

土壤质地分类的近红外光谱分析方法研究

曾庆猛¹, 孙宇瑞^{1*}, 严红兵²

1. 中国农业大学信息与电气工程学院, 北京 100083

2. 北京英贤仪器有限公司, 北京 100070

摘要 依据所收集的25种土样,采用两种不同精度仪器、三种光谱方法以及3个不同谱区,在四种分类标准条件下进行土壤质地分类分析方法研究。结果表明:(1)土壤化学组成的信息主要体现在近红外光谱的谱峰特征,而质地等物理信息主要反映在光谱的斜率、截距参数上,且二者在不同谱区的相对强度不同;(2)近红外光谱方法对土壤质地的分辨能力较低,随着分类粗化而有所提高;(3)在4组分类标准中土壤质地最高预测准确度为72%,其中在砂粒<70%和粘粒<40%条件下,预测准确度达到85%;(4)样本顶部漫反射光谱方法与扩展谱区范围均可有效提高质地预测准确度,而高精度仪器并不具备明显优势。

关键词 土壤质地;分类;近红外光谱

中图分类号: S152.3, O657.3 **文献标识码**: A **DOI**: 10.3964/j.issn.1000-0593(2009)07-1759-05

引言

近红外(NIR)光谱分析技术在农业领域得到了广泛应用,其主要信息基础是被测物质有机分子的各级含氢基团倍频与合频在近红外谱区的吸收^[1]。在刻画土壤特性的诸多参数中,含水率、有机质等化学组成分析成为当前光谱研究热点^[2-8]。对于诸如粒度等物理信息的光谱特性,在传统的化学成分分析中被视为由漫反射引起的干扰而采用滤波、微分、多元校正等各种预处理方法加以消除^[9-12]。作为土壤最基本的物理参数之一,土壤质地(Soil texture)被定义为砂粒(Sand)、粉粒(Silt)和粘粒(Clay)3个粒径在土壤构成中的重量百分比。筛分和沉降^[13]是确定土壤质地的经典方法,该方法不仅过程繁琐,而且试验周期长。鉴于农田质地快速分类的客观需求,土壤粒度构成与不同光谱漫反射相关规律性的研究逐渐成为一个新的、有潜在应用价值的探索领域。本文采用三种近红外光谱方法获取25种不同质地土壤样本的光谱信息,运用偏最小二乘法(PLS)^[14]建立质地预测模型和粒度回归模型,从谱线特征、样本漫反射方式、仪器精度和谱区范围不同角度进行性能对比分析,为未来仪器化设计提供参考依据。

1 材料与方法

1.1 土壤质地分类标准

美国农业部(USDA)根据土壤在农田中的持水保肥、通气透水特点,把土壤质地划分为4组12级,如图1所示,图中三个坐标轴对应的粒径分别为砂粒(粒径:2~0.05 mm)、粉粒(粒径:0.05~0.002 mm)和粘粒(粒径:<0.002 mm)。4组类分别为砂土组、壤土组、粘壤土组和粘土组。同时针对土壤剖面研究,根据土壤粒径、矿物性质、温度等特点把土壤质地划分为7级^[15]。此外,把图1中相邻级类视为同一亚类,可把土壤分成相互重叠的28个亚类。在本研究中,应用了上述四种土壤质地分类标准,分别简称为12级类、7级类、28亚类和4组类,在分类程度上呈现由细到粗的趋势。

图1同时显示25种试验土样的分布情况,样本由经典方法^[13]确定土壤质地类型,在本研究前通过烘干方法进行了脱水处理。

1.2 光谱分析仪

研究采用FT(傅里叶变换型)和色散型两种不同近红外光谱分析仪,其中FT光谱仪为MPA近红外光谱仪(BRUKER,德国,简称MPA);色散型近红外漫反射土壤质地光谱分析试验装置由USB2000光导纤维光谱仪(Si-

收稿日期:2008-06-16,修订日期:2008-09-18

基金项目:国家高技术研究发展计划“863”项目(2006AA10Z205,2006AA10A301)资助

作者简介:曾庆猛,1980年生,中国农业大学信息与电气工程学院博士研究生 e-mail: zengqingmeng@yahoo.com.cn

*通讯联系人 e-mail: pal@cau.edu.cn

CCD 阵列检测器, Ocean Optics, 美国) 以及光学系统与计算机等组成(简称 USB2000), 如图 2 所示。

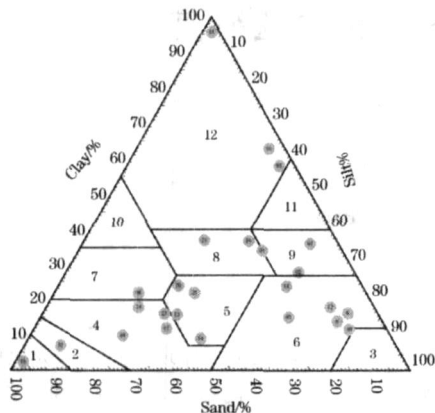


Fig. 1 USDA soil texture classification (12-subclasses) and soil samples distribution

