土壤有机质含量估测及其影响因素的光谱分析

胡文生[1]，任红艳[1,2*]，庄大方[1]，史学正[3]，刘绍贵[4]，黄耀欢[1]，于信芳[1]

(1. 中国科学院地理科学与资源研究所 资源环境与信息系统国家重点实验室，北京 100101；
2. 华东师范大学地理信息科学教育部重点实验室，上海 200062；3. 中国科学院南京土壤研究所，南京 210008；
4. 江苏省扬州市土壤肥料站，扬州 225101)

摘要：土壤颗粒大小差异使土壤反射光谱产生相应变化，影响土壤有机质含量等属性的光谱预测精度。本研究准备了颗粒粒径分别为 2、0.25 和 0.15mm 的土样，测定土壤有机质(Soil Organic Matter，SOM)含量，并于室内模拟条件下测定其反射光谱。通过分析不同粒径土样的原始(Raw)、多次散射校正(Multiple scattering correction，Msc)、一阶微分(First derivative，Fd)、连续统去除(Continuum removal，Cr)光谱与 SOM 含量之间的关系，筛选出与 SOM 含量相关性最强的 Fd 光谱单波段(2 250nm，$r=0.82$，$P<0.01$)，并建立线性回归模型；利用全波段光谱反射率，以偏最小二乘回归(Partial least square regression，PLSR)方法，确立 2mm 土样 Msc 处理光谱的 PLSR 模型为最优模型(RPD=3.56、$R^2=0.90$、RMSEP=1.96g/kg)。土壤颗粒粒径对土壤光谱反射率变化有明显影响，但二者之间并非简单的线性关系，可能存在一个转折点；单变量(单波段光谱反射率)线性回归模型的预测能力，明显低于全波段反射光谱(Msc 处理)-PLSR 模型；土样样本容量对 SOM 含量预测精度有显著影响。因此，根据样本容量大小，选择合适的土壤颗粒粒径与光谱预处理方法组合可以提高预测精度。

关键词：土壤颗粒粒径；反射光谱；有机质含量；光谱预处理

DOI：10.3724/SP.J.1047.2012.00258

1 引言

土壤有机质含量及其快速有效的测定，对于土壤固碳潜力、生态系统固碳能力，以及全球变化等研究具有重要意义[1]。利用土壤的可见-近红外反射光谱，可以快速、准确地实现土壤有机质、湿度等土壤物理性状和土壤氮素、重金属含量等土壤属性的测定，同时也是土壤属性遥感监测的重要基础，越来越受到研究人员的关注[2-5]。与最常见的物理化学分析方法相比，光谱估测法提高了土壤属性估测速度，但估测精度有所降低。光谱估测法以土壤光谱反射率间接地估算土壤属性，土壤水分、土壤类型、土壤质地、颗粒粒径大小等影响土壤光谱反射率的因素在不同程度上导致土壤属性估测精度的下降。土地耕作过程中的反复碾压使土壤粗糙度、团聚体和土块大小发生变化，土壤的反射、颗粒间的散射会随之变化。土壤表层颗粒越粗，土壤颗粒反射的光线与其他颗粒拦截的可能性越大，这就会整体降低土壤的光谱反射率[6-8]。Bänninger 和 Flühler 通过颗粒大小与反射率的差异试验发现，土壤颗粒从 0.063～0.071mm 增加到 0.45～0.55mm，土壤光谱反射率会明显降低[9]。有人认为，小于 2mm 的原状土与 0.1mm 粒径水平土样的土壤属性的光谱预测精度相近或者更细颗粒大小(0.2mm、0.5mm 或者 1mm)难以显著提高 2mm 土壤属性光谱预测的精度[10-11]，即准备 2mm 的土壤样品即可满足预测精度要求。目前，有关土壤属性或者其他组分光谱预测的研究中，所采用的颗粒大小不尽一致，除采用原位土壤样品光谱外[12-14]，有人采用 2 mm[15-17]、0.5mm[18] 等；然而，关于土样颗粒粒径对土壤属性光谱估测精度的影响，以及适合不同土壤属性光谱估测的颗粒粒径水平的研究与报道鲜见。利用不同光谱预处理方法，本文拟开展不同颗粒粒径水平和样本量的土壤有机质含量的光谱估测研究，探讨实现较高精度预测的合理土壤颗粒粒径、样本量与预处理方法组合，为土壤有机

