

引用格式: Yin Fang, Feng Kai, Wu Mengmeng, *et al.* A Remote Sensing Estimation Method for Heavy Metals in Soil based on Piecewise Partial Least Squares Model[J]. Remote Sensing Technology and Application, 2021, 36(6): 1321-1328. [尹芳, 封凯, 吴朦朦, 等. 一种基于分段偏最小二乘模型的土壤重金属遥感反演方法[J]. 遥感技术与应用, 2021, 36(6): 1321-1328.]  
DOI: 10.11873/j.issn.1004-0323.2021.6.1321

## 一种基于分段偏最小二乘模型的土壤重金属 遥感反演方法

尹芳<sup>1</sup>, 封凯<sup>2</sup>, 吴朦朦<sup>2</sup>, 拜得珍<sup>3</sup>, 王蕊<sup>3</sup>, 周园园<sup>3</sup>, 尹春涛<sup>2</sup>,  
尹翠景<sup>2</sup>, 刘磊<sup>2</sup>

(1. 长安大学土地工程学院, 陕西 西安 710054;

2. 长安大学地球科学与资源学院, 陕西 西安 710054;

3. 青海省环境科学研究设计院有限公司, 青海 西宁 810000)

**摘要:** 土壤中重金属由于其毒性而成为最有害的环境污染物之一, 利用遥感进行土壤重金属检测和分布制图是目前最为高效的手段。采用哨兵二号(Sentinel-2)多光谱影像与实测样品光谱数据, 对山西省铜矿峪铜矿尾矿库及其周边农田土壤的铜(Cu)含量进行估算, 利用68个土壤样品的反射光谱, 优选出适合土壤铜含量预测的波段, 结合分段偏最小二乘法(Piecewise Partial Least Squares Regression, P-PLSR), 对土壤铜含量进行估算, 将模型用于Sentinel-2影像获得了Cu含量的空间分布。通过P-PLSR对实测样品光谱建模反演Cu含量的决定系数( $R^2$ )为0.89, 预测偏差比(RPD)为2.82; 利用Sentinel-2多光谱影像获得了该区域Cu元素含量空间分布, 其Cu含量的估算精度 $R^2$ 为0.74, RPD为1.73, Cu含量高值区空间分布与尾矿库关系密切。Sentinel-2多光谱数据具有高空间分辨率(10、20和60 m)、高时间分辨率和幅宽大(290 km)等优势, 通过敏感波段选择并建立反演模型, 可实现大范围土壤环境制图。

**关键词:** 土壤重金属; 多光谱遥感; 分段偏最小二乘; 定量反演

**中图分类号:** X833; X87 **文献标志码:** A **文章编号:** 1004-0323(2021)06-1321-08

### 1 引言

矿山开采所形成的尾矿库不仅破坏土地资源, 且使尾矿库周边土壤中的重金属元素异常富集, 导致土壤生态环境恶化<sup>[1]</sup>。土壤重金属污染调查的传统方法基于野外样品化学分析结果进行空间插值以获得重金属元素的空间分布<sup>[2]</sup>所得结果无法准确描述污染细部特征, 且调查成本高、耗时长<sup>[3]</sup>。遥感技术与传统土壤调查方法相比具有非破坏性、可重

复性、高效、大尺度观测等特点<sup>[4-7]</sup>。

利用遥感技术可以有效估算农业土壤和河流沉积物中的重金属浓度<sup>[8-10]</sup>。Kooistra等<sup>[11]</sup>基于土壤重金属含量与土壤反射光谱特征波段间的高相关性, 建立偏最小二乘回归(Partial Least Squares Regression, PLSR)模型对莱茵河流域冲积平原区域土壤中的Cd、Zn含量进行反演。Chodak等<sup>[12]</sup>利用多元逐步回归法对森林土壤有机质层的重金属含量进行反演。Siebielec等<sup>[13]</sup>对波兰Tamowskie Gory

收稿日期: 2020-10-10; 修订日期: 2021-10-15

基金项目: 青海省重大科技专项“湟水流域水—气—土一体化环境管理体系及污染控制关键技术集成与示范”(2018-SF-A4), 国家自然科学基金项目(42071258), 中央高校基本科研业务费专项资金(300102270204, 300102278303)。

作者简介: 尹芳(1983—), 女, 河北河间人, 博士, 副教授, 主要从事遥感与GIS应用研究。E-mail: yinf@chd.edu.cn

通讯作者: 拜得珍(1981—), 男, 青海互助人, 副研究员, 主要从事环境规划和环境科研等技术工作。E-mail: 282785316@qq.com

地区的表层土壤样品建立了Cu、Cr、Ni等含量估算模型。Ferreira等<sup>[14]</sup>运用人工神经网络法建模,对巴西某地区土壤中Cu含量进行预测。Laakso等<sup>[15]</sup>通过分析短波红外光谱与矿物、元素含量的关系,预测了某些金属的空间分布特征。偏最小二乘模型由于具有简化高光谱数据结构、可改善信息冗杂的优势,且可解决光谱波段多重相关性问题,在土壤参数反演中应用最为广泛<sup>[16-19]</sup>。

