

# 基于 CARS-RUN-ELM 算法的水稻叶片氮磷含量协同反演方法

许童羽<sup>1,2</sup>, 金忠煜<sup>1</sup>, 郭忠辉<sup>1</sup>, 杨柳<sup>1</sup>, 白驹驰<sup>1</sup>, 冯帅<sup>1</sup>, 于丰华<sup>1,2</sup>

(1. 沈阳农业大学信息与电气工程学院, 沈阳 110866; 2. 辽宁省农业信息化工程技术中心; 沈阳 110866)

**摘要:** 同时反演氮、磷元素含量相对于单一元素反演可以更加全面地表达水稻的营养状况, 为快速、准确获取水稻叶片氮、磷含量和精准变量施肥提供依据。该研究基于不同氮肥处理的田间小区试验, 获取水稻叶片氮、磷含量数据, 采用竞争性自适应重加权采样法 (Competitive Adaptive Reweighted Sampling, CARS) 筛选氮素与磷素共同特征波长, 以特征波长反射率为输入, 以化学方法测得叶片氮、磷元素含量为输出, 分别使用反向传播神经网络、极限学习机 (Extreme Learning Machine, ELM)、龙格-库塔算法优化极限学习机 (RUNge Kutta optimizer-Extreme Learning Machine, RUN-ELM) 构建水稻叶片氮、磷含量反演模型并分析。结果表明: 采用 CARS 方法有效去除了高光谱中大量无用、冗余信息, 得到 5 个氮、磷元素共同特征波长, 去除具有共线性的特征波长, 最后筛选出的特征波长分别是 451、488、780、813 nm。使用筛选后的特征波长反射率构建 RUN-ELM 水稻叶片氮、磷含量反演模型效果最好, 氮素训练集的决定系数  $R^2$  为 0.690, 均方根误差为 0.669 mg/g, 磷素训练集的决定系数  $R^2$  为 0.620, 均方根误差为 0.027 mg/g。通过对比, RUN-ELM 在预测能力、模型稳定性上优于反向传播神经网络以及 ELM 模型。综上研究, 基于 CARS-RUN-ELM 的水稻叶片氮、磷含量反演模型可以快速、准确获取水稻叶片氮、磷含量, 可为水稻精准施肥提供参考。

**关键词:** 氮素; 磷素; 遥感; 协同反演; 特征提取; 高光谱; 机器学习; 水稻

doi: 10.11975/j.issn.1002-6819.2022.10.018

中图分类号: O657.3; S511

文献标志码: A

文章编号: 1002-6819(2022)-10-0148-08

许童羽, 金忠煜, 郭忠辉, 等. 基于 CARS-RUN-ELM 算法的水稻叶片氮磷含量协同反演方法[J]. 农业工程学报, 2022, 38(10): 148-155. doi: 10.11975/j.issn.1002-6819.2022.10.018 http://www.tcsae.org

Xu Tongyu, Jin Zhongyu, Guo Zhonghui, et al. Simultaneous inversion method of nitrogen and phosphorus contents in rice leaves using CARS-RUN-ELM algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2022, 38(10): 148-155. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2022.10.018 http://www.tcsae.org

## 0 引言

氮、磷元素作为调控水稻生长发育的主要营养元素, 对水稻的结构、产量、品质等具有至关重要的作用<sup>[1]</sup>。水稻中氮元素的匮乏会导致水稻生理性状发生变化, 进而导致水稻生长缓慢, 叶绿素合成受阻, 叶片颜色淡黄<sup>[2-3]</sup>。水稻缺磷后, 稻株生长缓慢, 稻苗细瘦, 叶片直立不披, 叶色暗绿或灰绿带紫且抽穗困难, 穗粒少且不饱满<sup>[4]</sup>。近年来, 随着环保意识加强和绿色发展的倡导, 化肥零增长和负增长成为硬性要求, 对化肥施用提出更高要求。准确获取水稻叶片中氮元素、磷元素的含量, 快速准确检测作物的营养健康状况, 实现按需精准施肥是重要路径, 对水稻田间管理以及返青期、分蘖期精准施肥具有十分重要的意义。因此, 选题具有重要理论意义和实用性。相比具有伤害性的化学测量方法, 遥感监测方法具有快速、无损等优点, 在作物营养检测方面具有更好的应用前景<sup>[5]</sup>。高光谱数据中包含多种营养元素信息, 单一

元素反演建模会导致高光谱数据得不到高效的利用, 只能片面地表达水稻的营养情况。通过单一元素反演建模的方法获取水稻叶片多元素含量, 会存在模型与模型之间难以协作、效率低等问题。同时反演氮、磷元素含量可以高效利用高光谱数据, 减少建模时间, 更加全面地体现水稻的营养情况, 帮助种植者进行施肥决策, 使水稻能够高产、增效。

高光谱遥感技术已经普遍应用于水稻等作物的营养监测, 利用高光谱数据对水稻主要营养元素、叶面积指数等进行估测, 进而诊断水稻营养状况已经成为国内外农业信息化领域专家学者研究的主要内容之一<sup>[6]</sup>。Yang 等<sup>[7]</sup>利用玉米叶片光谱反射率构建了玉米冠层叶绿素含量估计模型。李金敏等<sup>[8]</sup>利用地物光谱获取的高光谱数据, 构建了一种深度森林模型进行叶片氮含量的反演。Lu 等<sup>[9]</sup>将地面和无人机的高光谱反射率进行非负矩阵分解, 利用其从冠层混合光谱中提取纯植被光谱, 然后基于提取的植被光谱和水稻植物钾积累建立偏最小二乘回归模型, 构建水稻植物钾积累估计模型。目前, 估测水稻叶片中营养元素含量的遥感监测方法主要基于敏感波段筛选和光谱指数等方法建立水稻叶片营养元素含量估算模型。秦占飞等<sup>[10]</sup>基于成像高光谱数据, 以组合波段

收稿日期: 2022-03-05 修订日期: 2022-05-10

基金项目: 辽宁省教育厅重点攻关项目 (LSNZD202005)

作者简介: 许童羽, 教授, 研究方向为农业信息化。

Email: xutongyu@syau.edu.cn

738 和 522 nm 光谱反射率的一阶导数构成的比值光谱指数, 构建水稻叶片全氮含量估测模型。Mahajan 等<sup>[11]</sup>构建植被指数“ $VI_{S\_670\_1090} ((R_{1090}-R_{670})/(R_{1090}+R_{670}))$ ”用于监测杂交水稻氮、磷、硫元素的含量。

