

基于非线性变换与分数阶微分的土壤有机碳光谱估测

刘璇皓¹, 张晓萍¹, 文艳², 胡志斌¹, 段辰¹, 孙贯芳¹, 陈庆伟²

(1. 西北农林科技大学水土保持科学与工程学院(水土保持研究所), 水土保持与荒漠化整治全国重点实验室, 陕西 杨凌 712100; 2. 国网山东省电力公司经济技术研究院, 济南 250021)

摘要: [目的] 以陕西省延安市吴起县和志丹县的 213 个林地表层(0~20 cm)土壤样本为研究对象, 采集可见-近红外(Vis-NIR)光谱及实测土壤有机碳(SOC)质量分数, 探究非线性数学变换结合分数阶微分(FOD)处理对光谱特征挖掘及模型性能的影响, 为提高黄土高原地表 SOC 质量分数的光谱估测精度及智能化辨识提供支持。[方法] 对原始光谱(R)采取倒数($1/R$)、对数($\lg R$)、对数倒数($1/\lg R$)、平方根($\sqrt{R}$)等 4 种非线性变换, 及 0~2 阶(步长 0.1)的 FOD 处理, 利用竞争性自适应重加权采样(CARS)和连续投影算法(SPA)筛选特征波段, 构建并对比偏最小二乘回归(PLSR)、随机森林(RF)和极限梯度提升(XGBoost)3 种模型的反演效果。[结果] 研究区林地土壤 SOC 质量分数为 2.50~23.11 g/kg, 平均值为 7.70 g/kg。在原始反射光谱(R)中, SOC 质量分数较高的样本整体反射率水平较低, 光谱反射率随 SOC 升高呈整体降低趋势, 光谱特征波段集中在 550~800 nm 及 900 nm 附近。相较于整数阶和 >1.5 的高阶微分, 采用 0.3~0.9 阶的 FOD 数据预处理, 能有效去除基线漂移并凸显细部光谱特征。对数与倒数的非线性变换可以显著提升光谱对 SOC 的辨识程度。相比 SPA, CARS 在测试集上 R^2 平均提升 0.052, RMSE 平均降低 0.147 g/kg, RPD 平均提升 0.132。在所有模型中, PLSR 的预测性能整体优于 RF 和 XGBoost, 基于非线性变换与 FOD 处理的 CARS-PLSR 模型测试集 R^2 可达 0.843。[结论] 研究证实, 非线性变换结合中低阶分数阶微分及 CARS-PLSR 建模策略, 可有效提升黄土高原土壤有机碳的光谱估测精度。

关键词: 土壤有机碳; 可见-近红外光谱; 分数阶微分; 非线性变换; 黄土高原

中图分类号: S153.6; S127

文献标识码: A

文章编号: 1009-2242(2026)05-0225-13

Spectral Estimation of Soil Organic Carbon Based on Nonlinear Transformations and Fractional Differentiation

Liu Xuanhao¹, Zhang Xiaoping¹, Wen Yan², Hu Zhibin¹, Duan Chen¹, Sun Guanfang¹, Chen Qingwei²

(1. State Key Laboratory of Soil and Water Conservation and Desertification Control, College of Soil and Water Conservation Science and Engineering (Institute of Soil and Water Conservation), Northwest A&F University, Yangling, Shaanxi 712100, China; 2. Economic & Technology Research Institute of State Grid Shandong Electric Power Company, Jinan 250021, China)

Abstract: [Objective] Using 213 forest soil samples (0–20 cm depth) from Wuqi and Zhidan counties in Yan'an City, Shaanxi Province, visible-near infrared (Vis-NIR) spectra and measured soil organic carbon (SOC) content are collected to investigate the effects of nonlinear mathematical transformations combined with fractional-order differential (FOD) processing on spectral feature extraction and model performance. This study aims to provide support for improving the accuracy of spectral estimation and intelligent identification of surface SOC content in the Loess Plateau. [Methods] Four nonlinear transformations-reciprocal ($1/R$), logarithmic ($\lg R$), reciprocal logarithmic ($1/\lg R$), and square root ($\sqrt{R}$)-were applied to raw spectra (R), followed by fractional-order differential (FOD) processing at orders 0–2 (increments of 0.1). The competitive adaptive reweighted

收稿日期: 2026-01-16

修回日期: 2026-03-03

录用日期: 2026-03-10

网络首发日期(www.cnki.net): 2026-05-21

资助项目: 国家电网有限公司总部管理科技项目(5200-202416091A-1-1-ZN)

第一作者: 刘璇皓(2000—), 男, 硕士研究生, 主要从事遥感应用研究。E-mail: lxh00113@nwafu.edu.cn

通信作者: 张晓萍(1971—), 女, 博士, 教授, 主要从事水土保持及遥感应用研究。E-mail: zhangxp@ms.iswc.ac.cn

http://stbcbx.alljournal.com.cn

sampling (CARS) and successive projection algorithm (SPA) were employed to screen feature bands. Partial least squares regression (PLSR), random forest (RF), and extreme gradient boosting (XGBoost) models were constructed, and their performance was compared. [Results] The soil organic carbon (SOC) content in the forestland of the study area ranged from 2.50 g/kg to 23.11 g/kg, with an average value of 7.70 g/kg. In the raw reflectance spectrum (R), samples with higher SOC content exhibited lower overall reflectance levels. Spectral reflectance showed an overall decreasing trend with increasing SOC, with the spectral feature bands concentrated around 550 — 800 nm and 900 nm. Compared to integer-order and higher-order differentials (>1.5), preprocessing with FOD of orders 0.3 — 0.9 more effectively eliminated baseline drift and highlighted subtle spectral features. Nonlinear transformations, such as logarithmic and reciprocal transformations, significantly enhanced the spectral identification of soil organic carbon. Compared to SPA, CARS achieved an average R^2 improvement of 0.052, an average RMSE reduction of 0.147 g/kg, and an average RPD increase of 0.132 on the test set. Among all models, PLSR consistently outperformed RF and XGBoost in predictive performance. The CARS-PLSR model, incorporating nonlinear transformations and FOD processing, attained an R^2 of 0.843 on the test set. [Conclusion] The findings confirm that combining nonlinear transformations with low-to-medium-order fractional differentials and the CARS-PLSR modeling strategy effectively enhances the spectral estimation accuracy for soil organic carbon in the Loess Plateau.

Keywords: soil organic carbon; visible-near-infrared spectroscopy; fractional-order differential; nonlinear transformation; Loess Plateau

Received: 2026-01-16

Revised: 2026-03-03

Accepted: 2026-03-10

Online(www.cnki.net): 2026-05-21

土壤是陆地生态系统最大的碳库,其碳储量是植被碳库的 2.5 倍、大气碳库的 3.0 倍,是全球碳循环的核心^[1]。土壤有机碳(SOC)是土壤中有有机质(SOM)所含碳的部分,在改善土壤理化性质、提高生物多样性、提升土壤质量、维持土地生产力等方面发挥着关键作用^[2]。土壤有机碳质量分数的快速估测对土地的碳库估算、精准农业、智慧管理及高效利用等有着重要意义^[3]。传统的化学分析方法虽然精度较高,但需耗费大量人力、物力和财力,难以实现大范围快速监测。可见-近红外光谱估测技术具有快速、无损、便捷等优势,已成为 SOC 质量分数估测的有效方法^[4]。

基于光谱学原理预测土壤要素含量主要包括光谱数据预处理、特征波段筛选和构建模型 3 个步骤^[5]。其中,光谱数据预处理和特征波段选取,决定着所建模型的预测能力和可靠性,是光谱分析中重要的应用方向^[6]。诸多学者在光谱数据预处理和敏感波段筛选方面开展了相关研究。常见的光谱预处理方法包括多元散射校正、标准正态变换、非线性数学变换、小波变换^[7]和光谱微分等。光谱微分作为常用的信号处理技术,能有效减轻光谱曲线的基线漂移,在光谱数据处理中得到广泛应用^[8-10],其主要包括整数阶微分(一阶微分、二阶微分)和分数阶微分(fractional-order differential, FOD)。郭艳萍等^[11]对原始光谱数据进行倒数、对数和对数倒数等数学变换,并对变换后的结果进行一阶和二阶微分处理,得

