

基于近红外光谱分析的土壤全氮含量估测研究^①

张娟娟^{1,2}, 熊淑萍¹, 时雷^{1,2}, 马新明^{1,2}, 王高³

(1 河南粮食作物协同创新中心, 郑州 450002; 2 河南农业大学信息与管理科学学院, 郑州 450002;

3 上海出入境检验检疫局, 上海 201108)

摘要:应用近红外光谱分析技术对比研究基于土壤风干样本和鲜样来预测全氮含量的可行性。选取水稻土为研究对象,首先分析了不同水分土壤的光谱特征,显示随水分含量增加,吸光度升高,且鲜样的吸光度高于干样。通过比较不同预处理方法,对土壤干鲜样分别采用逐步多元回归(SMLR)和偏最小二乘法(PLSR)建立了相应的近红外模型。结果表明,利用近红外光谱均可预测干鲜土壤样本的全氮含量,特别是利用偏最小二乘法建立的标定模型,预测精度高,反演性较好,鲜样和干样外部验证决定系数分别达到 0.89 和 0.91,相对误差仅为 6.92% 和 5.92%,研究结果可以为田间土壤全氮含量的估测提供技术依据和参考。

关键词:近红外光谱;土壤;全氮;PLSR;SMLR

中图分类号:S153.6

土壤是植物生长的基质,为植物生长提供各种营养物质和水分。氮是植物生长发育所必需的营养元素,其丰缺程度直接影响植物的长势及产量和品质。目前传统的湿式化学法如 Kjeldahl 或 Bremner 法仍被广泛地用于测定土壤中的全氮、无机氮和有效氮^[1],但这些方法通常耗时、费工、有害或有污染,且难以满足实时监测土壤全氮的需要。近红外(NIR)技术是 20 世纪 60 年代兴起的一种快速分析技术,已在农业及其他许多领域得到广泛应用^[2-5]。土壤在近红外波段的光谱特征与官能团的振动模式有关,具有强分子链原子的泛频和合频,如 C-H、N-H、O-H 结合的化学键,因而可以利用 NIR 技术测量土壤的某些参数,如水分、有机质和氮等。

目前关于应用近红外监测土壤氮含量的研究已有较多报道。Dalal 和 Henry^[6]发现 1 700 ~ 2 100 nm 区域的土壤光谱与氮素之间关系密切;Chang 和 Laird^[7]显示 NIRS 能够很好地预测土壤有机氮含量;于飞健等^[8]、彭玉魁等^[9]基于多元线性回归应用 NIRS 预测土壤全氮含量;徐永明等^[10]、徐丽华等^[11]、潘涛等^[12]和张娟娟等^[13]采用偏最小二乘法等方法估算土壤的全氮含量。上述研究的对象多是风干土样,为了开展田间土壤养分的实时分析,许多研究直接对

土壤鲜样进行光谱测定并试图建立土壤养分含量的预测模型^[14-15]。然而,所得研究结果并不一致,可能是因为水分含量对土壤光谱曲线的影响,另外光谱预处理方法和建模方法不同也会影响预测精度。为此,本文的目的是分别采用土壤鲜样和风干后样本研究水分影响下土壤近红外光谱特征的变化,并选取逐步多元回归和偏最小二乘法对比分析 NIR 预测鲜样和风干土样全氮含量的精度,以期明确 NIRS 预测土壤氮含量的可行性,为实现土壤养分的近红外快速监测提供技术依据。

1 材料与方法

1.1 土样采集

土壤样品取自江苏南京江宁所设的氮肥试验。供试土壤为水稻土,水稻品种为武香粳 14 和 27123,设置 3 个施氮水平: N 0、270、360 kg/hm²,试验设计为随机区组试验,重复 3 次,于水稻抽穗期采集土样,每小区土壤取样深度分别为 0 ~ 10、10 ~ 20、20 ~ 30、30 ~ 50 cm,土样采集采用梅花法用土钻采集 5 点的混合样本,剔除测试过程中无法测试或有误差的样本,共选取 84 个样本进行分析。表 1 为样品全氮及水分含量的统计特征描述。

基金项目:河南科技攻关项目(112102110030)、河南省教育厅项目(14A210002)、国家十二五科技支撑计划项目(2014BAD10B06)和河南农业大学科技创新基金项目(KJXC2015A12)资助。

