

DOI: 10.11766/trxb201707170240

# 基于可见-近红外反射光谱的土壤碳酸钙含量与反演效果关系研究\*

林 卡<sup>1, 2</sup> 李德成<sup>1†</sup> 刘 峰<sup>1</sup> 张甘霖<sup>1, 2</sup>

(1 土壤与农业可持续发展国家重点实验室(中国科学院南京土壤研究所), 南京 210008)

(2 中国科学院大学, 北京 100049)

**摘 要** 土壤可见-近红外反射光谱中包含了大量的土壤属性信息, 研究人员根据土壤属性信息在光谱上的特征, 对土壤属性进行定量反演。是否属性值越高, 反演精度越高? 目前对于属性含量与反演效果的定量关系尚不清楚。采集了我国西北地区黑河流域 69 个代表性干旱土剖面(292 个发生层土样), 以气量法测定其碳酸钙含量, 使用 Cary 5000 分光光度计测定其可见-近红外光谱反射率, 以样本量和离散度(变异系数)作为数据集划分标准, 分别建立了 11 个相同样本量子集(A)和 5 个相近离散度子集(B), 应用偏最小二乘回归(PLSR)算法对各子集进行土壤碳酸钙含量反演, 以此探究碳酸钙含量与反演效果的定量关系。结果表明, 碳酸钙可增加可见-近红外波段的光谱反射率, 但利用可见近红外光谱反演土壤碳酸钙含量, 其反演效果与碳酸钙含量关系不显著。

**关键词** 可见-近红外反射光谱; 碳酸钙含量; 反演效果

**中图分类号** S153.4 **文献标识码** A

土壤光谱反射特性是土壤基本性质之一, 与土壤理化性质有密切关系。不同物质的电子过程及振动过程使其在光谱曲线上表现出特征吸收带, 这些特征吸收带或其组合是进行物质识别及定量估算的依据<sup>[1]</sup>。换句话说, 土壤组成物质的多样性以及土壤中每种组成物质独特的光谱特性均会在光谱曲线上体现, 因此土壤光谱反射特征能综合反映土壤理化性状指标<sup>[2]</sup>。基于此, 能够利用土壤光谱从某种程度上反演出土壤的属性, 如有机质、铁氧化物、机械组成、水分含量、碳酸钙含量以及质地和母质等<sup>[3-4]</sup>。研究人员定性地认为, 土壤属性含量越高, 其在光谱反射曲线上的特征越明显, 利用光谱反演其含量的精度越高, 但对于土壤属性含量

的大小与反演精度的定量关系尚不清楚。

为探索光谱属性反演效果与属性含量的定量关系, 本研究选择碳酸钙作为目标属性, 一方面由于碳酸钙具有重要的土壤学意义<sup>[5-9]</sup>, 另一方面因为目前利用可见-近红外光谱预测碳酸钙的研究较少, 且精度不高<sup>[10-11]</sup>。本文针对典型研究区的碳酸钙含量数据分布特征, 根据样本数和离散度(变异系数=标准差/均值)两个指标, 将样本数据集分为若干不同碳酸钙含量的子集, 基于偏最小二乘回归(PLSR)算法对每个子集进行建模分析, 探究光谱反演碳酸钙含量的效果与碳酸钙含量的定量关系, 并分析其影响因素, 为光谱定量反演土壤属性提供借鉴。

\* 国家自然科学基金项目(41371224, 41130530)资助 Supported by the National Natural Science Foundation of China (Nos. 41371224, 41130530)

† 通讯作者 Corresponding author, E-mail: dcli@issas.ac.cn

作者简介: 林 卡(1993—), 女, 浙江温州人, 硕士研究生, 土壤学专业。E-mail: klin@issas.ac.cn

收稿日期: 2017-07-17; 收到修改稿日期: 2017-09-22; 优先数字出版日期(www.cnki.net): 2017-11-10

## 1 材料与方 法

### 1.1 土壤样品来源

研究区位于黑河流域, 是西北地区第二大内陆流域, 跨越青海、甘肃及内蒙古三省区, 面积约为 14.3 万  $\text{km}^2$ , 其景观类型复杂、流域规模适中, 社会生态环境问题突出, 近年来一直是我国重点研究的内陆区域之一<sup>[12-15]</sup>。2012—2013 年的 7—8 月调查采集黑河流域 69 个代表性干旱土剖面 (图 1 为样点空间分布情况), 包括 292 个发生层的土样。样点布设采用代表性调查方法, 综合考虑了气候、地形地貌、海拔、成土母质、土地利用、土壤类型和可达性<sup>[7]</sup>。研究区中无样点或样点极少的东部地区多为沙漠, 西北地区多为禁区。剖面挖掘深度为 1.2 m 或至基岩上界面, 发生层划分依据《野外土壤调查与描述手册》<sup>1)</sup>

