

关联类标准化样品特征波段的矿区土壤重金属 Pb 高光谱反演

周 茉^{1,2}, 邹 滨^{1,2*}, 涂宇龙^{1,2}, 夏吉品^{1,2}

1. 中南大学有色金属成矿预测与地质环境监测教育部重点实验室, 地球科学与信息物理学院, 湖南 长沙 410083
2. 中南大学地球科学与信息物理学院, 湖南 长沙 410083

摘 要 目前土壤重金属 Pb 的高光谱反演实验样本多源于研究区采样, 忽略了自然样本中复杂组分的光谱混杂问题, 所建立的模型精度偏低, 重金属光谱响应机理解释仍不明确。基于此, 选取湖南省某铅锌矿区采集相对清洁背景土壤, 运用控制变量方法, 制作 Pb 类标准化样品 40 个, 同时采集矿区 86 个自然污染样本。首先利用类标准化样品建立偏最小二乘回归 (PLSR) 定量反演模型, 提取 Pb 光谱反射特征波段, 然后基于特征波段进行矿区自然污染样本土壤 Pb 定量回归建模, 对比全波段建模结果, 验证特征波段的有效性。实验结果表明: (1) 在背景土壤性质稳定的前提下, 制作的类标准化样品具有总体一致的光谱形态, 同时 Pb 含量作为引起反射率变化的唯一驱动因子证实了土壤中重金属 Pb 光谱响应信号的存在; (2) 相较于自然样本光谱反演结果, 类标准化样品能显著提升土壤 Pb 含量模型估算能力, 决定系数 (R_p^2) 为 0.85, 相对分析误差 (RPD) 2.30, 有效排除其他组分对于光谱建模的干扰, 模型性能良好; 使用变量重要性投影指标 (VIP) 与 PLSR 系数提取的类标准化样品特征波段集中在 970~1 000, 1 700~2 080 和 2 220~2 400 nm 区间; (3) 与全波段建模对比, 基于特征波段的矿区自然污染样本的 PLSR 建模结果, R_p^2 由 0.32 提升至 0.55, RPD 由 1.20 提升至 1.44, 表明特征波段能够较好地表征 Pb 的响应信号, 波段提取有助于过滤噪声, 减少数据冗余同时提高模型反演精度。提出的基于类标准化样品特征波段反演自然土壤样本 Pb 含量, 有效解决自然样本成分复杂, 重金属光谱信号微弱的问题, 丰富了土壤重金属遥感监测理论与应用实例。

关键词 类标准化样品; 土壤重金属; Pb; 特征波段; VIP; 偏最小二乘回归

中图分类号: O657.3 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)07-2182-06

引 言

矿产资源开发引起的土壤重金属污染问题, 正严重威胁着矿区周边自然生态环境的可持续发展和人体健康^[1]。铅是重金属污染的主要元素之一, 易蓄积在土壤中通过食物链进入人体对神经系统和造血系统产生危害, 也是我国“十三五”规划中重金属防治的重点关注污染物之一^[2]。开展污染防治治理工作的首要前提是查明我国的土壤重金属污染状况。传统基于野外取样的实验室化学分析方法虽精度较高, 但耗时费力、成本高昂, 无法实现污染的宏观时空动态和周期性监测。而高光谱遥感作为一种高效、环境友好的探测手段, 能够动态、实时地获取具有连续波段和高光谱分辨率的地物反射光谱曲线, 为实现连续时空、大范围的土壤重金属污染监

测提供了可靠的技术手段^[3]。

目前利用高光谱技术对土壤重金属含量展开研究已取得了一定成果, 国内外学者通过筛选出土壤重金属的特征光谱波段, 建立经验统计意义上的定量反演模型, 证实了矿区土壤中多种重金属含量与土壤光谱的响应关系: 如 Liu 等^[4]使用不同预处理方法反演土壤 Pb 含量, 证实基于室内外高光谱反演重金属的可行性; 江振蓝等^[5]考虑土壤空间异质性, 基于 GWR 建立 Cr, Cu, Zn 和 Pb 四种金属光谱预测模型; Cheng 等^[6]探究了可见光-近红外波段反演郊区土壤中 Cd 和 Pb 等 6 种金属的可行性, 提出了一种联合主成分分析双标图和偏相关系数的重金属反演机制探究方案。