然而,目前土壤重金属遥感反演的相关研究多基于实验室实测的土壤反射光谱数据,仅有少数利用多光谱遥感影像反演土壤参数的实例,多集中于土壤PH、有机质等特征<sup>[16]</sup>。多光谱遥感影像由于波段较少、光谱分辨率较低,加之土壤重金属含量低、光谱特征微弱,利用多光谱遥感影像反演土壤重金属含量的可行性尚需研究。

因此,以山西省铜矿峪铜矿尾矿库及周边农田为研究区,利用土壤样品化学分析结果、可见光-短波红外区间反射光谱及Sentinel-2多光谱影像,建立研究区铜元素含量反演模型,探索基于多光谱影像的土壤重金属含量高精度反演方法。

## 2 研究区和数据

### 2.1 研究区概况

研究区位于山西省中条山西北部,垣曲县西南

5 km处(图1)。气候类型属温带半干旱、半湿润大陆性气候。铜矿峪铜矿属特大型斑岩型铜矿,矿产在开采、选冶过程中,矿渣堆积形成一个巨大的尾矿库(图1)。尾矿库内部较高的Cu元素含量对周围农田造成了巨大威胁,且尾矿库在2000年曾发生过溃坝,导致周边农田Cu污染严重。

### 2.2 数据

#### 2.2.1 土壤样品采集与光谱数据获取

在铜矿峪铜矿尾矿库内部及周边农田进行样品定点采集,采样时根据当地农田的分布、与尾矿库距离、可到达程度、在影像中未被植被覆盖等因素综合选取了68个样点,采样点分布如图1所示。本研究共采集68个深度为0~10 cm的浅层土壤样品,重约500 g,其中包括尾矿库内部9个样品以及周边农田59个样品,并对采样点情况做好记录。

除去土壤样品中的石块、植物根系、凋落物等后晾晒风干。土壤样品的实验室光谱测量采用SR-3500便携式光谱仪,其光谱范围为350~2 500 nm,光谱采样间隔为1 nm。利用两盏溴钨灯为光源,采用视场为25°的光纤探头测量土壤光谱,使用白色硫酸钡面板提供反射校准参考。每个土壤样品测量5次,并使用平均光谱代表样品光谱<sup>[20]</sup>。

#### 2.2.2 遥感数据与预处理

Sentinel-2卫星携带的高分辨率多光谱成像仪

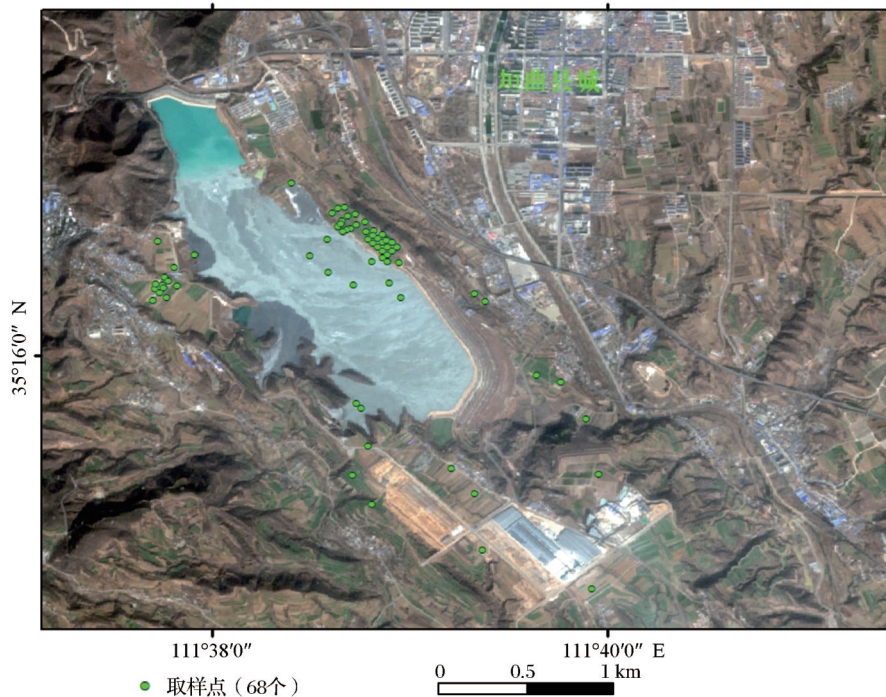


图1 2018年研究区Sentinel-2自然彩色合成影像(RGB432)和土壤样品分布

Fig.1 Sentinel-2 natural color composite image (RGB432) and soil sample distribution in the study area in 2018

在可见光—短波红外范围内有13个波段,幅宽达290 km,地面分辨率分别为10、20和60 m。实验选取的影像获取时间为2018年1月1日,无云、雪且植被覆盖较少。利用欧空局的Sen2Cor工具对Sentinel-2影像进行大气校正,之后将空间分辨率为10 m和20 m共10个波段像元统一重采样至10 m。

