DOI:10.20176/j.cnki.nxdz.20260701

# 基于分数阶微分与多维光谱指数的土壤盐分高光谱反演
## ——以宁夏平罗县为例

贾萍萍[1,2,3]，赵小宁[1,2,3*]，徐士雯[1,2,3]，唐　冲[1,2,3]，戚杰军[1,2,3]，王文婷[1,2,3]

(1.陕西理工大学 秦岭国家公园研究院，陕西 汉中 723001；
2.陕西理工大学 秦巴生物资源与生态环境省部共建国家重点实验室(培育)，陕西 汉中 723001；
3.陕西理工大学 生物科学与工程学院，陕西 汉中 723001)

**摘　要**：土壤盐渍化是全球土壤退化的重要诱因。为提高土壤盐分反演精度，以宁夏平罗县为研究区，基于光谱仪实测数据，系统分析不同程度盐渍化土壤的光谱特征；采用分数阶(0～2.00阶)微分方法增强光谱有效信息，结合二维与三维光谱指数与Boruta特征选择算法，构建随机森林(RF)、极端随机树(ERT)及极端梯度提升(XGBoost)3类土壤盐分反演模型。结果表明，土壤光谱在1 400、1 900、2 200 nm附近出现典型吸收谷；分数阶微分方法能够有效增强光谱吸收特征，其中，0.75阶微分处理下，二维、三维光谱指数与土壤盐分的相关性最强，相关系数($r$)分别为－0.786 8、－0.849；Boruta算法筛选出27个关键光谱特征。模型对比分析结果显示，4种建模策略下，RF模型最稳定且精度最高，当以SG、FOD、Boruta及光谱指数为输入特征时，RF模型的拟合效果最佳($R^2$=0.92，$e_{rms}$=0.84)。特征重要性分析结果显示，特征TBI2是模型预测的最主要驱动因子(重要性得分为0.14)，SHAP解释分析进一步证实TBI2对土壤盐分预测呈正向贡献(平均值为0.56)，而特征RI呈负向贡献。偏依赖图(PDP)与个体条件期望曲线(ICE)揭示，TBI2值在2附近时，模型预测结果呈阶跃式正向响应，而R1590则表现为平滑负向响应，多种解释方法的分析结果一致。该研究为盐渍化土壤盐分的高光谱反演提供了重要参考。

**关键词**：土壤盐渍化；高光谱反演；分数阶微分；机器学习；可解释方法；宁夏平罗

**分类号**：(中图)S155；TP79　　**文献标志码**：A

盐渍化土壤广泛分布于干旱、半干旱区域。土壤盐渍化是全球范围内土地与生态系统退化的重要驱动因素之一，对农业生产力与生态安全构成严重威胁[1-2]。据统计，全球盐渍化土壤(0～30 cm)面积约为4.24亿$hm^2$，且呈现持续扩大趋势[3]。土壤盐渍化给土地管理和农业可持续发展带来重大挑战，导致全球粮食产量减少50%[4-5]。因此，土壤中盐分的快速、准确监测与动态评估，对于盐渍化土壤防治及农业可持续发展具有重要意义。传统土壤盐分测定主要依赖野外采样与实验室分析。尽管该类方法的精度较高，但存在分析成本高、耗时长及空间覆盖有限等不足，难以满足大范围连续监测需求[6]。随着遥感技术发展，土壤光谱学能够为土壤盐分预测提供有效信息[7-8]，且近地高光谱遥感技术具有分辨率高和光谱波段连续的优势，在土壤盐分监测方面展现出巨大潜力[9-10]。通常，土壤光谱信号较弱，且易受到土壤水分、地表覆盖及背景噪声等多种因素干扰，导致在直接使用土壤原始光谱数据建模时难以准确表征土壤盐分信息[11]。为增强光谱特征的表达效果，分数阶微分(fractional order derivative，FOD)作为一种光谱预处理方法，近年来受到广泛关注。潘昊等[12]基于环境减灾二

**收稿日期**：2026-04-22
**基金项目**：国家自然科学基金资助项目(U24A20631)；陕西理工大学校级科研项目(X2025018、X20260044)
**作者简介**：贾萍萍(1996—)，女，博士，讲师，主要从事土壤遥感监测研究，(电子信箱)18392310624@163.com。
*通信联系人：赵小宁(1978—)，女，教授，博士，主要从事陆地地表水碳过程与模拟、环境遥感等研究，(电子信箱)zhaoxiaoning@snut.edu.cn。
**引用格式**：贾萍萍，赵小宁，徐士雯，等.基于分数阶微分与多维光谱指数的土壤盐分高光谱反演：以宁夏平罗县为例[J].宁夏大学学报(自然科学版中英文)，2026，47(4)：289-301.

号卫星高光谱数据,开展0～2.0阶FOD变换分析,发现在0.9阶FOD处理下,不同程度盐渍化土壤光谱的差异最显著,相关系数($r$)达0.58,且光谱反演精度最高($R^2$=0.78)。周凤玺等[13]系统比较Grünwald-Letnikov、Riemann-Liouville和Caputo这3种FOD定义形式在盐渍化土壤电导率高光谱反演中的适用性。结果显示,经0.6、0.7阶Grünwald-Letnikov处理后,光谱的$r$较原始光谱分别提升9.5%、6.7%。与传统整数阶微分相比,FOD能够在保留低频信息的同时,强化光谱细节特征,在土壤属性反演中展现出良好应用潜力[14]。然而,现有研究多集中于单一微分阶次与特定定义形式分析,缺乏不同阶次FOD处理对土壤盐分敏感特征响应机制的系统性分析。因此,不同阶次FOD处理对土壤盐分的最优应用模式仍不明确。此外,单一数据难以全面刻画土壤盐分的复杂空间变异特征。近年来的研究结果显示,多源数据(如光谱特征、光谱指数及环境协变量)融合与机器学习方法结合,能够显著提升土壤属性的反演精度[15-16]。然而,如何有效融合FOD特征,构建稳定且具有良好泛化能力的土壤盐分预测模型,仍是当前研究面临的重要问题。

平罗县是宁夏重要粮食生产基地,农业以引黄灌溉为主,土壤盐渍化已成为制约该地农田生产力和生态可持续发展的主要因素[14]。文中基于地面高光谱数据,采用FOD处理方法,系统分析不同阶次FOD处理对土壤光谱特征与盐分相关性的提升作用;融合FOD特征、优选光谱指数及环境变量构建多策略预测模型,并比较各模型的土壤盐分反演效果,确定最优反演方案,以期为区域尺度土壤盐渍化精准监测提供理论依据与方法支持。

## 1 研究区域与研究方法

### 1.1 研究区概况

研究区位于宁夏北部平罗县(105°57′42″～106°58′02″E、38°36′18″～39°51′13″N),隶属石嘴山市(图1)。黄河纵贯其境,西依贺兰山,东与内蒙古鄂托克前旗毗邻,全县总面积2 634 km$^2$,其中,耕地面积820 km$^2$[17]。平罗县地势西高东低,属典型的大陆性干旱气候,蒸发量与降水量的比值为10∶1。受自然条件和引黄灌溉影响,平罗县土壤盐渍化问题突出,是银川平原盐碱土的主要分布区之一,其中,轻度、中度和重度盐渍化土壤面积分别占该县总面积的25.2%、39.8%和2.7%[18];该区域主要土壤类型以灌淤土、潮土、盐土和碱土为主[19],土壤质地变化较大,普遍呈碱性;该区域主要种植作物为小麦、水稻、玉米等粮食作物,近年来逐步发展了蔬菜、制种等特色农业。

