

## 顾及粗糙度的土壤有机碳成像高光谱估测模型

徐璐<sup>1</sup>, 陈奕云<sup>1, 2, 3\*</sup>, 洪永胜<sup>1</sup>, 魏钰<sup>1</sup>, 郭龙<sup>4</sup>, Marc Linderman<sup>5</sup>

1. 武汉大学资源与环境科学学院, 湖北 武汉 430079
2. 自然资源部数字制图与国土信息应用重点实验室, 湖北 武汉 430079
3. 土壤与农业可持续发展国家重点实验室, 江苏 南京 210008
4. 华中农业大学资源与环境学院, 湖北 武汉 430070
5. Geographical and Sustainability Sciences, The University of Iowa, Iowa City, IA 52246, USA

**摘要** 可见近红外非成像光谱分析技术已被广泛用于土壤有机碳(SOC)含量估测,然而该技术的使用受土壤粗糙度的影响,对样本的前处理要求较高,导致模型的实用性受限。针对这一问题,以美国爱荷华州农田土壤为研究对象,使用成像及非成像光谱仪获取土壤样本研磨前后的可见近红外反射光谱,采用去包络线(CR)、吸光度变换(AB)、S-G平滑(SG)、标准正态变换(SNV)、多元散射校正(MSC)5种光谱预处理手段,利用偏最小二乘回归(PLSR)和支持向量回归(SVR)算法构建并对比土壤SOC光谱估算模型,探究利用成像光谱数据估测高粗糙度样本SOC含量的可行性。实验结果表明,使用成像光谱数据能够实现高粗糙度样本的SOC含量估算,而使用非成像光谱数据则无法估算高粗糙度样本的SOC含量;基于成像光谱数据建立的高粗糙度SOC最优PLSR估算模型 $R^2$ 能够达到0.739以及最优SVR估算模型 $R^2$ 为0.712,而基于非成像光谱数据建立的高粗糙度SOC最优PLSR和SVR估算模型 $R^2$ 仅仅分别为0.344和0.311。基于AB、SG、SNV和MSC这4种预处理手段之后的成像光谱数据建立的土壤样本研磨前的PLSR模型性能优于样本研磨之后建立的PLSR模型,而SVR模型性能正好相反。而对于非成像光谱数据来说,土壤样本研磨后建立PLSR和SVR模型精度总是强于样本研磨前建立的模型精度。对于这两种光谱数据和两个估算模型而言,不同的光谱预处理方法提高模型估算精度的能力不同。土壤样本研磨前后,基于成像光谱数据建立的PLSR和SVR模型性能均优于非成像光谱数据所构建的模型。成像光谱技术能够增强高粗糙度土壤样本可见近红外光谱与SOC的相关性,从而提高模型估算精度;能够克服土壤粗糙度的影响;为野外大尺度估测SOC含量提供了新的手段。

**关键词** 成像光谱技术;土壤粗糙度;可见近红外光谱;光谱预处理;土壤有机碳

**中图分类号:** O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)09-2788-07

## 引言

土壤有机碳(soil organic carbon, SOC)是土壤的重要组成部分,在碳循环与全球气候变化、土壤健康与粮食安全等方面起着至关重要的作用<sup>[1-2]</sup>。SOC信息的高效获取对于可持续土壤资源管理与政策制定具有重要意义<sup>[3-4]</sup>。因此,探索快速且精确的SOC含量估测方法成为了一个热点问题。随着遥感技术的发展,可见近红外光谱技术逐渐被用于SOC含量估测研究<sup>[5]</sup>。与传统的化学方法相比,此方法具有快速高

效、低成本、无污染的特点。

已有研究表明,利用非成像的可见近红外光谱技术能够较好地估测SOC含量<sup>[6]</sup>。但是这些研究大多需要对野外采集到的土壤进行风干、研磨、过筛等预处理,以降低土壤粗糙度对土壤属性光谱建模的不利影响。自然状况下的土壤颗粒因团聚作用形成不同直径的土壤团聚体,呈现非定向空间分布,具有一定程度的随机粗糙度<sup>[7]</sup>。这会导致入射光在土壤样本的不同位置发生不同程度的散射、吸收和反射;同时,土壤颗粒大小及颗粒间的阴影部分也会对光谱反射率造成影响。颗粒粒径越大,颗粒间阴影部分面积越大,光谱反

