

武汉大学学报(信息科学版)

Geomatics and Information Science of Wuhan University

ISSN 1671-8860, CN 42-1676/TN

《武汉大学学报(信息科学版)》网络首发论文

题目: 结合野外光谱与实验室光谱的土壤 Pb 含量反演方法研究
作者: 张霞, 丁松滔, 岑奕, 孙伟超, 王晋年
DOI: 10.13203/j.whugis20200386
收稿日期: 2020-11-15
网络首发日期: 2021-08-23
引用格式: 张霞, 丁松滔, 岑奕, 孙伟超, 王晋年. 结合野外光谱与实验室光谱的土壤 Pb 含量反演方法研究. 武汉大学学报(信息科学版).
<https://doi.org/10.13203/j.whugis20200386>



网络首发: 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式(包括网络呈现版式)排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

出版确认: 纸质期刊编辑部通过与《中国学术期刊(光盘版)》电子杂志社有限公司签约, 在《中国学术期刊(网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊(网络版)》是国家新闻出版广电总局批准的网络连续型出版物(ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

DOI: 10.13203/j.whugis20200386

引文格式: ZHANG Xia, DING Songtao, CEN Yi, et al. Soil Heavy Metal Pb Content Inversion Method by Combining Field with Laboratory Spectra[J]. *Geomatics and Information Science of Wuhan University*, 2021, DOI: 10.13203/j.whugis20200386 (张霞, 丁松滔, 岑奕, 等. 结合野外光谱与实验室光谱的土壤 Pb 含量反演方法研究[J]. 武汉大学学报·信息科学版, 2021, DOI: 10.13203/j.whugis20200386)

结合野外光谱与实验室光谱的土壤 Pb 含量反演方法研究

张霞¹ 丁松滔^{1,2} 岑奕¹ 孙伟超¹ 王晋年^{3,4}

1 中国科学院空天信息创新研究院, 北京 100101

2 中国科学院大学, 北京 100049

3 广州大学地理科学与遥感学院, 广州 510006

4 中科遥感(深圳)卫星应用创新研究院有限公司, 深圳 518000

摘要: 重金属污染日益加剧, 重金属在土壤中的聚集不仅破坏了生态平衡, 也对人类的健康生活造成了影响, 因此快捷、准确地获取土壤中的重金属含量成为土壤污染监测与治理的重要环节。高光谱遥感技术的发展, 使得快速低成本反演土壤重金属含量成为可能。针对野外光谱受环境因素(土壤粒径、含水量等)的影响, 且现有研究中普遍存在样本量不足的问题, 本文提出结合野外光谱与实验室光谱构建机理模型的方法, 首先, 采用直接校正(direct standardization, DS)算法对野外光谱进行环境因素校正, 其次, 通过引入实验室光谱联合建模的方式, 提高样本的差异性, 进一步地, 提取铁氧化物特征谱段用于建模以增加反演的机理性。利用河北雄安一般农作区的 70 个土壤样本野外光谱数据研究表明, 未经 DS 校正的野外光谱全谱段单独建模反演精度 R^2 仅为 0.22, 而本文方法的反演精度 R^2 可达 0.91, 模型具有出色的估算能力, 表明在去除环境因素对野外光谱影响基础上综合利用野外光谱与实验室光谱的铁氧化物特征谱段建模能够显著提高 Pb 含量反演精度。

关键词: 高光谱遥感; 土壤重金属; DS 算法; 铁氧化物特征谱段; 环境因素去除

0 引言

近年来, 土壤重金属污染已经成为一个严重的环境问题, 尤其是在城市快速扩张、工业迅速发展的国家和地区^[1], 研究表明, 我国耕地土壤重金属污染的概率约为

收稿日期:2020-11-15
项目资助:国家重点研发计划(2017YFB0503800); 高分辨率对地观测系统专项(30-H30C01-9004-19/21)。
第一作者简介:张霞, 研究员, 主要从事高光谱定量遥感研究。 zhangxia@radi.ac.cn
通讯作者简介:丁松滔, 博士生。 dingst425@163.com

16.67%，土壤重金属污染的面积大约占据耕地总面积的六分之一。Pb 是土壤中常见的一种重金属污染物，渗透进土壤中后移动性较差，残留时间长，不仅影响农作物质量，人类的健康^[2]，而且会影响整个生态系统的安全。因此，开展土壤重金属 Pb 含量调查和监测方法研究具有重要意义。高光谱遥感通过获取连续且精细的土壤反射光谱，能够体现出土壤成分的细微变化，具有低成本大范围快速监测土壤重金属状况的潜力^[3]。

