

OLI 与 6SV 的褐土带煤炭开采沉陷区土壤有机碳反演

赵鑫¹, 徐占军^{1*}, 尹建平², 毕如田¹, 冯俊芳¹, 刘培³

1. 山西农业大学资源环境学院, 山西 晋中 030801
2. 中煤平朔集团有限公司节能环保部, 山西 朔州 036006
3. 河南理工大学测绘与国土信息工程学院, 河南 焦作 454000

摘要 遥感反演已广泛应用于区域土壤理化性质的动态监测, 但是鲜有针对有机碳含量低、下垫面不均一等土壤光谱特性不显著区域的研究。黄土高原褐土带地形多样, 丘陵广布, 有机碳含量低。采煤活动引起大面积土壤退化, 土壤光谱特性受到强烈干扰, 制约了区域尺度土壤有机碳(soil organic carbon)含量遥感反演精度。以山西省褐土带典型采煤沉陷区为例, 借助地表反射率和室外实地采集的样本数据对褐土带煤矿开采沉陷区土壤有机碳含量进行反演。采用结合高空间、时间分辨率辅助气象数据的 6SV(second simulation of a satellite signal in the solar spectrum-vector)模型和 FLAASH(fast line-of-sight atmospheric analysis of spectral hypercubes)模型对研究区 Landsat8 OLI 影像的大气校正方法进行对比实验, 分析其对褐土带采煤沉陷区土壤光谱曲线及有机碳含量的影响, 识别敏感波段。选择原始光谱反射率(R)和平方根(\sqrt{R})、倒数的对数 $\log(1/R)$ 、一阶微分(R')等数学变换形式, 利用多元线性回归(MLR)、BP 神经网络(BP neural net)和偏最小二乘回归(PLSR)建立土壤有机碳反演模型。结果表明: 6SV 模型大气校正的效果要优于 FLAASH 模型, 可以有效消除大气、地形对于反射率的干扰, 可见光波段反射率降低而近红外波段明显上升, 不同有机质含量等级土壤反射光谱特性分明; 640~670, 850~880, 1 570~1 600 和 2 110~2 290 nm 波段对土壤有机碳含量指示性强; 相较于多元线性回归(决定系数 R^2 为 0.765)、BP 神经网络(R^2 为 0.767), 偏最小二乘回归模型反演精度最高(R^2 为 0.778); 结合高空间、时间分辨率辅助气象数据的 6SV 大气校正模型与偏最小二乘回归建模能显著提高褐土带采煤沉陷区土壤有机碳的反演精度。在此基础上预测研究区 2013 年—2015 年土壤有机碳含量, 研究发现: 研究区土壤有机碳含量中部高, 两侧低, 复垦使土壤有机碳含量得到恢复。研究结果可用于揭示黄土高原褐土带采煤沉陷区土壤有机碳含量的时空分布特征, 为改进区域土壤光谱分析、土地复垦评价、建立褐土带采煤沉陷区碳通量观测网络和土壤碳库估算提供理论和技术支持, 对研究区域甚至全球范围褐土带生态可持续发展提供依据。

关键词 土壤有机碳; OLI; 6SV; 褐土带; 采煤沉陷区

中图分类号: TP79 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)03-0886-08

引言

土壤作为陆地生态系统中储量最大的碳库, 在储存 SOC、封存和释放 CO_2 方面发挥重要作用。而农田土壤作为降低陆地生态系统碳排放的最大潜在因素^[1-2], 受到煤炭开采的高强度扰动, 碳汇能力减弱而碳源能力增强。中国等世界煤炭开采大国必须将动态观测与定量研究相结合, 探求采

煤活动对区域农田碳储量的作用规律, 更精准地对采煤区土壤有机碳库进行可持续化管理, 完成煤炭绿色低碳开采。因此, 寻找合适的采煤沉陷区 SOC 含量空间预测方法意义重大。目前, 传统土壤性质分析方法因成本、效率问题难以满足连续调查与评估采煤沉陷区 SOC 时空变化的要求。

多光谱影像数据具有信息丰富、空间尺度大、成本低等特点, 运用多光谱遥感影像数据既可以建立精度较高的回归模型, 又能将反演结果通过影像可视化地表达, 利于探究宏

收稿日期: 2018-05-24, 修订日期: 2018-10-16

基金项目: 国家自然科学基金项目(51304130), 山西省青年科技研究基金项目(2015021125), 国土资源部公益性行业科研专项经费课题(201411007)资助

作者简介: 赵鑫, 1994 年生, 山西农业大学资源环境学院硕士研究生 e-mail: ZX18235446528@163.com

* 通讯联系人 e-mail: zjxul63@126.com

观尺度 SOC 含量时空分异规律。国内外学者基于不同的遥感数据源在 SOC 反演上取得了一定的成效。Summers^[3]等在 0.4~2.5 μm 范围内建立光谱反演模型预测南澳大利亚州 Belalie 草原土壤的有机碳含量。刘焕军等^[4]通过对多光谱影像与实测土壤样本数据进行相关性分析提取出土壤有机质敏感波段进行建模；顾晓鹤等^[5]研究发现，土壤有机质与 HJ1A-HSI 高光谱遥感影像 0.546, 0.664, 0.673 和 0.855 μm 处的反射率具有相关性。

