

西平县高标准农田Zn含量高光谱反演研究

蔡太义, 王志刚, 杨刘帅, 王群, 黄会娟, 于海洋, 张传忠, 张灿, 刘鹏, 冯玉庆, 贺成龙, 张合兵

引用本文:

蔡太义, 王志刚, 杨刘帅, 王群, 黄会娟, 于海洋, 张传忠, 张灿, 刘鹏, 冯玉庆, 贺成龙, 张合兵. 西平县高标准农田Zn含量高光谱反演研究[J]. 农业环境科学学报, 2022, 41(10): 2223-2231.

在线阅读 View online: <https://doi.org/10.11654/jaes.2022-0851>

您可能感兴趣的其他文章

Articles you may be interested in

基于SD-SWT的铜胁迫下玉米光谱奇异性甄别与污染监测

李艳茹, 杨可明, 韩倩倩, 高伟, 张建红

农业环境科学学报. 2020, 39(9): 1869-1877 <https://doi.org/10.11654/jaes.2020-0468>

玉米/大豆间作的镉累积规律初探

李涵, 黄道友, 黄山, 汤彬, 郭欢乐, 曹钟洋, 陈松林, 陈志辉

农业环境科学学报. 2020, 39(9): 1900-1907 <https://doi.org/10.11654/jaes.2020-0401>

我国20种典型土壤中锌对白符跳虫的毒性阈值及其预测模型

王巍然, 林祥龙, 赵龙, 张家乐, 樊文华, 侯红

农业环境科学学报. 2021, 40(4): 766-773 <https://doi.org/10.11654/jaes.2020-1220>

麦玉轮作区农田土壤重金属调查及评价

李鹏, 张惠娟, 徐莉, 李辉信, 胡锋, 焦加国

农业环境科学学报. 2022, 41(1): 46-54 <https://doi.org/10.11654/jaes.2021-0624>

丝瓜络固定生物氧化锰吸附重金属离子

孟佑婷

农业环境科学学报. 2021, 40(4): 859-865 <https://doi.org/10.11654/jaes.2020-1033>



关注微信公众号, 获得更多资讯信息

蔡太义, 王志刚, 杨刘帅, 等. 西平县高标准农田 Zn 含量高光谱反演研究[J]. 农业环境科学学报, 2022, 41(10): 2223-2231.

CAI T Y, WANG Z G, YANG L S, et al. Hyperspectral inversion model of Zn in high standard farmland soil in Xiping County[J]. *Journal of Agro-Environment Science*, 2022, 41(10): 2223-2231.



开放科学 OSID

西平县高标准农田 Zn 含量高光谱反演研究

蔡太义¹, 王志刚¹, 杨刘帅^{1,2}, 王群³, 黄会娟¹, 于海洋¹, 张传忠⁴, 张灿⁵, 刘鹏⁶, 冯玉庆¹, 贺成龙¹, 张合兵^{1*}

(1. 河南理工大学测绘与国土信息工程学院, 河南 焦作 454000; 2. 自然资源部第四地形测量队, 哈尔滨 150000; 3. 河南农业大学农学院/河南省农田生态系统野外科学观测研究站, 郑州 450046; 4. 河南省土壤调理与修复工程技术研究中心, 河南 商丘 476000; 5. 中向旭曜科技有限公司, 江苏 无锡 214400; 6. 河南省地质局矿产资源勘查中心, 郑州 450053)

摘要:为实现高标准农田土壤重金属 Zn 含量的快速测定, 本文以西平县土壤 Zn 为研究对象, 通过采集 168 个土壤样本进行室内实验, 获得土壤高光谱数据(400~2 400 nm)并进行 Savitzky-Golay 平滑后, 利用 5 种光谱变换, 结合连续投影算法识别最佳特征波段, 采用偏最小二乘回归方法构建 Zn 元素最佳反演模型。结果表明:二阶微分(SD)在 1 409 nm 波段的相关性(-0.502)最大, 一阶微分(FD)在 2 323 nm 波段的相关性(0.491)最大, 去包络线(CR)在 2 439 nm 波段的相关性(0.476)最大; 倒数对数(LOG)、一阶微分、二阶微分、平滑曲线(SG)、去包络线的拟合度(R^2)在 0.65~0.70 之间, 相对分析误差(RPD)处于 1.71~2.29 之间, 其中去包络线的拟合度($R^2=0.70$ 、 $RPD=2.29$)最高。经过 5 种光谱变换后的光谱反射率可有效突出光谱反射率的变化特征, 并用于构建反演模型; 土壤重金属 Zn 的最佳模型是以去包络线光谱变换为最佳偏最小二乘模型。

关键词:高标准农田; 高光谱反演; 偏最小二乘法; 连续投影算法; Zn

中图分类号: X87; X833 文献标志码: A 文章编号: 1672-2043(2022)10-2223-09 doi:10.11654/jaes.2022-0851

Hyperspectral inversion model of Zn in high standard farmland soil in Xiping County

CAI Taiyi¹, WANG Zhigang¹, YANG Liushuai^{1,2}, WANG Qun³, HUANG Huijuan¹, YU Haiyang¹, ZHANG Chuazhong⁴, ZHANG Can⁵, LIU Peng⁶, FENG Yuqing¹, HE Chenglong¹, ZHANG Hebing^{1*}

(1. School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China; 2. The Fourth Topographic Survey Team of the Ministry of Natural Resources, Harbin 150000, China; 3. College of Agronomy, Henan Agricultural University / Henan Province Agro-ecosystem Field Observation and Research Station, Zhengzhou 450046, China; 4. Henan Province Soil Conditioning and Repair Engineering Technology Research Center, Shangqiu 476000, China; 5. View Sino Orise Technology Co., Ltd., Wuxi 214400, China; 6. Mineral Resources Exploration Center of Henan Geological Bureau, Zhengzhou 450053, China)

Abstract: To rapidly determine the heavy metal Zn content in high-standard farmland soil, we collected and analyzed Zn in soil collected in Xiping County. Through indoor experiments in which 168 soil samples were collected, soil hyperspectral data (400-2 400 nm) were obtained and smoothed using the Savitzky-Golay method. Five types of spectral transformations and continuous projection algorithms were used to identify the best characteristic bands, and the partial least square regression method was used to construct an optimal inversion model of Zn. The correlation of the second-order differential (-0.502) was highest at the 1 409 nm band; the correlation of the first-order