作者简介:张娟娟(1979—),女,河南博爱人,博士,讲师,主要从事农业遥感监测研究。E-mail: zhangjuan_2003@126.com

表 1 土壤参数特征统计
Table 1 Statistic characteristics of soil parameters

| 参数 | 样品个数 | 最大值 | 最小值 | 平均值 | 标准差 |
|-------------|------|-------|-------|-------|------|
| 全氮 (g/kg) | 84 | 1.67 | 0.69 | 1.23 | 0.28 |
| 水分含量 (g/kg) | 84 | 315.2 | 156.8 | 223.5 | 32.5 |

1.2 光谱测量

1.2.1 鲜样光谱测量 鲜样测量时,将土样尽量混匀,放入盛样皿时尽量使底部平整。近红外光谱仪使用美国 Thermo Nicolet 公司生产的 Antaris 型傅立叶变换近红外光谱仪,光谱测量范围为 $10\,000 \sim 4\,000\text{ cm}^{-1}$,分辨率为 8 cm^{-1} ,波数精度为 0.1 cm^{-1} 。测试时采用积分球旋转,将处理过的样品盛放在盛样皿里,稍稍压实。每个样品扫描次数为 64,每个样品扫描 3 次,求其平均值。

1.2.2 干样光谱测量 样品经自然风干,木棒压磨后粗略去除植物残体,然后过 20 目筛子混合均匀后进行测量。光谱仪与测试方法同上。

1.3 土壤指标测定方法

土壤全氮采用凯氏定氮法测定;土壤水分含量为质量含水量,按湿基测定方法计算^[16]。

1.4 数据分析方法

首先采用不同的方法对光谱进行预处理,包括平滑、导数变化、多元散射校正等。建模方法包括逐步多元回归和偏最小二乘法。选取 58 个样本进行标定,26 个样本用来检验。模型精度衡量指标主要包括定标决定系数 (R^2)、标准差 (RMSEC),交互验证相关系数 (R_{CV}^2)、交互验证标准差 (RMSECV),外部验证决定系数 (R_v^2)、外部验证标准差 (RMSEP) 及相对误差 (RE)。标准差越小决定系数越高,模型预测效果越好;当 $RE < 20\%$ 时,认为所建模型具有良好的适用性。试验数据采用仪器自带的 TQ Analyst 光谱定量分析软件对模型进行优化。

2 结果与分析

2.1 水分影响下土壤光谱特征变化

图 1 为两个样本风干前后的光谱曲线,由图可知,鲜样比干样表现出较高的吸光度,显著波峰在 $7\,140\text{ cm}^{-1}$ ($1\,400\text{ nm}$) 和 $5\,260\text{ cm}^{-1}$ ($1\,900\text{ nm}$),与水分吸收密切相关。鲜样风干后,与原有光谱的变化规律并不一致,也就是说由于水分的存在有可能掩盖土壤的其他信息。

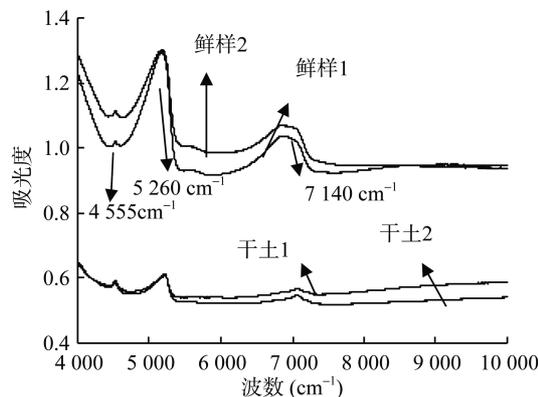


图 1 不同水分含量的土壤光谱变化
Fig. 1 Spectral characteristics of different soil water moistures

2.2 光谱预处理

由于土壤氮素在光谱上并无明显吸收,难以用常规光谱分析方法进行解析,必须对光谱经过一些变换处理后才能提取有效信息。常用的预处理方法有平滑、导数变换、多元散射校正等。本研究选择不同光谱预处理方法进行比较后发现,对鲜样光谱均以平滑后的吸光度结果最优,而干样光谱预测全氮以多元散射校正+平滑+一阶导数效果较好。