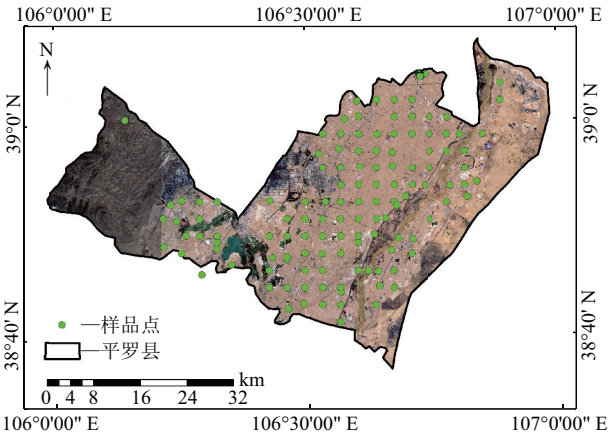



**图1 研究区位置及采样点**

**Fig. 1 Location of the study area and sampling points**

### 1.2 土壤样本的采集与处理

2024年3月26日—4月11日,在平罗县耕地采集土壤样品(裸土)。使用手持全球定位系统(global positioning system, GPS)记录采样点的经纬度、土地类型和地表覆盖等信息。每个采样点取0～20 cm表层土样,共采集到126个土壤样品。一般,土壤样品数量要兼顾插值精度与调查成本。已有研究结果显示,在县域尺度上进行土壤盐分空间变异分析时,满足克里格插值精度要求的样品数量为107个(在1 000 km$^2$内)[20]。文中采样密度与上述要求接近,显示该研究中采集到的样品数量合理。将一部分土壤样品装入密封袋并标记,送至实验室,采用烘干法测定土壤含水率。剩余样品经自然风干后过筛。取部分风干土壤,经研磨后过2 mm筛。按照土壤与水的质量比为1∶5,配制土壤浸提液,采用电导法测定土壤电导率($\sigma$),并将$\sigma$换算为土壤含盐量($w_s$)。文中土壤样本盐渍化等级见表1(表1中,$w_{s,av}$为土壤平均含盐量)。

**表1 土壤样本盐渍化分级**

**Tab. 1 Salinization classification of soil samples**

| $w_s$ | 等级 | 样本数/个 | $w_{s,av}$ | 变异系数/% |
|---|---|---|---|---|
| 0～2 | 非盐渍化 | 91 | 0.96 | 47.92 |
| >2～4 | 轻度盐渍化 | 30 | 2.80 | 18.93 |
| >4～8 | 中度盐渍化 | 4 | 5.67 | 12.17 |
| >8～16 | 重度盐渍化 | 1 | 10.22 | 0 |

### 1.3 土壤高光谱数据的获取与处理

采用FieldSpec4型野外便携式地物光谱仪(美国Analytical Spectral Devices公司)测定土壤原位

光谱，探测波段为 350～2 500 nm，重采样间隔为 1 nm。其中，350～1 000、1 000～1 500、1 500～2 100 nm 波段光谱的分辨率分别为 3.5、10、7 nm。测定时，天气晴朗无云无风，光线充足，测定时间为 10:00—14:00；光谱仪探头垂直朝下且距地面约 30 cm。每个样点重复测定 5 次，每次测定前采用标准白板校正，最后取 5 次的平均值作为该样点的光谱反射率。异常曲线去除、断点校正及光谱反射率平均值计算均在 ViewSpecPro 软件中完成。

高光谱属于纳米级连续地物光谱，其波段划分精细。在计算光谱指数时，需从众多波段中筛选出敏感波段，并将其作为代表性波段计算光谱指数。为了消除仪器噪声和环境背景干扰，需要去除噪声过大的边缘波段（350～399 nm、2 401～2 500 nm）。采用 Savitzky-Golay（SG）算法对 400～2 400 nm 波段光谱进行平滑处理，然后以 10 nm 为间隔进行重采样处理，最终得到由 201 个波段组成的原始反射率（original reflectance，OR）曲线。

### 1.4　环境协变量（environmental variables, ENVI）

文中地形数据从 http://search.asf.alaska.edu/#/网站下载，数据集为 ALOS 卫星（Advanced Land Observing Satellite）搭载的 PALSAR 雷达（L 波段相控阵合成孔径雷达）拍摄的遥感数据，其空间分辨率为 12.5 m。在 ArcGIS 10.4 中，通过 Spatial Analyst Tools 中 Extract Multi Values to Points 工具，从数字高程模型（digital elevation model，DEM）中提取每个采样点的高程、坡向（aspect）、坡度（slope）、平面曲率（plane curvature）、剖面曲率（profile curvature）及地形起伏指数（terrain ruggedness index，TRI）。

### 1.5　分数阶微分（fractional order derivative, FOD）

相较于整数阶微分处理方法，FOD 处理能够有效捕获数据的微小变化和整体信息[14]。目前，Grünwald-Letnikov（GL）算法的 FOD 是常用方法之一，是从 $n$ 阶导数扩展而来。对于任意实数 $n$，若 $f(x)$ 在区间 $[s,t]$ 存在 $n$ 阶连续导数，则其在 $t$ 点的 $n$ 阶导数可表示为

$$ {}_{s}\mathrm{D}_{t}^{n} f(x)=\lim_{h\to 0}\frac{1}{h^{n}}\sum_{m=0}^{\left[\frac{t-s}{h}\right]}(-1)^{m}\times \frac{\Gamma(n+1)}{m!\ \Gamma(n-m+1)} f(x-mh)。 \tag{1} $$

式中：$f(x)$ 为随着波长 $x$ 变化的高光谱反射率函数；$n$ 为微分阶数（任意实数）；$h$ 为步长（即波长间隔，文中 $h=0.25$）；$s$ 为微分下限；$t$ 为微分上限；$m$ 为累加索引数。当 $n$ 为正实数时，式（1）为 $f(x)$ 在 $t$ 点的 $n$ 阶导数；当 $n$ 为负实数时，式（1）为 $f(x)$ 在 $t$ 点的 $n$ 阶积分。

当 $h=0.25$，$s$、$t$ 分别为起止波长时，$f(x)$ 的 $\alpha$ 阶导数（$\alpha$ 为微分阶数）可简化为

$$ \frac{\mathrm{d}^{\alpha} f(x)}{\mathrm{d}x^{\alpha}}\approx f(x)+(-\alpha) f(x-1)+\frac{(-\alpha)(-\alpha+1)}{2} f(x-2)+\cdots+\frac{\Gamma(-\alpha+1)}{n!\ \Gamma(-\alpha+n+1)} f(x-n)。 \tag{2} $$

由式（2）可知，当 $\alpha=0$、1、2 时，分别对应原始光谱反射率、一阶微分与二阶微分；而当 $\alpha$ 为其他任意实数（如 0.2、0.5、0.8 等）时，则对应高光谱 FOD。

### 1.6　光谱指数（spectral index）

光谱指数是由两个及以上波段的光谱反射率经数学运算变换得到，可增强特定目标与光谱之间的相关性[21]，在遥感研究中，常被用于土壤盐分的间接反演估算[22]。平罗县耕地土壤盐分空间变异显著，西大滩白僵土中硫酸盐含量较高，土壤盐渍化呈轻度至重度梯度分布，且只灌不排地区植被退化明显。针对这些特点，光谱差值指数能够反映盐斑区域，光谱比值指数能够区分土壤中盐分种类，光谱归一化指数差值适用于盐分等级划分，光谱三波段指数可分别用于去植被化、识别伴生矿物和协同监测盐分与植被胁迫分析。文中采用的光谱指数及其计算公式见表 2。

**表 2　光谱指数及其计算公式**

**Tab. 2　Spectral indices and their calculation formulas**

| 光谱指数 | 符号 | 计算公式 | 文献 |
|---|---|---|---|
| 差值指数(difference index) | DI | $R_{\lambda1}-R_{\lambda2}$ | [22] |
| 比值指数(ratio index) | RI | $R_{\lambda1}/R_{\lambda2}$ | [22] |
| 归一化指数(normalized difference index) | NDI | $(R_{\lambda1}-R_{\lambda2})/(R_{\lambda1}+R_{\lambda2})$ | [22] |
| 一波段指数(single-band index1) | TBI1 | $R_{\lambda1}/R_{\lambda2}R_{\lambda3}$ | [23] |
| 二波段指数(two-band index2) | TBI2 | $R_{\lambda1}/(R_{\lambda2}+R_{\lambda3})$ | [23] |
| 三波段指数(three-band index3) | TBI3 | $(R_{\lambda1}-R_{\lambda3})/(R_{\lambda1}+R_{\lambda3})$ | [23] |

注：$R_{\lambda1}$、$R_{\lambda2}$、$R_{\lambda3}$ 分别为 400～2 400 nm 波段随机可用光谱反射率；所选波段为光谱指数与土壤盐分相关性较强波段。

### 1.7 光谱特征筛选

文中采用Boruta算法进行光谱特征筛选。Boruta是一种基于随机森林(random forest, RF)的特征筛选算法。该算法的核心原理是构建阴影特征(原始特征的随机打乱副本)作为特征参照,然后比较真实特征与阴影特征在RF模型中的重要性得分,最终筛选出重要性分值显著高于阴影特征且与目标变量具有相关性的特征[24]。该算法的参数设置:显著性水平($P$值)为0.01;采用多重检验校正(mcAdj=TRUE)控制假阳性率;追踪等级(doTrace)为2,用于输出详细迭代过程信息;最大运行次数(maxRuns)为400,即算法在400轮迭代内若仍未达到收敛则自动终止。同时,执行过程中设置随机种子以确保结果可重复。

