算法优选的特征波段信息统计结果及分布情况如和图 8 和表 4 所示。CA 处理下，CA－CARS－OR 和 CA－CARS－MCX 数据均分布于近红外范围；CA－CARS－SNV 数据包含的波段数最多，大多数分布于近红外范围，少量位于短波红外内，CA－CARS－MSC 数据大多数分布于短波红外范围，近红外范围有少量分布；SPA 处理下，SPA－CARS－SNV 数据分布于紫外线及紫光范围内，其余 3 组数据均分布于近红外波段范围内。

3.3　模型构建与评价

完成特征波段优选后，使用 PLSR、BPNN 和

表 3 挑旗期 CA/SPA－CARS 提取特征波段数目

Tab. 3 Number of characteristic bands extracted by CA/SPA－CARS in flag leaf stage

数据集	波段数
CA－CARS－OR	25
CA－CARS－MSC	18
CA－CARS－SNV	13
CA－CARS－MCX	4
SPA－CARS－OR	3
SPA－CARS－MSC	7
SPA－CARS－SNV	20
SPA－CARS－MCX	3

监测要求，因此从结果中去除。比较挑旗期各模型验证集 RMSE，PLSR 模型 RMSE 最低，其次为 ELM 和 BPNN 模型，因此在挑旗期条锈病 DI 估测模型中，PLSR 模型表现最佳。

ELM 算法对各特征数据集进行 DI 估测模型构建，各生育期不同模型验证集精度统计情况如表 5 和表 6 所示。由表 5 可知，挑旗期各模型验证精度情况，PLSR 模型中 CA－CARS－MSC 数据建模精度最高，验证集 R^2、RMSE 和 RPD 分别为 0.787、0.022 和 2.167；BPNN 模型中 SPA－CARS－SNV 数据建模精度最高，验证集 R^2、RMSE 和 RPD 分别为 0.692、0.026 和 1.802；ELM 模型中 CA－CARS－MSC 数据建模精度最高，验证集 R^2、RMSE 和 RPD 分别为 0.719、0.025 和 1.885，SPA－CARS－OR 数据构建的 ELM 模型验证精度极低，无法满足条锈病