收稿日期: 2022-08-26 录用日期: 2022-09-21

作者简介: 蔡太义(1972—), 男, 河南唐河人, 博士/博士后, 副教授, 主要从事多源遥感、土壤健康及三维结构研究。E-mail: caity2008@hpu.edu.cn

*通信作者: 张合兵 E-mail: jzitzhb@hpu.edu.cn

基金项目: 国家自然科学基金项目(41671225); 河南省重大科技专项(181100110400)

Project supported: The National Natural Science Foundation of China(41671225); The Key Science and Technology Project of Henan Province, China(181100110400)

differential(0.491) was largest in the 2 323 nm band; and the correlation of the de-envelope(0.476) was the highest in the 2 439 nm band. The fitting degree of a reciprocal logarithm, first-order differential, second-order differential, smooth curve, and de-envelope was 0.65–0.70, and the residual predictive deviation (RPD) was 1.71–2.29. The de-envelope showed the highest fitting degree ($R^2=0.70$, RPD=2.29). The five types of spectral transformation can highlight variations in spectral reflectance and can be used to construct an inversion mode. The best model for soil heavy metal Zn is the de-enveloping spectral transformation, which is a partial least square model.

Keywords: high standard farmland; hyperspectral inversion; partial least squares; continuous projection algorithm; Zn

土壤重金属是土壤复杂生态系统中重要的组成成分,影响着土壤环境质量^[1]。锌(Zn)是一种重要的土壤重金属元素^[2],是在大气沉降、工业污染和施肥等过程中缓慢积累,且大部分处于结合状态的金属离子,并且其不易降解^[3]。我国土壤中Zn的环境背景含量在2.60~593.00 mg·kg⁻¹之间^[4]。Zn含量过低会减少作物激素的生成,过高则会对作物产生毒害作用^[5]。因此,Zn含量的研究是高标准农田建设过程中的一个重要内容。传统重金属测量与监测方法具有耗时长、范围窄和实时性差的缺点,不能满足反映生态系统结构变化的要求^[6]。高光谱遥感数据可以快速检测土壤重金属属性信息^[7],能节省人力与物力,同时具有成本低、无污染和无损的优势^[8],是土壤重金属含量研究的新趋势^[9]。

前人对Zn元素含量研究已取得重大进展。陶超等^[10]利用定量回归模型和定性反演模型比较采样点重金属铅(Pb)与Zn的可迁移能力,证明在快速检测土壤重金属污染状况的问题上,定性分类是一种更加切实可行的方式。王金凤等^[11]运用电感耦合等离子体质谱和分光光度计测定土壤样品,分析土壤光谱和Zn含量的映射关系,实现了Zn含量高效快速反演,一定程度上揭示了喀斯特地区土壤重金属Zn的光谱吸收特性。程先锋等^[12]通过构建土壤中Zn、Pb、砷(As)、镉(Cd)4种重金属含量及土壤光谱之间的多元线性回归模型,证明土壤重金属含量与光谱特征具有较强的相关性。宋婷婷等^[13]建立了基于乘积变换的Zn含量最优预测模型并制作了污染分布图,表明Zn污染和矿业活动间具有显著相关性。TAN等^[14]在伊通县矿区对选定的光谱特征和土壤样品的重金属数据进行反演并建立反演模型,结果表明CARS叠加法能较好地预测研究区As、铬(Cr)、Pb和Zn的含量。LIN等^[15]利用黑龙江省黑土样本的室内光谱数据,分析了土壤有机质、铁(Fe)、Zn含量与6种光谱反射率的相关性;然后在 $\alpha=0.01$ 水平进行相关系数显著性检验,采用主成分分析(KPCA)算法,结合随机森林(RF)和支持向量机(SVM)组合模型应用于Zn、Fe含

量的定量反演。结果表明在土壤重金属Zn、Fe的预测中,KPCA-RF模型优于RF模型,其在土壤成分含量的定量反演中具有广阔的应用前景。

目前,对于土壤Zn的研究,国内外多侧重于矿区Zn的污染及土壤Zn的回归模型的精度与迁移能力,而较少关注耕地质量提升与治理中Zn的反演研究,从而制约了国家土壤健康及高标准农田建设的稳步推进。因此,本研究以黄淮海典型区域(西平县)高标准农田土壤重金属Zn作为研究对象,于实验室测得野外采集土样的Zn含量,并利用ASD Field Spec便携式光谱仪采集土壤光谱数据;运用5种方式对反射率光谱进行变换以增强光谱特征信息,并与土壤Zn含量实测值进行皮尔逊(Pearson)相关分析,初次筛选出显著特征波段;再利用连续投影算法(SPA)二次筛选出最佳特征波段,以构建Zn含量偏最小二乘法回归模型(PLS);检验5个模型的均方根误差(RMSEC、RMSECV、RMSEP)、校正决定系数(R^2)和相对分析误差(RPD)并进行比较,选取Zn含量最佳反演模型。本研究旨在为预测高标准农田建设区域土壤重金属Zn含量及其治理提供理论参考。

1 材料与方法

1.1 研究区概况

驻马店市西平县位于北纬33°10′~33°32′,东经113°36′~114°13′(图1),河南中南部区域,其土壤类型分布及面积占比见表1。县境总面积为1 089.77 km²,划分为3个街道办事处、8镇、8乡和老王坡管理委员会等行政区。该地气候属亚湿润大陆性季风型,全年气候温和。西平县资源丰富,农作物物种丰盛,是我国粮食生产先进县与粮食主产区之一,同时其西部山区拥有森林覆盖率95%以上的国家级森林公园。

1.2 土样采集

土样采集前对西平县研究区的土壤类型、空间特征、行政单元和地形因素进行分析,以设定精确定位的采样点。利用对角线采样法采样,野外采样深度为

表1 西平县土壤类型分布及面积占比

Table 1 Distribution of soil types and area share in Xiping County

类型 Type	分布 Distribution	面积 Area/万 hm ²	占比 Percent/%
黄棕壤土 Yellow-brown loam	西部乡镇低山丘陵区、冈丘区及部分高地	2.93	35.5
潮土 Tidal soil	洪河、柳堰河及其支流两侧	1.72	20.7
砂姜黑土 Sand ginger black soil	淤泥河两岸及坡洼地带	3.63	43.8