当前利用多光谱数据在预测黑土、荒漠化土、红土、潮土等 SOC 固然取得良好效果，但研究地区地貌类型单一，对土壤光谱反射特性因剧烈人为活动而呈现紊乱状态的采煤沉陷区遥感反演研究不足，鲜有对居于黄土与棕壤过渡带、铁氧化物含量高的褐土带 SOC 含量的研究。遥感影像大气校正方法多基于 MODTRAN4+辐射传输模型，容易受研究区特点及目标类型等干扰，且输入参数局限于几种固定模式，空间、时间分辨率较低，校正结果精度受限。以往研究多数采用了 MLR，而 BP 神经网络、PLSR 等方法虽有应用但较少，建模方法也缺乏横向之间的精度对比。本研究区地处黄土高原褐土带，煤矿开采等人为活动引起了大范围地面沉降和生态破坏，直接作用于耕地并使土壤理化性质剧烈变化^[6-7]，土壤的反射光谱特性也随有机质含量、质地、水分含量等变化受到影响^[8]。在矿区土地复垦与生态修复过程中，选择 Landsat8 OLI 多光谱影像为数据源，配合针对 OLI 影像的 6SV 大气校正模型以及高空间、时间分辨率的辅助气象数据，在考虑到研究区地形起伏较大的特点上提取地表反射率，并建立 MLR, PLSR 和 BP 神经网络模型，开展对黄土高原褐土带采煤沉陷区农田 SOC 的定量预测与评估。

1 实验部分

1.1 研究区概况

长河流域位于山西省晋城市泽州县境内(35°30′—35°31′N, 112°36′—112°39′E)，包括川底、大东沟、下村三个乡镇 47 个村庄，属于温带大陆性气候，年均温 11.2 °C，最高温 40.2 °C，最低温 -18.7 °C，年均降水量为 616.8 mm。该区处于黄土高原东南端，境内以山地、丘陵为主，地形起伏多变，地势西高东低，海拔 730~1 185 m。研究区处于黄土向棕壤过渡地带，土壤成土母质主要为黄土性和黄土状母质，土壤类型主要为褐土，是黄土高原南部主要分布土类，易粘化，板结效应明显^[9]。区域内主要粮食作物是玉米、小麦、农地利用方式主要为玉米-小麦旱作方式。长河贯穿南北全境，属于沁河支流。流域内有 12 个正在开采的煤矿几乎覆盖流域西部，煤矿的生产规模均在 45 万 $\text{t} \cdot \text{a}^{-1}$ 以上。受长期采煤影响，流域内地表沉陷严重，土壤碳库受到极大扰动。

1.2 样品采集

研究区土壤类型为褐土，质地类别是壤土，土壤粘粒含量为 12.90%~47.70%；土壤 pH 范围 6.49~8.51，均值为 7.97；土壤有机质含量为 7.80~20.60 $\text{g} \cdot \text{kg}^{-1}$ ，均值为 13.18 $\text{g} \cdot \text{kg}^{-1}$ 。于 2015 年 7 月，以当地矿区为中心，在考虑地形图和实地勘探基础上，紧邻矿区的四周采用“Z”形路线

采样，使用网格法进行布点，网格规格为 1 km×1 km，在网格中选取相对平整的土壤进行样本采集，在 5 m 半径内，随机选择 3 个不同位置采集表层土壤组成 1 个混合样，采集深度为 0~20 cm，若遇到房屋、道路、河流等不适合采样的情况则选择临近同类土壤采取，共采集 117 个表层土壤样品，每份土样约 1 kg。使用 Garmin GPS72 记录采样点坐标，以确定其于遥感影像中的对应位置。为了减少混合像元的影响，提高端元丰度，在采样后又重新在研究区进行调研，根据地理坐标对样本点进行筛选，在保证样本点覆盖所有地貌类型的前提下，剔除影像单个像元范围内(30 m×30 m)出现建设用地的样本点，避免建设用地平滑表面导致的过高反射率对同像元内耕地反射率造成噪声干扰，使得提取的褐土带耕地反射率真实反映耕地的光谱特性。筛选后采样点共 63 个，如图 1 所示。带回室内，自然风干，研磨后剔除植物根系、小石块、动物残体等杂质，过 2 mm 筛子，然后采用重铬酸钾($\text{K}_2\text{Cr}_2\text{O}_7$)—外加热法测定样本 SOC 含量^[10]。

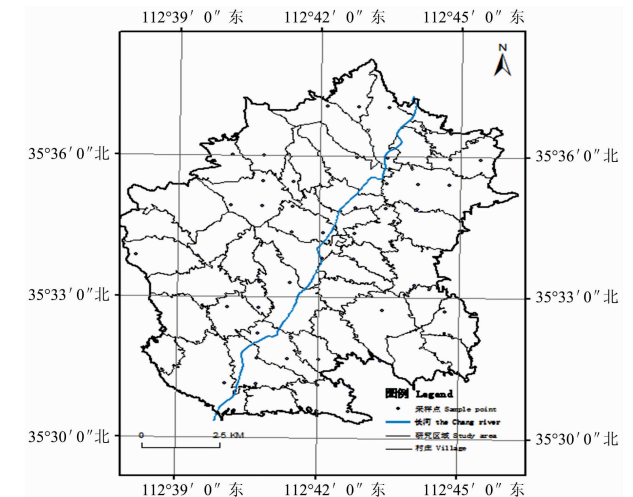


图 1 研究区土壤采样点分布

Fig. 1 Distribution of soil sampling points in study area

为满足经典统计分析建模的要求，对 SOC 数据进行正态性检验，结果服从正态分布。

1.3 影像选取与预处理

使用的 Landsat8 OLI 影像于 <https://earthexplorer.usgs.gov/> 下载。Landsat8 卫星 OLI 传感器相较于 Landsat7 的 ETM+ 传感器增加了 Coastal 波段，波长更短，利于气溶胶性质的检测。考虑到 7 月份遥感影像中植被、云量对于 SOM 反射光谱造成的信息干扰，选用 2013 年—2015 年春季、秋季的影像，轨道号 125-035，云量低于 2%，植被较少，地表裸露，可以较好的还原 SOM 原有的光谱信息，有利于提高 SOM 的反演精度。使用 ENVI5.3 对影像进行包括裁剪、几何校正(校正误差小于 1 个像元)的预处理。