由于高光谱与化学元素之间存在非线性关系, 而传统线性回归方法处理复杂的非线性问题时存在一定的缺陷, 但神经网络具有较强的非线性处理能力, 已被广泛应用于高光谱反演相关研究中<sup>[12]</sup>。Luo 等<sup>[13]</sup>利用高光谱数据, 采用多重逐步回归 (Multi-Scale Shape Regression, MSR)、偏最小二乘法 (Partial Least Squares, PLS)、反向传播神经网络 (Back Propagation Neural Network, BPNN) 和随机森林等方法构建了牡丹叶片花青素反演模型。冯帅等<sup>[14]</sup>选取特征波段和植被指数作为模型输入, 构建非支配的精英策略遗传算法优化极限学习机水稻叶片氮素含量反演模型。Annala 等<sup>[15]</sup>通过使用一维卷积神经网络反演叶片光学特性的随机模型, 估测植被叶绿素含量。Zhang 等<sup>[16]</sup>构建了一种基于迁移学习的高光谱冬小麦叶绿素含量估计方法。许童羽等<sup>[17]</sup>基于多植被指数组合, 构建天牛须搜索算法优化极限学习机水稻冠层叶片氮素含量反演模型。神经网络具有多输入多输出的优势, 具备同时反演多种元素含量的能力。但目前极少有这方面的相关研究, 所以本文使用神经网络进行多元素含量的反演。

高光谱数据具有数据量大、维度多的特点, 但也因此存在冗余信息, 导致计算、建模困难。使用高光谱降维方法对采集到的高光谱数据进行降维, 可以减少冗余信息, 加快建模速度<sup>[18]</sup>。为更全面表达水稻生长状况, 避免单一元素反演模型之间难以协作的问题, 本研究基于不同氮素水平的水稻田小区试验, 应用竞争自适应重加权采样法 (Competitive Adaptive Reweighted Sampling, CARS) 对高光谱数据进行降维, 使用龙格-库塔优化算法 (RUNge Kutta optimizer, RUN) 对极限学习机进行优化, 探索利用高光谱数据对水稻叶片氮素、磷素含量同时估算的可行性, 评估不同方法对同时反演氮素、磷素含量的准确性, 以期为描述水稻生长状况及水稻田间精准施肥提供理论依据和技术支撑。

## 1 材料与方法

### 1.1 研究区概况

研究区位于辽宁省鞍山市海城市耿庄镇, 地处东经 122°39'18", 北纬 40°58'58"。耿庄镇地处辽宁省海城市西北部, 属温带大陆性季风气候, 四季分明, 雨热同季, 降水集中, 日照时间长, 温差较大。年平均降水量 652 mm, 年平均气温 8.4 °C, 年平均日照时数约 3 663 h, 无霜期年平均 165 d, 封冻期 150 d 左右。该地区土地肥沃, 为水稻生长提供了良好的地理条件。

### 1.2 试验设计

试验设计为小区试验, 所用水稻品种为“北粳 1705”。该品种为粳型常规水稻品种, 生育期 157 d, 穗长 16.5 cm, 每公顷有效穗数 426 万穗, 结实率 84.7%, 适宜在辽宁中熟稻区种植, 试验区分布如图 1 所示。小区试验设 4 个

基本施氮水平, 分别为零氮肥、低氮肥、中氮肥和偏高氮肥, 对应数值分别为 0、750、1 125、1 500 kg/hm<sup>2</sup>, 分别记为 N0、N1、N2、N3。共划分 11 个小区, 其中零氮肥 3 组重复 (试验田编号为 N0T1、N0T2、N0T3)、低氮肥 3 组重复 (试验田编号为 N1T1、N1T2、N1T3)、中氮肥 3 组重复 (试验田编号为 N2T1、N2T2、N2T3)、偏高氮肥 2 组重复 (试验田编号为 N3T1、N3T2)。基本氮肥分两次施入, 分别为基肥 80%, 返青肥 20%。各小区磷肥、钾肥、硅肥施用量相同, 分别在返青期和分蘖期施用。由于施用氮肥也会对植物磷元素含量产生影响, 故磷肥未做变量处理。小区施肥和田间管理都按照当地正常作业水平进行。

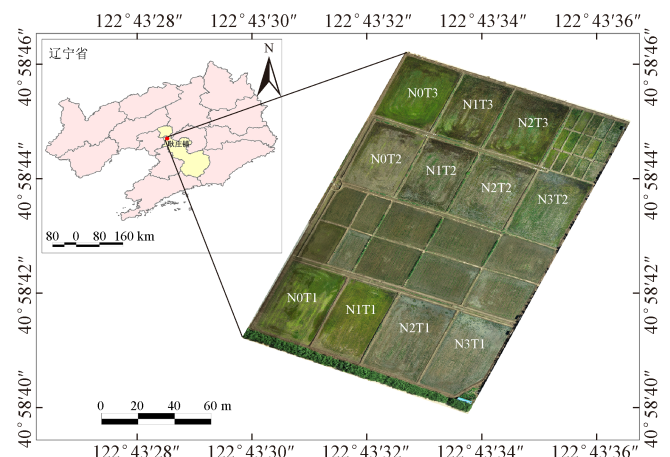


图 1 研究区概况

Fig.1 Overview of the study area

在水稻生长全生育期, 每隔 7~10 d 进行数据采集, 在每小区内选择具有代表性的 2~3 穴水稻进行叶片高光光谱数据以及叶片氮素、磷素含量数据的获取。

### 1.3 数据获取

#### 1.3.1 叶片氮素、磷素含量测定

对全生育期氮素含量进行测定, 在每个试验小区采样点进行破坏性采样, 为保证叶片鲜活, 采样时将水稻根部及根部的土壤一起挖出, 装入标记好试验编号、日期及序号的自封袋中, 放入低温保温箱并带回实验室。进行光谱采集后, 将采集的水稻所有新鲜叶片剪下分别装入自封袋中, 清洗叶片表面灰尘后在 105 °C 条件下杀青 30 min, 杀青完成后, 在 90 °C 的烘干箱中进行烘干, 烘干至恒质量后称量并粉碎。采用凯氏定氮法对水稻叶片氮素含量进行检测, 采用钒钼黄比色法对水稻叶片磷素含量进行测定。

#### 1.3.2 叶片高光光谱数据采集

采用美国海洋光学公司生产的光纤光谱仪 HR2000+, 其测定光谱范围为 400~1 000 nm, 光谱分辨率为 0.461 nm, 对带回实验室的新鲜水稻在 2 h 内进行叶片光谱采集。每次测量前, 均使用标准白板对仪器进行校正; 测量时, 探头紧压于叶片上, 通过自带 Ocean View 软件完成水稻叶片高光光谱数据的采集。

将试验中获得的 325 组数据, 使用 Excel 2010 进行整理, 采用标准差的方式分别进行异常值剔除, 同时采

用蒙特卡洛算法对异常光谱数据剔除, 最终得到 317 个样本。使用 KS (Kolmogorov-Smirnov) 检验方法, 将数据集随机按照 7:3 的比例进行划分, 划分后训练集 224 个, 验证集 93 个。如表 1 所示。

