

# 基于高光谱的北京铁矿区土壤重金属 镍元素含量反演研究

郭学飞<sup>1</sup>, 曹颖<sup>1\*</sup>, 焦润成<sup>1</sup>, 南贇<sup>1</sup>, 赵云峰<sup>2</sup>, 丁霞<sup>2</sup>

(1. 北京市地质研究所, 北京 100120; 2. 北京市地质工程设计研究院, 北京 101500)

**摘要:** 运用高光谱数据对北京典型铁矿区土壤重金属镍含量进行建模反演, 探索高光谱遥感技术在土壤重金属污染快速监测上应用的可行性。使用便携式地物光谱仪采集研究区土壤样本光谱反射率数据, 光谱反射率数据经多种数学变换后, 经逐步回归方法筛选最佳特征波段, 利用多元线性回归 (SLR) 和偏最小二乘回归 (PLSR) 方法建立模型以光谱反射数据对土壤重金属镍元素含量进行反演。基于光谱二阶微分的多元线性回归模型 (SD-MLR) 的稳定性和精度最高 ( $R^2 = 0.842$ ,  $RMSE = 4.474$ ), 能够良好地预测研究区土壤镍元素含量。光谱数据数学变换能够有效提高其与土壤镍元素含量间的相关性。不同的光谱变换形式建立模型的预测能力和精度有如下关系, 光谱二阶微分 > 光谱倒数对数一阶微分 > 光谱一阶微分 > 光谱倒数对数 > 光谱连续统去除 > 原始光谱。采用光谱二阶微分建立多元线性回归模型为研究区土壤镍元素含量反演的最佳模型, 可为土壤重金属污染快速监测提供技术参考。

**关键词:** 土壤重金属; 高光谱; 相关分析; 反演建模

**中图分类号:** TP79      **文献标识码:** A      **文章编号:** 0564-3945(2021)04-0960-08

DOI: 10.19336/j.cnki.trtb.2020050902

郭学飞, 曹颖, 焦润成, 南贇, 赵云峰, 丁霞. 基于高光谱的北京铁矿区土壤重金属镍元素含量反演研究 [J]. 土壤通报, 2021, 52(4): 960-967

GUO Xue-fei, CAO Ying, JIAO Run-cheng, NAN Yun, ZHU Yun-feng, DING Xia. An Inversion of Soil Nickel Contents with Hyperspectral in Iron Mine Area of Beijing[J]. Chinese Journal of Soil Science, 2021, 52(4): 960-967

土壤是生态环境的重要组成部分, 是人类赖以生存和发展的关键物质基础, 土壤环境的健康与否关系到生态环境安全和人类自身健康。随着工农业生产的快速发展, 大量污染物进入土壤环境, 造成土壤污染日益严重, 其中以生物毒性显著的重金属污染尤其突出<sup>[1-2]</sup>。土壤重金属污染具有隐蔽性、滞后性和难降解性, 对生态环境和人类健康存在严重的威胁<sup>[3-4]</sup>, 因此开展对土壤重金属的监测研究具有重大的现实意义。

高光谱遥感具有光谱范围宽、光谱分辨率高和“图谱合一”的特点, 克服了传统的实验室化学分析检测方法检测环节多、耗时长、成本高的缺点, 能够快速、宏观地监测地表信息, 逐渐被应用于土壤重金属污染监测领域<sup>[5-6]</sup>。如解宪丽等利用土壤可见光-近红外反射光谱分析了镍等多种土壤重金属含量与反射光谱之间存在的相关性<sup>[7]</sup>; Riedel Frank 等基于实验室采集的土壤高光谱反射数据, 对镍、砷、铅

等重金属元素含量进行了估测<sup>[8]</sup>; 古力扎提·艾买提等采用偏最小二乘法对新疆煤矿区土壤铅含量开展了高光谱估算<sup>[9]</sup>; 王金凤等利用土壤高光谱数据对卡斯特地区土壤锌元素含量进行了反演<sup>[10]</sup>; 有学者在总结土壤重金属指标及其光谱学机理的基础上建立了土壤重金属信息提取系统, 并计算了东北黑土地硒、锰、铁等元素的含量, 表明了该系统良好的实际应用效果<sup>[11]</sup>。

重金属元素在自然界含量较低且在土壤光谱曲线上反映极为微弱, 因此一般区域的重金属元素含量监测难以取得理想效果<sup>[12]</sup>。本研究选择北京地区典型铁矿区为研究区, 土壤重金属镍 (Ni) 元素在矿区及周边区域有明显的富集现象, 使得利用高光谱数据反演土壤 Ni 元素含量成为可能。在对土壤原始光谱数据进行多种变换的基础上, 利用相关分析和逐步回归算法确定反演土壤 Ni 元素含量的特征波段, 分别利用多元线性回归和偏最小二乘回归建立反演模型,

