基于特征光谱提取的有机基质全氮含量快速检测方法*

朱咏莉　李萍萍　孙德民　毛罕平
（江苏大学现代农业装备与技术省部共建教育部重点实验室，镇江 212013）

【摘要】 以醋糟有机基质为研究对象，采用便携式可见/近红外光谱仪获得基质样品的光谱信息，经过归一化和一阶微分预处理后，采用逐步回归法提取对有机基质全氮反应敏感的特征光谱，建立基于特征波长组合的线性回归模型。其中，以 1 699、746、1 864 和 2 154 nm 为特征波长的四元回归模型为最佳，其预测相关系数和预测均方根误差分别为 0.933 4 和 1.04。结果表明，利用可见/近红外光谱技术，通过特征光谱的提取并建立相应的回归模型，可以实现对有机基质全氮含量的快速准确检测。

关键词：有机基质　氮素　检测　逐步回归法　特征波长　可见/近红外光谱
中图分类号：S129；O657.33　　**文献标识码**：A　　**文章编号**：1000-1298(2011)05-0175-03

Total Nitrogen Content Detection in Organic Substrate Using Visible-near-infrared Spectroscopy

Zhu Yongli　Li Pingping　Sun Demin　Mao Hanping
(*Key Laboratory of Modern Agricultural Equipment and Technology*, *Ministry of Education & Jiangsu Province*, *Jiangsu University*, *Zhenjiang* 212013, *China*)

Abstract

For the purpose of developing a method for rapid detection of total nitrogen content in organic vinegar residue substrates, visible-near-infrared spectroscopy was used to collect the spectral data of the representative samples. And the spectra were pretreated by the method of first derivative after normalization. In order to choose the characteristic wavelengths for the total nitrogen content in the substrates, stepwise regression (SWR) method was applied twice. The results indicated that the characteristic wavelength was at 1 699 nm, 746 nm, 1 864 nm and 2 154 nm, respectively. Based on these wavelengths, a linear regression model was performed. For comparing the predictive precision of the model, the correlationship between prediction total nitrogen content and that obtained by chemical method was analyzed. It was confirmed that the multivariate equation was the best with correlation coefficient 0.933 4 and root mean squared error of prediction (RMSEP) 1.04. Therefore, it was demonstrated that visible-near-infrared spectroscopy with characteristic wavelengths selecting could be an available method for rapid detection of total nitrogen content in organic substrate.

Key words　Organic substrate, Total nitrogen, Detection, Stepwise regression method, Characteristic wavelength, Visible-near-infrared spectroscopy

引言

随着设施园艺的迅速发展，固体栽培基质（尤其是有机基质）的开发和应用得到了较大的促进，成为设施作物栽培的重要介质类型。在基质肥力中，氮素是最活跃的因素，也是生产中最重要的限制因子之一，需要对其含量的动态变化进行及时、准确的监测。目前现有的全氮含量检测方法为化学法，

收稿日期：2010-08-05　修回日期：2010-09-02
* 国家星火计划资助项目(2010GA690001)、江苏高校优势学科建设工程资助项目和江苏省教育厅项目(08KJD210006)
作者简介：朱咏莉，副研究员，博士，主要从事农业环境信息检测及农业资源利用研究，E-mail：zhuyongli76@126.com
通讯作者：李萍萍，教授，博士生导师，主要从事农业资源利用研究，E-mail：lipingping@ujs.edu.cn

过程复杂,很难满足有机基质栽培实时调控的需求。针对有机基质,有必要研究一种适宜的氮素含量快速检测方法。

可见/近红外光谱技术(NIRS)具有快速、无损、易操作及稳定性好等特点,已被广泛用于有机物料中氮含量指标的检测[1~6]。目前对有机物料氮素含量的研究大部分是在实验室条件下采用高档光谱仪完成的,需要采集大量的光谱数据,对环境条件的要求也较为苛刻,而所建立的预测模型均为全光谱模型,包含信息过于庞杂,因此难以应用于田间的实时检测。而小型滤光片式光谱仪以其结构简单、稳定性高等特点[7],日益受到青睐,具有广阔的应用前景。但如何选取滤光片波长是该类仪器研制中面临的首要问题。

本文以醋糟有机栽培基质为对象,通过便携式可见/近红外光谱仪获取其光谱信息,采用逐步回归法提取对基质全氮敏感的特征光谱,建立基于多个特征波长组合的数学模型,以反演基质全氮含量,其目的是寻找一种快速准确地检测有机基质全氮含量的方法,并为小型滤光片式光谱仪进一步的研制提供依据。