### 3 研究方法

#### 3.1 土壤铜元素含量分析

利用电感耦合等离子体原子发射光谱(ICP-AES)等方法测试了样品Cu、Pb、Zn、Cd、Cr、Ni、As、Sb、Hg共9种元素含量。根据中国生态环境部2018年发布的土壤环境质量标准(GB 15618-2018),68个样品中共有16个样品(24%)的Cu含量超过风险筛选值(100 mg/kg),仅有1个样品Cd含量超过风险筛选值(0.6 mg/kg),其余元素基本无超过风险筛选值的样品,因此选择土壤Cu含量开展研究。

土壤样品Cu含量的基本统计特征如表1所示,Cu含量变化范围为28.40~413.63 mg/kg。将采集的68个土壤样品按照约3:1的比例划分为建模集和验证集,从样品中随机抽取48个样品作为建模样品,剩余20个样品作为验证样品,验证集样品具有较大的均值(94.87 mg/kg)和变异系数(1.05)。

表1 68件土壤样品中Cu含量的描述性统计信息(mg/kg)

Table 1 Descriptive statistics of Cu content for the 68 soil samples (mg/kg)

重金属	最小值	最大值	均值	标准差	变异系数	样品个数
全部	28.40	413.63	83.23	74.64	0.90	68
建模集	28.40	302.81	79.53	63.29	0.80	48
验证集	28.60	413.63	94.32	98.87	1.05	20

#### 3.2 土壤光谱曲线特征分析

土壤样品反射光谱在400~800 nm、1 500~1 800 nm、2 100~2 400 nm范围内存在明显的差异(图2(a))。400~800 nm处的光谱差异反映了不同样品的有机质和铁氧化物含量存在差异<sup>[21]</sup>;在1 500~1 800 nm处虽然无特征性吸收,但不同样品光谱曲线坡度差异明显,这些差异与样品的矿物组成、粒度等有关;大部分样品在2 200 nm均存在弱吸收带,为粘土矿物的特征吸收带<sup>[22-23]</sup>;2 335 nm附近的吸收可能为绿泥石的吸收带,与斑岩铜矿床伴有绿泥石化蚀变吻合,导致铜含量高的样品光谱与其他样品光谱在此区间差异明显<sup>[24]</sup>。由于实验的目的是基于Sentinel-2影像反演研究区土壤Cu元素含量分布,因此在筛选特征波段前,利用ENVI软件的Spectral Library Resampling功能将实测光谱重采样至Sentinel-2影像波段范围(如图2(b)),重采样后样品光谱的主要特征得以保留。

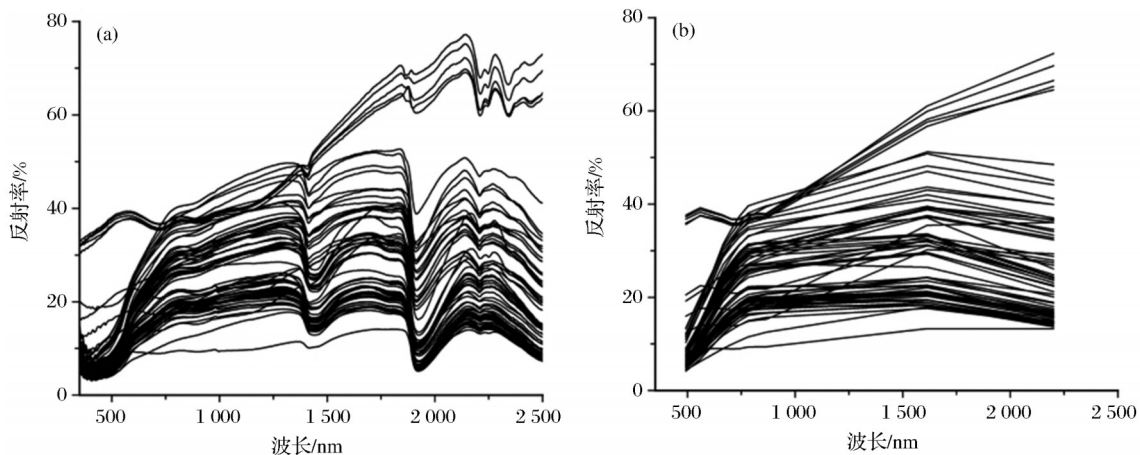


图2 68个土壤样品经平滑处理后的实验室光谱和土壤样品实测光谱重采样至Sentinel-2影像光谱范围

Fig.2 Laboratory spectra of 68 soil samples after smoothing and The measured spectrum of soil samples resampling to the Sentinel-2 image bandpass

#### 3.3 特征波段筛选

由于原始光谱与Cu含量间相关性较低,实验将重采样后的反射光谱进行均值规格化处理,并将Cu含量进行对数变换,有效提高了Cu含量与样品光谱之间的相关性(表2)。根据原始光谱和均值规

格化光谱与Cu含量的相关性,结合光谱的吸收特征,在10个波段中选择B2、B3、B7、B8、B12共5个波段作为特征波段,这5个波段经均值规格化处理各波段与Cu含量间的相关性均显著提高,且波段中心波长与光谱特征分析中的吸收谱带吻