### 1.2 土壤样品测定分析

土壤碳酸钙含量测定采用气量法<sup>[16]</sup>。

可见-近红外光谱测定采用安捷伦公司的 Cary5000 分光光度计进行 (土样风干后过 60 目 /0.25 mm 筛, 在 45℃ 的烘箱中烘 24h), 以 1 nm 间隔采样, 波长采集范围为 350~2 500 nm, 其中 350~700 nm 分辨率  $\leq 0.048$  nm, 700~2 500 nm 分辨率  $\leq 0.2$  nm。Cary5000 分光光度计仪器参数及

光谱测定方法的细节参见 Zeng 等<sup>[17]</sup> 研究。

### 1.3 光谱数据预处理

为消除光谱曲线存在的许多“毛刺”噪声和包含的其他无关信息和噪音<sup>[2]</sup>, 利用 Uncrambler 9.7 对光谱曲线进行 Savitzky-Golay (SG) 平滑处理<sup>[18]</sup> (2 次多项式, 11 个点), 将平滑后的数据作为原始数据 R。

### 1.4 数据集的划分

以确定的样本数为划分标准, 将数据集按碳酸钙含量由小到大排列, 设定子集碳酸钙上限分别为 50、60、70、80、90、100、110、120、130、140、150  $\text{g kg}^{-1}$ , 分别记为 A-1、A-2、A-3、A-4、A-5、A-6、A-7、A-8、A-9、A-10、A-11, 相邻子集间互有交集, 每个子集样本数为 60。

以相近的离散度 (变异系数=标准差/均值) 为划分标准, 将数据集按照碳酸钙含量由小到大排列, 综合考虑样本总体的数据分布后, 设定子集碳酸钙上限分别为 50、70、100、130、236  $\text{g kg}^{-1}$ , 分别记为 B-1、B-2、B-3、B-4、B-5, 相邻子集间互有交集, 每个子集离散度在 0.15~0.19 之间。

将每个子集样本按碳酸钙含量从低到高排列, 按 3:1 的比例将每个子集分为建模集和预测集, 即每隔 3 个样本选择 1 个样本进入预测集, 剩余样本进

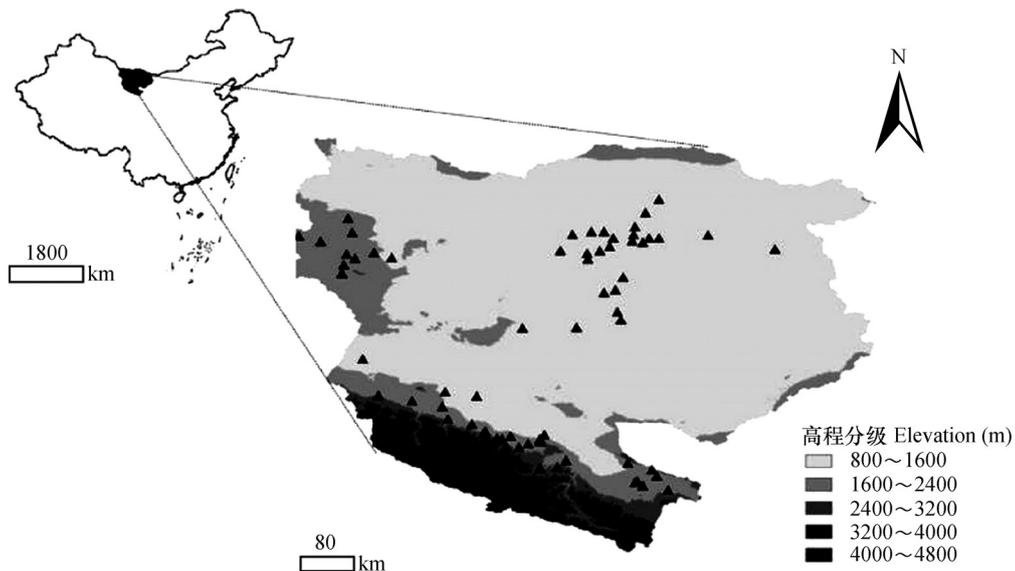


图1 我国西北地区黑河流域干旱土典型剖面空间分布

Fig. 1 Spatial distribution of the selected typical Aridosols profiles in the Heihe River Basin of Northwest China