### 1.8 模型的构建与验证

选择随机森林(random forest, RF)、极端随机树(extremely randomized trees, ERT)和极端梯度提升(extreme gradient boosting, XGBoost)3种算法进行建模。RF通过自助重抽样从原始数据中生成多个样本,并为每个样本构建决策树;在节点分裂时仅从随机选取的若干特征中选择分割特征。根据回归任务,最终整合所有决策树的预测结果并求平均值。该方法不易过拟合,对数据集适应性较强[25]。同样,ERT以决策树为基本单元,但直接使用整个数据集训练每棵树,并在节点划分时随机为每个特征指定分割阈值,再从中选出最优属性。该方法提升了模型随机性,增大了决策树间差异[26]。XGBoost则结合回归树与提升算法,在目标函数中加入树模型复杂度作为正则项,并利用二阶泰勒展开优化损失函数。同时,XGBoost借鉴RF采样策略,支持样本抽样和多核并行计算,可有效处理稀疏和缺失数据。该方法降低了内存需求并提升了大规模数据训练速度[27]。

采用决定系数($R^2$)、均方误差($e_{rms}$)、均方根误差($e_{me}$)评估模型模拟效果。采用5折交叉验证方法对模型进行验证。通过Python建模及绘图,使用网格搜索方法(Grid Search)[28]进行超参数调优。

为探究光谱特征对土壤盐渍化反演模型的影响,文中设计4种递进式建模策略(表3)。所有策略均采用SG滤波与FOD预处理,其中,SG平滑窗口设为9,多项式阶数取3;FOD阶数为0~2.00、阶数步长为0.25;光谱波段为400~2 400 nm,边界数据采用单边截断处理方式。采用Boruta算法筛选光谱特征,并剔除冗余信息。策略Ⅱ额外引入光谱指数,以增强土壤光谱特征;策略Ⅲ引入地形因子,这是因为平罗县灌区地形起伏会影响地表积水、地下水位埋深及土壤表层盐分聚集程度;策略Ⅳ则融合全部特征。对比分析4种策略下模型精度,可定量评估光谱指数与环境变量各自的贡献及其交互作用,从而为平罗县土壤盐渍化遥感监测筛选出最优输入特征组合。

**表3 建模策略与输入特征**

**Tab. 3 Modeling strategies and input variables**

| 建模策略 | 输入特征 |
|---|---|
| 策略Ⅰ | SG、FOD、Boruta |
| 策略Ⅱ | SG、FOD、Boruta、spectral index |
| 策略Ⅲ | SG、FOD、Boruta、ENVI |
| 策略Ⅳ | SG、FOD、Boruta、spectral index、ENVI |

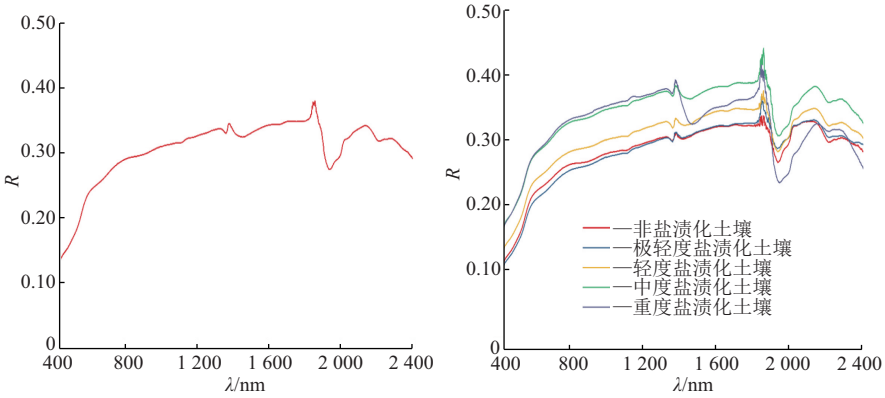
### 1.9 模型解释方法

夏普利加性解释(SHapley Additive exPlanations, SHAP)基于博弈论中的Shapley值,将模型预测结果分解为不同特征的贡献之和,可为每个样本提供符合一致性与可加性要求的特征归因解释,具有良好理论基础和直观性[29]。部分依赖图(partial dependence plot, PDP)通过固定其他特征值不变,改变目标特征值并计算模型的预测平均值,能够直观展示该特征对预测结果的全局边际效应,包括线性、非线性或交互作用[30]。个体条件期望图(individual conditional expectation plot, ICE)则是每个样本预测值随着特征值变化的曲线,能够揭示PDP中平均曲线下方区域隐藏的个体异质性,尤其在特征之间存在交互效应时,具有较高分析价值[31]。

## 2 结果与分析

### 2.1 土壤光谱特征

对研究区土壤样品的高光谱曲线进行分析(图2),发现不同程度盐渍化土壤的光谱曲线在形态上一致,总体变化平缓但有一定波动性:在1 400、1 900、2 200 nm附近各有一处吸收谷;光谱反射率($R$)在400~800 nm内快速增加,在800~1 400 nm内缓慢增加,在400~1 400 nm内随着土壤盐碱度增大呈现规律性变化,即随着土壤盐碱度增大而增大,但在1 400 nm后,随着土壤盐碱度增大而发生相应变化。这可能是因为研究区土壤多为龟裂碱土,盐碱度较大,且含盐量较低土壤的pH值较大。




（a）全部土壤样本　　（b）不同程度盐渍化土壤

**图 2　不同程度盐渍化土壤光谱**

**Fig. 2　Spectral curves of salinized soils**

采用FOD方法处理光谱数据，当FOD的阶次从0增加到2.00时，1 400、1 900 nm处吸收谷越来越明显（图3），1 900 nm处吸收谷逐渐变为负吸收峰，整体上可见光区域光谱的吸收特征不断增强。图中蓝色区域的变化显示土壤光谱之间差异逐渐减小。随着FOD阶次增加，光谱强度逐渐降低，但光谱细节特征变得明显。因此，FOD方法处理可有效识别光谱中正负峰，但高阶FOD处理的光谱极易受到噪声干扰。