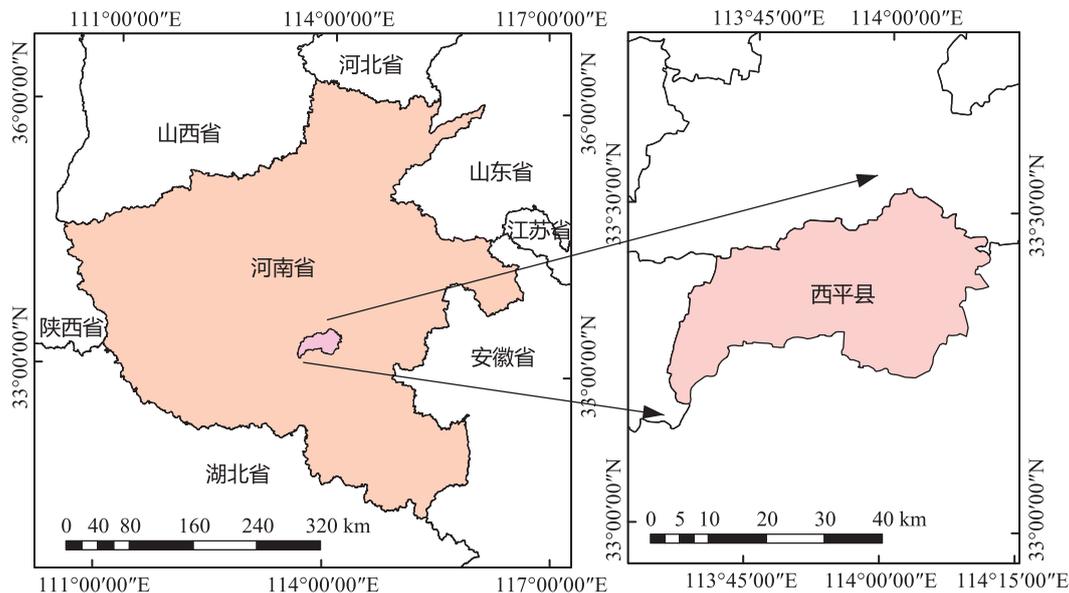


图1 研究区位置示意图

Figure 1 Location diagram of the study area

0~30 cm 土壤层,同时记录详细地形特征。从田里采集的土样及时摊开风干,倒在干净的容器内,去除石块、植物根茎残体及其余杂物。取研磨均匀后通过1 mm 孔径孔筛的土壤样本,利用四分法取样,用于土壤光谱数据的采集和Zn含量的测定^[16]。

1.3 土样光谱室内测定

土样光谱使用ASD Field Spec 便携式光谱仪测定,测量光谱数据波段350~2 500 nm。在土壤样本光谱测定开始前,将盛装在黑色器皿中的土壤样本表面刮平,采用合适功率的卤化灯作为测量光源,转动器皿以合适的距离重复10次采集4个方向的土样光谱,以降低土壤样本光谱受各向异性因素的影响,取10次土壤样本光谱反射率平均值作为原始光谱反射率值。人为去除350~400 nm和2 400~2 500 nm 两段受外界噪声影响较大,导致光谱数据不稳定的光谱反射率数据^[17]。

1.4 光谱变换方法

倒数对数光谱变换可以增强光谱反射率数据在可见光区域的特征差异,并能减少因光照强度和地形

起伏变化等因素对光谱特征曲线引起的负面影响^[18],其变换公式见表2。

光谱微分技术是一种简单而有效的确定光谱曲线极值点波长位置的方法,是通过反射率光谱数据进行阶数不同的数学微分模拟来提取特征光谱参数,该技术能够消除平缓背景部分环境或基线漂移的影响,扩大样本之间的光谱特征差异,有助于吸收特征的提取^[19](表2)。

包络线是利用直线连接每个随波长不断变化的吸收谷或反射峰凸出的特征点,并使连接线的外角大于180°的线^[20]。去包络线是以相应波段上实际光谱与包络线两者反射率值进行比值计算,从而使之归一化^[21](表2)。

1.5 相关分析与显著性检验

相关分析是用于判断多个随机变量间是否存在统计学关联的有效统计分析方法^[22],且能深入分析存在关联的变量间的关联强度与方向,反映事物间的统计学相关程度。两个变量之间相关关系的强弱程度,可以用数值的形式揭示,其数值又称为相关系数。皮

表2 光谱变换公式

Table 2 Spectral transformation equations

光谱变换 Spectral transformation	公式 Formula
倒数对数 Inverse logarithmic	$\log \frac{1}{R(\lambda_i)}$
去包络线 Removal of envelopes	$\frac{R(\lambda_i)}{R_{start} + \frac{R_{end} - R_{start}}{\lambda_{end} - \lambda_{start}} (\lambda_i - \lambda_{start})}$
一阶微分 First-order differentiation	$\frac{R(\lambda_{i+1}) - R(\lambda_{i-1}))}{2\Delta\lambda}$
二阶微分 Second-order differentiation	$\frac{R(\lambda_{i+2}) - 2R(\lambda_{i+1}) + R(\lambda_i)}{\Delta\lambda^2}$

注： λ_i 表示波长； λ_{start} 、 λ_{end} 表示在吸收曲线里的起始点波长和末端点波长； $\Delta\lambda$ 表示相邻2个波长的间距； $R(\lambda_i)$ 表示光谱反射率值； R_{start} 、 R_{end} 表示在吸收曲线里的起始点和末端点的原始光谱。

Notes: λ_i denotes wavelength; λ_{start} and λ_{end} represent the starting and end wavelengths in the absorption curve; $\Delta\lambda$ represents the distance between two adjacent wavelengths; $R(\lambda_i)$ represents the spectral reflectance value; R_{start} and R_{end} represent the original spectra of the starting and end points in the absorption curve.