收稿日期: 2020-05-09; 修订日期: 2021-04-11

基金项目: 北京市总工会 2018 年助推创新项目“土壤重金属污染空间分异性高光谱遥感监测示范研究”资助

作者简介: 郭学飞 (1990-), 男, 河北省邢台市人, 硕士, 工程师, 主要从事地质环境遥感方面的研究。E-mail: xuefeigu@163.com

\*通讯作者: E-mail: cy\_998@163.com

再对模型的反演效果进行评价, 以期为北京铁矿区土壤重金属 Ni 元素含量的快速监测提供科学参考。

## 1 实验数据

### 1.1 研究区概况

研究区地处北京北部山区, 潮河流域下游, 密云水库的北部 (图 1), 覆盖地理范围  $117^{\circ}0'33'' \sim 117^{\circ}10'25''$  E,  $40^{\circ}32'41'' \sim 40^{\circ}37'8''$  N, 总面积  $114 \text{ km}^2$ 。研究区属于温带大陆性季风气候, 年平均气温  $10.8 \text{ }^{\circ}\text{C}$ , 年平均降水量  $661.3 \text{ mm}$ 。该区域地势北高南低, 为丘陵地貌。土地利用方式主要为旱地、林地、建设用地、矿业用地。区内存在多处铁矿开采区, 虽部分矿点已停产关闭, 但历史上常年的采矿活动堆积的固体废弃物, 造成了周边区域的土壤

重金属污染, 其中以镍 (Ni) 元素污染较为突出。

### 1.2 土壤样品采集与处理

通过对研究区的野外踏勘, 综合考虑矿区位置、地形地貌和土地利用现状等因素布设采样点 (图 1), 共采集土壤样品 74 个。采样时用高精度 GPS 记录采样点位置信息, 采集深度为  $0 \sim 20 \text{ cm}$  的表层土壤样品, 去除土壤中粒径较大的石块以及杂草等杂物, 保留约  $1 \text{ kg}$  土壤样品。采回实验室的样品, 经风干、研磨、过 100 目筛后, 将样品一分为二, 一份用于实验室高光谱测定, 另一份用于化学分析测定土壤重金属含量。土壤样品经  $\text{HNO}_3\text{-HCl-HClO}_4$  微波消解后, 用电感耦合等离子体原子发射光谱法 (ICP-AES) 测定土壤 Ni 元素的含量, 测定过程用国家标准样本 GSB04—1767—2004 进行质量控制。

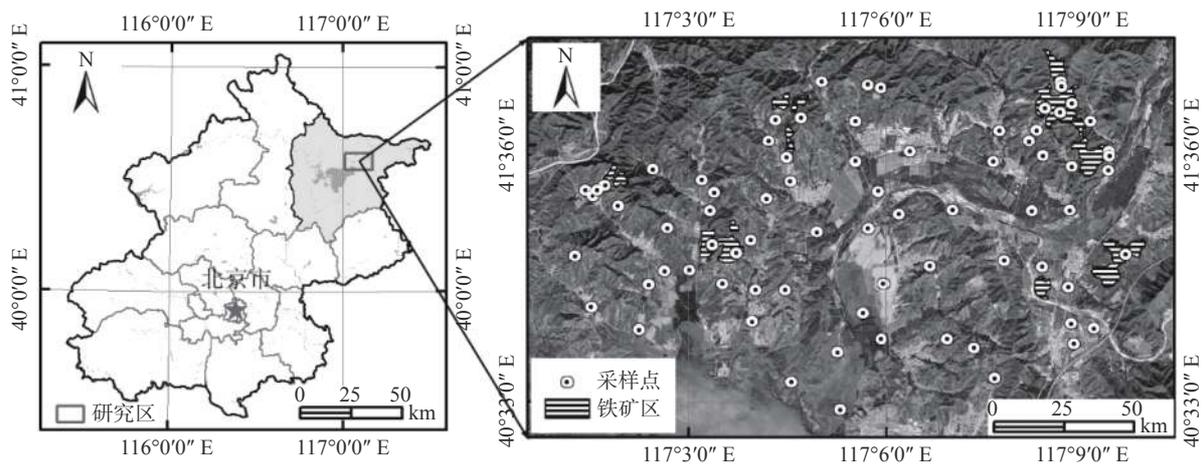


图 1 研究区位置与采样点分布

Fig.1 Location of the study area and distribution of sampling sites

### 1.3 土壤光谱采集

利用 ASD 公司设计制造的 FieldSpec4 型号光谱仪采集土壤样品高光谱反射率数据。采集的光谱范围为  $350 \sim 2500 \text{ nm}$ , 在  $350 \sim 1000 \text{ nm}$  之间的采样间隔为  $1.4 \text{ nm}$ , 光谱分辨率为  $3 \text{ nm}$ ; 在  $1000 \sim 2500 \text{ nm}$  范围内的采样间隔为  $2 \text{ nm}$ , 光谱分辨率为  $10 \text{ nm}$ ; 经过间隔为  $1 \text{ nm}$  光谱重采样, 两者输出的总波段数为 2151。为了避免外界光线的干扰, 选择夜晚在暗室条件下采集土壤光谱。将处理后的土壤样品放进直径  $15 \text{ cm}$ 、深  $2 \text{ cm}$  的透明玻璃皿内铺平, 测量时以功率  $1000 \text{ W}$  的卤素灯作为光源, 光源高度  $20 \text{ cm}$ 、天顶角  $30^{\circ}$ , 选择  $5^{\circ}$  视场角探头垂直土壤表面且距样品高度约  $10 \text{ cm}$  处采集光谱数据。每个土壤样品采集 5 次光谱数据, 去掉异常值后取平均值作为土壤样品的实际反射率光谱数据 (图 2)。