然而, 尽管目前已有的研究一定程度上证实了重金属含量与实测光谱的相关性, 但由于自然样本成分复杂、光谱信噪比低, 仍存在以下问题阻碍土壤重金属遥感监测应用发

收稿日期: 2019-06-07, 修订日期: 2019-10-12

基金项目: 国家自然科学基金项目 (41871317), 中南大学“创新驱动计划”项目 (2018CX016), 2019 年研究生自主探索创新项目 (2019zzts078) 资助

作者简介: 周 茉, 1997 年生, 中南大学地球科学与信息物理学院硕士研究生 e-mail: zhoumo@csu.edu.cn

* 通讯联系人 e-mail: 210010@csu.edu.cn

展:首先,难以得到稳定的高精度反演模型;其次,模型多基于间接反演,而不同重金属与有机质、铁氧化物的响应规律存在差异,土壤重金属光谱响应机制仍有待探究;同时,土壤结构、组分的空间分异也会导致同种重金属在不同地区的光谱响应差异^[7-8],不同研究得到的重金属光谱特征波段各异,如 Cheng 等^[7]得出 Pb 特征波段在 580~660 和 2 020~2 400 nm,而 Liu 等^[4]认为 Pb 光谱特征在 300~500 和 1 900~2 000 nm 波段。已有学者尝试引入控制变量实验^[9],利用皮尔逊相关系数分析重金属溶液浓度与光谱的相关关系,但缺少对于变量间非线性相关性的探讨。在此背景下,本研究尝试引入类标准化样品探究土壤重金属光谱响应规律,在相对清洁背景中人为添加不同剂量 Pb 标准溶液,控制变量得到其余物化成分一致的自制类标准化样品,识别特征波段并应用于自然污染样本反演建模。

1 实验部分

1.1 研究区与土壤取样分析

研究区位于湖南省某铅锌矿区,该矿区矿冶历史悠久并因有色金属闻名世界,地处湖南境内湘江上游,地形以山地为主,土壤主要为红壤;历史采矿活动产生的矿渣废液通过大气沉降、管道运输、河流扩散进入土壤环境,破坏生态环境并严重损害人类健康。

土壤采集包括类标准化样品背景土壤采样和矿区污染土壤采样。考虑矿区地形地势及污染源分布,选取矿区南面高处同一地点相对清洁的土壤,采集表层范围(0~20 cm)内土壤约 50 kg,混匀装入专用密封塑料袋带回作为类标准化样品背景土壤;考虑矿区作业运输路线、土地利用类型及污染扩散因素布设采样点。采用五点取样法采集样点 10 m² 范围内矿区污染表层土样 86 个,装入专用密封塑料袋带回,使用手持式 GPS 按顺序记录 WGS84 框架下采样点经纬度坐标,位置如图 1 所示。所有土样均在室内阴凉通风处风干,剔除石块、植物碎片,研磨,过 100 目尼龙筛,按编号分装。对于矿区污染土壤,使用 HNO₃-HF-HClO₄ 混合酸分解,电感耦合等离子发射光谱法(USEPA 6010C:2007)测定土壤重金属 Pb 含量。