尔逊相关系数(积差相关系数)是应用最广泛的相关系数,其公式如下:

$$R = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

式中： x_i 、 y_i ($i=1, 2, \dots, n$)为两个随机变量 x 和 y 的样本值； \bar{x} 、 \bar{y} 为两个随机变量样本值的平均值； n 为两个变量样本值的个数。 R 为相关系数,且 $-1 \leq R \leq 1$ ； $|R|$ 越接近1,说明两变量间线性相关性越强, $|R|=1$,为完全线性相关； $|R|$ 越接近0,说明两变量间线性相关性越弱, $|R|=0$,为完全不具有相关性； $R>0$,说明两变量为正相关； $R<0$,说明两变量为负相关。在科研应用中,可以划分几个区间取值来表示不同的相关程度。

样本间是否存在显著线性相关关系不能直接通过相关性分析得到的系数值确定,其原因在于随机抽样的性质或样本数量的不足等,因此必须使用假设检验的方式对总体的显著性进行相关分析显著性检验^[23]。

1.6 连续投影算法

连续投影算法(SPA)是将向矢量空间共线性最小化的前向变量选择算法^[24]。SPA通过投影将光谱

数据的波长映射到其他波长上进行向量的投影分析,从而比较出所投影向量的大小,选择投影向量值最大的波长作为优选波长,然后经过矫正模型选取出所需要的特征波长^[25]。

1.7 偏最小二乘法回归模型反演与精度检验方法

在构建偏最小二乘法回归模型(PLS)^[26]中,从自变量中提取的主成分要求尽可能体现出光谱曲线的原特征信息^[27],另外从因变量中提取的独立成分要求与主成分存在最大的相关性。将提取出的成分与因变量通过多元回归方法构建回归方程^[28]。 $RMSE$ 、 R^2 和 PRD ^[29]的公式及定义见表3。

2 结果与分析

2.1 土壤Zn属性特征

土壤重金属Zn含量范围在57.13~306.40 $\text{mg} \cdot \text{kg}^{-1}$ 之间(表4),变异系数^[30]为0.401,为中等变异性,适宜构建模型。Zn元素的质量分数偏态分布密度曲线表现出土壤Zn含量多集中于125~130 $\text{mg} \cdot \text{kg}^{-1}$ 之间(图2),极端值分布较少。《土壤环境质量 农用地土壤污染风险管控标准(试行)》(GB 15618—2018)规定^[31],当土壤 $\text{pH} \leq 6.5$ 时,土壤重金属Zn含量的风险筛选值为200 $\text{mg} \cdot \text{kg}^{-1}$ 。调查研究区土壤平均 pH 值为5.9,样本数据表明该区整体Zn污染风险较低,少数样本极端值表明该区局部Zn污染风险较高,应当加强土壤

表3 精度校验方法公式

Table 3 Precision calibration method formula

公式 Formula	定义 Definition
$R^2 = 1 - \frac{\sum_{i=1}^n (Y_i - Y_i^{pre})^2}{\sum_{i=1}^n (Y_i - Y_i^{avg})^2}$	用来评价构建的回归方程的拟合程度及优劣程度, R^2 随着变量个数的增加而增大
$RMSEC = \sqrt{\frac{\sum_{i=1}^n (Y_i - Y_i^{pre})^2}{n}}$	用来判断观测值与真实值间偏差大小
$PRD = \frac{\sum_{i=1}^n (Y_i - Y_i^{avg})}{\sum_{i=1}^n (Y_i - Y_i^{pre})^2}$	验证集标准差与均方根误差的比值

注： Y_i 、 Y_i^{pre} 和 Y_i^{avg} 分别为样本*i*的真实值、预测值和平均值； n 为样本数。

Note: Y_i , Y_i^{pre} and Y_i^{avg} are the true, predicted and mean values of sample i respectively; n is the number of samples.

表4 土壤重金属Zn含量分析

Table 4 Analysis of soil properties for heavy metal Zn content

最小值 Minimum values/ $(\text{mg} \cdot \text{kg}^{-1})$	最大值 Maximum values/ $(\text{mg} \cdot \text{kg}^{-1})$	均值 Average/ $(\text{mg} \cdot \text{kg}^{-1})$	标准差 Standard deviation/ $(\text{mg} \cdot \text{kg}^{-1})$	变异系数 Coefficient of variation	样本数 Sample size
57.13	306.40	123.86	49.79	0.401	168

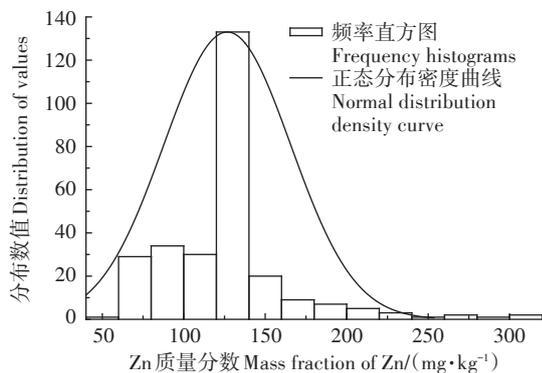


图2 Zn质量分数描述统计图

Figure 2 Descriptive statistics of Zn mass fraction

环境监测和农产品协同检测。

2.2 土壤属性光谱反射特征

土壤光谱曲线一般为突出弧线,土壤样本间的光谱差异性主要体现在相同波段上的光谱反射率数值不同(图3)。土壤样本在可见光带400~780 nm范围处的反射率较低,但呈大幅度上升的趋势,光谱曲线陡峭,土壤有机质在600、800 nm附近有微弱的反射峰,在780~2 300 nm处近红外波段区域曲线较平缓,在2 300~2 400 nm处逐渐下降。1 850~2 140 nm光谱曲线呈现明显的“V”字形。图中3个水分吸收带表现出的波谷与黏土中的多种矿物质组成有关,1 420 nm为羟基(—OH)谱带、1 920 nm附近为H₂O谱带、2 200 nm波段范围为羟基伸缩振动与Al—OH和Mg—OH弯曲振动的合谱带^[32]。4条平均光谱曲线整体趋势相似,丘岗地区的平均光谱曲线较低,山地丘陵区的平均光谱曲线最高,而县区与平原区的平均光谱曲线居中,几乎吻合。

2.3 相关分析与显著性检验

从5种光谱变换后的光谱数据与土壤Zn含量的

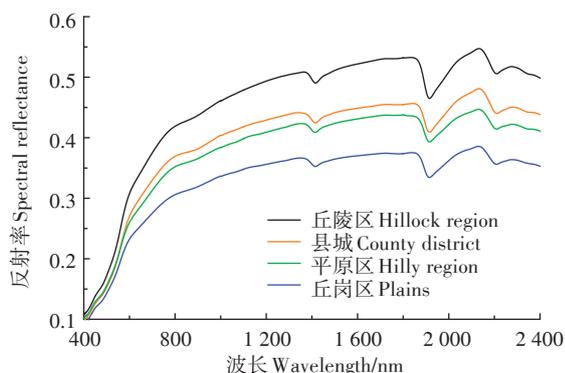


图3 西平县不同地类土壤平均光谱反射率曲线

Figure 3 Average spectral reflectance curves of different land types of soil in Xiping County

