

基于可见-近红外光谱技术预测茶鲜叶全氮含量

胡永光¹, 李萍萍^{1*}, 母建华¹, 毛罕平¹, 吴才聪^{1,3}, 陈 斌²

1. 现代农业装备与技术教育部重点实验室(江苏大学), 江苏大学江苏省现代农业装备与技术重点实验室, 江苏 镇江 212013
2. 江苏大学食品与生物工程学院, 江苏 镇江 212013
3. 北京大学遥感与地理信息系统研究所, 北京 100871

摘 要 为快速无损监测茶树氮素营养及其生长状况, 基于可见-近红外光谱技术建立了茶鲜叶全氮含量的预测模型。以茶鲜叶为对象, 田间试验使用便携式光谱仪采集叶片漫反射光谱信息, 通过不同预处理和统计分析, 建立茶鲜叶全氮含量预测的光谱模型。试验共采集 111 个样品, 其中 86 个样品作校正集, 25 个样品作预测集。通过一阶导数与滑动平均滤波相结合的预处理方法, 用 7 个主成分建立的偏最小二乘模型最好, 其校正集均方根误差(RMSEC)为 0.097 3, 预测集的相关系数为 0.888 1, 预测均方根误差(RMSEP)为 0.130 4, 预测的平均相对误差为 4.339 %。研究结果表明, 利用可见-近红外光谱技术可以很好地预测茶鲜叶全氮含量, 对于快速实时监测茶树长势和施肥管理具有重要指导意义。

关键词 可见-近红外光谱; 偏最小二乘回归; 茶鲜叶; 全氮含量; 预测

中图分类号: S571.1; O433.4 **文献标识码**: A **DOI**: 10.3964/j.issn.1000-0593(2008)12-2821-05

引 言

氮素在茶叶内平均含量约占干重的 1.5 %, 是合成蛋白质和叶绿素的重要组成部分, 又参与酶的合成, 它在多方面直接或间接影响茶树的代谢活动和生长发育, 特别是对茶叶有效成分氨基酸含量的多寡具有重要意义^[1]。因此, 监测氮素营养状况是茶树栽培管理过程中的主要内容之一。

为及时掌握作物的生长情况, 人们常根据经验进行外观诊断, 但其可靠性不高; 或采用基于土壤和作物的实验室分析, 而这些分析普遍要求破坏土壤和植被样本, 且测量费时费力, 过程复杂。传统的氮素营养诊断、无损测试方法主要有肥料窗口法和叶色卡片法, 这些方法均属于定性或半定量的方法^[2]。采用快速、简便、非破坏性的现代分析方法测定氮含量是极其需要的。日本研制的叶绿素仪 SPAD-502 能测定作物叶片的叶绿素相对含量, 而叶绿素含量与作物叶片氮含量又有相关关系^[2], 所以可通过叶绿素仪读数(SPAD 值)间接了解作物叶片的氮含量。目前, 国外利用 SPAD-502 诊断水稻氮素营养的研究比较成熟^[3]。国内有不少学者研究作物如油菜、小麦、水稻等的 SPAD 值、叶片叶绿素含量和叶

片氮含量之间的相关关系^[4,5]。虽然叶绿素仪读数和叶片氮含量之间具有线性关系, 但这种线性关系受作物品种、发育阶段、测定的叶位及叶片上测点位置的影响而存在较大差异。因此, 根据叶绿素仪读数来指示作物氮含量的可靠性、精度不高。另一种快速无损的氮含量检测方法—近红外光谱分析法目前也得到了广泛应用。鲍一丹等分析了土壤含水量和土壤粒径大小对氮含量光谱预测模型的影响^[6]; 孙建英等选取几个特征波长点处的光谱信息进行了全氮含量的多元回归分析^[7]。但土壤成分复杂, 氮含量的测定受多种因素干扰, 且土壤中氮含量不能直接反映作物氮的吸收情况, 而作物叶片氮含量则能更直观反映作物氮素营养状况。张金恒等指出水稻叶片氮含量的敏感波段为绿光(525 ~ 605 nm)、黄光(605 ~ 655 nm)和短波近红外光(750 ~ 1 100 nm)^[8]; 薛利红、李映雪等主要研究了不同施氮水平下小麦叶片氮含量及氮积累量与冠层反射光谱之间的相关关系^[9,10], 但未建立氮素营养光谱诊断模型。由以上分析可知, 通过作物叶片反射光谱估测氮含量是可行的。

