

ASD Field Spec3 野外便携式高光谱仪诊断冬小麦氮营养

刘昌华¹, 方 征¹, 陈志超^{1*}, 周 兰², 岳学智¹,
王 哲¹, 王春阳¹, Yuxin Miao³

(1. 河南理工大学测绘与国土信息工程学院, 焦作 454000; 2. 中国农业大学资源与环境学院, 北京 100193;
3. Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN, 55108, USA)

摘 要: 氮素营养诊断关键在于氮营养指数 (nitrogen nutrient index, NNI) 预测。对于冬小麦氮营养指数预测模型而言, 如何选取预处理方法和建模方法不一而足, 不同预处理和模型选取对预测结果精度的影响程度目前还不清楚。该研究以 ASD Field Spec3 野外便携式高光谱仪采集乐陵市冬小麦冠层高光谱数据, 采用 10 种光谱预处理方法并结合 3 种模型 (偏最小二乘回归、BP 神经网络和随机森林算法) 建立多种冬小麦氮营养指数高光谱预测模型。对比模型预测精度表明最佳的高光谱建模方法为随机森林算法结合 SG 卷积平滑预处理所建模型 (预测集 $R^2=0.795$, RMSE=0.125, RE=11.7%) 精度高、可靠性强, 是筛选出最佳的冬小麦氮营养指数高光谱预测模型。该研究结果对冬小麦氮营养指数高光谱预测建模具有科学价值, 为筛选最优高光谱预处理方法和预测模型提供技术参考。

关键词: 光谱分析; 氮; 诊断; 冬小麦; 模型; 氮营养指数

doi: 10.11975/j.issn.1002-6819.2018.19.021

中图分类号: S127

文献标志码: A

文章编号: 1002-6819(2018)-19-0162-08

刘昌华, 方 征, 陈志超, 周 兰, 岳学智, 王 哲, 王春阳, Yuxin Miao. ASD Field Spec3 野外便携式高光谱仪诊断冬小麦氮营养[J]. 农业工程学报, 2018, 34(19): 162-169. doi: 10.11975/j.issn.1002-6819.2018.19.021 http://www.tcsae.org
Liu Changhua, Fang Zheng, Chen Zhichao, Zhou Lan, Yue Xuezhi, Wang Zhe, Wang Chunyang, Yuxin Miao. Nitrogen nutrition diagnosis of winter wheat based on ASD Field Spec3[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(19): 162-169. (in Chinese with English abstract)
doi: 10.11975/j.issn.1002-6819.2018.19.021 http://www.tcsae.org

0 引 言

对作物快速而准确的氮素营养诊断可以为精准农业管理提供技术支持, 对实现作物高产与氮肥的高效利用有重要意义^[1]。冬小麦传统的氮素营养诊断的方法主要有外观诊断、化学诊断和叶绿素计法等^[2-4], 然而这些方法破坏小麦植株取样工作量大, 室内分析化验操作复杂, 时效性差^[5]。因此遥感作为快速、准确、实时的无损氮营养诊断手段得到了学者们的广泛应用与研究。

已有研究表明高光谱采集的光谱数据除样品自身的光谱信息外还会包含样品背景、杂散光和电磁噪音等因素的影响^[6-7], 消除噪音等影响可以有效提高模型的预测能力^[8]。目前已有学者基于高光谱成像仪对油菜^[9-10]、生菜^[11]、柑橘^[12]和玉米^[13]等不同作物进行研究, 通过分析比较 Savitzky-Golay (SG) 卷积平滑滤波、基线校正、多元散射校正、标准正态变换、一阶导数和正交信号校正等预处理方法使作物叶片氮素含量估测模型预测性能得

到明显提升, 然而不同作物对于最佳的预处理也不尽相同。乔星星等^[14]对冬小麦冠层高光谱数据进行 SG 平滑的预处理建立偏最小二乘模型 (partial least squares regression, PLSR) 模型估测冬小麦生物量, 通过对 6 种不同平滑点数的比较得出 9 点平滑处理为估测生物量的最佳平滑点数, 但是真正提高光谱监测模型的预测精度还需在此基础上, 再结合其他预处理方法和化学计量学模型的优化方法。Li 等^[15]通过标准正态变换 (standard normal variate transformation, SNV) 技术对不同地区的冬小麦高光谱数据进行光散射校正结合 SG 平滑建立预测氮含量的 PLSR 模型, 建模精度 R^2 都达到 0.8 以上。姚霞等、汤守鹏等^[16-17]通过多元散射校正 (multiple scatter correction, MSC)、SG 和导数等预处理基础上建立 PLSR 模型和 BP 神经网络模型估测冬小麦氮含量, 比较不同预处理和不同模型的结果表明 MSC+SG 二阶导的 BP 神经网络建模精度与预测精度都达到最佳。

对于冬小麦氮营养指数预测模型而言, 如何选取预处理方法和建模方法不一而足, 不同预处理和模型选取对预测结果精度的影响程度目前还不清楚。以往对冬小麦氮含量预测模型的研究都使用一种或少数几种的预处理比较分析。为了更为系统地比较不同预处理方法对冬小麦氮营养指数预测模型精度的影响程度, 本研究选用 10 种预处理方法结合偏最小二乘、BP 神经网络和随机森林算法 3 种建模方法建立冬小麦氮素营养指数的高光谱预测模型, 对比筛选出预测氮营养指数的最佳的预处理方法和

收稿日期: 2018-05-03 修订日期: 2018-08-23

基金项目: 国家自然科学基金资助项目 (41371105); 河南省软科学研究计划项目 (162400410058); 河南省高等学校重点科研项目 (18A420001); 河南省智慧中原地理信息技术协同创新中心“开放课题 (2016A002)