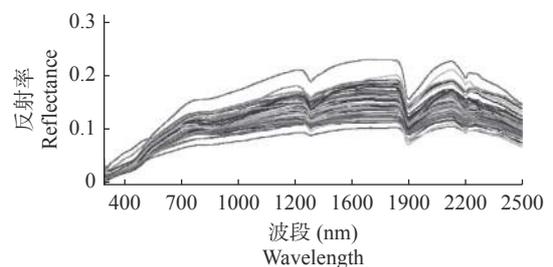


图 2 土壤样品光谱反射率曲线

Fig.2 Spectral reflectance curve of soil samples

## 2 研究方法

### 2.1 光谱数据处理

受仪器自身的影响采集的土壤反射率光谱数据在  $350 \sim 399 \text{ nm}$  和  $2400 \sim 2500 \text{ nm}$  两个谱段包含大量的噪声, 因此剔除不稳定噪声波段, 保留  $400 \sim$

2399 nm 光谱波段用于实验分析。由于在样品的制备、光谱采集等过程中很难做到条件一致,光谱仪自身产生的噪声也会引起干扰误差,为了消除随机因素的干扰,减小实验误差,采用 9 点加权移动平均法对光谱反射率数据进行平滑处理。为了进一步降低环境背景对光谱数据的干扰,增强光谱数据与重金属相关的光谱信息,对光谱反射率数据 (Reflectance spectrum, R) 进行一阶微分 (First Derivative, FD)、二阶微分 (Second Derivative, SD)、连续统去除 (Continuum Removed, CR)、倒数对数 (Absorbance Transformation, AT)、倒数对数一阶微分 (Absorbance Transformation First Derivative, ATFD) 共 5 种变换处理<sup>[13-14]</sup>。

## 2.2 特征波段选择

土壤重金属反演建模涉及的波段范围覆盖了可见光至热红外区域的诸多波段,通过计算反射率光谱数据各波段与重金属元素含量的相关系数,初步选出相关系数相对较高的波段,然后将波段带入逐步回归方程中,进一步筛选出对建模贡献大的波段作为建立反演模型的最佳变量。研究显示经过光谱变换处理后的光谱曲线相比原始的光谱曲线可以更有效地提取出光谱的特征波段<sup>[15]</sup>。

## 2.3 反演模型建立

土壤光谱是土壤中各种物质成分光谱性质的综合反映,由于土壤成分复杂,很难直接建立某种重金属元素含量与土壤特征光谱波段间的数学物理模型,目前基于土壤高光谱建立重金属元素含量估算模型的方法通常为经验统计法<sup>[16]</sup>。本文采用多元线性回归 (Multiple Linear Regression, MLR) 和偏最小二乘回归 (Partial Least Squares Regression, PLSR) 两种回归方法进行反演建模。MLR 是一种传统的单因变量对多自变量的回归建模方法,通常采用逐步回归方法筛选并剔除引起多重共线性的变量,适合于解决变量个数过多引起的数据冗余问题;PLSR 是一种多因变量对多自变量的回归建模方法,适合于变量个数多而样本数量较少的问题。本研究土壤样本较少且光谱数据存在冗余,因此选择以上两种方法进行建模,通过对比分析,选择最佳反演模型。

利用 SPSS 25.0 软件对光谱反射率数据与重金属元素含量进行相关分析并筛选特征波段,建立多元线性回归模型和偏最小二乘回归模型。

## 2.4 模型检验

从模型的拟合度和预测精度两方面对模型进行检验,拟合度表示预测值与实测值的拟合程度,预测精度表示预测结果的精密程度。采用决定系数 ( $R^2$ ) 的大小来表示模型的拟合度,  $R^2$  越大,模型的拟合效果越好,模型的稳定性越强。研究显示  $R^2 < 0.5$  时,模型的拟合效果差、稳定性低,模型不具有预测能力;  $0.5 \leq R^2 < 0.7$  时,模型的拟合效果较好、稳定性较强,模型能够进行初步预测;  $R^2 \geq 0.7$  时,模型的拟合效果好、稳定性强,模型具有良好的预测能力<sup>[17]</sup>。采用均方根误差 (RMSE) 来表示模型的预测精度, RMSE 越小,模型的预测精度越高。因此选择  $R^2$  相对较大、RMSE 相对较小的模型作为反演土壤 Ni 元素含量的最佳模型。决定系数  $R^2$  和均方根误差 RMSE 的计算公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (u_i - \bar{u}_i)^2} \quad (2)$$

式中,  $R^2$  为决定系数,  $n$  为样本数,  $\hat{y}_i$  为预测样本值,  $y_i$  为实测样本值,  $\bar{y}$  为实测样本的平均值; RMSE 为均方根误差,  $u_i$  为实测样本值,  $\bar{u}_i$  为预测样本值。