收稿日期: 2021-07-19, 修订日期: 2021-11-21

基金项目: 国家重点研发项目(2018YFD1100801-01)资助

作者简介: 徐璐,女,1996年生,武汉大学资源与环境科学学院硕士研究生 e-mail: xuluwh@whu.edu.cn

\* 通讯作者 e-mail: chenyy@whu.edu.cn

射率越低。这些特性导致高粗糙度的土壤样本表面具有复杂的光谱散射与反射特征。

非成像光谱技术仅能够获取土壤某个点处的反射光谱信息,无法同时获取土壤光谱的空间连续分布信息,且容易受到土壤粗糙度对光谱测量的影响,因此难以成功估测高粗糙度样本的 SOC 含量。土壤样本经过风干、研磨、过筛的预处理后,其颗粒大小与形状相对均匀,从而降低了粗糙度,使得同一个土壤样本的表面各点具有近似一致的光谱反射特征。这为非成像光谱数据实现 SOC 含量的精确估算提供了保障<sup>[8]</sup>。但这一预处理过程需要耗费大量时间、人力、物力,具有较高的成本,时效性差,难以实现大尺度的野外土壤光谱测量及 SOC 含量的实时监测。因此,开展能够克服土壤粗糙度影响的 SOC 光谱反演具有重要的科研应用价值。

成像光谱技术融合了光谱技术和成像技术,能够同时获取土壤样本中每个像素点的光谱信息及其空间分布信息。已有少数研究将成像光谱技术成功应用于低粗糙度样本的 SOC 估算的报道<sup>[9]</sup>。但是成像光谱技术估算高粗糙度样本 SOC 含量的潜力仍有待验证。

因此,以研磨前后不同粗糙度下的土壤样本为研究对象,提取高光谱影像中的土壤光谱信息,基于 5 种光谱预处理策略,利用偏最小二乘回归(PLSR)和支持向量回归

(SVR)算法建立 SOC 估测模型,并与利用非成像光谱建立的 SOC 估测模型性能进行对比,考察成像光谱技术估测高粗糙度下土壤样本 SOC 含量的能力,并分析其估测机理,为土壤有机碳储量实时监测提供新技术。

## 1 实验部分

### 1.1 研究区概况及样本采集

研究区位于美国爱荷华州东南部。该地区海拔位于 237~266 m 之间,坡度为 0°~9.7°,年平均气温为 9.88 °C,年平均降雨量为 903 mm。土壤类型主要为粉质粘壤土、粘壤土和粉砂壤土。土地利用类型以耕地为主。该研究区主要种植玉米、大豆、黑麦等农作物。

2015 年 10 月,在研究区利用网格采样的策略共采集了 60 个土壤样本(图 1)。土壤样本采样间隔为 130 m。每个样本是由五个子样本组成的复合样本(子样本采集于 1 m×1 m 网格的四个顶点及中心点)。采样前,将土壤表面的植被、杂草和其他杂物去除。采样时,将土壤样本按照采样顺序装入带有标签的塑料袋中并密封。采样后,将采集到土壤样本带回实验室进行光谱测量及化学分析。

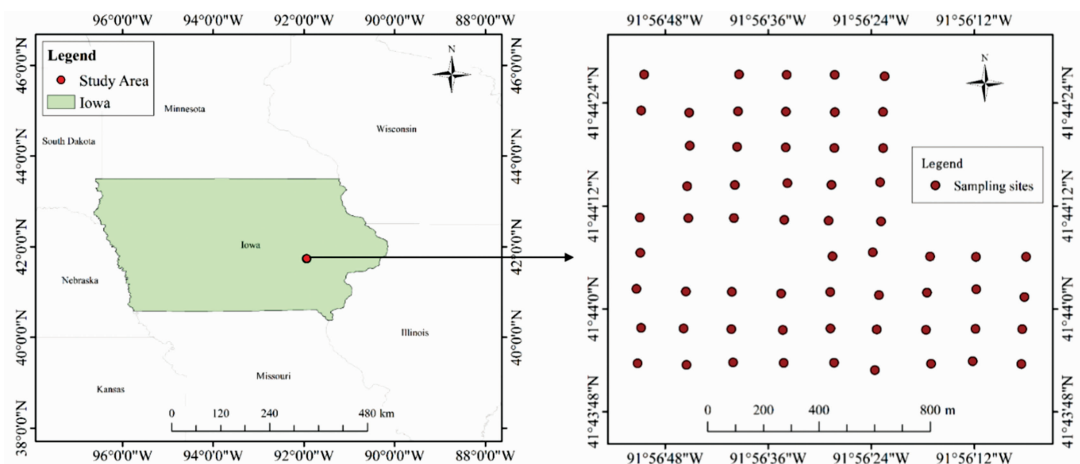


图 1 研究区地理位置及采样点分布情况

Fig. 1 The geographical location of study area and the distribution of soil samples

### 1.2 光谱测量及化学分析

在实验室中将采集的土壤样本进行自然风干。首先测量风干后未研磨的土壤样本的成像光谱与非成像光谱。之后将所有样本进行研磨和过筛(2 mm)预处理,再分别进行成像与非成像光谱测量。最后,将样本继续研磨、过筛(0.15 mm),用于测量 SOC 含量。