迄今为止，大多数研究聚焦于实验室光谱建模，很少用到野外光谱进行建模^[4]。野外光谱会受到环境因素（土壤粒径和含水量等）的影响，且由于土壤样本难以获取，样本数一般偏少，导致数据差异性不足。目前去除环境因素的光谱转换算法主要包括：Spiking^[5]、外部参数校正^[6]（External Parameter Orthogonalization, EPO）算法以及直接校正（Direct Standardization, DS）算法，这些环境因素除去算法在土壤有机质（碳）和粘土含量等属性参量高光谱反演方面有较好的应用，但在土壤重金属反演方面的应用还鲜有触及^[7]。

现有研究大多以全谱段建模为主，未能充分利用土壤组分对土壤重金属的吸附作用机理。土壤光谱曲线上很难直接探测到土壤重金属的自身光谱特征，直接通过土壤重金属的光谱响应特征反演重金属含量不可行^{[8][9]}，然而通过借助重金属元素与土壤光谱活性物质（有机质、黏土矿物、铁氧化物）之间的吸附或赋存关系，可以间接反演土壤重金属元素含量^[10]，此为土壤重金属的高光谱间接反演机理。Sun 等^[11]提出基于土壤光谱活性物质特征谱段的重金属反演方法，提取对 Ni 起主要吸附作用的有机质和黏土矿物的特征谱段反演 Ni 含量，将精度 R^2 由全谱段建模的 0.67 提升到 0.98。Zhang 等^[12]提取有机质的特征谱段对 Cd 含量反演，显著提升了野外光谱的估算精度。但这种方法对土壤其他重金属反演的有效性还有待验证。

本文以河北雄安一般农作区为例，利用野外光谱开展重金属 Pb 含量反演的研究，本文在去除环境因素对野外光谱影响的基

础上，通过结合野外光谱与实验室光谱建模的方法，提高样本的差异性，以期提高反演精度。同时，研究中提取了对重金属 Pb 起主要吸附作用的土壤光谱活性物质的特征谱段作为模型的输入，探讨了特征谱段建模相对于全谱段建模的有效性。

1 研究数据与预处理

1.1 研究数据

土壤样本采集的时间为 2018 年 9 月，采集地点主要在河北雄安新区的雄县和安新县的一般农作区农田。野外光谱采集使用的是 SVC HR1024i 地物光谱仪，测量波长范围为 350~2500nm，共获取 70 个样点的土壤样本及配套野外土壤光谱。将土壤样本在实验室风干研磨后分别过 20 目和 100 目筛制成标准样，一部分用于电感耦合等离子体质谱仪 ICP-MS（7500a, Agilent Technologies, USA）测定 Pb 含量，另一部分用于测定实验室光谱，测量仪器同野外光谱采集。

1.2 数据预处理

野外光谱测量时，大气中的水汽在 1400nm 和 1900nm 处存在强吸收，且 1900~2500nm 区间的光谱存在较严重噪声^[13]。为去除大气水汽等噪声影响，并保留尽可能多的光谱区间，将 1800nm 之后的光谱波段从土壤野外光谱中剔除，对 350~1800nm 光谱使用分段 SG 滤波进行去噪处理，其中 890~1020nm 和 1330~1520nm 为中等噪声，采用窗口大小为 15 的二次多项式，其余区间为轻度噪声，采用窗口大小为 7 的二次多项式。利用去除噪声后的 350~1800nm 的野外光谱进行土壤铁氧化物特征谱段提取及重金属含量估算。

实验室光谱在 350~1800nm 区间存在不同程度的光谱噪声，同样采取分段 SG 滤波处理，350~500nm 和 1000~1800nm 为中度噪声区间，采用窗口大小为 15 的二次多项式进行噪声去除；剩余的 500~900nm 为轻度噪声区间，采用窗口大小为 7 的二次多项式进行噪声去除。

2 研究方法

2.1 铁氧化物特征谱段提取

土壤中对 Pb 起主要吸附作用的物质为铁氧化物^[14], Galv ǎn^[15]和徐彬彬^[16]的研究表明,铁氧化物的吸收特征在 500nm 和 950nm 附近的吸收峰。因此,本文从雄安 70 条有效土壤光谱中提取 500nm 和 950nm 为中心的铁氧化物吸收峰,提取波段区间为 $[B_m - W/4, B_m + W/4]$, 其中 B_m 是最大的吸收波段, W 是吸收区域的宽度,最终铁氧化物特征谱段提取出的光谱波段范围为 450.7~523.2nm 以及 914.2~1027.9nm。