收稿日期：2011-12-06；修回日期：2012-02-13.
基金项目：地理信息科学教育部重点实验室开放研究基金资助项目(KLGIS2011A13)和国家自然科学基金项目(41001279)。
作者简介：胡文生(1973-)，女，博士。主要研究方向为资源环境与遥感应用。E-mail：huwsh@lreis.ac.cn
*通讯作者：任红艳(1979-)，男，博士，助理研究员，主要研究方向为环境遥感应用。E-mail：renhy@lreis.ac.cn

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

质含量的有效、快速估测提供理论与方法支持。

2　研究方法与数据预处理

真武镇地处扬州市东郊、江都市北部,归江都市行政管辖。江都在区域地质构造上位于苏北盆地高邮凹陷的南半部。江都市气候属副热带湿润气候区,年平均气温 14.9℃,降水量 978.7mm,四季分明,无霜期较长。真武镇土壤类型主要为江河冲积物上发育而来的潴育型水稻土。目前,大田农作物主要有小麦、油菜、水稻。

2.1　土壤样品采集与有机质含量测定

土壤样品采集于 2009 年 3 月上旬,每点采集大约 200g 的农田表层土壤(0～20 cm),。共采集 146 个土壤样品。样品带回实验室之后,除去样品中的石块、植物根系和凋落物等后在室内风干。首先,进行研磨并分别过 2 mm、0.25 mm、0.15 mm 筛子(研磨后土壤颗粒粒径分别记为 P1、P2 与 P3):所有样品均过 2 mm 筛子;然后,用 4 分法取对角样品继续研磨过 0.25 mm 筛子;最后,以同样的方法将所有过 0.25mm 筛子的样品研磨再过 0.15 mm 筛子。过 0.15 mm 筛子样品进行土壤有机质含量分析,采用重铬酸钾容量法-外加热法[19]测定样品土壤有机质含量。

2.2　土壤光谱测定与光谱数据预处理

在大气环境相对稳定的室内条件下,利用光谱测量仪器采集 P1、P2 和 P3 土壤样品的反射光谱。光谱测量仪器采用 ASD_Fieldspec (350-2500),仪器设备参数及土壤光谱的具体测量方法参见文献[2]。利用多次散射校正(Multiple scatter correction, Msc)、一阶微分(First derivative, Fd)和连续统去除(Continuum removal, Cr)等 3 种方法[2]对土壤光谱数据进行预处理。

偏最小二乘回归(PLSR, Partial least square regression)通过压缩高度相关的波长变量解决了多重共线性问题,充分利用并概括了这些光谱信息,而且使这些压缩而生成的独立变量(成分 PC,Principal components)可以很好地解释因变量变化。在模型建立和预测过程中,会分别给出建模(Calibration)和预测(Prediction)的均方根误差:RMSEC (Root mean square error of calibration)和 RMSEP (Root mean square error of prediction)。模型的确定取决于相对最高的决定系数(R^2)和相对最低的 RMSE;此外,结合因变量的标准误差(Standard Deviation, SD)与模型预测结果的 RMSEP 的比值,即 RPD(the ratio of SD of samples to RMSEP)是否大于 2 对 PLSR 模型的预测精度进行评价和比较[5]。