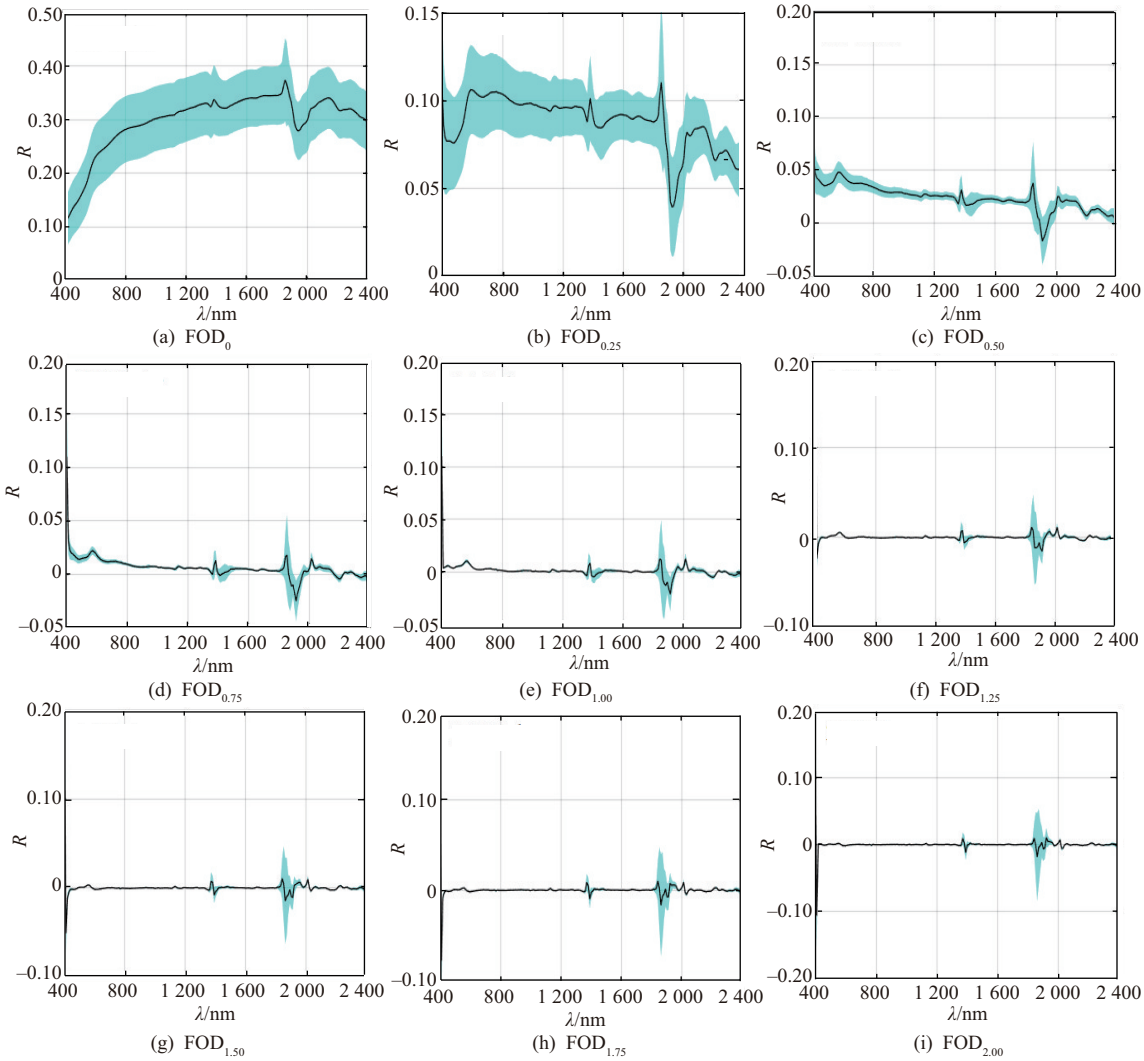



**图 3　分数阶微分下光谱吸收特征曲线**

**Fig. 3　Variation characteristics of fractional-order differential spectral curves**

### 2.2 土壤盐分与光谱特征的相关性

不同阶次FOD处理对土壤盐分与光谱特征相关性的影响显著(图4)。在0～1.00阶次FOD处理下,400～1 200 nm波段$R$与土壤盐分均为正相关,1 200 nm后正负相关性均存在。当FOD阶次在1.50～1.75时,1 400～2 400 nm波段$R$与土壤盐分的正相关性增强。此外,随着FOD阶次增加,土壤盐分与光谱特征的相关系数($r$)发生显著变化,说明FOD处理能够突显特定波段的光谱特征。

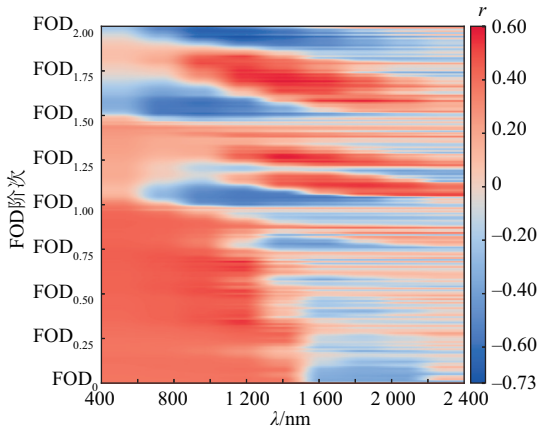



**图4 分数阶微分下光谱反射率与土壤盐分的相关性**

**Fig. 4 Correlation between reflectance and soil salinity under fractional order differential spectra**

二维和三维光谱指数与土壤盐分的相关系数($r$)均在$FOD_{0.75}$处理下最大(表4～表5)。基于实测高光谱数据,分析$FOD_{0.75}$处理下土壤盐分与二维光谱指数(DI、RI、NDI)和三维光谱指数(TBI1、TBI2、TBI3)的相关性(图5～图6)。二维光谱指数

**表4 二维光谱指数与土壤盐分的相关性**

**Tab. 4 Correlation between two-dimensional spectral indices and soil salinity**

| 分数阶FOD | $r$ | | |
|---|---|---|---|
| | DI | RI | NDI |
| $FOD_0$ | 0.779 6 | 0.779 1 | 0.774 4 |
| $FOD_{0.25}$ | 0.775 1 | −0.736 0 | 0.731 5 |
| $FOD_{0.5}$ | 0.781 5 | −0.754 3 | 0.757 4 |
| $FOD_{0.75}$ | 0.760 2 | −0.786 8 | 0.784 0 |
| $FOD_{1.00}$ | 0.716 5 | 0.716 4 | 0.735 8 |
| $FOD_{1.25}$ | 0.655 6 | −0.715 0 | 0.735 4 |
| $FOD_{1.50}$ | 0.662 3 | −0.735 4 | 0.747 8 |
| $FOD_{1.75}$ | 0.618 5 | 0.747 8 | 0.756 1 |
| $FOD_{2.00}$ | 0.598 7 | −0.696 9 | 0.681 6 |

**表5 三维光谱指数与土壤盐分的相关性**

**Tab. 5 Correlation between three-dimensional spectral indices and soil salinity**

| 分数阶FOD | $r$ | | |
|---|---|---|---|
| | TBI1 | TBI2 | TBI3 |
| $FOD_0$ | −0.638 | 0.785 | −0.794 |
| $FOD_{0.25}$ | 0.756 | 0.809 | −0.805 |
| $FOD_{0.50}$ | −0.799 | 0.817 | −0.810 |
| $FOD_{0.75}$ | −0.803 | −0.827 | 0.849 |
| $FOD_{1.00}$ | 0.749 | 0.804 | 0.823 |
| $FOD_{1.25}$ | 0.770 | 0.786 | −0.814 |
| $FOD_{1.50}$ | 0.766 | 0.775 | 0.798 |
| $FOD_{1.75}$ | 0.755 | 0.772 | 0.798 |
| $FOD_{2.00}$ | 0.766 | 0.753 | −0.771 |

中,RI与土壤盐分的相关性最强($r$=−0.786 8),其次为NDI($r$=0.784 0);三维光谱指数中,TBI3与土壤盐分的相关性最强($r$=0.849)。

$FOD_{0.75}$处理下光谱$R$与土壤盐分的Boruta分析结果显示(图7),筛选得到的特征为R880、R900、R950、R1050、R1070、R1590、R1610、R1780、R2110、R2130、R2160、R2170、R2180、R720、R730、R740、R750、R820、R890、R910、R920、R1080、R1430、R1470、R1600、R2150和R2330。取上述特征进行建模。

### 2.3 模型的建立

表6为建模策略(策略Ⅰ～Ⅳ)下模型(RF、ERT、XGBoost)对土壤盐分预测的拟合性能。由表6可知,策略Ⅰ下,RF模型的拟合效果最好($R^2$=0.91),而ERT、XGBoost模型的拟合效果较差;策略Ⅱ提升了所有模型的拟合性能,其中,RF模型的拟合效果最好($R^2$=0.92);策略Ⅲ下,RF模型的拟合效果依然较好($R^2$=0.91),但ERT、XGBoost模型的拟合性能显著下降;策略Ⅳ下,RF模型的拟合效果也较好($R^2$=0.90)。因此,各策略下RF模型的拟合效果均较好,尤其在策略Ⅰ、Ⅱ、Ⅲ下表现最佳,显示RF模型在土壤盐分预测中具有较强的泛化能力和适用性。

策略Ⅰ下RF模型虽然在拟合效果上表现较好,但其泛化性能稍差,而ERT模型的交叉验证结果较稳定。策略Ⅱ显著提高了模型的泛化性能和拟合性能,尤其是RF、ERT模型,表明策略Ⅱ能够有效优化数据特征与模型构建流程,适用于土壤盐分预测。策略Ⅳ提升了模型的拟合能力,尤其是提