到更为丰富的敏感波段,提高土壤盐分估算模型的精度,表明对光谱曲线进行非线性数学变换结合整数阶微分是增强光谱数据特征的有效途径之一。传统的整数阶微分虽然可以增强光谱特征,但在数据处理过程中容易遗漏、丢失部分光谱数据信息。与整数阶微分相比,分数阶微分拥有更小的阶次间隔且变化可控,可以突出高光谱数据中细微的光谱特征,增强光谱反射率信息。现有土壤 SOC 光谱反演研究在预处理环节多采用少数常见的光谱变换进行组合处理,方法体系相对单一。将多种数学变换与分数阶微分联合用于光谱预处理的工作仍较少,且其对后续敏感波段提取及模型构建效果的影响尚缺乏系统评估。

鉴于高光谱数据具有波段数量多、信息冗余度高和维度庞大等特征,已有研究^[12-13]常引入主成分分析(principal component analysis, PCA)、遗传算法(genetic algorithm, GA)、偏最小二乘(partial least squares regression, PLSR)、连续投影算法(successive projections algorithm, SPA)及竞争性自适应重加权采样(competitive adaptive re-weighted sampling, CARS)等变量选择方法开展特征波段筛选,并在此基础上建立反演模型,普遍取得较好的预测精度。还有学者^[14]将多种特征变量筛选算法的优点相结合,采用混合变量算法进行特征波段筛选。

有研究^[15]表明,土壤的反射光谱和土壤分类系

统一样类型众多而复杂。受生态背景、母质与理化性质差异影响, SOC-光谱关系具有显著区域性特征。在目标区域内构建并优化 SOC 反演模型, 是实现黄土高原区有机碳快速测定的基础。针对黄土高原地区, 已有研究人员采用不同建模策略, 开展光谱快速预测土壤养分含量的研究。Zhu 等^[16]利用小波变换进行光谱预处理, 在太原盆地建立土壤有机质和土壤全氮的 PLSR 预测模型; 周磊等^[7]在陕州区建立农田表层土壤有机碳含量的 CWT-CARS-PLSR 模型, 其决定系数(R^2)达 0.91, 相对分析误差 RPD 为 3.28; 倪玲娜等^[17]以黄土高原 7 个小流域的 126 个土壤样品和 108 个泥沙样品为研究对象, 分别建立土壤和泥沙中碳酸钙含量的 PLSR 预测模型, 模型 R^2 分别为 0.948 和 0.875, 实现光谱对土壤碳酸钙含量的快速无损测量; 王可琳^[18]以皇甫川流域为研究区, 采用高光谱结合多光谱的方法建立土壤有机碳的 RF 模型, 其测试集 RMSE 为 0.20 g/kg。

为探究数学变换与分数阶微分联合预处理后的光谱数据在黄土高原地区 SOC 定量估算中的表征能力与适用性, 本研究使用 Ocean Optics SR6 光谱仪获取延安地区 0~20 cm 表层土壤光谱信息, 并结合室内测定的 SOC 含量, 分别采用倒数、对数、对数倒数及平方根 4 类数学变换与分数阶微分的组合方案对光谱特征进行对比分析, 进一步选取 CARS 与 SPA 2 种变量筛选方法提取各处理条件下的关键波段, 并将其作为输入特征用于 SOC 反演建模, 从而综合评估 PLSR、RF 与 XGBoost 等不同模型框架对 SOC 含量的预测能力与应用潜力。本研究结论可为研究区及其他地区 SOC 的快速无损估算提供方法借鉴与技术支持。

1 材料与方法

1.1 研究区概况

本研究选取陕西省延安市境内吴起县与志丹县(36°21'~37°24'N, 107°38'~109°03'E)为研究区(图 1)。该区位于黄土高原腹地, 为典型梁峁沟壑丘陵区, 具有高原残塬与沟壑地貌特征。受温带半干旱大陆性季风气候影响, 春季干燥多风, 夏季旱涝交替, 秋季温凉湿润, 年平均气温 8.5℃, 多年平均降水量 450~550 mm, 年内时空分配不均, 集中于 7—9 月。土壤类型以黄绵土为主, 土质疏松, 矿物组成以石英、长石等碎屑矿物为主, 且高度同质, 为利用可见光近红外光谱技术进行土壤成分反演提供较好的基质背景。

作为国家“退耕还林(草)”工程的核心示范区, 20 多年来吴起县与志丹县的林地覆盖率显著提升。

受植被类型、群落结构、枯落物与根系养分周转的影响, 土壤有机碳在空间分布上表现出明显的异质性。长期的人工林演替显著改变土壤的物理化学性质及其反射光谱特征, 为建立土壤有机碳光谱预测模型提供了丰富的信息。

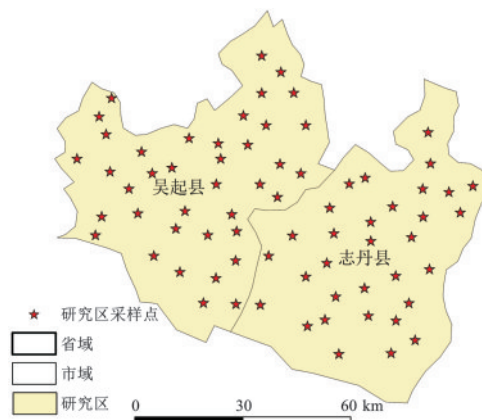


图1 研究区及采样点示意

Fig. 1 Schematic diagram of study area and sampling points

1.2 土壤样品采集及有机碳含量测定

本研究野外取样工作在 2025 年 8—9 月开展, 挑选天气晴好且前期无降雨的时段进行采样。结合地貌条件与通达性在研究区均匀选取 71 块样地。为增加土壤有机质含量差异, 增强光谱表现的辨识性特征, 选择发育不同阶段的以山杏、刺槐及油松为建群种的林地作为样地。每个样地按 10 m 间隔在对角线上取 3 个样点, 最终在研究范围内获取 213 处样点进行土壤样品采集。野外取土时, 采用“梅花式”五点布点方式, 剥离地表枯落物使土壤裸露, 采集 0~20 cm 表层土样。每个点位同步取得 5 份平行土样后充分混匀, 合并为 1 份样品, 随后分装、编号并置于洁净样品袋中, 单份样品保留量约 300 g, 同时记录样点中心坐标信息。样品运回后在室内进行自然阴干, 清除根系与砾石等非土体成分, 研细并通过 100 目筛后保存备用。土壤有机碳(SOC)含量采用重铬酸钾滴定法(外加热消解)完成测定^[19]。

1.3 光谱数据采集与处理

土壤反射光谱由美国 Ocean Optics 公司生产的 SR6 通用光纤光谱仪获取, 其工作波段覆盖 180~1 050 nm, 光谱分辨率为 0.5 nm。为尽可能消除环境杂散光对测量的影响, 全部测试在暗室条件下完成, 暗室照明采用该公司 HL-2000-HP-FHSA 高功率卤钨灯作为稳定的模拟光源。测量时, 将光纤探头垂直置于样品上方约 10 cm 处。正式采集前先进行标准白板标定, 每份样品重复采集 5 条光谱曲线, 并以其算术均值作为该样品的代表反射曲线。随后借助 Ocean

View 软件对所得曲线进行滑动平均平滑处理。鉴于可见-近红外光谱在两端区间出现的高频噪声,本研究剔除 180~399、1 001~1 050 nm 的数据。为降低数据维度并避免光谱混叠,对保留的 400~1 000 nm 波段按 1 nm 步长进行重采样。最终获得 601 个波段作为原始反射光谱数据(R)用于进一步分析。

1.4 光谱变换处理

1.4.1 非线性变换 为进一步突出土壤反射光谱局部细节信息,本研究对 213 个土壤样本的原始光谱反射率(R)进行平方根($\sqrt{R}$)、倒数($1/R$)、对数($\lg R$)和对数倒数($1/\lg R$)4 种非线性变换^[20]。平方根变换是一种温和的非线性变换,可以很好地缓解异方差;倒数变换可以增强反射率小处的细节;对数变换可以压缩动态范围,突出弱特征,而对数倒数变换同时吸收这 2 种优点。

1.4.2 分数阶微分 分数阶微分是关于函数任意阶微分和积分数学理论的拓展运用方法,更是对

$$\frac{d^v f(x)}{dx^v} \approx f(x) + (-v)f(x-1) + \frac{(-v)(-v+1)}{2} f(x-2) + \dots + \frac{\Gamma(-v+1)}{m! \Gamma(-v+m+1)} f(x-m) \quad (2)$$