作者简介: 刘昌华, 男, 湖南衡东人, 博士, 教授, 博士生导师, 主要从事遥感应用与土地生态方面的研究与教学工作。Email: lchhj@163.com

*通信作者: 陈志超, 副教授, 主要从事农业遥感与精准农业方面的研究。Email: logczc@163.com

模型,为基于高光谱的氮营养指数快速预测提供技术参考。

1 材料与方法

1.1 试验设计

试验于 2016—2017 年在山东省乐陵市南夏家村农业局基地进行 (37°41'58.59"N, 117°08'41.65"E)。试验小区设置 2 个品种供试作物分别为济麦 22、鲁原 502, 试验共有 6 个氮梯度处理, 施氮量分别为不施氮肥 (N_1), 120 (N_2), 180 (N_3), 240 (N_4), 300 (N_5), 280 kg/hm² (N_6), 其中 N_6 为农民施肥模式。每个处理 3 次重复, 共 6×3×2 个小区, 小区面积 7 m×10 m, 采用区组随机排列, 取样与采集光谱数据日期为 2016 年 4 月 1 日 (返青期)、2016 年 4 月 17 日 (拔节期)、2016 年 4 月 28 日 (孕穗期)、2016 年 5 月 7 日 (扬花期)、2017 年 3 月 27 日 (返青期)、2017 年 4 月 15 日 (拔节期)、2017 年 4 月 22 日 (孕穗期)、2017 年 5 月 9 日 (扬花期)。

1.2 数据获取

1.2.1 农学参数获取

每个试验小区选择具有代表性的 1 m 双行冬小麦植株, 清理泥土后置于烘箱中以 105 °C 杀青 30 min, 再以 75 °C 恒温烘干至质量恒定, 测定植株生物量 (干质量)。将冬小麦植株烘干后的植株样本研磨粉碎后, 用凯氏定氮仪测定植株氮浓度。

临界氮浓度 (N_c) 是根据氮浓度稀释模型计算出小麦获得地上部最大生物量所需的最低氮浓度值, 华北平原冬小麦的临界氮浓度根据 Yue 等^[18]提出的公式进行计算, 计算公式如下:

$$N_c = 4.15W^{-0.38} \quad (1)$$

式中 W 为地上部生物量, Mg/hm², 当地上部生物量在 1~10 Mg DM/ha 时, 可用式 (1) 计算, 当地上部生物量小于 1Mg DM/ha 时, 则临界氮浓度为定值即 $N_c=4.15\%$ DM。

氮营养指数^[19] (nitrogen nutrient index, NNI) 可以定性定量诊断作物氮营养的充足、过剩以及不足的程度, 表 1 为冬小麦建模样本的氮营养指数 (NNI) 的基本情况。

$$NNI = \frac{Na}{N_c} \quad (2)$$

式中 NNI 是氮营养指数, Na 是植株实际氮浓度。

若 $NNI < 1$ 时, 表示植株氮肥使用量不足, 缺少氮素; $NNI=1$ 时, 表示植株氮肥使用量适宜; $NNI>1$, 表示植株氮肥使用量过剩, 对氮过量吸收。

表 1 冬小麦样本集氮营养指数统计

Table 1 Statistical characteristics of nitrogen nutrient index (NNI) for winter wheat samples

样本集 Sample set	样本数 Sample number	最小值 Min	最大值 Max	均值 Mean	标准差 Standard deviation	变异系数 Coefficient of variation/%
建模集 Calibration set	175	0.28	1.72	0.89	0.31	34
预测集 Prediction set	88	0.31	1.52	0.86	0.28	32
总样本集 Whole set	263	0.28	1.72	0.88	0.30	34

1.2.2 光谱数据获取

高光谱测定采用美国 ASD (Analytical Spectral Devices) 公司的 ASD Field Spec3 野外便携式高光谱仪, 其波长范围在 350~2 500 nm, 采样间隔为 1.4 nm (350~1 000 nm) 和 2 nm (1 000~2 500 nm), 重采样间隔为 1 nm, 视场角为 25°。冠层光谱测量时选择在晴朗无风的天气, 每次测量时间为 10:00—14:00。每次采集光谱前先以白板定标, 仪器探头垂直向下, 仪器在冠层上方 80 cm 处, 对每个试验小区选定具有代表性的植株冠层, 采集 10 次光谱值, 剔除异常值后平均光谱, 作为该试验小区的最终光谱, 每个关键生育时期采集 36 小区的冠层光谱值, 如图 1 所示。

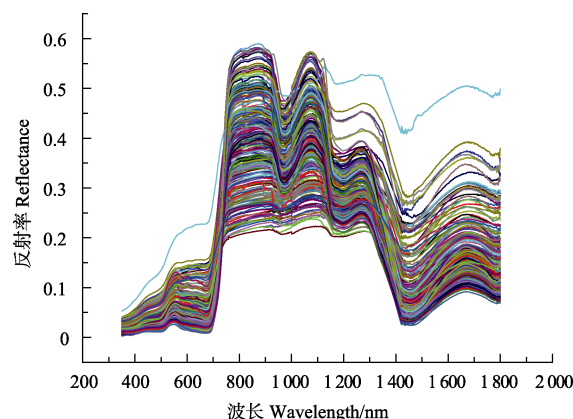


图 1 总样本集的光谱反射率

Fig.1 Spectral reflectance of whole sample set

1.3 光谱参数定义

根据上述小区测得高光谱原始光谱反射率, 本文采用冬小麦冠层高光谱常见的 19 个高光谱特征参数^[20] (表 2), 其特征参数是从原始光谱和一阶微分中选取绿峰、红谷和“三边” (红边, 黄边, 蓝边) 的光谱位置和反射率作为高光谱位置变量; 选取“三边”一阶微分波段值得总和作为高光谱面积变量; 使用绿峰、红谷、“三边”反射率以及“三边”面积变量的比值和归一化运算作为高光谱植被指数变量, 本文建模的自变量为 19 个高光谱特征参数。

1.4 光谱数据预处理

本文选用的高光谱特征参数为建模的自变量, 高光谱特征参数只用到 350~800 nm 的波段, 因此去掉 800 nm 以后的波段。为了更好的系统比较不同的预处理方法对反演模型预测精度的影响, 本文对原始光谱数据采用 Savitzky-Golay 卷积平滑滤波、基线校正、多元散射校正、标准正态变换等常用的算法进行预处理, 预处理方法如表 3 所示。Savitzky-Golay 平滑可以提高信噪比, 过滤噪声^[21]; 基线校正可以消减仪器背景或基线波动对光谱的影响^[22]; 多元散射校正对光谱散射的影响可以有效消减^[23]; 标准正态变换可以减弱表面散射以及光程变化对漫反射光谱的影响^[24], 预处理后的光谱曲线如图 2 所示。