2.3 模型构建

表 2 给出了采用两种建模方法得到的标定结果。从表中可以看出,鲜样光谱采用逐步多元回归选取 5 个波段,风干土样光谱选取 3 个波段时相关系数可以达到 0.86 以上,相关系数低于用偏最小二乘法建立的校正模型,内部交互验证的均方根偏差也低于后者,说明采用近红外光谱区全谱区的光谱信息建立的数学模型优于采用逐步多元线性回归方法建立的数学模型。另外,在文中用鲜样光谱建立全氮模型的结果低于利用干样的建模结果,可能是由于水分的存在掩盖了部分信息。为了明确风干后水分对筛选出的全氮敏感波段的影响,进一步分析了风干水分与土壤全氮含量之间的相关关系,结果显示二者相关系数为 0.32,无相关性,表明风干后土壤水分对筛选的波段并无较大影响。此外,对于偏最小二乘法,建模主成分的选择也有不同,成分数选择地越多,校正标准误差越小,且回归系数越大;理论上只要成分数选得足够多,则校正标准误差可以趋向于 0,且回归系数趋向于 1。但成分数过多会造成模拟过度,使得模型个性增强而共性减小,反而使预测误差增大。如鲜样用偏最小二乘法建模选择的主成分数是 9,而风干土样选择的主成分是 5,但当风干土样 > 5 时内部交互验证预测误差反而会增加(图 2)。

表 2 不同建模方法校正模型结果
Table 2 Results of calibration set with different calibration methods

| 参数 | 回归方法 | 建模谱区选择(cm^{-1})/主因子数 | 校正模型 | | 内部交互验证 | |
|------|------|-----------------------------------|-------|------|------------|--------|
| | | | R^2 | RMSE | R_{CV}^2 | RMSECV |
| 鲜样氮 | SMLR | 9 242, 7 077, 6 610, 4 027, 9 141 | 0.89 | 0.09 | 0.86 | 0.10 |
| | PLSR | 9 | 0.90 | 0.02 | 0.94 | 0.08 |
| 风干样氮 | SMLR | 5 333, 4 539, 7 102 | 0.93 | 0.07 | 0.92 | 0.07 |
| | PLSR | 5 | 0.96 | 0.05 | 0.94 | 0.06 |

注：PLSR 选取主因子数，鲜样为 9，风干样为 5。

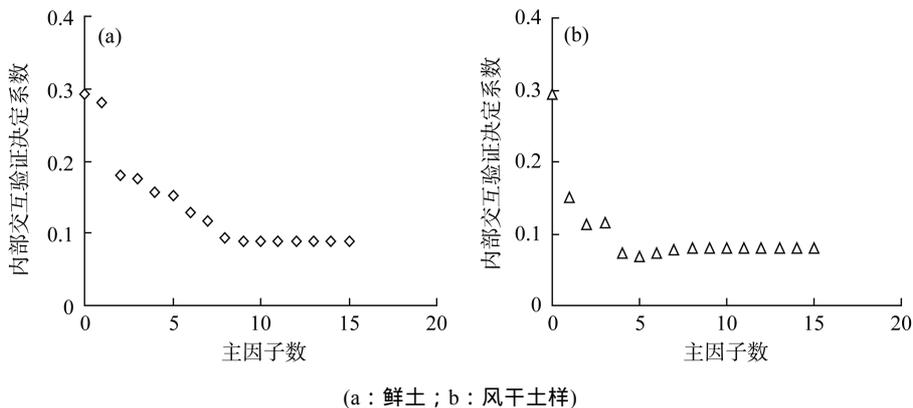


图 2 主因子数对 PLSR 定量分析模型 RMSECV 的影响

Fig. 2 The cross-validated root mean square error (RMSECV) of prediction against the number of factors (NF)

2.4 模型验证

图 3 和图 4 分别为采用逐步多元回归和偏最小二乘法对鲜样和风干土样全氮的外部验证结果。由图可知，利用逐步多元回归建立的模型用来外部检验时决定系数要低于用偏最小二乘法，而均方根偏差高于后者(表 3)。对鲜样和风干土样来说，风干土样预测结果要优于鲜样。另外，模型的 RE 值也是衡量模型好坏的一个重要指标，决定系数越高，RE 值越小，说明模型的预测能力越强，适用性越好。上述表明，这些近红外校正模型的建模效果与预测效果较好，可以用来对未知土壤样品中的氮含

量进行预测。

2.5 讨论

土壤光谱对水分很敏感，含水量对近红外频谱的干扰会掩盖土壤的一些其他信息。由于水分的影响，鲜样和风干土水分利用逐步多元回归筛选的波段具有较大差别。为了明确风干后土壤水分对筛选出的全氮敏感波段的影响，进一步分析了风干水分与土壤全氮含量之间的相关关系，结果显示二者相关系数为 0.32，无相关性，表明筛选的波段与土壤水分并无明显关系。风干后选取的 3 个波段 5 333、4 539 和 7 102 cm^{-1} (1 875、2 203 和 1 408 nm)，与前人研究^[6,17]选取的波段较为一致。