式中: v 为阶次; m 为光谱数据长度, $m = \frac{t-a}{h}$ 。

本研究将分数阶微分的起始和终止阶次设置为 0 阶和 2 阶,阶次间隔设置为 0.1,即有 21 种分数阶微分处理方式。相关研究中多将阶次间隔设置为 0.2 或 0.5。为对比及捕获光谱细节差异,本文选取 0.1 的阶次间隔^[23-24]。分数阶微分的处理由 Spyder (Python3.13)软件实现。

1.5 特征波段选择

变量相关分析可用来度量 SOC 含量与各类光谱指标之间的线性耦合程度。本文分别针对 5 类光谱反射率数据(含原始光谱),将其分数阶微分后的结果与土壤 SOC 含量逐一进行相关检验,从而获得各波段反射率与 SOC 的相关系数分布,为识别敏感波段和建立 SOC 反演模型提供依据。

高光谱数据处理中特征波段的提取方法,可划分为基于光谱特征的方法和基于数据驱动的方法。基于光谱特征的方法通过分析光谱信号的物理、统计特性直接从光谱数据中筛选出与目标成分相关的特征波段。这种方法利用光谱数据自身结构,不需要大量训练数据与复杂机器学习模型,适用于数据样本有限的场景。连续投影算法(SPA)即是基于光谱特征的方法,可以很好地消除共线性信息变量,揭示光谱成分的响应机制。基于数据驱动的方法多依赖于大规模的数据集和机器学习模型,通过自动化

传统整数阶微分的结合和延展。分数阶微分变换可以有效减少基线变化的影响,加强光谱信号和分离紧密相邻的重叠峰^[21]。根据定义方式,常用的分数阶微分算法的表达形式主要包括 Riemann-Liouville (R-L)、Grunwald-Letnikov (G-L) 和 Caputo^[22]。对于高光谱数据处理的研究多采用 G-L 分数阶微分定义形式,它作为一种离散定义方法,使用加权和表示函数,运算过程相对简单、易行。其具体定义为:

$$f(x) = \lim_{h \rightarrow 0} \frac{1}{h^q} \sum_{m=0}^{\frac{t-a}{h}} (-1)^m \times \frac{\Gamma(q+1)}{m! \Gamma(q-m+1)} \times f(x-mh) \quad (1)$$

式中: q 为微分阶数; h 为微分步长; t 为微分的上限; a 为微分的下限; Γ 为 Gamma 函数。设定光谱重采样间隔为 1 nm,因此令 $h=1$,推导出函数 $f(x)$ 的 v 阶分数阶微分的表达式为:

的训练优化波段选择。这种方法对光谱特征的依赖性不强,而是通过训练数据来捕捉与目标成分相关的特征波段,适用于处理多维、复杂数据集。竞争性自适应重加权采样(CARS)即是基于数据驱动的方法,通过迭代评估波段重要性、逐步剔除不敏感波段,从而获取最优特征组合^[25]。

CARS 用于高维光谱场景下的特征筛选,能够提取与土壤有机碳相关的关键波段组合,同时减少信息冗余并降低建模复杂度。本文将蒙特卡洛抽样设定为 50 次迭代,单次抽样比例为 80%。考虑到随机抽样带来的波动性,对每组光谱数据重复运行 5 次,并以精度最优时得到的变量子集作为最终的最优波段集合^[26]。

SPA 可显著减少参与建模的波段规模,降低无关信息对模型的干扰,从而在保证精度的同时提高计算效率。本文将候选变量的上限设定为 50,迭代过程中每一步从正交子空间中选取投影值更高的向量作为新的起始方向,随后结合回归计算获得对应的 RMSE。以 RMSE 最小为准则,确定最优的初始变量及其匹配的变量数量。

1.6 反演模型构建

为比较不同模型在 SOC 光谱估测中的适用性,选取偏最小二乘回归(PLSR)、随机森林(RF)与极限梯度提升(XGBoost)等模型进行对比。PLSR 和 RF 2 种模型在文献中多见且效果较好。PLSR 适于处理高维共线光谱数据,具有较好的稳定性和可解释性。

RF 通过多树集成刻画非线性关系,鲁棒性较强。XGBoost 在梯度提升框架下引入正则化以抑制过拟合,兼顾拟合能力与泛化性能,值得尝试^[27]。因此,本文选取 3 类模型进行对比评估。

为评估模型的预测能力,将全部样本按照 7:3 随机划分为训练集(149 个)与独立测试集(64 个),为确保每次划分结果一致,设定随机种子为 42。经 5 次交叉检验,训练集与测试集 SOC 平均值差为 0.53 g/kg,约占训练集平均值的 7%,标准差差异为 0.09 g/kg,表明 2 组样本在 SOC 水平与波动幅度上基本一致。模型超参数优化仅在训练集内进行,并采用 5 折交叉验证($K=5$)评估参数组合的平均表

现。超参数优化采用网格搜索(grid search CV)实现,并以交叉验证平均 RMSE 最小作为最优参数选择准则。PLSR、RF 及 XGBoost 的模型超参数及其搜索范围见表 1。

在变量输入方面,分别以 CARS 与 SPA 筛选的特征波段作为模型输入,组合形成 CARS-PLSR、CARS-RF、CARS-XGBoost、SPA-PLSR、SPA-RF 与 SPA-XGBoost 共 6 种建模策略。对光谱数据进行 5 种变换(含原始光谱),考虑 21 个微分阶次条件和 6 种建模策略,共构建 630 个模型。模型训练与评估基于 Python 实现(scikit-learn 与 xgboost 库),并行计算参数设为 $n_jobs=-1$ 。

表 1 选用的 3 类模型超参数搜索范围及设置理由
Table 1 Hyperparameter search ranges and rationale for three types of models used in this study

模型	超参数	搜索范围	设置理由
PLSR	n_components	[2, 3, 5, 7, 10, 15, 20, 25, 30]	潜变量数量,控制降维后模型复杂度与信息提取能力
	n_estimators	[100, 200, 300, 500]	森林中树的数量,影响方差与稳定性
	max_depth	[10, 15, 20, 25, None]	单棵树最大深度,控制模型复杂度以抑制过拟合
RF	min_samples_split	[2, 5, 10]	节点最小分裂样本数,限制过度分裂以提高泛化能力
	min_samples_leaf	[1, 2, 4]	叶节点最小样本数,减少高方差分裂,增强鲁棒性
	max_features	['sqrt', 'log2', None]	每次分裂可用特征数,增加随机性以降低树间相关与过拟合风险
XGBoost	n_estimators	[100, 200, 300, 500]	提升迭代次数(弱学习器数量),决定整体拟合能力
	max_depth	[3, 5, 7, 9]	基学习器复杂度,深度越大越易过拟合,需与学习率共同权衡
	learning_rate	[0.01, 0.05, 0.1, 0.2]	学习率,控制每轮更新幅度
	subsample	[0.6, 0.8, 1.0]	行采样比例,引入随机性以降低过拟合
	colsample_bytree	[0.6, 0.8, 1.0]	列采样比例,降低特征依赖,提升泛化能力
	min_child_weight	[1, 3, 5]	叶节点最小样本权重和,提高分裂保守性,抑制过拟合

1.7 模型精度评价

本研究选取 3 种在光谱分析领域广泛应用的统计指标来定量评价模型的性能,分别是决定系数(coefficient of determination, R^2)、均方根误差(root mean square error, RMSE)和相对分析误差(relative percent deviation, RPD)。

R^2 用于衡量模型预测值与实测值之间的拟合程度,表征模型对土壤有机碳含量变异信息的解释能力,其值越接近 1,表明模型拟合效果越好,解释能力越强。均方根误差(RMSE)反映模型预测值与样本实测值之间的平均离散程度,具有与原始数据相同的量纲,其值越小,表明模型的预测误差越小,精度越高。相对分析误差(RPD)为测试集样本实测值的标准差(standard deviation, SD)与测试集均方根误差(RMSE)的比值。RPD 是一个综合性指标,用于评价模型的实际预测能力和稳定性。参照相关研

究的分类标准,当 $RPD<1.4$ 时,模型预测性能较差,不可靠,当 $1.4\leq RPD<2.0$ 时,模型具有一般的预测能力,可用于粗略估算,当 $RPD\geq 2.0$ 时,模型具有极佳的预测能力和较高的稳定性,可用于定量分析反演。

计算公式^[14]为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

(3)

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - \hat{x}_i)^2}{n}}$$

(4)

$$RPD = \frac{SD}{RMSE}$$

(5)

式中: n 为样本数; x_i 为第 i 个样本的 SOC 含量实测值; $\hat{x}_i$ 为第 i 个样本的 SOC 含量估算值; $\bar{x}$ 为 SOC 含量的平均值。