利用 K 折交叉验证法<sup>[18]</sup> (K-folded Cross Validation, 简称 K-CV) 对模型进行内部验证。将 74 组土壤 Ni 元素含量数据近似均分成 8 组,将每一组子数据作为一次验证集,其余的 7 组子数据作为建模集,这样会得到 8 个模型以及对应的决定系数,取决定系数的平均值作为此 K-CV 下建立模型的决定系数  $R^2$ 。

## 3 结果与分析

### 3.1 土壤重金属含量分析

测定研究区土壤样本 Ni 元素含量并进行统计分析 (表 1), Ni 元素含量的平均值为  $53.08 \text{ mg kg}^{-1}$ , 最大值与最小值之间的差异性较为明显, Ni 元素的变异系数为 21%, 属于中等程度变异,说明研究区土壤重金属 Ni 元素出现了一定程度的富集<sup>[19]</sup>。

### 3.2 重金属元素含量与光谱相关性分析

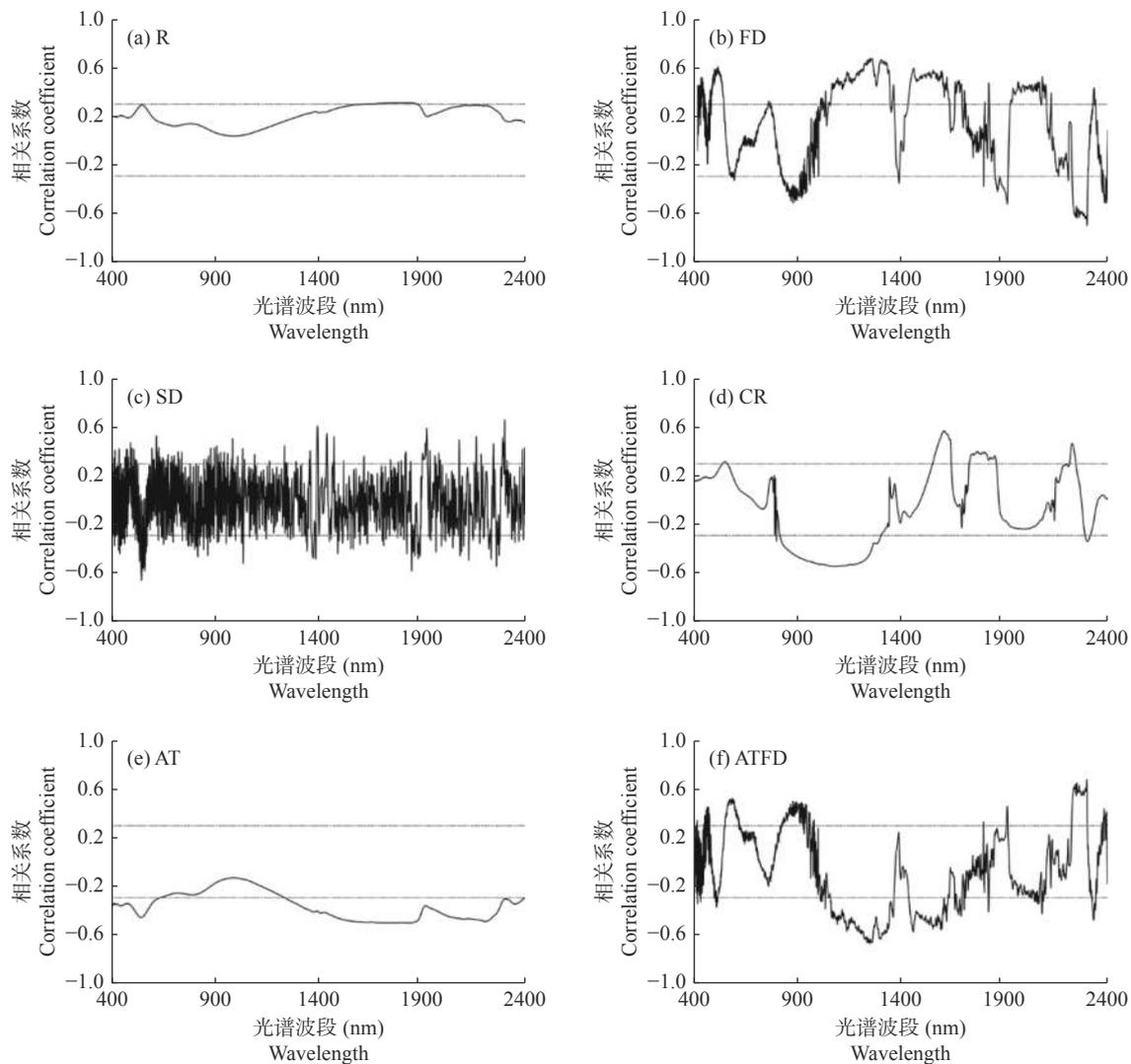
为了分析土壤重金属元素含量与光谱之间的相关关系,将重金属元素含量与原始光谱和 8 种变换