整体技术路线流程图见图 2。

1.4 研究方法

1.4.1 大气校正

使用 ENVI5.3 对预处理后的影像进行大气校正，同时利用 RSD(remote sensing desktop)针对 Landsat8 OLI 的 6SV

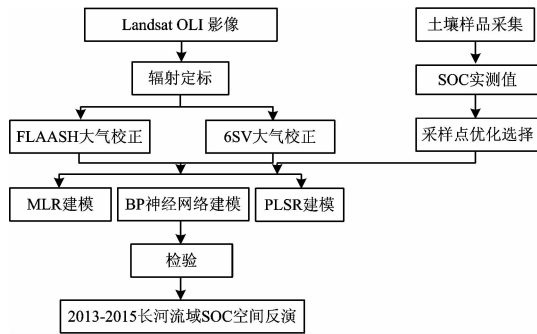


图 2 整体技术路线图

Fig. 2 Flow chart of overall technology

大气校正模块,调用成像当天辅助气象数据,对影像进行大气校正。

ENVI 中的 FLAASH 大气校正模型是基于 MORTRAN4+内核对定标后的影像进行大气校正。它假设太阳波谱范围内(不包括热辐射)地表为标准的平面朗伯体^[11],模型中辅助气象数据的输入局限于几种固定的模式,对于地形复杂区域大气校正的精度受限。

RSD 引进了 USGS 基于 6SV2.1 模型为 Landsat8 设计的大气校正程序 Landsat 8 Surface Reflectance Code (LaSRC)。首先引进了 LaSRC 的核心校正代码并使之适应 Windows 窗口环境,改造和重写了属性文件解析、辅助文件导入以及全部 I/O 部分代码,使之在 Windows 环境顺利运行。6SV 模型是 6S 模型的矢量版本,在精确地模拟卫星和平面观测高度、合理考虑异向性和非均一地表面、非朗伯体等因素的基础上,可以对连续有序的散射进行模拟,计算辐射极化,进而求解辐射传输方程,能够更加有效地消除瑞利散射和气溶胶的影响。

辅助气象数据由 USGS 通过 MODIS 逐日为 LaSRC 计算而得,与 FLAASH 相比,大气压、水蒸气、臭氧等参数有更高的时间、空间分辨率。此外 Levy 和 Vermote 等发现红、蓝波段比中红外、蓝波段反演气溶胶精度更高,以此结合 Landsat8 OLI Coastal 波段可以进行更高精度的大气校正。6SV 模型传感器接收到大气反射率为

$$R^*(\theta_s, \theta_v, \Phi_v) = T_g(\theta_s, \theta_v) \left\{ R_a(\theta_s, \theta_v, \Phi_v) + \frac{T(\theta_s)}{1 - R_p S} [R_c e_{\mu v}^- + R_p T_d(\theta_v)] \right\} \quad (1)$$

式中, $R^*(\theta_s, \theta_v, \Phi_v)$ 为传感器所接受的大气顶层表观反射率, $R_a(\theta_s, \theta_v, \Phi_v)$ 是瑞利散射和气溶胶散射引起的路径辐射, θ_s 为太阳天顶角, θ_v 为观测天顶角, Φ_v 为方位角; S 是大气球面反射率; $T(\theta_s)$ 为下行辐射总透射率, $T_d(\theta_v)$ 为上行辐射总透射率, $e_{\mu v}^-$ 是直接投射到传感器的上行辐射, $\mu v = \cos\theta_v$, 为卫星天顶角余弦; τ 是大气光学厚度; R_p 为非均一目标反射率, R_c 表示邻近效应^[12]。

1.4.2 光谱特征分析及特征波段选择

依据泽州县长河流域土地利用现状图和实地调查资料,

对 FLAASH, 6SV 影像进行 SOC 的光谱曲线分析,掌握矿区褐土 SOC 的光谱特征。利用波段数据进行 PLSR, 根据回归系数进行波段选择。最终在光谱特征分析和回归系数等信息的基础上,选择对 SOC 反映较好的特征光谱波段。

1.4.3 模型构建与精度比较

选取 50 个样点构成建模样本集, 13 个样本点构成验证样本集, 在所选特征波段基础上, 利用两种大气校正影像各波段反射率及其变换形式, 分别建立 MLR, BP 神经网络和 PLSR 的 SOC 反演模型。经过建模比较, FLAASH 大气校正影像 MLR, BP 神经网络模型选用平方根(\sqrt{R}), PLSR 模型选用 $\log(1/R)$, 6SV 大气校正影像选用一阶微分(R')的变换形式。MLR 在 SPSS 22 中实现, BP 神经网络在 MATLAB R2015b 中实现, PLSR 在 SIMCA-P 13.0 中实现。用验证集对模型精度进行检验, MLR 模型精度评价指标为决定系数 R^2 , 统计量 F 及显著性水平 Sig。BP 模型精度评价指标为决定系数 R^2 、均方根误差 RMSE。以决定系数 R^2 、累积交叉有效性系数 Q^2 评价 PLSR 模型的精度。大量模拟实验后确定 BP 模型参数设置如表 1 所示。

表 1 BP 模型主要参数设置

Table 1 Main parameters of BP model

BP 数据源	BP 神经网络	
	隐含层节点数	学习速率
FLAASH 影像	3	0.01
FLAASH' 影像	3	0.01
6SV 影像	4	0.01
6SV' 影像	4	0.01