2 结果与分析

2.1 土壤有机碳含量的统计分析

参照第二次全国土壤普查关于土壤有机质的分级体系,并结合有机质与有机碳的换算关系,可将既有的有机质等级阈值折算为SOC分级标准^[7]。基于该标准,本文对213份样品进行等级划分,共得到5个等级,各等级样本数量统计见表2。结果显示,研

究样品SOC含量为2.50~23.11 g/kg,平均值为7.70 g/kg,对应4级水平,与黄土高原半干旱区人工林土壤常见的肥力状况相一致^[28]。

样地SOC总体变异系数为0.42,变异水平处于中等范围,表明虽然同为林地,但受植被种类、恢复年限、地形特征等的影响,有机碳质量分数的空间异质性显著。

表 2 吴起县和志丹县境内林地土壤样品统计特征
Table 2 Statistical characteristics of forest soil samples in Wuqi and Zhidan Counties

分级及标准/(g·kg ⁻¹)	样本数/个	有机碳含量/(g·kg ⁻¹)				
		最小值	最大值	平均值	标准差	变异系数
二级(<23.20)	1	23.11	23.11	23.11	—	—
三级(<17.40)	22	11.68	17.26	14.34	1.76	0.12
四级(<11.60)	132	5.81	11.48	7.84	1.61	0.21
五级(<5.80)	52	3.50	5.79	4.79	0.67	0.14
六级(<3.48)	6	2.50	3.45	3.04	0.39	0.13
总体	213	2.50	23.11	7.70	3.23	0.42

注:“—”表示不涉及。

2.2 土壤光谱特征分析

经过平滑处理后的213个土壤样本的室内光谱反射率曲线见图2。由图2可知,所有土壤样本的光谱反射率均表现出一致的延展趋势,即随波长的增长总体上增大,而在可见光波段(400~760 nm),反射率曲线斜率较大,呈快速上升趋势。进入近红外波段(760~1 000 nm)后,上升趋势逐渐平缓,在900 nm附近,光谱曲线出现一个微弱的吸收谷,主要是由土壤中铁氧化物(如针铁矿、赤铁矿等)的电子跃迁引起^[29]。

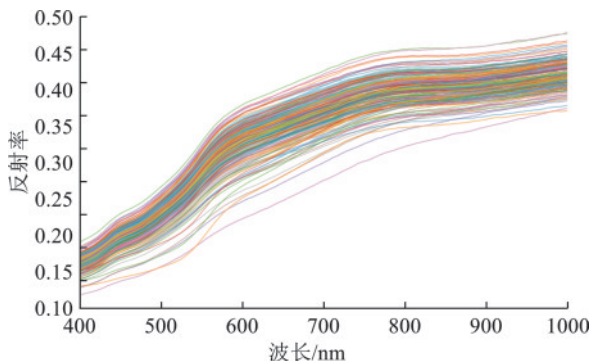


图 2 经过平滑处理的原始光谱反射率曲线(R)
Fig.2 Raw spectral reflectance curves after smoothing (R)

为更深入了解SOC含量与光谱反射率之间的关系,分别从不同SOC含量等级的样本中选取1条光谱进行成图。由图3可知,随着SOC含量的增加,光谱反射率曲线表现出整体下移的趋势,即SOC含量越高,光谱反射率越低。这是由于有机质具有强烈的吸光性,它包裹在土壤矿物颗粒表面,减少光的散

射。同时可以观察到,不同SOC等级间的反射率差异在550~800 nm最为显著,为后续建模和反演SOC含量提供依据。

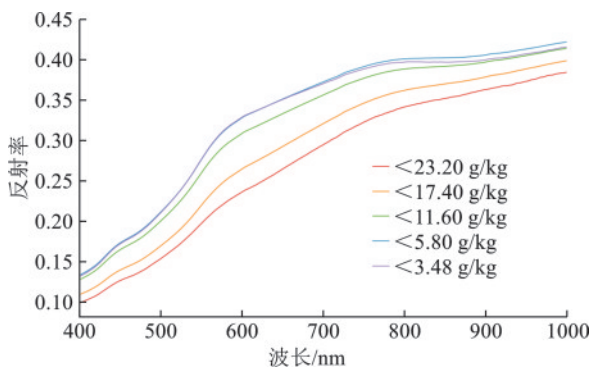


图 3 不同SOC含量原始光谱反射率曲线
Fig.3 Raw spectral reflectance curves for different SOC contents

为增强土壤反射光谱在局部区域的细节表征,在原始光谱及其4类非线性变换光谱的基础上,引入分数阶微分进行处理,其响应曲线见图4。鉴于0~2阶内部分阶次的谱形差异不大,图中选取 α 值分别为0.1、0.5、1.0、1.5、2.0作为代表进行显示。

由图4可知,不论采用何种变换形式,微分阶次提高后各波段反射率幅值整体逐步收敛并接近0,说明分数阶微分对基线漂移具有明显抑制作用。在低阶区间(0~1阶),微分后吸收特征更易被凸显。在450、550、730、900 nm附近可观察到较清晰的峰、谷变化。相较而言,当阶次提升至1~2阶时,

各处理得到的谱线形态趋于一致,典型吸收谷的辨识度下降,表明较高阶微分对有效信息的突出能力弱于低阶微分。

就整数阶微分表现而言,原始反射曲线在450、550、900 nm附近仅呈现较弱吸收迹象,经过1阶微分

后,相应吸收特征明显增强。进一步采用2阶微分时,相比1阶微分结果吸收特征数目有所减少,虽然峰形得到一定锐化,但同时引入更强的高频噪声。因此,微分阶次并非越高越能强化所有吸收信息,需在特征增强与噪声放大之间进行权衡。