1) 张甘霖, 李德成. 野外土壤调查与描述手册 (待版)

入建模集，以保证建模样本和验证样本的碳酸钙值均匀分布在子集样本的碳酸钙含量范围内。

### 1.5 数据统计分析与建模方法

采用PLSR算法进行光谱反演建模，这是由于其借鉴了普通多元回归、主成分分析和典型相关分析的思想，较好地解决了自变量多重共线性问题，能在样本数少于自变量的情况下进行回归建模，提高了模型的应用能力、精度和可靠性<sup>[19-21]</sup>。本研究将验证集的相对分析误差（RPD）作为预测精度评价的唯一参数。交叉验证方法采用Full-Cross-Validation。

PLSR建模分析过程在Unscrambler 9.7(Camo Software AS)软件中操作，数据统计及图表制作在IBM Statistics SPSS 20.0和Microsoft Office Excel

2007中完成。

## 2 结果

### 2.1 数据集土壤碳酸钙含量分布特征

利用标准偏差法对碳酸钙含量数据异常值进行剔除，利用残差图对异常光谱数据进行剔除，最终得到69个干旱土代表性剖面共285个发生层土样。图2是样本总体的频数分布直方图。黑河流域干旱土碳酸钙含量介于4.86~236.03g kg<sup>-1</sup>之间，平均值为103.07g kg<sup>-1</sup>。以30~60g kg<sup>-1</sup>和120~150g kg<sup>-1</sup>的土样最多，分别占土样数的21.4%和32.6%。总体而言，碳酸钙含量较高，符合干旱区土壤富含碳酸钙的特征。

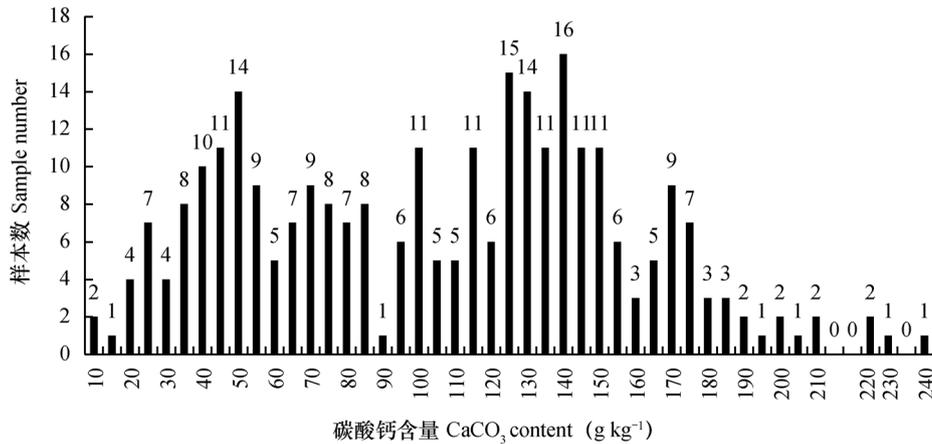


图2 样本总体频数分布直方图

Fig. 2 Histogram of frequency distribution of the whole samples

### 2.2 相同样本量子集的模式构建

以相同样本数作为数据集的划分标准，划分出11个子集，每个子集碳酸钙含量的描述统计结果见表1。每个子集样本数为60，相邻子集间互有交集，从A-1到A-11，碳酸钙从35.77g kg<sup>-1</sup>均值逐渐

增大至137.53g kg<sup>-1</sup>，标准差范围为6.71~13.38g kg<sup>-1</sup>，从A-2到A-8，数据集的变异系数相近，保持在13.40%~18.61%之间，说明这几个数据集之间的离散程度相近。

表1 相同样本量子集碳酸钙含量的描述统计

Table 1 Descriptive statistics of soil CaCO<sub>3</sub> contents in subsets the same in sample size