表 1 水稻叶片氮素、磷素质量分数统计

Table 1 Statistical of nitrogen and phosphorus contents in rice leaves (mg·g<sup>-1</sup>)

样本集 Sample set	样本数量 Sample size	元素类别 Element type	最小值 Minimum value	最大值 Maximum value	平均值 Mean value	标准差 SD
训练集 Training set	224	氮素	0.995	6.801	3.331	1.202
		磷素	0.188	0.427	0.325	0.045
验证集 Validation set	93	氮素	1.145	5.583	3.365	1.116
		磷素	0.214	0.443	0.329	0.053

## 1.4 数据处理与建模方法

### 1.4.1 高光谱曲线预处理

高光谱原始曲线预处理包括重采样和高斯平滑两个步骤。由于海阳光学光纤光谱仪采集的高光谱曲线分辨率为 0.461 nm, 本研究使用插值的方法对高光谱曲线进行重采样, 重采样后的高光谱曲线分辨率为 1 nm。由于使用仪器采集的高光谱曲线会因震动产生一定范围内的误差。因此, 使用 Matlab 中的 smooths 函数对高光谱曲线进行高斯平滑滤波处理。经过预处理后的高光谱曲线较预处理前的高光谱曲线更加平滑, 接近于真实光谱曲线, 如图 2 所示。

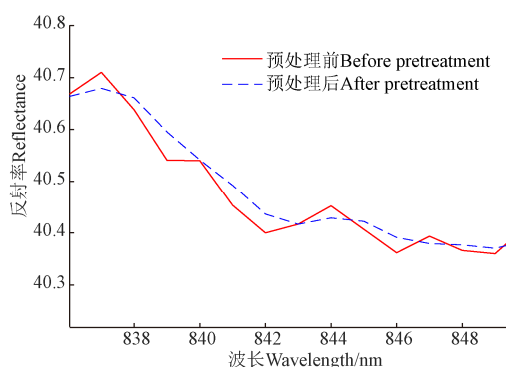


图 2 预处理前后高光谱曲线对比

Fig.2 Comparison of hyperspectral curves before and after pretreatment

### 1.4.2 高光谱降维方法

竞争性自适应重加权采样法是一种结合蒙特卡洛采样与偏最小二乘 (Partial Least Squares, PLS) 模型回归系数的特征变量选择方法, 模仿达尔文理论中的“适者生存”的原则<sup>[19]</sup>。CARS 算法中, 每次通过自适应加权采样 (Adaptive Reweighted Sampling, ARS) 保留 PLS 模型中回归系数绝对值权重较大的点作为新的子集, 去掉权重较小的点, 然后基于新的子集建立 PLS 模型, 经过多次计算, 选择 PLS 模型交互验证均方根误差 (Root Mean Square Error of Cross Validation, RMSECV) 最小的子集中的波长作为特征波长。

1) 采用蒙特卡洛采样法, 每次随机从校正集中选择 80% 的样本进入建模集。剩余 20% 作为预测集建立 PLS

模型, 并记录每一次采样过程中 PLS 模型回归系数的绝对值权重<sup>[20]</sup>。

$$w_i = |b_i| / \sum_{i=1}^m |b_i| \quad (1)$$

式中  $|b_i|$  为第  $i$  个变量的回归系数绝对值,  $w_i$  为第  $i$  个变量的回归系数绝对值权重,  $m$  为每次采样中剩余的变量数。

2) 利用指数衰减函数 (Exponentially Decreasing Function, EDF) 强行去除回归系数绝对值权重相对较小的波长。在第  $j$  次基于 MC 采样建立 PLS 模型时, 根据 EDF 得到保留的波长点的比例  $R_j$  为

$$R_j = \mu e^{-kj} \quad (2)$$

式中  $\mu$  和  $k$  是常数, 可以按照以下两种情况计算: 在第一次采样并进行相应计算时, 所有的波长都参与了建模分析, 因此此时保留的波长点的比例为 1; 在最后一次采样在 (第  $N$  次) 完成并进行相应计算时, 只剩下两个波长参与 PLS 建模, 此时保留的波长点的比例为  $2/n$ , 其中  $n$  是原始波长点数。由以上最初及最后一次采样的情况可知,  $\mu$  和  $k$  的计算公式为

$$\mu = \left(\frac{n}{2}\right)^{\frac{1}{N-1}}, k = \frac{\ln\left(\frac{n}{2}\right)}{N-1} \quad (3)$$

3) 在每次采样时, 都从上一次采样时的变量数中采用自适应加权采样 (Adaptive Reweighted Sampling, ARS) 选择数量为  $R_j \times n$  个波长变量, 进行 PLS 建模, 计算 RMSECV。

4) 在  $N$  次采样完成之后, CARS 算法得到了  $N$  组候选的特征波长子集, 以及对应的 RMSECV 值, 选择 RMSECV 最小值所对应的波长变量子集为特征波长。

### 1.4.3 反演模型的构建

极限学习机 (Extreme Learning Machine, ELM) 是一类基于前馈神经网络构建的机器学习方法, 已广泛应用于遥感反演领域<sup>[21-24]</sup>。但使用单纯的极限学习机方法用于反演, 仍然存在精度并不高、误差较大的问题。智能优化算法一般都是建立在生物智能或物理现象基础上的随机搜索算法, 算法一般不要求目标函数和约束的连续性与凸性, 对计算中数据的不确定性也有很强的适应能力。因此使用智能优化算法对极限学习机方法进行优化, 进而提高反演精度, 减小误差。RUN 是由 Ahmadianfar 等<sup>[25]</sup>于 2021 年提出的, 该算法利用计算梯度搜索概念来指导寻优, 具有寻优能力强、收敛速度快等特点。同时, 为了评估运行算法的效率, 使用两个著名的排名测试 (即 Friedman 和 Quade 测试) 与其他 5 个优化器 (灰狼优化算法 (Grey Wolf Optimizer, GWO)<sup>[26]</sup>、布谷鸟算法 (Cuckoo Search, CS)<sup>[27]</sup>、水循环算法 (Water Cycle Algorithm, WCA)<sup>[28]</sup>、鲸鱼优化算法 (Whale Optimization Algorithm, WOA)<sup>[29]</sup>、入侵杂草优化算法 (Invasive Weed Optimization, IWO)<sup>[30]</sup>) 进行了比较。研究表明, RUN 提供了非常有竞争力的结果, 并且在大多数测试函数中都优于其他优化器。



本研究需同时反演水稻叶片氮素和磷素含量，因此选用可以支持多元回归输出的神经网络模型，分别是 BP 神经网络、极限学习机（Extreme Learning Machine, ELM）、龙格-库塔算法优化极限学习机（RUNge Kutta optimizer-Extreme Learning Machine, RUN-ELM）3 种方法建模，通过决定系数  $R^2$  以及均方根误差（Root Mean Square Error, RMSE）来评价模型的精准度和可靠性，挑选最优的水稻叶片氮、磷元素反演模型。