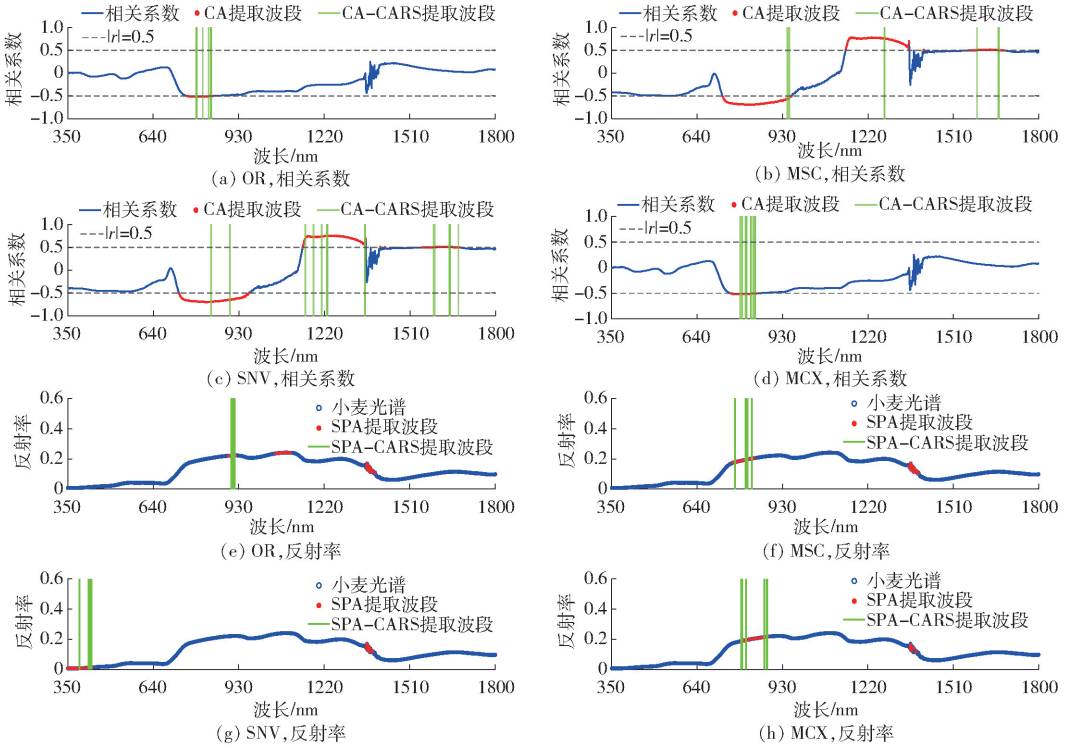

图 8 灌浆期提取特征波段分布情况

Fig. 8 Distribution of extracted characteristic bands in filling stage

表 4 灌浆期 CA/SPA－CARS 提取特征波段数目

Tab. 4 Number of characteristic bands extracted by CA/SPA－CARS in filling stage

数据集	波段数
CA－CARS－OR	7
CA－CARS－MSC	7
CA－CARS－SNV	15
CA－CARS－MCX	10
SPA－CARS－OR	5
SPA－CARS－MSC	5
SPA－CARS－SNV	5
SPA－CARS－MCX	7

由表 6 可知，灌浆期各模型验证精度，PLSR 模型中 CA－CARS－MSC 数据建模精度最高，验证集 R^2、RMSE 和 RPD 分别为 0.861、0.094 和 2.687；BPNN 模型中 CA－CARS－OR 数据建模精度最高，验证集 R^2、RMSE 和 RPD 分别为 0.775、0.119 和 2.106；ELM 模型中 CA－CARS－MSC 数据建模精度最高，验证集 R^2、RMSE 和 RPD 分别为 0.788、0.116 和 2.172。比较灌浆期各模型验证集 RMSE，PLSR 模型 RMSE 最低，因此在灌浆期条锈病 DI 估测模型中，PLSR 模型表现最优。

由上述模型评价情况可知，PLSR 模型表现最佳。比较各 PLSR 模型评价指标，多个模型验证精度相近且精度较高，模型精度相近的情况下，输入变量少的模型相对稳定，受无关因素或噪声数据影响小，泛化能力较强，且数据获取与处理成本较低，数

表 5　挑旗期不同模型验证集精度

Tab. 5　Accuracy of validation set for different models in flag leaf stage

数据集	PLSR			BPNN			ELM		
	R^2	RMSE	RPD	R^2	RMSE	RPD	R^2	RMSE	RPD
CA－CARS－OR	0. 739	0. 024	1. 959	0. 612	0. 030	1. 605	0. 680	0. 027	1. 768
CA－CARS－MSC	0. 787	0. 022	2. 167	0. 679	0. 027	1. 764	0. 719	0. 025	1. 885
CA－CARS－SNV	0. 781	0. 022	2. 135	0. 575	0. 031	1. 534	0. 625	0. 029	1. 633
CA－CARS－MCX	0. 734	0. 025	1. 941	0. 618	0. 029	1. 617	0. 636	0. 029	1. 657
SPA－CARS－OR	0. 729	0. 031	1. 919	0. 567	0. 031	1. 519			
SPA－CARS－MSC	0. 783	0. 022	2. 146	0. 580	0. 031	1. 544	0. 710	0. 026	1. 858
SPA－CARS－SNV	0. 764	0. 023	2. 060	0. 692	0. 026	1. 802	0. 694	0. 026	1. 807
SPA－CARS－MCX	0. 782	0. 022	2. 140	0. 635	0. 029	1. 654	0. 713	0. 025	1. 867