2.2 DS 转换算法

DS 算法是一种常用的环境因素去除算法,它的目的是为了通过转换数据集,根据对应的野外光谱和实验室光谱之间的关系,计算出转换矩阵并对野外光谱进行转换,去除野外光谱上环境因素的影响,如水分、颗粒大小和温度的影响^[17]。DS 算法计算步骤如下:

野外光谱 (X_{field}) 和实验室光谱 (X_{lab}) 的矩阵大小都是 $m \times p$, m 是转换光谱的数量, p 是波段的数量。DS 算法的模型如式 (1) 所示:

$$X_{lab} = X_{field}B + E \quad (1)$$

式中, B 是有 $p \times p$ 个未知参数构成的转换矩阵,它由实验室光谱和野外光谱矩阵共同决定; E 为残差矩阵,表示为:

$$E = \lambda d_s^T \quad (2)$$

式中, d_s^T 由于基线差异产生的 $p \times 1$ 的矩阵, λ 是所有列向量值都为 1 的 $p \times 1$ 的矩阵,因此公式 (1) 可以被写为:

$$X_{lab} = X_{field}B + \lambda d_s^T \quad (3)$$

为了计算未知的转换矩阵 B ,首先需要定义一个大小为 $m \times m$ 的中心化矩阵 C_m ,可表示为:

$$C_m = I_m - \left(\frac{1}{m}\right) \lambda * \lambda^T \quad (4)$$

其中, I_m 是一个 $m \times m$ 的单位矩阵。然后将 (3) 式的两边同时乘以 C_m , 因 $C_m \times \lambda d_s^T$ 等于 0, 并且 $C_m \times X_{field}$ 和 $C_m \times X_{lab}$ 分别为野外

光谱矩阵和实验室光谱矩阵的中心化矩阵, 分别用 \bar{X}_{lab} 和 \bar{X}_{field} 来对其进行表示后, 公式 (3) 可以表示为:

$$\bar{X}_{lab} = \bar{X}_{field}B \quad (5)$$

经过最小二乘后变为:

$$B = \bar{X}_{field}^+ \bar{X}_{lab} \quad (6)$$

式中, $+$ 表示 \bar{X}_{field} 的广义逆矩阵。

基于转换矩阵 B , 可以计算出残差矩阵 E 的值, 根据公式 (2), 首先将式 (3) 两边同时乘以 $(1/m) \lambda^T$ 来计算 d_s :

$$d_s = \bar{X}_{lab}^T - B^T \bar{X}_{field}^T \quad (7)$$

式中, \bar{X}_{field} 和 \bar{X}_{lab} 分别是由野外光谱 (X_{field}) 和实验室光谱 (X_{lab}) 对每列求均值得到的 $1 \times p$ 的矩阵。最终野外光谱可以根据下式进行转换, 得到去除环境因素影响的野外光谱:

$$X_{field}' = X_{field}B + E \quad (8)$$

在 DS 转换的过程中, 需要选择有代表性和差异性的土壤样本作为转换集, 用于计算转换矩阵, 因此本文用 Kennard-Stone 算法来筛选样本, 具体步骤如下^[18]:

- 1) 首先计算两两样本之间距离, 选择距离最大的两个样本;
- 2) 然后分别计算剩余样本与已选两样本之间的距离;
- 3) 对于每个剩余样本而言, 计算其与已选各样本之间的最短距离, 选择这些最短距离中相对最大的距离所对应的样本, 作为新入选的样本;
- 4) 重复步骤 3), 直至所选样本的个数等于事先设定的数目为止。

2.3 GA-PLSR 建模

在高光谱建模中, 波段选择有利于降低模型复杂度, 同时提高模型估算精度。遗传算法 (genetic algorithm, GA) 是一种随机的全局寻优算法, 在偏最小二乘回归 (partial least squares regression, PLSR) 建模中认为是一种有效的波段选择算法^[19]。GA-PLSR 已经被用于土壤反射光谱估算有机碳和重金属含量研究^{[20][21]}, 取得了降低模型复杂度和提高模型估算精度的效果。因此, 本文采用 GA-PLSR 构建土壤重金属含量反演模型。