注: ' 表示对原始反射率进行数学变换

Note: ' indicates mathematical transformation of the original data

2 结果与讨论

2.1 影像质量的评价和土壤光谱特征分析

为突出 SOC 变动特征, 以土壤有机质含量等级变化揭示有机碳含量变化特征。由表 2 可知, 研究区土壤有机质含量为 $7.80 \sim 20.60 \text{ g} \cdot \text{kg}^{-1}$, 均值为 $13.18 \text{ g} \cdot \text{kg}^{-1}$ 。根据第二次土壤普查养分分级标准, 以三级 ($20 \sim 30 \text{ g} \cdot \text{kg}^{-1}$)、四级 ($10 \sim 20 \text{ g} \cdot \text{kg}^{-1}$)、五级 ($6 \sim 10 \text{ g} \cdot \text{kg}^{-1}$) 3 级别划分土壤有机质含量。含量水平主要集中于四、五级两级, 总体水平较低。建模集和验证集的褐土样本统计特征与总集接近, 说明两个子集可以很好代表总体样本。根据野外采样点实测数据, 绘制并分析 2015 年三种影像上各级土壤有机质含量的

表 2 土壤有机质含量统计特征

Table 2 Statistical characteristics of SOM content

样本集	样本数	最小值/ ($\text{g} \cdot \text{kg}^{-1}$)	最大值/ ($\text{g} \cdot \text{kg}^{-1}$)	均值/ ($\text{g} \cdot \text{kg}^{-1}$)
总集	63	7.80	20.60	13.18
建模集	50	7.80	20.60	13.26
验证集	13	9.00	16.80	12.86

反射率,结果如图 3 所示。

总体而言:(1)图 3(a)和(b)两幅影像与图 3(c)影像相比,可见光波段反射率均有一定程度的降低,而红外波段反射率均得到一定程度的提高,对光学窗口内的波段而言,辐射亮度失真的关键在于散射,受大气氧分子、水分子吸收作用的影响并不明显,但在更长的红外窗口波段,大气的主要影响是吸收,而非散射。瑞利散射和无选择性散射直接作用于可见光波段,一部分未到达地面的辐射经大气散射后直接进入传感器,传感器接收到无效信号,导致图像质量下降。而水汽、CO₂ 和 O₃ 等对于红外波段的强吸收,使得传感器接收到该波段的反射能量衰减,真实反射光谱信息丢失严重。通过大气校正摒除了大气散射、吸收的影响后,无效辐射能量对光谱细节干扰明显减弱。在可见光波段,土壤反射率降低;在近红外和短波红外波段,土壤反射率增高。说明 FLAASH 大气校正和 6SV 大气校正均可减弱大气吸收和散射的影响;(2)即使不同采样点土壤有机质含量不同,但是大气校正以后它们的土壤光谱曲线形态十分相似,且总体呈递增趋势。

图 3(a)为 FLAASH 影像各级土壤有机质含量的光谱反射率,其反射率在可见光波段较低,五级有机质的反射率相较于三、四级有机质低,而四级有机质反射率相较于三级有机质含量来说要高,呈现出土壤有机质含量与反射率不相对应的状态,各级土壤有机质可见光波段反射率可分性差,这是因为 FLAASH 大气校正假设在太阳波谱范围内(不包括热辐射)地表为标准的平面朗伯体,并没有考虑地表非均一的状态,对于地形起伏较大、地物类型多样的环境校正效果并不准确,并且 FLAASH 大气校正对于气溶胶、大气模型等参数相较 6S 大气校正分为几种固定模式,并没有影像当天详细的气象数据,在微观区域上进行大气校正精度不足。四级土壤有机质红外波段与可见光波段差值仅为 0.08,而波动

最大的三级土壤有机质红外波段与可见光波段差值不超过 0.19,说明 FLAASH 影像包含的信息量较少,区分不同等级土壤有机质含量的能力较低。

图 3(b)为 6SV 影像各级土壤有机质含量的光谱反射率,其反射率在可见光波段较低,三级土壤有机质可见光波段反射率最低,五级土壤有机质可见光波段反射率最高,土壤有机质含量与反射率各自对应,反射率可分性强。这是因为 6SV 模型不仅考虑了地表为非均一朗伯平面体的情况,并且结合从 MODIS 影像中计算得到的当天的气象辅助数据,对于大气环境、气溶胶等参数有精确的记录,可以计算极化辐射,能大幅度提高地表反射率的提取精度。土壤有机质含量与土壤可见光(1~4 Band: 0.43~0.45, 0.45~0.51, 0.53~0.59, 0.59~0.67 μm)反射率总体上呈负相关关系,反射率随土壤有机质的增加而减少,沙普明等的研究也反映了这个规律。一方面因为有机质含量低,湿度低、质地粗的土壤对可见光的吸收率低,反射率高,而研究区土壤为褐土,湿度低,质地粗,故而会出现低含量土壤有机质反射率高于高含量土壤有机质的现象;另一方面由于采矿活动中穿孔、凿岩、爆破、装运、破碎等工序的影响,耕地表面矿物粉尘覆盖增多,且褐土本身铁元素会发生水解,氧化和迁移比较明显,氧化铁含量较未扰动土壤增加,会导致可见光波段土壤反射率的明显下降,这也在一定程度上说明了 6SV 影像可以精准反映褐土光谱特性。图 3(b)中三级土壤有机质含量因为大于 2%,所引起的土壤反射率的下降可以掩盖其他成分的光谱特征。导致近红外波段受到铁氧化物的吸收作用影响减弱,反射率得到提高。由图 3(b)可知各级土壤有机质可见光波段与红外波段差值均达到 0.18,且波动最大的三级土壤有机质红外波段与可见光波段差值达到 0.25,说明 6SV 影像包含的信息量较多,6SV 影像区分不同等级土壤有机质能力较高。

