

近红外光谱结合 ANN 法快速测定水稻叶片氮含量

周 萍,张广才,王 俊,周崇俊,韩晓日

(沈阳农业大学 土地与环境学院,辽宁 沈阳 110866)

摘要:应用近红外(NIR)光谱和误差反传人工神经网络(BP-ANN)方法建立了水稻叶片氮素含量的定量分析模型。首先对近红外光谱进行 Savitzky-Golay 求导处理,然后通过相关系数法选择波长范围,采用偏最小二乘回归 PLS 降维并输入 BP-ANN 建立校正模型,用验证样品对校正模型进行验证。结果表明:BP-ANN 最佳模型的预测相关系数(RP)为 0.974 7,预测标准误差(SEP)为 4.005,预测相对标准差(RPD)为 3.109。表明 BP-ANN 模型稳健可靠,可较好地用于水稻叶片氮素的快速测定。

关键词:人工神经网络;近红外光谱;水稻叶片;氮素

中图分类号:S511

文献标识码:A

文章编号:1002-2767(2011)04-0022-04

快速而准确地测量作物氮营养含量是植物营养诊断中的一项重要而又关键的工作,如何实时、快速、无损、准确地获取作物营养信息,已成为世界各国精准农业研究中的热点之一^[1]。常规化学分析测试方法存在费时繁琐、破坏植株和实时性差等缺点,田间速测法虽实时性强,但精度不高。随着现代遥感技术的发展,植被的光谱遥感技术已涉及包括氮素在内的作物生化组分的估测^[2-3],

并在理论和实践上为评价作物长势提供了可靠的保证。近红外光谱技术是近红外光谱仪与化学计量学的有机结合,具有快速、准确等优点,在农业领域得到了广泛的应用^[4],具有无需样品前处理、高效、快速和无污染的优点,是其它分析方法不具备的独特优势,如果能将其成功地应用于植物营养快速诊断中将具有重要的理论和实践意义。

水稻是我国最重要的粮食作物之一。全国水稻种植面积约占粮食作物面积的 30%,产量接近粮食总产量的一半。目前,有关基于近红外光谱快速测定作物营养成分的研究主要集中在烟草、小麦、棉花、苹果、油菜等作物上^[5-8],他们利用近红外光谱构建模型来快速监测作物的生化组分和籽粒品质,取得了可喜进展。但是,利用近红外光

收稿日期:2011-01-09

基金项目:教育部留学回国人员科研启动基金资助项目

第一作者简介:周萍(1985-),女,吉林省松原市人,硕士,从事植物营养与优化配方施肥研究。E-mail: xingfudexingfu@163.com。

通讯作者:张广才(1972-),男,博士,副教授,从事植物营养诊断与优化配方施肥研究。E-mail: guangcaizhang@163.com。

Effects of Different Sowing Date on Starch Content and Yield of Potato

ZHANG Rong-hua, XU Qing-fen, YANG Yan-hua, XU Ning

(Industrial Crops Institute of Heilongjiang Academy of Land Reclamation Sciences, Harbin, Heilongjiang 150038)

Abstract: Taking the mid-maturing potato variety Kexin18 as experimental material, the experiment of different sowing date was conducted in Western Farm in Heilongjiang to discuss the effect on growth, starch content and yield of potato. The results showed that: the potatoes should not be planted too early, it may be vulnerable to low temperature to make the seedlings become weak, thus affecting the yield and quality; Potato suitable sowing date was in early and mid May in Heilongjiang province, which would obtain high yields, yield difference reached a significant level with other sowing dates. The best sowing date was May 6th, the maximum yield reached 49 555 kg·hm⁻², starch content reached 15.51%, starch content difference reached a significant level with other sowing dates.

Key words: sowing date; potato; starch content; yield

谱技术预测水稻叶片氮素营养状况的研究报道较少,有必要进一步研究估算水稻叶片氮素营养状况的技术和方法,以确立解释性和精确性兼备的水稻氮素营养估测模型。

通过筛选最佳光谱预处理方法,建立 ANN 模型,以期实现水稻叶氮素含量快速、无损、实时的测定。

1 材料与方法

1.1 测定仪器及软件

采用 FPI(聚光科技股份有限公司)的 Sup-NIR-2700 近红外分析仪(NIR analyzer),通过漫反射方法测定水稻叶片粉末,获得样品的光谱数据,应用仪器自带的化学计量学软件 CM-2000 进行数据分析和建模。

1.2 样品来源及参考值的测定

采用盆栽试验,在沈阳农业大学土地与环境学院的试验网室进行,供试水稻品种为沈农号,试验共设 19 个处理,5 次重复,供试土壤为棕壤。其肥力状况为:土壤有机质 1.789%,全氮 0.039%,碱解氮 $78.5 \text{ mg} \cdot \text{kg}^{-1}$,速效磷 $10.35 \text{ mg} \cdot \text{kg}^{-1}$,速效钾