1 材料与方法

1.1 样品采集

试验材料采自新鲜醋糟发酵堆制过程的不同阶段,共计 106 个样品。样品采集后混匀,一部分立即装入保鲜袋中,保存于 4℃冰箱,用于光谱采集以及硝态氮和铵态氮含量的测定;另一部分晾干、粉碎过 1 mm筛后用于光谱采集和全氮含量的测定。

1.2 试验方法

1.2.1 仪器设备

采用美国 ASD(analytical spectral device)公司生产的 Field Spec 3 型便携式光谱分析仪。该仪器光谱测量范围 350 ~2 500 nm;在 350 ~1 000 nm 光谱区采样间隔为 1.4 nm,分辨率为 3 nm;在 1 000 ~2 500 nm 光谱区采样间隔为 2 nm,分辨率为 10 nm;采样频率 10 Hz。

1.2.2 光谱反射率测量

将光谱探头及光源固定于三角架上,传感器探头位于工作台上方 7 cm 左右处,垂直于被测物,试验设定视场为 25°。测量前进行系统配置优化标准白板标定。测定时,将样品装于 1.5 cm 深的圆形器皿中。为消除光谱采样误差,取每 10 次扫描数据的平均值作为该样品的原始光谱。在整个光谱采集过程中,保持相同的测定条件(如外界光照等)以及较为一致的装样厚度和紧密性等。

1.2.3 样本全氮含量的测定

样品采用开氏消化法进行消化,流动注射分析仪(英国 SEAL Analytical 公司)测定其全氮含量,并以干基进行表示。

1.3 数据处理方法

将所采集的样品光谱反射率(R)进行反射吸光度(A)的转换($A = \lg(1/R)$),再对数据进行预处理。光谱数据的预处理采用江苏大学自主开发的 NIRSA 软件系统进行。特征光谱波长的提取、光谱数据与建模的分析则结合 Sigma Plot 13.0 进行。

2 结果与分析

2.1 有机基质样品的全氮含量测定及其光谱特征

试验总样本数为 106 个,其中 66 个作为校正集建立校正模型,其余 40 个作为预测集来验证模型稳健性,其全氮含量的统计结果如表 1 所示。总样品集全氮含量在 23.0 ~35.6 g/kg 之间变化,其平均值、标准差和变异系数分别为 28.5 g/kg、2.9 g/kg 和 10.1%。

表 1 样品氮素含量的统计结果

Tab.1 Statistic parameters of the total nitrogen content of samples

项目	样本数	最小值/g·kg^{-1}	最大值/g·kg^{-1}	平均值/g·kg^{-1}	标准差/g·kg^{-1}	变异系数/%
总样本	106	23.0	35.6	28.5	2.9	10.1
校正集	66	23.0	35.6	28.4	3.0	10.5
预测集	40	23.7	34.9	28.6	2.8	9.6

醋糟主要为谷糠和稻壳,其全氮组分主要包括蛋白质和其他少量的有机含氮化合物以及铵态氮和硝态氮等无机氮。在近红外区域,与氮素有关的主要为 N—H 官能团[8~9]。从醋糟基质粉碎干样的原始光谱图(图 1)可以看出,在 350 ~1 000 nm 吸光度较高,且随波长的增加基本呈平滑的下降趋势。在 1 160 ~1 265、1 400 ~1 540、1 690 ~1 800、1 890 ~1 990、2 040 ~2 150 和 2 250 ~2 350 nm 区间则表现为一定强度的吸收峰。文献[10]表明,1 400 ~1 540 nm和 1 890 ~1 990 nm 主要为样品中少量水分子中 O—H 键与胺类和酰胺类等 N—H 键的一级倍频和组合频吸收,其他 4 个区间则主要为 N—H 键

的二级倍频和组合频吸收。

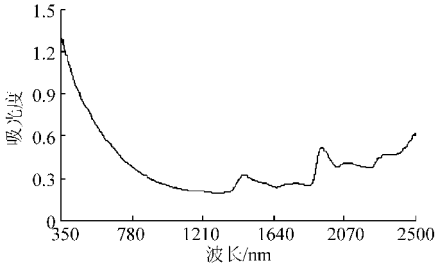

图 1　醋糟基质干样的典型光谱

Fig. 1　Typical reflectance spectra of dried samples

2.2　有机基质全氮特征光谱的提取

全光谱包含的信息过于庞杂，难以在实际生产中得到应用。为了得到简便的校正模型，采用有进有出的逐步回归法（SWR）进行特征波长的提取，初次提取设置引入自变量数小于等于 15，检验值大于等于 3 时引入，小于 3 时剔除。光谱数据的预处理方法为先归一化、后一阶微分。经初步提取，得到的特征波长数为 15，其中 1 500 ~ 2 400 nm 间的波长有 11 个，占提取波长总数的 70% 以上。