合(表2)。

表2 光谱变量与Cu含量相关性分析结果

Table 2 Results of correlation analysis between spectral variables and Cu content

特征波段(nm) 相关性	原始光谱		均值规格化光谱	
	实测Cu含量	Cu含量对数	实测Cu含量	Cu含量对数
B2~490	0.45	0.48	0.68	0.65
B3~560	0.36	0.39	0.67	0.63
B7~783	-0.09	-0.08	-0.71	-0.79
B8~865	-0.10	-0.10	-0.73	-0.70
B12~2 190	0.34	0.36	0.58	0.59

### 3.4 模型建立

#### 3.4.1 偏最小二乘回归模型

偏最小二乘回归(PLSR)是一种可同时实现典型相关分析、主成分分析和多元线性回归分析等基本功能的一种多元统计数据分析方法<sup>[25]</sup>。偏最小二乘回归模型可包含更多的光谱信息从而具有较高的反演精度,可有效反演土壤重金属含量<sup>[26-28]</sup>。利用预测均方根误差(RMSEP)、决定系数( $R^2$ )、预测偏差比(RPD)对铜含量预测精度进行评价,高精度的预测模型具有高 $R^2$ 和RPD以及低RMSEP<sup>[29-31]</sup>。

当 $RPD \leq 1.4$ 时,表明模型无法对样品进行预测;  $1.4 < RPD < 2$ 时,模型可以对样品进行粗略评估;  $RPD \geq 2$ 时,模型具有极好的预测能力<sup>[30]</sup>。 $R^2 > 0.90$ ,表明模型预测精度极高;  $0.82 < R^2 \leq 0.90$ ,预测精度良好;  $0.66 < R^2 \leq 0.81$ ,模型可实现近似定量预测;  $0.50 < R^2 \leq 0.65$ ,模型仅能区分高低值;  $R^2 \leq 0.50$ ,模型无法进行有效预测<sup>[32-33]</sup>。

利用PLSR对48个样品的Cu含量值与重采样后5个特征波段的反射光谱进行建模,对研究区土壤Cu含量进行预测,回归模型如下:

$$Y = 108.2175 + B2 * 13.7848 + B3 * 2.8460 - B7 * 4.3600 - B8 * 4.2889 + B12 * 1.5516$$

建模结果如图3(a)所示,利用基本偏最小二乘法建立的反演模型决定系数 $R^2$ 仅为0.46,均方根误差RMSE较高(43.79 mg/kg),且 $RPD < 1.4$ ,表明该模型无法对样品进行预测。

#### 3.4.2 分段偏最小二乘回归模型

为解决基本PLSR无法有效反演本区土壤Cu含量的问题,提出了分段偏最小二乘回归法(Piecewise Partial Least Squares Regression, P-PLSR)。根据中国生态环境部2018年发布的土壤环境质量标准(GB 15618-2018),土壤中Cu浓度的背景值为

100 mg/kg,高于100 mg/kg则划定为污染区。本区土壤光谱分析表明,本区Cu浓度高值样品均存在2 200、2 250、2 335 nm的吸收特征,而低值样品则仅在2 200 nm存在吸收特征(图2),高值、低值样品光谱曲线形态差异较大,无法用统一的模型进行精确预测。因此将样品以100 mg/kg为界限,分为低于100 mg/kg(未污染)和高于100 mg/kg(污染)两部分,将两部分样品分别采用PLSR建立模型。

将48个样品以100 mg/kg为阈值分为两部分,包括9个高于背景值的样品(污染)以及39个低于背景值的样品(未污染)。将两组样品的Cu含量值与重采样后5个特征波段反射光谱分别利用PLSR进行建模,得到回归模型 $Y_{\text{高值}}$ 与 $Y_{\text{低值}}$ ,再将回归模型的结果进行整合。回归模型如下:

$$Y_{\text{低值}} = 65.8127 + B2 * 1.7099 + B3 * 0.8088 - B7 * 1.4339 - B8a * 1.3671 + B12 * 1.3274$$

$$Y_{\text{高值}} = 176.7753 + B2 * 6.2382 + B3 * 2.7489 - B7 * 3.0351 - B8a * 3.0564 + B12 * 2.1678$$

图3(b)为基于分段P-PLSR获得的研究区土壤Cu浓度实测值与预测值散点图。由图3(b)可知,利用分段偏最小二乘法进行建模比利用基本偏最小二乘法建模拥有更高的 $R^2$ (0.89)、更小的RMSE(20.19 mg/kg),以及更高的RPD(2.82)。表明利用分段偏最小二乘法建立的模型具有高 $R^2$ 和RPD以及低RMSE,即该模型稳定性较好、拟合程度较高,且具有极好的预测能力。