参考已有 GA-PLSR 建模研究^[22],将 GA 参数设置为:染色体个数 20,迭代次数 1000,代际间隙 90%,基因变异概率 10%。本文中用均方根误差 (Root Mean Square Error, RMSE) 作为 GA 算法中适应度的评判标准,即目标函数,用 RMSE 值小的个体替换父一代中 RMSE 值大的个体,不断反馈产生新解进行下一轮迭代,直到达到最大的迭代次数,最终将 RMSE 最小的个体作为最优解输出,认为此时建立的 PLSR 为最优的模型。

2.4 模型评价标准

精度评定采用预测均方根误差 (root mean square error of prediction, RMSEP)、剩余估算偏差 (residual predictive deviation, RPD) 和 决 定 系 数 (coefficient of determination, R^2) 3 个评价指标。3 个评价指标, RMSEP 值越小, RPD 值越大, R^2 值越接近 1, 说明反演模型的精度越高; 反之, RMSEP 值越大, RPD 值越小, R^2 值越小, 说明反演模型的精度越低, 本文模型优劣参考现有的土壤属性含量高光谱估算的评价标准^{[23][24]}: 出色模型, $R^2 > 0.9$; 良好模型, $0.9 > R^2 > 0.8$; 近似模型, $0.8 > R^2 > 0.65$; 具有一定估算能力, $0.65 > R^2 > 0.50$; 不具备估算能力, $0.50 > R^2$ 。

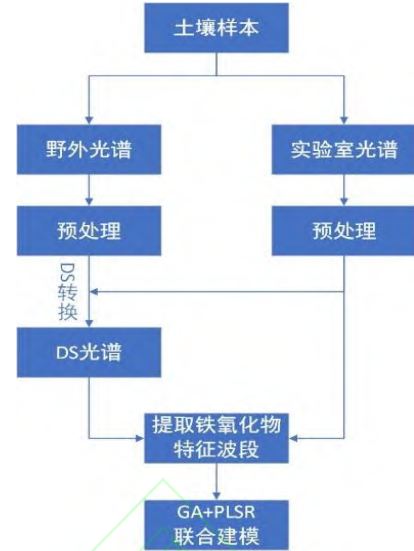


图 1 方法技术路线

Fig.1 Flowchart of the proposed method

3 实验结果与分析

3.1 样本集划分

本文将经过数据预处理后的 70 个样本按照 2: 1 的比例划分为训练集和测试集, 训练集包括 47 个样本, 测试集包括 23 个样本, 训练集用于进行 PLSR 模型建立时训练模型, 测试集用于带入训练后的模型中估算重金属 Pb 的含量, 并计算评价指标, 评估模型估算能力。样本集的数据统计如表 1 所示, 从表中可看出划分的训练集和测试集与总样本集的统计特征基本一致。

表 1 样本集 Pb 含量统计 (mg/kg)

Table 1 Statistics of Pb content in sample set (mg/kg)

样本集	最小值	最大值	平均值	标准差
总样本集	3.57	28.59	15.59	5.80
训练集	3.57	28.59	15.46	5.71
测试集	3.96	26.06	15.85	6.10

3.2 DS 转换集大小设定

DS 转换算法涉及到转换集的大小设定, 而转换集选择的样本个数影响反演模型的精度, 本文实验了不同转换集大小对应的反演精度, 结果如图 2 所示 (由于三种精度指标同步变化, 此处仅展示 R^2), 从图中可以看出, 当转换集的样本个数为 30 时, 反演模型的精度最高, 因此本文选取 30 个样本