极限学习机模型随机产生输入层与隐含层间的连接权值及隐含层神经元的阈值，在训练过程中只需设置隐含层神经元的个数，便可获得唯一的最优解，然而这会导致 ELM 模型的稳定性和泛化能力较差。本研究使用寻优能力强、收敛速度快的龙格-库塔智能优化算法对 ELM 进行优化，算法具体优化步骤如下<sup>[25,31]</sup>：

步骤 1：初始化种群，设置龙格-库塔智能优化算法的种群大小，随机生成  $L$  个位置并计算初始位置。

步骤 2：搜索机制，基于龙格-库塔（Runge-Kutta, RK）方法，使用一组随机解搜索决策空间，并实现适当的全局和局部搜索，使用随机参数来增加优化过程中最佳解的重要性。

步骤 3：更新解决方案，以一组随机解开始优化过程。在每次迭代中，使用 RK 方法更新最优解位置。

步骤 4：解质量增强(Enhanced Solution Quality, ESQ)使用 RK 搜索机制生成位置  $x_{n+1}$ ，然后使用 ESQ 机制在搜索空间中探索有希望的区域。

步骤 5：判断是否获得最优解，若获得到最优解，则进行下一步，若未获得最优解，则重复步骤 2~步骤 4。

步骤 6：根据最优解设置 ELM 基本参数。

步骤 7：根据最优解设置最有权值和阈值。

步骤 8：计算误差，若误差在允许范围内，则输出结果，若误差在允许范围外，则重复步骤 6、步骤 7。

整体优化流程图如图 3 所示。

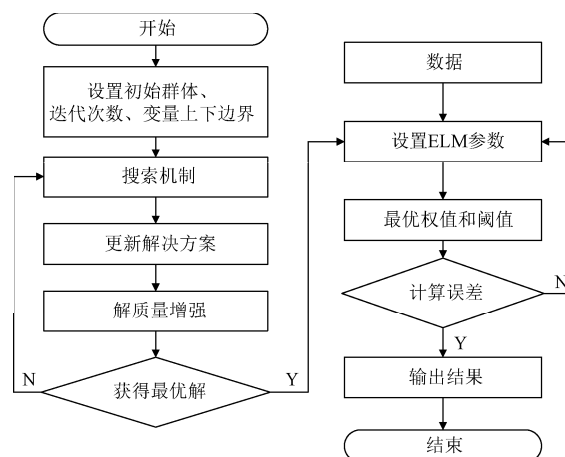


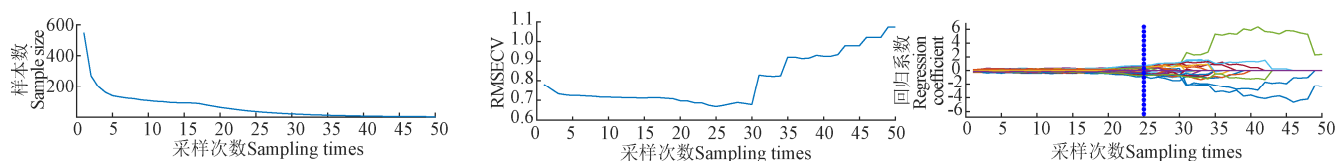
图 3 基于 RUN 优化 ELM 的流程图

Fig.3 Flow chart of Extreme Learning Machine (ELM) optimization based on RUNge Kutta

## 2 结果与分析

### 2.1 特征波长的选择

本研究同时对水稻叶片氮素、磷素进行反演建模。在特征波长选择过程中，使用 Matlab 软件平台，采用竞争自适应重加权采样法分别对氮素和磷素提取特征波段。由于采集的光谱数据 400~450 nm 处噪声较大，故将筛选范围定为 451~1 000 nm。将磷素与氮素分别进行 50 次 CARS 变量选择过程。反复迭代采样次数并比较每次采样的 RMSECV 值，直至找到最小的 RMSECV 值所包含的最优变量子集<sup>[32]</sup>。由图 4、图 5 看出，随着采样次数增加，被选择波长数逐渐减少，下降趋势逐渐趋于平缓。同时，RMSECV 值缓慢下降至最低点后，逐渐上升，变量回归系数绝对值不断增大。表明在筛选过程中，首先剔除了与氮素含量、磷素含量无关的波长，使 RMSECV 值减小后，剔除了与组分相关的波长，信息丢失导致 RMSECV 增大。



注：RMSECV 为交叉验证均方根误差。下同。

Note: RMSECV is the root mean square error of cross validation. Same below.

图 4 筛选氮素特征波长过程

Fig.4 Process of screening characteristic wavelengths of nitrogen

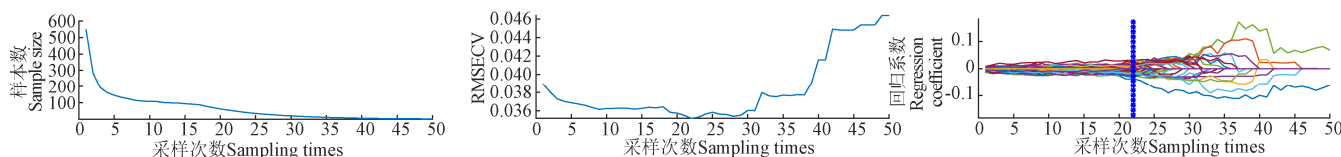


图 5 筛选磷素特征波长过程

Fig.5 Process of screening characteristic wavelengths of phosphorus

图 4、图 5 中垂直蓝点标记的位置对应筛选过程中 RMSECV 值最小，对应的变量数即为筛选后最优波长数。最后，使用 CARS 对氮素提取了 25 个特征波长，对磷素

提取了 22 个特征波长。提取结果如图 6 所示。将提取出的特征波长做交集，筛选出 5 个共同的特征波长，分别是 451、488、780、781、813 nm。考虑到相邻波长存在