表 2 高光谱参数的定义
Table 2 Definition of hyperspectral parameters

分类 Classification	定义 Definition	参数符号 Parameter symbols
基于高光谱位置变量	蓝边 (490–530 nm) 内一阶微分光谱中的最大值	D_b
	蓝边 D_b 对应的波长位置 (nm)	λ_b
	黄边 (550–582 nm) 内一阶微分光谱中的最大值	D_y
	黄边 D_y 对应的波长位置 (nm)	λ_y
	红边 (680–780 nm) 内一阶微分光谱中的最大值	D_r
	红边 D_r 对应的波长位置 (nm)	λ_r
	波长 510–560 nm 范围内最大的波段反射率	R_g
	R_g 对应的波长位置 (nm)	λ_g
	波长 640–680 nm 范围内最小的波段反射率	R_r
	R_r 对应的波长位置 (nm)	λ_o
基于高光谱面积变量	蓝边波长范围内一阶微分波段值的总和	SD_b
	黄边波长范围内一阶微分波段值的总和	SD_y
	红边波长范围内一阶微分波段值的总和	SD_r
基于植被指数变量	绿峰反射率(R_g)与红谷反射率(R_r)的比值	R_g/R_r
	绿峰反射率(R_g)与红谷反射率(R_r)的归一化值	$(R_g-R_r)/(R_g+R_r)$
	红边内一阶微分的总和(SD_r)与蓝边内一阶微分的总和(SD_b)的比值	SD_r/SD_b
	红边内一阶微分的总和(SD_r)与黄边内一阶微分的总和(SD_y)的比值	SD_r/SD_y
	红边内一阶微分的总和(SD_r)与蓝边内一阶微分的总和(SD_b)的归一化值	$(SD_r-SD_b)/(SD_r+SD_b)$
	红边内一阶微分的总和(SD_r)与黄边内一阶微分的总和(SD_y)的归一化值	$(SD_r-SD_y)/(SD_r+SD_y)$

表 3 光谱数据预处理方法
Table 3 Pre-processing methods applied to the spectral curves

预处理方法 Pre-processing method	简写 Abbreviation
Savitzky-Golay 卷积平滑 Savitzky-Golay smoothing	SG
基线校正 Baseline correction	BC
多元散射校正 Multiple scatter correction	MSC
标准正态变换 Standard normal variate transformation	SNV
去趋势算法 De-trending	D
Savitzky-Golay 平滑+基线校正 Savitzky-Golay smoothing+ Baseline correction	SG+BC
Savitzky-Golay 平滑+多元散射校正 Savitzky-Golay smoothing+ Multiple scatter correction	SG+MSC
Savitzky-Golay 平滑+标准正态变换 Savitzky-Golay smoothing+ Standard normal variate transformation	SG+SNV
Savitzky-Golay 平滑+去趋势算法 Savitzky-Golay smoothing+ De-trending	SG+D
标准正态变换+去趋势算法 Standard normal variate transformation+ De-trending	SNV+D

1.5 模型构建

本文选用偏最小二乘回归模型、反向传播神经网络回归模型以及随机森林算法 3 种方法进行建模, 其中构建模型的自变量为 19 个高光谱特征参数, 因变量为氮营养指数, 依据检验模型的确定系数 R^2 、相对误差值(relative error, RE) 以及均方根误差值 (root mean square error, RMSE) 来检验模型的精度与可靠性, 筛选最佳模型。

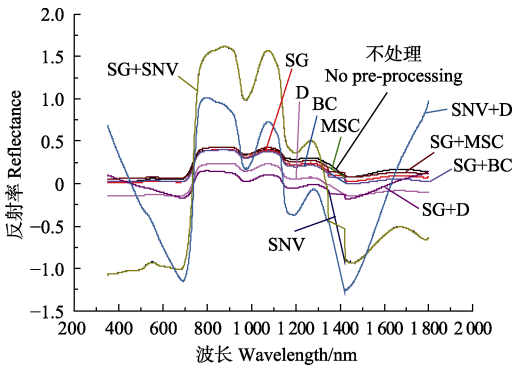


图 2 预处理后的光谱反射率
Fig.2 Spectral reflectance after pre-processing

1.5.1 偏最小二乘模型

偏最小二乘模型 (partial least squares regression, PLSR) 是一种集主成分分析, 多元线性回归分析和最小二乘回归方法于一体的建模方法, 在建模过程中利用主成分分析来判断加入建模的自变量能否显著提高预测能力, 因此该方法能够对多个高光谱特征之间的多重自相关性进行解释并利用所有有效数据构建回归模型^[25], 故在作物高光谱应用领域受到较大关注。PLSR 模型在 Unscrambler 9.7 (CAMO ASA, Trondheim, Norway) 软件中构建, 采用内部交叉验证均方根误差 (RMSE) 及决定系数 (R^2) 来优化建模参数。

1.5.2 BP 神经网络

反向传播神经网络 (back-propagation artificial neural network, BPANN) 是目前应用最为广泛的学习算法, 通过正向传播与反向传播的 2 个循环过程不断把误差分摊给各层单元进行权值修正, 直到网络输出误差减少到可接受的程度, 从而实现变量的回归。BP 神经网络具有高度自学习和自适应的能力, 适合于求解内部机制复杂的问题, 是解决非线性问题的有效方法^[26]。BP 神经网络在 MATLAB R2014b 软件中构建, 神经网络采用常用的 3 层结构, 输入层到隐含层激励函数为 tansig 型函数, 隐含层到输出层激励函数为 purelin 函数, 学习函数为梯度下降动量权重函数, 训练函数 Levenberg-Marquardt 算法, 隐含层节点数 10, 迭代次数 100 次, 最大确认失败次数 15, 学习率 0.01, 学习目标为 0.000 1。