表 1 土壤 Ni 含量统计分析  
Table 1 Statistical analysis of soil Ni content

元素 Element	样本数 Sample number	最大值 (mg kg <sup>-1</sup> ) Maximum	最小值 (mg kg <sup>-1</sup> ) Minimum	平均值 (mg kg <sup>-1</sup> ) Mean	标准差 (mg kg <sup>-1</sup> ) STD	变异系数 (%) CV
Ni	74	85.3	33.9	53.08	10.87	21

光谱数据的各个波段进行皮尔森 (Pearson) 相关分析, 计算得到的土壤重金属 Ni 元素含量与光谱反射率的相关系数如图 3 所示, 达极显著水平 ( $P < 0.01$ ) 的相关系数临界值为  $\pm 0.298$ , 对应的自由度为 72。原始光谱反射率与重金属元素含量的相关性较差, 只有少数波段与重金属含量存在显著相关且相关程度为弱相关。经过变换后的土壤光谱数据与土壤重

金属元素含量的相关性有了不同程度的提升, 显著相关的波段数量也明显增多。其中, 一阶微分变换的最大相关系数达 0.7, 显著性波段达 1158 个, 分布于可见光波段至短波红外波段。可见, 光谱变换处理能够有效降低背景噪声干扰, 提高光谱数据与重金属元素含量之间的相关性。



注: 图中虚直线表示 0.01 显著水平下的相关系数临界值:  $\pm 0.298$ 。

图 3 土壤 Ni 元素含量与光谱反射率及其变换的相关系数

Fig.3 Correlation coefficient between soil Ni content and spectral radiance and its transformation

### 3.3 特征波段

根据土壤光谱与重金属 Ni 元素含量相关系数的

大小, 结合 Ni 元素的光谱响应特征, 初步筛选出各种光谱数据变换形式下的特征波段 (表 2)。利用皮

表 2 初选特征波段及相关系数  
Table 2 Primary feature bands and correlation coefficient

光谱变换 Transformed spectrum	初选特征波段及对应的相关系数 $b(r)^{\text{D}}$ Primary feature bands and correlation coefficient
R	1691 (0.30), 1815 (0.30), 1895 (0.30)
FD	434 (0.51), 510 (0.56), 1139 (0.54), 1249 (0.66), 1303 (0.65), 1461 (0.56), 1555 (0.54), 1598 (0.54), 1913 (-0.51), 2249 (-0.62), 2297 (-0.66)
SD	546 (-0.67), 618 (0.56), 985 (0.50), 1037 (-0.53), 1396 (0.6), 1441 (0.55), 1475 (0.52), 1849 (0.59), 1922 (0.59), 2126 (0.6), 2162 (0.52), 2259 (-0.55), 2299 (0.66)
CR	1078 (-0.55), 1604 (0.57), 1767 (0.40), 2223 (0.47)
AT	545 (-0.46), 1379 (-0.41), 1412 (-0.42), 1693 (-0.50), 1744 (-0.50), 2204 (-0.49)
ATFD	588 (0.53), 879 (0.50), 1094 (-0.51), 1139 (-0.57), 1250 (-0.67), 1300 (-0.64), 1461 (-0.57), 1555 (-0.55), 1593 (-0.56), 2249 (0.65), 2297 (0.68)

注: ①b代表波段, 单位nm; r代表达0.01显著水平的相关系数值。

尔森相关系数初步筛选出的特征波段虽然与 Ni 元素含量间存在相关性, 但并非所有波段都对建立 Ni 元素含量反演模型有贡献, 因此需要通过逐步回归算法进一步筛选出对反演建模贡献大的波段, 作为最终反演建模的特征波段。

将初步筛选的特征波段逐个引入回归模型, 并进行显著性检验, 当显著性  $P < 0.05$  时, 引入该波段参与建模, 当显著性  $P > 0.1$  时, 则将其剔除, 直至回归方程中不能引入也不能剔除波段为止<sup>[20-21]</sup>, 最终将贡献显著的波段 ( $P < 0.05$ ) 作为参与建模的特征波段。从重金属 Ni 元素含量反演的最佳特征波段来看 (表 3), 除少数几个波段位于 500 ~ 600 nm 外, 其他波段均集中在近红外波段范围。

### 3.4 模型建立与分析

将土壤重金属 Ni 元素含量作为因变量, 选定的特征波段作为自变量, 分别利用多元线性回归法和偏最小二乘回归法建立各种光谱变换形式下的土壤重金属元素高光谱反演模型, 模型均通过交叉验证检验。模型稳定性和精度相关统计量见表 4。

从模型稳定性来看, 光谱二阶微分变换 (SD) 建模的稳定性最高, 利用 MLR 和 PLSR 两种方法建

表 3 参与建模的特征波段  
Table 3 Modeling feature bands

光谱变换 Transformed spectrum	特征波段 (nm) Feature bands
R	1815
FD	2297, 2249
SD	546, 1849, 1396, 1475, 1441, 1037
CR	1078, 1604
AT	1379, 1412, 1693, 1744, 2204
ATFD	588, 2297