相关系数折线图(图4)可以看出,经过光谱变换,各波段与土壤Zn含量的相关系数也有所不同,呈现多元化的变化趋势。倒数对数光谱变换(LOG)与平滑光谱变换(SG)的相关系数曲线较为平滑,且呈现以数值0为对称轴的对称显示,显著相关系数的波段数、最大相关波段基本一致;二者中一者递增,另一者则递减,最显著波段相关系数呈现相反关系,其绝对值十分相近,并均接近0.3。

去包络线光谱变换(CR)的相关系数曲线出现上下频繁波动,在600~1 400、1 900~2 400 nm范围内有多处波峰和波谷,数值变化范围介于-0.3~0.3,曲线变化形似于一阶微分变换(FD)后的光谱相关系数曲线。SG光谱反射率相关系数曲线与CR光谱相关系数曲线拥有相似的波峰位置。FD与二阶微分变换(SD)光谱的相关系数数值变化折线图不再呈现近似单一数值的平滑线,而是表现为在正负值上下频繁的波动,FD光谱的相关系数曲线在580 nm附近出现一个波峰,在1 900 nm附近出现一个波谷,SD的相关系数曲线变化较为剧烈,连续波段相邻的相关系数在0上下跳跃。

FD、SD、CR变换光谱与Zn含量显著相关的波段数量较SG交换光谱的显著相关波段数明显减少,其中SD变换光谱相关系数曲线的显著相关波段数最少。5种变换光谱对应的最显著波段中,SD变换光谱的相关系数最大,FD变换和CR变换光谱的相关系数次之,LOG变换和SG变换光谱对应的相关系数最小(表5)。

综上所述,经过LOG、FD、SG等5种光谱变换后

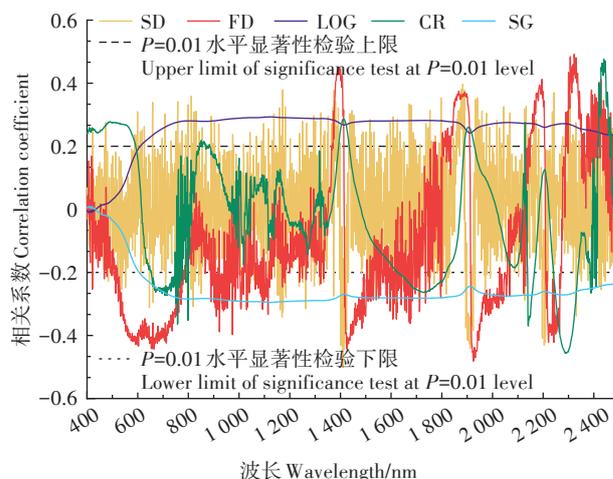


图4 5种变换光谱反射率相关系数显著性曲线

Figure 4 Significance curves for the correlation coefficients of the five transformed spectral reflectances

的光谱反射率能有效突出光谱反射率的变化特征,经检验光谱变换下通过 $\alpha=0.01$ 水平显著性检验的相关波段可作为第一次特征波段数据,用于进行第二次特征波段的筛选,从而构建土壤Zn含量的反演模型。

2.4 最佳特征波段的选取

利用 Matlab 软件结合连续投影算^[31]法分别对 SG、LOG、FD、SD 和 CR 变换后的光谱数据进行波段的第二次筛选,得出最佳特征波段见表6。

实验没有考虑其他重金属元素会干扰光谱反演分析过程,本研究认为每种重金属都有很多敏感波段且不尽相同,但在通过优选特征波段的组合后可逐步降低与其他重金属敏感波段的重叠程度。通过第一次筛选和第二次筛选后,将从300~2 500 nm整谱波段中筛选出的5~10个最佳特征波段进行组合,使光谱数据与Zn含量的相关性进一步上升,很大程度上降低了与其他重金属的关联性。

2.5 土壤重金属Zn元素反演模型

Zn分别以FD、LOG、SG、SD、CR光谱变换为自变量的模型^[34]整体上具有较高的 \bar{R}^2 ,其值均大于0.60,最高为0.70, RMSEC都在20范围上下(表7)。Zn以SG光谱变换为自变量的模型交叉验证系数 \bar{R}^2 最小,为0.61,以CR光谱变换的交叉验证系数 \bar{R}^2 最高,为0.68。

对比建模集结果和交叉验证结果,Zn含量以CR光谱变换为自变量的模型建模效果较好。检验的

\bar{R}_v^2 、RMSEP均在验证集检验的结果中有不同的变化,但建模效果比较好的模型在验证集的验证中精度仍然较高。进行RPD数据分析发现,5种光谱变换模型中CR光谱变换模型的RPD值最高为2.29,其余4种回归模型的RPD值介于1.71~2.02之间,表明以SD、LOG和SG光谱变换构建的偏最小二乘法回归模型具有定量预测能力,而CR光谱变换模型的RPD值最高, \bar{R}_v^2 最大, RMSEP最小,表明CR具有很好的模型预测能力。在Zn含量实测值与预测值的拟合散点图(图5)中也可以看出,拟合点更多位于 $y=x$ 直线下方,说明模型的预测值更多的是小于实际值;而基于CR变换构建的模型中预测值与实测值的拟合直线(图5a)与 $y=x$ 直线夹角最小,说明其是5种模型中的最优模型。

3 结论

(1)研究区土壤Zn含量呈一定程度的偏态分布,大部分土壤样品Zn含量介于60~160 $\text{mg}\cdot\text{kg}^{-1}$,均值为123.86 $\text{mg}\cdot\text{kg}^{-1}$,表明研究区整体Zn污染风险较低;但存在极端值大于300 $\text{mg}\cdot\text{kg}^{-1}$ 的样品,说明研究区内局部地点Zn污染风险高。

(2)通过光谱反射率平滑处理消除仪器不同探测元件对光谱数据的影响,使反射率曲线变平滑,提高信噪比。通过光谱变换即倒数对数(LOG)、一阶微分

表5 光谱变换后反射率与Zn含量显著相关的波段

Table 5 The wavelengths with significant correlation between reflectivity and Zn content after spectral transformation

光谱变换 Spectral transformations	波段数 Number of wavelengths	最显著波段 The most significant wavelength/nm	相关系数 Correlation coefficient
SG	1 675	1 122	-0.295
FD	1 183	2 323	0.491
SD	454	1 409	-0.502
CR	795	2 439	0.476
LOG	1 653	1 123	0.293