1.5.3 随机森林算法

随机森林 (Random forest, RF) 算法是通过多个决策树组成的一个随机建立的森林, 是采用 bootstrap 重抽样法从训练集中抽取 m 个样本构建 m 个回归树, 在回归树的每个节点从 n 个 (小于自变量个数) 随机抽取的自变量中选择分割点对变量空间进行递归二叉分割, 未抽到成为袋外样本, 使误差的计算能够处理袋外样本数据, 将这些树的平均数作为因变量预测的结果。RF 算法不仅训练和预测速度快^[27], 还对多元共线性不敏感, 对异常值和噪声具有较好的容忍度^[28-29]。RF 算法在 MATLAB R2014b 软件中构建, 随机森林算法中的分类树的个数为 2000, 分割变量为 2。

2 结果与分析

2.1 偏最小二乘回归

采用 PLSR 结合不同预处理建立氮素营养指数估测模型的建模精度与预测精度如表 4 所示，建模集模型决定系数 R^2 平均系数为 0.683，最大决定系数为 0.789，均方根误差 RMSE 最小值为 0.142，相对中误差 RE 最小值为 12.3%；预测集模型决定系数 R^2 平均系数为 0.588，最大决定系数为 0.717，均方根误差 RMSE 最小值为 0.15，相对中误差 RE 最小值为 12.8%。

比较得知，使用偏最小二乘法构建模型时使用 SG、SNV、SG+SNV 和 SG+BC 的预处理方法效果较好，其中效果最优的预处理为 SG 平滑。图 3 为 PLSR 模型结合最佳预处理建立的氮素营养指数高光谱模型精度结果。

表 4 偏最小二乘回归氮营养指数 (NNI) 预测模型统计结果
Table 4 Statistics results for spectral models of nitrogen nutrient index(NNI) by partial least squares regression(PLSR)

预处理 Pre-processing method	建模 Calibration			预测 Validation		
	R^2	RMSE	RE/%	R^2	RMSE	RE/%
不处理	0.643	0.187	16.7	0.543	0.190	17.9
SG	0.789	0.142	12.3	0.717	0.150	12.8
BC	0.719	0.164	14.3	0.635	0.171	15.0
MSC	0.624	0.189	17.4	0.552	0.189	17.5
SNV	0.728	0.161	14.1	0.518	0.196	18.1
D	0.627	0.189	17.5	0.556	0.188	17.9
SG+BC	0.717	0.165	14.4	0.663	0.164	15.0
SG+MSC	0.619	0.191	17.6	0.531	0.194	18.3
SG+SNV	0.759	0.152	13.3	0.609	0.177	16.5
SG+D	0.656	0.182	15.8	0.556	0.189	17.6
SNV+D	0.632	0.188	16.9	0.585	0.182	16.6

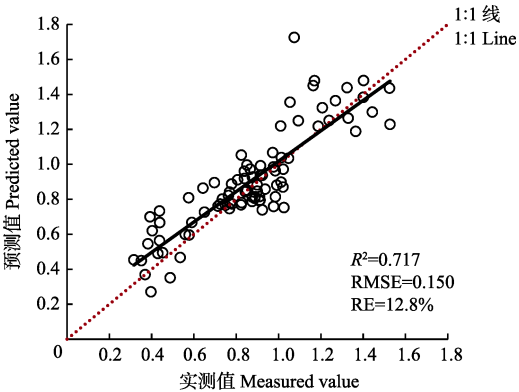


图 3 SG-PLSR 模型预测氮素营养指数与实际指数验证结果
Fig.3 Result of validation between measured and predicted NNI by using SG-PLSR model

2.2 误差反向传播神经网络

采用 BP 神经网络方法结合不同预处理建立氮素营养指数估测模型的建模精度与预测精度如表 5 所示，建模集模型决定系数 R^2 平均系数为 0.834，最大决定系数为 0.959，均方根误差 RMSE 最小值为 0.115，相对中误差 RE 最小值为 9.8%；预测集模型决定系数 R^2 平均系数为

0.714，最大决定系数为 0.780，均方根误差 RMSE 最小值为 0.133，相对中误差 RE 最小值为 12.3%。

从表 5 的建模精度结果可以得知 BP 神经网络构建的回归模型相比 PLSR 模型的建模精度显著提高，通过 BP 神经网络结合 10 种预处理的建模模型决定系数都为 0.8 以上，平均决定系数更是从 PLSR 模型的 0.683 提升到 0.834。其中结合 SG、SG+MSC、SG+SNV 和 SNV+D 的预处理建模模型精度都达到 0.85 以上，预测效果除 BC、MSC 和 SNV 外决定系数都达到 0.7 以上。结合 SG 预处理的建模效果最好。图 4 为 BP 神经网络模型结合最佳预处理建立的氮素营养指数高光谱模型精度结果。

表 5 BP 神经网络氮营养指数 (NNI) 预测模型统计结果
Table 5 Statistics results for spectral models of nitrogen nutrient index(NNI) by back-propagation(BP) artificial neural network

预处理 Pre-processing method	建模 Calibration			预测 Validation		
	R^2	RMSE	RE/%	R^2	RMSE	RE/%
不处理	0.811	0.135	11.3	0.678	0.159	14.9
SG	0.861	0.115	9.8	0.780	0.133	12.3
BC	0.813	0.134	11.1	0.695	0.156	14.3
MSC	0.805	0.137	11.5	0.659	0.165	15.4
SNV	0.824	0.130	11.2	0.699	0.155	14.6
D	0.846	0.122	10.3	0.702	0.154	14.6
SG+BC	0.805	0.137	11.7	0.725	0.148	13.6
SG+MSC	0.852	0.119	10.1	0.745	0.143	12.7
SG+SNV	0.855	0.118	9.7	0.736	0.145	13.2
SG+D	0.846	0.121	9.9	0.714	0.151	14.2
SNV+D	0.852	0.119	9.9	0.718	0.150	14.0

