

DOI: 10.11766/trxb201602250646

苏北滨海土壤碳酸钙含量反射光谱估算研究*

洪长桥 郑光辉[†] 陈昌春

(南京信息工程大学地理与遥感学院, 南京 210044)

摘要 土壤属性的快速、精确测定是实现现代精细农业的基础。本研究分析了江苏省北部滨海土壤的属性特征以及碳酸钙的可见-近红外反射光谱特征, 探讨利用可见-近红外光谱估算滨海土壤碳酸钙含量的可行性, 比较不同光谱反射率数据集、不同预处理方法以及不同建模方法定量反演的优劣。结果表明: (1) 苏北滨海土壤有机质含量较低、碳酸钙含量较高, 其光谱曲线在 2 340 nm 处有明显的碳酸钙吸收特征; (2) 滨海土壤碳酸钙含量与土壤的可见-近红外波段反射率呈正相关, 且碳酸钙含量高低对于土壤的近红外波段反射率的影响高于可见光波段; (3) 可见-近红外反射光谱可用于估算滨海土壤碳酸钙含量。就建模结果而言, 381~2 459 nm 波段反射光谱数据集、log(1/R) 预处理、偏最小二乘回归三者结合的效果比较理想。

关键词 碳酸钙; 有机质; 反射光谱; 滨海土壤

中图分类号 S151.9 **文献标识码** A

碳酸钙 (CaCO₃) 是土壤中碳酸盐的主要成分, 对土壤酸碱度、土壤胶体性状及养分状况等土壤一系列物理、化学和生物性质具有重要影响^[1-2], 例如: 高含量的碳酸钙通过降低土壤磷的有效性进而影响植物生长^[3]、土壤碳酸钙影响土壤对重金属的吸附容量等^[4]。传统的土壤碳酸盐测定主要基于化学方法, 具有测定成本高、速度慢、破坏性等不足, 难以满足现代精细农业的需要。土壤反射光谱分析技术具有成本低、分析速度快、无破坏、无危险、可同时对多种成分进行估算等特点, 为土壤学研究提供了新的方法和手段^[5]。土壤反射光谱是对各种土壤属性的综合反映, 其中包含了碳酸钙含量等信息, 这是光谱分析技术被应用于土壤碳酸钙含量估算的理论基础。但不同土壤成分反射光谱信息之间相互影响^[6], 通常是基于各土壤属性的光谱特征、采用化学计量学方法来反演各

土壤属性值。当前用来估算土壤属性的化学计量学方法较多, 如偏最小二乘回归方法 (PLSR)、主成分回归方法 (PCR) 以及多元线性回归方法 (MLR) 等。

近年来, 学界在碳酸钙的反射光谱特征及土壤碳酸钙反射光谱估算方面进行了有益的探索。在碳酸钙反射光谱特征方面, Gaffey^[7] 研究发现碳酸盐矿物由于碳酸根的振动过程, 在可见-近红外 (350~2 550 nm) 光谱区域有 7 个较强的吸收特征; 张婷婷^[8] 研究认为土壤碳酸盐在波段 2 230~2 270 nm 波段具有独特的光谱特征, 2 300 nm 处则是其在近红外区域的特征吸收波段; Stenberg 等^[9] 研究发现土壤碳酸盐在 2 335 nm 处具有较强的吸收特征, 在 1 870 nm、1 990 nm 和 2 160 nm 处则吸收较弱, 且易受水的干扰; 修连存等^[10] 以化学成分为碳酸钙的方解石为研究对象,

* 国家自然科学基金项目 (41201215)、国家级大学生实践创新训练计划 (201310300041Z) 和南京信息工程大学“优秀本科毕业论文 (设计) 支持计划”项目联合资助 Supported by the National Natural Science Foundation of China (No. 41201215), the National Undergraduate Training Programs for Innovation and Entrepreneurship of China (No. 201310300041Z), and the Support Plan of Outstanding Undergraduate Thesis from Nanjing University of Information Science & Technology, China

[†] 通讯作者 Corresponding author, E-mail: zgh@nuist.edu.cn

作者简介: 洪长桥 (1993—), 男, 湖南衡阳人, 硕士研究生, 主要从事土壤遥感研究。E-mail: hcquinst@126.com

收稿日期: 2016-02-25; 收到修改稿日期: 2016-06-10; 优先数字出版日期 (www.cnki.net): 2016-07-05

发现其在近红外波段的4个典型吸收特征峰分别大致位于1 874 nm、1 996 nm、2 154 nm以及2 338 nm处；王冬等^[11]分析了105 °C高温干燥处理前后纯碳酸钙的反射光谱特征，发现在近红外反射光谱区域峰形基本不变，该研究表明碳酸钙光谱特征与结晶水、其他吸附水关系不大，其吸收波段分别位于1 878 nm、1 996 nm以及2 339 nm处。碳酸钙的可见-近红外反射光谱特征可作为其含量反演的理论依据。在土壤碳酸钙反射光谱估算方面，Ben-Dor和Banin^[12]基于近红外反射光谱（1 000 ~ 2 500 nm），结合多元线性回归方法估算了以色列干旱、半干旱地区的土壤碳酸钙含量，并指出光谱反射率及其一阶微分、光谱反射率倒数的对数及其一阶微分和二阶微分5种预处理方法中，基于光谱反射率的估算精度最好（验证 R^2 为0.69）；Chang等^[13]基于近红外反射光谱，运用主成分回归方法成功估算了土壤碳酸钙含量；Gomez等^[14]利用可见-近红外反射光谱（400 ~ 2 500 nm）估算了法国南部裸土碳酸钙含量，结果表明偏最小二乘回归方法估算精度高于连续统去除法；Volkan Bilgili等^[15]利用350 ~ 2 500 nm波段的反射光谱，结合偏最小二乘法成功估算了土耳其北部半干旱地区土壤碳酸钙含量。以上研究表明，反射光谱分析技术在估算部分地区一些土壤类型的碳酸钙含量是可行的。