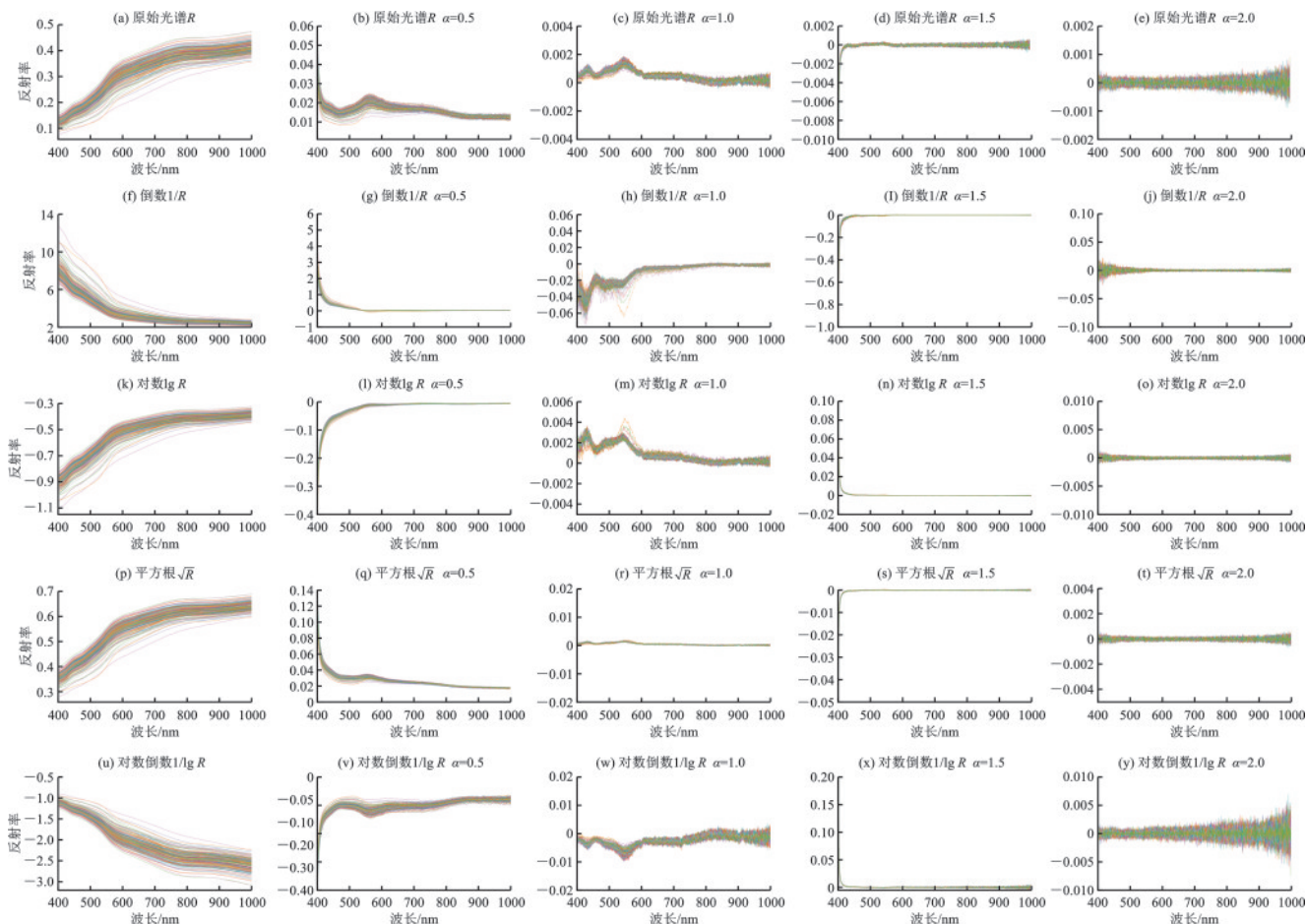


图4 5种数据处理(包括原始谱)不同微分阶次变换下的光谱反射率曲线

Fig.4 Spectral reflectance curves under different differential orders for five data processing methods (including raw spectra)

2.3 相关性分析与特征波段筛选

原始反射光谱及4类变换光谱在微分处理后与土壤有机碳(SOC)含量之间的相关性见图5。整体而言,随着微分阶次逐渐升高,5种光谱序列的相关系数幅度均不断减弱并趋近于0。同时,在接近一阶微分的位置,相关性符号出现由正到负(或由负到正)的交替变化。

从不同变换方式的对比来看,平方根与对数变换得到的相关曲线形态与未变换光谱较为接近,尤其在540~600 nm与690~800 nm波段表现更突出。对数变换的相关性提升最为明显,在546 nm处相关系数达到-0.794。倒数与对数倒数处理的变化规律也较一致,但倒数变换的表现更优,其在866 nm处的相关系数为-0.812。不同变换条件下,微分后出现的最大相关波段及其对应系数汇总见表3。

由表3可知,各方法的相关性随阶次变化呈先增强后减弱的规律,而峰值多集中出现在0.6~0.8阶的分数阶微分范围内。进一步检验表明,表3所列各变换在对应阶次下的最大相关系数均达到极显著水平($p < 0.001$),说明敏感波段与SOC之间的统计关联具有可靠性。

本研究基于CARS和SPA 2种降维算法筛选出的特征波段位置分布见图6。5种数据变换条件下,CARS筛选特征波段有着较为相似的分布规律,即随微分阶次的增加,筛选特征波段的数量均随之增加。在0~1阶微分中,特征波段的数量较少且分布较散,基本分布在550~1 000 nm内;在1~2阶微分中,随着微分阶次增加特征波段逐渐涵盖400~1 000 nm。SPA筛选波段与CARS筛选波段相比更为分散,筛选波段的数量大幅减少,集中在400~450 nm及550 nm周围。

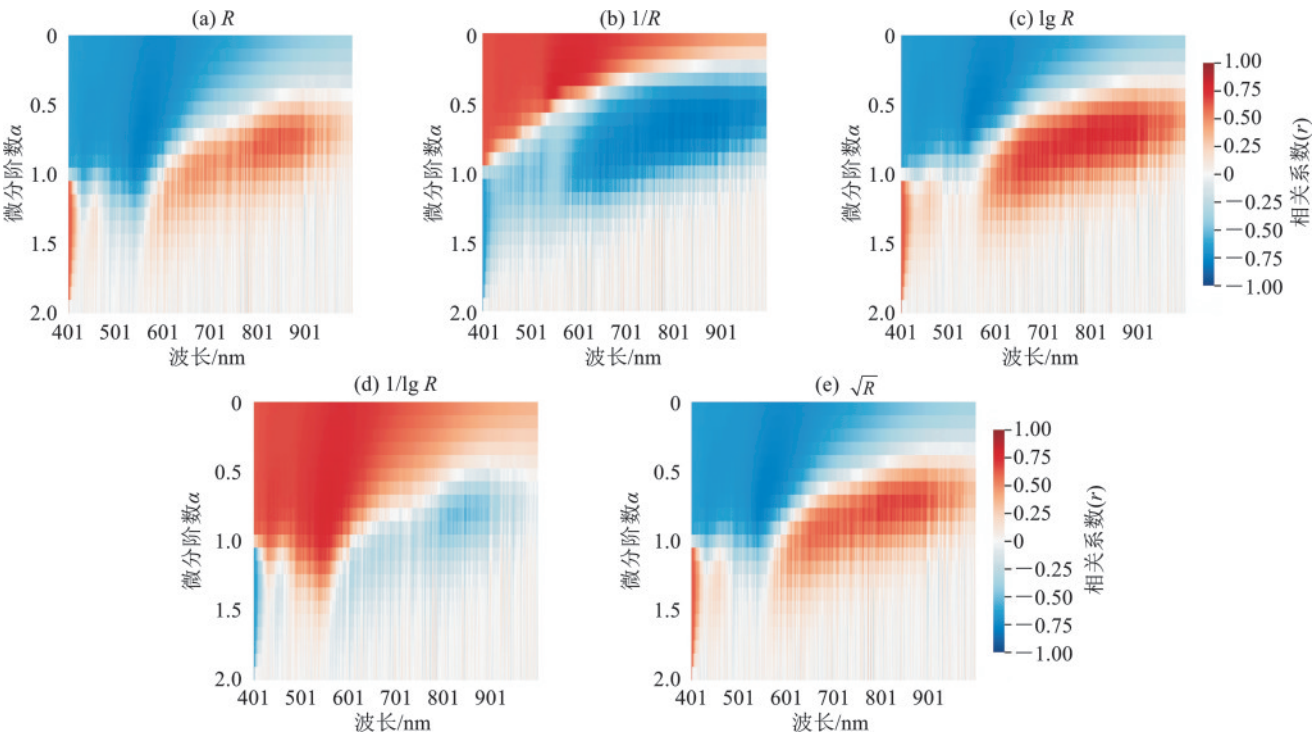


图 5 不同变换下光谱反射率与 SOC 含量的相关系数矩阵

Fig.5 Correlation coefficient matrices of spectral reflectance and SOC content under different transformations

表 3 不同变换下土壤 SOC 含量最大相关波段及相关系数对比

Table 3 Comparison of most strongly correlated bands and correlation coefficients of SOC under different transformations

变换	R		1/R		$\sqrt{R}$		lg R		1/lg R	
微分阶次	最大相关波段/nm	相关系数	最大相关波段/nm	相关系数	最大相关波段/nm	相关系数	最大相关波段/nm	相关系数	最大相关波段/nm	相关系数
0	586	-0.734***	592	0.736***	586	-0.737***	587	-0.739***	586	0.728***
0.1	586	-0.740***	586	0.753***	586	-0.745***	586	-0.748***	579	0.732***
0.2	579	-0.746***	578	0.774***	579	-0.753***	579	-0.760***	579	0.737***
0.3	579	-0.753***	562	0.790***	571	-0.762***	566	-0.772***	566	0.742***
0.4	566	-0.762***	546	0.775***	566	-0.774***	566	-0.786***	566	0.749***
0.5	565	-0.770***	892	-0.786***	559	-0.783***	551	-0.793***	565	0.754***
0.6	559	-0.778***	866	-0.812***	551	-0.790***	546	-0.794***	559	0.760***
0.7	559	-0.783***	866	-0.812***	551	-0.793***	795	0.783***	559	0.764***
0.8	551	-0.785***	795	-0.793***	542	-0.780***	795	0.769***	551	0.766***
0.9	546	-0.775***	697	-0.785***	537	-0.739***	697	0.760***	551	0.763***
1.0	546	-0.739***	697	-0.771***	697	0.665***	697	0.724***	546	0.747***
1.1	540	-0.675***	649	-0.711***	402	0.654***	402	0.659***	546	0.706***
1.2	402	0.652***	402	-0.662***	402	0.657***	402	0.660***	402	-0.653***
1.3	402	0.651***	402	-0.661***	402	0.656***	402	0.659***	402	-0.651***
1.4	402	0.650***	402	-0.660***	402	0.655***	402	0.658***	402	-0.651***
1.5	402	0.650***	402	-0.659***	402	0.655***	402	0.658***	402	-0.651***
1.6	402	0.650***	402	-0.659***	402	0.655***	402	0.658***	402	-0.651***
1.7	402	0.650***	402	-0.658***	402	0.655***	402	0.657***	402	-0.651***
1.8	402	0.646***	401	-0.657***	402	0.651***	402	0.654***	402	-0.647***
1.9	400	0.639***	401	-0.656***	400	0.645***	400	0.649***	400	-0.639***
2.0	400	0.639***	400	-0.653***	400	0.645***	400	0.649***	400	-0.640***