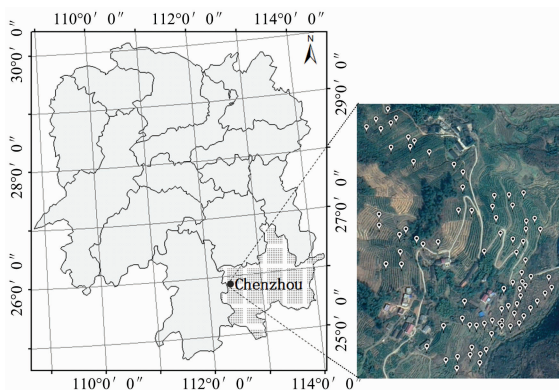


图 1 研究区及采样点位置

Fig. 1 Location of study area and soil sampling points

1.2 类标准化 Pb 污染土壤样品制作及测定

为保证土壤类标准化样品稳定性,采用小单元包装,将采回的背景土壤分装于专用塑料袋,每份样品 100 g。考虑自然土壤 Pb 含量,设置 40 个浓度等级,范围为 50~8 000 mg·kg⁻¹,非等间距分布的浓度设置符合实际情况且满足 0.05 显著水平下的正态分布。类标准化样品背景土壤试剂添加方案如表 1 所示。使用浓度为 1 000 μg·mL⁻¹ 的铅标准溶液(GSBG62071—90(8201)),根据方案设定,量取相应剂量的标准溶液,稀释。最后将稀释溶液添加至 100 g 背景土壤中,静置,等待样品中的溶液渗透背景土壤,待自然风干后再次研磨、混匀,类标准化样品 Pb 含量测定与自然污染样本所述步骤相同。

表 1 类标准化样品背景土壤 Pb 含量设计方案(mg·kg⁻¹)
Table 1 Designed scheme of Pb content in the background soil of near standard samples

编号	含量	编号	含量	编号	含量	编号	含量
1	50.0	11	2 825.0	21	3 975.0	31	5 300.0
2	450.0	12	2 975.0	22	4 050.0	32	5 400.0
3	900.0	13	3 100.0	23	4 175.0	33	5 725.0
4	1 200.0	14	3 275.0	24	4 325.0	34	6 050.0
5	1 450.0	15	3 475.0	25	4 475.0	35	6 400.0
6	1 750.0	16	3 600.0	26	4 600.0	36	6 600.0
7	1 950.0	17	3 675.0	27	4 725.0	37	6 800.0
8	2 200.0	18	3 775.0	28	4 875.0	38	7 000.0
9	2 450.0	19	3 850.0	29	5 025.0	39	7 600.0
10	2 650.0	20	3 900.0	30	5 150.0	40	8 000.0

1.3 光谱测定与预处理

使用 PSR-3500 便携式地物光谱仪测定 86 个矿区样本和 40 个类标准化样品室内全波段(350~2 500 nm)范围反射率,光谱分辨率为 1.5 nm(350~1 000 nm),3.8 nm(1 000~1 900 nm),2.5 nm(1 900~2 500 nm);光源使用 1 000 W 卤素灯,5°视场角,光源照射方向与垂直方向呈 15°夹角;光源距离 30 cm,探头距离 5 cm。开始测量前及每间隔 5 个样本进行白板辐射定标,每个样本重复测定 10 次,取平均反射率为实际光谱反射数据。

为放大光谱与土壤组分相关的细节特征,消除数据无关信息,进行光谱去噪、平滑及光谱增强。首先,去除可见-近红外区域边缘低信噪比区间:350~399 和 2 400~2 500 nm,降低噪声同时提高计算效率;选择 Savitzky-Golay 卷积平滑光谱,设置三阶多项式,移动窗口宽度 11 进行多项式最小二乘拟合;使用多元散射校正(multiplicative scatter correction, MSC)消除样品间散射带来的基线漂移效应,增强光谱特征。

1.4 特征波段选择与模型建立

偏最小二乘回归(partial least squares regression, PLSR)作为一种经典多元统计分析模型能够有效地规避变量共线性问题,在自变量存在多重相关性的情况下进行建模,是土壤重金属 Pb 高光谱反演的可靠方式^[10]。实验基于 Matlab R2018a,使用简单交叉验证方法检验模型性能,利用决定系数(R^2)、均方根误差(RMSE)及相对分析误差(RPD)评价模

型精度, R^2 越大, RMSE 越小, 表明模型精度越高, 拟合验证效果越好。当 $1.4 < \text{RPD} < 2$ 时, 模型可对样本作出粗略预估, 当 $\text{RPD} > 2$ 时, 模型对于样本有极好的预测能力。

变量重要性投影指标(variable importance in projection, VIP)是 PLSR 模型的一种滤波型变量选择方法, 研究表明结合 VIP 和 PLSR 系数(即 b-coefficient)的变量筛选方法能够有效识别重金属光谱特征波段, 揭示光谱响应机制^[6]。VIP 的思想是累积计算逐个变量 j 被成分 w 影响的重要性, 即变量重要性的投影, 如式(1)所示。

$$\text{VIP}_j = \sqrt{p \sum_{k=1}^h (\text{SS}(b_{ktk}) (\omega_{jk} / \|\omega_k\|)^2) / \sum_{k=1}^h \text{SS}(b_{ktk})} \quad (1)$$