然而不同区域的土壤类型多样、成分复杂、空间异质性高，导致不同土壤类型之间的高光谱特征以及估算模型差异明显^[16]。目前对于不同类型土壤碳酸钙光谱特征及其估算研究尚处于探索阶段。滨海地区滩涂土壤（简称滨海土壤）由沿海地区的盐渍淤泥发育而成，是我国重要的后备土壤资源。近年来由于耕地占补平衡政策，通常利用苏北滩涂开垦的耕地来补充苏南一些建设用地占用耕地，具有显著的区域特征^[17]。加强滨海土壤反射光谱特征及土壤属性高精度反演研究对于沿海地区精细农业的发展以及土壤学理论的进一步研究具有一定的积极意义。学界在一些类型土壤碳酸钙反射光谱估算方面取得了不少的研究成果，然而有关滨海土壤中的碳酸钙含量的反射光谱估算尚少见报道。此外，受土壤异质性影响，学者们得到的不同光谱预处理方法的结论有所差异，不同区域同一土壤属性以及同一区域不同土壤属性的最佳预处理方法会有所不同^[18]；在建立土壤碳酸钙含量估算模型时，不同的建模方法之间的估算精度存在一定的差异，

尚缺少系统比较以及评价；再者，不同的光谱波段所包含的土壤碳酸钙反射光谱信息不尽相同，基于不同波段的反射光谱数据集对于土壤属性估算的影响尚有待进一步研究。

鉴于上述原因，本研究拟针对以下几个问题进行探索：（1）研究利用可见-近红外反射光谱估算我国苏北地区滨海土壤碳酸钙含量的可行性；（2）评估不同波段反射光谱数据集以及不同预处理方法对于滨海土壤碳酸钙含量估算精度的影响；（3）比较偏最小二乘回归、主成分回归以及多元逐步回归三种建模方法的滨海土壤碳酸钙含量估算能力。

1 材料与方 法

1.1 样本采集与分析

东台市位于苏北沿海地区，属北亚热带暖湿季风气候，土壤类型主要为范公堤以西的水稻土和以东的滨海盐土。本研究主要采集滨海盐土，于2013年6月采用GPS定位，在东台市范公堤以东地区采集17个剖面样点和8个0 ~ 5 cm表土样。剖面采样深度分别为0 ~ 5 cm、5 ~ 10 cm、10 ~ 20 cm、20 ~ 30 cm、30 ~ 40 cm、40 ~ 60 cm、60 ~ 80 cm、80 ~ 100 cm。由于有两个剖面分别遗失1个0 ~ 5 cm和80 ~ 100 cm土样，共得到142个样品。土样在室内经风干、研磨，过100目筛，每个样品分成两份，一份用于化学分析，一份用于光谱测量。土壤中碳酸钙含量采用中和滴定法^[19]测定，有机质（soil organic matter, SOM）含量采用重铬酸钾法^[20]测定。

1.2 光谱测量

采用美国分析光谱仪器公司（ASD）FieldSpec 3便携式光谱仪，波长范围350 ~ 2 500 nm，采样间隔1.4 nm（350 ~ 1 000 nm）和2 nm（1 000 ~ 2 500 nm），光谱分辨率为3 nm（350 ~ 1 000 nm）和10 nm（1 000 ~ 2 500 nm），重采样间隔1 nm，输出波段数为2 151。光源为一盏50 W卤素灯，光照角度为15°，光源距离为30 cm，容器半径为2.5 cm，采用5°视场探头角，探头距离为15 cm，探测半径为0.655 cm。测量步骤为：去除暗电流、利用标准参照板获取绝对反射率、容器土样表面刮平、在样本跟随自动旋转平台匀速转动360°的同时获取光谱曲线20条，算术平均后得到该土样的光谱

曲线。

1.3 光谱预处理

为消除不同探测元件造成的误差,利用ASD自带软件ViewSpecPro对光谱曲线进行断点校正。为消除由光谱仪波段间能量响应差别造成的毛刺噪音,利用Unscrambler 9.7对光谱曲线进行Savitzky-Golay平滑处理(2次多项式,11个点)。最终得到355~2 495 nm波段的光谱反射率,将其作为原始光谱数据。

1.4 建模方法

偏最小二乘回归方法(partial least squares regression, PLSR)是Wold等^[21]于20世纪80年代提出的新型多元统计数据分析方法。该方法可以在自变量多重相关的条件下,有效构造出对系统解释最强的子空间,并且能在样本数少于自变量数的情况下进行回归建模,大大提高了模型的应用能力、精度以及可靠性^[22]。

主成分回归方法^[13](principal component regression, PCR)的原理是利用主成分分析将所有自变量组合成多个相互无关的新变量,选取尽可能少的组合变量代表原有自变量信息,再与因变量建立回归模型。

多元逐步回归方法^[23](stepwise multiple linear regression, SMLR)的基本思路是将全部变量按重要性逐步选入回归方程,期间利用 F 统计量选择或剔除自变量,最终建立自变量与因变量之间的回归关系。

模型精度通过决定系数 R^2 、均方根误差RMSE、样本值标准差与估算均方根误差比RPD^[24]进行评价,建模、交叉验证和预测RMSE越小, R^2 、RPD越大,表明建模、交叉验证、预测精度越高,模型越好。