3　土壤有机质含量的预测及其含量变化光谱响应分析

3.1　不同粒径土壤可见-近红外光谱

随机选取土样 40 个(分别研磨并过 2mm、0.25mm 和 0.15mm 尼龙筛),不同粒径水平样品的波长-反射率曲线特征如图 1 所示。P1 粒径土样在 350～2500nm 区间的光谱反射率变化范围明显大于 P2(图 1(b))和 P3 粒径样品(图 1(a)),这表明颗粒较细土壤的光谱反射稳定性高于较粗土壤;各粒径水平样品光谱反射率取平均后(图 1(d)),P3 样品在 350～2500nm 区间的光谱反射率明显高于 P2 和 P1 样品(图 1(c)),但 P1 样品的光谱反射率略大于 P2 样品。这说明土壤光谱反射率高低与土壤颗粒粒径大小有关,即颗粒较小土样的光谱反射率较高,但并非颗粒越小光谱反射率就越高,可能存在一个阈值。

3.2　SOM 含量与土壤光谱反射率

P1、P2、P3 土壤的光谱反射率(Raw)与 SOM 含量之间的相关系数变化趋势一致(图 2(a)),反射率与 SOM 含量呈极显著负相关,这说明土壤光谱反射率随着 SOM 含量的增加而降低。在 350～600nm 区间,不同颗粒土壤的光谱反射-SOM 含量相关性差异不明显,但在 600nm 左右达到最大负相关;在 600～2500nm 区间,反射率-SOM 含量相关性随波长增加而逐步减弱,但随着土壤颗粒粒径的增加而降低,即 $|r_{P3}|>|r_{P2}|>|r_{P1}|$,这表明土壤研磨可以提高土壤 600～2500nm 区间光谱反射率,对 SOM 含量变化的响应。

从图 2(a)、(b)、(d)可知,原始光谱-SOM 含量相关性差异不大的光谱区间(如 350～600nm),Msc、Cr 和 Fd 光谱预处理的作用不明显;而在相关性差异较大的 600～2500nm 区间,光谱预处理可以明显地缩小不同粒径土样光谱反射-SOM 含量相关性差异,如在 1250nm(Cr 预处理)、2200～2250nm