采用 Cubert UHD185 型成像光谱仪测量土壤样本研磨前后的成像光谱。该成像光谱仪是由德国 Cubert GmbH 公司研发,具有实时成像的特点,光谱范围为 450~998 nm,光谱分辨率为 4 nm<sup>[10]</sup>。光谱测量在暗室中进行,以卤素灯作为唯一光源,入射光与土壤样本成 45°夹角,与样本相距 30 cm。UHD185 型成像光谱仪被固定于土壤样本正上方 12 cm 处。测量土壤样本光谱前,需要进行白板测量以校正光

谱仪。

采用 ASD FieldSpec3 便携式光谱仪测量土壤样本研磨前后的非成像光谱。整个光谱测量环境与成像光谱测量环境相同,在暗室中进行。卤素灯作为唯一光源,与土壤样本成 45°夹角,并与样本相距 30 cm。光纤放置在样本正上方,距离样本 12 cm。在测量光谱之前利用白板校正光谱仪。每个样本测量 10 次光谱,取平均值作为最终的样本光谱。光谱范围为 350~2 500 nm,光谱分辨率为 1 nm。

采用 Vario MACRO cube 元素分析仪测量 SOC 含量<sup>[11]</sup>。测量前,需要对仪器进行气密性检验。测量时,将用锡箔纸包好的样本按顺序放入元素分析仪中测量 SOC 含量。元素分析仪中氧气和氦气的气瓶减压阀的输出压力分别调至 0.20~0.22 和 0.20~0.23 MPa,燃烧管温度设置为 1 150

℃, 还原管温度设置为 850 ℃。

### 1.3 光谱预处理

测量得到的单个样本成像光谱影像大小为  $50 \times 50$  个像素点(图 2)。为了降低周边测量环境的影响(如盛放土壤样本的培养皿等), 提取红色框中  $14 \times 20$  共 280 个像素点的平均光谱作为土壤样本的成像光谱数据。由于仪器固有特性以及实际测量误差, 导致  $450 \sim 489$  nm 范围内信噪比较低, 故选择  $490 \sim 998$  nm 范围内的成像光谱数据用于建模分析。成像光谱与非成像光谱数据波段范围与光谱分辨率不一致。为了有效对比两者数据估算 SOC 的潜力, 本文选择  $490 \sim 998$  nm 范围内非成像光谱数据进行建模, 并且将非成像光谱反射率重采样为 4 nm。此外, 通过原始成像与非成像光谱进行去包络线(CR)、吸光度变换(AB)、S-G 平滑(SG)、标准正态变换(SNV)、多元散射校正(MSC)5 种光谱预处理后, 对原始光谱及预处理后光谱开展了 SOC 估测建模。

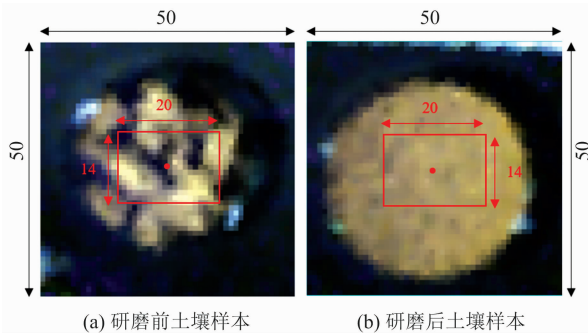


图 2 单个土壤样本研磨前及研磨后成像光谱影像

R: 638 nm, G: 550 nm, B: 470 nm

Fig. 2 RGB images of the single soil sample before and after grinding

R: 638 nm, G: 550 nm, B: 470 nm

### 1.4 模型构建及验证

本研究中采用偏最小二乘回归(PLSR)算法和支持向量回归(SVR)算法建立不同预处理后成像与非成像光谱的 SOC

含量估测模型。PLSR 是一种典型的线性回归算法, 由 Wold 提出, 被广泛应用于光谱分析研究。该方法将原始可见近红外光谱波段投影到一个新的空间, 形成新的特征变量, 利用新的特征变量建立回归模型。它集合了多元线性回归、主成分分析和典型相关分析优点, 能够有效降低高光谱波段的多重共线性和信息冗余。SVR 是利用核函数将原始可见近红外光谱投影至高维线性空间寻找最优估算结果的算法, 对高维数据具有良好的鲁棒性。常见的核函数包括线性核函数、高斯核函数、多项式核函数和径向基核函数。其中基于径向基核函数建立的 SVR 模型为非线性模型, 效果最佳。因此, 选择径向基核函数的 SVR 算法建立 SOC 估算模型。