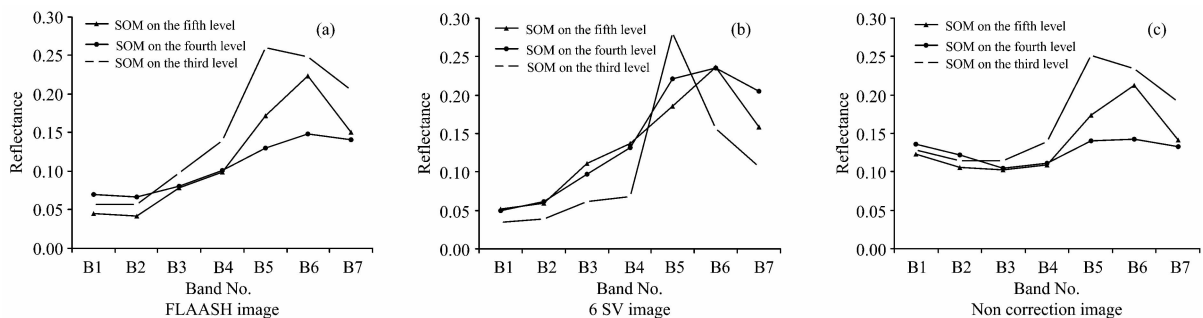


图 3 褐土光谱曲线

Fig. 3 Spectral curves of cinnamon soil

2.2 特征波段的选择

以 PLSR 分别对 FLAASH 影像、6SV 影像 7 个波段的反射率与土壤有机质含量之间的关系进行分析,并绘制回归系数曲线,如图 4 所示。FLAASH 影像的曲线于 1—3 波段平缓上升,3—5 波段上升速度加快,5—7 波段迅速下降,最高峰位于 5 波段;6SV 影像曲线于 1—2 波段短暂上升后,在 2—4 波段连续下降,4—5 波段又转为上升,5—7 波段间下降上升波动强烈,最高峰同样在 5 波段。两幅影像曲线在

4—7 波段均存在波峰或波谷,与图 3 光谱曲线特征相近,可见 4—7 波段对于反映土壤有机质含量作用明显。此外将 FLAASH 影像与 6SV 影像回归系数取绝对值对比,后者明显大于前者,即 6SV 大气校正所得地表反射率对于褐土带矿区土壤有机质信息细节捕捉更优,除 5 波段外,两影像波段回归系数绝对值差值较小,说明波段间存在相关性。因此以 6SV 影像为主要参考,兼顾 FLAASH 影像,综合考虑曲线变化特征波段、回归系数与邻近波段的相关性,最终确定土

壤有机质的特征波段 4—7。

2.3 建模结果

对于 FLAASH 影像和 6SV 影像,以选择的特征波段反

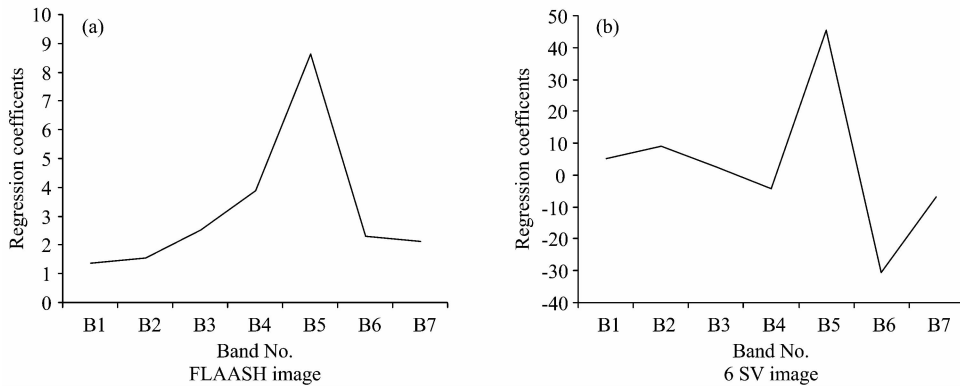


图 4 偏最小二乘回归系数

Fig. 4 Partial least squares regression coefficients

表 3 SOC 反演模型与精度

Table 3 SOC retrieval models and accuracies

模型名称	评价指标	FLAASH 影像	FLAASH' 影像	6SV 影像	6SV' 影像
多元线性回归	R^2	0.063	0.067	0.765	0.765
	F	1.818	1.879	40.879**	40.820**
BP 神经网络	R^2	0.385	0.389	0.764	0.767
	RMSE	0.133	0.116	0.190	0.189
偏最小二乘回归	R^2	0.060	0.141	0.775	0.778
	Q^2	0.003	0.093	0.706	0.764

注: ** 表示 $p < 0.01$ 水平显著

Note: ** means significant at $p < 0.01$ level

在未进行数学变换的情况下,FLAASH 的 MLR 模型的 R^2 仅为 0.063,稍优于 PLSR 模型,最佳模型为 BP 模型, R^2 为 0.385, RMSE 为 0.133。6SV 影像的 MLR 模型 R^2 为 0.765,优于 BP 模型,PLSR 模型 R^2 高于 MLR 模型,为最优模型。