的转换集进行转换矩阵计算。

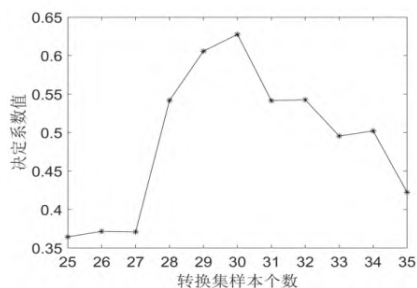


图 2 不同转换集大小对应建模精度
Fig.2 The accuracy change with conversion dataset size

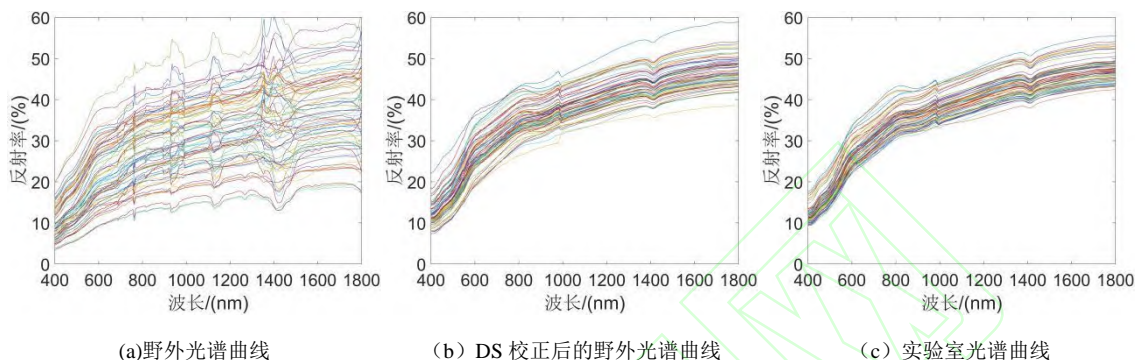


图 3 反射光谱曲线
Fig.3 Reflectance spectrum

将转换集中对应的野外光谱和实验室光谱按照 2.2 节中的步骤与公式进行计算得到转换矩阵,用转换矩阵对所有 70 个样本的野外光谱进行处理,获得 DS 校正后的野外光谱。校正效果如图 3 所示

由图 3 可看出,经过 DS 转换后的野外光谱,波形与实验室反射光谱相似,光谱也变得更为平滑,表明通过 DS 转换,能够较好地去除环境因素(土壤粒径、含水量等)对野外光谱的影响。

3.3 DS 校正及联合建模精度分析

本节分别采用原始野外光谱、DS 校正后的野外光谱,联合 DS 校正的野外光谱与实验室光谱进行建模,比较三种光谱的重金属 Pb 估算精度,以及全谱段建模与铁氧化物特征谱段建模的反演精度。

为了避免实验的偶然性,每组建模实验均运算五次,取精度最优值作为模型的反演精度。

3.3.1 不同光谱数据建模精度对比分析

图 4 为未经 DS 校正的野外光谱单独建模的结果, R^2 仅为 0.5137,属于具有一定估算能力的模型,可见直接利用野外光谱进行

重金属 Pb 含量反演时精度较低。

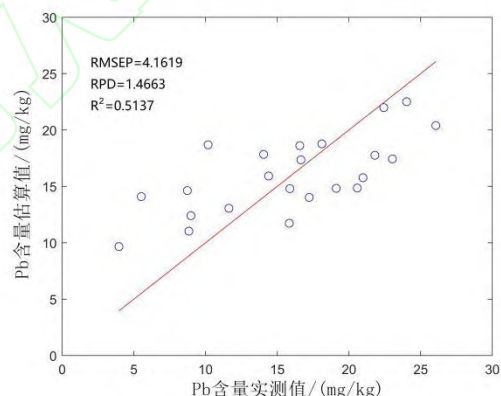


图 4 原始野外光谱 Pb 含量估算结果

Fig.4 Pb content estimation results by original field spectra

图 5 为 DS 校正后的野外光谱独立建模结果, R^2 由原始野外光谱建模的 0.5137 提高到了 0.6276,精度有了明显提升,说明 DS 算法去除野外光谱的环境因素影响的有效性。但是,所建模型仍属于具有一定估算能力的模型,从散点图中可看出,部分样本的预测值和测量值之间还存在着较大的差距,精度尚有待提升。

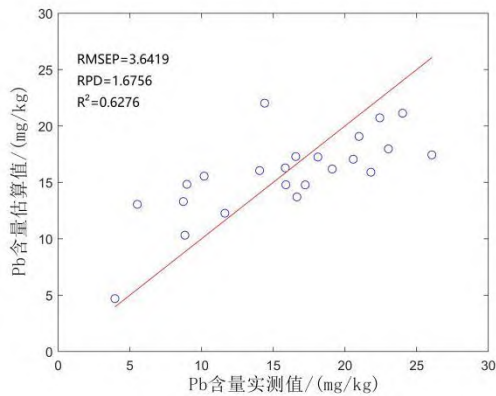


图 5 DS 野外光谱 Pb 含量估算结果
Fig.5 Pb content estimation results by DS corrected field spectra