根据浓度梯度法, 土壤样本按 2:1 比例划分为建模集和验证集。建模集被用来建立 PLSR 和 SVR 估测模型, 而验证集被用来评价模型性能。评价指标包括决定系数( $R^2$ )和均方根误差(RMSE)。 $R^2$  越大, RMSE 越小, 模型性能越好。

## 2 结果与讨论

### 2.1 SOC 数据分布特征

表 1 描述了研究区土壤样本全集、建模集和验证集的 SOC 含量的统计特征。样本全集有 60 个样本, 建模集有 40 个样本, 验证集有 20 个样本。样本全集与建模集中 SOC 含量范围介于  $1.21\% \sim 2.97\%$  之间, 验证集中 SOC 含量最大值为  $2.83\%$ , 最小值为  $1.24\%$ 。样本全集、建模集和验证集的均值皆为  $2.22\%$ , 中位数为  $2.33\%$ 。样本全集、建模集、验证集的标准差分别为 0.43, 0.44 和 0.42; 变异系数分别为 0.19, 0.20 和 0.19, 属于中等变异。三个样本集的偏度系数分别为  $-0.72$ ,  $-0.71$  和  $-0.82$ , 峰度系数分别为  $-0.28$ ,  $-0.29$  和  $0.04$ , 为峰值较小的非正态分布。根据 SOC 的统计特征可知, 建模集和验证集的 SOC 含量特征与样本全集的 SOC 含量统计特征相似。这表明建模集和验证集可以代表样本全集, 能够被用来建立和验证 SOC 估测模型。

表 1 SOC 含量统计性特征描述表

Table 1 The statistical description of SOC content

样本集	样本量/个	土壤有机碳含量/%				标准差	变异系数	偏度系数	峰度系数
		最大值	最小值	平均值	中位数				
样本全集	60	2.97	1.21	2.22	2.33	0.43	0.19	-0.72	-0.28
建模集	40	2.97	1.21	2.22	2.33	0.44	0.20	-0.71	-0.29
验证集	20	2.83	1.24	2.22	2.33	0.42	0.19	-0.82	0.04

### 2.2 光谱特征

图 3 展示了在  $490 \sim 998$  nm 范围内, 研磨前后土壤成像光谱与非成像光谱平均反射率特征。其中, 成像光谱反射率是所有样本中所有像素点的平均光谱反射率。图 3 中  $d, b$  曲线分别代表了土壤样本研磨前后成像光谱的反射率。土壤样本研磨前的成像光谱反射率在  $490 \sim 974$  nm 范围内不断增加, 在  $974 \sim 998$  nm 之间开始下降。土壤样本研磨后的成像光谱反射率的变化趋势与研磨前的光谱反射率变化趋势相

似。土壤样本研磨后的光谱反射率在  $490 \sim 978$  nm 范围内不断增加, 在  $978 \sim 998$  nm 范围内不断降低。在  $490 \sim 998$  nm 范围内, 土壤样本研磨前的光谱反射率总是低于研磨后的光谱反射率, 且随着光谱波长不断增加, 反射率变化越来越大。图 3 中  $c, a$  曲线分别代表了土壤样本研磨前后非成像光谱反射率。从图中可以看出, 非成像光谱反射率在  $490 \sim 998$  nm 范围内不断增加, 无下降趋势。

土壤样本研磨前的成像光谱与非成像光谱反射率总是低

于研磨后的光谱反射率。随着光谱波长增加，土壤样本研磨前后光谱反射率差异逐渐变大，土壤研磨前后非成像光谱反射率差异小于成像光谱反射率的差异。

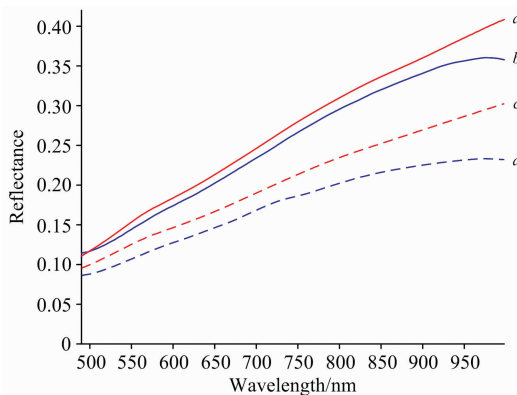


图 3 土壤样本研磨前后成像光谱 (*d, b*) 与非成像光谱 (*c, a*) 平均反射率

Fig. 3 The mean reflectances of imaging spectra (*d, b*) and non-imaging spectra (*c, a*) of the soil samples before and after grinding