模的决定系数  $R^2$  均超过 0.7, 模型具有良好的预测能力; 其次为光谱倒数对数一阶微分 (ATFD) 建模, MLR 和 PLSR 建模的决定系数  $R^2$  均超过 0.5, 模型具有初步的预测能力; 其他光谱变换方式建模的决定系数均小于 0.5, 模型没有预测能力, 其中原始光谱 (R) 建模的稳定性最差, 决定系数  $R^2$  数值不足 0.1。

从模型预测精度来看, 对于同一种光谱变换形式, MLR 和 PLSR 两种方法建模的精度相似, 影响模型精度的主要因素是光谱变换形式。以 MLR 为例, 光谱二阶微分变换建模 (SD-MLR) 的均方根误差最小为 4.474, 模型精度最高, 在实测值与预测值散点图上 (图 4) 可看出验证样本点基本分布在 1 : 1 直

表 4 MLR 和 PLSR 方法土壤 Ni 元素含量建模精度统计  
Table 4 Modeling accuracy statistics of Ni with MLR and PLSR

光谱变换 Transformed spectrum	多元线性回归 MLR		偏最小二乘回归 PLSR	
	决定系数 $R^2$	均方根误差 RMSE	决定系数 $R^2$	均方根误差 RMSE
R	0.094	10.327	-	-
FD	0.488	7.890	0.478	7.802
SD	0.842	4.474	0.819	4.559
CR	0.418	8.341	0.417	8.170
AT	0.475	8.094	0.475	7.759
ATFD	0.540	7.409	0.540	7.258

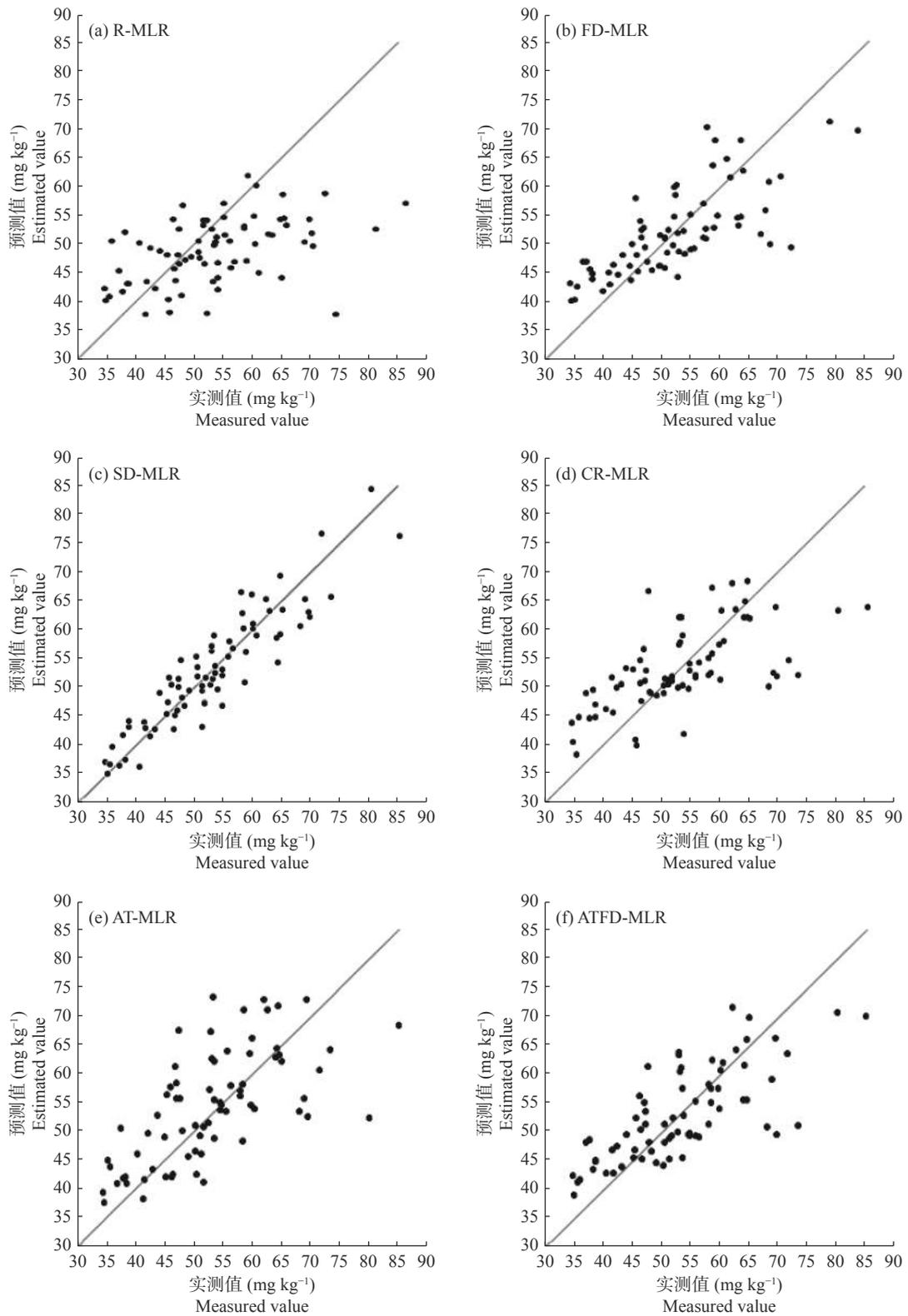


图4 土壤 Ni 元素实测值与模型(MLR)预测值比较  
 Fig.4 Comparison of measured and estimated values of soil Ni modeling by MLR