表6 土壤Zn属性最佳光谱特征波段

Table 6 Best spectral feature wavelengths for soil Zn properties

光谱变换 Spectral transformations	波段数 Number of Wavelengths	最佳特征波段 Best feature wavelength/nm
SG	5	1 122, 1 413, 1 801, 2 124, 2 190
FD	5	995, 1 746, 1 801, 2 260, 2 487
SD	10	469, 829, 975, 1 069, 1 097, 1 384, 1 886, 1 889, 1 978, 2 445
CR	5	751, 799, 853, 2 139, 2 277
LOG	5	1 120, 1 403, 1 795, 2 120, 2 200

表7 土壤Zn含量的偏最小二乘法模型的建模与验证

Table 7 Modelling and validation of partial least squares models for soil Zn content

光谱变换 Spectral transformations	建模集 Modelling set(n=127)		交叉验证 Cross-validation(n=127)		验证集 Validation set(n=41)		
	\bar{R}^2	RMSEC	\bar{R}^2	RMSECV	\bar{R}_v^2	RMSEP	RPD
SG	0.65	22.51	0.61	26.44	0.63	24.33	1.86
LOG	0.67	21.60	0.66	22.12	0.64	23.27	1.95
FD	0.64	23.26	0.62	25.53	0.61	26.43	1.71
SD	0.67	21.61	0.63	24.33	0.65	22.47	2.02
CR	0.70	18.69	0.68	20.21	0.69	19.83	2.29

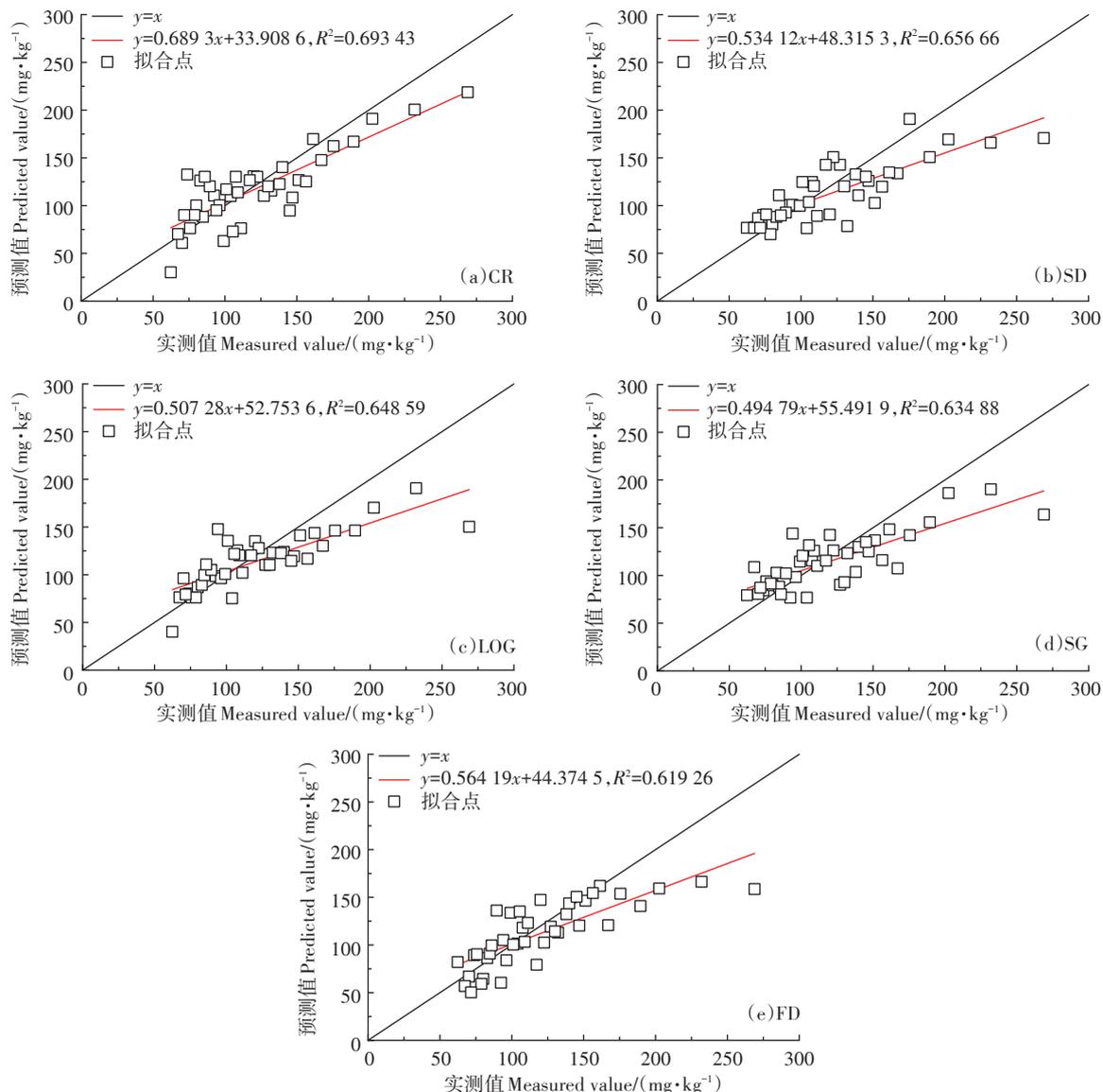


图5 土壤Zn含量实测值与预测值的拟合散点图

Figure 5 Fitting scatter plots of measured and predicted soil Zn content

(FD)、二阶微分(SD)和去包络线(CR)等方法提高提取的特征波段的准确性和模型的预测能力,有效突出光谱曲线的吸收和反射特征,从而提高了光谱的灵敏性并增强了有用信息。

(3)以5种反射率变换光谱数据为基础,提取出最佳特征波段作为模型自变量,对应土壤Zn含量作为模型因变量,构建了基于偏最小二乘法的西平县高标准基本农田建设区域土壤Zn属性反演模型。土壤Zn的最佳模型是以CR光谱变换为最佳的偏最小二乘模型。

本研究仅比较了通过不同光谱变换而挑选出的特征波段经同一建模方法建立反演模型的精度,没有

进行多种光谱变换方式、多种不同建模方式间的组合比较,土壤光谱数据还有待深入挖掘。此外本模型精度有待进一步提高,且当前模型仅适用于研究区,而对其他地区的适应性有待研究验证。

参考文献:

- [1] 杜鹰,张秀青,夏友仁,等.着力解决发展瓶颈问题 力促小麦生产再上新台阶——河南省小麦生产调研报告[J].价格理论与实践,2021(2):4-7,29. DU Y, ZHANG X Q, XIA Y R, et al. Efforts to solve the stagnate problem of development, and strive to promote wheat production to a higher level: Henan Province wheat production survey report[J]. *Price: Theory & Practice*, 2021(2):4-7, 29.
- [2] 李勤之,孙丽,杜俊洋,等.农用地土壤重金属锌的生态安全阈值研

- 究[J]. 环境科学学报, 2022, 42(7):408-420. LI M Z, SUN L, DU J Y, et al. Soil ecological safety thresholds for zinc in agricultural land[J]. *Acta Scientiae Circumstantiae*, 2022, 42(7):408-420.
- [3] 李金林. 铅锌污染土壤修复治理应用技术研究[J]. 资源节约与环保, 2020(2):53-54. LI J L. Research on application technology of lead-zinc contaminated soil remediation treatment[J]. *Resources Economization & Environmental Protection*, 2020(2):53-54.
- [4] 赵丹, 吴畏达, 王骥, 等. 土壤环境基线相关概念辨析[J]. 环境保护, 2020, 48(24):51-55. ZHAO D, WU W D, WANG B, et al. Differentiation and analysis of the concepts related to soil environmental baseline[J]. *Environmental Protection*, 2020, 48(24):51-55.
- [5] 梁红霞, 侯克斌, 陈富荣. 安徽池州地区富锌土壤资源分布特征及成因分析[J]. 安徽农业科学, 2022, 50(9):59-64. LIANG H X, HOU K B, CHEN F R. Analysis on distribution characteristics and causes of zinc-rich soil resources in Chizhou of Anhui Province[J]. *Journal of Anhui Agricultural Sciences*, 2022, 50(9):59-64.
- [6] HORTA A, AZEVEDO L, NEVES J, et al. Integrating portable X-ray fluorescence (pXRF) measurement uncertainty for accurate soil contamination mapping[J]. *Geoderma*, 2021, 382:114712.
- [7] 郭飞, 许镇, 马宏宏, 等. 基于PCA的土壤Cd含量高光谱反演模型对比研究[J]. 光谱学与光谱分析, 2021, 41(5):1625-1630. GUO F, XU Z, MA H H, et al. A comparative study of the hyperspectral inversion models based on the PCA for retrieving the Cd content in the soil[J]. *Spectroscopy and Spectral Analysis*, 2021, 41(5):1625-1630.
- [8] 郭斌, 白昊睿, 张波, 等. 基于RF和连续小波变换的露天煤矿土壤锌含量高光谱遥感反演[J]. 农业工程学报, 2022, 38(10):138-147. GUO B, BAI H R, ZHANG B, et al. Inversion of soil zinc contents using hyperspectral remote sensing based on random forest and continuous wavelet transform in an opencast coal mine[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, 38(10):138-147.
- [9] 周茱, 邹滨, 涂宇龙, 等. 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[J]. 光谱学与光谱分析, 2020, 40(7):2182-2187. ZHOU M, ZOU B, TU Y L, et al. Hyperspectral modeling of Pb content in mining area based on spectral feature band extracted from near standard soil samples[J]. *Spectroscopy and Spectral Analysis*, 2020, 40(7):2182-2187.
- [10] 陶超, 王亚晋, 邹滨, 等. 土壤重金属铅、锌高光谱反演模型可迁移能力分析研究[J]. 光谱学与光谱分析, 2018, 38(6):1850-1855. TAO C, WANG Y J, ZOU B, et al. Assessment and analysis of migrations of heavy metal lead and zinc in soil with hyperspectral inversion model[J]. *Spectroscopy and Spectral Analysis*, 2018, 38(6):1850-1855.
- [11] 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. WANG J F, WANG S J, BAI X Y, et al. Prediction soil heavy metal zinc based on spectral reflectance in karst area[J]. *Spectroscopy and Spectral Analysis*, 2019, 39(12):3873-3879.
- [12] 程先锋, 宋婷婷, 陈玉, 等. 滇西兰坪铅锌矿区土壤重金属含量的高光谱反演分析[J]. 岩石矿物学杂志, 2017, 36(1):60-69. CHENG X F, SONG T T, CHEN Y, et al. Retrieval and analysis of heavy metal content in soil based on measured spectra in the Lanping Zn-Pb mining area, western Yunnan Province[J]. *Acta Petrologica et Mineralogica*, 2017, 36(1):60-69.
- [13] 宋婷婷, 付秀丽, 陈玉, 等. 云南个旧矿区土壤锌污染遥感反演研究[J]. 遥感技术与应用, 2018, 33(1):88-95. SONG T T, FU X L, CHEN Y, et al. Remote sensing inversion of soil zinc pollution in Gejiu mining area of Yunnan[J]. *Remote Sensing Technology and Application*, 2018, 33(1):88-95.
- [14] TAN K, MA W, CHEN L, et al. Estimating the distribution trend of soil heavy metals in mining area from HyMap airborne hyperspectral imagery based on ensemble learning[J]. *Journal of Hazardous Materials*, 2021, 401:123288.
- [15] LIN N, LIU H, YANG J, et al. Hyperspectral estimation of soil composition contents based on kernel principal component analysis and machine learning model[J]. *Journal of Applied Remote Sensing*, 2020, 14(3):034507.
- [16] 郭鹏, 李华, 陈红艳, 等. 基于光谱指数优选的土壤盐分定量光谱估测[J]. 水土保持通报, 2018, 38(3):193-199, 205. GUO P, LI H, CHEN H Y, et al. Quantitative spectral estimation of soil salinity based on optimum spectral indices[J]. *Bulletin of Soil and Water Conservation*, 2018, 38(3):193-199, 205.
- [17] 张秋霞, 张合兵, 刘文锴, 等. 高标准基本农田建设区域土壤重金属含量的高光谱反演[J]. 农业工程学报, 2017, 33(12):230-239. ZHANG Q X, ZHANG H B, LIU W K, et al. Inversion of heavy metals content with hyperspectral reflectance in soil of well-facilitated capital farmland construction areas[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2017, 33(12):230-239.
- [18] 陆频频, 戴激光, 施先忠. 基于高光谱遥感的四种典型道路光谱特征分析[J]. 测绘与空间地理信息, 2019, 42(5):141-144. LU P P, DAI J G, SHI X Z. Analysis of spectral characteristics of four typical roads based on hyperspectral remote sensing[J]. *Geomatics & Spatial Information Technology*, 2019, 42(5):141-144.
- [19] 杨菲菲, 刘升平, 诸叶平, 等. 基于高光谱遥感的冬小麦涝渍胁迫识别及程度判别分析[J]. 智慧农业(中英文), 2021, 3(2):35-44. YANG F F, LIU S P, ZHU Y P, et al. Identification and level discrimination of waterlogging stress in winter wheat using hyperspectral remote sensing[J]. *Smart Agriculture*, 2021, 3(2):35-44.
- [20] 郭熙, 谢碧裕, 叶英聪, 等. 高光谱特征辨别渥育型麻沙泥田和潮沙泥田水稻土[J]. 农业工程学报, 2014, 30(21):184-191. GUO X, XIE B Y, YE Y C, et al. Discrimination between hydromorphic alluvial sandy mud paddy and tide sandy mud paddy based on hyperspectral characteristics[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2014, 30(21):184-191.
- [21] 韩雨, 刘焕军, 张新乐, 等. 基于冠层光谱的水稻穗颈病害程度预测模型[J]. 光谱学与光谱分析, 2021, 41(4):1220-1226. HAN Y, LIU H J, ZHANG X L, et al. Prediction model of rice panicles blast disease degree based on canopy hyperspectral reflectance[J]. *Spectroscopy and Spectral Analysis*, 2021, 41(4):1220-1226.
- [22] 林达. 基于典型相关分析的高分辨率遥感图像特征融合与分类方法研究[D]. 武汉: 武汉大学, 2016. LIN D. Feature fusion algorithm for high resolution remote sensing image classification based on canonical correlation analysis[D]. Wuhan: Wuhan University, 2016.