表 6　灌浆期不同模型验证集精度

Tab. 6　Accuracy of validation set for different models in filling stage

数据集	PLSR			BPNN			ELM		
	R^2	RMSE	RPD	R^2	RMSE	RPD	R^2	RMSE	RPD
CA－CARS－OR	0. 853	0. 097	2. 604	0. 775	0. 119	2. 106	0. 741	0. 128	1. 964
CA－CARS－MSC	0. 861	0. 094	2. 687	0. 745	0. 127	1. 980	0. 788	0. 116	2. 172
CA－CARS－SNV	0. 859	0. 094	2. 667	0. 767	0. 122	2. 070	0. 723	0. 133	1. 898
CA－CARS－MCX	0. 788	0. 116	2. 171	0. 754	0. 125	2. 016	0. 683	0. 142	1. 777
SPA－CARS－OR	0. 765	0. 122	2. 063	0. 601	0. 159	1. 584	0. 440	0. 188	1. 337
SPA－CARS－MSC	0. 839	0. 101	2. 495	0. 719	0. 133	1. 887	0. 731	0. 131	1. 927
SPA－CARS－SNV	0. 754	0. 125	2. 016	0. 526	0. 173	1. 452	0. 677	0. 143	1. 760
SPA－CARS－MCX	0. 813	0. 109	2. 309	0. 702	0. 137	1. 833	0. 759	0. 123	2. 039

据更新也更便捷。

为探究小麦条锈病在不同生育期的最优 PLSR 监测模型，比较了不同模型验证集的 RMSE 以及输入特征波段数目，如图 9 所示。挑旗期时，CA－CARS－MSC、SPA－CARS－MSC、CA－CARS－SNV 和 SPA－CARS－MCX 数据构建的 PLSR 模型验证集 RMSE 较低且接近，但 SPA－CARS－MCX 数据包含的特征波段数目远小于其余数据集，因此选择 SPA－CARS－MCX 数据构建的 PLSR 模型作为小麦挑旗期条锈病监测模型，特征波段为：842、850、858 nm；灌浆期时，CA－CARS－MSC 数据构建模型的 RMSE 最小且波段数目较少，所以以此数据构建的 PLSR 模型作为小麦灌浆期条锈病监测模型，特征波段为：947、953、1 275、1 277、1 590、1 663、1 665 nm。

4　讨论

小麦受到条锈病胁迫后，不同生育期生长光谱曲线发生改变，主要出现在可见光、近红外和短波红

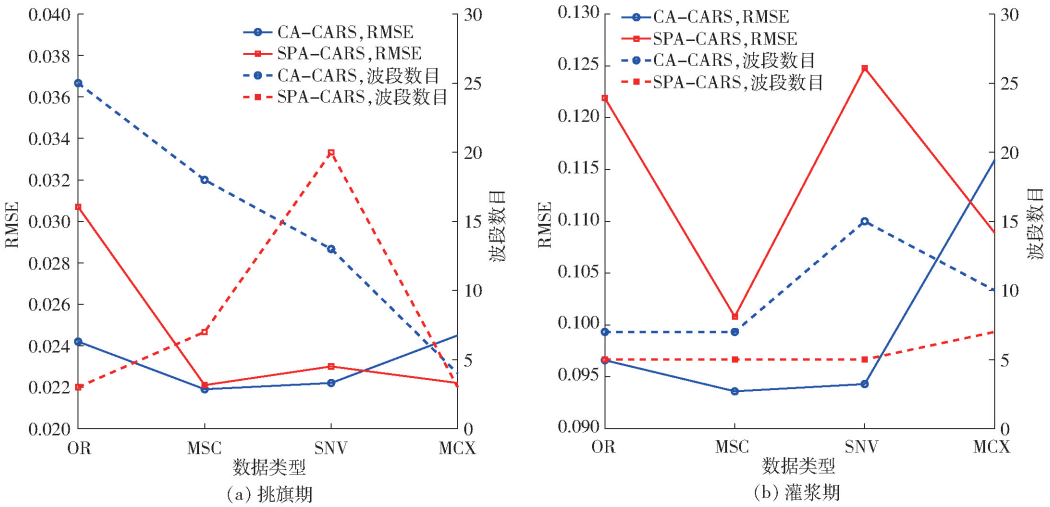

图 9　PLSR 模型验证集 RMSE 及特征波段数目

Fig. 9　RMSE and number of characteristic bands for validation set of PLSR model

外波段范围内。当小麦处于挑旗期时,植株的病害症状较低,病害植株在可见光范围光谱发生轻微变化,主要是由于条锈病的侵染使得小麦叶片叶绿素含量发生改变,对可见光的吸收产生影响[21],同时条锈病的胁迫可能改变了小麦叶片细胞结构,影响了水分吸收,导致叶片含水率上升,使得近红外和短波红外(1 350 nm 左右)范围的小麦冠层反射率发生显著变化,分别出现反射率上升和下降的情况[32-33]。当处于灌浆期时,病害区小麦病害症状严重,DI 远高于对照区植株,反射率差异主要出现在近红外范围内,是由于植被内部结构破坏导致发病小麦在近红外范围反射率降低[34]。