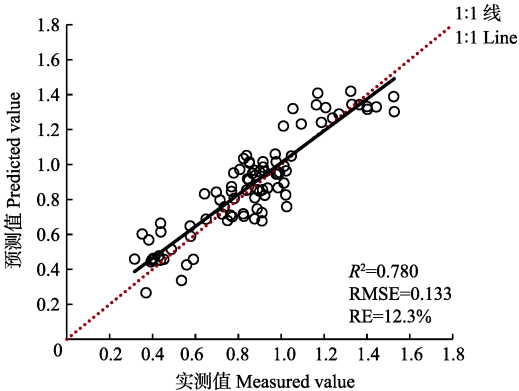


图 4 SG-BP 模型预测氮素营养指数与实际指数验证结果
Fig.4 Result of validation between measured and predicted NNI by using SG-BP model

2.3 随机森林回归

使用 RF 算法结合不同预处理建立氮素营养指数估测模型的建模精度与预测精度如表 6 所示，建模集模型决定系数 R^2 平均系数为 0.945，最大决定系数为 0.959，均方根误差 RMSE 最小值为 0.061，相对中误差 RE 最小值为 5.3%；预测集模型决定系数 R^2 平均系数为 0.742，最大决定系数为 0.795，均方根误差 RMSE 最小值为 0.125，相对中误差 RE 最小值为 11.7%。

表 6 随机森林氮营养指数 (NNI) 预测模型统计结果
Table 6 Statistics for spectral models of nitrogen nutrient index by random forest (RF)

预处理 Pre-processing method	建模 Calibration			预测 Validation		
	R^2	RMSE	RE/%	R^2	RMSE	RE/%
不处理	0.941	0.075	6.5	0.736	0.146	13.4
SG	0.959	0.061	5.3	0.795	0.125	11.7
BC	0.945	0.073	6.2	0.732	0.147	13.5
MSC	0.939	0.077	6.6	0.717	0.150	13.7
SNV	0.943	0.074	6.3	0.734	0.146	13.3
D	0.947	0.071	6.2	0.743	0.143	13.3
SG+BC	0.943	0.074	6.4	0.741	0.144	13.4
SG+MSC	0.940	0.076	6.5	0.735	0.146	13.4
SG+SNV	0.945	0.072	6.1	0.758	0.139	12.6
SG+D	0.944	0.071	6.4	0.738	0.144	13.0
SNV+D	0.950	0.065	5.8	0.728	0.140	13.1

从表 6 的建模精度结果可以得知 RF 构建的回归模型相比 PLSR 模型和 BP 神经网络的建模精度皆有明显提高, 通过 RF 结合不同的预处理的建模模型决定系数都达到 0.9 以上, 预测模型精度 R^2 都达到 0.7 以上, 综合各预处理精度, RF 算法预测集平均 R^2 为 0.742, RMSE 为 0.143, RE 为 13.1%, 比偏最小二乘模型与 BP 神经网络的平均决定系数 R^2 有所提高, 均方根误差 RMSE 降低, 相对误差 RE 依次降低了 3.4 与 0.9 个百分点。其中结合 SG 预处理的建模效果最好。图 5 为 RF 算法模型结合最佳预处理建立的氮素营养指数高光谱模型精度结果。通过 3 种模型建模比较, SG 卷积平滑比其他 9 种预处理的预测集模型平均 R^2 提高范围为 0.054~0.121, RMSE 平均降低范围为 0.016~0.032, RE 平均降低范围为 1.7~3.1 个百分点。

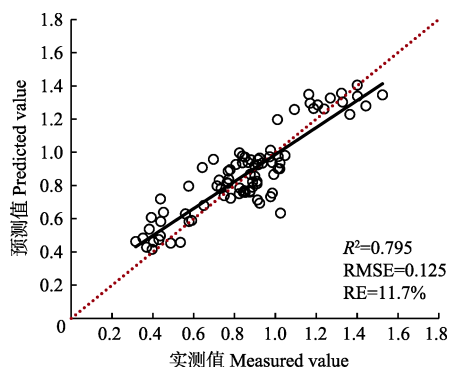


图 5 SG-RF 模型预测氮素营养指数与实际指数验证结果

Fig.5 Result of validation between measured and predicted NNI by using SG-RF model

3 讨论

国内外学者对基于遥感的氮素营养诊断研究多采用氮浓度、吸氮量、叶面积指数以及生物量等指标为依据指导施肥, 然而这些农学参数会由于冠层结构、种植密度、田间气候等因素的影响会对氮素营养状况的诊断产生偏差^[30], 采用基于氮浓度曲线提出的氮营养指数 (NNI)

可以对冬小麦的氮素是否充足、过剩以及不足的状态给以准确的诊断, 因此快速估测 NNI 不仅可以实时的对冬小麦氮营养状况进行诊断评价, 还可以对冬小麦的氮肥施用量进行合理调控, 这无疑是对精准农业的技术发展提供重要的技术支持。

冬小麦高光谱数据进行预处理是过滤噪声、提高信噪比和提升建模精度的关键手段, 然而不同预处理方法的建模精度有所差异。对比本文使用的 10 种预处理方法结合 PLSR、BP 神经网络和 RF 算法 3 种建模方法得到的预测结果精度可知, SG 卷积平滑预处理得到的预测结果精度皆优于其他 9 种预处理方法, 这是由于 SNV、MSC 的预处理主要功能是消除表面微小颗粒的多元散射造成的影响, 由于冬小麦冠层叶片表面较为平滑, 采集过程中仪器信号不稳定和背景干扰造成的谱线漂移和噪声对建模精度影响更为突出, 而本次建模采用的是高光谱特征参数作为自变量, 其中“三边”参数的导数运算就有消除谱线漂移误差功能, 同时结合 SG 卷积滤波平滑消除光谱上的细小噪声, 提高光谱的整体信噪比还可以最大程度的保留光谱有效信息。对于单一预处理而言, 结合 SG 卷积平滑后再进行其他预处理变换, 也可提升其单一预处理时的模型预测精度, 这与郭斗斗等得出的结论一致^[31]。在本研究中 SG 卷积滤波平滑建立的反演冬小麦氮素营养指数模型的建模精度与预测精度都达到最佳效果。因此在采集冬小麦冠层高光谱的实际工作中, SG 卷积平滑可以有效地解决采集过程中背景噪声的影响。