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

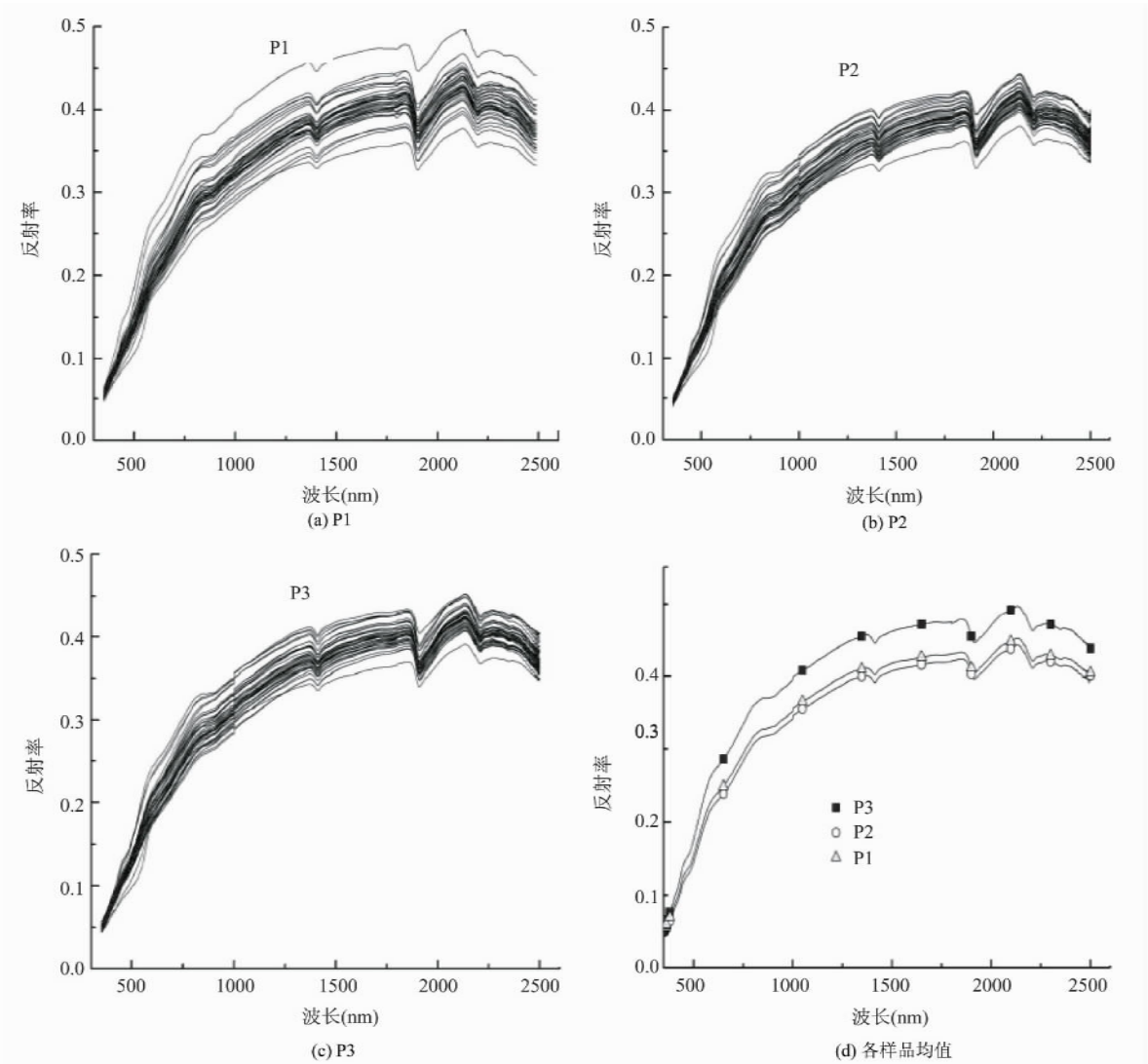

图 1 P1、P2 和 P3 土样的原始反射光谱对比

Fig. 1 Reflectance spectra of soil at the particle sizes of P1, P2 and P3

(Cr 与 Fd 预处理)、1400nm(Fd 预处理)等波段处，P1 与 P3 样品的光谱反射-SOM 含量相关性相近。此外，与较为平缓的光谱反射-SOM 含量相关性(Raw)曲线相比，经过光谱预处理的反射-SOM 含量相关性曲线出现明显的峰和谷，而且曲线的振幅(最大正负相关系数绝对值之和) 均大于原始光谱。由此可知，Fd、Cr、Msc 等光谱预处理可以显著提高土壤反射光谱对 SOM 含量变化的响应，亦可以降低土壤颗粒大小导致的光谱反射-SOM 含量相关性差异。

3.3 单波段光谱反射率的 SOM 含量预测

根据 3.2 中的 P1、P2 和 P3 土样的光谱反射—SOM 含量线性相关分析，挑选出 Raw 光谱与 Fd、Msc、Cr 等预处理下的最大相关系数(见表 1)。从表 1 可知，与 Raw 光谱反射-SOM 含量最大相关对比，Fd能明显提高 P1、P2 和 P3 3 个颗粒粒径土样反射光谱对 SOM 含量变化的响应，而 Cr 预处理则仅在 P1 粒径水平上提高其响应，但 Msc 作用不明显。因此，就提高土壤反射光谱对 SOM 含量变化响应能力(线性相关系数)而言，Fd 的效果最佳。然而，随着土壤颗粒粒径从 P2 降低到 P3，Fd 光谱反射-SOM 含量变化的最大相关并未随之增加。

在实际应用过程中，准备颗粒较细的土样需要耗费更多的人力与物力，而合适的光谱预处理(Fd)