1: Sand; 2: Loamy sand; 3: Silt; 4: Sandy loam;
5: Loam; 6: Silt loam; 7: Sandy clay loam; 8: Clay loam;
9: Silty clay loam; 10: Sandy clay; 11: Silty clay; 12: Clay

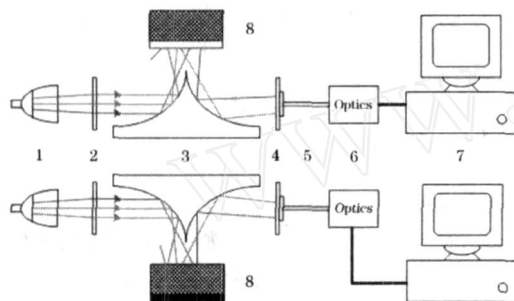


Fig. 2 Soil SNIR spectral analysis device based on USB2000, with methods of diffuse reflection from (a) bottom and (b) top of sample

1: Gilded tungsten; 2: Optic filter; 3: Reflector; 4: Fresnel lens;
5: Optic fiber; 6: USB 2000; 7: PC; 8: Sample cell

图 2 中镀金反射碗钨灯 (1) 发出的光经过滤光片 (2) 用以提高短波近红外谱区相对强度后, 再通过反射镜组 (3) 在样本 (8) 的顶部或底部表面产生漫反射光, 由菲涅尔透镜 (4) 会聚, 经光纤 (5) 导入光谱仪 USB2000 (6), 光谱由计算机 (7) 进行数据处理。

近红外光谱仪 MPA 和 USB2000 的主要性能差异如表 1 所示, 可见 MPA 具有更宽谱区、更高分辨率和精度等特点。

Table 1 Basic parameters of MPA and USB2000

名称	谱区/nm	分辨率	波长/精确度	波长/准确度
MPA	833~2 560	0.3 nm @ 1 250 nm	0.05 cm ⁻¹	0.1 cm ⁻¹
USB2000	750~1 100	10 nm	0.1 nm	0.2 nm

1.3 光谱分析方法

上述两种光谱仪应用三种不同光谱方法分析土壤样本, 分别是:

方法 1 为 FT 型全谱近红外、样本底部漫反射方式, 在 MPA 上直接测定, 样本厚度 2 mm。

方法 2 为色散型短波近红外、样本底部漫反射方式, 如图 2(a) 所示, 样本厚度为 2 mm。

方法 3 为色散型短波近红外、样本顶部漫反射方式, 如图 2(b) 所示, 样本厚度为 5 mm。与方法 2 的区别在于样本置于反射镜组下方, 入射光不经过玻璃直接入射到样本顶部。

在以上三种光谱方法中, 每个样本重复测量三次, 第三次重新装样, 结果取平均。

1.4 数据处理方法

数据处理由 JMP 软件中的 PLS 平台实现。在分析、预测土壤质地时, 首先采用 PLS 建立粒级含量的定量预测模型, 然后通过不同质地分类标准确定预测后样本所属类型, 建模过程采用 Leave-one-out 抽样方式。

2 结果与讨论

2.1 土壤样本不同漫反射方法的光谱特征

随机选择的相同两个土样用三种光谱方法扫描其近红外漫反射光谱如图 3 所示。

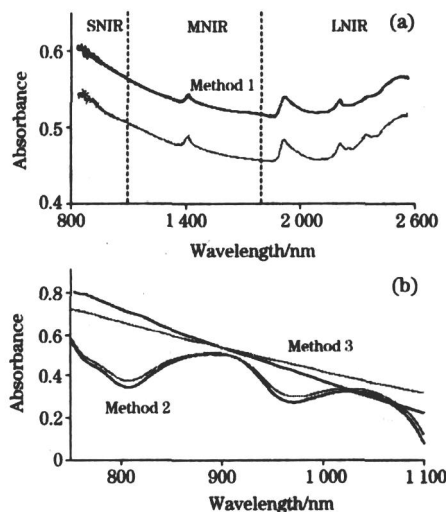


Fig. 3 NIR spectrum of soil samples
(a): Method 1; (b): Method 2 and 3

由漫反射光谱理论^[1]可知, 样本的漫反射光谱决定于吸收系数 K 与散射系数 S , 前者主要反映样本的化学组成或结构信息, 在近红外谱区主要是含氢基团的倍频与合频; 后者主要反映样本的物理结果信息, 因此漫反射光谱同时包含了样本的化学信息和物理信息两方面, 但在不同谱区两者的相对强度不同。