注:相关系数为 Pearson 相关系数,采用双侧显著性检验($n=213$);***表示在该变换下波段与 SOC 在 0.001 水平上显著相关。

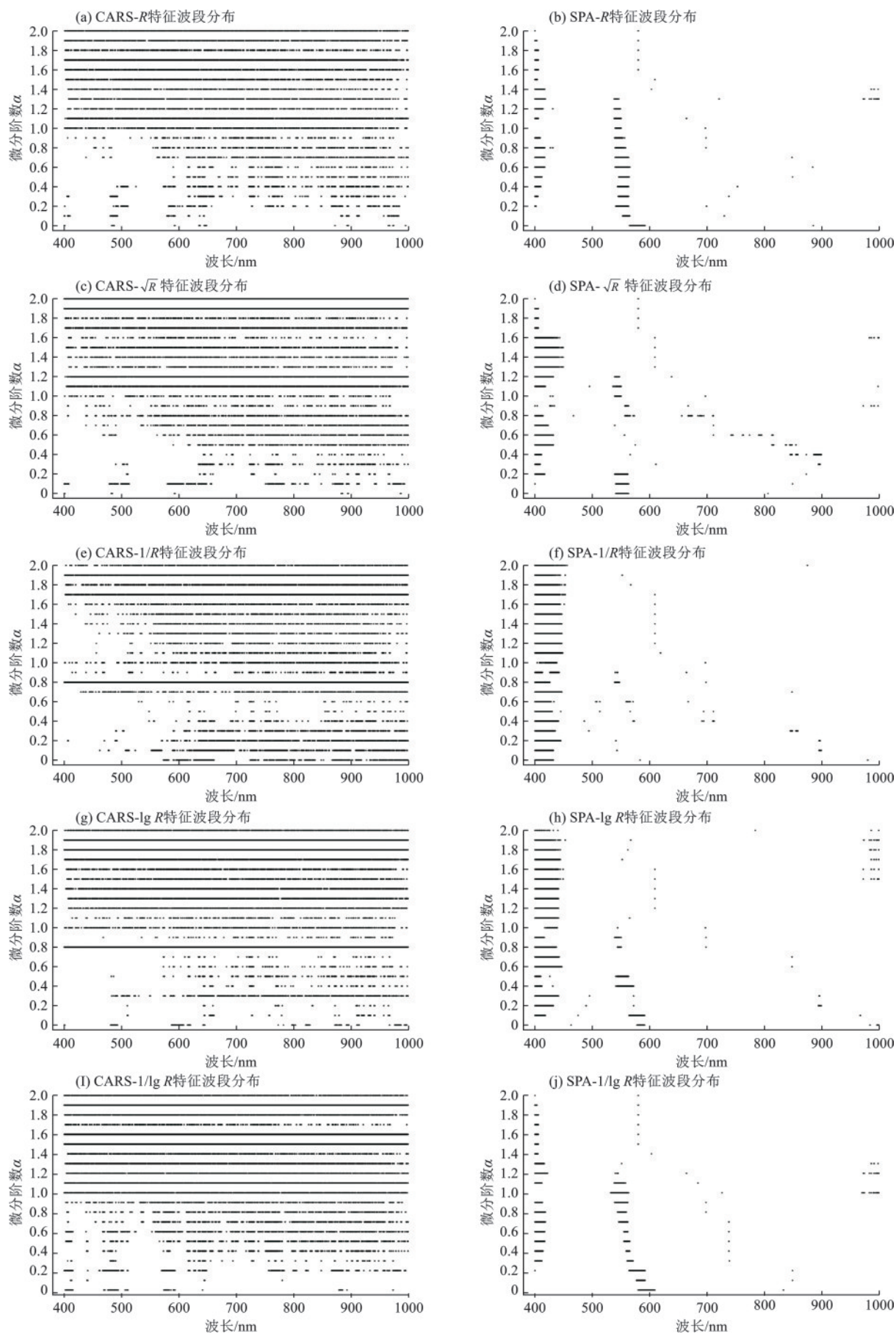


图6 CARS、SPA筛选后的特征波段位置分布

Fig.6 Distribution of feature band locations selected by CARS and SPA

2.4 模型构建与精度分析

本研究依次选取不同光谱变换的分数阶微分处理波段,基于 2 种特征筛选算法筛选出的特征光谱波段分别作为输入变量,并基于相同的训练集和测试集构建反演模型,共有 6 种建模策略。相较于原始光谱, CARS-PLSR、CARS-RF、CARS-XGBoost、SPA-PLSR、SPA-RF 和 SPA-XGBoost 等 6 种建模策略下模型的 R^2 分别提升 0.022(对数,1.1 阶微分)、0.038(对数倒数,0.7 阶微分)、0.007(对数,0.9 阶微分)、0.034(平方根,0.6 阶微分)、0.002(平方根,0.4 阶微分)和 0.033(对数,0.3 阶微分),与之对应的 RPD 分别为 2.525、2.184、2.162、2.254、1.990 和 2.042。除 SPA-RF 策略下的最佳模型 RPD<2 以外,其余 5 种策略的最佳模型 RPD 均>2,说明 5 种建模策略下的最佳模型均具有良好的性能,可用于定量反演。不同建模策略下,0~1 阶微分的模型精度优于 1~2 阶微分,当微分阶次超过 1.5 时,无论何种变换与建模策略,模型的 R^2 均在 0.1~0.6 波动,表明 1.5~2 阶的分数阶微分对于提高本研究区土壤有机碳质量分数的反演精度效果不佳。

PLSR 模型的精度高于 RF 和 XGBoost 模型。PLSR 模型测试集的平均 R^2 为 0.649,平均 RMSE 为 1.924 g/kg,平均 RPD 为 1.803;RF 模型测试集的平均 R^2 为 0.564,平均 RMSE 为 2.169 g/kg,平均 RPD

为 1.572;XGBoost 模型测试集的平均 R^2 为 0.546,平均 RMSE 为 2.208 g/kg,平均 RPD 为 1.552。利用 CARS 筛选波段建模的效果明显优于 SPA,CARS 筛选波段建模的测试集平均 R^2 为 0.612,平均 RMSE 为 2.027 g/kg,平均 RPD 为 1.710;SPA 筛选波段建模的测试集平均 R^2 为 0.560,平均 RMSE 为 2.174 g/kg,平均 RPD 为 1.578。

CARS-PLSR 建模策略在 R 、 $1/R$ 、 $\lg R$ 、 $1/\lg R$ 和 $\sqrt{R}$ 等数据的不同处理形式下,依次在 0.9、0.9、1.1、0.3、0.9 阶微分下达到最大精度, R^2 分别为 0.821、0.830、0.843、0.805、0.828,RPD 分别为 2.423、2.427、2.525、2.264、2.411,证明 CARS-PLSR 模型具有良好的稳定性。

对比不同建模策略下的模型最佳表现见表 4。RF 与 XGBoost 在训练集上拟合程度更高,尤其 XGBoost 出现训练 R^2 接近 1 的情况,而测试集提升幅度有限,提示其在本样本规模下存在更高的过拟合风险。相比之下,PLSR 训练/测试性能差距相对较小,泛化表现更稳定。

综合结果表明,不论采用哪类光谱预处理方式或建模方案,性能最优的模型对应的分数阶微分范围常出现在 0.3~0.9 阶。说明引入分数阶微分能够在一定程度上进一步提取与 SOC 相关的有效谱段信息,从而提升土壤有机碳含量反演模型的预测精度。

表 4 不同建模策略下最佳模型训练集与测试集性能对比
Table 4 Comparison of optimal model performance on training and test sets under different modeling strategies