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

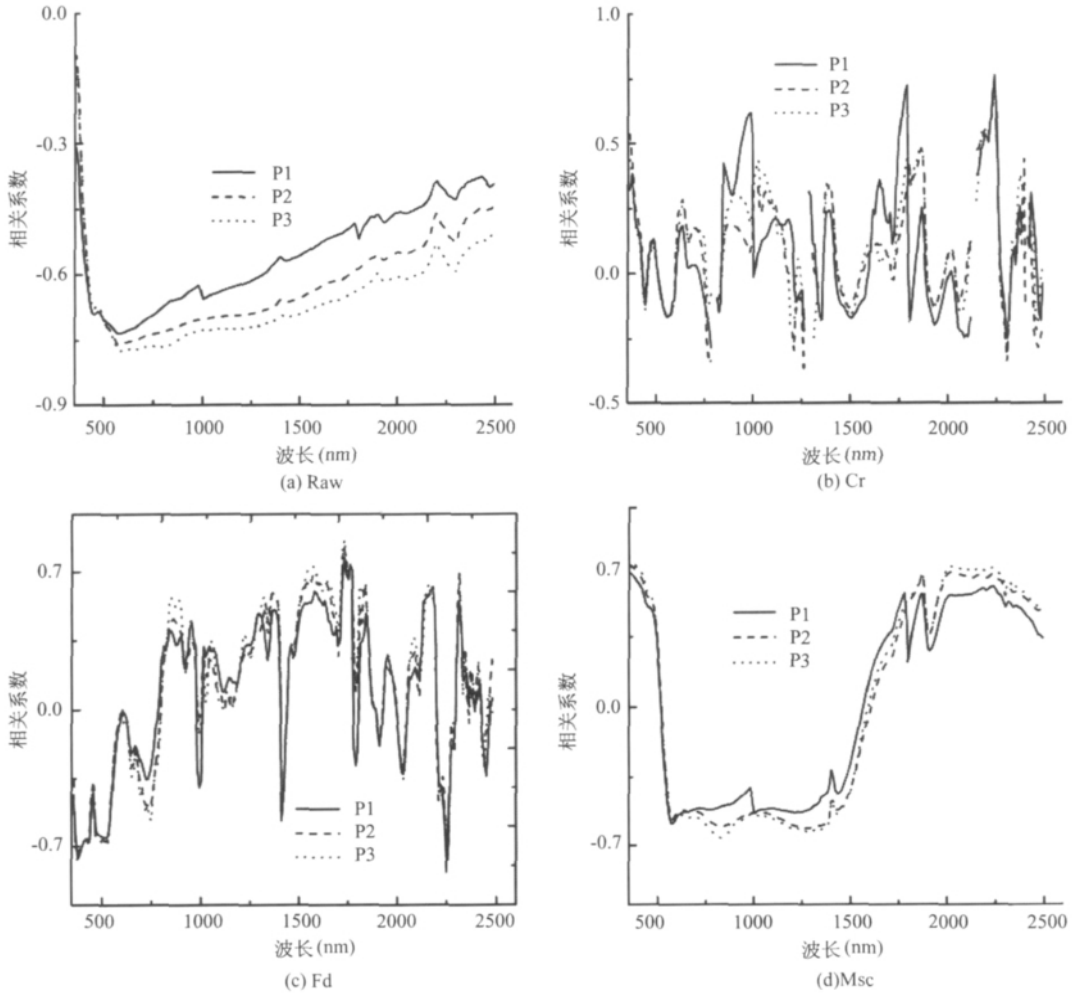

图 2　不同颗粒粒径土壤光谱反射率－SOM 含量相关性

(a)原始光谱;(b)连续统去除;(c)一阶微分;(d)多次散射校正

Fig. 2　Correlation coefficients between SOM contents and reflectance spectra of soils at 3 particle sizes

a: Raw; b: Cr; c: Fd; d: Msc

不但可以显著提高土壤反射光谱对 SOM 含量变化的响应,还可以减弱颗粒粒径对这种响应的影响。因此,选择 2mm 土样经 Fd 预处理光谱在 2 250nm 处的反射率为自变量,SOM 含量为因变量,建立回归方程(表 2),其调整 R^2 达到 0.69($P<0.01$)。

表 1　不同颗粒大小土壤反射光谱与 SOM 含量最大相关波段

Tab. 1　Wavelengths with highest correlation coefficient between SOM content and pretreated reflectance spectra of soils at various particle sizes