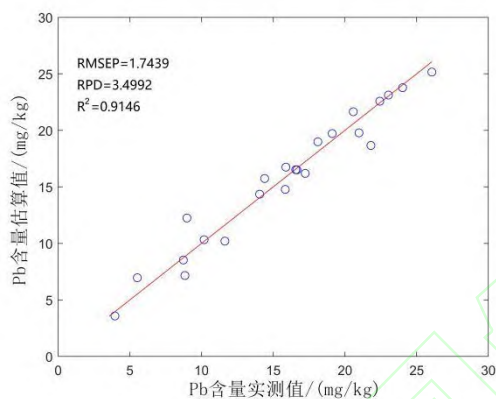


图 6 DS 野外光谱与实验室光谱联合建模的 Pb 含量估算结果

Fig.6 Pb content estimation results by combining DS

corrected field spectra with laboratory spectra

结合 DS 校正的野外光谱和实验室光谱的建模结果如图 6，样本点紧密分布于拟合线附近，精度 R^2 可以达到 0.9146，属于出色模型，由此可见，通过联合建模增强了样本的差异性从而使得反演精度得到大幅提高。

3.3.2 铁氧化物特征谱段建模的有效性分析

表 2 中展示了三类光谱数据分别以全谱段和铁氧化物特征谱段作为 GA-PLSR 输入时的反演精度。除了 DS 校正野外光谱的全谱段与铁氧化物特征谱段反演精度相当外，其他两类光谱数据的铁氧化物特征谱段反演精度均高于全谱段反演精度，根据 2.4 节的模型评价标准，对于未进行 DS 校正的野外光谱而言，通过铁氧化物特征谱段提取，建模精度 R^2 从 0.22 提升到 0.5137，精度得到了大幅的提升，反演模型从不具备估算能力提升为具有一定估算能力；联合建模的精度 R^2 从 0.8952 提升到 0.9146，使得反演模型从良好模型提升为出色模型，证明了提取 500nm 和 950nm 的铁氧化物特征谱段对于提高土壤重金属 Pb 反演精度的有效性。

表 2 反演精度对比

Table 2 Comparison of modeling accuracy

光谱波段	光谱类型	RMSEP	RPD	R^2
全谱段	野外光谱	5.2711	1.1577	0.2200
	DS 野外光谱	3.4934	1.7468	0.6574
	联合光谱			
	(DS 野外光谱与实验室光谱)	1.9324	3.1580	0.8952
铁氧化物特征谱段	野外光谱	4.1619	1.4663	0.5137
	DS 野外光谱	3.6419	1.6756	0.6276
	联合光谱			
	(DS 野外光谱与实验室光谱)	1.7439	3.4992	0.9146

4 结论

本文针对环境因素对土壤野外光谱乃至土壤重金属含量高光谱遥感反演的影响、

土壤样本不足及全谱段建模机理性不足问题，以土壤重金属 Pb 为例，开展了基于野外光谱的土壤重金属含量反演方法研究，首先采用 DS 算法去除环境因素对野外光谱的影响，在此基础上联合野外与实验室光谱，提取对 Pb 起主导吸附作用的铁氧化物的特

征谱段用于反演建模,在增强样本的差异性的同时,增强反演的机理性。以雄安一般农作区土壤数据集为例,分析了该方法对提高Pb含量反演精度的有效性。

研究表明,DS转换算法能够去除环境因素对野外光谱的影响,与原始野外光谱建模相比,反演精度 R^2 从0.5137提升到了0.6276;联合野外光谱与实验室光谱建模的方法,有利于提高样本的差异性,引入噪声极少的先验知识后,反演精度得到大幅提升, R^2 达到0.9146,使得反演模型达到了出色模型的标准。

除了DS校正的野外光谱,全谱段反演精度与铁氧化物特征谱段反演精度大致相当外,其他光谱类型的结果皆为铁氧化物特征谱段反演精度显著优于全谱段反演精度,表明通过提取对重金属Pb起主要吸附作用的铁氧化物的特征谱段,降低了数据的冗余

度,提高了反演的机理性,对于提高模型反演精度是有效的,且本文提取的铁氧化物波段在500nm和950nm的吸收特征对于重金属Pb来说确实是敏感有效的特征谱段。

由于野外光谱与同步或准同步获取的航空/卫星高光谱图像在土壤自然状态上最为接近,基于野外光谱的反演方法研究将为高光谱遥感图像反演土壤重金属含量奠定基础。本文方法通过环境因素去除以及基于铁氧化物特征谱段建模,联合野外与实验室光谱显著提高了土壤Pb含量反演精度,未来可尝试将该方法应用于同一研究区或同类的多个研究区的多源光谱数据(实验室、野外、图像光谱),探讨提高基于高光谱图像反演土壤重金属含量精度的可行性。