在茶叶成分分析方面, 国内外学者先后利用近红外光谱方法定性和定量地分析了茶叶中咖啡碱、氨基酸、茶多酚、蛋白质以及水分的含量^[11-15], 这些成果有利于茶叶品质的快

收稿日期: 2007-08-08, 修订日期: 2007-11-26

基金项目: 国家自然科学基金项目(60575020), “十一五”国家科技支撑计划项目(2006BAD11A13), 江苏大学博士创新基金项目和现代农业装备与技术重点实验室开放基金项目(NZ200602)资助

作者简介: 胡永光, 1976 年生, 江苏大学农业工程研究院博士生 e-mail: deerhu@163.com

*通讯联系人 e-mail: lipingping@ujs.cn

速分析以及品质控制。还有一些学者进行了茶叶种类、真伪的识别方面的研究^[16-21], 其结果较好, 说明利用近红外光谱技术可以快速、准确、无损的对茶叶品种进行鉴别。以上都是以加工成品茶为研究对象, 而陈斌对茶鲜叶中叶绿素含量作了初步探讨^[22], Hu 等基于 SPAD 值建立了茶鲜叶相对氮含量的光谱预测模型, 但不能用此模型预测叶片氮含量精确值^[23]。

本研究以茶树活体的鲜叶为研究对象, 采用便携式光谱仪在田间原位采集叶片漫反射光谱, 建立茶鲜叶全氮含量与可见-近红外光谱之间的关系模型, 为茶鲜叶快速无损氮含量的田间在线检测和施肥管理提供切实可靠的依据和基础。

1 材料与方法

1.1 样品的选择

本试验于 2007 年 4 月在江苏省丹阳市进行, 茶鲜叶全部采自丹阳迈春茶场。由于 SPAD-502 型叶绿素仪能相对定量叶片的氮含量, 故根据茶树种类、茶树年龄、叶位、不同地理位置、土壤肥力等情况, 先用其测定叶绿素含量来估测氮含量并进行标记, 以便获得范围大、分布均匀且具有一定氮含量梯度的样品集。试验中共标记 111 个样品。

1.2 光谱的采集

1.2.1 仪器设备

美国 ASD(analytical spectral device)的 FieldSpec 3 便携式光谱仪, 光谱范围为 350 ~ 2500 nm, 采样间隔为 1.4 nm (350 ~ 1 000 nm 区间)和 2 nm (1 000 ~ 2 500 nm 区间), 分辨率为 3 nm (350 ~ 1 000 nm 区间)和 10 nm (1 000 ~ 2 500 nm 区间), 采样频率 10 Hz。该光谱仪可通过无线网络由笔记本电脑操作控制。

1.2.2 光谱的扫描

测量选择在晴朗无云的天气进行, 测量时间为 9:30 ~ 14:30, 测量方式为漫反射式, 选用 8 视场角的探头。测量时保持光谱仪距离样本表面 2 ~ 3 cm, 以确保视场域范围在叶片上。测定前进行系统配置优化和白板校正。每个样本重复测量 3 次, 每条谱线经过 10 次扫描。

1.3 全氮含量的测定

样本在 105 °C 鼓风干燥箱中先烘 10 ~ 30 min 进行杀青, 再降温到 65 °C 保持一段时间, 直至逐尽水分。烘干的样品粉碎后全部过筛, 再采用凯氏定氮法测定全氮含量。测得全氮含量浓度范围为 1.423 % ~ 3.515 %。

2 模型结果与分析

2.1 可见-近红外光谱波段分析

所有采集到的反射光谱先经过 ASD ViewSpec Pro 4.05 软件进行有效性检查, 剔除测量的粗大误差, 并取平均值作为样本的光谱测量值。将数据以文本格式导出, 用于建模。

茶鲜叶的典型光谱曲线如图 1 所示, 由图可知茶叶叶片反射光谱的基本特征。在可见光区域, 叶片的光谱特征受色素, 主要是叶绿素的影响, 主要吸收蓝光和红光, 因此在 680

nm 附近的红光区域和 460 nm 的蓝光区域有吸收谷, 550 nm 附近是叶绿素的强反射峰^[13]。在 680 ~ 750 nm 区域内, 由于受叶片和冠层散射的影响, 反射率值随着波长的增加而急剧升高。茶叶叶片受组织结构影响, 对近红外区域的 750 ~ 1 300 nm 波段吸收较少, 反射和透射可高达 50 %, 因此这一区间出现一个较高的反射平台。在长波近红外光谱带 (1 300 ~ 2 500 nm) 的 1 450 和 1 940 nm 附近, 曲线有两个水吸收作用引起的吸收谷。由于大气中水分吸收光的作用, 1 400 和 1 800 nm 附近出现了干扰, 导致出现反射率远大于 1 的现象, 并在曲线上形成了大量垂直的干扰直线。为保证模型精度, 建模时将 1 350 ~ 1 400, 1 800 ~ 1 920 和 2 400 ~ 2 500 nm 波段去除。