颗粒	Raw		Cr		Fd		Msc	
粒径	*r*	波长	*r*	波长	*r*	波长	*r*	波长
P1	0.77	660	0.80	2230	0.82	2250	0.71	2230
P2	0.81	900	0.78	2230	0.83	2250	0.78	1870
P3	0.81	830	0.76	2230	0.83	1730	0.79	2230

表 2　基于 2mm 土样 2 250nm 处光谱反射率(Fd 预处理)的 SOM 含量回归方程

Tab. 2　Linear regression equation on fitting Fd spectral reflectance (2mm) and heavy metal concentration

	波长(nm)	回归方程	校正系数
SOM	2 250	SOM=－2.68×104×Fd2250+43.44	0.69(＊＊)

＊＊ Correlation is significant at the 0.01 level; 95% of confidence bounds

3.4　全波段光谱反射率的 SOM 含量预测

通过对全波段的光谱反射-SOM 含量关系的分析,挑选了最大相关波段进入回归方程,这种单变量回归方程可能造成其他敏感波段的丢失。针对该问题,何挺等人[13]建议采用逐步回归分析,即根

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

据波段对回归分析的贡献逐步加入一些敏感波段，但由于敏感波段之间的“自相关”，建立的回归方程受共线性问题和过度拟合等问题干扰，因此，本文采用 PLSR 进行多元统计分析。同时，为了考察样本容量对 SOM 含量 PLSR 预测结果精度的影响，分别对 146 个(大样本容量)和其中 40 个随机样本(小样本容量)的 SOM 含量进行了建模与预测，并对其精度变化进行对比分析。

在大样本土样中，随机选择 87 个样本作为训练集(Calibration set)，其余 59 个作为验证样本集(Validation set)；针对小样本土样而言，运用逐一法建模和交叉验证的方法，以 P1、P2、P3 土样的光谱信息和 SOM 含量分别为自变量和因变量，建立 PLSR 模型并验证。从模型的精度(表 3)可知：对于相同的光谱预处理而言，小样本容量的 PLSR 模型精度明显低于大样本容量。与大样本容量 Raw 光谱相比，Msc 预处理可以分别提高 P1、P2 和 P3 土样的 PLSR 模型预测精度，其中，P1 土样的 Msc－PLSR 模型预测精度最高(RPD＝3.56、R^2＝0.90、RMSEP＝1.96 g/kg)；大样本 P1、P2 和 P3 土样 PLSR 模型预测精度随着颗粒粒径的增加而增加。就小样本容量土样而言，仅 P3 土样的 Fd－PLSR 模型预测精度较高(RPD＝2.33、R^2＝0.82、RMSEP＝2.31g/kg)；与大样本相反，随着土样颗粒粒径的增加，光谱估测小样本土样 SOM 含量的精度略有降低。

表 3 不同样本量 SOM 含量的 PLSR 预测结果及其精度

Tab. 3 Parameters of PLS regression models for predicting SOM (g/kg) with different samples

光谱预处理方法	Raw			Cr			Fd			Msc		
颗粒粒径 (mm)	2	0.25	0.15	2	0.25	0.15	2	0.25	0.15	2	0.25	0.15
					n＝146							
因子数	16	15	12	17	19	15	9	12	3	16	14	12
决定系数	0.89	0.85	0.76	0.75	0.82	0.58	0.72	0.65	0.46	0.90	0.85	0.78
均方根误	2.08	2.4	3.06	3.10	2.64	4.06	3.29	3.72	4.61	1.96	2.41	2.92
标准差与均方根误比值	3.36	2.91	2.28	2.25	2.64	1.72	2.12	1.88	1.51	3.56	2.90	2.39
					n＝40							
因子数	3	2	9	9	9	11	3	3	10	4	14	6
决定系数	0.61	0.65	0.75	0.65	0.78	0.81	0.67	0.63	0.82	0.6	0.76	0.67
均方根误	3.41	3.22	2.71	3.23	2.54	2.37	3.13	3.3	2.31	3.45	2.66	3.15
标准差与均方根误比值	1.58	1.67	1.98	1.66	2.12	2.27	1.72	1.63	2.33	1.56	2.02	1.71