## 4 研究结果与分析

### 4.1 地面样品光谱分析结果

将由20个样品组成的验证集分为低于100 mg/kg(未污染)和高于100 mg/kg(污染)两部分,并分别基于回归模型 $Y_{\text{低值}}$ 与 $Y_{\text{高值}}$ 进行验证,得到验证集土壤Cu浓度实测与预测值散点图(图3(c))。由图3(c)可知,验证集具有较高的决定系数 $R^2$ (0.96),可以对样品进行预测评估。验证结果表明建立的分段偏最小二乘回归模型可以对研究区土壤样品Cu含量进行高精度预测。

### 4.2 土壤铜含量遥感反演

将Sentinel-2数据相应波段带入回归模型 $Y_{\text{高值}}$ 与 $Y_{\text{低值}}$ ,再将回归模型的结果以100 mg/kg为阈值进行整合,获得研究区土壤Cu元素含量空间分布结果(图4)。对结果中研究区北部、西部、南部主要建设用地进行了掩膜处理(图4中白色区域),仅部

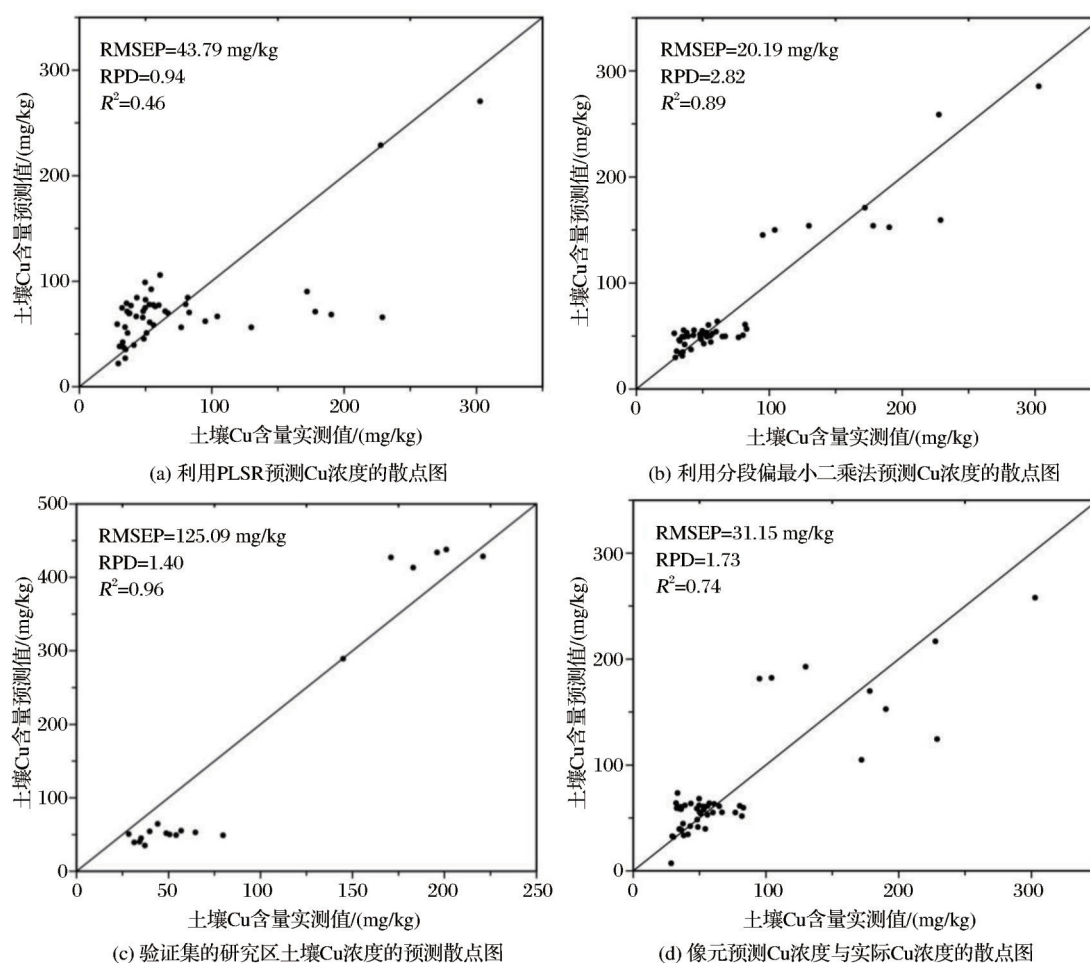


图3 土壤Cu含量反演结果

Fig.3 The scatter plot of Cu concentration

分建设用地未完全掩膜。结果中有作物生长的农田范围、山体阴影、林地和草地均反演为低值,而高值则主要为尾矿库和一些裸露农田土壤,表明结果较为合理。

根据野外取样点坐标在结果图像上提取各点的反演值(即模型预测值),获得了训练集48个反演值与相应Cu含量实测值的散点图(图3(d))。由图3(d)可以得知,利用分段偏最小二乘法建模处理Sentinel-2遥感影像所获得的像元反演值与相应Cu含量实测值散点图的 $R^2$ 为0.74, RMSE为31.15 mg/kg, RPD为1.73。结果表明,利用分段偏最小二乘法建立的模型处理Sentinel-2影像所获得的铜含量结果,具有高 $R^2$ 和RPD以及低RMSE,即该模型应用到Sentinel-2遥感影像上也同样具有较好的稳定性、较高拟合程度以及较好的预测能力,证实了利用Sentinel-2遥感影像反演土壤铜浓度的可行性。