$80.7 \text{ mg} \cdot \text{kg}^{-1}$, pH 6.8。设定不同的施肥量使样品中待测养分含量尽可能涵盖实际生产过程中水稻叶片中待测养分含量的浓度范围,每盆装烘干土 10 kg。氮磷钾肥施用比例为 $\text{N}:\text{P}_2\text{O}_5:\text{K}_2\text{O}=1.0:0.5:0.5$ 。肥料种类分别为尿素(N 46%)、磷酸二铵(N 18%, P_2O_5 46%)、硫酸钾(K_2O 50%)。在水稻分蘖期采集样品 87 个,将样品放入 70℃ 鼓风干燥箱内烘干 48 h,用磨样机粉碎干燥的样品,并全部通过 0.5(35 目)的尼龙筛,充分混合均匀,放入干燥器中保存,用来测定化学参考值和光谱数据。水稻叶全氮的测定采用德国 VARIO EL III 元素分析仪(见表 1)。

表 1 水稻叶片氮含量的数据统计

样本数	最大值 / $\text{g} \cdot \text{kg}^{-1}$	最小值 / $\text{g} \cdot \text{kg}^{-1}$	平均值 / $\text{g} \cdot \text{kg}^{-1}$	标准偏差 SD	变异系数 CV
87	51.495	12.981	34.234	12.45	3.638

1.3 近红外光谱的采集

光谱采集平均次数设为 10,样品吸光度光谱见图 1,经 Savitzky-Golay 一阶导数和均值中心化预处理后的光谱见图 2。

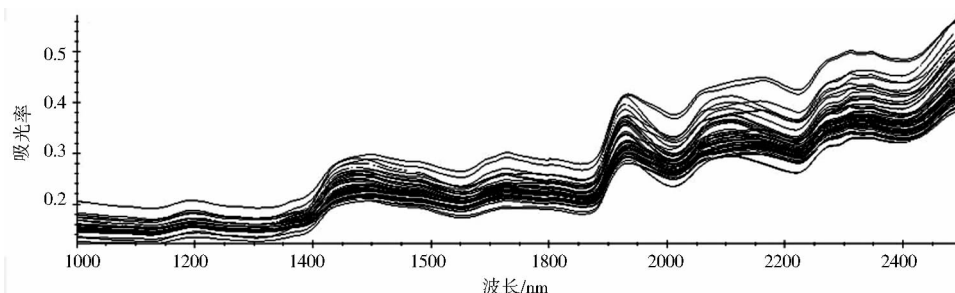


图 1 水稻叶片近红外吸收光谱

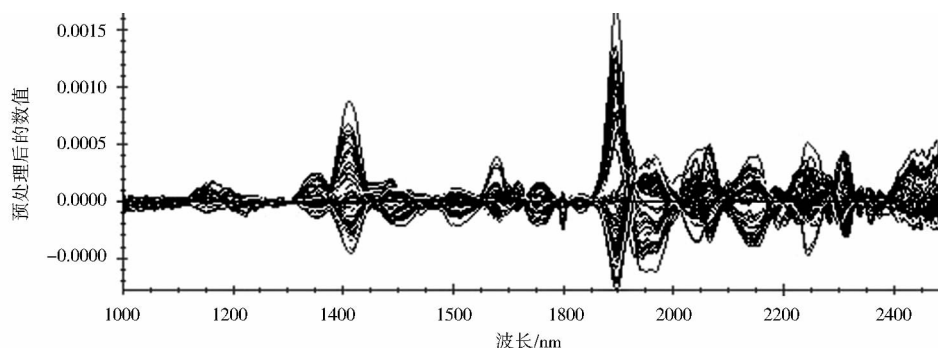


图 2 水稻叶片近红外 S-G 一阶导数光谱

1.4 BP-ANN 原理

BP-ANN 是应用最多的一种神经网络方法^[9],它具有良好的非线性映射能力。BP 神经网络由 3 部分组成:输入层、隐含层和输出层。每层

都有许多节点,每个节点的输出被送到下一层的所有节点,将这些处理单元组成层互相连接起来,并对联接进行加权,从而形成神经网络的拓扑结构。为减小网络规模以缩短训练时间,去除噪音

并保留有效信息,在应用 ANN 的过程中,常使用主成分分析或偏最小二乘法对光谱数据降维,获得的主成分得分作为网络的输入变量。

1.5 数学模型性能评价

使用校正集样品建立 BP-ANN 校正模型,并采取内部交叉验证法进行检验。检验中所采用的评估参数包括:校正相关系数(RC)、校正标准误差(SEC)、预测相关系数(RP)、预测标准误差(SEP)、预测相对标准差(RPD)。其中预测相对标准差(RPD)是指预测样品集中化学参考值的标准偏差(SD)与模型预测标准误差(SEP)的比值。RPD 的数值越大,模型的稳定性越好。理想的 RPD 值至少应该大于等于 3^[10]。