参考文献

- [1]Wong, S., Li, X., Zhang, G., Qi, S., Min, Y. Heavy metals in agricultural soils of the Pearl River Delta, South China[J]. Environmental Pollution,2002,119(1), 33-44.
- [2]Gmochowska, W., Pietranik, A., Tyszk, R., Ettler, V., MihaljevičGmochowska, W., Pietranik, A., Tyszk, R., Ettler, V., Mihaljevič. Distribution of Pb, Cd and Hg in Wroclaw location of Pb, Cd and Hg in W., Tyszk, R., Ettler, V. [J]. Mihalje Geochemistry, 2019, 79(3), 434-445.
- [3] Khosravia, V., DoulatiArdejani, F., Yousefic, S., Aryafarc, A., 2018. Monitoring soil lead and zinc contents via combination of spectroscopy with extreme learning machine and other data mining methods[J]. Geoderma, 318, 29-41.
- [4] PENG Xiaoting, GAO Wenxiu, WANG Junje. Inversion of soil parameters from hyperspectra based on continuum removal and partial least squares regression[J]. Geomatics and Information Science of Wuhan University, 2014,39 (7) :862-866.(彭小婷, 高文秀,王俊杰. 基于包络线去除和偏最小二乘的土壤参数光谱反演[J]. 武汉大学学报 汉大信息科学版, 2014, 39(7): 862-866.)
- [5] Guerrero, C., Zornoza, R., Gomez, I. & Mataix-Beneyto, J. 2010. Spiking of NIR regional models using samples from target sites: effect of model size on prediction accuracy[J]. Geoderma, 158, 66-77.
- [6]Ge, Y., Morgan, C.L.&Ackerson, J.P. 2014. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact[J]. Geoderma, 221,61, 22
- [7] Zou, B.,JiangX.,Feng H.,Tu,Y.,Tao C. Multisource spectral-integrated estimation of cadmium concentrations in soil using a direct standardization and Spiking algorithm.[J]. The Science of the Total Environment, 2020,701.[Doi: 10.1016/j.scitotenv.2019.134890]
- [8] Rathod, P.H., Rossiter, D.G., Noomen, M.F., van der Meer, F.D., 2013. Proximal spectral sensing to monitor phytoremediation of metal-contaminated soils[J]. International Journal of Phytoremediation 15(5), 405-426.
- [9] Wu, Y., Chen, J., Ji, J., Gong, P., Liao, Q., Tian, Q., Ma, H., 2007. A mechanism study of reflectance spectroscopy for investigating heavy metals in soils[J]. Soil Science Society of America Journal 71(3), 918.
- [10] Malley, D., Williams, P., 1997. Use of near-infrared reflectance spectroscopy in prediction of heavy metals in freshwater sediment by their association with organic matter[J]. Environmental Science & Technology 31(12), 3461-3467.
- [11] Sun, W., Xia, Z., Sun, X., Sun, Y., & Yi, C. Predicting nickel concentration in soil using reflectance spectroscopy associated with organic matter and clay minerals[J]. Geoderma, 2018, 327, 25-35. [DOI: 10.1016/j.geoderma.2018.04.019]