4 结论与讨论

本文通过对 2mm、0.25mm、0.15mm 颗粒粒径的土样反射光谱曲线分析，发现：(1)土壤颗粒粒径对土壤光谱反射率变化有明显影响，但二者之间并非简单的线性关系，可能存在一个转折点；就 SOM 含量的最大相关波段光谱反射率回归模型的预测精度而言，一阶微分光谱预处理的效果最佳，其 2mm 土样的预测调整系数(R^2)即达到 0.69($P<0.01$)；然而，其预测能力明显低于 2mm 土样 Msc 处理的全波段反射光谱-PLSR 模型(RPD＝3.56、R^2＝0.90、RMSEP＝1.96g/kg)。(2)当样本容量较大时，Msc 预处理作用明显，而颗粒粒径对预测结果精度影响较小，2mm 土样的光谱即可达到较高的预测精度，从而降低了土壤样品准备工作的要求，提高了 SOM 含量光谱估测的工作效率；但在样本容量小的情况下，尽管光谱预处理作用均明显，但是准备更细(0.15mm)颗粒的土样也难以达到大样本容量的预测精度，这就说明样本容量对 SOM 含量的估测精度有着一定的影响，但是多少样本数为限，尚待进一步研究。

在土样样本容量较大时，SOM 含量的 PLSR 预测模型精度要高于较小样本容量，且受颗粒大小影响较小。因此，若条件允许，尽可能多采集些样品，提高样本容量，降低土壤样品准备工作难度，提高 SOM 含量光谱预测的效率；若受条件限制仅能

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

采取有限样本,本文建议准备 0.15mm 的土样进行光谱测量,并采用一阶微分光谱建立 PLSR 预测模型,从而提高 SOM 含量的估测精度。

参考文献:

[1]Shi X Z, Yang R W, Weindorf D C, *et al*. Simulation of organic carbon dynamics at regional scale for paddy soils in China[J]. Climatic Change, (doi:10.1007/s10584-009-9704-1)2009.

[2] 任红艳, 庄大方, 邱冬生,等. 矿区农田土壤砷污染的可见-近红外反射光谱分析研究[J]. 光谱学与光谱分析, 2009, 29 (1): 114-118.

[3] Ben-dor E, Banin A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties [J]. Soil Science Society of America Journal, 1995, 59: 364-372.

[4] 王润生. 高光谱遥感的物质组分和物质成分反演的应用分析[J]. 地球信息科学学报, 2009, 11 (3):261-267.

[5] Wu Y Z, Chen J, Ji J F, *et al*. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils[J]. Soil Science Society American Journal, 2007, 71: 918-926.

[6] Irons J R, Campbell G S, Noman J M, *et al*. Prediction and measurement of soil bidirection reflectance[J]. IEEE Transcactions on Geoscience and Remote Sensing, 1992, 30: 249-260.

[7] Jacquemoud S, Baret F, Hanocq J F. Modelling spectral and bidirectional soil reflectance[J]. Remote Sensing of Environment, 1992, 41: 123-132.

[8] Cierneiewski J, Karnieli A. Virtural surfaces simulating the bidirectional reflectance of semiarid soils[J]. International Journal of Remote Sensing, 2003, 24(7), 1469-1486.

[9] Bänninger D, Flühler H. Modelling light scattering at soil surfaces[J]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(7), 1462-1471.

[10] Cañasveras J C, Barrón V, Campillo M C, *et al*. Estimation of aggregate stability indices in Mediterranean soils by diffuse reflectance spectroscopy[J]. Geoderma. doi:10.1016/j.geoderma,2009.

[11] Nduwamungu C, Ziadi N, Tremblay G,*et al*. Near-infrared reflectance spectroscopy prediction of soil properties: effects of sample cups and preparation[J]. Soil Science Society American Journal, 2009, 73 (6): 1896-1903.