3 种建模方法中 RF 算法建立的冬小麦氮素营养指数模型效果达到最佳。采用机器学习建模方法明显优于线性回归建模方法的预测精度。这是由于 PLSR 的多元线性回归模型虽然解决了数据变量多和变量之间的重相关性等问题, 但也会损失部分有效信息, 而且在建模过程和实际应用中存在许多非线性问题, 而线性回归模型解决非线性问题时表现能力较弱。机器学习算法的建模方法可以通过内核函数和自学习能力有效解决连续的非线性问题。通过不同预处理的 2 种机器学习算法比较, RF 算法的模型精度都略优于 BP 神经网络的模型精度, 这是由于 BP 神经网络在引入样本训练学习时容易学习过多的样本细节, 这导致训练出来的模型不能最好地映射出数据本身的特定规律, 而 RF 算法引入 2 大随机变量的策略, 可以对噪声和异常值都有较强的容忍度, 同时有研究表明, 当应对大数据量时 RF 算法比 BP 神经网络的训练与预测速度快、效率高^[27], 在诊断大面积多样本的冬小麦氮营养状况时有更为明显的优势, 因此 RF 算法在实际应用中可以更好的为未来精准农业提供实时高效的技术服务。

4 结论

本研究以华北平原乐陵市南夏家村为研究区域, 使用高光谱技术建立冬小麦氮素营养指数的预测模型时, 通过综合对比 10 种光谱预处理方法和 3 种建模方法发现, 选取不同预处理和建模方法对建模的精度有较大影响, 得出如下结论:

1) 本研究中最优的高光谱建模方法为 RF 算法。通过 RF 算法获得的预测集模型比偏最小二乘模型与 BP 神经网络的平均决定系数和均方根误差均低, 相对误差 RE 依次降低了 3.4 与 0.9 个百分点。因此可以看出 10 种预处理方法结合 RF 模型都获得的较高的预测精度, 是一种反演氮营养指数较稳健的建模方法。

2) 对光谱进行预处理可以很好地提升建模精度, 本次试验最佳的光谱预处理方法为 SG 卷积平滑。通过 3 种模型建模比较, SG 卷积平滑比其他 9 种预处理的预测集模型平均 R^2 提高范围为 0.054~0.121, RMSE 平均降低范围为 0.016~0.032, RE 平均降低范围为 1.7~3.1 个百分点。

3) 对单一预处理而言, 结合 SG 卷积平滑后再进行其他预处理变换, 也可提升其单一预处理时的模型预测精度。

[参 考 文 献]

- [1] 赵春江, 薛绪掌, 王秀, 等. 精准农业技术体系的研究进展与展望[J]. 农业工程学报, 2003, 19(4): 7—12.
Zhao Chunjiang, Xue Xuzhang, Wang Xiu, et al. Advance and prospects of precision agriculture technology system[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2003, 19(4): 7—12. (in Chinese with English abstract)
- [2] Roth G W, Fox R H, Marshall H G. Plant tissue tests for predicting nitrogen fertilizer requirements of winter wheat[J]. Agron, 1989(81): 502—507.
- [3] Follett R H, Follett R F, Halvorson A D. Use of a chlorophyll meter to evaluate the nitrogen status of dryland winter wheat[J]. Communications in Soil Science and Plant Analysis, 1992, 23(7/8): 687—697.
- [4] 张树兰, 同延安, 赵护兵, 等. 冬小麦-夏玉米轮作氮肥施量与氮营养诊断[J]. 西北农业学报, 2000(2): 104—107.
- [5] 孙钦平, 李吉进, 邹国元, 等. 应用数字图像技术对有机肥施用后玉米氮营养诊断研究[J]. 光谱学与光谱分析, 2010, 30(9): 2447—2450.
Sun Qiping, Li Jijin, Zou Guoyuan, et al. Nitrogen status diagnosis by using digital photography analysis for organic fertilized maize[J]. Spectroscopy and Spectral Analysis, 2010, 30(9): 2447—2450. (in Chinese with English abstract)
- [6] 毛博慧, 李民赞, 孙红, 等. 冬小麦苗期叶绿素含量检测光谱学参数寻优[J]. 农业工程学报, 2017, 33(增刊 1): 164—169.
Mao Bohui, Li Minzan, Sun Hong, et al. Optimization of spectroscopy parameters and prediction of chlorophyll content at seeding stage of winter wheat[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017, 33(Supp.1): 164—169. (in Chinese with English abstract)
- [7] 程萌, 张俊逸, 李民赞, 等. 用于微小型光谱仪的冬小麦抽穗期叶绿素含量诊断模型[J]. 农业工程学报, 2017, 33(增刊 1): 157—163.
Cheng Meng, Zhang Junyi, Li Minzan, et al. Chlorophyll content diagnosis model of winter wheat at heading stage applied in miniature spectrometer[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017, 33(Supp.1): 157—163. (in Chinese with English abstract)
- [8] 褚小立, 袁洪福, 陆婉珍. 近红外分析中光谱预处理及波长选择方法进展与应用[J]. 化学进展, 2004, 16(4): 528—542.
- [9] 张筱蕾, 刘飞, 聂鹏程, 等. 高光谱成像技术的油菜叶片氮含量及分布快速检测[J]. 光谱学与光谱分析, 2014, 34(9): 2513—2518.
Zhang Xiaolei, Liu Fei, Nei Pengcheng, et al. Rapid detection of nitrogen content and distribution in oilseed rape leaves based on hyperspectral imaging[J]. Spectroscopy and Spectral Analysis, 2014, 34(9): 2513—2518. (in Chinese with English abstract)
- [10] 黄敏. 基于光谱及数据挖掘技术的油菜养分及品质信息的无损检测研究[D]. 杭州: 浙江大学, 2009.
Huang Min. Non-destructive Detection of Rapeseed Growth and Quality Information Based on Spectral And Data Mining Technology[D]. Hangzhou: Zhejiang University, 2009. (in Chinese with English abstract)
- [11] 孙俊, 金夏明, 毛罕平, 等. 基于高光谱图像的生菜叶片氮素含量预测模型研究[J]. 分析化学, 2014(5): 672—677.
Sun Jun, Jin Xiaming, Mao Hanping, et al. A model for predicting nitrogen of lettuce leaves based on hyperspectral imaging[J]. Chinese Journal of Analytical Chemistry, 2014(5): 672—677. (in Chinese with English abstract)
- [12] 李金梦. 基于高光谱成像技术的柑橘叶片含氮量快速检测方法研究[D]. 杭州: 浙江大学, 2014.
Li Jinneng. Rapid Detection of Nitrogen Content in Citrus Leaves Based on Hyperspectral Imaging Technology[D]. Hangzhou: Zhejiang University, 2014. (in Chinese with English abstract)
- [13] 王丽凤, 张长利, 赵越, 等. 高光谱成像技术的玉米叶片氮含量检测模型[J]. 农机化研究, 2017(11): 140—147.
- [14] 乔星星, 冯美臣, 杨武德, 等. SG 平滑处理对冬小麦地上干生物量光谱监测的影响[J]. 山西农业科学, 2016(10): 1450—1454.
- [15] Li F, Misteale B, Hu Y, et al. Reflectance estimation of canopy nitrogen content in winter wheat using optimised hyperspectral spectral indices and partial least squares regression[J]. European Journal of Agronomy, 2014, 52(Part B): 198—209.
- [16] 姚霞, 汤守鹏, 曹卫星, 等. 应用近红外光谱估测小麦叶片氮含量[J]. 植物生态学报, 2011(8): 844—852.