- [23] 冯正南, 宋文福. 浅析相关系数的显著性检验[J]. 大学物理实验, 2009, 22(3):85-87, 90. FENG Z N, SONG W F. Analysis of correlation coefficient significance test[J]. *Physical Experiment of College*, 2009, 22(3):85-87, 90.
- [24] 张运轲, 任登凤, 韩玉阁, 等. 基于连续投影算法提取特征波长的空中目标参考光谱选取[J]. 红外与激光工程, 2021, 50(4):240-250. ZHANG Y K, REN D F, HAN Y G, et al. Air target reference spectrum selection based on characteristic wavelengths extracted by successive projections algorithm[J]. *Infrared and Laser Engineering*, 2021, 50(4):240-250.
- [25] PENG X, SHI T, SONG A, et al. Estimating soil organic carbon using VIS/NIR spectroscopy with SVMR and SPA methods[J]. *Remote Sensing*, 2014, 6(4):2699-2717.
- [26] 王启元, 赵艳玲, 房铄东, 等. 基于多光谱遥感的裸土土壤含水量反演研究[J]. 矿业科学学报, 2020, 5(6):608-615. WANG Q Y, ZHAO Y L, FANG S D, et al. Inversion of soil moisture in bare soil based on multi-spectral remote sensing[J]. *Journal of Mining Science and Technology*, 2020, 5(6):608-615.
- [27] 赵明松, 谢毅, 陆龙妹, 等. 基于高光谱特征指数的土壤有机质含量建模[J]. 土壤学报, 2021, 58(1):42-54. ZHAO M S, XIE Y, LU L M, et al. Modeling for soil organic matter content based on hyperspectral feature indices[J]. *Acta Pedologica Sinica*, 2021, 58(1):42-54.
- [28] 赵超越, 张友华. 小麦赤霉病神经网络和偏最小二乘预测模型[J]. 吉首大学学报(自然科学版), 2017, 38(4):47-52. ZHAO C Y, ZHANG Y H. Prediction models of Neural network and partial least squares for wheat scab[J]. *Journal of Jishou University(Natural Sciences Edition)*, 2017, 38(4):47-52.
- [29] 钟燕. 基于HSI高光谱数据的耕地土壤重金属镉、铅含量遥感反演[D]. 雅安:四川农业大学, 2016. ZHONG Y. Remote sensing estimation of heavy metal Cd and Pb in cultivated soil based on HSI hyperspectral data[D]. Ya'an:Sichuan University, 2016.
- [30] 袁宇志, 郭颖, 张育灿, 等. 亚热带典型小流域景观格局对耕地土壤酸化的影响[J]. 土壤, 2019, 51(1):90-99. YUAN Y Z, GUO Y, ZHANG Y C, et al. Impacts of landscape patterns on farmland soil acidification in typical subtropical small watersheds of China[J]. *Soils*, 2019, 51(1):90-99.
- [31] 雷国建, 刘千钧, 陈志良, 等. 不同行业污染土壤重金属污染特征比较研究[J]. 土壤, 2013, 45(6):1023-1027. LEI G J, LIU Q J, CHEN Z L, et al. Comparative study on characters of soil heavy metal pollution in different industries[J]. *Soils*, 2013, 45(6):1023-1027.
- [32] 乔娟峰, 熊黑钢, 王小平, 等. 新疆阜康荒地土壤有机质高光谱特征及其反演模型研究[J]. 干旱地区农业研究, 2018, 36(5):207-214. QIAO J F, XIONG H G, WANG X P, et al. Wasteland soil organic matter hyperspectral characteristics and inversion model research in Fukang, Xinjiang[J]. *Agricultural Research in the Arid Areas*, 2018, 36(5):207-214.
- [33] 刘庆杰, 蔺启忠, 王钦军, 等. 基于连续快速傅里叶变换的红外光谱处理技术[J]. 光谱学与光谱分析, 2009, 29(12):3279-3282. LIU Q J, LIN Q Z, WANG Q J, et al. Continuum based fast fourier transform processing of infrared spectrum[J]. *Spectroscopy and Spectral Analysis*, 2009, 29(12):3279-3282.
- [34] 刘润, 张绍良, 侯湖平, 等. 基于思维进化优化BP神经网络的大豆叶片叶绿素含量高光谱反演[J]. 江苏农业科学, 2018, 46(13):212-216. LIU R, ZHANG S L, HOU H P, et al. Hyperspectral retrieval model of chlorophyll contents in soybean leaves based on mind evolutionary algorithm optimization - BP neural network[J]. *Jiangsu Agricultural Sciences*, 2018, 46(13):212-216.

(责任编辑:李丹)