2 结果

2.1 滨海土壤属性特征

对东台市滨海土壤142个土壤样本的碳酸钙、有机质含量进行统计分析,结果如表1所示。样本碳酸钙含量均在均值 ± 2 倍标准差以内,总体上碳酸钙含量高,有机质含量低(平均值为6.95 g kg⁻¹),区域性特征较强。由于土壤中有机质含量大于20 g kg⁻¹可能会掩盖其他光谱特征,小于20 g kg⁻¹时对土壤光谱影响相对较小,难以掩盖碳酸钙、氧

化铁等因素对土壤光谱的影响^[25]。本研究区样品中仅个别样品有机质含量高于20 g kg⁻¹,因而,在该研究区土壤有机质对碳酸钙含量与光谱特征关系的研究影响相对较小。

表1 土壤碳酸钙、有机质含量统计

土壤属性	最大值	最小值	平均值	标准差	变异系数
Soil properties	Max	Min	Mean	Std	CV
	(g kg ⁻¹)	(%)			
碳酸钙CaCO ₃	63.67	0.90	32.54	16.27	50.00
有机质SOM	33.38	1.39	6.95	5.97	85.90

2.2 土壤碳酸钙对土壤光谱特征的影响

由滨海土壤部分土样的可见-近红外波段光谱反射率曲线(图1)可以看出:各土样的光谱曲线形态大致相同,可见光波段的光谱反射率低于近红外波段,吸收特征位置基本相同,深度略有差异;除具有大部分土壤类型中所共有的较明显的1 400 nm、1 900 nm、2 200 nm处来自OH、H₂O、Al-OH等的特征吸收以外,滨海土壤还具有较明显的2 340 nm处主要来自CaCO₃的特征吸收;随着有机质含量的增加,土壤在可见-近红外波段的整体光谱反射率会有相应的减少,且其减少程度在可见光区域表现更明显,这与前人研究结果^[26]一致;在有机质含量较低且相近的情况下,碳酸钙含量越高,其在2 340 nm处的吸收特征越明显;随着碳酸钙含量的增加,土壤在可见-近红外波段的整体光谱反射率也有相应的提高,其中,在近红外波段的提高程度要高于可见光波段。

基于142个土样,对土壤碳酸钙、有机质含量与可见-近红外波段反射率进行了相关分析,结果如图2所示。在可见-近红外波段,土壤有机质含量与光谱反射率呈负相关,且在可见光波段的相关性高于近红外波段;而土壤碳酸钙含量与光谱反射率则呈正相关,且在近红外波段的相关性高于可见光波段;在可见光波段及少部分近红外波段,有机质含量与光谱反射率的相关性大于碳酸钙,然而在大部分近红外波段,碳酸钙含量与光谱反射率的相关性则大于有机质。这在一定程度上反映了滨海土壤反射光谱曲线的变化特征,同时也说明了土壤中碳酸钙的含量与土壤可见-近红外波段光谱反射率的关系密切。

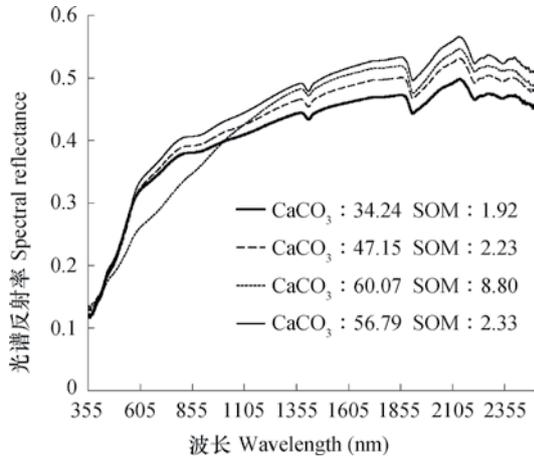
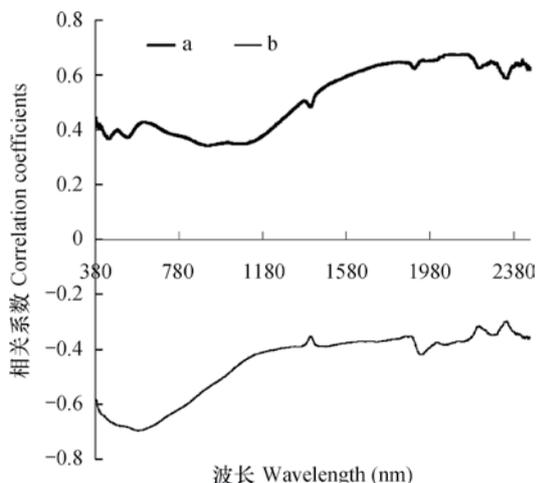


图1 部分样本的反射光谱曲线
(CaCO_3 和SOM含量单位: g kg^{-1})

Fig. 1 Soil reflection spectral curves of some soil samples



注: a、b 分别代表 CaCO_3 、SOM 与光谱反射率的相关系数
Note: a and b stand for correlation coefficients of spectral reflectance and CaCO_3 and correlation coefficients of spectral reflectance and SOM, respectively

图2 CaCO_3 、SOM 与光谱反射率相关系数

Fig. 2 Correlation coefficients of spectral reflectance vs CaCO_3 and SOM

2.3 基于不同波段反射光谱数据集的模型构建

本研究随机选取112个样品作为建模集, 其余30个样品作为验证集; 交叉验证方法采用Full Cross-Validation。估算 R^2 大于0.5、RPD大于1.4时估算精度较好, RPD大于2.5时表明模型或估算结果极好^[13]。