- Yao Xia, Tang Shoupeng, Cao Weixing, et al. Estimating the nitrogen content in wheat leaves by near-infrared reflectance spectroscopy[J]. Chinese Journal of Plant Ecology, 2011(8): 844—852. (in Chinese with English abstract)
- [17] 汤守鹏. 基于近红外光谱技术的小麦叶片氮素营养及籽粒蛋白质含量监测研究[D]. 南京: 南京农业大学, 2009.
- Tang Shoupeng. Monitoring Leaf Nitrogen Status and Grain Protein Content in Wheat using Near Infrared Spectroscopy[D]. Nanjing: Nanjing Agricultural University, 2009. (in Chinese with English abstract)
- [18] Yue S, Meng Q, Zhao R, et al. Critical nitrogen dilution curve for optimizing nitrogen management of winter wheat production in the North China Plain[J]. Agronomy Journal, 2012, 104(2): 523—529.
- [19] Lemaire G, Jeuffroy M, Gastal F. Diagnosis tool for plant and crop N status in vegetative stage[J]. European Journal of Agronomy, 2008, 28(4): 614—624.
- [20] 王秀珍, 黄敬峰, 李云梅, 等. 水稻生物化学参数与高光谱遥感特征参数的相关分析[J]. 农业工程学报, 2003, 19(2): 144—148.
- Wang Xiuzhen, Huang Jingfeng, Li Yunhai, et al. Correlation between chemical contents of leaves and characteristic variables of hyperspectra on rice field[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2003, 19(2): 144—148. (in Chinese with English abstract)
- [21] 王雪, 吉海彦. 基于便携式光谱仪和反射、透射光谱技术的小麦叶片叶绿素定量分析研究[J]. 中国农学通报, 2011, 27(21): 39—43.
- Wang Xue, Ji Haiyan. Quantitative analysis of chlorophyll in wheat leaf based on reflection spectroscopy and transmission spectroscopy using portable spectrometer[J]. Chinese Agricultural Science Bulletin, 2011, 27(21): 39—43. (in Chinese with English abstract)
- [22] 潘冉冉, 骆一凡, 王昌, 等. 高光谱成像的油菜和杂草分类方法[J]. 光谱学与光谱分析, 2017, 37(11): 3567—3572.
- Pan Ranran, Luo Yifan, Wang Chang, et al. Classifications oilseed rape and weeds based on hyperspectral imaging[J]. Spectroscopy and Spectral Analysis, 2017, 37(11): 3567—3572. (in Chinese with English abstract)
- [23] Zhang J, Han W, Huang L, et al. Leaf Chlorophyll content estimation of winter wheat based on visible and near-infrared sensors[J]. Sensors, 2016, 16(4): 1—11.
- [24] 于雷, 洪永胜, 勇周, 等. 高光谱估算土壤有机质含量的波长变量筛选方法[J]. 农业工程学报, 2016, 32(13): 95—102.
- Yu Lei, Hong Yongsheng, Zhou Yong, et al. Wavelength variable selection methods forestimation of soil organic matter content using hyperspectral technique [J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(13): 95—102. (in Chinese with English abstract)
- [25] 王圆圆, 李贵才, 张立军, 等. 利用偏最小二乘回归从冬小麦冠层光谱提取叶片含水量[J]. 光谱学与光谱分析, 2010, 30(4): 1070—1074.
- Wang Yuanyuan, Li Guicai, Zhang Lijun, et al. Retrieval of leaf water content of winter wheat from canopy hyperspectral data using partial least square regression [J]. Spectroscopy and Spectral Analysis, 2010, 30(4): 1070—1074. (in Chinese with English abstract)
- [26] 崔日鲜, 刘亚东, 付金东. 基于机器学习和可见光光谱的冬小麦叶片氮积累量估算[J]. 光谱学与光谱分析, 2016, 36(6): 1837—1842.
- Cui Rixian, Liu Yadong, Fu Jindong. Estimation of winter wheat leaf nitrogen accumulation using machine learning algorithm and visible spectral[J]. Spectroscopy and Spectral Analysis, 2016, 36(6): 1837—1842. (in Chinese with English abstract)
- [27] Grimm R, Behrens T, Märker M, et al. Soil organic carbon concentrations and stocks on Barro Colorado Island-Digital soil mapping using random forests analysis[J]. Geoderma, 2008, 146(1): 102—113.
- [28] 岳继博, 杨贵军, 冯海宽. 基于随机森林算法的冬小麦生物量遥感估算模型对比[J]. 农业工程学报, 2016, 32(18): 175—182.
- Yue Jibo, Yang Guijun, Feng Haikuan. Comparative of remote sensing estimation models of winter wheat biomass based on random forest algorithm[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(18): 175—182. (in Chinese with English abstract)
- [29] 王丽爱, 马昌, 周旭东, 等. 基于随机森林回归算法的小麦叶片 SPAD 值遥感估算[J]. 农业机械学报, 2015, 46(1): 259—265.
- Wang Liai, Ma Chang, Zhou Xudong, et al. Estimation of wheat leaf SPAD value using RF algorithmic model and remote sensing data[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(1): 259—265. (in Chinese with English abstract)
- [30] 陈鹏飞, 孙九林, 王纪华, 等. 基于遥感的作物氮素营养诊断技术_现状与趋势[J]. 中国科学: 信息科学, 2010, 40(增刊): 21—37.
- [31] 郭斗斗, 黄绍敏, 张水清, 等. 多种潮土有机质高光谱预测模型的对比分析[J]. 农业工程学报, 2014(21): 192—200.
- Guo Doudou, Huang Shaomin, Zhang Shuiqing, et al. Comparative analysis of various hyperspectral prediction models of fluvo-aquic soil organic matter[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(21): 192—200. (in Chinese with English abstract)