图 3(a) 为土样在近红外全谱区的光谱, 包括短波近红外 SNIR (833~1 100 nm) 谱区、中波近红外 MNIR (1 100~1 800 nm) 谱区与长波近红外 LNIR (1 800~2 560 nm) 谱区。在 1 450, 1 940 和 2 250 nm 等处分别为土样中结合水的三倍频、二倍频与合频的谱峰, 可以看出其谱峰强度随着倍频的倍次增加而降低^[1]。相对而言, LNIR 谱区比 SNIR 谱区

包含了更多的土样化学信息,而对于土样的物理信息,在不同谱区的差异不是很明显。另外,本研究应用的 MPA 光谱仪配置的检测器对 SNIR 谱区响应相对较低,因此该谱区的信噪比较低。

图 3(b) 为土样在 SNIR 谱区的光谱,可以认为方法 2 和 3 的光谱差异是由于方法 2 样本池底部玻璃以及玻璃上沉积的细微颗粒(粘粒或粉粒)而引起了漫反射差异造成的。由方法 3 的光谱可见,土壤的物理信息主要反映在谱线的斜率、截距参数上,由此显见土壤样本近红外光谱中物理结构的信息量远比化学信息低。近红外光谱采用 PLS 算法分析土样物理结构(粒级含量),建立数学模型时,各 PLS 主成分对土壤粒级信息的累积贡献率(或称为“解释率”, Percent Explained %)如图 4 所示。

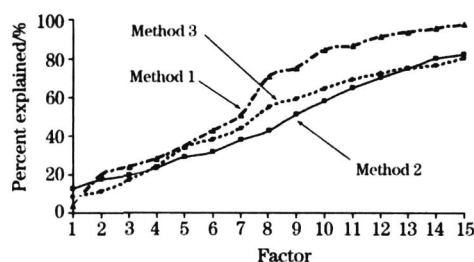


Fig 4 Percent variation explained by PLS factor

由图 4 可知,随着主成分数目的增加,样本土壤粒级信息的累积贡献率增速缓慢,达到 15 个主成分时,三种光谱方法的累积贡献率分别为 97%, 82% 和 80%。近红外光谱分析为了避免模型的过拟合,一般只能运用尽量少的 PLS 主成分;对于土壤样本化学待测量信息,运用 3~5 个 PLS 主成分,其累积贡献率即可达到 90% 以上^[2-8],对于土壤样本物理结构信息,5 个 PLS 主成分的累积贡献率只有 40% 以下,难以得到较好的分析结果。这也进一步说明近红外光谱中样本的物理信息相对比化学信息弱得多,并预示了本研究应用近红外光谱分析土壤物理结构相对于其他化学组成分析将比较困难,对土壤质地的分辨能力有限。但近红外光谱方法研究土壤物理结构在便捷性等方面比传统方法有明显优势,对此需要进行深入的研究。

由图 4 进一步可知,随着主成分数目的增加,三种光谱方法的样本土壤粒级信息的累积贡献率的增速是不同的,其中方法 1 的累积贡献率增加速度最快,预示了该光谱方法可能比另外两种方法的效果更好;方法 3 在一定范围内其主成分累积贡献率增速比方法 2 高,也预示了方法 3 在质地分辨上可能优于方法 2。

2.2 质地预测性能比较

由交叉验证(Cross validation)确定了 4 个 PLS 成分并进行粒级含量的建模,此时三种光谱方法的累积贡献率分别为 28%、23% 和 24%。由预测模型的粒级含量便可确定土壤的质地类型,在四种质地分类标准中正确预测土壤质地的个数如图 5 所示。另外对方法 1 中的三个分谱区 SNIR、MNIR 和 LNIR 采用相同的数据处理方法,结果如图 6 所示。

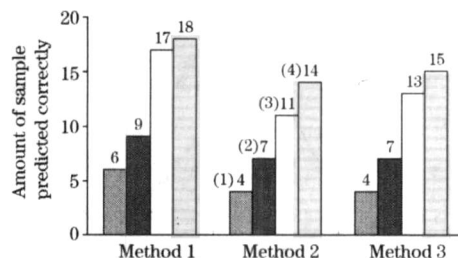


Fig. 5 Amount of samples predicted correctly in 3 spectral methods and 4 texture systems