在特征波段初步提取中,使用 SPA 算法对每个数据集进行处理时,将所有波段依次作为初始变量运行一次 SPA 并统计了结果中出现频率最高的前 100 个波段,相比于只进行一次 SPA 处理,更加全面地探索了特征空间,减少遗漏重要信息的概率以及随机性影响,提供了更具代表性的特征。在通过 CA/SPA-CARS 算法完成整个特征波段优选工作后,各个数据集由最初 350~1 800 nm 共 1 451 个波段缩减为 5~25 个特征波段,占原始数据集 0.3%~1.7%,表明了通过 CA/SPA-CARS 算法的联合特征变量提取,极大程度上去除了无信息与低信息变量,降低了变量维度,提升了后续建模效率。

使用的 CA 与 SPA 两种特征波段初步提取算法,对于模型构建均发挥了重要作用。从模型预测精度来看,CA-CARS 算法优于 SPA-CARS,原因可能为:CA 通过计算变量之间的线性相关系数,能够发现小麦条锈病 DI 与敏感波段之间的特定线性关系,全面考虑了所有波段与病情指数的相关性,从而筛选出对条锈病敏感且具有强相关性的波段组合[35];而 SPA 主要基于变量投影,可能会因为过于注重消除共线性而忽略了一些与条锈病高度相关但存在一定共线性的波段;从模型稳定性来看,经过 SPA 处理的数据虽然建立模型精度略低于 CA,但需要特征波段数更少,冗余度较低,使最终建立的模型具有简单稳定、数据处理成本低和可解释性强等特点,且 SPA 能更好地识别变量非线性关系,这是 CA 算法有所欠缺的部分[36],所以在小麦条锈病监测工作中,可根据实际情况选择合适的特征变量提取方法。

选取 PLSR、BPNN 和 ELM 3 种模型构建方法,PLSR 模型表现最优。高光谱数据在完成特征波段初步提取工作后,进一步使用 CARS 算法优选特征波段,而 CARS 算法内置 PLS 模型,通过 PLS 模型评估变量重要性,所以 CARS 算法筛选出的变量是在 PLS 模型框架下被认为对目标变量解释能力较强的变量,因此 CARS 算法优选出的特征波段相比 BPNN 和 ELM 模型,对 PLSR 模型有更高的适配性。

5 结论

(1)小麦处于挑旗期时,PLSR 模型精度最高,各特征数据集构建的 PLSR 模型验证集 R^2 均达到 0.7 以上,综合对比各模型精度以及输入特征波段数量,以 SPA-CARS-MCX 数据构建的 PLSR 模型最优,共选择 3 个特征波段(842、850、858 nm),验证集 R^2 达到 0.782,RMSE 和 RPD 分别为 0.022 和 2.140,可以满足早期小麦条锈病监测要求。

(2)小麦在灌浆期时,PLSR 模型精度最高,且各数据集建模精度整体优于挑旗期,其中以 CA-CARS-MSC 数据集构建的 PLSR 模型最优,共提取 7 个特征波段(947、953、1 275、1 277、1 590、1 663、1 665 nm),验证集 R^2 达到 0.861,RMSE 和 RPD 分别为 0.094 和 2.687,实现了对小麦生育中期条锈病的良好监测。

参 考 文 献

[1] 廖娟, 陶婉琰, 臧英, 等. 农作物病虫害遥感监测关键技术研究进展与展望[J]. 农业机械学报, 2023, 54(11): 1-19.
LIAO Juan, TAO Wanyan, ZHANG Ying, et al. Research progress and prospect of key technologies in crop disease and insect pest monitoring[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 1-19. (in Chinese)