2.3 SOC 与光谱相关性分析

图 4 展示了样本研磨前后不同粗糙度条件下土壤成像光谱、非成像光谱与 SOC 的皮尔逊相关系数曲线。图 4 中 *d, b* 曲线分别代表了土壤样本研磨前、后成像光谱与 SOC 的相关系数。研磨前成像光谱反射率与 SOC 的相关系数绝对值介于 0.49~0.70 之间，在 606 nm 处相关系数绝对值最大。研磨后成像光谱反射率与 SOC 的相关系数绝对值范围为 0.75~0.85，在 666 nm 处相关系数绝对值最大。

对于非成像光谱，研磨前光谱与 SOC 的相关系数最大绝对值为 0.56，最小绝对值为 0.38。598 nm 处光谱与 SOC 的

相关性最强。研磨后光谱与 SOC 的相关系数最大值为 0.84，最小值为 0.75。668 nm 处光谱与 SOC 的相关性最强。

总体来说，土壤样本研磨前后成像光谱与非成像光谱与 SOC 的相关性皆为负相关。在 490~998 nm 范围内，土壤样本研磨前，成像光谱与 SOC 的相关性明显强于非成像光谱与 SOC 的相关性；土壤样本研磨后，成像光谱与 SOC 的相关性稍微强于非成像光谱与 SOC 的相关性。

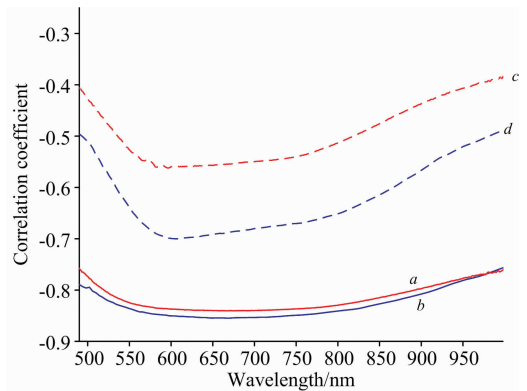


图 4 土壤样本研磨前后成像光谱 (*d, b*)、非成像光谱 (*c, a*) 与 SOC 相关系数图

Fig. 4 The correlation coefficient between imaging spectra (*d, b*)/non-imaging spectra (*c, a*) and SOC for the soil samples before and after grinding

2.4 模型性能对比

表 2 展示了两种粗糙度下的土壤样本(研磨前后)，基于成像光谱与非成像光谱，利用其原始光谱及 5 种预处理光谱，建立的 PLSR 和 SVR 模型性能对比表。

从表中可以看出，土壤样本研磨前后，不同光谱预处理下，基于成像光谱建立的 PLSR 和 SVR 模型性能总是优于基于非成像光谱建立的估测模型。

表 2 土壤样本研磨前后成像光谱与非成像光谱在不同预处理下 PLSR 和 SVR 模型性能表  
Table 2 PLSR and SVR model performances based on imaging spectra/non-imaging spectra of the soil samples before and after grinding

样本预处理	光谱预处理	光谱类型	偏最小二乘回归				支持向量回归			
			建模集		验证集		建模集		验证集	
			R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE	R <sup>2</sup>	RMSE
未研磨样本	None	Imaging	0.788	0.040	0.739	0.045	0.804	0.037	0.712	0.050
		Non-imaging	0.703	0.056	0.344	0.114	0.721	0.053	0.311	0.119
	CR	Imaging	0.797	0.039	0.651	0.060	0.789	0.040	0.705	0.051
		Non-imaging	0.948	0.010	0.196	0.139	0.672	0.062	0.469	0.092
	AB	Imaging	0.785	0.041	0.719	0.049	0.808	0.036	0.699	0.052
		Non-imaging	0.735	0.050	0.411	0.102	0.788	0.040	0.281	0.125
	SG	Imaging	0.788	0.040	0.739	0.045	0.804	0.037	0.712	0.050
		Non-imaging	0.699	0.057	0.347	0.113	0.717	0.054	0.299	0.122
	SNV	Imaging	0.795	0.039	0.719	0.049	0.802	0.038	0.735	0.046
		Non-imaging	0.703	0.056	0.396	0.105	0.719	0.053	0.410	0.102
	MSC	Imaging	0.794	0.039	0.718	0.049	0.797	0.039	0.724	0.048
		Non-imaging	0.703	0.056	0.398	0.104	0.700	0.057	0.394	0.105