(1): 12-subclasses; (2): 7-subclasses;
(3): 28-clusters; 4: (4)-classes

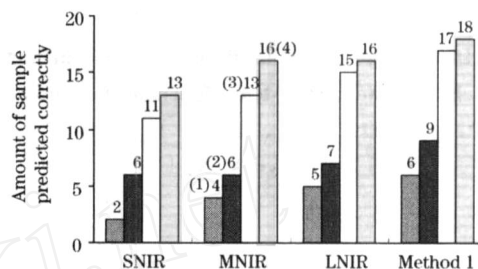


Fig. 6 Amount of samples predicted correctly in method 1 and 4 texture systems at sub-spectrum

(1): 12-subclasses; (2): 7-subclasses;
(3): 28-clusters; 4: 4-classes

分析图 5 和图 6 中结果可知:

(1) 在相同质地分类标准下方法 1 的预测准确度均为最高,分别是: 12 级类为 24 % (6/25), 7 级类为 36 % (9/25), 28 亚类为 68 % (17/25) 和 4 组类为 72 % (18/25), 说明近红外光谱方法对土壤质地的分辨能力较低,因此随着质地分类的由细到粗,而光谱的预测准确度有所提高,主要原因在于光谱携带的物理信息较少。从粒级含量角度看,以方法 1 在 4 组分类标准为例,预测正确的样本主要位于砂粒 < 70 % 以及粘粒 < 40 % 范围,如图 7 所示,在此范围预测准确度达到 85 % (17/20)。从预测结果分布看,与壤土和粘壤土组类结果相比,对砂土和粘土组类的预测准确度较低,部分原因与样本数量及分布均匀程度有关。

(2) 在图 5 所示结果中,尽管方法 2、方法 3 在 12 级和 7 级分类标准上预测结果相同(4 和 7),但在 28 亚类和 4 组类标准上方法 3 的预测准确度却略高于方法 2,表明顶部漫反射比底部漫反射方式更有利于土壤质地的预测。这种差异显然与方法 2 中样本池底部玻璃有关,同时验证了 PLS 主成分累积贡献率的差异与土壤质地分辨能力之间的关系。

(3) 图 6 中 SNIR 谱区和图 5 中方法 2 的所在谱区、样本反射方式类似,但方法 1 所用光谱仪的精度虽然远高于方法 2,但其栅测结果对比(12 级类: 4 vs. 2; 7 级类: 7 vs. 6; 28 亚类: 11 vs. 11; 4 组类: 14 vs. 13)表明,高精度光谱仪在质地预测上并未具备明显优势。

(4) 由图 3(a) 可知方法 1 在 MNIR 和 LNIR 的信噪比较低,对比图 6 中 MNIR, LNIR 和全谱区的预测结果(12 级类: 4 vs. 5 vs. 6; 7 级类: 6 vs. 7 vs. 9; 28 亚类: 13 vs.

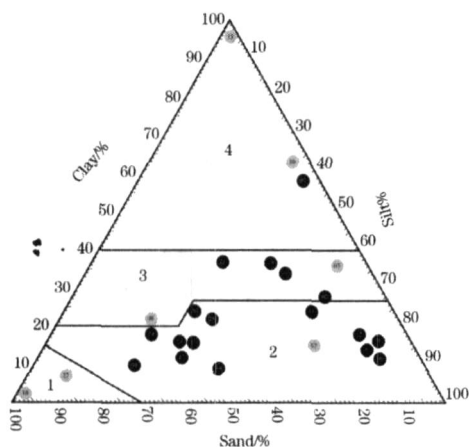


Fig 7 Distribution of predicted samples from method 1 in 4 classes of texture system (● denotes the sample predicted correctly)

1: Sand; 2: Loamy; 3: Clay loamy; 4: Clay

15 vs 17; 4 组类: 16 vs 16 vs 18), 可见全谱区预测结果优于分谱区, 表明谱区的展宽有利于土壤质地预测。

3 结 论

通过两种不同精度仪器、三种光谱方法、3 个谱区条件, 基于四种质地分类标准进行土壤质地分类研究, 表明近红外光谱分析方法用于土壤质地分类研究的可行性。不同光谱方法的光谱特征表明谱线的斜率、截距参数反映了土壤样本的质地信息, 且相对于化学信息, 物理信息比较少。通过 Leave-one-out 抽样方法、4 个 PLS 成分建立三种光谱方法以及方法 1 在 3 个分谱区的预测模型的比较分析, 表明光谱方法对土壤质地的分辨能力较低; 在砂粒 < 70 % 和粘粒 < 40 % 条件下, 对于四组类的土壤质地分类标准, 最高预测准确度达到 85 %; 不同光谱方法条件下的预测结果表明样本顶部漫反射、展宽谱区均能提高土壤质地预测准确度, 而高精度光谱仪并未体现明显优势。