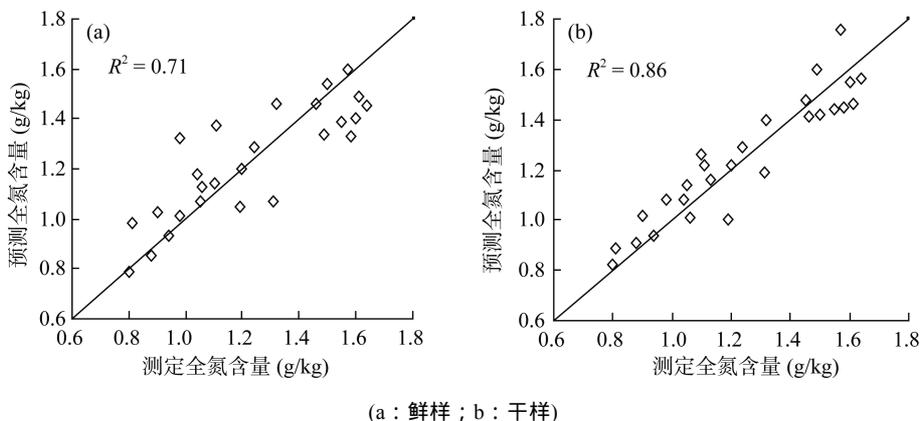


图 3 基于 SMLR 的全氮外部验证结果

Fig. 3 External-validation between measured and estimated values of soil N contents based on SMLR

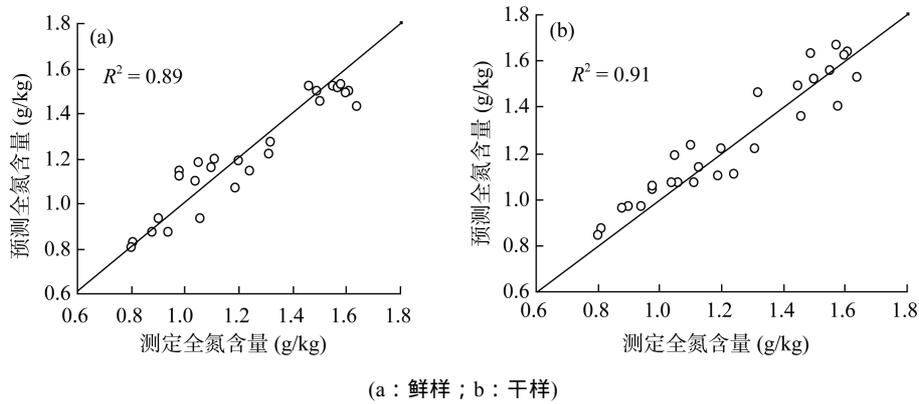


图 4 基于 PLSR 的全氮外部验证结果

Fig. 4 External-validation between measured and estimated values of soil N contents based on PLSR

表 3 基于不同方法建模检验结果
Table 3 External-validation of N contents based on SMLR and PLSR

| 参数 | 回归方法 | 外部验证 | | 相对误差 RE (%) |
|------|------|---------|-------|----------------|
| | | R_v^2 | RMSEP | |
| 鲜样氮 | SMLR | 0.71 | 0.15 | 11.38 |
| | PLSR | 0.89 | 0.09 | 6.92 |
| 风干样氮 | SMLR | 0.86 | 0.11 | 7.54 |
| | PLSR | 0.91 | 0.08 | 5.92 |

在本研究中,尽管水分有较大差异,但采用逐步多元回归和偏最小二乘法结合最优的预处理方法和合适的建模因子数,用鲜样和干样光谱均能实现对全氮含量的有效预测,鲜样相对预测结果较差。这与何勇等^[15]的研究结果不太一致,其认为当土壤保存自然状态时,微量的氮元素很难被预测出来,这可能与其采用的光谱预处理方法有较大关系,选择适当的预处理方法将较多的影响和误差消除后,方能建立较好的预测模型。

3 结论

利用近红外光谱均可预测干鲜土壤样品的全氮含量,特别是利用偏最小二乘法建立的标定模型,预测精度高,反演性较好,鲜样和干样外部验证决定系数分别达到0.89和0.91,相对误差仅为6.92%和5.92%,该结果可以为田间土壤全氮含量的估测提供技术依据和参考。

参考文献:

[1] 李生秀. 关于土壤供氮指标的研究[J]. 土壤学报, 1990, 27(3): 233-240
[2] 汪善勤, 舒宁, 蔡崇法. 基于 PCA 和 SRRI 的潮土土壤属性与田间光谱关系研究[J]. 土壤, 2008, 40 (6): 966-970

[3] 何绪生. 近红外反射光谱分析在土壤学的应用及前景[J]. 中国农业科技导报, 2004, 6(4): 71-76
[4] Cambule AH, Rossiter DG, Stoorvogel JJ, Smaling EMA. Building a near infrared spectral library for soil organic carbon estimation in the Limpopo National Park, Mozambique[J]. Geoderma, 2012, 183-184: 41-48
[5] 任红艳, 史学正, 庄大方, 徐新良, 江东, 于信芳. 基于可见-近红外反射光谱的油区农田土壤石油烃含量估测研究[J]. 土壤, 2013, 45(2): 295-300
[6] Dalal RC, Henry RJ. Simultaneous determination of moisture, organic carbon, and total nitrogen by near infrared reflectance spectrophotometry[J]. Soil Science Society of America Journal, 1986, 50: 120-123
[7] Chang C, Laird A. Near infrared reflectance spectroscopic analysis of soil C and N[J]. Soil Science, 2002, 167: 110-116
[8] 于飞健, 闵顺耕, 巨晓棠. 近红外光谱法分析土壤中的有机质和氮素[J]. 分析实验室, 2002, 1(3): 49-51
[9] 彭玉魁, 张建新, 何绪生. 土壤水分、有机质和总氮含量的近红外光谱分析研究[J]. 土壤学报, 1998, 35(4): 553-559
[10] 徐永明, 蔺启忠, 王璐. 基于高分辨率反射光谱的土壤营养元素估算模型[J]. 土壤学报, 2006, 43(9): 709-716
[11] 徐丽华, 谢德体, 魏朝富, 李兵. 紫色土土壤全氮和全磷含量的高光谱遥感预测[J]. 光谱学与光谱分析, 2013, 33(3): 723-727
[12] 潘涛, 吴振涛, 陈华舟. 土壤总氮近红外光谱分析的波段优选[J]. 分析化学, 2012, 40(6): 920-940
[13] 张娟娟, 田永超, 姚霞, 曹卫星, 马新明, 朱艳. 土壤全氮的高光谱估测研究[J]. 自然资源学报, 2011, 26(5): 881-890
[14] 孙建英, 李民赞, 唐宁. 东北黑土的光谱特性及其与土壤参数的相关性分析[J]. 光谱学与光谱分析, 2007, 27(8): 62-65
[15] 鲍一丹, 何勇, 方慧. 土壤的光谱特征及氮含量的预测研究[J]. 光谱学与光谱分析, 2007, 27(1): 62-65
[16] 鲍士旦. 土壤农化分析[M]. 北京: 中国农业出版社, 1999: 42-49
[17] 张娟娟, 田永超, 姚霞, 曹卫星, 马新明, 朱艳. 同时估测土壤全氮、有机质和速效氮含量的光谱指数研究[J]. 土壤学报, 2012, 49(1): 50-59

Soil Nitrogen Content Prediction with Near Infrared Spectroscopy

ZHANG Juan-juan^{1,2}, XIONG Shu-ping¹, SHI Lei^{1,2}, MA Xin-ming^{1,2}, WANG Gao³

(1 Collaborative Innovation Center of Henan Grain Crops, Zhengzhou 450002, China; 2 College of Information and Management, Henan Agricultural University, Zhengzhou 450002, China; 3 Shanghai Entry-Exit Inspection and Quarantine Bureau, Shanghai 201108, China)

Abstract: This paper was focused on analyzing the feasibility of predicting nitrogen contents in humid and dry soils with Near Infrared Spectroscopy (NIRS). The spectral characteristics of paddy soils in southern China were studied and the effects of different soil moisture contents on soil spectrum were analyzed. With moisture increasing, spectral absorbance hoisted and humid soil showed higher absorbance than dry one by adopting the pretreatment methods of spectrum, the calibration models were established with PLSR and SMLR. The result showed that total nitrogen in humid and dry soil samples both could be predicted by using near infrared spectroscopy, particularly by using the model established by PLSR. From external validation, R_v^2 and RMSEP of dry soil and humid soil were 0.89 and 0.91, 6.92% and 5.92%, respectively. It may provide technical support for measuring soil total nitrogen content in the field.

Key words: Near Infrared Spectral; Soil; Nitrogen; PLSR; SMLR