2 结果与分析

2.1 近红外光谱预处理

由于原始光谱数据含有来自高频随机噪音、样本不均匀、基线漂移、光散射等不利因素的影响,需对光谱进行预处理。所测的样品近红外光谱数据有比较应用均值中心化、Savitzky-Golay 平滑(窗口宽度 13 数据点,拟合多项式次数 2)、Savitzky-Golay 一阶导数求导(窗口宽度 21 数据点,拟合多项式次数 3,求导次数 1),以寻找最佳光谱预处理方法。

2.2 建模波段选择

不同波段范围显著影响模型的精度和预测能力,为了建立叶片氮含量最佳预测模型,采用相关系数法^[11]选择波段范围 1 000~1 300 nm,1 550~1 700 nm,2 050~2 400 nm 作为建模波长区域。

2.3 神经网络参数选择

在 BP 神经网络的建立过程中,采用偏最小

二乘法(PLS)方式对数据进行网络计算之前的分解,PLS 降维的最佳主成分数为 10^[12],经过反复训练,确定 ANN 隐含层的节点数目为 10,输入层到隐含层的初始权重、隐含层到输出层的初始权重为-0.5~0.5 的随机数,隐含层、输出层的传递函数选择 S 型非线性函数,初始学习速率为 0.1,动量项为 0.9。

2.4 BP-ANN 模型

从 87 个水稻叶片样品中,剔除 4 个重复数据,进一步采用 K-S 分类的方法取 59 个作为校正样品,剩余 24 个作为预测样品。首先对光谱数据进行不同的预处理方法,然后建立相应的人工神经网络(BP-ANN)模型,并使用内部交叉验证法对模型进行可靠性检验,因为在样品数量比较少的情况下,内部交叉验证法^[13-14]可以有效利用有限的样品,但是计算速度较之外部检验法要缓慢一些。通过模型的参数比较,确定最佳预处理方法和模型。对光谱数据采用了 S-G 平滑(13,2)+S-G 一阶导数(21,3),均值中心化+S-G 一阶导数(21,3),均值中心化+S-G 平滑(13,2),均值中心化+S-G 平滑(13,2)+S-G 一阶导数(21,3)共 4 种预处理方法。

表 2 是水稻叶风干粉末样品的氮营养元素含量与相应的近红外光谱数据之间的人工神经网络模型的各种参数结果,由表 2 可以看出最好的预处理方法是均值中心化+S-G 一阶导数(21,3),其人工神经网络模型结果是:校正相关系数(RC)与预测相关系数(RP)分别为 0.950 9、0.974 7,均大于 0.95,校正标准误差(SEC)与预测标准误差(SEP)分别为 4.393、4.005,说明建立的模型

表 2 不同预处理方法的 ANN 模型结果

预处理方法	校正相关 系数 RC	校正标准 误差 SEC	统计指标		
			预测相关 系数 RP	预测标准 误差 SEP	预测相对 标准差 RPD
S-G 平滑(13,2)+S-G 一阶导数(21,3)	0.8774	6.390	0.8338	8.269	1.506
均值中心化+S-G 一阶导数(21,3)	0.9509	4.393	0.9747	4.005	3.109
均值中心化+S-G 平滑(13,2)	0.8350	7.561	0.9028	5.790	2.150
均值中心化+S-G 平滑(13,2)+S-G 一阶导数(21,3)	0.9258	5.356	0.8219	8.701	1.431

与化学值有很好的相关性。

图 3 是使用 BP-ANN 方法获得的水稻叶片校正集样品和验证集样品的氮素化学值(参考值)与相应预测值的最佳相关曲线,横轴为化学值,纵轴为预测值。

从图 3 可见,校正样本和验证样本的各点均

分布在曲线的附近,说明采用近红外光谱结合 BP-ANN 方法获得的水稻叶氮素含量的预测值与化学值之间均具有良好的相关性。结合表 2 可知,预测相对标准差(RPD)大于 3,说明模型效果良好,预测精度高,所建立的 BP-ANN 校正模型可应用于实际分析。