其中, $\text{SS}(b_{ktk}) = b_{ktk}^2 t_{ktk}$, 表示第 k 个成分的平方和, ω 是权值矩阵, h 是最优 PLSR 模型中的潜在变量个数。VIP 得分的阈值通常设置为 1, b-coefficient 的阈值是其总体的统计标准差, 满足 $\text{VIP} > 1$ 且 b-coefficient 大于总体标准差的波段为特征波段。

表 2 类标准化样品背景土壤组分实测值

Table 2 Measured values of background soil components

土壤组分	pH	有机质/ ($\text{g} \cdot \text{kg}^{-1}$)	全铁/ ($\text{mg} \cdot \text{kg}^{-1}$)	铅/ ($\text{mg} \cdot \text{kg}^{-1}$)	砷/ ($\text{mg} \cdot \text{kg}^{-1}$)	镉/ ($\text{mg} \cdot \text{kg}^{-1}$)
第一组	4.11	33.2	21.1	35.8	15.1	0.108
第二组	4.07	32.3	23.3	36.3	17.2	0.061
第三组	4.07	33.3	21.8	36.6	15.0	0.061
第四组	4.05	33.3	22.0	40.4	15.0	0.099
第五组	4.09	29.4	21.5	34.9	14.0	0.066
平均值	4.08	32.30	21.94	36.80	15.26	0.080
风险管制值*	—	—	—	400	200	1.5

注: 风险管制值:《土壤环境质量农用地土壤污染风险管控标准》(试行)(GB15618—2018)风险管制值^[11]

Note: Risk intervent values for soil contamination of agricultural land in GB15618—2018^[11]

表 3 矿区土壤 Pb 含量描述性统计数据

Table 3 Statistic of soil Pb content in Chenzhou mining area

元素	最小值	最大值	均值	标准差	变异系数	偏度	峰度	二级标准值
Pb	54.60	10 054.00	753.42	1 522.85	2.02	4.11	21.49	250

2.2 类标准化样品与矿区土壤样品光谱特征分析

图 2 是类标准化样品、自然污染土样的原始和预处理后的光谱曲线图。由原始光谱图[图 2(a)和(b)]可知, 不同地区土壤光谱曲线总体形态相似, 在 $< 400 \text{ nm}$ 和 $> 2 400 \text{ nm}$ 波段区间光谱变化剧烈, 噪声较大; 土壤光谱在 $< 1 000 \text{ nm}$ 波段间随波长增加反射率迅速上升, 在 $1 400, 1 900$ 和 $2 200 \text{ nm}$ 波段附近有明显的水吸收谷, 反射率降低。光谱预处理结果[图 2(c)和(d)]表明, SG 滤波能有效消除光谱数据噪声, 预处理后光谱更为平滑, 同时多元散射校正一定程度上修正了光谱基线漂移现象。两组数据的相似性为基于类标准化样品提取的光谱响应特征波段, 建立自然样本土壤 Pb 反演模型提供支撑。

2 结果与讨论

2.1 土壤 Pb 含量描述性统计分析

类标准化样品背景土壤、矿区 86 个自然污染样本 Pb 含量统计结果如表 2、表 3 所示。在混匀的背景土壤中随机抽取五组样品送检, 五组样品结果均显示 Pb, As 和 Cd 重金属含量低于国家《土壤环境质量农用地土壤污染风险管控标准》(试行)(GB15618—2018)^[11] 风险管制值, Pb 平均值 $36.8 \text{ mg} \cdot \text{kg}^{-1}$, 土壤 pH 值, 有机质、全铁及重金属含量稳定。因此, 基于背景土壤制备的类标准化样品能够减轻土壤多变量引起的光谱变化对重金属光谱反演的干扰, 为控制变量探究单一重金属 Pb 响应机制提供了良好的条件。矿区土壤重金属 Pb 平均含量超过风险管制值一倍以上, 最大值 $10 054.00 \text{ mg} \cdot \text{kg}^{-1}$, 达风险管制值的 25 倍, 样点超标率达 37.2%, 重金属含量严重超标; 同时, 由于矿区不同位置尾矿堆积、矿渣排放相异, 土壤 Pb 标准差和变异系数较大, 矿区内重金属浓度分异度较高, 不服从正态分布, 数据离散化程度高。