共线性, 将特征波长反射率与氮素含量、磷素含量进行相关性分析, 将相关性较低的 781 nm 去除, 将剩余 4 个波长反射率作为模型输入。

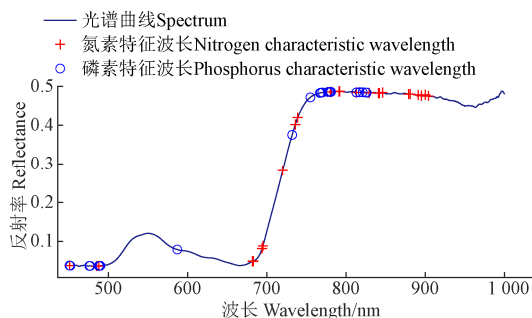


图 6 特征波长筛选结果图

Fig.6 Screening results of characteristic wavelength

## 2.2 水稻叶片氮素、磷素含量反演建模与分析

本研究将使用 CARS 筛选出的 4 个特征波长对应反射率为模型输入, 以水稻叶片氮素含量以及磷素含量作

为模型输出, 构建水稻叶片氮素、磷素含量 BP 神经网络、ELM、RUN-ELM 反演模型。将模型参数调整至最佳状态后, 建模结果如图 7 所示。由图 7 可知, RUN-ELM 反演模型结果优于 BP 神经网络反演模型以及 ELM 反演模型, 氮素训练集、验证集决定系数  $R^2$  分别为 0.690、0.596, RMSE 分别为 0.669、0.774 mg/g; 磷素训练集、验证集决定系数  $R^2$  为 0.620、0.422, RMSE 分别为 0.027、0.041 mg/g。其中, ELM 反演模型结果最差, 训练集与验证集之间相差较多, 氮素训练集、验证集决定系数  $R^2$  分别为 0.727、0.491, RMSE 分别为 0.628、0.876 mg/g; 磷素训练集、验证集决定系数  $R^2$  为 0.606、0.327, RMSE 分别为 0.028、0.044 mg/g, 模型出现了不稳定的现象。BP 神经网络反演模型结果介于二者之间, 氮素训练集、验证集决定系数  $R^2$  分别为 0.586、0.590, RMSE 分别为 0.774、0.734 mg/g; 磷素训练集、验证集决定系数  $R^2$  为 0.416、0.358, RMSE 分别为 0.034、0.042 mg/g。由此可见, RUN-ELM 提高了模型的稳定性和反演精度。

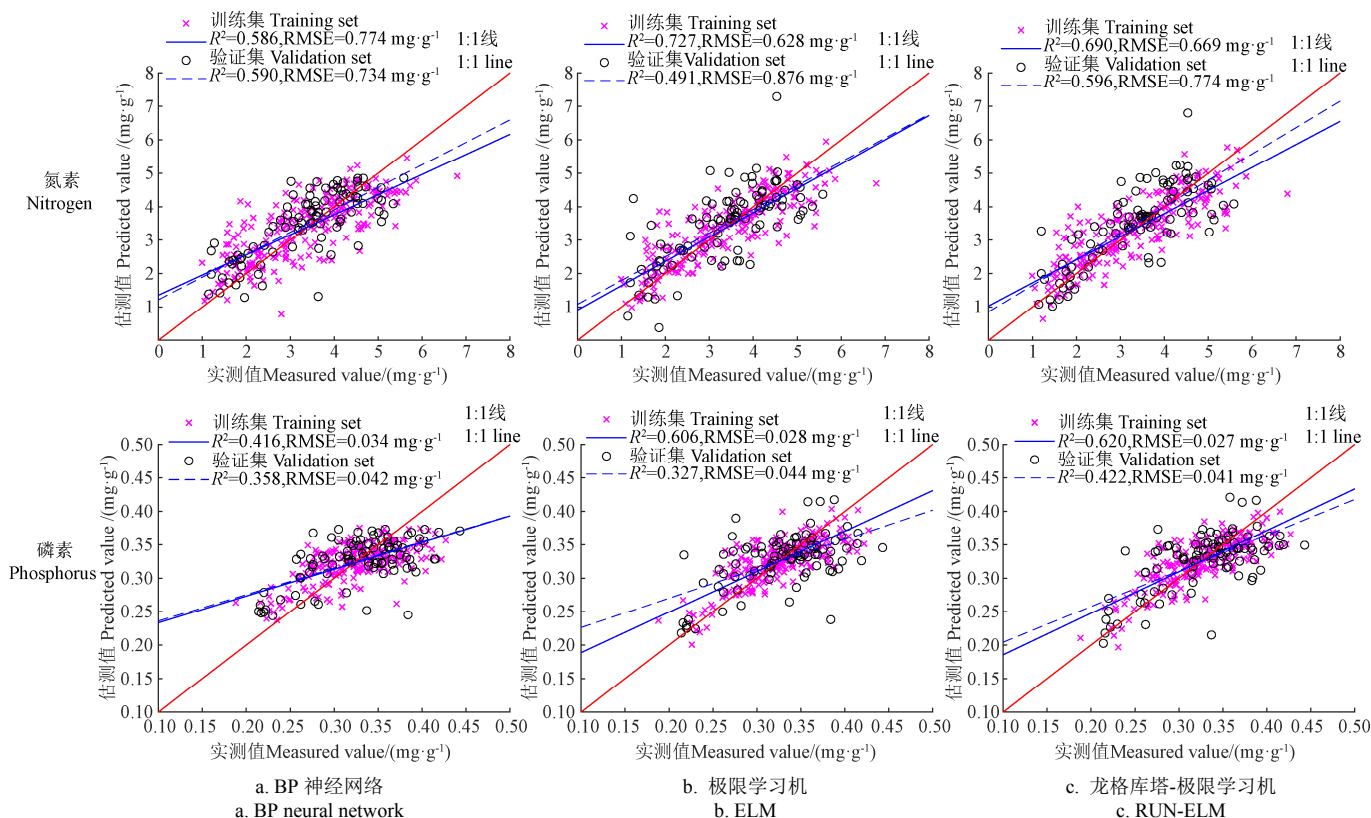


图 7 3 种水稻叶片氮、磷含量反演模型检验结果

Fig7 Test results of three inversion models for nitrogen and phosphorus contents in rice leaves

## 3 讨论

磷是植物生长发育所必需的大量元素之一, 也是制约水稻产量形成的重要因素。氮作为组成重要大分子 (如核酸、蛋白质和叶绿素等) 的关键成分, 同样是水稻生长不可或缺的大量元素<sup>[33]</sup>。氮、磷元素同时反演, 可以更加全面的了解水稻营养情况及生长状况。本文以东北地区水稻为研究对象, 利用海洋光学光纤光谱仪采集高光谱数据, 运用竞争自适应重加权采样法筛选水稻叶片氮、磷元素共有特征波长, 构建了 BP 神经网络、极限学

习机、RUN-ELM 3 种水稻叶片氮、磷含量反演模型。在高光谱数据降维方面, 有专家学者在研究中表明 CARS 和 Random Frog 算法在土壤有机碳的光谱反演中各具优势<sup>[34]</sup>, 后续研究中将进一步探讨 Random Frog 算法是否在水稻叶片营养元素反演中也具备潜在的应用价值。使用 CARS 算法提取后的氮素特征波长在 450~900 nm 范围内, 磷素特征波长在 450~826 nm 范围内。叶片内部组织结构的变化, 会影响水稻叶片氮素含量的变化, 这些变化均会在高光谱反射率上体现出来<sup>[35-36]</sup>。而水稻在轻度缺磷时, 叶绿素浓度有可能升高<sup>[37]</sup>。因此, 氮素与