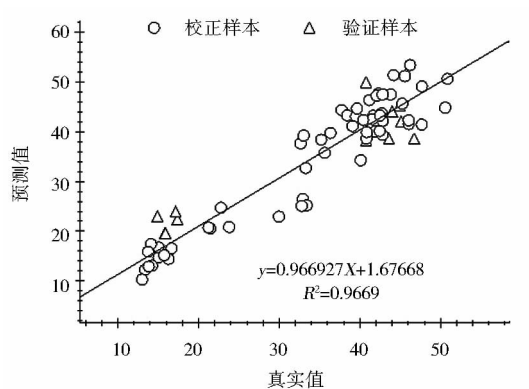


图3 ANN 的最佳预测模型

3 结论与讨论

定标模型的优劣取决于所建模型相关系数的高低及预测误差的大小,较高的相关系数和较小的预测误差对于模型的可靠性是十分重要的。研究的最佳模型校正相关系数(RC)与预测相关系数(RP)分别为 0.950 9、0.974 7,校正标准误差(SEC)与预测标准误差(SEP)分别为 4.393、4.005,预测相对标准差(RPD)为 3.109。表明应用 BP-ANN 建模方法建立的水稻叶氮素含量模型准确可靠,可以用于实际生产中快速准确测定水稻叶片中全氮含量。彭玉魁^[15]认为样品数量的多少直接影响分析的准确性。该研究的样品数目只有 87 个,因此还需要不断扩大定标样品的数量和代表性以完善模型。

参考文献:

- [1] 薛利红,曹卫星,罗卫红,等.基于冠层反射光谱的水稻群体叶片氮素状况监测[J].中国农业科学,2003,36(7):807-812.
- [2] Tarpley L, Reddy K R, Gretchen FSC. Reflectance indices with precision and accuracy in predicting cotton leaf nitrogen concentration[J]. Crop Science, 2000, 40: 1814-1819.

- [3] 朱艳,李映雪,周冬琴,等.稻麦叶片氮积累量与冠层反射光谱的定量关系[J].植物生态学报,2006,30(6):983-990.
- [4] 严衍禄.近红外光谱分析基础与应用[M].北京:中国轻工业出版社,2005.
- [5] Petisco C G, Cfiado B, Mediavilla S, et al. Near-infrared reflectance spectroscopy as a fast and non-destructive tool to predict foliar organic constituents of several woody species[J]. Analytical and Bioanalytical Chemistry, 2006, 386: 1823-1833.
- [6] Read J J, Tarpley L, Mckinon J M. Narrow-waveband reflectance ratios for remote estimation of nitrogen status in cotton[J]. Journal of Environmental Quality, 2002, 31: 1442-1452.
- [7] 王渊,黄敬峰,王福民,等.油菜叶片和冠层水平氮素含量的高光谱反射率估算模型[J].光谱与光谱分析,2008,28(2): 273-277.
- [8] 张俊华,张佳宝.不同生育期小麦光谱特征对叶绿素和氮素的响应研究[J].土壤通报,2008,39(3):586-592.
- [9] 徐广通,袁洪福,陆婉珍.现代近红外光谱技术及应用进展[J].光谱学与光谱分析,2000,20(2):134-142.
- [10] Williams P C, Sobering D C. Comparison of commercial near infrared transmittance and reflectance instruments for analysis of whole grains and seeds[J]. Journal of Near Infrared Spectroscopy, 1993, 1: 25-32.
- [11] 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.
- [12] 许禄.化学计量学一些重要方法的原理及应用[M].北京:科学出版社,2004.
- [13] Stone M. Cross-validatory choice and assessment of statistical predictions[J]. Journal of the Royal Statistical Society, Series B (Methodological), 1974, 36(2): 111-147.
- [14] 曹干,谭中文,梁计南,等.蔗汁品质成分的傅里叶变换近红外分析数学模型[J].中国农业科学,2003,40(3): 254-258.
- [15] 彭玉魁.近红外光谱分析技术及其在农业中的应用[J].陕西农业科学,2000(3):25-27.

Rapid Analysis of Rice Leaf Nitrogen Using Near Infrared Spectroscopy and Artificial Neural Network

ZHOU Ping, ZHANG Guang-cai, WANG Jiao, ZHOU Chong-jun, HAN Xiao-ri

(Land and Environment College of Shenyang Agricultural University, Shenyang, Liaoning 110866)

Abstract: The models of quantitative analysis of nitrogen in the rice leaf were established by using near infrared spectroscopy(NIS)coupled with the back propagation-artificial neural network method(BP-ANN). Firstly, the data of original spectra were pretreated by Savitzky-Golay derivative. Secondly, the wavelength range of model was optimized by using correlation coefficient method. Finally, PLS dimension-reducing was input into BP-ANN. The calibration models were established by calibration set and validated by prediction set. The results showed that the related coefficient(RP)of the best prediction for nitrogen was 0.974 7, the standard errors of prediction(SEP)for nitrogen was 4.005, and ratio of performance deviation(RPD)was 3.109. Therefore, the method could be applied to fast and accurate determination of nitrogen in the rice leaf.

Key words: artificial neural network; near infrared spectroscopy; rice leaf; nitrogen