数据集Data set	样本数Sample size	最小值Min	最大值Max	均值Mean	标准差Std	变异系数C.V.(%)
A-1	60	8.27	49.99	35.77	10.69	29.89
A-2	60	29.04	59.98	43.78	8.15	18.61
A-3	60	39.05	69.34	51.70	9.39	18.16
A-4	60	44.72	78.94	60.60	10.70	17.66
A-5	60	48.56	89.17	66.18	11.57	17.48
A-6	60	57.20	99.90	78.67	13.16	16.73
A-7	60	65.86	109.9	85.98	13.38	15.56
A-8	60	75.86	118.9	98.47	13.19	13.40
A-9	60	98.85	129.7	116.7	9.31	7.98
A-10	60	115.7	139.5	129.1	6.75	5.23
A-11	60	126.1	149.2	137.5	6.71	4.88

为研究碳酸钙含量与建模效果的关系, 本研究基于PLSR算法对11个子集的碳酸钙含量进行建模预测, 图3是11个子集模型预测结果的变化曲线。由图3可见, 各子集的验证集的RPD值介于

0.92 ~ 1.04之间, 在1上下浮动, 未见明显的变化特征, 表明碳酸钙含量高低对可见-近红外光谱预测土壤碳酸钙含量的效果影响不显著。

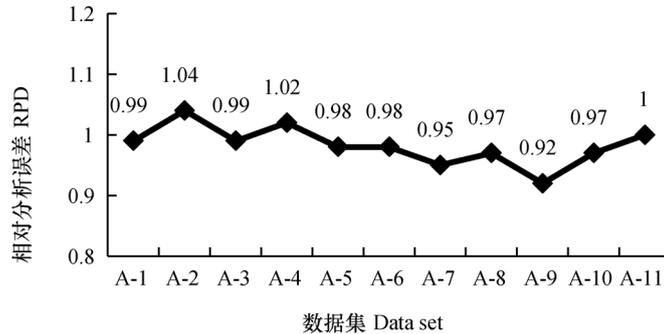


图3 相同样本量子集的模型预测效果变化曲线

Fig. 3 Variation curves of the prediction effect of the model using subsets the same in sample size

### 2.3 相近离散度子集的模型构建

以相近离散度作为数据集的划分标准, 划分出5个子集, 每个子集碳酸钙含量的描述统计结果见表2。从B-1到B-5, 相邻子集间互有交集, 均值从

39.21 g kg<sup>-1</sup>逐渐增大至157.96 g kg<sup>-1</sup>, 各子集变异系数相近, 变化范围为15.15% ~ 18.80%, 受样本总体数据分布的影响, 子集样本数从50个逐渐增加至97个。

表2 相同离散度子集碳酸钙含量的描述统计

Table 2 Descriptive statistics of soil CaCO<sub>3</sub> contents in subsets similar in dispersion degree

数据集 Data set	样本数 Sample number	最小值 Min	最大值 Max	均值 Mean	标准差 Std	变异系数 CV(%)
B-1	50	23.46	49.99	39.21	7.37	18.80
B-2	60	39.05	69.34	51.70	9.39	18.16
B-3	60	57.20	99.90	78.67	13.16	16.73
B-4	98	70.00	129.7	103.8	18.75	18.06
B-5	97	130.1	236.0	158.0	23.93	15.15

基于PLSR算法对5个子集的碳酸钙含量进行建模预测, 其RPD值分别为0.88、0.99、0.98、0.89、1.08, 尽管5个子集的碳酸钙含量均值变化更大, 但各子集预测效果变化仍不显著, 同样表明碳酸钙含量高低对可见-近红外光谱预测土壤碳酸钙含量的效果影响不显著。

## 3 讨论

以上分析结果表明, 碳酸钙含量的高低不是影响光谱预测效果变化的主要因素, 这与研究人员已有的定性认识不一致。本研究根据相同样本数和相近离散度两个指标对数据集进行划分, 分别建立了

11个(相同样本数)A子集和5个(相近离散度)B子集, 此外A中有7个子集(A-2到A-8)既具有相同样本数又保持了相近的离散度, 从而降低了数据集的自由度、离散度等因素对建模预测效果的影响, 增加了预测效果的可比性。

土壤是一个非常复杂而庞大的生态系统, 其在形成过程中的物质转化、迁移和累积特征均会直接或间接地反映在相应的土壤反射光谱中。此外, 其包含的各种物质之间相互联系、相互影响, 各成分差异和含量变化也会影响光谱反射特性。例如土壤中某些成分含量的增加会降低反射率, 如水分、有机质、铁等<sup>[22]</sup>; 而一些成分含量的增加会提高土壤的反射率, 如黏粒含量、矿质粉粒等。这是A、