不同学者利用不同波段反射率数据建立的同一土壤属性估算模型结果不尽相同^[12, 14-15], 然而鲜有学者对基于不同波段反射光谱数据集建立的估

算模型的效果进行比较, 针对某土壤属性甄选较优的建模波段。本研究为获取典型波段的反射光谱数据集, 参考了已有的代表性研究。Volkan Bilgili等^[15]研究土壤碳酸钙含量估算时采用的是可见-近红外全波段数据(350~2 500 nm), 因而利用已采集的355~2 495 nm波段代表全波段数据, 作为本研究数据集之一; 由于碳酸钙在近红外波段有明显的吸收特征, 因此也选择781~2 495 nm波段反射率数据作为研究的数据集之一; Ben-Dor和Banin^[12]认为用于碳酸钙建模的最佳波段并不是其吸收特征波段, 较高含量的碳酸钙一般导致土壤在可见-近红外反射光谱波段具有更高的反射率, 因此也选择355~780 nm波段反射率作为本研究数据集之一; Shepherd和Walsh^[27]研究表明, 350~380 nm和2 460~2 500 nm波段的信噪比较低, 因此本研究也选择去除二者后的381~2 459 nm波段的光谱反射率作为数据集之一; 基于滨海土壤可见-近红外光谱反射率与碳酸钙含量的相关关系, 还选择了相关系数较高的波段即1 393 nm、1 418 nm、1 475 nm、2 186 nm、2 195 nm、2 314 nm以及2 365 nm等作为敏感波段。基于选择的355~2 495 nm、355~780 nm、781~2 495 nm、381~2 459 nm以及7个敏感波段等5个数据集, 结合偏最小二乘回归方法, 分别建立回归模型, 结果如表2所示。

根据表2可以看出, 利用5种反射光谱数据集建模得到的结果差异较大。基于355~2 495 nm、781~2 495 nm、381~2 459 nm波段建立的模型效果较好(R^2 在0.70以上, $\text{RPD} > 1.80$); 而利用7个敏感波段结合偏最小二乘回归方法建立的模型, 尽管建模以及交叉验证精度较好($R^2 > 0.70$), 但模型稳定性较差, 其估算 R^2 可能大于建模 R^2 ; 利用355~780 nm波段建立的模型效果最差, 而基于381~2 459 nm波段建立的模型效果最好, 其建模、交叉验证以及估算的 R^2 均大于0.85, RPD为2.66。

2.4 基于不同预处理方法的模型构建

为探讨不同预处理方法对模型精度的影响, 本研究基于经过Savitzky-Golay平滑的381~2 459 nm波段光谱反射率(Reflectance, R), 分别进行一阶微分(FD)、多元散射校正(MSC)、标准归一化(SNV)、对数的倒数($1/\log R$)、倒数的对数($\log 1/R$)等预处理, 结合偏最小二乘回归方法

表2 不同数据集的建模、交叉验证、估算结果表

Table 2 Calibration, cross validation and prediction based on different data sets

光谱数据集 Spectroscopy Data sets	建模 Calibration			交叉验证 Cross validation		估算 Prediction		
	潜变量数 Number of PC	均方根误差 RMSE (g kg ⁻¹)	R ²	均方根误差 RMSE (g kg ⁻¹)	R ²	均方根误差 RMSE (g kg ⁻¹)	R ²	RPD
	355 ~ 2 495 nm	8	4.26	0.92	5.29	0.88	7.47	0.83
355 ~ 780 nm	5	6.78	0.79	7.38	0.76	13.80	0.42	1.32
781 ~ 2 495 nm	7	4.51	0.91	5.56	0.86	9.92	0.70	1.83
381 ~ 2 459 nm	8	4.29	0.92	5.21	0.88	6.84	0.86	2.66
A	5	7.23	0.76	7.74	0.73	8.17	0.80	2.22

注: A代表1 393、1 418、1 475、2 186、2 195、2 314、2 365 nm 7个敏感波段 Note: A stands for 7 sensitive spectral wave

表3 不同预处理方法的建模、交叉验证、估算结果表

Table 3 Calibration, cross validation and prediction using different pre-processing methods

预处理 Pre-processing	建模 Calibration			交叉验证 Cross validation		估算 Prediction		
	潜变量数 Number of PC	均方根误差 RMSE (g kg ⁻¹)	R ²	均方根误差 RMSE (g kg ⁻¹)	R ²	均方根误差 RMSE (g kg ⁻¹)	R ²	RPD
	R	8	4.29	0.92	5.21	0.88	6.84	0.86
MSC	7	5.06	0.88	6.09	0.84	7.04	0.85	2.58
FD	4	4.79	0.90	6.16	0.83	13.62	0.44	1.33
SNV	8	4.87	0.89	6.05	0.84	7.01	0.85	2.59
log (1/R)	10	4.05	0.93	5.22	0.88	6.55	0.87	2.77
1/log R	9	4.23	0.92	5.29	0.88	7.13	0.85	2.55

注: R、MSC、FD、SNV、1/R、log R分别代表反射率及其多元散射校正、一阶微分、标准归一化、倒数、对数等预处理方法 Note: R, MSC, FD, SNV, 1/R and log R stand for reflectance and its pre-processing methods "multiplicative signal correction, first derivative, standard normal variate, reciprocal and logarithm", respectively