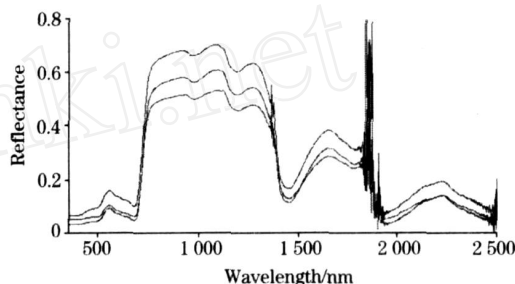


Fig 1 Typical visible-near infrared reflectance spectra of fresh tea leaf samples

2.2 可见-近红外光谱波段分析

本试验数据的处理采用江苏大学自主开发的 NIRSA 软件系统(计算机软件著作权登记号 2007SR06801)。该系统专门用于近红外光谱数据分析和建模, 由数据编辑、数据预处理、样品分析、谱图比较、模型校正、样品预测、报告编辑和谱图显示效果设置等功能模块组成。其中模型校正模块提供多元线性回归、逐步回归、主成分回归、偏最小二乘回归、主成分/偏最小二乘与神经网络集成的建模算法。

2.2.1 预处理方法

为了去除来自高频随机噪声、基线漂移、光程变化等的影响, 需要对样本的光谱数据进行预处理。这里采用了导数、滑动平均滤波(MAF)、归一化等预处理方法。

2.2.2 模型的建立与预测

本研究尝试用逐步回归(SWR)、主成分(PCAR)、偏最小二乘(PLSR)、偏最小二乘与神经网络相结合(PLS + ANN)等方法试建模型, 以回判系数、回判均方根误差(RMSEC)作为衡量模型优劣的指标。采用 SWR, PCAR 和 PLS + ANN 建立的模型其回判系数都小于 0.8, 而 PLSR 模型的回判系数都在 0.9 以上, 且 RMSEC 较小, 故选用偏最小二乘回归方法建立校正模型。

采用 PLSR 模型对预测集中的样本进行预测, 比较预测集中的全氮含量实测值与预测值的相关系数 r 及预测均方根误差(RMSEP), 最后得到预测集中相关系数最高、均方根误差最小的模型即为最佳模型。

表 1 给出了不同预处理方法、不同主成分因子数时 PLSR 预测模型的结果。用预测均方根误差及预测集中实测值与预测值之间的相关系数 r 等指标来评价模型的优劣。

由表 1 可知一阶导数比二阶导数的预测结果好。因平滑处理可有效平滑高频噪声, 又将滑动平均滤波(MAF)与导数处理相结合, 采用一阶导数与 MAF 相结合所得预测模型最佳, 其预测相关系数为 0.888 1, 均方根误差为 0.130 4。但将其与归一化的预处理方法相结合, 结果并未得到优化, 其预测精度反而有所下降。

采用一阶导数与滑动平均滤波相结合的预处理方法, 选用 7 个主成分时建立的偏最小二乘回归模型最佳, 校正集均方根误差 RMSEC 为 0.097 3。用此模型对已测定全氮含量的 25 个样品进行预测, 所得预测结果见表 2。由表可知, 预

测的平均相对误差为 4.339%, 表明模型可以满足实际应用。

Table 1 Calibration result of total nitrogen content with different preprocessing methods and principal component factors

预处理方法	主成分数	$r(\text{预测})$	RMSEP
一阶导数	6	0.824 3	0.160 9
二阶导数	5	0.740 2	0.193 9
一阶导数 + MAF	7	0.888 1	0.130 4
二阶导数 + MAF	6	0.802 5	0.170 7
一阶导数 + MAF + 归一化	6	0.856 5	0.150 5