在对土壤反射率数据进行相应的数学变换之后,FLAASH 影像所有模型的精度均得到提高;6SV 影像 MLR 模型 F 检验值低于变换之前,模型精度略有下降,除此之外其他模型精度均有相应改善。这在一定程度上说明对土壤反射率进行适当的数学变换可以降低影像预处理过程中没有消除掉的噪声。FLAASH 影像的 PLSR 模型 R^2 为 0.141,优于 MLR 模型,最佳模型仍为 BP 模型, R^2 为 0.389, RMSE 为 0.116。6SV 影像的 MLR 模型的 R^2 为 0.765,为三种模型中最低。BP 模型略优于 MLR 模型,最优模型仍为 PLSR 模型, R^2 高达 0.778, Q^2 为 0.764,接近 1,说明模型预测效果与逼近效果相近,建模效果良好。

由结果可知,6SV 影像的建模效果远远优于 FLAASH 影像。这表明 6SV 大气校正因为考虑到了非均一朗伯体的问题而较 FLAASH 在微观区域的大气校正上精度有所提高。此外由于气溶胶、大气环境等辅助性数据的输入并非仅仅局

限于几种固定的模式进行套用,而是通过 USGS 官网所提供的相关辅助数据中进行特定区域特定时间的气象信息读取,在空间和时间分辨率上较普通的 FLAASH 大气校正有了明显的提高,所以大气校正后地表反射率的精度更为接近真实的地表反射率,在反演相应的土壤理化性质上更有优势。

再者,由上述比较可知对土壤反射率进行适当的数学变换一定程度上可以提高模型的精度,这是因为数学变换可以降低影像预处理过程中没有消除掉的噪声,使得土壤有机质与地表反射率之间的关系更为显著。

从具体模型上来看,MLR 在三种模型中建模效果最差,这可能是由于不能完全消除特征波段中存在的线性相关问题,导致函数逼近过程中一部分重合信息被多次利用,降低了模型的精度。BP 模型优于 MLR 模型而低于 PLSR 模型,主要因为 BP 神经网络对于样本量的要求低于经典的统计分析模型,所以在样本量较少的情况下 BP 模型精度要高于 MLR 模型,但是 BP 模型样本依赖性严重且一定程度上会有过拟合现象的出现,所以在预测精度上要低于 PLSR 模型。而由于 PLSR 模型对数据进行降维后可以更有效地把握数据的特征信息,更容易辨识系统信息与噪声,甚至一些非随机噪声,避免了因子内部高度线性相关、因子结构不确定性和模型不能识别等问题,在提取出自变量最大变异信息的同时,还具有一个响应矩阵对因变量进行最优模拟,有效克服了典型相关分析、多元线性相关分析和主成分分析的不足,并继承其优点^[13-14],因此它的预测效果要优于一般的模型。

由图 5 基于 6SV 影像的三种模型验证散点图可知,PLSR 模型检验样本拟合效果最好, R^2 为 0.702, BP 神经网络次之, R^2 为 0.696, MLR 模型验证效果最差, R^2 为 0.675。

2.4 多时相褐土带采煤沉陷区反演结果图像

利用 6SV 影像和样本数据建立 PLSR 模型,对研究区 2013 年—2015 年土壤有机质含量进行预测,并乘以 Van Bemmelen 因数,得到有机碳反演结果,如图 6、表 4 所示。

褐土矿区有机质含量中部高、两侧低,有机碳含量最低

区域集中在西北部。西北部地区海拔高，丘陵密布，重力作用加速地表物质的迁移，土层薄弱，抗蚀抗冲性差，为矿区最早开发的地区，扰动时间长，扰动强度大，地表沉陷、水土流失严重，地表裂缝分布多，SOC 含量降低；中部地区开发晚，扰动强度较低，故而含量较高。由表可知，自 2013 年矿区进行复垦以来，SOC 含量 $< 5 \text{ g} \cdot \text{kg}^{-1}$ 的土地面积逐年

减少，而 SOC 含量 $5 \sim 15 \text{ g} \cdot \text{kg}^{-1}$ 的土地面积逐年增多，截至 2015 年，SOC 含量 $< 5 \text{ g} \cdot \text{kg}^{-1}$ 的土地由 0.14% 降低至 0.09%，而 SOC 含量 $10 \sim 15 \text{ g} \cdot \text{kg}^{-1}$ 的土地由 0.04% 增加至 15.87%。结果表明，长河流域褐土矿区复垦稳固甚至增加了 SOC 的含量，复垦对于矿区生态稳定和环境恢复具有积极作用。