磷素共同影响着一些波段反射率的变化。故本研究将氮素与磷素的特征波长做交集处理, 筛选出共同特征波长, 分别为 451、488、780、781、813 nm, 分布在可见光和近红外范围内, 其中 780 与 781 nm 相邻, 存在共线性, 因此将相关性较低的 781 nm 去除。已有研究表明, 在作物生长过程中, 氮肥促进叶绿素增加, 作物光合作用增强, 对蓝光和红光的吸收增强, 红光与蓝光反射率逐渐减小<sup>[38]</sup>。本研究筛选的特征波长中 451 nm 位于蓝光范围内, 780 nm 位于红光范围。因此, 本研究挑选出的特征波长符合农学客观事实。

将 451、488、780、813 nm 特征波长反射率作为输入, 化学实验测得的氮、磷含量作为输出, 分别采用 BP 神经网络、ELM、RUN-ELM 对氮、磷元素含量同时反演。从反演效果来看, BP 神经网络反演模型在同时反演两种营养元素时, 氮元素反演精度一般, 磷元素反演精度较差, 决定系数  $R^2$  小于 0.5。ELM 反演模型训练集氮元素、磷元素反演精度均较好, 但验证集与训练集之间相差较多, 出现了模型不稳定、过拟合的现象。RUN-ELM 反演模型氮、磷两种元素训练集决定系数  $R^2$  均大于 0.6, 并且训练集与验证集决定系数的差值要小于 ELM 反演模型训练集与验证集决定系数的差值。其原因是 ELM 的初始权值和阈值都是随机产生, 每次产生的初始权值和阈值具有盲目性。而 RUN 优化算法依据 RK 方法中提出的计算梯度的概念, 利用计算出的斜率作为搜索逻辑, 搜索空间中有前途的区域, 采用增强解质量 (Enhanced Solution Quality, ESQ) 机制, 避免了局部最优解, 提高了收敛速度, 进而可以为 ELM 提供一个更好的权值和阈值, 提高了模型的精度和稳定性。

本研究基于叶片尺度高光谱数据构建氮素、磷素反演模型, 尽管 RUN-ELM 模型整体反演效果较好, 但仍然存在不足。磷素反演效果相对氮素反演效果较差, 究其原因可能是由于磷素数据分布不均匀。在后续的研究中, 将进一步探讨如何提高多元素反演模型的精度。

## 4 结 论

本文基于不同氮肥施用量处理的水稻小区试验, 破坏性采样法获取全生育期水稻叶片, 使用化学实验方法获取叶片氮素、磷素含量, 采用竞争自适应重加权采样法 (Competitive Adaptive Reweighted Sampling, CARS) 提取高光谱特征波长, 使用 3 种方法构建水稻叶片氮、磷元素含量反演模型并比较其差异, 主要结论如下:

1) 针对高光谱数据, 使用竞争性自适应重加权采样法筛选的特征波长分别是 451、488、780、813 nm, 用于同时反演磷素和氮素。

2) 分别采用 BP 神经网络 (Back Propagation Neural Network, BPNN)、极限学习机 (Extreme Learning Machine, ELM)、龙格-库塔算法优化极限学习机 (RUNge Kutta optimizer-Extreme Learning Machine, RUN-ELM) 构建水稻叶片氮、磷元素含量反演模型。结果表明, RUN-ELM 模型的估测效果优于 BP 神经网络以及 ELM 模型, 氮素训练集与验证集  $R^2$  分别为 0.690、0.596, 均

方根误差分别为 0.669、0.774 mg/g; 磷素训练集与验证集  $R^2$  分别为 0.620、0.422, 均方根误差分别为 0.027、0.041 mg/g。本文建立的水稻叶片氮、磷元素含量反演模型可以同时水稻叶片两种营养元素进行估测, 大大提高了获取水稻叶片营养元素含量的效率, 并且能够保证估测结果的精度和模型的稳定性, 为准确获取水稻氮、磷元素含量提供了一种新方法。

## 【参 考 文 献】

- [1] 王洁, 肖瑶, 张国豪, 等. 氮磷调控对杂交水稻产量及品质的影响[J]. 广东农业科学, 2021, 48(8): 131-138.  
Wang Jie, Xiao Yao, Zhang Guohao, et al. Effects of Nitrogen and phosphorus regulation on the yield and quality of hybrid rice[J]. Guangdong Agricultural Sciences, 2021, 48(8): 131-138. (in Chinese with English abstract)
- [2] 吴言, 唐宁, 张边江. 缺氮对不同粳稻品种光合特性的影响[J]. 湖北农业科学, 2014, 53(8): 1762-1764, 1828.  
Wu Yan, Tang Ning, Zhang Bianjiang. Effects of nitrogen deficiency on the photosynthetic characteristics of different varieties of japonica rice[J]. Hubei Agricultural Sciences, 2014, 53(8): 1762-1764, 1828. (in Chinese with English abstract)
- [3] 顾清, 邓劲松, 陆超, 等. 基于光谱和形状特征的水稻扫描叶片氮素营养诊断[J]. 农业机械学报, 2012, 43(8): 170-174, 159.  
Gu Qing, Deng Jinsong, Lu Chao, et al. Diagnosis of rice nitrogen nutrition based on spectral and shape characteristics of scanning leaves[J]. Transactions of the Chinese Society for Agricultural Machinery, 2012, 43(8): 170-174, 159. (in Chinese with English abstract)
- [4] 李成, 杨晓新, 李施杨. 水稻缺素症状的研究[J]. 农业与技术, 2003(5): 85-90.  
Li Cheng, Yang Xiaoxin, Li Shiyang. Studies on the nutritional deficiency symptom of rice[J]. Agriculture and Technology, 2003(5): 85-90. (in Chinese with English abstract)
- [5] 印玉明, 王永清, 马春晨, 等. 利用日光诱导叶绿素荧光监测水稻叶片叶绿素含量[J]. 农业工程学报, 2021, 37(12): 169-180.  
Yin Yuming, Wang Yongqing, Ma Chunchen, et al. Monitoring of chlorophyll content in rice canopy and single leaf using sun-induced chlorophyll fluorescence[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(12): 169-180. (in Chinese with English abstract)
- [6] 陈晓凯, 李粉玲, 王玉娜, 等. 无人机高光谱遥感估算冬小麦叶面积指数[J]. 农业工程学报, 2020, 36(22): 40-49.  
Chen Xiaokai, Li Fenling, Wang Yuna, et al. Estimation of winter wheat leaf area index based on UAV hyperspectral remote sensing[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(22): 40-49. (in Chinese with English abstract)
- [7] Yang H Y, Ming B, Nie C W, et al. Maize canopy and leaf chlorophyll content assessment from leaf spectral reflectance: estimation and uncertainty analysis across growth stages and vertical distribution[J]. Remote Sensing, 2022, 14(9): 2115.
- [8] 李金敏, 陈秀青, 杨琦, 等. 基于高光谱的水稻叶片氮含量估计的深度森林模型研究[J]. 作物学报, 2021, 47(7): 1342-1350.  
Li Jinmin, Chen Xiuqing, Yang Qi, et al. Deep learning models for estimation of paddy rice leaf nitrogen concentration based on canopy hyperspectral data[J]. Acta Agronomica Sinica, 2021, 47(7): 1342-1350. (in Chinese with English abstract)