Nitrogen nutrition diagnosis of winter wheat based on ASD Field Spec3

Liu Changhua¹, Fang Zheng¹, Chen Zhichao^{1*}, Zhou Lan², Yue Xuezhi¹,

Wang Zhe¹, Wang Chunyang¹, Yuxin Miao³

(1. Institute of Surveying-mapping and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China;

2. College of Resources and Environmental Sciences, China Agricultural University, Beijing 100193, China;

3. Department of Soil, Water, and Climate, University of Minnesota, St. Paul, MN, 55108, USA)

Abstract: For crop's prediction model of nitrogen nutrition index (NNI), how to select the pretreatment and modeling method is unclear as well as different pretreatments and their influence degrees on prediction accuracy. So it is of great significance to take more systematic related research for building crop nitrogen nutrition diagnosis rapidly and accurately, which can provide important technical support for precision agriculture management, and realize high yield with high efficiency of nitrogen utilization. Taking Nanxia Village, Laoling City in North China Plain as the research area, based on ASD Field Spec3, the prediction model of winter-wheat nitrogen nutrition index was established with hyperspectral technology in this study. PLSR combined with different pretreatments was applied to establish an prediction model of winter-wheat nitrogen nutrition index, whose average value of model-set model decision coefficient R^2 was 0.683, with the maximum one 0.789, the minimum root mean square error (RMSE) 0.142, and the minimum of relative medium error (RE) 12.3%. The prediction-set model's mean value of decision coefficient R^2 is 0.588, with the maximum one 0.717, the minimum value of root mean square error (RMSE) 0.150, and the minimum of relative medium error (RE) 12.8%. The comparison shows that the pretreatment methods with SG(Savitzky-Golay), SNV(standard normal variate transformation), SG+SNV and SG+BC(baseline correction) are effective when partial least square method is used to build the model, especially SG smoothing is the optimal one as mentioned above with the R^2 of 0.789, the RMSE of 0.142, the RE of 12.3%, and the R^2 of the prediction accuracy of 0.717. Meanwhile, BP neural network method combined with different pretreatments was used to establish an prediction model of nitrogen nutrition index, whose average value of model-set model decision coefficient R^2 was 0.834, with the maximum one 0.861, the minimum RMSE 0.115, and the minimum of RE 9.8%. The prediction-set model's mean value of decision coefficient R^2 was 0.714, with the maximum one 0.780, the minimum value of RMSE 0.133, and the minimum of RE 12.3%. It can be known that the regression model constructed by BP neural network is significantly more accurate than the one constructed by PLSR model. The decision coefficient of all the models pretreated by BP neural network was above 0.8, while the average one was increased to 0.834 from 0.683 under PLSR model. All pretreatment modeling accuracy combined with SG, SG+MSC(multiple scatter correction), SG+SNV and SNV+D(De-trending) reached above 0.85, when the predictive effect reached above 0.7 except BC, MSC and SNV. SG pretreatment R^2 with the best modeling effect reached 0.861, with mean square root error 0.115, relative error 9.8% and predicted effect R^2 0.780 as mentioned above. The NNI estimation model RF algorithm combined with different pretreatments was used to establish an prediction model of NNI, whose average value of model-set model decision coefficient R^2 was 0.945, with the maximum one 0.959, the minimum RMSE 0.061, and the minimum of RE 5.3%. The prediction-set model's mean value of decision coefficient R^2 is 0.742, with the maximum 0.795, the minimum value of RMSE 0.125, and the minimum of RE 11.7%. It can be known that the regression model constructed by RF is significantly improved compared with PLSR model and BP neural network. The decision coefficient of all the models by RF are all above 0.9, and the prediction model accuracy R^2 is above 0.7. After 10 spectral pretreatment methods and 3 modeling ones have been comprehensively compared in this study, it is found that different pretreatment and modeling methods have great impacts on modeling precision. The optimal hyperspectral modeling method is RF(random forest) algorithm. The average value of decision coefficient R^2 for the prediction-set model obtained through RF algorithm was higher than the biased least squares model and BP(back-propagation) neural network respectively, with lower RMSE and RE. Therefore, it can be seen that 10 pretreatment methods combined with RF model have higher prediction accuracy, which is a robust modeling method to invert nitrogen nutrition index. From above, to preprocess the spectrum for winter wheat can improve modeling accuracy. The best spectral pretreatment method in this experiment is SG convolution smoothing. Therefore, by comparing three models, the average R^2 increase range of SG convolution smoothing compared with the other nine pre-processed prediction set models is 0.054~0.121, with average RMSE decrease range 0.016~0.032 and average RE decrease range 1.7~3.1 percentage points. In terms of single preprocessing, other preprocessing transformations can be carried out after combining with SG convolution smoothing, which can also improve the prediction accuracy of its single preprocessing model.

Keywords: spectrum analysis; nitrogen; diagnosis; winter wheat; diagnose model; nitrogen nutrition index (NNI)