B两种数据集预测效果变化不显著的主要原因。

因此为探究碳酸钙对土壤光谱反射率的影响,做出各子集的平均光谱反射率曲线图(图4),由于A数据集中有7个子集的大小和分布相似,因此这里只画出A数据集的平均光谱曲线进行说明。虽然土壤成分复杂,但从图4中依然可以清楚地看出反射光谱曲线的基本特征。含不同等级碳酸钙的干旱土的光谱曲线变化大致相同,可见光波段(355~780 nm)的光谱反射率明显低于近红外波段(781~2 495 nm),吸收特征位置基本相同,但吸收深度稍有差异。这主要是因为土壤的可见-近红外光谱曲线是由土壤组成物质的分子吸收外界光源电磁辐射能量,发生能级跃迁而产生的,由于土壤中不同组分的跃迁能极差不同,因此土壤光谱曲线存在不同的吸收特征波段,研究人员以此来定量分析出土壤组成成分含量。基本的土壤成分对反射光谱的影响在图4中均有体现,如较为明显的1 400 nm、1 900 nm和2 200 nm吸收峰主要与水分和黏土矿物有关<sup>[23]</sup>,本研究样品在获取光谱数据前对样品进行了烘干处理,在一定程度上消除了水分对光谱曲线的影响。

从A-1到A-11,碳酸钙含量升高,干旱土的可见-近红外波段光谱反射率也逐渐升高,该趋势在近红外波段更加明显。可见碳酸钙含量对光谱反射率是有影响的,但本研究中光谱反演效果与碳酸钙含量关系不显著,可能的原因是:一方面,在可见光波段特别是350~590 nm波段,11个子集的光谱反射率曲线交叉重合,说明碳酸钙含量的高低对可见光波段反射率的影响较小。而在建模过程中,这个波段的反射率参与建模,从而减少了各模型预测效果之间的差异。另一方面,在近红外波段,A-1、A-2、A-3、A-4、A-5、A-6的光谱反射率接近,这可能是由于这几个子集的碳酸钙含量较低,土壤的其他成分对其产生的影响掩盖了不同碳酸钙含量本身的影响。如有研究表明当土壤有机质含量 $>20 \text{ g kg}^{-1}$ 时,有机质会掩盖碳酸钙的光谱特征<sup>[24]</sup>。本研究中涉及的干旱土土样有机质含量最高为 $12.3 \text{ g kg}^{-1}$ ,但当碳酸钙含量较低时,其对土壤碳酸钙光谱特征的影响仍不可忽略。从而将会影响到碳酸钙光谱预测模型的精度,所以低碳酸钙的子集预测结果相近。

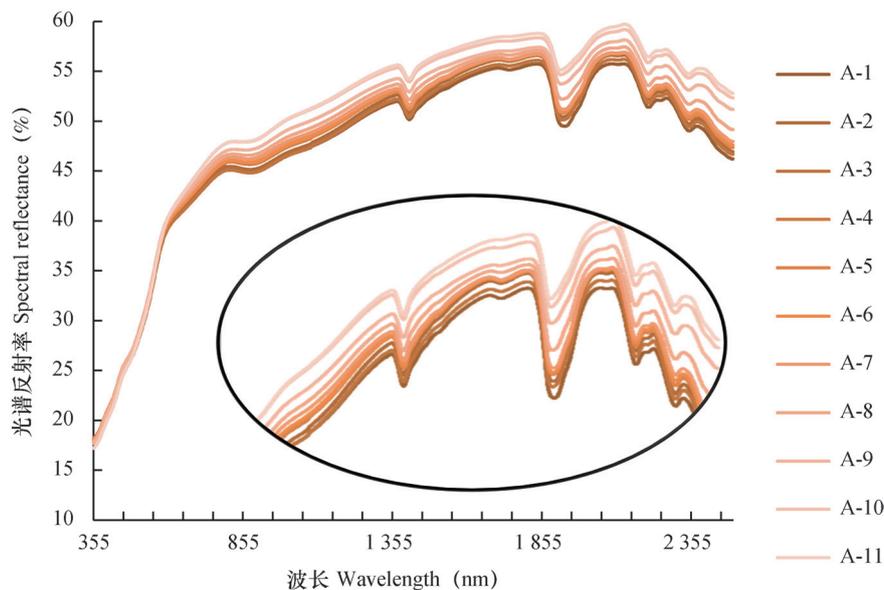


图4 相同样本量子集平均光谱反射率曲线

Fig. 4 Curves of the average spectral reflectance of the subsets the same in sample size