根据研究区Sentinel-2遥感影像获得的土壤Cu含量反演结果(图4),表明铜元素在研究区内的分

布具有如下特征:①尾矿库的内部整体呈红色,说明铜含量整体偏高,明显高于周边土壤中铜元素含量,与实际情况相符;②在尾矿库周边的部分农田里,土壤铜浓度已超过标准背景值100 mg/kg,尤其在研究区的东南部,农田土壤铜污染较为严重且展布具有沿尾矿库向东南方向延伸的明显规律,与该尾矿库东南部2000年发生的溃坝事件所导致的带状延伸的污染条带吻合。

实验尝试利用Sentinel-2多光谱遥感影像和少量非均匀分布的土壤样品,获得了基于遥感影像的土壤Cu含量反演结果且精度较高,证实了利用Sentinel-2多光谱影像进行土壤铜含量反演具有可行性,本结果可以为区内土壤污染调查、监测和治理提供一定的依据。然而,实验中的样品空间分布不均匀,特别是高值样品相对较少,分为建模集、验证集后各自的高值样品更少,后续研究可以适当增加高值样品数量,确保结果在统计意义上更加稳健。

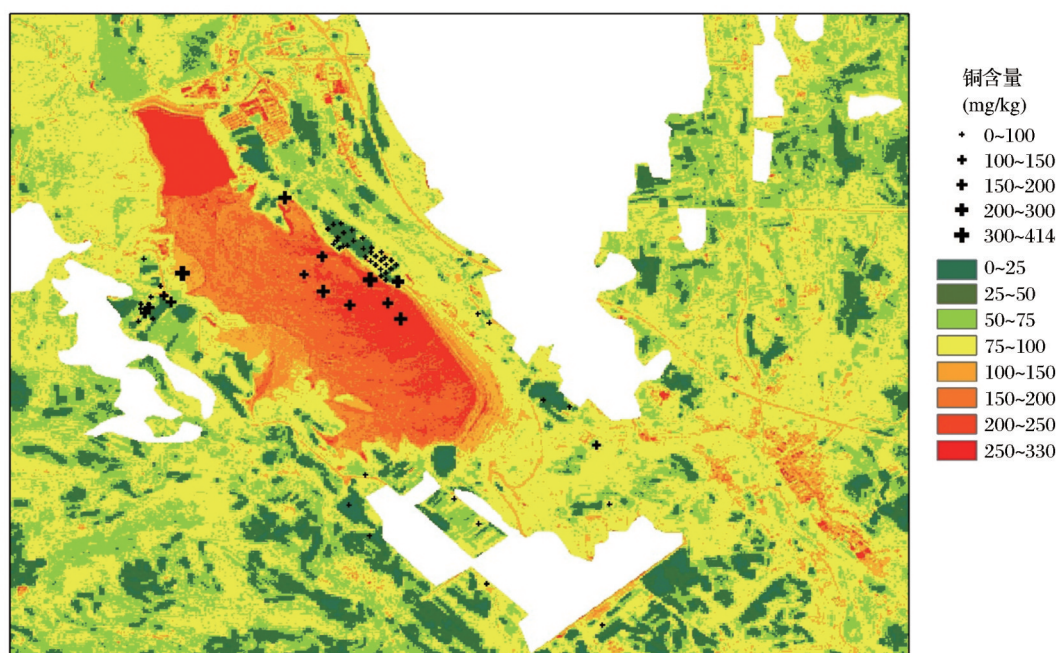


图4 研究区 Sentinel-2 遥感数据土壤 Cu 浓度反演结果图

Fig.4 Predicting result of soil Cu concentration from Sentinel-2 remote sensing data in the study area

## 5 结 论

(1)对实测土壤光谱反射率进行均值规格化处理,对 Cu 含量进行对数变换,可以有效的提高土壤 Cu 含量与反射光谱之间的相关性。其中波段 B7 的光谱反射率与 Cu 含量的相关性从-0.09 提高到-0.79。

(2)利用分段偏最小二乘法建模,有效改善了基本偏最小二乘法无法预测本区 Cu 含量的问题,提高了模型的预测精度,使最终的模型稳定性更好、拟合程度更高、预测精度更好。利用分段偏最小二乘法处理重采样至 Sentinel-2 波段的实测光谱建模反演 Cu 含量的估算精度  $R^2$  达到 0.89 (RPD=2.82)。

(3)通过分段偏最小二乘法建立的模型还可以扩展应用于 Sentinel-2 多光谱遥感影像,预测结果的  $R^2$  为 0.74,实现基于多光谱影像的土壤 Cu 含量制图,为研究区大面积、高精度土壤 Cu 元素含量调查提供可行性方法。