为了进一步降低模型的复杂度，删除不重要的解释变量，采用逐步回归法（设置 95% ≤P≤99%，同时进行共线性诊断，P 表示置信概率），对已提取的 15 个波长进行二次提取，按引入自变量的不同，可得到 15 个校正模型。其中，以引入单变量的模型的相关系数最低（0. 922），校正均方根误差最高（约 1. 1），引入 15 个自变量的模型相关系数最高（0. 995），校正均方根误差仅约为 0. 3。

从图 2 可以看出，随引入自变量数的增加，校正均方根误差呈接近线性的下降趋势，而其预测均方根误差却变化较为缓慢，自变量数为 1 ~ 4 时，其预测均方根误差从 1. 3 降至 1. 0 左右，随后自变量数目达到 4 以上时，预测均方根误差没有明显的变化，基本维持在 1. 0 左右。这说明，增加自变量数尽管可有效降低模型的校正均方根误差，但对预测均方根误差的降低效果却不明显。因此，综合模型对校正集和预测集样本的预测效果，认为包含有 4 个自变量的四元模型为理论上的最佳模型，其特征波长分别为 1 699、746、1 864 和 2 154 nm。

2.3　有机基质全氮的光谱模型及其预测效果

基于以上特征光谱波长的醋醣基质全氮含量预测模型可以表示为

$$T_N = 41.098 - 365\,163.3x_{1\,699\ nm} + 85\,137.1x_{746\ nm} - 66\,821.9x_{1\,864\ nm} + 111\,060.2x_{2\,154\ nm}$$

式中　T_N——醋糟基质全氮含量，g/kg

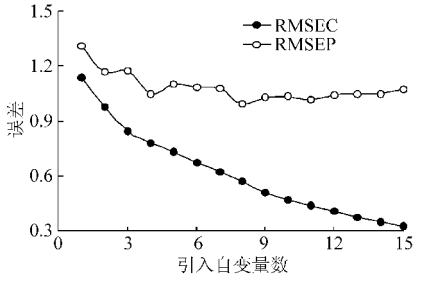

图 2　模型引入自变量数与校正均方根误差（RMSEC）和预测均方根误差（RMSEP）间的关系

Fig. 2　Root mean squared errors of calibration and prediction *vs* independent variable numbers

x——不同波长对应的反射吸光度经归一化和一阶微分处理后的数值，下角标为波长

该模型的校正相关系数和校正均方根误差分别为 0. 965 9 和 0. 75，其对预测集样品的预测效果如图 3 所示，预测相关系数为 0. 933 4，预测均方根误差为 1. 04。并且，对于试验所取预测集样品，全氮含量在 23. 7 ~ 34. 9 g/kg 之间，预测值的平均相对误差仅为 2. 2%。由此可见，采用该四元模型可以对醋醣基质全氮含量进行较为准确的预测。

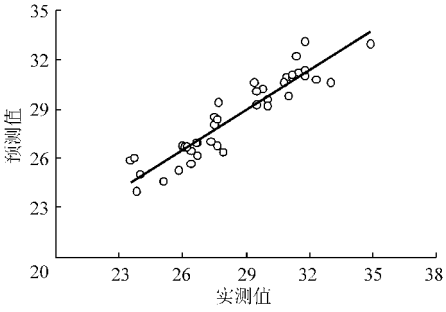

图 3　全氮含量预测值与实测值的关系

Fig. 3　Measured values *vs* spectra prediction values for total nitrogen content

但与全光谱波段的模型（其预测相关系数和预测均方根误差分别为 0. 971 7 和 0. 71）相比，尽管该四元模型更为简单、实用，预测精度却仍偏低。如何通过一些技术手段进一步提高该模型的预测精度，建立更为精确的实用模型则是进一步需要研究的内容。

3　结束语

通过采用逐步回归分析法提取出对醋糟有机基质全氮敏感的特征光谱，并通过特征波长的组合建立了可以预测有机基质全氮含量的回归模型。其中，四元模型具有较高的预测精度，预测效果较好。因此，利用可见/近红外光谱技术，通过特征光谱的提取和建立基于该特征波长的回归模型，可以实现对有机基质全氮含量的快速准确检测。

（下转第 192 页）

improved genetic algorithm neural network [J]. Transactions of the Chinese Society of Agricultural Engineering, 2009, 25(10):340 ~ 344. (in Chinese)