续表 2

研磨样本	None	Imaging	0.743	0.049	0.648	0.061	0.848	0.029	0.716	0.049
		Non-imaging	0.774	0.043	0.645	0.062	0.799	0.038	0.706	0.051
	CR	Imaging	0.733	0.051	0.666	0.058	0.830	0.032	0.724	0.048
		Non-imaging	0.858	0.027	0.564	0.076	0.836	0.031	0.708	0.051
	AB	Imaging	0.794	0.039	0.709	0.051	0.849	0.029	0.731	0.047
		Non-imaging	0.782	0.041	0.693	0.053	0.751	0.047	0.614	0.067
	SG	Imaging	0.743	0.049	0.648	0.061	0.848	0.029	0.716	0.049
		Non-imaging	0.774	0.043	0.645	0.062	0.798	0.038	0.706	0.051
	SNV	Imaging	0.727	0.052	0.626	0.065	0.783	0.041	0.738	0.045
		Non-imaging	0.768	0.044	0.657	0.059	0.788	0.040	0.693	0.053
MSC	Imaging	0.722	0.053	0.622	0.066	0.782	0.041	0.736	0.046	
	Non-imaging	0.767	0.044	0.655	0.060	0.788	0.040	0.692	0.053	

土壤样本研磨前, 基于成像光谱建立的 PLSR 模型的  $R_p^2$  介于 0.651~0.739, 其中基于原始成像光谱建立的 PLSR 模型的估测精度最高 ( $R_p^2=0.739$ ,  $RMSE_p=0.045\%$ ), 其他光谱预处理算法并未提高 PLSR 模型预测精度。基于成像光谱建立的 SVR 模型的  $R_p^2$  介于 0.699~0.735, 其中利用 SNV 和 MSC 的算法可以提高 SVR 模型的预测精度。而基于非成像光谱建立的 PLSR 和 SVR 模型的  $R_p^2$  分别介于 0.196~0.411 和 0.281~0.469, 模型性能整体较差, 不能被用于 SOC 估测。

土壤样本研磨后, 对于成像光谱, 仅有 AB 预处理算法能够提高 PLSR 模型估测精度, 并成为最优的 PLSR 模型,  $R_p^2=0.709$ ,  $RMSE_p=0.051\%$ ; 5 种预处理方法均能提高 SVR 模型估测精度, 其中 SNV 预处理算法效果最为显著。基于 SNV 算法建立的 SVR 模型的  $R_p^2$  达到 0.738。对于非成像光谱, 样本研磨能够有效提高 PLSR 和 SVR 模型估测精度, 其中最优模型为基于 SNV 预处理算法建立的 SVR 估测模型,  $R_p^2$  达到 0.708, 使得模型从无法估测 SOC 转变为精确估测 SOC。

总的来说, 对于成像光谱, 土壤样本研磨前后建立的估测模型皆能被用于估测 SOC 含量。土壤样本研磨前建立 PLSR 模型性能总是优于样本研磨之后建立的 PLSR 模型, 而 SVR 模型性能正好相反。对于非成像光谱, 仅有土壤样本研磨后建立的 PLSR 和 SVR 模型能够成功估测 SOC 含量, 但估测精度总是低于成像光谱建立的估测模型。

### 2.5 粗糙度对光谱反射率的影响

土壤在不同粗糙度下的光谱反射率特征不同。整体表现为粗糙度大, 反射率低, 反之, 粗糙度小, 反射率高。土壤样本研磨前粗糙度高, 研磨后粗糙度低。本研究探讨了应用不同光谱仪测量时, 土壤样本在处于研磨前后的这两种粗糙度状态下的反射率特征。

成像与非成像光谱仪测量得到的研磨前的土壤光谱平均反射率总是小于研磨后的土壤光谱平均反射率。这是由于研磨前, 土壤颗粒在团聚作用的驱动下形成不同大小的团聚体, 具有非定向空间分布, 导致较高的粗糙度。测量光谱时, 光散射现象明显, 导致较低的反射率。

土壤样本研磨前后, 成像光谱平均反射率均小于非成像光谱平均反射率。这是因为成像光谱可以获得土壤样本阴

影部分与非阴影部分像素点的反射率。阴影部分土壤反射率低, 故而成像光谱平均反射率低于非成像光谱平均反射率。研磨后土壤样本阴影部分面积减少, 因此研磨后的成像光谱反射率略微低于非成像光谱反射率。这说明成像光谱数据能够更加真实地反映土壤粗糙度对土壤光谱反射率的影响。

### 2.6 粗糙度对 SOC 光谱响应波段的影响

不同的土壤粗糙度下, 可见近红外范围内的光谱波段在估测 SOC 含量时具有不同的重要性。变量投影重要性 (variable importance in projection, VIP) 指示每个光谱波段在估测 SOC 含量中的重要性。图 5 展示了样本研磨前后这两种粗糙度状态下土壤成像光谱与非成像光谱的 VIP 图。其中  $VIP > 1$ , 光谱波段被视为重要波段, 表明在估测 SOC 含量时起关键作用。采用成像光谱技术估测 SOC 含量时, 样本研磨前重要波段为 614~766 与 902~998 nm。其中, 766~902 nm 范围内的光谱波段在样本研磨前为非重要波段, 在样本研磨后为重要波段。利用非成像光谱建模时, 767~871 nm 范围内的光谱波段在样本研磨前为非重要波段, 在样本研磨后为重要波段。样本研磨后重要波段为 686~998 nm。样本研磨前后, 成像光谱与非成像光谱波段重要性曲线相似, 再一次证明了成像光谱技术被用于估测 SOC 含量的可行性。