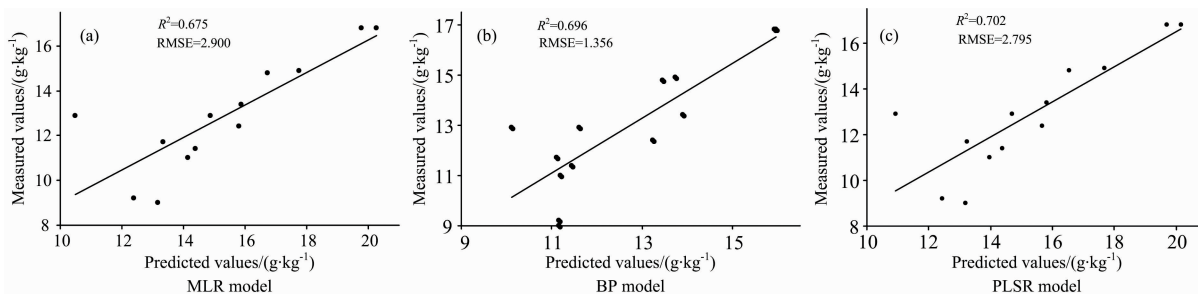


图 5 验证散点图

Fig. 5 Scatter plot of validation

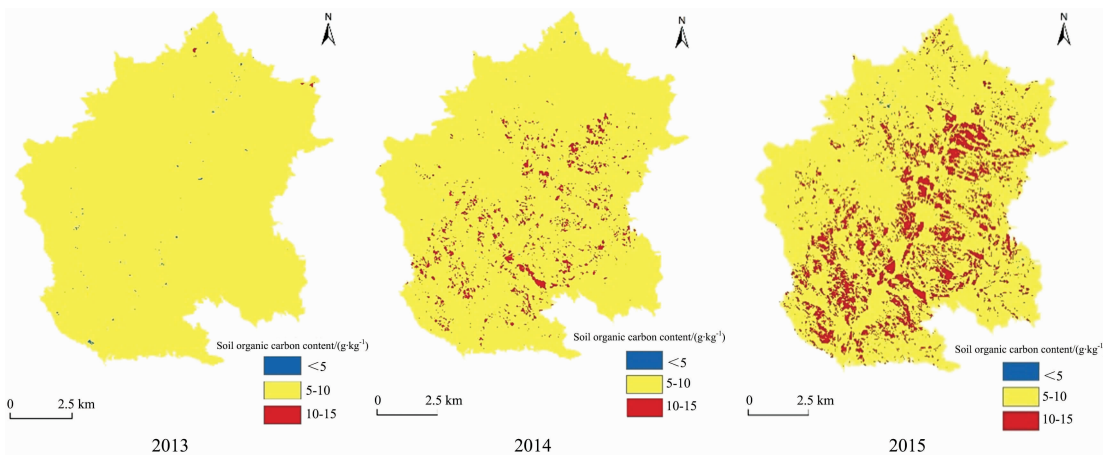


图 6 2013 年—2015 年 SOC 反演结果

Fig. 6 SOC retrieval results from 2013 to 2015

表 4 土壤有机碳反演结果统计表

Table 4 Statistical table of SOC results

土壤有机碳含量 / $(\text{g} \cdot \text{kg}^{-1})$	占图幅百分比/%		
	2013	2014	2015
< 5	0.14	0.04	0.09
$5 \sim 10$	99.82	95.75	84.04
$10 \sim 15$	0.04	4.21	15.87

3 结 论

尝试基于 Landsat8 OLI 的 6SV 大气校正模型提取地表反射率，在建立三种数据模型的基础上，利用 PLSR 模型对 SOC 进行反演建模，取得良好的效果，主要结论如下：

(1) 基于 Landsat8 OLI 的 6SV 大气校正优于 FLAASH 大气校正。利用 6SV 大气校正的影像对 SOC 进行定量反演，取得优于 FLAASH 大气校正的精度，说明针对 Landsat8

OLI 的 6SV 大气校正模型在考虑到微观区域地表不均一性以及结合高空间、时间分辨率的气象辅助数据后，比一般的 FLAASH 大气校正模型更能真实反映矿区 SOC 的反射率特征，此方法切实可行。

(2) 在非线性关系建模方面，PLSR 和 BP 神经网络模型要优于 MLR 模型。对于黄土高原褐土矿区 SOC 的反演，PLSR 与 BP 神经网络模型相较于 MLR 无论是逼近效果还是预测效果都有一定程度的提高。说明 PLSR 和 BP 神经网络模型能更好的模拟 SOC 和地表反射率之间复杂的非线性关系。三种模型中 PLSR 模型精度最高，一定程度上说明 PLSR 作为一种新型的建模方法，性能较好且稳定。

(3) 适当的数据处理有助于提高模型反演精度。对地表反射率数据进行相应的数学变换后，模型精度得到改善。

(4) 长河流域褐土矿区 SOC 含量中部高、两侧低，含量最低区域集中在西北部。复垦稳固甚至增加了 SOC 的含量，复垦对于矿区生态稳定和环境恢复具有积极作用。

(5) 基于 Landsat8 OLI 的 6SV 大气校正模型适合矿区等

地形复杂区域,在探究褐土带土壤光谱特性、明晰土壤理化性质(如 SOC 含量)、掌握土壤碳库动态变化规律、优化区域

环境监测等方面应用前景广阔,对区域甚至全球范围褐土带生态可持续发展的研究具有借鉴意义。

References

- [1] Miller B A, Koszinski S, Wehrhan M, et al. *Soil*, 2015, 1(1): 217.
- [2] Zhang G S, Ni Z W. *Soil & Tillage Research*, 2017, 165: 294.
- [3] Summers D, Lewis M, Ostendorf B, et al. *Ecological Indicators*, 2011, 11(1): 123.
- [4] LIU Huan-jun, ZHAO Chun-jiang, WANG Ji-hua, et al(刘焕军, 赵春江, 王纪华). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2011, 27(8): 211.
- [5] GU Xiao-he, WANG Kun, PAN Yu-chun, et al(顾晓鹤, 王 堃, 潘瑜春, 等). *Geography and Geoinformation Science(地理与地理信息科学)*, 2011, 27(6): 69.
- [6] LI Li, WANG Yue-xin, WANG Wei-bin(栗 丽, 王曰鑫, 王卫斌). *Chinese Journal of Soil Science(土壤通报)*, 2010, 41(5): 1237.
- [7] JIAO Xiao-yan, WANG Li-ge, LU Chao-dong, et al(焦晓燕, 王立革, 卢朝东, 等). *Journal of Soil and Water Conservation(水土保持学报)*, 2009, 23(4): 123.
- [8] PENG Jie, ZHANG Yang-zhu, ZHOU Qing, et al(彭 杰, 张杨珠, 周 清, 等). *Chinese Journal of Soil Science(土壤通报)*, 2009, 40(5): 1204.
- [9] LIU Yao-zong, ZHANG Ji-yuan, et al(刘耀宗, 张纪元, 等). *Soil of Shanxi(山西土壤)*. Beijing: Science Press(北京: 科学出版社), 1992. 84.
- [10] BAO Shi-dan(鲍士旦). *Soil Agricultural Chemistry Analysis(土壤农化分析)*. Beijing: China Agriculture Press(北京: 中国农业出版社), 2013. 30.
- [11] Research System Inc. FLAASH Module User's Guide, ENVI FLAASH Version 4.1, September, 2004 Edition.
- [12] Eric Vermote, Chris Justice, Martin Claverie, et al. *Remote Sensing of Environment*, 2016, 185: 46.
- [13] Savitzky A, Golaym J E. *Journal of Arid Environments*, 2010, 74(2): 229.
- [14] Hu X Y. *Pedosphere*, 2013, 23(4): 417.