另一方面, 比较两类样品光谱, 相较于类标准化样品[图 2(a)和(c)]光谱的统一趋势, 自然样本原始光谱稳定性较差, 存在许多毛刺噪声。由于土壤光谱特性与土壤组分中多种物质含量有关, 多成分的非线性同步变化导致自然样品光谱缺乏统一特征, 噪声杂乱; 而制备的类标准化样品, Pb 含量变化作为光谱反射率变化的唯一驱动因子, 不同样本的反射率差异初步验证了 Pb 光谱响应的存在, 同时预处理后光谱具有较为一致的光谱吸收峰谷, 为基于类标准化样品探究光谱响应特征波段提供了可行性。

2.3 类标准化样品 Pb 含量 PLSR 建模及特征波段提取

图 3 显示了类标准化样品 Pb 含量的 PLSR 建模和特征波段提取结果。由图 3(a)可知, PLSR 模型 $R_c^2 = 0.98$, $R_p^2 = 0.85$, $\text{RMSEP} = 843.87$, $\text{RPD} = 2.30$, 模型精度较高, 对于

高值和低值的预测稳定,表现良好,可用于土壤 Pb 的定量预测。相较于以往研究中自然土壤样本反演结果^[6, 8],类标准化样品建模精度显著提升,说明解混复杂光谱曲线是提升反演精度的途径之一。而自然样本由于组分混杂,Pb 光谱信

号特征被掩盖,难以提取有效的光谱波段建立可靠的反演模型。建模结果证实了土壤 Pb 存在一定光谱响应信号,类标准化样品能够排除其他组分对于光谱的干扰,提升重金属含量反演模型精度。