从以上分析可知土壤的光谱反射率明显随着碳酸钙含量的升高而升高。为了深入探求两者关系,逐波段对土壤光谱反射率和碳酸钙含量进行了相关性分析(图5)。可以看出,除A-3、A-4

以外,其他子集的相关系数曲线在近红外波段均为正值,这与在可见-近红外波段,尤其是近红外波段,土壤碳酸钙含量与光谱反射率成正相关这一结论相一致。在可见光波段上,各子集相关系

数曲线相互交叉, 有正有负, 主要是因为土壤中的铁氧化物、羟基、黏土矿物等组分的差别使土壤样品的表观颜色也明显有差别, 这些差别主要在可见光谱上表现出来, 其他官能团的信息均十分微弱。Stenberg等<sup>[25]</sup>发现土壤碳酸盐在1 870 nm、1 990 nm和2 160 nm处有较弱的吸收特征, 仅在2 335 nm处有较强吸收特征。因此碳酸根在可见-近红外光谱上的微弱振动特征, 容易被土壤其他组分特征掩盖。此时需要借助相关光谱数据分析软件, 对数据进行转换, 从而放大其微弱的信

息波段。鉴于本文的研究目的并不是对如何提高光谱预测效果进行探讨, 因此这里未深入研究碳酸钙的吸收特征。

鉴于目标属性的光谱信息相对微弱, 这对研究结果的普适性造成局限。但研究样本的土壤类型相对单一, 可以降低各样本间的差异, 放大碳酸钙含量的差别, 在一定程度上弥补了上述缺陷。同时各子集对样本数和离散度作了限定, 使不同方案所得结果具有可比性, 对照不同方案所得结果均可表明光谱反演效果与土壤属性值的大小无关。

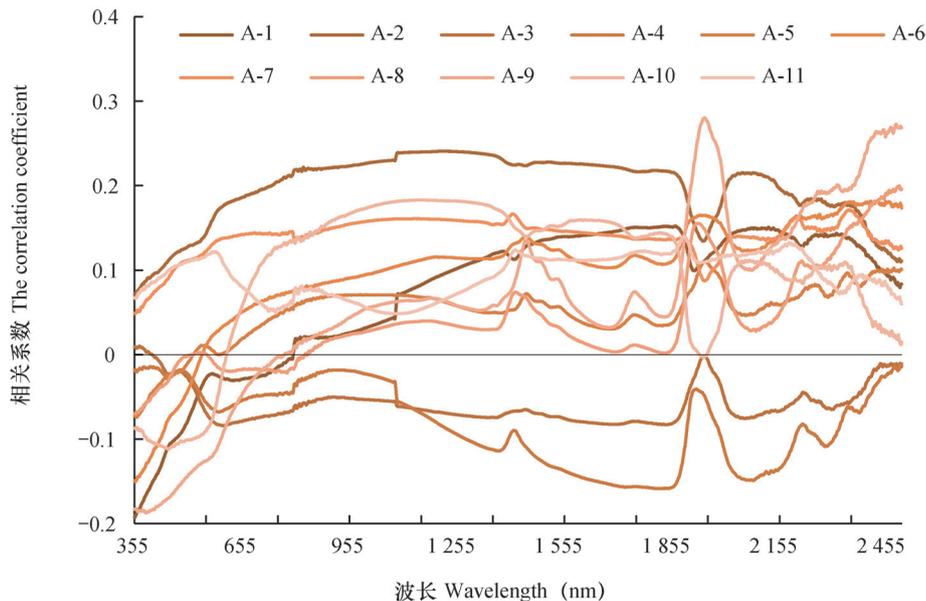


图5 相同样本量子集碳酸钙与反射率的相关系数曲线

Fig. 5 Curves of correlation coefficient of calcium carbonate content and spectral reflectance in subsets the same in sample size

## 4 结 论

本文以相同样本量和相近离散度作为数据集划分标准, 分别建立11个相同样本量子集和5个相近离散度子集, 应用偏最小二乘回归(PLSR)算法进行不同等级土壤碳酸钙含量反演, 结果表明碳酸钙可增加可见-近红外波段的光谱反射率, 但利用可见近红外光谱反演土壤碳酸钙, 其反演效果与碳酸钙含量关系不明显, 因此在利用光谱预测土壤碳酸钙含量时不需对不同含量的碳酸钙样本进行划分。该结论对其他土壤属性的适用性值得进一步实践, 同时如何提高目标属性的预测精度也将是下一步的工作重点。