- [12] Zhang,X., Sun, W., Cen, Y., Zhang, L., Wang, N. Predicting cadmium concentration in soils using laboratory and field reflectance spectroscopy[J], Science of Total Environment, 2019, 650: 321-334.
- [13] Kooistra, L., Salas, E. A. L., Clevers, J., Wehrens, R., Leuven, R., Nienhuis, P. H., &Buydens, L. M. C. Exploring field vegetation reflectance as an indicator of soil contamination in river floodplains[J]. Environmental Pollution, 2004, 127(2), 281-290. [Doi: 10.1016/s0269-7491(03)00266-5]
- [14] Ge Yunsi. Study on the influence of iron oxides on the adsorption characteristics of Pb in soils[D]. East China Norma University, 2016.(葛云思.土壤中铁氧化物对铅吸附特性的影响研究[D].华东师范大学,2016.)
- [15]Galvve, L.S., Vitorello, G Variability of Laboratory Measured Soil Lines of Soils from Southeastern Brazil[J]. Remote Sensing of Environment,1998,63(2), 166-181.
- [16]Xu B B, Ji G S, Zhu Y H. A preliminary research of geographic regionalization of China land background and spectral reflectance characteristics of soil[J]. Remote Sensing of Environment, 1991, 6 (5) :142-152.(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 (5) :142-152.)
- [17] Ji, W., et al. Accounting for the effects of water and the environment on proximally sensed vis-NIR soil spectra and their calibrations[J]. European Journal of Soil Science 66. 3 (2015): 555-565.
- [18] Zhou Bin, Xu Yulong, Jiang Xiaolu, Tao Chao, Zhou Lai, Xiong Liwei. Estimation of Cd content in soil using combined laboratory and field DS spectroscopy[J]. Spectroscopy and Spectral Analysis, 2019, 39(10): 3223-3231.(邹滨, 涂宇龙, 姜晓璐, 陶超, 周莱, 熊立伟. 土壤 Cd 含量实验室与野外 DS 转换光谱联合反演[J]. 光谱学与光谱分析, 2019, 39(10): 3223-3231.)
- [19]Leardi, R., Amparo, L.G. Genetic algorithms applied to feature selection in PLS regression: how and when to use them[J]. Chemometrics & Intelligent Laboratory Systems, 1998, 41(2):195-207.
- [20] Vohland, M., Besold, J., Hill, J., FrJnd, H.-C., 2011. Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy[J]. Geoderma 166(1), 198-205.
- [21] Wang, J., Cui, L., Gao, W., Shi, T., Chen, Y., Gao, Y., 2014. Prediction of low heavy metal concentrations in agricultural soils using visible and near-infrared reflectance spectroscopy[J]. Geoderma 216, 1-9.
- [22]Wang,J., Cui,L., Gao,W., Shi,T., Chen,Y., Gao,Y. Prediction of low heavy metal concentrations in agricultural soils using visible and near-infrared reflectance spectroscopy[J]. Geoderma, 2014, 216.
- [23]Vohland, M., Besold, J., Hill, J., & Fr,nd, H. C. Comparing different multivariate calibration methods for the determination of soil organic carbon pools with visible to near infrared spectroscopy[J]. Geoderma, 2011, 166(1), 0-205. [DOI: 10.1016/j.geoderma.2011.08.001]
- [24] 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.[DOI: 10.1016/j.biosystemseng.2005.05.001]

Soil heavy metal Pb content inversion method by combining field with laboratory spectra

ZHANG Xia¹ DING Songtao^{1,2} CEN Yi¹ SUN Weichao¹

WANG Jinnian^{3,4}

1 Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100101, China

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

3 School of Geography and Remote Sensing, Guangzhou University, Guangzhou 510006, China

Abstract: The pollution of heavy metal has become increasingly serious in recent years. The accumulation of heavy metals in the soil will be a threat to ecological balance and human health. Therefore, we need to obtain heavy metal content in soil quickly and accurately. With the development of Hyperspectral technique, it makes it possible to estimate heavy metal content fast and at low cost by spectral data. Because the field spectra tends to be affected by environmental factors, and the number of samples is insufficient in existing studies, this paper proposed a method of constructing model by combining field spectra with laboratory spectra. First of all, Direct Standardization algorithm (DS) was employed to eliminate the environmental factors of the field spectra. Secondly, laboratory spectra was introduced to joint modeling to enhance the diversity of the samples. Finally, the characteristic spectra of iron oxide were extracted for modeling to increase the rationality of model. This method was validated by the spectral data of 70 soil samples from Xiongan farming area in Hebei province. The model established by field spectra without DS correction and characteristic spectrum obtained an accuracy (R^2) of only 0.22. However, the R^2 of the proposed method in this paper reached 0.91, and the model had excellent estimation ability. It indicates that spectral modeling of Pb content can be significantly improved by removing the influence of environmental factors on the field spectra, extracting the iron oxide characteristic spectra of the combined field and laboratory spectra.

Key words: hyperspectral remote sensing; soil heavy metal; DS conversion algorithm; iron oxide; environmental factors removal

First author: ZHANG Xia, Professor, focuses on quantitative applications of hyperspectral remote sensing. E-mail: zhangxia@radi.ac.cn

Corresponding author: Ding Songtao, PhD. student, focuses on soil metal content retrieval by hyperspectral remote sensing. E-mail: dingst425@163.com; Wang Jinnian, professor, engages in multidisciplinary remote sensing applications. E-mail: wangjn@radi.ac.cn

Foundation support: National Key R&D Program of China (2017YFB0503800); Special Project of China High-Resolution Earth Observation System (30-H30C01-9004-19/21).