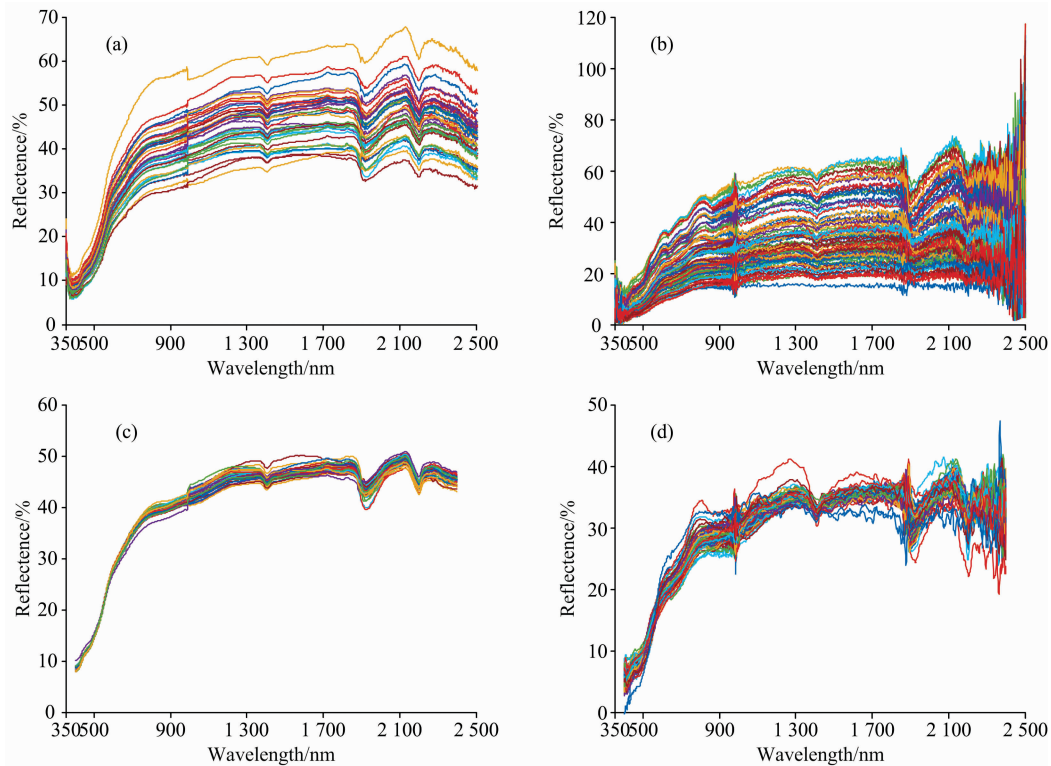


图 2 土壤光谱曲线图

(a), (c): 类标准化样品 Pb 污染土壤原始和预处理后光谱; (b), (d): 矿区自然土壤样本原始光谱和预处理后光谱

Fig. 2 Spectra of soil samples

(a), (c): The original and pretreatment spectra of near standard soil samples;

(b), (d): The original and pretreatment spectra of mining area soil samples

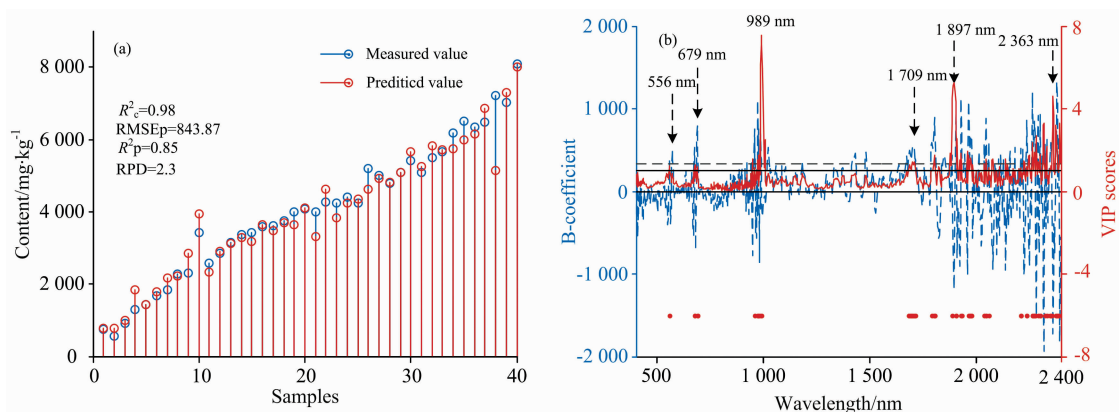


图 3 类标准化样品 Pb 含量 PLSR 建模结果图

(a): 建模结果茎叶图; (b): VIP 评分结果

Fig. 3 PLSR modeling results for near standard soil samples Pb contents

(a): Stem plots of modeling; (b): VIP scores

图 3(b)表明了对类标准化样品建模过程中,结合 VIP 和 b-coefficient 得到的 Pb 在 400~2 400 nm 的光谱特征波段, VIP 阈值为 1, b-coefficient 标准差为 338.581 6。由图可

知,光谱对土壤 Pb 含量的变化均产生了不同程度的响应; VIP 与 b-coefficient 响应曲线基本成轴对称状态,在 400~940 和 1 000~1 880 nm 区间曲线变化平稳,前者较低呈现

弱信号状态; 1 900~2 400 nm 内 VIP 值与 b-coefficient 增幅明显, 尤其后者响应剧烈, 信号显著增强; VIP 和 b-coefficient 的峰值分别出现在 989 和 2 320 nm 附近, 通过两者筛选出的特征波段集中在 970~1 000, 1 700~2 080 以及 2 220~2 400 nm。筛选结果直观地反映了特征波段的光谱位置分布和对模型的贡献率, 特征波段提取结果再次证实土壤 Pb 光谱响应信号的存在, 并且控制变量的类标准化样品能够有力佐证 Pb 含量是驱动光谱信号变化的唯一因子。