[2] 韩西红. 小麦条锈病发生规律及综合防治策略[J]. 农家参谋, 2022, 23(12): 84-86.
HAN Xihong. Occurrence law of wheat stripe rust and comprehensive control strategy[J]. The Farmers Consultant, 2022, 23(12): 84-86. (in Chinese)

[3] 王静, 景元书, 黄文江, 等. 冬小麦条锈病严重度不同估算方法对比研究[J]. 光谱学与光谱分析, 2015, 35(6): 1649-1653.
WANG Jing, JING Yuanshu, HUANG Wenjiang, et al. Comparative research on estimating the severity of yellow rust in winter wheat[J]. Spectroscopy and Spectral Analysis, 2015, 35(6): 1649-1653. (in Chinese)

[4] 胡根生,谢一帆,鲍文霞,等. 基于轻量型网络的无人机遥感图像中茶叶枯病检测方法[J]. 农业机械学报,2024, 55(4): 165-175.
HU Gensheng, XIE Yifan, BAO Wenxia, et al. Lightweight network for tea leaf blight detection in UAV remote sensing images

[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(4): 165－175. (in Chinese)
[5] 冯子恒, 李晓, 段剑钊, 等. 基于特征波段选择和机器学习的小麦白粉病高光谱遥感监测[J]. 作物学报, 2022, 48(9): 2300－2314.
FENG Ziheng, LI Xiao, DUAN Jianzhao, et al. Hyperspectral remote sensing monitoring of wheat powdery mildew based on feature band selection and machine learning[J]. Acta Agronomica Sinica, 2022, 48(9): 2300－2314. (in Chinese)
[6] 刘潭, 李子默, 冯帅, 等. 基于 LMPSO－SVM 的高光谱水稻稻瘟病害分级检测[J]. 农业机械学报, 2023, 54(11): 208－216.
LIU Tan, LI Zimo, FENG Shuai, et al. Classification detection of hyperspectral rice blast disease based on LMPSO－SVM[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(11): 208－216. (in Chinese)
[7] 谢亚平, 仝晓刚, 王晓慧. 基于高光谱的水稻稻曲病早期监测研究[J]. 农业机械学报, 2023, 54(9): 288－296.
XIE Yaping, TONG Xiaogang, WANG Xiaohui. Early monitoring of rice koji disease based on hyperspectroscopy[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(9): 288－296. (in Chinese)
[8] MUSTAFA G, ZHENG H, KHAN H I, et al. Enhancing fusarium head blight detection in wheat crops using hyperspectral indices and machine learning classifiers[J]. Computers and Electronics in Agriculture, 2024, 218(3): 108663.
[9] 黄亦其, 刘祥焕, 黄震宇, 等. 基于高光谱成像的甘蔗叶片早期轮斑病与锈病识别技术[J]. 农业机械学报, 2023, 54(4): 259－267.
HUANG Yiqi, LIU Xianghuan, HUANG Zhenyu, et al. Identification of early wheel spot and rust on sugarcane leaves based on spectral analysis[J]. Transactions of the Chinese Society for Agricultural Machinery, 2023, 54(4): 259－267. (in Chinese)
[10] ZHANG Hansu, HUANG Linsheng, HUANG Wenjiang, et al. Detection of wheat fusarium head blight using UAV-based spectral and image feature fusion[J]. Frontiers in Plant Science, 2022, 13(9): 1004427.
[11] MA Huiqin, HUANG Wenjiang, DONG Yingying, et al. Using UAV-based hyperspectral imagery to detect winter wheat fusarium head blight[J]. Remote Sensing, 2021, 13(15): 3024.
[12] 李杨, 李翠玲, 王秀, 等. 高光谱成像的黄瓜病虫害识别和特征波长提取方法[J]. 光谱学与光谱分析, 2024, 44(2): 301－309.