线附近; 原始光谱建模的均方根误差最大为 10.327, 模型精度最低, 在实测值与预测值散点图上验证点分布最为分散, 模型没有预测能力。

整体来看, 原始光谱数据经过数学变换后, 参与模型建立的特征波段的数量有所增加, 所建立模型的稳定性和精度随之提高, 尤其以微分变换建

模的稳定性和精度的提高最为明显；MLR 和 PLSR 两种方法建立的最佳反演模型精度相差不大，均具有良好的预测能力，但 MLR 建模的拟合度略优于 PLSR 建模，说明在研究区 MLR 建模具有更高的稳定性和可靠性。

## 4 讨论

土壤原始光谱曲线反射率在光谱波段上的分异性不明显，无法直接用于土壤重金属 Ni 元素含量的反演。通过对原始光谱数据进行预处理以及多种形式的数学变换，发现不同的变换方式对光谱与重金属 Ni 元素含量相关性有不同的提升效果，其中以光谱微分变换的提升效果最为明显，一方面增强了相关系数的大小，最大相关系数由原始光谱的 0.3 提高至 0.7，另一方面有效增强了 Ni 元素光谱特征信息，反射率与光谱波段的分异性明显增强。表明微分变换能够有效突出光谱之间的差异特征，增强有效光谱信息，提高模型预测精度，这与陈银莹等人<sup>[13, 16, 22]</sup>的研究结果一致。

对于模型建立方法，有学者认为多元逐步回归法是建模效果最好的方法<sup>[10, 15, 23]</sup>，也有学者应用偏最小二乘法建模取得了最佳建模效果<sup>[24-25]</sup>。在本研究中基于光谱微分、倒数对数、连续统去除等数学变换光谱，在逐步回归方法选取特征波段的基础上，利用多元线性回归法和偏最小二乘回归法进行建模，最佳模型均是光谱二阶微分变换，模型  $R^2$  均超过 0.8，达到良好预测能力；利用多元线性回归法建立模型的  $R^2$  较偏最小二乘回归略有提高，但最大增幅仅 0.23，表明多元回归法和偏最小二乘回归法在本研究中均是较为稳定的重金属元素含量反演方法。

土壤重金属元素反演模型的建立有赖于合理的光谱数学变换，而不同的变换方式得到的特征波段有所不同，从而导致建立的模型存在一定的不确定性。后续研究应加强土壤重金属含量与反射光谱间的机理研究，从机理上找出与某种重金属元素或元素组合相对应的光谱波段，有利于构建适应性更加广泛的土壤重金属元素含量反演模型，实现土壤重金属污染高光遥感监测技术的大范围推广应用。

## 5 结论

(1) 光谱数学变换能够有效降低光谱背景噪声干扰，增强有效光谱信息，提高土壤光谱数据与重金属元素含量间的相关性。

(2) 不同的光谱变换形式对模型预测能力和精度的提升效果排序如下：光谱二阶微分 > 光谱倒数对数一阶微分 > 光谱一阶微分 > 光谱倒数对数 > 光谱连续统去除 > 原始光谱。