### 参考文献 (References):

- [1] Han Ping, Wang Jihua, Lu Anxiang, *et al.* Distribution and environment quality evaluation of heavy metals in soil in Shunyi of Beijing, China[J]. Journal of Agro-Environment Science, 2012, 31(1): 106-112.[韩平,王纪华,陆安祥,等.北京顺义区土壤重金属分布与环境质量评价[J].农业环境科学学报, 2012,31(1):106-112.]
- [2] Leenaers H, Okx J P, Burrough P A. Employing elevation data for efficient mapping of soil pollution on floodplains[J]. Soil Use Manage, 1990(6): 105-114.
- [3] Siebielec G, Mccarty G W, Stuczynski T I, *et al.* Near- and mid-infrared diffuse reflectance spectroscopy for measuring soil metal content [J]. Journal of Environmental Quality, 2004, 33(6): 2056-2069.
- [4] Viscarra Rossel R A, Walvoort D J J, McBratney A B, *et al.* Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties[J]. Geoderma, 2006, 131(1-2): 59-75.
- [5] Viscarra Rossel R A, McBratney A B. Laboratory evaluation of a proximal sensing technique for simultaneous measurement of soil clay and water content[J]. Geoderma, 1998, 85(1): 19-39.
- [6] Guerrero C, Viscarra Rossel R A, Mouazen A M. Special issue 'diffuse reflectance spectroscopy in soil science and land resource assessment'[J]. Geoderma, 2010, 158(1-2): 1-2.
- [7] He Junliang, Cui Junli, Zhang Shuyuan, *et al.* Hyperspectral estimation of heavy metal Cu content in soil based on partial least square method[J]. Remote Sensing Technology and Application, 2019, 34(5): 998-1004.[贺军亮,崔军丽,张淑媛,等.基于偏最小二乘的土壤重金属铜含量高光谱估算[J].遥感技术与应用, 2019,34(5):998-1004.]
- [8] Chen T, Chang Q, Clevers J G P W, *et al.* Rapid identification of soil cadmium pollution risk at regional scale based on visible and near-infrared spectroscopy[J]. Environmental Pollution, 2015(206): 217-226.
- [9] Wang J, Cui L, Gao W, *et al.* Prediction of low heavy metal concentrations in agricultural soils using visible and near-infrared reflectance spectroscopy[J]. Geoderma, 2014(216): 1-9.
- [10] Wu Y, Chen J, Wu X, *et al.* Possibilities of reflectance Spec-