建模策略	变换	阶次	特征数	训练集			测试集		
				R^2	RMSE/ ($\text{g}\cdot\text{kg}^{-1}$)	RPD	R^2	RMSE/ ($\text{g}\cdot\text{kg}^{-1}$)	RPD
CARS-PLSR	$\lg R$	1.1	194	0.969 3	0.548 4	5.750 7	0.843 1	1.317 3	2.524 7
CARS-RF	$1/\lg R$	0.7	257	0.968 5	0.555 0	5.637 7	0.790 0	1.524 0	2.182 4
CARS-XGBoost	$\lg R$	0.9	97	1.000 0	0.013 8	227.107 2	0.786 1	1.538 1	2.162 4
SPA-PLSR	$\sqrt{R}$	0.6	49	0.796 2	1.412 3	2.215 4	0.803 3	1.475 0	2.254 9
SPA-RF	$\sqrt{R}$	0.4	36	0.882 6	1.072 3	2.917 9	0.747 5	1.671 2	1.990 2
SPA-XGBoost	$\lg R$	0.3	46	0.974 7	0.497 5	6.288 8	0.760 1	1.629 1	2.041 6

3 讨论

3.1 非线性变换与分数阶微分联合预处理及最优阶次选取

高光谱数据具有信息量大、波段连续但冗余与噪声并存等特点,预处理策略在 SOC 反演中往往起决定作用。本研究在延安林地表层土壤样本上系统比较 4 类非线性变换与 0~2 阶分数阶微分(阶次为 0.1)的组合效应,并进一步探究其对波段筛选与模型泛化的影响。

研究表明,低阶 FOD(0.3~0.9)与非线性变换的联合处理能够在抑制基线漂移与散射效应的同时突出与 SOC 相关的谱形差异,进而提升建模精度,而当阶次较高(>1.5)时,模型性能显著下降。而李武耀等^[30]在绿洲区研究中得到“1.2 阶微分+RF 最优”;赵启东等^[23]在渭库绿洲区 SOC 估测中发现较高阶次(1.6 阶)可取得较好效果,而本文最优阶次整体偏低。

不同研究的最优阶次不一致,可能与波段范围差异、土壤背景和阶次步长的设定不同有关。波段

范围可改变SOC信息的主要载体(如包含SWIR时有机官能团吸收带与水分吸收带可能对建模的贡献更为显著),而本文波段限定在400~1 000 nm。土壤背景与干扰因子不同(盐分、含水量、质地)可改变光谱噪声结构及谱形响应机制。阶次步长设置不同会影响最优区间的定位精度。这些原因导致FOD最优阶次具有明显的数据依赖性与区域性,需要在目标区域内通过阶次扫描与交叉验证加以确定。

3.2 敏感波段确定与波段筛选

本文识别的SOC敏感信息主要集中在550~800 nm及近900 nm波段(图5),与史舟等^[15]的研究结果一致。可见光至近红外短波段中,SOC含量升高增强吸收,降低反射率,使谱线整体下移。550~800 nm差异更显著,反映有机质吸光与土壤颜色变化对反射率的综合影响。900 nm附近的特征则与矿物背景(如铁氧化物)相关的吸收特性有关,且在黄土高原黄绵土矿物组成相对同质的条件下更易呈现稳定响应^[29]。

在特征波段筛选方面,CARS总体优于SPA。其原因在于SOC的光谱响应往往体现为宽波段协同变化,CARS通过迭代保留对回归贡献较大的波段集合,能够在高维共线背景下保留成组的有效信息,而SPA以降低共线性为目标,输出波段数量较少且更分散,可能导致SOC相关的宽谱信息被过度压缩,从而限制模型性能。类似结论在土壤参数高光谱估测研究中亦有报道,刘佳琳等^[26]在盐渍土含盐量估算中指出,CARS筛选波段的PLSR模型可获得更高精度,说明在连续谱形信息占主导的场景下,保留协同波段对提升反演表现更为关键。

3.3 模型选择与PLSR最优的原因

在模型层面,PLSR在本研究中整体优于RF与XGBoost,并在非线性变换结合低阶FOD及CARS条件下获得最优预测效果。该现象可从数据结构与模型复杂度2个方面进行认识。Vis-NIR光谱变量存在显著多重共线性,PLSR通过潜变量提取在降维的同时最大化解释SOC相关协方差,能够以较少自由度获得稳定拟合,因此在较小样本规模下具有较强泛化优势。而RF与XGBoost虽具备非线性拟合能力,但对样本量、参数设置与正则化更敏感。从训练/测试集对照可见,XGBoost在训练集上可达到近乎完全拟合而测试集提升有限,表明其存在更高的过拟合风险,需要更严格的正则化与早停策略或更大样本支持。相对而言,PLSR训练/测试集表现更

可控,说明在本研究的SOC响应条件下,其更具适配性。值得注意的是,已有研究^[30]在特定区域与数据条件下也观察到RF优于PLSR的情况,提示模型优劣受研究对象、噪声结构及特征筛选策略影响明显。因此,模型选择应与预处理方式、变量筛选结果及样本规模协同考虑,而非简单依赖某一类算法^[31]。

4 结 论

1)研究区林地土壤SOC含量为2.50~23.11 g/kg,平均为7.70 g/kg。

2)在400~1 000 nm的可见-近红外范围内,不同SOC含量土壤的光谱特征在550~800 nm最为显著,900 nm附近存在稳定吸收特征。

3)分数阶微分能够有效削弱基线漂移并增强与SOC相关的谱形信息。最优模型普遍出现在0~1阶且主要集中于0.3~0.9阶。

4)与原始光谱相比,非线性变换整体提升模型表现,其中对数与倒数变换在相关性增强与建模精度提升方面更为突出,CARS筛选波段建模效果整体优于SPA。

5)模型对比显示PLSR建模总体优于RF与XGBoost,CARS-PLSR在不同变换下均获得最佳预测性能, R^2 高达0.843, RMSE为1.317 g/kg, RPD为2.525。

本研究表明,“ $\lg R$ (或 $1/R$)+0.3~0.9阶FOD+CARS+PLSR”可作为黄土高原林地表层SOC室内快速估测的优选技术,并为区域尺度SOC快速调查、土壤质量评价与碳汇管理提供可复现的方法参考,并为后续在不同土类、水分条件及更宽谱段数据上的推广应用奠定坚实基础。

参考文献:

- [1] Basile-Doelsch I, Balesdent J, Pellerin S. Reviews and syntheses: The mechanisms underlying carbon storage in soil[J]. *Biogeosciences*, 2020, 17(21): 5223-5242.
- [2] McBratney A, Field D J, Koch A. The dimensions of soil security[J]. *Geoderma*, 2014, 213: 203-213.
- [3] Chen Yun, Wang Jinliang, Liu Guangjie, et al. Hyper-spectral estimation model of forest soil organic matter in northwest Yunnan Province, China[J]. *Forests*, 2019, 10(3): e217.
- [4] Li Hongyi, Yang Jiangtao, Hu Bifeng, et al. Improving *in-situ* spectral estimation of wetland soil organic carbon by integrating multiple optimization strategies[J]. *Catena*, 2025, 255: e109078.
- [5] Odebiri O, Odindi J, Mutanga O. Basic and deep learn-