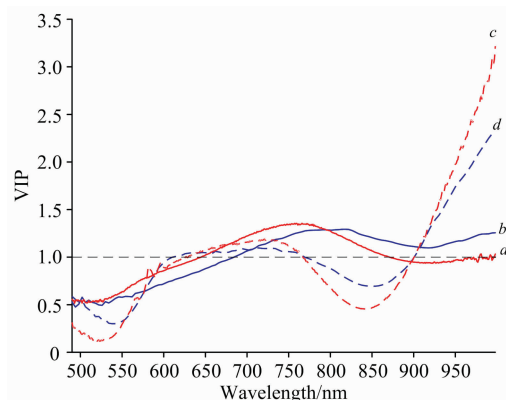


图 5 土壤样本研磨前后成像光谱 (d, b) 与非成像光谱 (c, a) VIP 图

Fig. 5 The VIP of imaging spectra (d, b)/non-imaging spectra (c, a) of the soil samples before and after grinding

## 2.7 成像光谱技术的优势

将土壤样本研磨前后的成像光谱与非成像光谱数据用于SOC含量的估测, 并对比了不同数据源及光谱预处理对PLSR和SVR模型性能的影响。结果显示成像光谱数据在土壤样本研磨前后具有较为优秀的表现。而使用非成像光谱数据所构建的模型仅能用于研磨后土壤样本SOC含量估测, 这是由于研磨后, 降低了粗糙度, 增强了非成像光谱与SOC的相关性, 从而提高了PLSR模型的估测能力。土壤样本研磨前后, 成像光谱与SOC的相关性较强, 说明成像光谱数据能够克服土壤粗糙度的影响, 可以被用来建立良好的SOC估测模型。

此外, 利用成像光谱建立的估测模型有潜力被推广应用于卫星遥感数据。国内现有的高光谱遥感卫星有高分五号、珠海一号等。高分五号获取的高光谱卫星遥感影像在可见近红外光谱范围内共有150个波段, 光谱分辨率为4.28 nm<sup>[4]</sup>。珠海一号获取的高光谱卫星遥感影像在400~1 000 nm范围内共有32个波段, 光谱分辨率为3~8 nm<sup>[12]</sup>。本研究获取的成像光谱数据与高分五号、珠海一号高光谱遥感影像具有

相似的波段数量与光谱分辨率。可以进一步探索成像光谱与卫星遥感影像光谱间的传递性, 从而将近地成像光谱SOC估测模型迁移应用于卫星遥感影像中, 为大尺度的SOC含量估测提供新型手段。

## 3 结 论

以美国爱荷华州土壤样本为对象, 在实验室内测量了不同粗糙度下土壤样本的成像及非成像光谱反射率, 基于5种光谱预处理算法, 利用PLSR和SVR方法建立了SOC含量估测模型, 探究了利用成像光谱技术估测高粗糙度的土壤样本的SOC含量的潜力。结果表明成像光谱技术能够克服土壤粗糙度的影响, 成功预测高粗糙度样本的SOC含量。成像光谱技术能够同时获取土壤空间维度与光谱维度信息, 增强了光谱与SOC的相关性, 因此具有良好的估测能力。成像光谱技术作为一个新型光谱测量技术, 为野外光谱测量和大尺度SOC含量估测提供了技术支持。