参 考 文 献

- [1] YAN Yan-lu, ZHAO Long-lian, HAN Dong-hai, et al (严衍禄, 赵龙莲, 韩东海, 等). Foundation and Application of NIR Spectral Analysis (近红外光谱分析基础与应用). Beijing: China Light Industry Press (北京: 中国轻工业出版社), 2005. 13.
- [2] Morra M J, Hall M H, Freeborn L L. Soil Sci. Soc. Am. J. 1991, 55: 288.
- [3] Sudduth K A, Hummel J W. Trans. ASAE, 1993, 36(1): 187.
- [4] Sudduth K A, Hummel J W. Trans. ASAE, 1993, 36(6): 1571.
- [5] Viscarra Rossel R A, Walvoort D J J, McBratney A B, et al. Geoderma, 2006, 131: 59.
- [6] BAO Yi-dan, HE Yong, FANG Hui, et al (鲍一丹, 何勇, 方慧, 等). Spectroscopy and Spectral Analysis (光谱学与光谱分析), 2007, 27(1): 62.
- [7] GUO Qing-lin, ZHOU Yu-long, ZHANG Qiu-lin, et al (郭庆林, 周玉龙, 张秋琳, 等). Spectroscopy and Spectral Analysis (光谱学与光谱分析), 2008, 28(1): 200.
- [8] CHEN Peng-fei, LIU Liang-yun, WANG Ji-hua, et al (陈鹏飞, 刘良云, 王纪华, 等). Spectroscopy and Spectral Analysis (光谱学与光谱分析), 2008, 28(2): 295.
- [9] YAN Yan-lu, ZHAO Long-lian, LI Jun-hui, et al (严衍禄, 赵龙莲, 李军会, 等). Spectroscopy and Spectral Analysis (光谱学与光谱分析), 2000, 20(6): 777.
- [10] ZHAO Huan-huan, YAN Yan-lu (赵环环, 严衍禄). Spectroscopy and Spectral Analysis (光谱学与光谱分析), 2006, 26(5): 842.
- [11] MIN Shun-geng, XIE Xiu-juan, ZHOU Xue-qiu, et al (闵顺耕, 谢秀娟, 周学秋, 等). Chinese Journal of Analytical Chemistry (分析化学), 1998, 26(1): 34.
- [12] GAO Rong-qiang, FAN Shi-fu, YAN Yan-lu, et al (高荣强, 范世福, 严衍禄, 等). Spectroscopy and Spectral Analysis (光谱学与光谱分析), 2004, 24(12): 1563.
- [13] Gee G W, Bauder J W. Particle-Size Analysis, in Klute A (Ed.). Methods of Soil Analysis, Part 1 (2nd edition), Agronomy Monographs. No.9, Madison: American Society of Agronomy, 1986. 383.
- [14] WANG Hui-wen (王惠文). Partial Least-Squares Regression Method and Applications (偏最小二乘回归方法及其应用). Beijing: National Defense Industry Press (北京: 国防工业出版社), 1999. 152.
- [15] Ward Chesworth, Encyclopedia of Soil Science. Dordrecht: Springer, 2008, 705.

NIR Spectral Analysis for Soil Textural Classification

ZENG Qing-meng¹, SUN Yurui^{1*}, YAN Hong-bing²

1. College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

2. Beijing INCE Instrument Co., Ltd., Beijing 100070, China

Abstract Using 25 soil samples with known textural compositions, 2 types of NIR instruments, 3 spectral methods associated with 3 spectrum ranges and 3 sampling intervals, the approach to soil textural classification was investigated. From the results obtained, the following conclusions can be drawn: (1) The chemical information could be identified from the peak of the spectral curves, whereas the slope and intercept of spectral curves concerning soil texture resulted from the physical properties of soil samples. Moreover, the intensity of chemical and physical properties varied in different spectra; (2) The distinguishing ability of NIR was limited, depending on the classification criterion proposed; (3) Being tested with four classification criterions, the maximal predicting probability was 72%. In the case of sand < 70% and clay < 40%, the maximum was up to 85%; (4) Either acquiring scatter information from the surface of soil samples or extending spectral bands could improve the predicting probability.

Keywords Soil texture; Classification; Near infrared

(Received Jun. 16, 2008; accepted Sep. 18, 2008)

* Corresponding author