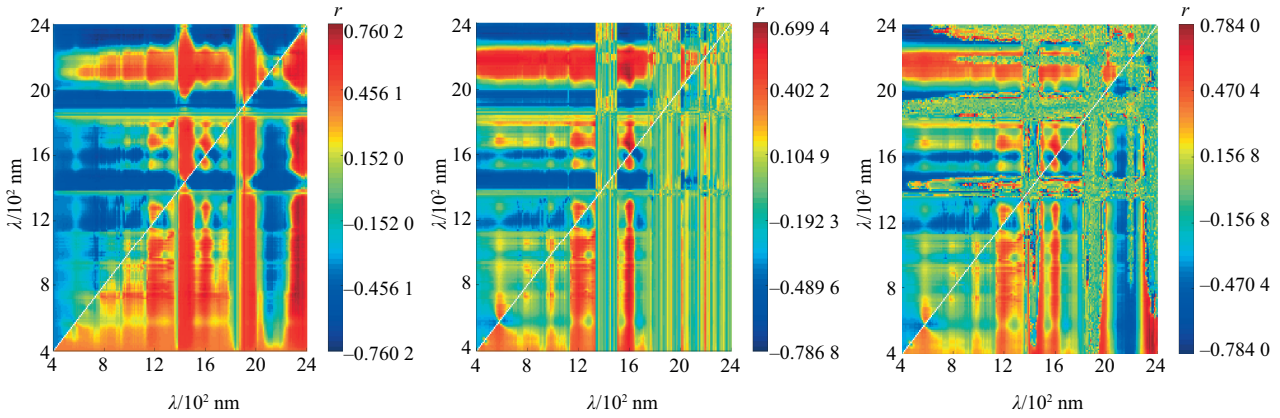



**图 5　$FOD_{0.75}$处理下二维光谱指数与土壤盐分的相关性**

**Fig. 5　Correlation between two-dimensional spectral indices and soil salinity under $FOD_{0.75}$**

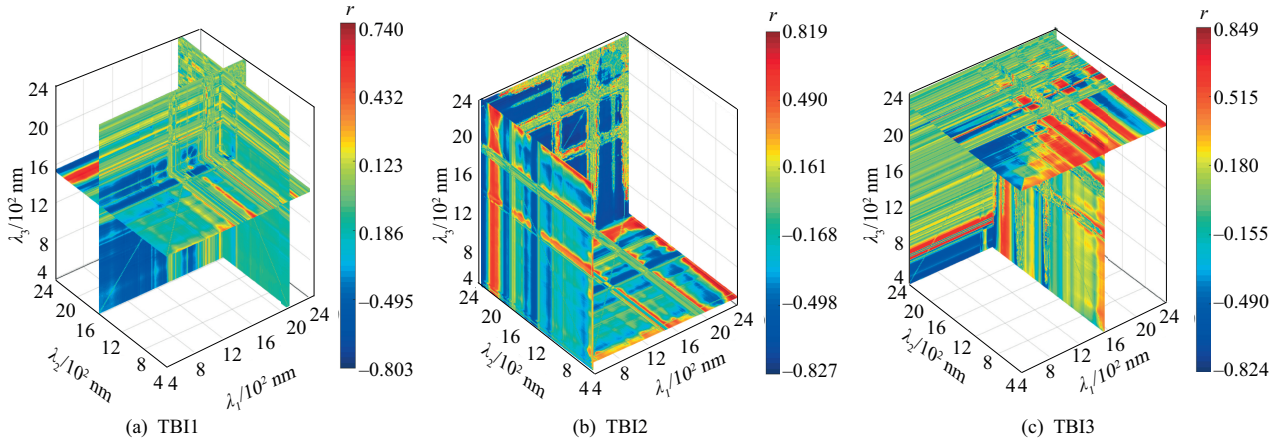



**图 6　$FOD_{0.75}$处理下三维光谱指数与土壤盐分的相关性**

**Fig. 6　Correlation between three-dimensional spectral indices and soil salinity under $FOD_{0.75}$**

升了 RF 模型的泛化能力和预测能力。在模型的拟合效果方面，RF 模型具有较强的泛化能力和鲁棒性，适合土壤盐分预测。另外，ERT 模型在泛化能力上稍优于 RF 模型，但需优化训练过程，以进一步提升模型的拟合能力。

策略Ⅱ下 RF 模型的特征重要性分析结果显示（图 8），各个特征对土壤盐分预测的贡献差异明显。其中，特征 TBI2 的重要性分值最高，接近 0.14，远高于其他特征，显示 TBI2 在模型预测中起主导作用；R1590、R1600、DI、TBI3 等特征的重要性分值依次减小，为 0.02～0.10，显示这些特征在模型预测中也有一定贡献；其他特征，如 R890、R2330、R740 等的重要性分值较低，均未超过 0.02，显示这些特征对模型预测结果影响小。总体来看，模型主要依赖少数关键光谱特征进行反演，且特征重要性呈现长尾分布特征。

采用 SHAP 分析对策略Ⅱ下 RF 模型进一步解释分析（图 9）。分析结果显示，TBI2 对土壤盐分预测的贡献度最高（0.56），RI 的贡献度次之。由数据点颜色可知：对于 TBI2，较高的 SHAP 值会提升模型预测值；而对于 RI，较高的 SHAP 值则对模型预测产生副作用。

为筛选出对模型预测有显著影响的关键特征，基于 PDP，计算各特征对预测结果的响应度（即 PDP 曲线中最大值与最小值之差），并据此对所有特征排序。选取响应度最大的前 4 个特征用于后续融合解释分析，以表征模型的主要驱动因素。采用 PDP、ICE 及 SHAP 方法进行模型融合解释分析，结果见图 10。3 种方法的分析结果显示，4 个特征的变化趋势基本一致，表明模型解释结果具有较好的稳定性和一致性。对于 TBI2，其 PDP 曲线呈明显阶跃式上升趋势，并在 TBI2 值为 2 附近出现显著增长；ICE 曲线整体变化规律与 PDP 一致，但不同样本间存在一定差异；SHAP 散点主要集中于变化区间，TBI2 贡献度随着取值增大显著提高。对于 R1590，PDP 曲线整体呈下降趋势，表明随着其取值增大，R1590 对模型预测结果的影响逐渐减小；ICE 曲线分布较集中，变化趋势也一致；SHAP 散点分析

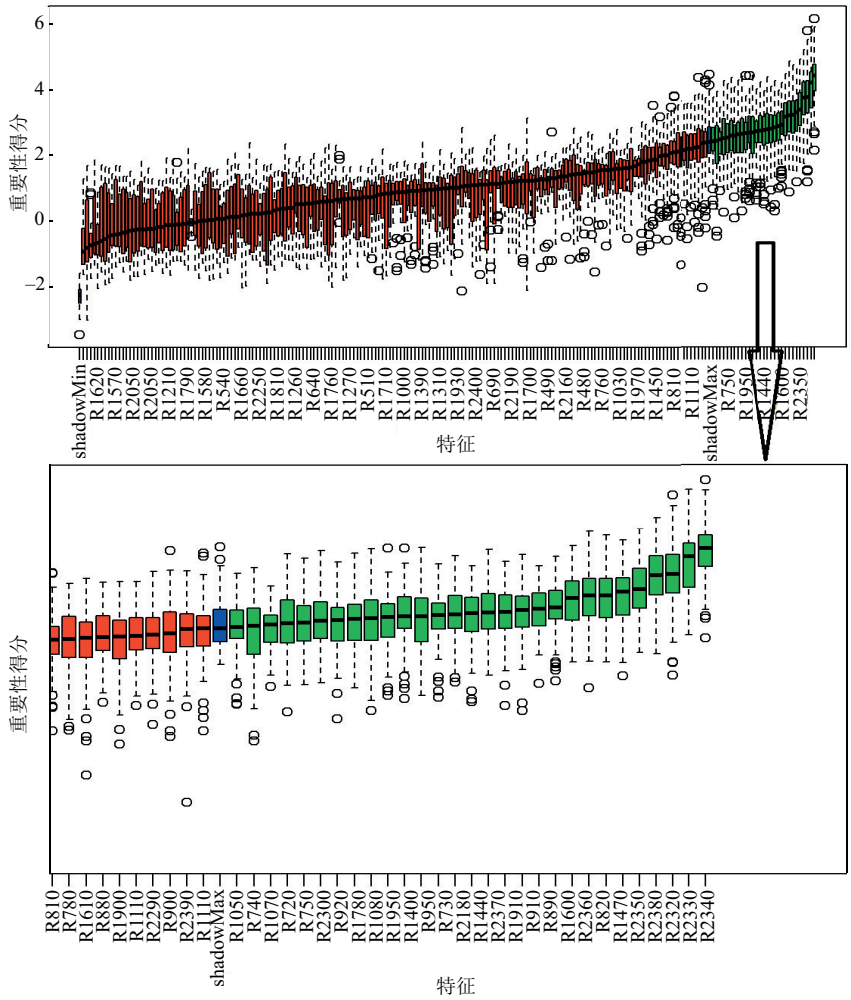



图 7 基于 Boruta 算法的土壤盐分 FOD 特征筛选

**Fig. 7 Fractional-order differential feature selection for soil salinity based on Boruta algorithm**

表 6 土壤盐分模型的反演结果

**Tab. 6 Inversion results of soil salinity model**

| 建模策略 | 模型 | $R^2$ | $e_{rms}$ | $e_{ms}$ |
|---|---|---|---|---|
| 策略Ⅰ | RF | 0.91 | 0.84 | 0.71 |
| | ERT | 0.69 | 1.59 | 2.53 |
| | XGBoost | 0.68 | 1.63 | 2.67 |
| 策略Ⅱ | RF | 0.92 | 0.84 | 0.70 |
| | ERT | 0.90 | 0.91 | 0.82 |
| | XGBoost | 0.89 | 0.94 | 0.88 |
| 策略Ⅲ | RF | 0.91 | 0.86 | 0.75 |
| | ERT | 0.47 | 2.10 | 4.40 |
| | XGBoost | 0.61 | 1.81 | 3.26 |
| 策略Ⅳ | RF | 0.90 | 0.76 | 0.58 |
| | ERT | 0.68 | 1.62 | 2.62 |
| | XGBoost | 0.77 | 1.37 | 1.86 |