## References

- [ 1 ] Zhang J, Zhang F. *Frontiers of Agricultural Science and Engineering*, 2020, 7(3): 233.
- [ 2 ] ZHANG Tong-yao, HU Yue-ming, REN Xiang-ning, et al(张童瑶, 胡月明, 任向宁, 等). *Journal of Agricultural Resources and Environment(农业资源与环境学报)*, 2020, 37(6): 805.
- [ 3 ] ZHAO Yong-cun, XU Sheng-xiang, WANG Mei-yan, et al(赵永存, 徐胜祥, 王美艳, 等). *Bulletin of Chinese Academy of Sciences(中国科学院院刊)*, 2018, 33(2): 191.
- [ 4 ] LIU Huan-jun, BAO Yi-lin, MENG Xiang-tian, et al(刘焕军, 鲍依临, 孟祥添, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2020, 36(12): 90.
- [ 5 ] ZHANG Gan-lin, SHI Zhou, ZHU A-xing, et al(张甘霖, 史舟, 朱阿兴, 等). *Acta Pedologica Sinica(土壤学报)*, 2020, 57(5): 1060.
- [ 6 ] GUO Jia-xin, ZHU Qing, ZHAO Xiao-min, et al(国佳欣, 朱青, 赵小敏, 等). *Chinese Journal of Applied Ecology(应用生态学报)*, 2020, 31(3): 863.
- [ 7 ] ZHANG Hui-juan, SUN Yu-rui(张慧娟, 孙宇瑞). *Transactions of the Chinese Society for Agricultural Machinery(农业机械学报)*, 2010, 41(3): 33.
- [ 8 ] Xu L, Hong Y, Wei Y, et al. *Remote Sensing*, 2020, 12(20): 3394.
- [ 9 ] JIAO Cai-xia, ZHENG Guang-hui, XIE Xian-li, et al(焦彩霞, 郑光辉, 解宪丽, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2020, 40(10): 3277.
- [ 10 ] WANG Wei-dong, CHANG Qing-rui, WANG Yu-na(王伟东, 常庆瑞, 王玉娜). *Arid Zone Research(干旱区研究)*, 2020, 37(5): 1362.
- [ 11 ] GAO Cui-ping, LI Yan, LIU Mei-ying, et al(高翠萍, 李岩, 刘美英, 等). *Journal of Northern Agriculture(北方农业学报)*, 2017, 45(1): 76.
- [ 12 ] LI Xian-yi, FAN Hai-sheng, PAN Shen-lin, et al(李先怡, 范海生, 潘申林, 等). *Satellite Application(卫星应用)*, 2019, (8): 12.

# Estimation of Soil Organic Carbon Content by Imaging Spectroscopy With Soil Roughness

XU Lu<sup>1</sup>, CHEN Yi-yun<sup>1,2,3\*</sup>, HONG Yong-sheng<sup>1</sup>, WEI Yu<sup>1</sup>, GUO Long<sup>4</sup>, Marc Linderman<sup>5</sup>

1. School of Resource and Environmental Science, Wuhan University, Wuhan 430079, China

2. Key Laboratory of Digital Mapping and Land Information Application, Ministry of Natural Resources, Wuhan 430079, China

3. State Key Laboratory of Soil and Sustainable Agriculture, Nanjing 210008, China

4. College of Resource and Environment, Huazhong Agricultural University, Wuhan 430070, China

5. Geographical and Sustainability Sciences, The University of Iowa, Iowa City, IA 52246, USA

**Abstract** Visible and near-infrared (VIS-NIR) non-imaging spectroscopy has been widely applied to estimate soil organic carbon (SOC) content. Due to the high demand for soil sample pretreatments, VIS-NIR non-imaging spectroscopy easily suffers from soil roughness in practical application. This study explored the potential of imaging spectroscopy to estimate SOC content with high soil roughness. With soil samples collected in Iowa State, United States, imaging spectra were utilized to measure the VIS-NIR spectra of soil samples with and without ground. With five spectral pre-processing including continuum removed (CR), absorbance transformation (AB), S-G smoothing (SG), standard normal variate (SNV), and multiplicative scatter correction (MSC), partial least squares regression (PLSR) and support vector regression (SVR) were used to build estimation models to analyze the potential of imaging spectra. Non-imaging spectra were also applied to build PLSR and SVR models as a comparison. Results demonstrated that imaging spectra could achieve SOC content estimation for soil samples with high roughness, but non-imaging spectra could not successfully estimate that. The best PLSR and SVR model developed by imaging spectra could reach 0.739 and 0.712 of  $R^2$  for SOC content estimation of soil samples with high roughness, while that established by non-imaging spectra could achieve 0.344 and 0.311 of  $R^2$ . Based on the imaging spectra after the four pre-processing methods of AB, SG, SNV, and MSC, the performance of the PLSR model established before soil sample grinding was better than that of the PLSR model established after soil sample grinding, while the performance of the SVR model was just the opposite. For non-imaging spectra, the accuracies of PLSR and SVR models established after soil samples grinding were always better than that of models established before soil samples grinding. For these two spectral data and the two estimation models, different spectral pre-processing methods had different abilities to improve the estimation accuracy of the model. The performance of imaging spectroscopy outperformed non-imaging spectra before or after being ground soil samples. Imaging spectra could enhance the correlation coefficient between VIS-NIR spectra and SOC for soil samples with high roughness, there by improving PLSR model's performance. Our findings provide a new way to estimate SOC content on large-scale yield because imaging spectra could overcome the influence of soil roughness.

**Keywords** Imaging spectroscopy; Soil roughness; Visible and near-infrared spectra; Spectra pre-processing; Soil organic carbon

(Received Jul. 19, 2021; accepted Nov. 21, 2021)

\* Corresponding author