分别建立模型,结果如表3所示。

从表3可以看出,除FD外,其他4种预处理方法建立的模型效果较好,建模、交叉验证和估算 $R^2 > 0.83$, RPD > 2.50 ;比较以上预处理方法,log (1/R)预处理的模型结果最好,建模、交叉验证、估算 R^2 分别为0.93、0.88、0.87,估算RPD为2.77, RMSE较小,介于4~6.6 g kg⁻¹,模型精度最好。利用原始光谱反射率R建立的模型,精度优于MSC、FD、SNV以及1/log R预处理方法,仅次于log (1/R)方法。

2.5 基于不同回归方法的模型构建

基于经过Savitzky-Golay平滑的381~2 459 nm

波段光谱反射率数据集以及log (1/R)预处理方法,比较偏最小二乘回归、主成分回归以及多元逐步回归三种方法的滨海土壤碳酸钙含量的估算能力。偏最小二乘回归、主成分回归在Unscrambler 9.7软件中进行。多元逐步回归在SPSS 20软件中进行,基本步骤是:基于log (1/R)预处理后的光谱数据与土壤碳酸钙含量进行相关分析,选择相关性较强($r > 0.6$)的1 812 nm、1 826 nm、2 141 nm、2 245 nm、2 282 nm、2 402 nm、2 424 nm以及2 452 nm等波段,利用多元逐步回归方法建立模型(通过显著性水平为0.001的F值检验):

$$Y = -12.313 - 5.535.931X_{2.141\text{nm}} + 5.258.436X_{2.245\text{nm}} \quad (1)$$

表4 不同回归方法的建模、交叉验证、估算结果表

Table 4 Calibration, cross validation and prediction using different regression methods

回归方法 Regression method	建模 Calibration		交叉验证 Cross Validation		估算 Prediction			
	潜变量数 Number of PC	均方根误差 RMSE (g kg ⁻¹)	R ²	均方根误差 RMSE (g kg ⁻¹)	R ²	均方根误差 RMSE (g kg ⁻¹)	R ²	RPD
	PLSR	10	4.05	0.93	5.22	0.88	6.55	0.87
PCR	6	5.85	0.85	6.41	0.82	7.85	0.81	2.31
SMLR		8.40	0.68			10.46	0.67	1.74

注：PLSR、PCR、SMLR分别代表偏最小二乘回归、主成分回归、多元逐步线性回归 Note: PLSR, PCR and SMLR stand for partial least squares regression, principal component regression and stepwise multiple linear regression, respectively

式中， Y 表示土壤碳酸钙含量估算值； $X_{2\ 141\text{nm}}$ 和 $X_{2\ 245\text{nm}}$ 分别表示2 141 nm、2 245 nm处的土壤光谱反射率。

建模结果如表4所示。根据建模结果可以看出：三种回归方法建模 R^2 介于0.68~0.93之间，RMSE较小，介于4.05~8.40 g kg⁻¹之间，建模效果均较好，其中PLSR最优，PCR其次；对于交叉验证RMSE，PLSR优于PCR，由于多元逐步回归一般利用显著性水平来进行验证，因此此处未进行交叉验证；三种回归方法的估算RMSE较小，估算 R^2 介于0.67~0.87，模型均较好，尤其是PLSR方法，其RPD > 2.5，精度较优。综合各项评价指标可知，PLSR、PCR和SMLR方法建模、估算结果等均较好，其中PLSR方法表现最优，其RMSE介于4.05~6.55 g kg⁻¹之间， R^2 介于0.87~0.93之间；此外，SMLR仅用两个波段建立了回归模型，并达到了较好的效果。

3 讨论

不同类型土壤的属性特征存在较明显的差异，与其他土壤类型^[28]不同的是，苏北滨海土壤中有有机质含量较少，而碳酸钙含量则较多。尽管许多土壤类型光谱曲线大致趋势相似，但是光谱曲线局部差异明显。苏北滨海土壤由于有机质含量较少、碳酸钙含量较多的原因，在2 340 nm处具有较明显的吸收特征，便于发现碳酸钙等其他属性对土壤光谱反射率的影响特征。

Volkan Bilgili等^[15]研究认为，土耳其北部半干旱地区土壤中的碳酸盐与近红外波段1 890 nm、

2 350 nm和2 360 nm具有显著的相关关系；而本研究中滨海土壤碳酸钙含量与1 393 nm、1 418 nm、1 475 nm、2 186 nm、2 195 nm、2 314 nm以及2 365 nm等7个波段相关性较强（ $r > 0.6$ ）。但这些波段均不是碳酸盐的特征吸收波段。研究还发现，随着碳酸钙含量的增加，滨海土壤在可见-近红外波段的整体光谱反射率也有相应的提高，这与Ben-Dor和Banin^[12]的研究结论相似；其中，在近红外区域的提高程度要高于可见光区域。

Ben-Dor和Banin^[12]在进行以色列干旱、半干旱地区土壤碳酸钙含量反演时，选入模型的波段是1 647.3 nm和1 482.5 nm；而本研究进行苏北滨海土壤碳酸钙含量反演时选用的则是2 141 nm和2 245 nm。二者选用的波段并不一致，即不同类型土壤的同一属性反射光谱估算结果存在差异。已有研究表明，被选用到回归模型用来估算碳酸钙的许多光谱特征仍无法解释^[14]，学界普遍认为，在对一些土壤属性与光谱反射率之间的物理机制不甚明晰的情况下，基于经验的模型同样可作为估算这些土壤属性成分的较好选择^[15]。