2.4 基于特征波段的矿区土壤含量高光谱反演

矿区自然污染土样 Pb 的 PLSR 光谱建模结果如图 4 所示。图 4(a)是全波段光谱建模结果, $R_c^2=0.24$, $R_p^2=0.32$,

$RMSEP=1\ 376.83$ 。相比类标准化样本建模结果, 自然样本建模精度偏低, 误差较大, 光谱低信噪比是影响模型效果的主要原因。图 4(b)是特征波段建模结果, 结合表 4 的对照分析可知, 去除非特征波段后建模精度有显著的提升, R_p^2 由 0.32 提高至 0.55, RPD 由 1.20 提高至 1.44, $RMSEP$ 由 1 376.83 降低至 1 207.70。结果说明基于类标准化样品提取的土壤 Pb 光谱特征波段能够较好表征重金属 Pb 的响应信号, 同时为从自然样本混合光谱中识别重金属指示性信号提供依据, 减少冗余的光谱数据对建模的干扰, 提高反演模型性能。

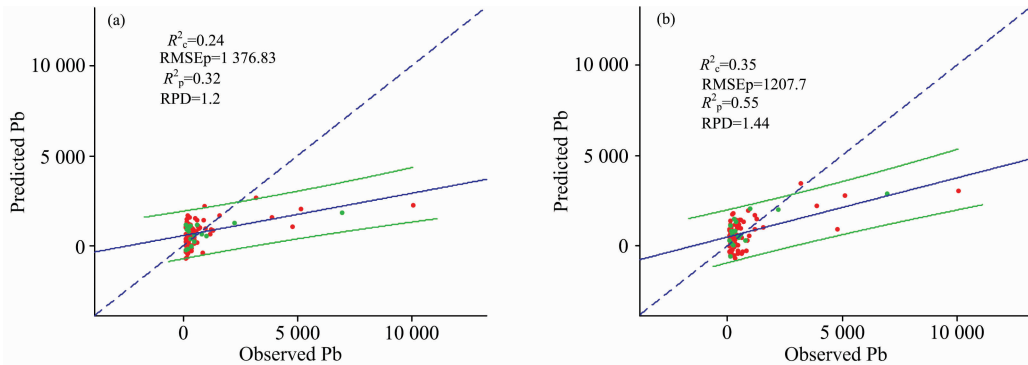


图 4 矿区土壤 Pb 偏最小二乘回归模型散点图

(a): 全波段建模散点图; (b): 特征波段建模散点图

Fig. 4 Scatter plots of PLSR modeling for Pb content in mining area

(a): Entire spectra band; (b): Characteristic spectra band

表 4 矿区土壤 Pb 全波段与特征波段建模结果对比

Table 4 Comparison of Pb content estimation model based on entire spectra band and characteristic spectra band

建模波段	RMSEP	R_p^2	RPD	LVs
全波段	1 376.83	0.32	1.20	2
特征波段	1 207.69	0.55	1.44	3

3 结论

选取湖南省某铅锌矿区为研究区, 制作土壤重金属 Pb 类标准化样品, 基于类标准化样品进行 Pb 光谱特征波段提取, 探讨 Pb 光谱响应机制; 并对比基于全波段和特征波段的自然样本偏最小二乘定量回归模型精度和性能, 结果表明:

(1)在背景土壤组分含量稳定情况下, 控制土壤 Pb 含量

为单一变量得到的 40 个类标准化样品, 光谱总体形态趋势一致, 反射率曲线随 Pb 含量变化总体上移或下降, 证实了土壤中 Pb 光谱响应信号的存在; (2)相较于自然样本, 类标准化样品光谱反演精度更高, 控制背景成分变化的类标准化样品能够有效排除其他组分对于光谱建模的干扰, 改善模型性能; 通过 VIP 和 PLSR 回归系数筛选的 Pb 特征指示波段集中在 970~1 000, 1 700~2 080 和 2 220~2 400 nm 区间, 提出的基于类标准化样品的土壤重金属特征波段识别为不同类别重金属的光谱响应机理探究提供了新的思路; (3)相比矿区样本全波段 PLSR 建模结果, 引入类标准化样品特征波段的建模模型表现更佳, 说明本研究得到的特征波段能够较好表征 Pb 的响应信号, 波段提取有助于过滤噪声光谱信息, 减少数据冗余同时提高模型反演精度, 有效解决自然样本成分复杂, 重金属光谱信号微弱的问题, 为发展基于高光谱遥感的土壤重金属监测评估提供了理论和技术支撑。