(3) 利用 MLR 和 PLSR 两种算法均可建立基于变换光谱的重金属 Ni 元素含量的反演模型，MLR 建模的稳定性和精度略优于 PLSR。

## 参考文献：

- [1] 孙鑫, 娄燕宏, 王会, 等. 重金属污染土壤的植物强化修复研究进展[J]. 土壤通报, 2017, 48(4): 1008 - 1013.
- [2] 庄国泰. 我国土壤污染现状与防控策略[J]. 中国科学院院刊, 2015, 30(4): 477 - 483.
- [3] 陈卫平, 杨阳, 谢天, 等. 中国农田土壤重金属污染防治挑战与对策[J]. 土壤学报, 2018, 55(2): 261 - 271.
- [4] Arao T, Ishikawa S, Murakami M, et al. Heavy metal contamination of agricultural soil and countermeasures in Japan[J]. *Paddy and Water Environment*, 2010, 8(3): 247 - 257.
- [5] 夏军. 准东煤田土壤重金属污染高光谱遥感监测研究[D]. 乌鲁木齐: 新疆大学, 2014.
- [6] Kooistra L, Wehrens R, Leuven R, 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.
- [7] 解宪丽, 孙波, 郝红涛. 土壤可见光—近红外反射光谱与重金属含量之间的相关性[J]. 土壤学报, 2007, 44(6): 982 - 993.
- [8] Frank R, Michael D, Ingo M, et al. Prediction of soil parameters using the spectral range between 350 and 15, 000 nm: A case study based on the Permanent Soil Monitoring Program in Saxony, Germany[J]. *Geoderma*, 2018, 315: 188 - 198.
- [9] 古力扎提·艾买提, 阿不都拉·阿不力孜, 茹克亚·沙吾提, 等. 准东煤田土壤铅含量高光谱估算[J]. 土壤通报, 2018, 49(5): 1233 - 1239.
- [10] 王金凤, 王世杰, 白晓永, 等. 基于高光谱反射率的喀斯特地区土壤重金属锌元素含量反演[J]. 光谱学与光谱分析, 2019, 39(12): 3873 - 3879.
- [11] 张东辉, 赵英俊, 陆冬华, 等. 高光谱在土壤重金属信息提取中的应用与实现[J]. 土壤通报, 2018, 49(1): 32 - 37.
- [12] 郭颖, 毕如田, 郑超, 等. 土壤重金属高光谱反演研究综述[J]. 环境科技, 2018, 31(1): 67 - 71.
- [13] 陈银莹, 柳云龙. 工业区土壤高光谱特性及其重金属含量预测研究——以上海市闵行区为例[J]. 生态环境学报, 2018, 27(11): 2156 - 2162.
- [14] 张秋霞, 张合兵, 张会娟, 等. 粮食主产区耕地土壤重金属高光谱综合反演模型[J]. 农业机械学报, 2017, 48(3): 148 - 155.
- [15] 沈强, 张世文, 葛畅, 等. 矿业废弃地重构土壤重金属含量高光谱反演[J]. 光谱学与光谱分析, 2019, 39(4): 1214 - 1221.
- [16] 贺军亮, 李志远, 李仁杰, 等. 基于有机质光谱特征的土壤重金

- 属Pb估算模型研究[J]. 环境污染与防治, 2019, 41(7): 753 - 757.
- [ 17 ] 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.
- [ 18 ] 汪庆华, 刘江伟, 张兰兰. 交叉验证K近邻算法分类研究[J]. 西安工业大学学报, 2015, 35(2): 119 - 124.
- [ 19 ] 李琼琼, 柳云龙. 城市居民区土壤重金属含量高光谱反演研究[J]. 遥感技术与应用, 2019, 34(3): 540 - 546.
- [ 20 ] 游士兵, 严 研. 逐步回归分析法及其应用[J]. 统计与决策, 2017, (14): 31 - 35.
- [ 21 ] 王自成, 朱家明, 陈华友. 基于逐步回归筛选的回归组合预测模型[J]. 统计与决策, 2019, (17): 75 - 78.
- [ 22 ] 聂 哲, 李秀芬, 吕家欣, 等. 东北典型黑土区表层土壤有机质含量高光谱反演研究[J]. 土壤通报, 2019, 50(6): 1285 - 1293.
- [ 23 ] 徐明星, 吴绍华, 周生路, 等. 重金属含量的高光谱建模反演: 考古土壤中的应用[J]. *红外与毫米波学报*, 2002, 30(2): 109 - 114.
- [ 24 ] 于 庆, 吴泉源, 姚 磊, 等. 山东典型污灌区冬小麦叶片重金属高光谱反演及空间分布特征[J]. 河南农业科学, 2018, 47(8): 54 - 62.
- [ 25 ] 张银玲, 姚 峰, 牛 婷, 等. 基于高光谱的土壤重金属含量预测模型建立与评价[J]. *新疆环境保护*, 2016, 38(3): 15 - 21.

## An Inversion of Soil Nickel Contents with Hyperspectral in Iron Mine Area of Beijing

GUO Xue-fei<sup>1</sup>, CAO Ying<sup>1\*</sup>, JIAO Run-cheng<sup>1</sup>, NAN Yun<sup>1</sup>, ZHU Yun-feng<sup>2</sup>, DING Xia<sup>2</sup>

(1. *Beijing Institute of Geology, Beijing 100120, China*; 2. *Beijing Geological Engineering Design Institute, Beijing 101500, China*)

**Abstract:** Inversion modeling of heavy metal nickel in soils of typical iron mine area in Beijing was carried out with hyperspectral data, in order to explore the feasibility of hyperspectral remote sensing technology in rapid monitoring of soil heavy metal pollution. The spectral reflectance data of soil samples in the study area were obtained by portable spectrometer, and the original spectral data were mathematically transformed. Optimum feature bands were extracted with Stepwise regression method. The inversion model of soil nickel content was established with Multiple Linear Regression (SLR) and Partial Least Squares Regression (PLSR). The multiple linear regression model based on spectral second derivative (SD-MLR) had the highest stability and accuracy ( $R^2 = 0.842$ , RMSE = 4.474), which was the best inverse of the nickel content in the soil of the study area. Spectral mathematical transformation can effectively improve the correlation between spectral data and element content. The following relationship exists in the modeling predictive power and accuracy of spectrum reflectance under different transformations: second derivative > absorbance transformation first derivative > first derivative > absorbance transformation > continuum removed > original spectra. The SD-MLR model could well predict the nickel content in the area and provide a technical reference for rapid monitoring of soil heavy metal pollution.

**Key words:** Soil heavy metals; Hyper-spectral; Correlation; Inversion modeling

[ 责任编辑: 韩春兰 ]