[12] 徐永明, 蔺启忠, 黄秀华,等. 利用可见光/近红外反射光谱估算土壤总氮含量的试验研究[J]. 地理与地理信息科学, 2005, 21(1), 19-22.

[13] 何挺, 王静,林宗坚,等. 土壤有机质光谱特征研究[J]. 武汉大学学报-信息科学版, 2006, 31(11),975-979.

[14] 扶卿华, 倪绍祥, 王世新,等. 土壤盐分含量的遥感反演研究[J]. 农业工程学报, 2007, 23(1), 48-54.

[15] 刘焕军, 张柏, 赵军等. 黑土有机质含量高光谱模型研究[J]. 土壤学报, 2007, 44(1), 27-33.

[16] Summers D, Lewis M, Ostendorf B, *et al*. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties[J]. Ecological Indicators, doi: 10.1016/j.ecolind,2009.

[17] Martin M Z, Labbe N, Andre N, *et al*. Novel multivariate analysis for soil carbon measurements using laser-induced breakdown spectroscopy[J]. Soil Science Scociety American Journal, 2010, 74(1), 87-93.

[18] Savvides A, Corstanje R, Baxter S J, *et al*. The relationship between diffuse spectral reflectance of the soil and its cation exchange capacity is scale-dependent[J]. Geoderma, 2010, 154, 353-358.

[19] 鲁如坤. 土壤农业化学分析方法[M]. 北京: 中国农业科技出版社, 2000.

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved.　http://www.cnki.net

Effects on Application of Spectroscopy in Estimating of Soil Organic Matter Content

HU Wensheng[1], REN Hongyan[1,2], ZHUANG Dafang[1], SHI Xuezheng[3], LIU Shaogui[4], HUANG Yaohuan[1] and YU Xinfang[1]

(1. *State Key Laboratory of Resource and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing* 100101, *China*;
2. *Key Laboratory of Geographic Information Science, Ministry of Education, East China Normal University, Shanghai* 200062, *China*; 3. *Institute of Soil Science, CAS, Nanjing* 210008, *China*;
4. *Yangzhou Soil and Fertilizer Extension Service Center, Yangzhou* 225101, *China*)

Abstract: Reflectance spectra of soil responds to the differences of soil particle sizes, which affects directly the capability of predicting soil organic matter (SOM) content by spectral reflectance. With soil samples of different particle sizes at 2mm, 0.25mm and 0.15mm, reflectance spectra was collected under the condition of simulated sunshine in the laboratory and SOM contents were acquired. Multiple scattering correction (Msc), first derivative (Fd) and continuum removal (Cr) were used as data pretreatment methods to raise the Signal-to-Noise of raw spectra (Raw) before analyzing the correlation between spectral reflectance and SOM contents. Single waveband at 2250nm is selected from Fd-treated reflectance spectra because of its maximal linear correlation coefficient ($r=0.82$, $P<0.01$) responding to SOM contents, and then a linear regression is built with a moderate precision ($R^2=0.69$). Furthermore, partial least square (PLS) was adopted to extract several principal components (PC) of full wavelength from 350 to 2500nm and to build prediction models. Only PLSR model of Msc-treated reflectance spectra of soil samples at 2mm particle size yielded better prediction for SOM content (RPD=3.56, $R^2=0.90$, RMSEP=1.96g/kg). It indicates that particle size of soil sample poses obvious effects on soil reflectance spectra. There is a conceivable turning point of particle size because the correlation between particle size and reflectance is not simple linear relation. Capability of predicting SOM content by univariate linear regression model is markedly lower than that of PLSR model. In addition, prediction precision is significantly affected by capacity of soil samples. Hence, satisfying prediction precision of SOM content can be acquired by effective combination of moderate particle size (2mm) and proper spectral pretreatment (Msc).

Key words: soil particle size; reflectance spectra; soil organic matter content; spectral pretreatment.

© 1994-2012 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net