- [9] Lu J S, Li W Y, Yu M L, et al. Estimation of rice plant potassium accumulation based on non-negative matrix factorization using hyperspectral reflectance[J]. Precision Agriculture, 2021, 22(1): 51-74.
- [10] 秦占飞, 常庆瑞, 谢宝妮, 等. 基于无人机高光谱影像的引黄灌区水稻叶片全氮含量估测[J]. 农业工程学报, 2016, 32(23): 77-85.  
Qin Zhanfei, Chang Qingrui, Xie Baoni, et al. Rice leaf nitrogen content estimation based on hyperspectral imagery of UAV in Yellow River diversion irrigation district[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(23): 77-85. (in Chinese with English abstract)
- [11] Mahajan G R, Pandey R N, Sahoo R N, et al. Monitoring nitrogen, phosphorus and sulphur in hybrid rice (*Oryza sativa* L.) using hyperspectral remote sensing[J]. Precision Agriculture, 2017, 18(5): 736-761.
- [12] 沈润平, 丁国香, 魏国栓, 等. 基于人工神经网络的土壤有机质含量高光谱反演[J]. 土壤学报, 2009, 46(3): 391-397.  
Shen Runping, Ding Guoxiang, Wei Guoshuan, et al. Retrieval of soil organic matter content from hyper-spectrum based on ANN[J]. Acta Pedologica Sinica, 2009, 46(3): 391-397. (in Chinese with English abstract)
- [13] Luo L L, Chang Q R, Gao Y F, et al. Combining different transformations of ground hyperspectral data with Unmanned Aerial Vehicle (UAV) images for anthocyanin estimation in tree peony leaves[J]. Remote Sensing, 2022, 14(9): 2271.
- [14] 冯帅, 曹英丽, 许童羽, 等. 高光谱和 NSGA2-ELM 算法的粳稻叶片氮素含量反演[J]. 光谱学与光谱分析, 2020, 40(8): 2584-2591.  
Feng Shuai, Cao Yingli, Xu Tongyu, et al. Inversion based on high spectrum and NSGA2-ELM algorithm for the nitrogen content of Japonica rice leaves[J]. Spectroscopy and Spectral Analysis, 2020, 40(8): 2584-2591. (in Chinese with English abstract)
- [15] Annala L, Honkavaara E, Tuominen S, et al. Chlorophyll concentration retrieval by training convolutional neural network for Stochastic model of Leaf Optical Properties (SLOP) inversion[J]. Remote Sensing, 2020, 12(2): 283.
- [16] Zhang Y, Hui J, Qin Q M, et al. Transfer-learning-based approach for leaf chlorophyll content estimation of winter wheat from hyperspectral data[J]. Remote Sensing of Environment, 2021, 267: 112724.
- [17] 许童羽, 邢思敏, 于丰华, 等. 基于多植被指数组合的 BAS-ELM 粳稻冠层氮含量反演方法[J]. 沈阳农业大学学报, 2021, 52(5): 577-585.  
Xu Tongyu, Xing Simin, Yu Fenghua, et al. A BAS-ELM Inversion method of japonica rice canopy nitrogen content based on a combination of multiple vegetation indices[J]. Journal of Shenyang Agricultural University, 2021, 52(5): 577-585. (in Chinese with English abstract)
- [18] Li H T, Xie X T, Yang X K, et al. An Integrated model of summer and winter for chlorophyll-a retrieval in the pearl river estuary based on hyperspectral data[J]. Remote Sensing, 2022, 14(9): 2270.
- [19] Li H D, Liang Y Z, Xu Q S, et al. Key wavelengths screening using competitive adaptive reweighted sampling method for multivariate calibration[J]. Analytica Chimica Acta, 2009, 648(1): 77-84.
- [20] 路皓翔, 张静, 李灵巧, 等. 最小角回归结合竞争性自适应重加权采样的近红外光谱波长选择[J]. 光谱学与光谱分析, 2021, 41(6): 1782-1788.  
Lu Haoxiang, Zhang Jing, Li Lingqiao, et al. Least angle regression combined with competitive adaptive re-weighted sampling for NIR spectral wavelength selection[J]. Spectroscopy and Spectral Analysis, 2021, 41(6): 1782-1788. (in Chinese with English abstract)
- [21] 许童羽, 郭忠辉, 于丰华, 等. 采用 GA-ELM 的寒地水稻缺氮量诊断方法[J]. 农业工程学报, 2020, 36(2): 209-218.  
Xu Tongyu, Guo Zhonghui, Yu Fenghua, et al. Genetic algorithm combined with extreme learning machine to diagnose nitrogen deficiency in rice in cold region[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(2): 209-218. (in Chinese with English abstract)
- [22] 陈媛媛, 王志斌, 王召巴, 等. 区间极限学习机结合遗传算法用于红外光谱气体浓度反演的研究[J]. 光谱学与光谱分析, 2014, 34(5): 1244-1248.  
Chen Yuanyuan, Wang Zhibin, Wang Zhaoba, et al. Research on concentration retrieval of gas FTIR spectra by interval extreme learning machine and genetic algorithm[J]. Spectroscopy and Spectral Analysis, 2014, 34(5): 1244-1248. (in Chinese with English abstract)
- [23] 陆军胜, 陈绍民, 黄文敏, 等. 采用 SE<sub>PLS</sub>-ELM 模型估算夏玉米地上部生物量和叶面积指数[J]. 农业工程学报, 2021, 37(18): 128-135.  
Lu Junsheng, Chen Shaomin, Huang Wenmin, et al. Estimation of aboveground biomass and leaf area index of summer maize using SE<sub>PLS</sub>-ELM model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(18): 128-135. (in Chinese with English abstract)
- [24] Guo L, Sun X R, Fu P, et al. Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas[J]. Geoderma, 2021, 398: 115118.
- [25] Ahmadianfar I, Heidari A A, Gandomi A H, et al. RUN beyond the metaphor: An efficient optimization algorithm based on Runge Kutta method[J]. Expert Systems with Applications, 2021, 181: 115079.
- [26] Sm A, Smm B, Al A. Grey wolf optimizer[J]. Advances in Engineering Software, 2014, 69: 46-61.
- [27] Rajabioun R. Cuckoo optimization algorithm[J]. Applied Soft Computing, 2011, 11(8): 5508-5518.
- [28] Eskandar H, Sadollah A, Bahreininejad A, et al. Water cycle algorithm-A novel metaheuristic optimization method for solving constrained engineering optimization problems[J]. Computers & Structures, 2012, 110: 151-166.
- [29] Mirjalili S, Lewis A. The whale optimization algorithm[J]. Advances in Engineering Software, 2016, 95: 51-67.
- [30] Rad H S, Lucas C. A recommender system based on invasive weed optimization algorithm[C]. 2007 IEEE Congress on Evolutionary Computation, Singapore: Institute of Electrical and Electronics Engineers, 2007: 4297-4304.
- [31] 冯建强, 孙诗一. 四阶龙格-库塔法的原理及其应用[J]. 数学学习与研究, 2017(17): 3-5.
- [32] 彭海根, 金楹, 詹筱国, 等. 近红外光谱技术结合竞争性自适应重加权采样变量选择算法快速测定土壤水解性氮含量[J]. 分析测试学报, 2020, 39(10): 1305-1310.  
Peng Haigen, Jin Ying, Zhan Youguo, et al. Quantitative determination of hydrolytic nitrogen content in soil by near infrared spectroscopy combined with competitive adaptive reweighted sampling variable selection algorithm[J]. Journal of Instrumental Analysis, 2020, 39(10): 1305-1310. (in Chinese with English abstract)
- [33] 孙雅菲, 宋科, 秦秦, 等. 磷酸盐转运蛋白 OsPT4 影响水稻氮磷积累与利用的机理研究[J]. 中国水稻科学, 2021, 35(6): 565-572.  
Sun Yafei, Song Ke, Qin Qin, et al. Research on the mechanism of OsPT4 regulating the accumulation and utilization of nitrogen and phosphorus in rice[J]. Chinese