- troscopy for the assessment of contaminant elements in suburban soils[J]. *Applied Geochemistry*, 2005, 20(6): 1051-1059.
- [11] Kooistra L, Wehrens R, Leuven R S E W, *et al.* Possibilities of visible - near-infrared spectroscopy for the assessment of soil contamination in river floodplains[J]. *Analytica Chimica Acta*, 2001, 446(1-2): 97-105.
- [12] Chodak M, Ludwig B, Khanna P, *et al.* Use of near infrared spectroscopy to determine biological and chemical characteristics of organic layers under spruce and beech stands[J]. *Journal of Plant Nutrition & Soil Science*, 2002, 165(1): 27-33.
- [13] Siebielec G, Mccarty G W, Stuczynski T I, *et al.* Near- and mid-infrared diffuse reflectance spectroscopy for measuring soil metal content [J]. *Journal of Environmental Quality*, 2004, 33(6): 2056-2069.
- [14] Ferreira E C, Milori D M B P, Ferreira E J, *et al.* Artificial neural network for Cu quantitative determination in soil using a portable Laser induced breakdown spectroscopy system [J]. *Spectrochimica Acta Part B Atomic Spectroscopy*, 2008, 63(10): 1216-1220.
- [15] Laakso K, Peter J M, Rivard B, *et al.* Short-wave infrared spectral and geochemical characteristics of hydrothermal alteration at the Archean Izok Lake Zn-Cu-Pb-Ag Volcanogenic massive sulfide deposit, Nunavut, Canada: application in exploration target vectoring[J]. *Economic Geology*, 2016, 111(5): 1223-1239.
- [16] Nawar S, Buddenbaum H, Hill J. Digital mapping of soil properties using multivariate statistical analysis and ASTER data in an arid region[J]. *Remote Sensing*, 2015, 7(2): 1181-1205.
- [17] Malley D F, Williams P C. Use of near-infrared reflectance spectroscopy in prediction of heavy metals in freshwater sediment by their association with organic matter[J]. *Environmental Science & Technology*, 1973, 31(12): 3461-3467.
- [18] Ge Liang, Wang Bin, Zhang Liming. Partial least squares based band selection for hyperspectral imagery[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2011, 23(11): 1844-1852. [葛亮, 王斌, 张立明. 基于偏最小二乘法的高光谱图像波段选择[J]. *计算机辅助设计与图形学学报*, 2011, 23(11): 1844-1852.]
- [19] Shamsoddini A, Raval S, Taplin R. Spectroscopic analysis of soil metal contamination around a derelict mine site in the Blue Mountains, Australia, 2014, II-7(1): 75-79.
- [20] Kemper T, Sommer S. Estimate of heavy metal contamination in soils after a mining accident using reflectance spectroscopy. [J]. *Environmental Science & Technology*, 2002, 36(12).
- [21] Xu Binbin, Ji Genshan, Zhu Yonghao. A preliminary research of geographic regionalization of China land background and spectral characteristics [J]. *Remote Sensing of Environment*, 1991, 6(2): 142-151. [徐彬彬, 季耿善, 朱永豪. 中国陆地背景和土壤光谱反射特性的地理分区的初步研究[J]. *环境遥感*, 1991, 6(2): 142-151.]
- [22] Bendor E, Banin A. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties [J]. *Soil Science Society America Journal*, 1995(59): 364-372.
- [23] Xu B, Li D, Shi X. A preliminary study on identification of clay minerals in soils with reference to reflectance spectra [J]. *Pedosphere*, 1995(2): 135-142.
- [24] Leone A P, Sommer S. Multivariate analysis of laboratory spectra for the assessment of soil development and soil degradation in the Southern Apennines (Italy) [J]. *Remote Sensing of Environment* 2000, 72(3): 346-359.
- [25] Shi T, Chen Y, Liu Y, *et al.* Visible and near-infrared reflectance spectroscopy—an alternative for monitoring soil contamination by heavy metals [J]. *Hazardous Materials*, 2014(265): 166-176.
- [26] Wold S, Martens H, Wold H. The Multivariate calibration problem in chemistry solved by the PLS method [J]. *Lecture Notes in Mathematics*, 1983(973): 286-293.
- [27] Groenigen J W V, Muters C S, Horwath W R, *et al.* NIR and DRIFT-MIR spectrometry of soils for predicting soil and crop parameters in a flooded field [J]. *Plant & Soil*, 2003, 250(1): 155-165.
- [28] Moros J, Martinezsánchez M J, Pérezsirvent C, *et al.* Testing of the region of murcia soils by near infrared diffuse reflectance spectroscopy and chemometrics [J]. *Talanta*, 2009, 78(2): 388.
- [29] Ren Hongyan, Zhuang Dafang, Qiu Dongsheng, *et al.* Analysis of visible and near-infrared spectra of as-contaminated soil in croplands beside mines [J]. *Spectroscopy and Spectral Analysis*, 2009, 29(1): 114-118. [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.]
- [30] Saeys W, Mouazen A M, Ramon H. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy [J]. *Biosystems Engineering*, 2002, 91(4): 393-402.
- [31] Bellon-Maurel V, Fernandez-Ahumada E, Palagos B, *et al.* Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy [J]. *TrAC Trends in Analytical Chemistry*, 2010, 29(9): 1073-1081.
- [32] Saeys W, Mouazen A M, Ramon H. Potential for onsite and online analysis of pig manure using visible and near infrared reflectance spectroscopy [J]. *Biosystems Engineering*, 2005, 91(4): 393-402.
- [33] Vohland M, Besold J, Hill J, *et al.* Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy [J]. *Geoderma*, 2011, 166(1): 198-205.

## A Remote Sensing Estimation Method for Heavy Metals in Soil based on Piecewise Partial Least Squares Model

Yin Fang<sup>1</sup>, Feng Kai<sup>2</sup>, Wu Mengmeng<sup>2</sup>, Bai Dezhen<sup>3</sup>, Wang Rui<sup>3</sup>,  
Zhou Yuanyuan<sup>3</sup>, Yin Chuntao<sup>2</sup>, Yin Cuijing<sup>2</sup>, Liu Lei<sup>2</sup>

(1.School of Land Engineering, Chang'an University, Xi'an 710054, China;

2.School of Earth Sciences and Resources, Chang'an University, Xi'an 710054, China;

3.Qinghai Research and Design institute of Environmental Sciences, Xining 810000, China)

**Abstract:** Heavy metals in soil are among the most harmful environmental pollutants due to their toxicity. Detecting and mapping the distribution of heavy metal using remote sensing technique is inexpensive and efficient. In this study, Sentinel-2 multispectral data and field spectroscopy were adopted to estimate soil copper (Cu) concentrations of the tailing reservoir of Tongkuangyu Copper deposit, Shanxi Province, China and the surrounding farmland soil. Sixty-eight soil samples were collected and their reflectance spectra were used to estimate Cu concentration in soil. Spectral index applicable to the prediction of Cu contents in soil was derived, united with piecewise partial least square regression (P-PLSR), the soil Cu contents were estimated. The coefficient of determination ( $R^2$ ) and residual prediction deviation (RPD) for the model developed using lab-measured spectra were 0.89 and 2.81. The model was applied to the Sentinel-2 multispectral data and the spatial distribution map of Cu content was predicted with relatively high  $R^2$  (0.83) and RPD (1.56). The result could facilitate the development of remediation strategies in terms of environmental protection. Sentinel-2 multispectral data, due to its high spatial resolution (10 m, 20 m and 60 m), and large swath width (290 km), could provide an alternative method for large-scale soil environment monitoring through reasonable selection of sensitive bands.

**Key words:** Soil heavy metals; Multi-spectral remote sensing; Piecewise partial least squares; Quantitative inversion