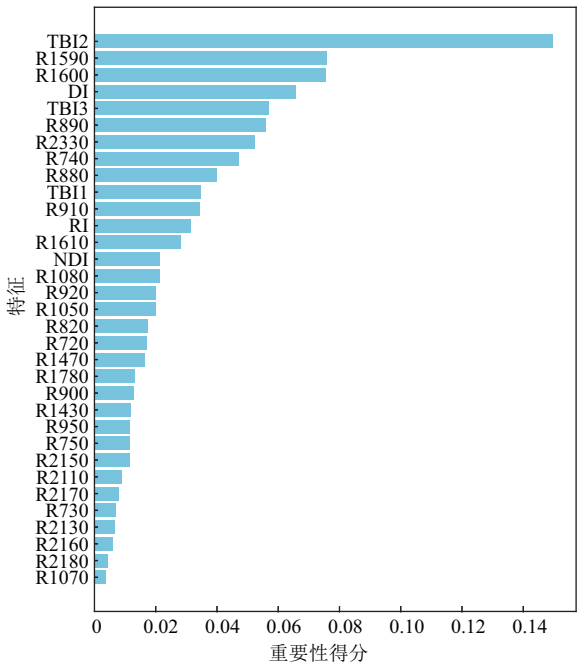
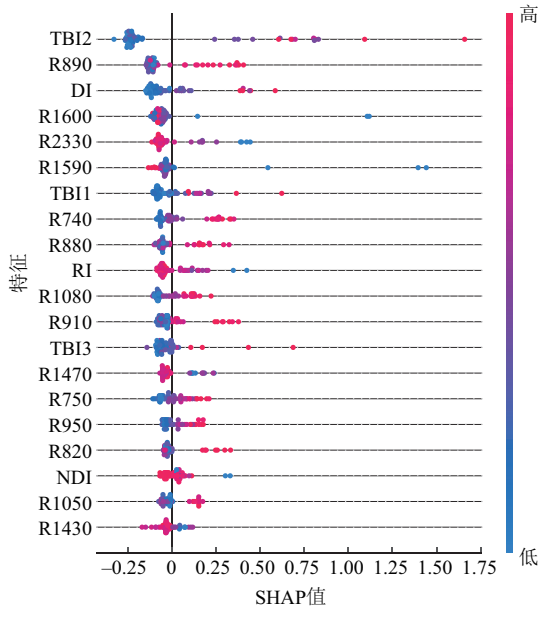
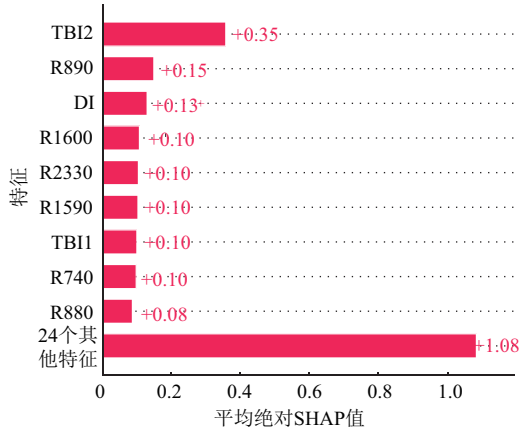



图 8 策略Ⅱ下 RF 模型的特征重要性分析

**Fig. 8 Feature importance analysis of RF model under Strategy Ⅱ**




(a)　SHAP摘要图(RF)




(b)　SHAP特征重要性

**图 9　策略Ⅱ下RF模型的特征解释和全局重要性分析**

**Fig. 9　Individual interpretation and global feature importance of RF model under Strategy Ⅱ**

亦呈现出较高取值对应较低贡献度特征。对于R890、R910，PDP曲线整体变化较为平缓，仅在局部区间出现轻微变化；ICE曲线分布相对集中，说明不同样本响应度差异较小；SHAP散点分布较分散，特征的整体贡献度较低。此外，ICE曲线上，部分特征呈现离散性，显示不同特征响应度存在差异，但总体趋势与PDP保持一致；SHAP散点分布进一步显示，所有特征的模型预测结果存在差异，但整体变化趋势与PDP分析结果一致。需要指出的是，SHAP摘要图与融合解释图的分析结果存在一定差异，这主要是由于这两种方法的分析侧重点不同：SHAP摘要图基于全样本平均绝对贡献值评估特征重要性并对其排序；而融合解释图更侧重刻画特征取值变化对模型预测结果的响应过程。SHAP分析方法能够捕获特征间的交互效应，而PDP在计算过程中可对其他特征作边际平均处理，因此两者在可视化表现上也存在差异。

## 3　讨论

### 3.1　FOD处理对土壤光谱响应的影响

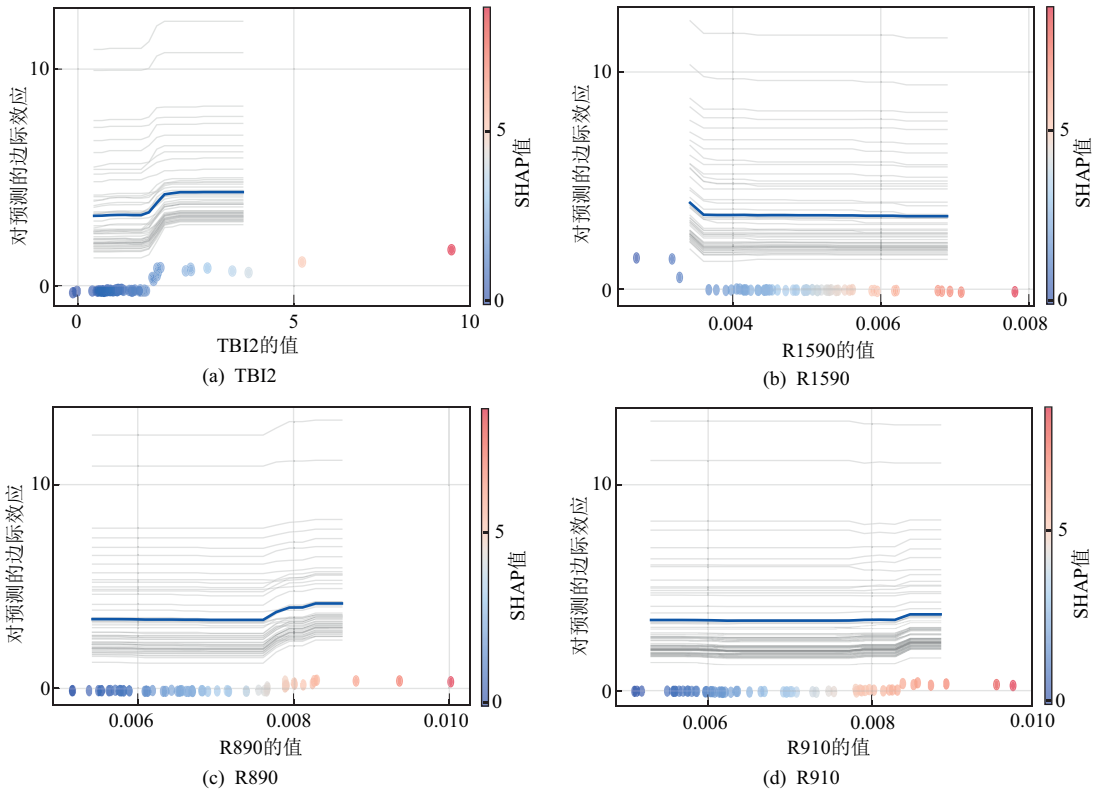
FOD处理可显著提升高光谱数据对土壤盐分的响应能力。相较于原始光谱，FOD处理在保留光谱低频信息同时，大大增强了光谱的局部细节特征，进而有效识别出光谱中反射峰和吸收谷等，从而提高了光谱对土壤属性变化的敏感性[32]。随着FOD阶次增加，1 400、1 900 nm附近的吸收特征逐渐增强，这些波段主要与土壤中O—H键的伸缩振动及水分吸收有关[33]。由于土壤中盐分与水分之间存在显著耦合关系，这些波段对土壤盐分变化也表现出较高敏感性[34]。该研究结果显示，$FOD_{0.75}$处理下预测模型性能最优，表明该阶次下光谱噪声抑制与特征增强达到了较好平衡。较低阶次难以有效突显光谱差异，而较高阶次虽然增强了光谱细节特征，但同时放大了噪声干扰，降低了模型稳定性。这与已有研究结果一致，即适中的FOD处理能够在光谱信息增强与噪声控制之间达到最佳平衡[35]。