基于滨海土壤碳酸钙建模结果可知，建模均方根误差小于交叉验证均方根误差，而交叉验证均方根误差小于估算均方根误差。建模 R^2 大于交叉验证 R^2 ，另外，除基于7个敏感波段数据集、MSC和SNV预处理建立的模型以外，交叉验证 R^2 大于估算 R^2 。根据基于不同光谱数据集建模的结果可知，利用381~2 459 nm波段数据集的建模效果优于355~2 495 nm，可能原因是较多的波段相应地增加了其受其他因素或成分影响的程度或可能性；而利用781~2 495 nm波段数据集的建模效果

低于381~2 459 nm波段,可能主要存在两方面的原因:一是碳酸钙对于光谱反射率的影响在近红外区受到其他分子和基团的干扰,二是同近红外波段一样,碳酸钙在可见光波段提高了土壤光谱反射率,建模时考虑这部分因素可提高估算的精度;381~780 nm波段数据集建模效果低于781~2 495 nm波段,原因可能与可见光波段反射率受碳酸钙含量变化的影响程度低于近红外波段有关;应用基于7个敏感波段的PLSR建模时,由于较少的光谱波段所包含的碳酸钙信息可能不足以及可能存在过度拟合现象等原因,导致模型的稳定性较差。

就本研究中不同预处理方法的效果而言,一阶微分FD在选用的5种预处理方法中表现最差,其他预处理方法效果都比较好,其中 $\log(1/R)$ 预处理最优,其次则是基于原始光谱反射率 R 的模型;而Ben-Dor和Banin^[12]对以色列干旱、半干旱地区土壤碳酸钙进行反演,结果表明在5种预处理方法中,基于原始光谱反射率 R 建立的模型效果最好。这表明土壤的高度异质性导致在对不同研究区的土壤属性反演时最佳预处理方法存在差异^[18]。

从不同建模方法的效果来看,本研究选用的3种回归方法均能较好地估算苏北滨海土壤碳酸钙含量,特别是PLSR方法达到了极好的精度水平,其次是PCR方法;SMLR方法比前两者稍差,但其所需波段较少,且大大节省了模型运行时间。产生此结果的原因可能是,SMLR方法虽考虑了少数主要变量,但忽略了其他次要变量以及可能存在的共线性问题等对模型估算结果的影响,最终导致其反演精度相对较低;PCR方法消除了自变量多重共线性问题,采用交叉验证方式,使得训练误差较低,从而提高了模型的反演精度;PLSR方法因为综合了典型相关分析、多元线性回归和PCR方法的优势,在土壤碳酸钙反射光谱估算方面表现出了较好的适用性。

由于不同区域滨海土壤形成过程存在差异^[29-31],未来有待进一步扩大研究区范围,比较不同时期、不同母质及不同区域的滨海土壤的属性以及光谱反射率等特征对于估算土壤碳酸钙含量的影响,为滨海地区农业生产及土壤改良等提供参考。

4 结 论

苏北滨海土壤有机质含量较低、碳酸钙含量较

高,在2 340 nm处具有较明显的碳酸钙吸收特征。土壤碳酸钙含量与土壤的可见-近红外波段光谱反射率呈较明显的正相关关系,土壤在可见-近红外波段的光谱反射率随碳酸钙含量增加而相应提高,且在近红外波段的提高程度高于可见光波段。利用可见-近红外反射光谱估算滨海土壤碳酸钙含量是可行的,就建模效果而言,381~2 459 nm波段在5个数据集中表现最佳; $\log(1/R)$ 预处理在光谱反射率 R 和5种预处理方法中表现最佳;PLSR、SMLR、PCR三种方法建立的模型效果均较好,其中,PLSR建立的模型表现最佳。

参 考 文 献

- [1] 陈留美, 张甘霖. 滨海沉积物发育的水稻土时间序列母质均一性判定与特性演变. 土壤学报, 2009, 46 (5): 753—763
Chen L M, Zhang G L. Parent material uniformity and evolution of soil characteristics of a paddy soil chronosequence derived from marine sediments (In Chinese). Acta Pedologica Sinica, 2009, 46 (5): 753—763
- [2] 鲁如坤. 土壤农业化学分析方法. 北京: 中国农业科学技术出版社, 2000
Lu R K. Analytical methods for soil and agro-chemistry (In Chinese). Beijing: China Agricultural Science and Technology Press, 2000
- [3] 赵雪莱, 何兴东, 薛萃萃, 等. 土壤碳酸钙/有效磷化学计量特征对油蒿群落植物密度的影响. 科学通报, 2012, 57 (1): 80—87
Zhao X L, He X D, Xue P P, et al. Effects of soil stoichiometry of the CaCO_3 /available phosphorus ratio on plant density in *Artemisia ordosica* communities (In Chinese). Chinese Science Bulletin, 2012, 57 (1): 80—87
- [4] 李虎, 贺婧, 刘冲, 等. 农田灰钙土中有机质和碳酸钙对Zn吸附-解吸行为的影响. 中国环境科学, 2015, 35 (6): 1786—1793
Li H, He J, Liu C, et al. Effect of organic matter and calcium carbonate on behaviors of zinc adsorption-desorption on/from Ningxia farmland sierozem soils (In Chinese). China Environmental Science, 2015, 35 (6): 1786—1793
- [5] Nocita M, Stevens A, van Wesemael B, et al. Soil spectroscopy: An alternative to wet chemistry for soil monitoring. Advances in Agronomy, 2015, 132: 139—159
- [6] 彭杰, 向红英, 周清, 等. 土壤氧化铁的高光谱响应研究. 光谱学与光谱分析, 2013, 33 (2): 502—506