- ing models in remote sensing of soil organic carbon estimation: A brief review [J]. *International Journal of Applied Earth Observation and Geoinformation*, 2021, 102:e102389.
- [6] 褚小立,袁洪福,陆婉珍.近红外分析中光谱预处理及波长选择方法进展与应用[J].*化学进展*,2004,16(4):528-542.
- Chu Xiaoli, Yuan Hongfu, Lu Wanzhen. Progress and application of spectral data pretreatment and wavelength selection methods in NIR analytical technique [J]. *Progress in Chemistry*, 2004, 16(4): 528-542.
- [7] 周磊,何秀芹,贾德伟,等.基于CWT-CARS的黄土高原东部陕州区农田表层有机碳含量高光谱反演[J].*遥感技术与应用*,2025,40(1):60-68.
- Zhou Lei, He Xiuqin, Jia Dewei, et al. Hyperspectral retrieval of organic carbon content in farmland surface in Shanzhou district of eastern Loess Plateau based on CWT-CARS[J].*Remote Sensing Technology and Application*, 2025, 40(1): 60-68.
- [8] 曾远文,范文武.不同光谱变换形式对土壤有机质偏最小二乘估算模型精度的影响[J].*安徽农学通报*,2025,31(15):89-93.
- Zeng Yuanwen, Fan Wenwu. Influence of different spectral transformation forms on the accuracy of partial least squares estimation model of soil organic matter [J]. *Anhui Agricultural Science Bulletin*, 2025, 31(15): 89-93.
- [9] 李天乐,赵泉华,贾淑涵,等.土壤有机质含量反演方法:土壤高光谱数据微分-瞬时频率变换[J].*地球信息科学学报*,2024,26(7):1733-1744.
- Li Tianle, Zhao Quanhua, Jia Shuhan, et al. Soil organic matter content inversion methods: Differential-instantaneous frequency transform of soil hyperspectral data [J]. *Journal of Geo-Information Science*, 2024, 26(7): 1733-1744.
- [10] Hong Yongsheng, Guo Long, Chen Songchao, et al. Exploring the potential of airborne hyperspectral image for estimating topsoil organic carbon: Effects of fractional-order derivative and optimal band combination algorithm [J]. *Geoderma*, 2020, 365:e114228.
- [11] 郭艳萍,王雪梅,赵枫,等.基于最优数学和小波变换的绿洲耕层土壤盐分RF高光谱反演[J].*农业工程学报*,2025,41(3):83-93.
- Guo Yanping, Wang Xuemei, Zhao Feng, et al. Hyperspectral inversion of the RF model for soil salinity in oasis tillage layer based on optimal mathematics and wavelet transform [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2025, 41(3): 83-93.
- [12] 邓昀,王君,陈守学,等.基于PCA-DBO-SVR的林地土壤有机质高光谱反演模型[J].*光谱学与光谱分析*,2025,45(2):569-583.
- Deng Yun, Wang Jun, Chen Shouxue, et al. Hyperspectral inversion model of forest soil organic matter based on PCA-DBO-SVR [J]. *Spectroscopy and Spectral Analysis*, 2025, 45(2): 569-583.
- [13] Morellos A, Pantazi X E, Moshou D, et al. Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy [J]. *Biosystems Engineering*, 2016, 152: 104-116.
- [14] 郑家兴,郭兴燕,郭金锋,等.基于分数阶微分和混合变量选择算法的水稻土有机碳含量反演[J].*激光与光电子学进展*,2026,63(2):389-402.
- Zheng Jiaying, Guo Xingyan, Guo Jinfeng, et al. Inversion of organic carbon content in paddy soil based on fractional order differential and hybrid variable selection algorithm [J]. *Laser and Optoelectronics Progress*, 2026, 63(2): 389-402.
- [15] 史舟,王乾龙,彭杰,等.中国主要土壤高光谱反射特性分类与有机质光谱预测模型[J].*中国科学:地球科学*,2014,44(5):978-988.
- Shi Zhou, Wang Qianlong, Peng Jie, et al. Classification of hyperspectral reflectance characteristics of main soils in China and spectral prediction model of organic matter [J]. *Scientia Sinica (Terrae)*, 2014, 44(5): 978-988.
- [16] Zhu H F, Xu Z J, Jing Y D, et al. Spatial variation and predictions of soil organic matter and total nitrogen based on VNIR reflectance in a basin of Chinese Loess Plateau [J]. *Journal of Soil Science and Plant Nutrition*, 2018, 18(4): 1126-1141.
- [17] 倪玲珊,岳荣,方怒放.基于中红外漫反射光谱法预测黄土碳酸钙含量[J].*水土保持研究*,2020,27(2):142-148.
- Ni Lingshan, Yue Rong, Fang Nufang. Estimation of calcium carbonate in loess by using mid-infrared diffuse reflectance spectroscopy [J]. *Research of Soil and Water Conservation*, 2020, 27(2): 142-148.
- [18] 王可琳.黄土高原典型流域土壤有机碳遥感反演及时空变化特征研究[D].陕西 杨凌:西北农林科技大学,2024.
- Wang Kelin. Remote sensing inversion and spatio-temporal variation of soil organic carbon in a typical basin of the Loess Plateau [D]. Yangling, Shaanxi: Northwest A&F University, 2024.
- [19] 鲍士旦.土壤农化分析[M].3版.北京:中国农业出版社,2000.
- Bao Shidan. Soil and agricultural chemistry analysis [M]. 3rd edition. Beijing: China Agriculture Press, 2000.

- [20] 熊超群, 荀咪, 李健, 等. 基于数学变换和连续小波变换的苹果园土壤总氮含量高光谱反演[J]. 湖北农业科学, 2025, 64(11): 165-170.
- Xiong Chaoqun, Xun Mi, Li Jian, et al. Hyperspectral inversion of soil total nitrogen content in an apple orchard based on mathematical transformation and continuous wavelet transform[J]. Hubei Agricultural Sciences, 2025, 64(11): 165-170.
- [21] 李长春, 施锦锦, 马春艳, 等. 基于小波变换和分数阶微分的冬小麦叶绿素含量估算[J]. 农业机械学报, 2021, 52(8): 172-182.
- Li Changchun, Shi Jinjin, Ma Chunyan, et al. Estimation of chlorophyll content in winter wheat based on wavelet transform and fractional differential[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(8): 172-182.
- [22] 丁启东, 王怡婧, 张俊华, 等. 基于高光谱信息的宁夏土壤有机质含量估算[J]. 生态学杂志, 2024, 43(8): 2531-2538.
- Ding Qidong, Wang Yijing, Zhang Junhua, et al. Estimation of soil organic matter content in Ningxia based on hyperspectral information[J]. Chinese Journal of Ecology, 2024, 43(8): 2531-2538.
- [23] 赵启东, 葛翔宇, 丁建丽, 等. 结合分数阶微分技术与机器学习算法的土壤有机碳含量光谱估测[J]. 激光与光电子学进展, 2020, 57(15): 245-253.
- Zhao Qidong, Ge Xiangyu, Ding Jianli, et al. Combination of fractional order differential and machine learning algorithm for spectral estimation of soil organic carbon content[J]. Laser and Optoelectronics Progress, 2020, 57(15): 245-253.
- [24] 国佳欣, 赵小敏, 郭熙, 等. 基于PLSR-BP复合模型的红壤有机质含量反演研究[J]. 土壤学报, 2020, 57(3): 636-645.
- Guo Jiaxin, Zhao Xiaomin, Guo Xi, et al. Inversion of organic matter content in red soil based on PLSR-BP composite model[J]. Acta Pedologica Sinica, 2020, 57(3): 636-645.
- [25] 梁永刚, 杨金燕. 高光谱土壤成分反演的特征波段提取研究进展[J]. 传感器与微系统, 2025, 44(11): 7-11.
- Liang Yonggang, Yang Jinyan. Research progress on feature band extraction for hyperspectral inversion of soil composition[J]. Transducer and Microsystem Technologies, 2025, 44(11): 7-11.
- [26] 刘佳琳, 王飞, 韩剑桥, 等. 基于分数阶微分的土壤含盐量光谱特征及定量估算研究[J]. 遥感技术与应用, 2025, 40(2): 344-358.
- Liu Jialin, Wang Fei, Han Jianqiao, et al. Study on spectral characteristics and quantitative estimation of soil salinity based on fractional order derivative[J]. Remote Sensing Technology and Application, 2025, 40(2): 344-358.
- [27] Bentéjac C, Csörgő A, Martínez-Muñoz G. A comparative analysis of gradient boosting algorithms[J]. Artificial Intelligence Review, 2021, 54(3): 1937-1967.
- [28] 许小明, 田起隆, 孙景梅, 等. 子午岭地区植被演替的土壤碳汇效应[J]. 水土保持学报, 2022, 36(3): 159-165.
- Xu Xiaoming, Tian Qilong, Sun Jingmei, et al. Effects on soil carbon sink of vegetation succession in the Ziwluling area[J]. Journal of Soil and Water Conservation, 2022, 36(3): 159-165.
- [29] 何挺, 王静, 程烨, 等. 土壤氧化铁光谱特征研究[J]. 地理与地理信息科学, 2006, 22(2): 30-34.
- He Ting, Wang Jing, Cheng Ye, et al. Study on spectral features of soil Fe_2O_3 [J]. Geography and Geo-Information Science, 2006, 22(2): 30-34.
- [30] 李武耀, 买买提·沙吾提, 买合木提·巴拉提. 基于分数阶微分的土壤有机质含量高光谱反演研究[J]. 激光与光电子学进展, 2023, 60(7): 404-411.
- Li Wuyao, Maimaiti·Sawuti, Maihemuti·Balati. Fractional differential-based hyperspectral inversion of soil organic matter content[J]. Laser and Optoelectronics Progress, 2023, 60(7): 404-411.
- [31] Yang Meihua, Xu Dongyun, Chen Songchao, et al. Evaluation of machine learning approaches to predict soil organic matter and pH using vis-NIR spectra[J]. Sensors, 2019, 19(2): e263.