## 参 考 文 献

- [ 1 ] 郑光辉. 江苏部分地区土壤属性高光谱定量估算研究. 南京: 南京大学, 2010  
Zheng G H. Study on the quantitative estimating of soil properties with hyper-spectrum in some parts of Jiangsu Province (In Chinese). Nanjing: Nanjing University, 2010
- [ 2 ] 史舟. 土地地面高光谱遥感原理与方法. 北京: 科学出版社, 2014  
Shi Z. The theory and method of soil surface hyperspectral remote sensing (In Chinese). Beijing: Science Press, 2014
- [ 3 ] Galvao L S, Pizarro M A, Epiphanyo J C N. Variations in reflectance of tropical soils: Spectral-chemical composition relationships from AVIRIS data.

- Remote Sensing of Environment, 2001, 75 (2): 245–255
- [ 4 ] 徐彬彬. 土壤剖面的反射光谱研究. 土壤, 2000, 32 (6): 281–287  
Xu B B. Study on reflectance spectra of soil profiles (In Chinese). Soils, 2000, 32 (6): 281–287
- [ 5 ] 马赵阳, 杜昌文, 周健民. 土壤碳酸钙中红外光谱特征及其应用. 光谱学与光谱分析, 2012, 32 (5): 1255–1258  
Ma Z Y, Du C W, Zhou J M. Characterization of soil calcium carbonate using mid-infrared photoacoustic spectroscopy (In Chinese). Spectroscopy and Spectral Analysis, 2012, 32 (5): 1255–1258
- [ 6 ] 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
- [ 7 ] 林卡, 李德成, 张甘霖. 西北黑河流域土壤pH与CaCO<sub>3</sub>相当物含量关系研究. 土壤学报, 2017, 54 (2): 344–353  
Lin K, Li D C, Zhang G L. Relationships between pH and content of calcium carbonate and equivalents in soil of the Heihe River Valley, Northwest China (In Chinese). Acta Pedologica Sinica, 2017, 54 (2): 344–353
- [ 8 ] Bissonnais Y L E. Aggregate stability and assessment of soil crustability and erodibility: I. Theory and methodology. European Journal of Soil Science, 1996, 47 (4): 425–437
- [ 9 ] 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
- [ 10 ] 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
- [ 11 ] 洪长桥, 郑光辉, 陈昌春. 苏北滨海土壤碳酸钙含量反射光谱估算研究. 土壤学报, 2016, 53 (5): 1120–1128  
Hong C Q, Zheng G H, Chen C C. Estimation of CaCO<sub>3</sub> content in coastal soil of North Jiangsu with reflectance spectroscopy (In Chinese). Acta Pedologica Sinica, 2016, 53 (5): 1120–1128
- [ 12 ] Yang F, Zhang G L, Yang J L, et al. Organic matter controls of soil water retention in an alpine grassland and its significance for hydrological processes. Journal of Hydrology, 2014, 519: 3086–3093
- [ 13 ] Cheng G D, Li X, Zhao W Z, et al. Integrated study of the water-ecosystem-economy in the Heihe River Basin. National Science Review, 2014, 1: 413–428
- [ 14 ] 怀保娟, 李忠勤, 孙美平, 等. 近50年黑河流域的冰川变化遥感分析. 地理学报, 2014, 69 (3): 365–377  
Huai B J, Li Z Q, Sun M P, et al. RS analysis of glaciers change in the Heihe River Basin in the last 50 years (In Chinese). Acta Geographica Sinica, 2014, 69 (3): 365–377
- [ 15 ] 宁宝英, 张志强, 何元庆. 基于文献统计的黑河流域研究重点和热点学科演变分析. 冰川冻土, 2013, 35 (2): 504–512  
Ning B Y, Zhang Z Q, He Y Q. Development of the research emphasis and main disciplines about the Heihe River Basin reviewed through bibliometric analysis (In Chinese). Journal of Glaciology and Geocryology, 2013, 35 (2): 504–512
- [ 16 ] 张甘霖, 龚子同. 土壤调查实验室分析方法. 北京: 科学出版社, 2012  
Zhang G L, Gong Z T. Soil survey laboratory methods (In Chinese). Beijing: Science Press, 2012
- [ 17 ] Zeng R, Zhao Y G, Li D C, et al. Selection of “Local” models for prediction of soil organic matter using a regional soil vis-NIR spectral library. Soil Science, 2016, 181 (1): 13–19
- [ 18 ] Savitzky A, Golay M J E. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry, 1964, 36 (8): 1627–1639
- [ 19 ] 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
- [ 20 ] Varmuza K, Filzmoser P. Introduction to multivariate statistical analysis in chemometrics. Florida, America: CRC press, 2016
- [ 21 ] Miller J N, Miller J C. Statistics and chemometrics for analytical chemistry. Pearson Education, 2005