### 3.2　光谱指数与FOD处理的协同作用机制

该研究结果显示，融合FOD处理与光谱指数方法，显著提升了土壤盐分的反演精度。在$FOD_{0.75}$处理下，二维及三维光谱指数与土壤盐分的相关性均较强，表明FOD处理能够有效增强构建光谱指数所依赖的关键波段信息。光谱指数通过多波段组合，削弱了光照条件、背景干扰等因素的影响，从而提高了光谱信号稳定性[36]。在此基础上，FOD处理进一步增强了光谱的局部变化特征，提升了光谱指数对微弱光谱差异的识别敏感度。二者结合既增强了光谱信息的表征效果，又提高了模型对复杂非线性关系的拟合能力[37]。因此，FOD处理与光谱指数协同作用，是提升土壤盐分反演精度的重要机制之一。

### 3.3　不同建模策略对模型性能的影响

不同建模策略对模型性能有显著影响。在策略Ⅱ下（融合SG、FOD、Boruta、光谱指数）各模型表现较好。其中，RF模型最稳定，这可能是因为RF模型在训练过程中使用随机属性选择机制，并基于数据差异和随机性提取特征，从而显著提高了模型的预测准确性[25]。该研究结果显示，合理的特征组合对于提升模型预测能力有关键作用。相较而言，仅使用光谱特征（策略Ⅰ），虽然能提高模




注:图中蓝色粗线为PDP曲线,表示特征值对预测结果的平均边际效应;灰色细线为ICE曲线,反映单个样本的响应规律;彩色散点为样本的SHAP值分布;色标从蓝色到红色代表特征值从低到高,直观呈现不同特征值下样本对预测结果的贡献差异与数据分布密度。

**图 10 关键特征对模型预测的部分依赖图**

**Fig. 10 Partial dependence plots of key features on model prediction**

型的拟合精度,但模型的泛化能力有限;在引入光谱指数后(策略Ⅱ),模型能够融合单波段与多波段组合信息,显著提升了其预测性能,这与已有研究得出的多源特征融合效果优于单一特征结论一致[38]。

此外,引入环境变量(策略Ⅲ、Ⅳ)并未显著提升模型性能,可能与空间尺度较大或环境变量与土壤盐分相关性较弱有关[32],这也与相关研究结果类似[10]。该研究结果显示,区域光谱信息仍是驱动土壤盐分变化的主要因素。

### 3.4 模型解释与关键特征贡献分析

RF模型特征重要性分析结果显示,各个特征的贡献度差异具有统计学意义。其中,特征TBI2的贡献度显著高于其他特征,是RF模型反演的核心特征;其次是R1590、R1600、DI;其他特征的贡献度较低。这与土壤盐分光谱响应以敏感波段为主导特征相符[39]。SHAP分析结果进一步揭示了特征作用方向差异。TBI2对模型预测结果具有显著正向促进作用,而RI则表现为负向贡献,表明不同光谱指数在盐分响应中的差异性,这与土壤盐分光谱响应的非线性特征一致[40]。已有研究结果显示,SHAP分析可有效量化特征的边际贡献并识别其正负作用[41]。PDP与ICE分析从响应过程角度补充解释结果。当TBI2在特定区间范围内,模型预测结果呈现明显非线性跃变,表现出阈值效应[33];R1590对模型预测结果表现出稳定负相关趋势,这与短波红外波段对水分、盐类矿物的吸收特性相关[32]。同时,ICE曲线的离散性反映了土壤盐分的空间异质性[42]。总体来看,SHAP与PDP方法的分析结果在变化趋势上具有一致性,但由于方法机制不同[28],二者在贡献分解与响应分析方面具有良好互补性。

## 4 结论

以宁夏平罗县盐渍化土壤为研究对象,基于实测高光谱数据,构建融合分数阶微分(FOD)与多维光谱指数的土壤盐分高精度反演模型;通过SHAP、PDP及ICE等可解释方法,分析特征贡献度与响应机制,并从非线性响应角度揭示土壤盐分与光谱间的复杂作用机制。

1)400~1 400 nm内光谱反射率随着土壤盐分升高而增大,在1 400、1 900、2 200 nm处存在明显

吸收谷，$FOD_{0.75}$处理可有效增强光谱细节特征，其中，三维光谱指数 TBI3 与土壤盐分的相关性最强（$r$=0.849）。

2）多源特征融合方法显著提升了土壤盐分的反演精度。以 SG、FOD、Boruta 与 spectral index 为输入特征时，随机森林（RF）模型的反演效果最好（$R^2$=0.92）；相比之下，引入环境变量对模型性能提升有限，表明光谱信息是决定土壤盐分变化的主导因素。

3）SHAP 分析结果显示，特征 TBI2 对模型预测结果呈正向主导作用，RI 则表现为负向贡献；PDP 与 ICE 曲线揭示了 TBI2 在特定取值范围内对土壤盐分呈阶跃式正向响应，R1590 则表现为平滑负向响应。关键特征对模型预测结果的影响表现出显著非线性特征与阈值效应。多种解释方法结论一致，表明所构建模型具有良好可解释性与稳定性。

## 参考文献：


[1] Lal R. Soil carbon sequestration impacts on global climate change and food security[J]. Science, 2004, 304: 1623-1627.

[2] Chi Yuan, Fan Minglei, Zhang Zhiwei, et al. Zoning the soil salinization levels in the northern China's coastal areas based on high-resolution soil mapping [J/OL]. Ecological Indicators, 2025, 172: 113303. https://doi.org/10.1016/j.ecolind.2025.113303.

[3] Deng Meihua, Zhang Chao, Tang Min, et al. Stacking ensemble learning coupled with multi-source remote sensing data: Enhancing soil salinity inversion accuracy in barley-cultivated salinized soils[J/OL]. Agricultural Water Management, 2025, 322: 109959. https://doi.org/10.1016/j.agwat.2025.109959.

[4] Ivushkin K, Bartholomeus H, Bregt A K, et al. Global mapping of soil salinity change[J/OL]. Remote Sensing of Environment, 2019, 231: 111260. https://doi.org/10.1016/j.rse.2019.111260.

[5] Shokri N, Hassani A, Sahimi M. Multi-scale soil salinization dynamics from global to pore scale: A review [J/OL]. Reviews of Geophysics, 2024, 62 (4). https://doi.org/10.1029/2023rg000804.

[6] Metternicht G I, Zinck J A. Remote sensing of soil salinity: Potentials and constraints[J]. Remote Sensing of Environment, 2003, 85(1): 1-20.

[7] Yu Bo, Yan Changxiang, Yuan Jing, et al. Prediction of soil properties based on characteristic wavelengths with optimal spectral resolution by using Vis-NIR spectroscopy[J/OL]. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 2023, 293: 122452. https://doi.org/10.1016/j.saa.2023.122452.

[8] Du Ruiqi, Chen Junying, Xiang Youzhen, et al. Timely monitoring of soil water-salt dynamics within cropland by hybrid spectral unmixing and machine learning models[J]. International Soil and Water Conservation Research, 2024, 12(3): 726-740.

[9] Bao Yilin, Ustin S, Meng Xiangtian, et al. A regional-scale hyperspectral prediction model of soil organic carbon considering geomorphic features [J/OL]. Geoderma, 2021, 403: 115263. https://doi.org/10.1016/j.geoderma.2021.115263.

[10] Jia Pingping, He Wei, Hu Yi, et al. Inversion of coastal cultivated soil salt content based on multi-source spectra and environmental variables [J/OL]. Soil Tillage Research, 2024, 241: 106124. https://doi.org/10.1016/j.still.2024.106124.

[11] Ben-Dor E, Metternicht G, Goldshleger N, et al. Review of remote sensing-based methods to assess soil salinity [M/OL]. Metternicht G, Zinck J A. Remote Sensing of Soil Salinization: Impact on Land Management. Boca Raton, CRC press, 2008. https://doi.org/10.1201/9781420065039.

[12] 潘昊，陈诗扬，李祎森，等．分数阶微分数据变换在滨海盐渍土盐分反演中的适用性[J]．农业工程学报，2025, 41(3): 73-82.