References

- [1] LIU Chang, SONG Bo, ZHANG Yun-xia, et al(刘畅, 宋波, 张云霞, 等). Environmental Science(环境科学), 2018, 39(2): 899.
- [2] SONG Bo, TIAN Mei-ling, CHEN Tong-bin, et al(宋波, 田美玲, 陈同斌, 等). Acta Pedologica Sinica(土壤学报), 2017, 51(1): 92.
- [3] Shi T Z, Chen Y Y, Liu Y L, et al. Journal of Hazardous Materials, 2013, 265C(2): 166.
- [4] Liu K, Zhao D, Fang J Y, et al. Journal of the Indian Society of Remote Sensing, 2016, 45(5): 805.
- [5] JIANG Zhen-lan, YANG Yu-sheng, SHA Jin-ming, et al(江振蓝, 杨玉盛, 沙晋明, 等). Acta Geographica Sinica(地理学报), 2017, 72

- (3): 533.
- [6] Cheng H, Shen R L, Chen Y Y, et al. *Geoderma*, 2019, 336: 59.
- [7] Wang J J, Cui L J, Gao W X, et al. *Geoderma*, 2014, 216: 1.
- [8] Zhang X, Sun W C, Cen Y, et al. *Science of The Total Environment*, 2019, 650: 321.
- [9] CHENG Hang, WAN Yuan, CHEN Yi-yun, et al(程 航, 万 远, 陈奕云, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2018, 38(3): 771.
- [10] Rossel R A V, Walvoort D J J, Mcbratney A B, et al. *Geoderma*, 2006, 131(1-2): 75.
- [11] Ministry of Ecology and Environment of the People's Republic of China(中华人民共和国生态环境部). GB15618—2018. *Soil Environmental Quality Risk Control Standard for Soil Contamination of Agricultural Land(GB15618—2018. 土壤环境质量 农用地土壤污染风险管控标准(试行)*, 2018.

Hyperspectral Modeling of Pb Content in Mining Area Based on Spectral Feature Band Extracted from Near Standard Soil Samples

ZHOU Mo^{1,2}, ZOU Bin^{1,2*}, TU Yu-long^{1,2}, XIA Ji-pin^{1,2}

1. The Key Laboratory of Metallogenic Prediction of Nonferrous Metals and Geological Environment Monitoring (Central South University), Ministry of Education, School of Geoscience and Info-Physics, Changsha 410083, China
2. School of Geosciences and Info-Physics, Central South University, Changsha 410083, China

Abstract Natural soil samples are the primary data source of heavy metal hyperspectral predicting models. However, their spectra are often confounded by complex components, resulting in poor performance model and unclear explanation for heavy metal spectral response mechanism. Near standard soil samples provide a promising method for the mechanism research. In this paper, 86 natural soils samples and relatively clean background soil were collected in a lead-zinc mine in Hunan province, and 40 near standard samples were made by artificially adding heavy metals in background soil using control variable method. Feature bands for soil Pb spectra were first selected based on near standard soil samples using partial least squares regressions(PLSR). The feature bands were used to calibrate the prediction model with PLSR for natural soil samples. The existence of Pb absorption features was confirmed by the overall consistent and change trend of near standard samples reflectance spectra. Near standard samples provided acceptable estimation accuracies of Pb concentrations, with the determination coefficient (R_p^2), and the ratio of prediction to deviation (RPD) values of 0.85 and 2.30. When compared with the entire-band PLSR model, feature-band model for natural soil samples increased the R_p^2 and the RPD from 0.32 and 0.20 to 1.55 and 1.44 by removing uninformative spectral variables. The mechanism investigation strategy we proposed could effectively solve the problem of complex sample composition and weak heavy metal spectral signal in previous research, and be applied in further soil heavy metal remote sensing monitoring.

Keywords Near standard samples; Soil heavy metal; Pb; Feature band; VIP; PLSR

(Received Jun. 7, 2019; accepted Oct. 12, 2019)

* Corresponding author