- Peng J, Xiang H Y, Zhou Q, et al. Influence of soil iron oxide on VNIR diffuse reflectance spectroscopy (In Chinese). *Spectroscopy and Spectral Analysis*, 2013, 33 (2): 502—506
- [7] Gaffey S J. Spectral reflectance of carbonate minerals in visible and near infrared (0.35~2.55 μm): Anhydrous carbonate minerals. *Journal of Geophysical Research*, 1987, 92: 1429—1440
- [8] 张婷婷. 基于PLS模型的农业土壤成分高光谱遥感反演研究. 长春: 吉林大学, 2010
Zhang T T. Partial least squares modeling of hyperspectral remote sensing for mapping agricultural soil properties (In Chinese). Changchun: Jilin University, 2010
- [9] Stenberg B, Viscarra Rossel R A, Mouazen A M, et al. Visible and near infrared spectroscopy in soil science. *Advances in Agronomy*, 2010, 107: 163—215
- [10] 修连存, 郑志忠, 俞正奎, 等. 近红外光谱分析技术在蚀变矿物鉴定中的应用. *地质学报*, 2007, 81 (11): 1584—1590
Xiu L C, Zheng Z Z, Yu Z K, et al. Mineral analysis technology application with near infrared spectroscopy in identifying alteration mineral (In Chinese). *Acta Geologica Sinica*, 2007, 81 (11): 1584—1590
- [11] 王冬, 马智宏, 潘立刚, 等. 近红外光谱法定量测定小麦粉中的石灰类添加物的研究. *光谱学与光谱分析*, 2013, 33 (1): 69—73
Wang D, Ma Z H, Pan L G, et al. Research on the quantitative determination of lime in wheat flour by near-infrared spectroscopy (In Chinese). *Spectroscopy and Spectral Analysis*, 2013, 33 (1): 69—73
- [12] Ben-Dor E, Banin A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Science Society of America Journal*, 1995, 59 (2): 364—372
- [13] Chang C W, Laird D A, Mausbach M J. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal*, 2001, 65 (2): 480—490
- [14] Gomez C, Lagacherie P, Coulouma G. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. *Geoderma*, 2008, 148: 141—148
- [15] Volkan Bilgili A, van Es H M, Akbas F, et al. Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey. *Journal of Arid Environments*, 2010, 74 (2): 229—238
- [16] 史舟, 王乾龙, 彭杰, 等. 中国主要土壤高光谱反射特性分类与有机质光谱预测模型. *中国科学: 地球科学*, 2014, 44 (5): 978—988
Shi Z, Wang Q L, Peng J, et al. Development of a national VNIR soil-spectral library for soil classification and prediction of organic matter concentrations (In Chinese). *Science China: Earth Science*, 2014, 44 (5): 978—988
- [17] 林晨, 吴绍华, 周生路. 滨海盐土遥感监测的发展趋势. *土壤学报*, 2011, 48 (5): 1072—1079
Lin C, Wu S H, Zhou S L. Development of remote sensing monitoring of coastal saline soil (In Chinese). *Acta Pedologica Sinica*, 2011, 48 (5): 1072—1079
- [18] 郑光辉, 周生路, 吴绍华. 土壤砷含量高光谱估算模型研究. *光谱学与光谱分析*, 2011, 31 (1): 173—176
Zheng G H, Zhou S L, Wu S H. Prediction of As in soil with reflectance spectroscopy (In Chinese). *Spectroscopy and Spectral Analysis*, 2011, 31 (1): 173—176
- [19] 刘光崧. 土壤理化分析与剖面描述. 北京: 中国标准出版社, 1997
Liu G S. Soil physical and chemical analysis & description of soil profiles (In Chinese). Beijing: Standards Press of China, 1997
- [20] 中国科学院南京土壤研究所. 土壤理化分析. 上海: 上海科学技术出版社, 1978: 81
Institute of Soil Science, Chinese Academy of Sciences. Soil physical and chemical analysis (In Chinese). Shanghai: Shanghai Science and Technology Press, 1978: 81
- [21] Wold S, Martens H, Wold H. The multivariate calibration problem in chemistry solved by the PLS Method //Matrix pencils. Berlin, Heidelberg: Springer, 1983: 286—293
- [22] 王惠文, 刘强, 屠永平. 偏最小二乘回归模型内涵分析方法研究. *北京航空航天大学学报*, 2000, 26 (4): 473—476
Wang H W, Liu Q, Tu Y P. Identification of optimal subspace from PLS regression (In Chinese). *Journal of Beijing University of Aeronautics and Astronautics*, 2000, 26 (4): 473—476
- [23] 姚艳敏, 魏娜, 唐鹏钦, 等. 黑土土壤水分高光谱特征及反演模型. *农业工程学报*, 2011, 27 (8): 95—100
Yao Y M, Wei N, Tang P Q, et al. Hyper-spectral characteristics and modeling of black soil moisture content (In Chinese). *Transactions of the CSAE*, 2011, 27 (8): 95—100
- [24] Fearn T. Assessing calibrations: SEP, RPD, RER and R^2 . *NIR News*, 2002, 13 (6): 12—14