Table 2 Prediction result for 25 unknown samples by PLS model

序号	样品编号	实测值/ %	预测值/ %	绝对误差/ %	相对误差/ %
1	3	2.867 0	2.965 0	0.098 0	3.417
2	9	2.495 0	2.683 5	0.188 5	7.554
3	13	2.398 0	2.482 0	0.084 0	3.503
4	20	2.300 0	2.530 8	0.230 8	10.36
5	24	2.464 0	2.478 9	0.014 9	0.605
6	33	1.724 0	1.904 0	0.180 0	10.443
7	39	3.115 0	3.202 4	0.087 4	2.805
8	43	3.048 0	2.878 3	-0.169 7	5.566
9	47	3.146 0	3.085 8	-0.060 2	1.915
10	49	3.254 0	3.140 9	-0.113 1	3.474
11	51	2.907 0	2.952 4	0.045 4	1.561
12	54	2.874 0	2.829 9	-0.044 1	1.535
13	57	2.903 0	3.159 8	0.256 8	8.846
14	61	2.511 0	2.606 2	0.095 2	3.789
15	64	2.688 0	2.534 5	-0.153 5	5.712
16	67	2.501 0	2.491 4	-0.009 6	0.384
17	70	2.478 0	2.529 3	0.051 3	2.069
18	72	2.590 0	2.631 5	0.041 5	1.603
19	75	2.713 0	2.719 1	0.006 1	0.225
20	88	2.593 0	2.717 6	0.124 6	4.807
21	92	2.668 0	2.651 5	-0.0165	0.619
22	98	2.100 0	2.170 7	0.070 7	3.369
23	101	2.373 0	2.154 7	-0.218 3	9.200
24	104	2.191 0	1.969 2	-0.221 8	10.122
25	112	1.807 0	1.903 2	0.096 2	5.324
预测结果统计分析			平均误差	0.107 1	4.339
			最大误差	0.256 8	10.443
			预测均方差	0.130 4	

3 结 论

本研究尝试用可见-近红外光谱技术来快速检测茶鲜叶全氮含量。选用便携式光谱仪, 通过田间试验获取茶鲜叶漫反射光谱信息。选取 86 个样品作校正集, 采用不同的预处理和统计分析方法建立了茶鲜叶全氮含量的预测模型。结果表

明, 利用一阶导数与滑动平滑滤波相结合的预处理方法, 选用 7 个主成分建立的偏最小二乘回归模型可以很好地预测全氮含量。其校正集均方根误差为 0.097 3。对 25 个样本预测所得相关系数为 0.888 1, 预测的均方根误差为 0.130 4。由此表明, 基于可见-近红外光谱预测茶鲜叶全氮含量具有很大潜力, 可用于指导茶树长势监测和施肥管理。

参 考 文 献

- [1] YANG Ya-jun(杨亚军). China Tea Tree Cultivation Science(中国茶树栽培学). Shanghai: Shanghai Scientific and Technical Press(上海: 上海科学技术出版社), 2005. 375.
- [2] JIA Liang-liang, CHEN Xin-ping, ZHANG Fu-suo(贾良良, 陈新平, 张福锁). World Agriculture(世界农业), 2001, (6): 36.
- [3] ZHANG Jin-heng, WANG Ke, WANG Ren-chao(张金恒, 王珂, 王人潮). Journal of Northwest University of Agriculture and Forestry Science and Technology, Natural Science Edition(西北农林科技大学学报·自然科学版), 2003, 31(2): 177.
- [4] ZHU Zhe-yan, BAO Yi-dan, HUANG Min, et al(朱哲燕, 鲍一丹, 黄敏, 等). Journal of Zhejiang University, Agriculture and Life Science(浙江大学学报·农业与生命科学版), 2006, 32(2): 152.
- [5] ZHU Xin-kai, SHENG Hai-jun, GU Jing, et al(朱新开, 盛海君, 顾晶, 等). Journal of Triticeae Crops(麦类作物学报), 2005, 25(2): 46.
- [6] BAO Yi-dan, HE Yong, FANG Hui(鲍一丹, 何勇, 方慧). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2007, 27(1): 62.
- [7] SUN Jian-ying, LI Min-zan, ZHENG Li-hua, et al(孙建英, 李民赞, 郑立华, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(3): 426.
- [8] ZHANG Jin-heng, WANG Ke, WANG Ren-chao, et al(张金恒, 王珂, 王人潮, 等). Journal of Zhejiang University, Agriculture and Life Science(浙江大学学报·农业与生命科学版), 2004, 30(3): 340.
- [9] XUE Li-hong, CAO Wei-xing, LUO Wei-hong, et al(薛利红, 曹卫星, 罗卫红, 等). Journal of Plant Ecology(植物生态学报), 2004, 28(2): 172.
- [10] LI Ying-xue, ZHU Yan, TIAN Yong-chao, et al(李映雪, 朱艳, 田永超, 等). Acta Agronomica Sinica(作物学报), 2006, 32(2): 203.
- [11] SUN Yao-guo, LIN Min, LÜ Jin, et al(孙耀国, 林敏, 吕进, 等). Chinese Journal of Spectroscopy Laboratory(光谱实验室), 2004, 21(5): 940.
- [12] Nahid Mashkouri Najafi, Ahmadi Seyed Hamid, Rajabi Khorrami Afshin. Microchemical Journal, 2003, 75: 151.
- [13] LUO Yi-fan, GUO Zhen-fei, ZHU Zhen-yu, et al(罗一帆, 郭振飞, 朱振宇, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2005, 25(8): 1230.
- [14] Zhang M H, Luypaert J, Fernández Pierna J A, et al. Talanta, 2004, 62(1): 25.
- [15] CHEN Hua-cai, LÜ Jin, CHEN Xing-dan(陈华才, 吕进, 陈星旦). Optics and Precision Engineering(光学精密工程), 2006, 14(1): 58.
- [16] ZHAO Jie-wen, CHEN Quan-sheng, ZHANG Hai-dong(赵杰文, 陈全胜, 张海东). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2006, 26(9): 1601.
- [17] LI Xiao-li, HE Yong, QIU Zheng-jun(李晓丽, 何勇, 裘正军). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2007, 27(2): 279.
- [18] Luypaert J, Zhang M H, Massart D L. Analytica Chimica Acta, 2003, 478: 303.
- [19] Chen Quansheng, Zhao Jiewen, Fang C H, et al. Spectrochimica Acta Part A, 2007, 66: 568.
- [20] He Yong, Li Xiaoli, Deng Xunfei. Journal of Food Engineering, 2007, 79: 1238.
- [21] Budinova G, Vlácil D, Mestek O, et al. Talanta, 1998, 47(2): 255.
- [22] CHEN Bin(陈斌). Journal of Jiangsu University of Science and Technology, Natural Science Edition(江苏理工大学学报·自然科学版), 1996, 17(4): 10.
- [23] Hu Yongguang, Li Pingping, Mao Hanping, et al. Agriculture and Hydrology Applications of Remote Sensing: Proceedings of Asia-Pacific Remote Sensing Symposium 2006, USA: SPIE, 2006. 64110G-1.