12 杨飞,祝诗平,邱青苗. 基于计算机视觉的花椒外观品质检测及其 MATLAB 实现[J]. 农业工程学报,2008,24 (1):198 ~ 202.
Yang Fei, Zhu Shiping, Qiu Qingmiao. Prickly ash appearance quality detection based on computer vision and its implementation in MATLAB [J]. Transactions of the Chinese Society of Agricultural Engineering, 2008, 24 (1):198 ~ 202. (in Chinese)

13 杨劲峰,陈清,韩晓日,等. 数字图像处理技术在蔬菜叶面积测量中的应用[J]. 农业工程学报,2002,18(4):155 ~ 158.
Yang Jinfeng, Chen Qing, Han Xiaori, et al. Measurement of vegetable leaf area using digital image processing techniques [J]. Transactions of the Chinese Society of Agricultural Engineering, 2002, 18(4):155 ~ 158. (in Chinese)

14 周平,刘俭英,王巧华,等. 鸡蛋图像检测方法与质量预测模型[J]. 农业机械学报,2007,38(11):80 ~ 82.
Zhou Ping, Liu Jianying, Wang Qiaohua, et al. Egg image detection method and weight prediction model [J]. Transactions of the Chinese Society for Agricultural Machinery, 2007, 38(11):80 ~ 82. (in Chinese)

15 Holland J H. Adaptation in natural and artificial systems [M]. Ann Arbor, MI: The M/T Press, 1992.

16 张秉森,王莹,李莉. 遗传算法改进 BP 网络对织物配色的优化研究[J]. 计算机工程与设计,2008,29 (19):5 033 ~ 5 035.
Zhang Bingsen, Wang Ying, Li Li. Optimization research of BP neural network with genetic algorithms used in textile color matching [J]. Computer Engineering and Design, 2008, 29(19):5 033 ~ 5 035. (in Chinese)

(上接第 177 页)

参 考 文 献

1 Albrecht R, Joffre R, Gros R, et al. Efficiency of near-infrared reflectance spectroscopy to assess and predict the stage of transformation of organic matter in the composting process [J]. Bioresource Technology, 2008, 99(2): 448 ~ 455.

2 黄光群, 王晓燕, 韩鲁佳. 基于支持向量机的有机肥总养分含量 NIRS 分析[J]. 农业机械学报, 2010, 41(2) : 93 ~ 98.
Huang Guangqun, Wang Xiaoyan, Han Lujia. Near infrared reflectance spectroscopy for total nutrient and analysis in organic fertilizer using linear support vector machine [J]. Transactions of the Chinese Society for Agricultural Machinery, 2010, 41(2): 93 ~ 98. (in Chinese)

3 Malley D, Mcclure C, Martin P. Compositional analysis of cattle manure during composting using a field portable near infrared spectrometer [J]. Communications in Soil Science & Plant Analysis, 2005, 36(4 ~ 6): 455 ~ 476.

4 樊霞, 韩鲁佳, 皇才进, 等. 基于近红外光谱技术的牛粪成分含量测定方法[J]. 农业机械学报, 2006, 37(3): 76 ~ 79.
Fan Xia, Han Lujia, Huang Caijin, et al. Determination of nutrient contents in beef manure with near infrared reflectance spectroscopy [J]. Transactions of the Chinese Society for Agricultural Machinery, 2006, 37(3): 76 ~ 79. (in Chinese)

5 Smidt E, Meissl K. The applicability of Fourier transform infrared (FT-IR) spectroscopy in waste management [J]. Waste Management, 2007, 27(2): 268 ~ 276.

6 Grube M, Lin J G, Lee P H, et al. Evaluation of sewage sludge-based compost by FT-IR spectroscopy [J]. Geoderma, 2006, 130(3 ~ 4): 324 ~ 333.

7 高洪智, 卢启鹏, 丁海泉, 等. 基于连续投影算法的土壤总氮近红外特征波长的选取[J]. 光谱学与光谱分析, 2009, 29(29): 2 951 ~ 2 954.
Gao Hongzhi, Lu Qipeng, Ding Haiquan, et al. Choice of characteristic near-infrared wavelengths for soil total nitrogen based on successive projection algorithm [J]. Spectroscopy and Spectral Analysis, 2009, 29(29): 2 951 ~ 2 954. (in Chinese)

8 陆婉珍. 现代近红外光谱分析技术[M]. 北京: 中国石化出版社, 2007.

9 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.

10 杰克·沃克曼, 洛伊斯·文依. 近红外光谱解析实用指南[M]. 褚小立, 许育鹏, 田高友, 译. 北京: 化学工业出版社, 2009.