- Journal of Rice Science, 2021, 35(6): 565-572. (in Chinese with English abstract)
- [34] Xu L, Hong Y S, Wei Y, et al. Estimation of organic carbon in anthropogenic soil by VIS-NIR spectroscopy: Effect of variable selection[J]. Remote Sensing, 2020, 12(20): 3394.
- [35] 杨峰, 范亚民, 李建龙, 等. 高光谱数据估测稻麦叶面积指数和叶绿素密度[J]. 农业工程学报, 2010, 26(2): 237-243.
- Yang Feng, Fan Yamin, Li Jianlong, et al. Estimating LAI and CCD of rice and wheat using hyperspectral remote sensing data[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2010, 26(2): 237-243. (in Chinese with English abstract)
- [36] 李永梅, 张立根, 张学俭. 水稻叶片高光谱响应特征及氮素估算[J]. 江苏农业科学, 2017, 45(23): 210-213.
- [37] 王磊, 白由路. 基于光谱理论的作物营养诊断研究进展[J]. 植物营养与肥科学报, 2006(6): 902-912.
- Wang Lei, Bai Youlu. Research advance on plant nutrition diagnosis based on spectral theory[J]. Journal of Plant Nutrition and Fertilizers, 2006(6): 902-912. (in Chinese with English abstract)
- [38] 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(4): 279-283.
- Liu Weidong, Xiang Yueqin, Zheng Lanfen, et al. Relationships between Rice LAI, CH. D and Hyperspectra Data[J]. National Remote Sensing Bulletin, 2000(4): 279-283. (in Chinese with English abstract)

## Simultaneous inversion method of nitrogen and phosphorus contents in rice leaves using CARS-RUN-ELM algorithm

Xu Tongyu<sup>1,2</sup>, Jin Zhongyu<sup>1</sup>, Guo Zhonghui<sup>1</sup>, Yang Liu<sup>1</sup>, Bai Juchi<sup>1</sup>, Feng Shuai<sup>1</sup>, Yu Fenghua<sup>1,2</sup>

(1. College of Information and Electrical Engineering, Shenyang Agricultural University, Shenyang 110866, China; 2. Liaoning Agricultural Information Technology Center, Shenyang 110866, China)

**Abstract:** Chemical fertilizers can be the zero and negative growth for the high requirement from the environmental and green development in recent years. Precise fertilization on demand can depend mainly on the rapid and accurate detection of crop nutrition and health status in fields. Among them, a synergistic inversion of nitrogen and phosphorus content can be expected to more comprehensively express the nutritional conditions of rice, compared with the single element inversion. It is also of great significance to the rice field management and accurate fertilization at the greening and tillering stage. In this study, a series of field plot experiments were conducted to realize the different treatments of nitrogen fertilizer. A chemical experiment was selected to obtain the nitrogen and phosphorus content in rice leaves, while a marine optical fiber spectrometer was used for the hyperspectral data of rice leaves. The data sets of nitrogen content were then sorted after measurement. A Kolmogorov-Smirnov test was also utilized to randomly divide the data sets into the 224 training and 93 verification sets, according to the ratio of 7:3. Competitive Adaptive Reweighted Sampling (CARS) was then used to screen the common characteristic wavelengths of nitrogen and phosphorus from the data sets. As such, the reflectivity of characteristic wavelengths was set as the input, whereas, the measured contents of nitrogen and phosphorus in the rice leaves were used as the output. A Back Propagation (BP) neural network, Extreme Learning Machine (ELM), and Runge-Kutta optimizer-Extreme Learning Machine (RUN-ELM) were used to construct the inversion models of nitrogen and phosphorus contents in the rice leaves. The results show that the CARS effectively removed a large number of redundant information in the hyperspectra data, where five common characteristic wavelengths of nitrogen and phosphorus were obtained to remove the collinearity characteristic wavelengths. After that, the characteristic wavelengths were selected as 451, 488, 780, and 813 nm. The best performance of the RUN-ELM model was achieved to retrieve the nitrogen and phosphorus contents in the rice leaves using the selected reflectance of characteristic wavelength. The determination coefficient and Root Mean Square Error (RMSE) of the nitrogen training set were 0.690 and 0.669 mg/g, respectively, while the determination coefficient and RMSE of the phosphorus training set were 0.620 and 0.027 mg/g, respectively. By contrast, the RUN-ELM model was superior to the BP neural network and ELM model in the prediction and simulation. Furthermore, the higher accuracy and stability of the ELM model were realized to improve the better weight and threshold, where the local optimal solution was avoided for the higher convergence speed than before. The reason was that the promising region was searched in the space using the calculated slope as the search logic and the Enhanced Solution Quality (ESQ) mechanism, according to the calculating gradient during Runge-Kutta (RK) optimization. To sum up, the CARS-RUN-ELM inversion model can rapidly and accurately extract the nitrogen and phosphorus content in the rice leaves. The high accuracy and stability of the model can greatly contribute to effectively gaining the nutrient element contents of rice leaves. The finding can provide a strong reference to detect the nitrogen and phosphorus content for the precise fertilization of rice on demand.

**Keywords:** nitrogen; phosphorus; remote sensing; cooperative inversion; feature extraction; hyperspectral; machine learning; rice