Retrieval of Soil Organic Carbon in Cinnamon Mining Belt Subsidence Area Based on OLI and 6SV

ZHAO Xin¹, XU Zhan-jun^{1*}, YIN Jian-ping², BI Ru-tian¹, FENG Jun-fang¹, LIU Pei³

1. College of Resources and Environment, Shanxi Agricultural University, Jinzhong 030801, China

2. Energy-Saving and Environment Protection Department, China Coal Pingshuo Group Co., Ltd., Shuozhou 036006, China

3. School of Surveying and Land Information Engineering, Henan Polytechnic University, Jiaozuo 454000, China

Abstract Remote sensing retrieval has been widely used for dynamic monitoring of the physical and chemical properties of regional soil, but there are few studies on the areas with low organic carbon content and uneven underlying surface which have unremarkable soil spectral characteristics. The cinnamon soil belt in Loess Plateau has multiplicity topography, widely distributed hills and low organic matter content. Large areas soil degradation caused by mining activities has resulted in the fact that soil spectroscopy characteristics are strongly disturbed, which has some inhibiting effect on the remote sensing retrieval accuracy of soil organic carbon content at the regional scale. Based on cinnamon soil belt with typical coal mining subsidence area in Shanxi Province as an example, this research used the surface reflectance and outdoor sample data from the field of coal mining subsidence area to retrieve soil organic carbon content. Conducting comparative experiments on the atmospheric correction methods of the Landsat8 OLI image in the study area by the FLAASH model and the 6SV model combined with high spatial and temporal resolution aided meteorological data to analyze the effect on soil spectral curve and organic carbon content in the mining subsidence area of the cinnamon soil belt and recognize sensitive bands. Multiple linear regression(MLR), BP neural network(BP) and partial least squares regression(PLSR) model were established to retrieve soil organic carbon content by using the original spectral reflectance R and mathematical transformation forms such as \sqrt{R} , $\log(1/R)$ and R' . The results showed that the atmospheric correction effect of the 6SV model was better than that of the FLAASH model which could effectively eliminate the interference of atmosphere and topography to reflectance. The reflectance of visible light decreased and the near-infrared rose obviously. The soil reflectance spectra of different organic matter content was clear. The bands of 640~670, 850~880, 1 570~1 600, 2 110~

2 290 nm were highly indicative of soil organic carbon content. Compared with multiple linear regression (Coefficient of determination R^2 was 0.765) and BP neural network (R^2 was 0.767), the partial least-squares regression model had the highest retrieval accuracy (R^2 was 0.778). It was found that the 6SV atmospheric correction model and partial least squares regression modeling combined with aided meteorological data which had high spatial and temporal resolution could significantly improve the retrieval accuracy of soil organic carbon in the mining subsidence area of the cinnamon belt. The soil organic carbon content in the study area from 2013 to 2015 was retrieved based on this model. Results showed that the soil organic carbon content in the middle of the study area was higher than that in both sides, and the soil organic carbon content was restored by reclamation. The results can be used to reveal the spatial-temporal distribution of soil organic carbon in the mining subsidence area of the cinnamon belt in the Loess Plateau, providing theoretical and technical support for improving regional soil spectral analysis, land reclamation evaluation, establishment of carbon flux observation network in mining subsidence area of the cinnamon belt and estimation of soil carbon pool, which provides the basis for the ecological sustainable development of the cinnamon belt in the regional and global scales.

Keywords Soil organic carbon; OLI; 6SV; Cinnamon soil; Mining subsidence area

(Received May 24, 2018; accepted Oct. 16, 2018)

* Corresponding author

关于《光谱学与光谱分析》调整审稿费收费标准的通知

尊敬的《光谱学与光谱分析》广大作者、读者：我刊自 2018 年 7 月 1 日以后登记的稿件向投稿作者收取审稿费 200 元/篇，在您投稿之前，为免受经济损失，请您必须考虑：

1. 没有创新的一般性稿件，请您不要投稿。
2. 没有国家级基金资助的稿件，请您不要投稿。
3. 不是光谱专业的稿件，请您不要投稿。
4. 与其他文章重合率超过 10% 的稿件，请您不要投稿。

所投稿件经初审通过后，作者会收到缴纳审稿费的通知。请作者及时从我刊网站(<http://www.gpxygpx.com>)查询稿件是否处于交审稿费状态，在收到通知后，请及时缴纳审稿费；如在 10 天之内没有收到您的审稿费，被视为自动放弃，我刊不再受理。交费后我刊开据增值税电子普通发票，并传至作者提供的电子邮箱，作者可自行打印。

联系电话：010-62181070, 62182998

电子邮箱：chngpxygpx@vip.sina.com

感谢您多年来对《光谱学与光谱分析》的支持和厚爱！

《光谱学与光谱分析》期刊社

2018 年 6 月 30 日