Determination of Total Nitrogen Content in Fresh Tea Leaf Using Visible-Near Infrared Spectroscopy

HU Yong-guang¹, LI Ping-ping^{1*}, MU Jian-hua¹, MAO Han-ping¹, WU Cai-cong^{1,3}, CHEN Bin²

1. Key Laboratory of Modern Agricultural Equipment and Technology, Ministry of Education(Jiangsu University), Jiangsu Provincial Key Laboratory of Modern Agricultural Equipment and Technology, Jiangsu University, Zhenjiang 212013, China
2. School of Food and Biological Engineering, Jiangsu University, Zhenjiang 212013, China
3. Institute of Remote Sensing and GIS, Peking University, Beijing 100871, China

Abstract To monitor tea tree growth and nitrogen nutrition in tea leaves, visible-near infrared spectroscopy was used to determine total nitrogen content. One hundred eleven fresh tea leaves of different nitrogen levels were sampled according to different tea type, plant age, leaf age, leaf position and soil nutrients, which covered a wide range of nitrogen content. Visible-near infrared reflectance spectra were scanned under the sunlight with a portable spectroradiometer (ASD FieldSpec 3) in field. The software of NIRSA developed by Jiangsu University was used to establish the calibration models and prediction models, which included spectra data editing, preprocessing, sample analysis, spectrogram comparison, calibration model and prediction model, analysis reporting and system configuration. Eighty six samples were used to establish the calibration model with the preprocessing of first/ second-order derivative plus moving average filter and the algorithm of PLS regression, stepwise regression, principal component regression, PLS regression plus artificial neural network and so on. The result shows that the PLS regression calibration model with 7 principal component factors after the preprocessing of first-order derivative plus moving average filter is the best and correspondingly the root mean square error of calibration is 0.973. Twenty five unknown samples were used to establish the prediction model and the correlation coefficient between predicted values and real values is 0.888 1, while the root mean square error of prediction is 0.130 4 with the mean relative error of 4.339 %. Therefore, visible-near infrared spectroscopy has a huge potential for the determination of total nitrogen content in fresh tea leaves in a rapid and nondestructive way. Consequently, the technique can be significant to monitoring the tea tree growth and fertilization management.

Keywords Visible-near infrared spectroscopy; Partial least squares regression; Fresh tea leaf; Total nitrogen content

(Received Aug. 8, 2007; accepted Nov. 26, 2007)

* Corresponding author