- [ 22 ] 李曦. 基于高光谱遥感的土壤有机质预测建模研究. 杭州: 浙江大学, 2013  
Li X. Study on prediction of soil organic matter based on hyperspectral remote sensing ( In Chinese ). Hangzhou: Zhejiang University, 2013
- [ 23 ] Hunt G R. Spectroscopic properties of rocks and minerals. Handbook of physical properties of rocks, 1982, 1: 295—385
- [ 24 ] Baumgardner M F, Silva L F, Biehl L L, et al. Spectral reflectance properties of soils. Advances in Agronomy, 1985, 38: 1—44
- [ 25 ] 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

## Study on Relationship between Soil Calcium Carbonate Content and Inversion Effect Based on Visible Near-infrared Reflectance Spectra

LIN Ka<sup>1, 2</sup> LI Decheng<sup>1†</sup> LIU Feng<sup>1</sup> ZHANG Ganlin<sup>1, 2</sup>

( 1 State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China )

( 2 University of Chinese Academy of Sciences, Beijing 100049, China )

**Abstract** 【Objective】 Soil visible near-infrared reflectance spectra contains large volumes of information on soil physical and chemical properties, which implies that it is feasible to use soil spectra to invert soil properties quantitatively. Is it the higher the property value, the higher the inversion accuracy? However, at present, it is still unclear how to relate quantitatively effects of inversions to soil property contents. 【Method】 Therefore, this study selected soil calcium carbonate content as the target attribute for exploration of quantitative relationship between spectral inversion effect and calcium carbonate content. A total of 292 soil samples were collected out of the genetic horizons of 69 typical Aridosols profiles in the Heihe River Basin, Northwest China, for analysis of calcium carbonate contents with the gasometric method and acquisition of visible near-infrared reflectance spectra with a Cary5000 spectrophotometer. Based on the characteristics of the distribution of calcium carbonate content in the typical study area, 11 identical sample size subsets ( A ) and 5 similar dispersion subsets ( B ) were established with sample size and data dispersion (coefficient of variation) as the criteria for dataset partitioning, and the partial least-squares regression (PLSR) method was used to invert calcium carbonate content from the spectral curves. 【Result】 Results show that calcium carbonate in the Aridosols of the Heihe River Basin varied in the range of 4.86 g kg<sup>-1</sup> ~ 236.03 g kg<sup>-1</sup> in content with an average of 103.07 g kg<sup>-1</sup>. Soil samples with calcium carbonate content varying in the range of 30 ~ 60 g kg<sup>-1</sup> and of 120 ~ 150 g kg<sup>-1</sup>, were in dominancy, accounting for 21.4% and 32.6% of the total, respectively. As a whole, the soil is high in calcium carbonate content, which is consistent with the characteristics of Aridosols being rich in calcium carbonate. With the PLSR, modeling was performed for prediction of calcium carbonate contents of the soil samples in the 11 A subsets. RPD of the validation set of each subset ranged between 0.92 and 1.04, fluctuating around 1 with no obvious features of variation, which indicates that calcium carbonate content does not have much impact on prediction or inversion of soil calcium carbonate content, using visible near-infrared reflectance spectra. Modeling was also done for prediction of calcium carbonate content in 5 B subgroups, with a

similar result. 【 Conclusion 】 Therefore, soil calcium carbonate content is not the main factor affecting the prediction using spectra, which is inconsistent with the qualitative knowledge the researchers already have in mind. Calcium carbonate can enhance spectral reflectance of visible near-infrared bands, but the effect is not so significantly reflected in using the visible near-infrared spectral reflectance to inverse soil calcium carbonate content. Therefore, it seems unnecessary to divide calcium carbonate samples by content of soil calcium carbonate when using spectra to predict calcium carbonate contents. Whether the conclusion is applicable to other soil properties needs to be further verified, and how to improve accuracy of the prediction of target attribute will be the focal point of the next phase of the study.

**Key words** Visible near-infrared reflection spectra; Calcium carbonate content; Inversion effect

(责任编辑: 檀满枝)