[13] 周凤玺，滕祥帅，郝君明，等．不同分数阶微分形式在盐渍土电导率高光谱反演中的适用性研究[J]．光谱学与光谱分析，2025, 45(1): 272-281.

[14] Leyden K, Goodwine B. Fractional-order system identification for health monitoring [J]. Nonlinear Dynamics, 2018, 92: 1-18.

[15] Xiao Chao, Ji Qingyuan, Chen Junqing, et al. Prediction of soil salinity parameters using machine learning models in an arid region of northwest China[J/OL]. Computers and Electronics in Agriculture, 2023, 204: 107512. https://doi.org/10.1016/j.compag.2022.107512.

[16] Moradpour M, Rostami Paydar G, Beiranvand Pour A, et al. Landsat-7 and ASTER remote sensing satellite imagery for identification of iron skarn mineralization[J]. Geocarto International, 2022, 37(7): 1971-1998.

[17] 桑娜．平罗县农村改革案例入选[EB/OL]．(2026-02-25) [2026-04-22]. http://www.nxpldj.gov.cn/

djyw/202602/t20260225_1089474.html.

[18] 陆宝金,田生昌,左忠,等.盐渍化土地可持续利用研究综述及展望[J].宁夏大学学报(自然科学版),2023,44(1):79-88.

[19] 陈睿华,尚天浩,张俊华,等.不同光谱类型对银川平原土壤含盐量反演精度的影响与校正[J].应用生态学报,2022,33(4):922-930.

[20] 张晓光,李士美.采样点数量对黄河三角洲垦利县土壤盐分空间变异的影响[J].资源科学,2016,38(12):2375-2382.

[21] Ge Xiangyu, Wang Jinzhe, Ding Jianli, et al. Combining UAV-based hyperspectral imagery and machine learning algorithms for soil moisture content monitoring[J/OL]. PeerJ,2019, 7:e6926. https://doi.org/10.7717/peerj.6926.

[22] Wang Nan, Peng Jie, Chen Songchao, et al. Improving remote sensing of salinity on topsoil with crop residues using novel indices of optical and microwave bands [J/OL]. Geoderma, 2022, 422: 115935. https://doi.org/10.1016/j.geoderma.2022.115935.

[23] 张子鹏,丁建丽,王敬哲,等.利用三维光谱指数定量估算土壤有机质含量:以新疆艾比湖流域为例[J].光谱学与光谱分析,2020,40(5):1514-1522.

[24] Kursa M B, Jankowski A, Rudnicki W R. Boruta: A System for Feature Selection[J]. Fundamenta Informaticae,2010, 101(4): 271-285.

[25] Breiman L. Random forests[J]. Machine learning, 2001, 45: 5-32.

[26] Geurts P, Ernst D, Weehenkel L. Extremely randomized trees[J]. Machine Learning, 2006, 63(1): 3-42.

[27] Elith J, Leathwick J R, Hastie T. A working guide to boosted regression trees[J]. Journal of Animal Ecology,2008, 77(4): 802-813.

[28] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[PP/OL]. V2. arXiv(2017-11-25) [2026-04-22]. https://doi.org/10.48550/arXiv.1705.07874.

[29] Friedman J H. Greedy function approximation: A gradient boosting machine[J]. Annals of Statistics,2001, 29:1189-1232.

[30] Goldstein A, Kapelner A, Bleich J, et al. Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation [J]. Journal of Computational and Graphical Statistics, 2015, 24(1): 44-65.

[31] Liashchynskyi P, Liashchynskyi P. Grid search, random search, genetic algorithm: A big comparison for NAS[PP/OL]. V1. arXiv(2019-12-12)[2026-04-22]. https://doi.org/10.48550/arXiv.1912.06059.

[32] 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):253-261.

[33] Ben-Dor E, Taylor R G, Hill J, et al. Imaging spectrometry for soil applications[J]. Advances in Agronomy,2009, 97: 321-392.

[34] Farifteh J F, van der Meer M, van der Meijde C, et al. Spectral characteristics of salt-affected soils: A laboratory experimen[J], Geoderma,2008, 145(3/4): 196-206.

[35] 黄华雨,丁启东,张俊华,等.基于地面高光谱的宁夏银北地区农田不同土层盐碱化信息反演[J].应用生态学报,2024,35(11):3073-3084.

[36] Xu Xibo, Wang Xiaoguang, Yang Peijie, et al. Strategy for mapping soil salt contents during the bare soil period through a satellite image: Optimal calibration set combined with random forest[J/OL]. Catena, 2023, 223: 106900. https://doi.org/10.1016/j.catena.2022.106900.

[37] 王怡婧,陈睿华,张俊华,等.基于分数阶微分技术的土壤水盐信息高光谱反演[J].应用生态学报,2023,34(5):1384-1394.

[38] 马国林,丁建丽,韩礼敬,等.基于变量优选与机器学习的干旱区湿地土壤盐渍化数字制图[J].农业工程学报,2020,36(19):124-131.

[39] Allbed A, Kumar L. Soil salinity mapping and monitoring in arid and semi-arid regions using remote sensing technology: A review[J]. Advances in Remote Sensing,2013, 2:373-385.

[40] Metternicht G, Zinck J A. Remote sensing of soil salinity: Potentials and constraints[J]. Remote Sensing of Environment,2003, 85(1):1-20.

[41] Huang Feini, Wei Shangguan, Li Qingliang, et al. Beyond prediction: An integrated post-hoc approach to interpret complex model in hydrometeorology[J/OL]. Environmental Modelling & Software, 2023, 167: 105762. https://doi.org/10.1016/j.envsoft.2023.105762.

[42] Molnar C. Interpretable machine learning: A guide for making black box models explainable[M/OL]. Raleigh: Lulu.com, 2020. https://originalstatic.aminer.cn/misc/pdf/Molnar-interpretable-machine-learning_compressed.pdf.

# Research on Hyperspectral Inversion of Soil Salinity Based on Fractional-Order Differentiation and Multidimensional Spectral Indices: A Case Study of Pingluo County, Ningxia

*JIA Pingping*[1,2,3], *ZHAO Xiaoning*[1,2,3*], *XU Shiwen*[1,2,3], *TANG Chong*[1,2,3], *QI Jiejun*[1,2,3], *WANG Wenting*[1,2,3]

(1. Qinling National Park Research Institute, Shaanxi University of Technology, Hanzhong 723001, China;
2. Qinba State Key Laboratory of Biological Resources and Ecological Environment (Cultivation), Shaanxi University of Technology, Hanzhong 723001, China;
3. School of Biological Science and Engineering, Shaanxi University of Technology, Hanzhong 723001, China)

**Abstract:** Soil salinization is a key issue of global soil degradation. To improve the inversion accuracy of soil salinity, this study takes Pingluo County in Ningxia as the study area. Based on the spectral data collected from spectrometers, the spectral characteristics of soils with different salinization degrees are systematically analyzed. The fractional-order (0-2.00 order) differentiation method is employed to enhance the effective information of the spectra, combined with two-dimensional and three-dimensional spectral indices and the Boruta feature selection algorithm to construct three types of soil salinity inversion models: Random Forest (RF), Extremely Randomized Trees (ERT) and Extreme Gradient Boosting (XGBoost). The results show that typical absorption valleys appear in the soil spectra around 1,400 nm, 1,900 nm and 2,200 nm. The fractional-order differentiation method effectively enhances the spectral absorption features, with the 0.75-order differentiation transformation showing the strongest correlation between the two-dimensional and three-dimensional spectral indices and soil salinity, with correlation coefficients ($r$) of −0.786 8 and 0.849, respectively. The Boruta algorithm identifies 27 key features. Comparative analysis of the models shows that among the four modeling strategies, the RF model is the most stable and accurate, achieving optimal fitting results ($R^2$=0.92, $e_{rms}$=0.84) when SG, FOD, Boruta, and Spectral Index are used as input features. Feature importance analysis indicates that feature TBI2 is the primary driving factor for model predictions (importance=0.14), and SHAP explanations further confirm that TBI2 contributes positively to soil salinity predictions (mean SHAP value=0.56), while feature RI contributes negatively. Partial Dependence Plots (PDP) and Individual Conditional Expectation (ICE) curves reveal that when TBI2 values are around 2, the model's predictions exhibit a stepwise positive response, while R1590 shows a smooth negative response. The analyses from various interpretative methods are consistent. This research provides important references for hyperspectral inversion of salinity in salinized soils.

**Key words:** soil salinization; hyperspectral inversion; fractional-order differentiation; machine learning; interpretable methods; Pingluo county, Ningxia

(责任编辑　高继红)