LI Yang, LI Cuiling, WANG Xiu, et al. Identification of cucumber disease and insect pest based on hyperspectral imaging [J]. Spectroscopy and Spectral Analysis, 2024, 44(2): 301－309. (in Chinese)
[13] ZHAO J, FANG Y, CHU G, et al. Identification of leaf-scale wheat powdery mildew (*Blumeria graminis* f. sp. Tritici) combining hyperspectral imaging and an SVM classifier[J]. Plants, 2020, 9(8): 936.
[14] HUANG Linshen, LIU Yong, HUANG Wenjiang, et al. Combining random forest and XGBoost methods in detecting early and mid-term winter wheat stripe rust using canopy level hyperspectral measurements[J]. Agriculture, 2022, 12(1):74.
[15] DENG Jie, WANG Rui, YANG Lujia, et al. Quantitative estimation of wheat stripe rust disease index using unmanned aerial vehicle hyperspectral imagery and innovative vegetation indices[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61(7): 4406111.
[16] GUO Anting, HUANG Wenjiang, DONG Yingying, et al. Wheat yellow rust detection using UAV-based hyperspectral technology[J]. Remote Sensing, 2021, 13(1): 123.
[17] 梅广源, 李荣, 梅新, 等. 基于 VSURF－CA 的小麦条锈病高光谱病情指数估测模型[J]. 中国农业科学, 2024, 57(3): 484－499.
MEI Guangyuan, LI Rong, MEI Xin , et al. A VSURF－CA based hyperspectral disease index estimation model of wheat stripe rust[J]. Scientia Agricultura Sinica, 2024, 57(3): 484－499. (in Chinese)
[18] YAO Zhifeng, LEI Yu, HE Dongjian. Early Visual detection of wheat stripe rust using visible/near-infrared hyperspectral imaging[J]. Sensors, 2019, 19(4): 952.
[19] 姚志凤,雷雨,何东健. 基于高光谱成像的小麦白粉病与条锈病识别[J]. 光谱学与光谱分析,2019,39(3):969－976.
YAO Zhifeng, LEI Yu, HE Dongjian. Identification of powdery mildew and stripe rust in wheat using hyperspectral imaging [J]. Spectroscopy and Spectral Analysis, 2019, 39(3):969－976. (in Chinese)
[20] 竞霞, 张腾, 白宗璠, 等. 融合 GA 与 SVR 算法的小麦条锈病特征优选与模型构建[J]. 农业机械学报, 2020, 51(11): 253－263.
JING Xia, ZHANG Teng, BAI Zongfan, et al. Feature selection and model construction of wheat stripe rust based on GA and SVR algorithm[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(11): 253－263. (in Chinese)
[21] 董锦绘, 杨小冬, 杨贵军, 等. 基于近地高光谱信息的小麦条锈病病情指数反演[J]. 麦类作物学报, 2016, 36(12): 1674－1680.
DONG Jinghui, YANG Xiaodong, YANG Guijun, et al. Inversion of wheat stripe rust disease index based on near ground hyperspectral data[J]. Journal of Triticeae Crops, 2016, 36(12): 1674－1680. (in Chinese)
[22] 田永超, 朱艳, 曹卫星, 等. 小麦冠层反射光谱与植株水分状况的关系[J]. 应用生态学报, 2004, 15(11): 2072－2076.
TIAN Yongchao, ZHU Yan, CAO Weixing, et al. Relationship between canopy reflectance and plant water status of wheat [J]. Chinese Journal of Applied Ecology, 2004, 15(11): 2072－2076. (in Chinese)
[23] COZZOLINO D, WILLIANS P J, HUFFMAN L L. An overview of pre-processing methods available for hyperspectral imaging applications[J]. Microchemical Journal, 2023, 193(10):109129.