- [25] Baumgardner M F, Silva L F, Biehl L L, et al. Spectral reflectance properties of soils. *Advances in Agronomy*, 1985, 38: 1—44
- [26] 彭杰, 周清, 张杨珠, 等. 有机质对土壤光谱特性的影响研究. *土壤学报*, 2013, 50 (3): 517—524
Peng J, Zhou Q, Zhang Y Z, et al. Effect of soil organic matter on spectral characteristics of soil (In Chinese). *Acta Pedologica Sinica*, 2013, 50 (3): 517—524
- [27] Shepherd K D, Walsh M G. Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal*, 2002, 66 (3): 988—998
- [28] 纪文君, 史舟, 周清, 等. 几种不同类型土壤的VIS-NIR光谱特性及有机质响应波段. *红外与毫米波学报*, 2012, 31 (3): 277—282
Ji W J, Shi Z, Zhou Q, et al. VIS-NIR reflectance spectroscopy of the organic matter in several types of soil (In Chinese). *Journal of Infrared and Millimeter Waves*, 2012, 31 (3): 277—282
- [29] 鞠兵, 吴克宁, 李玲, 等. 河南省典型土系的特定土层特征与分类研究. *土壤学报*, 2016, 53 (1): 48—57
Ju B, Wu K N, Li L, et al. Characteristic horizons and classification of soil series typical of Henan Province (In Chinese). *Acta Pedologica Sinica*, 2016, 53 (1): 48—57
- [30] 于淑会, 白璐, 刘慧涛, 等. 河北滨海平原土壤盐分空间异质性与淋排均质化作用研究. *土壤*, 2014, 46 (6): 1135—1141
Yu S H, Bai L, Liu H T, et al. Study on homogenizing effect to heterogeneous soil salinity in coastal plain of Hebei Province (In Chinese). *Soils*, 2014, 46 (6): 1135—1141
- [31] 韩光中, 王德彩, 谢贤健. 土壤退化时间序列的构建及其在我国土壤退化研究中的意义. *土壤*, 2015, 47 (6): 1015—1020
Han G Z, Wang D C, Xie X J. Soil degradation chronosequence and their significance in soil degradation study of China (In Chinese). *Soils*, 2015, 47 (6): 1015—1020

Estimation of CaCO₃ Content in Coastal Soil of North Jiangsu with Reflectance Spectroscopy

HONG Changqiao ZHENG Guanghui[†] CHEN Changchun

(School of Geography and Remote Sensing, Nanjing University of Information Science & Technology, Nanjing 210044, China)

Abstract 【Objective】Rapid and accurate measurement of soil properties is fundamental to modern precision agriculture. CaCO₃ is a major component of the carbonate in soil and has some important effects on a series of physical, chemical and biological properties of soil, such as soil pH, characteristics of soil colloids, soil nutrition and soil heavy metals absorption capacity. The traditional method for measuring soil CaCO₃ content is mainly based on chemical analysis, which is often rather costly, time-consuming and destructive. Therefore, the method is far from efficient to meet the requirement of modern precision agriculture. The technology of soil reflection spectroscopy can be used to make up the shortages of the traditional method, and provide a new approach for the study of pedology. This paper is oriented to explore feasibility of using visible-near infrared reflection spectra to estimate CaCO₃ content in coastal soil in North Jiangsu, to evaluate impacts of reflection spectra data sets, pre-processing methods on accuracy of the estimation of CaCO₃ content in coastal soil and to compare different modeling methods on estimating CaCO₃ content in coastal soil. 【Method】A total of 142 coastal soil samples were collected from North Jiangsu and analyzed for soil spectra with a portable FieldSpec 3. Soil properties of the soil samples were also determined with chemical methods and characterized, and the characteristics of visible-near infrared reflection spectra of the CaCO₃ in the soil were analyzed. On such a basis, reflection spectra data sets of various spectral bands (e.g. seven sensitive bands, 355 ~ 2 495 nm, 355 ~ 780 nm, 781 ~ 2 495 nm and 381 ~ 2 459 nm), reflectance (R) and its pre-processing methods (e.g. multiplicative signal correction (MSC)), first

derivative (FD), standard normal variate (SNV), $\log(1/R)$ and $1/\log R$, and three regressions methods (e.g. partial least squares regression method (PLSR), principal component regression method (PCR) and multiple stepwise linear regression method (SMLR)) were selected to perform modeling analysis. In addition, coefficient of determination (R^2), root mean square error (RMSE) and relative percent deviation (RPD) were used to evaluate accuracy of the models. **【Result】** Results shows that the coastal soil is quite low in soil organic matter (SOM), but very high in CaCO_3 content; that reflectance of the soil in visible-near infrared bands is positively related to soil CaCO_3 content. Moreover, the response of soil spectral reflectance to change of soil CaCO_3 content is stronger in near infrared bands than in visible bands. The estimation ($R^2 = 0.86$, $\text{RMSE} = 6.48$, $\text{RPD} = 2.66$) using the model based on spectral reflectance data sets with bands from 381 nm to 2 459 nm is higher in accuracy than those using the other four data sets; the estimations using the pre-processing methods, R , $\log(1/R)$, $1/\log R$, MSC and SNV are good ($0.85 \leq R^2 \leq 0.87$, $6.55 \leq \text{RMSE} \leq 7.13$, $2.55 \leq \text{RPD} \leq 2.77$); and the estimation using the model based on SMLR has R^2 being 0.67, while the estimations using the model based on PLSR and PCR are better than the former ($R^2 \geq 0.81$). **【Conclusion】** It could be concluded that it is feasible to use visible-near infrared reflection spectra in estimating CaCO_3 content in the coastal soil of North Jiangsu; what kind of reflection spectral data sets, pre-processing methods, and regression methods should be used has a certain impact on precision of the estimation of CaCO_3 content in the soil; and the application of the technology of soil reflection spectroscopy will have a bright prospect in determining properties of coastal soil.

Key words CaCO_3 ; Soil organic matter; Reflection spectroscopy; Coastal soil

(责任编辑：卢 萍)