[24] 辛波. 基于高光谱技术和机器学习的猕猴桃品质预测方法研究[D]. 雅安: 四川农业大学, 2022.
XIN Bo. Research on kiwifruit quality prediction method based on hyperspectral technology and machine learning[D]. Yaan: Sichuan Agricultural University, 2022. (in Chinese)
[25] 母雯竹, 张贵宇, 张维, 等. 基于 CARS-SPA 特征提取的黄水淀粉近红外光谱定量模型优化[J]. 食品科学, 2024, 45(19): 8-14.
MU Wenzhu, ZHANG Guiyu, ZHANG Wei, et al. Optimization of quantitative modeling of starch in Huangshui based on near-Infrared spectral feature extraction using competitive adaptive reweighted sampling combined with successive projections algorithm[J]. Food Science, 2024, 45(19): 8-14. (in Chinese)
[26] LI H, LIANG Y, XU Q, et al. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration[J]. Analytica Chimica Acta, 2009, 648(1): 77-84.
[27] 张源哲, 刘宇豪, 陆雨洁, 等. 基于偏最小二乘回归的荧光粉涂覆型白光 LED 的光谱预测研究[J]. 光谱学与光谱分析, 2022, 42(8): 2347-2352.
ZHANG Yuanzhe, LIU Yuhao, LU Yujie, et al. Study on the spectral prediction of phosphor-coated white LED based on partial least squares regression[J]. Spectroscopy and Spectral Analysis, 2022, 42(8): 2347-2352. (in Chinese)
[28] WANG L, ZENG Y, CHEN T. Back propagation neural network with adaptive differential evolution algorithm for time series forecasting[J]. Expert Systems with Applications, 2015, 42(2): 855-863.
[29] HUANG G,ZHU Q,SIEW C. Extreme learning machine: theory and applications[J]. Neurocomputing, 2005, 70(1): 489-501.
[30] 孔钰如, 王李娟, 张竞成, 等. 多源数据小麦条锈病预测研究[J]. 遥感技术与应用, 2022, 37(3): 571-579.
KONG Yuru, WANG Lijuan, ZHANG Jingcheng, et al. Research on prediction of wheat stripe rust with multi-source data[J]. Remote Sensing Technology and Application, 2022, 37(3): 571-579. (in Chinese)
[31] 刘凯, 王玉峰, 彭志青, 等. 基于机器学习的高光谱土壤养分特征波段提取方法[J]. 光学学报, 2025, 45(6): 0628001.
LIU Kai, WANG Yufeng, PENG Zhiqing, et al. A hyperspectral soil nutrient feature band extraction method based on machine learning[J]. Acta Optica Sinica, 2025, 45(6): 0628001. (in Chinese)
[32] 王小平, 赵传燕, 郭铌, 等. 黄土高原半干旱区春小麦冠层光谱对不同程度水分胁迫的响应[J]. 兰州大学学报(自然科学版), 2014, 50(3): 417-423.
WANG Xiaoping, ZHAO Chuanyan, GUO Ni, et al. Canopy hyperspectral reflectance response for spring wheat in different water stresses in semi-arid areas of the loess plateau[J]. Journal of Lanzhou University (Natural Sciences), 2014, 50(3): 417-423. (in Chinese)
[33] 商鸿生, 李玥仁. 小麦由水分胁迫诱导的抗条锈性[J]. 植物病理学报, 2004, 34(2):122-126.
SHANG Hongsheng, LI Yueren. Water stress-induced resistance to wheat stripe rust[J]. Acta Phytopathologica Sinica, 2004, 34(2): 122-126. (in Chinese)
[34] 蒋金豹, 陈云浩, 黄文江. 用高光谱微分指数监测冬小麦病害的研究[J]. 光谱学与光谱分析, 2007, 27(12): 2475-2479.
JIANG Jinbao, CHEN Yunhao, HUANG Wenjiang. Using hyperspectral derivative index to monitor winter wheat disease[J]. Spectroscopy and Spectral Analysis, 2007, 27(12): 2475-2479. (in Chinese)
[35] 丁玥, 张运, 王爱芳, 等. 基于无人机高光谱的水稻细菌条斑病监测[J]. 测绘与空间地理信息, 2022, 45(5): 44-47.
DING Yue, ZHANG Yun, WANG Aifang, et al. Surveillance of rice bacterial streak spot disease based on UAV hyperspectral [J]. Geomatics & Spatial Information Technology, 2022, 45(5): 44-47. (in Chinese)
[36] 凯楠, 刘咏梅, 李京忠, 等. 青海瑞香狼毒叶绿素含量高光谱预测模型[J]. 生态学杂志, 2017, 36(4): 1150-1157.
KAI Nan, LIU Yongmei, LI Jingzhong, et al. Hyperspectal predicting model of chlorophyll content of *Stellera chamaejasme* in Qinghai Province[J]. Chinese Journal of Ecology, 2017, 36(4): 1150-1157. (in Chinese)