

基于 CWT – Stacking 集成学习模型的灌区土壤水盐含量和 pH 值估算

李瑞平^{1,2} 孙伟琦¹ 元志辉¹ 马小茗^{1,3} 赵建伟¹ 阿莫尔¹
陈忠冉¹ 宝路¹ 王思楠⁴

(1. 内蒙古农业大学水利与土木建筑工程学院, 呼和浩特 010018;
2. 内蒙古农业大学旱区水工程生态环境全国重点实验室, 呼和浩特 010018;
3. 内蒙古自治区水利科学研究院, 呼和浩特 010010;
4. 水利部牧区水利科学研究所, 呼和浩特 010020)

摘要: 土壤含水率、含盐量、pH 值精确估算对于干旱半干旱农业灌区的水土环境监测与可持续土地利用至关重要。本文以河套灌区农田和其他植被覆盖区为研究对象,采用一阶导数(First-order derivative)、二阶导数(Second-order derivative)、包络线去除(Continuum removal, CR)、标准正态变换(Standard normal variate, SNV)以及连续小波变换(Continuous wavelet transform, CWT)对原始光谱进行预处理与特征增强,并选取偏最小二乘回归(Partial least squares regression, PLSR)、支持向量机(Support vector machine, SVM)、随机森林(Random forest, RF)、极端梯度提升(Extreme gradient boosting, XGBoost)和轻量级梯度提升机(Light gradient boosting machine, LightGBM)5种基学习器,结合 Stacking(堆叠泛化)、Voting(投票集成)、Bagging(装袋集成)3种集成策略,构建土壤水盐含量、pH 值综合估算模型。结果表明:小波变换在去除高频噪声与凸显局部特征方面表现最优,适用性评分达到4.7,显著提升了光谱数据与土壤水盐含量、pH 值的敏感性。相较于单一模型,集成学习策略有效降低了过拟合风险,并提高了模型泛化能力,其中 Stacking 集成策略综合表现最佳,对土壤含水率、含盐量、pH 值训练集决定系数分别提升4.9%、4.7%和5.1%。基于 CWT – Stacking 构建的最优模型实现了目标水盐含量、pH 值的高精度估算,土壤含水率、含盐量、pH 值验证集决定系数分别达到0.85、0.89和0.82,均方根误差(RMSE)分别为2.26%、0.61 g/kg和0.30。先进的光谱变换技术与复杂的集成学习算法相结合,为干旱和半干旱地区大尺度土壤健康状况动态监测提供了高精度、可行的技术手段。

关键词: 土壤水盐含量和 pH 值估计; 地物光谱; 集成学习; 光谱变换; 河套灌区

中图分类号: S278 **文献标识码:** A **文章编号:** 1000-1298(2026)20-0094-14

OSID:



Estimation of Soil Water – Salt Content and pH Value in Irrigation Districts Based on CWT – Stacking Ensemble Learning Model

Li Ruiping^{1,2} Sun Weiqi¹ Yuan Zhihui¹ Ma Xiaoming^{1,3} Zhao Jianwei¹
A Moer¹ Chen Zhongran¹ Bao Lu¹ Wang Sinan⁴

(1. College of Water Conservancy and Civil Engineering, Inner Mongolia Agricultural University, Hohhot 010018, China
2. State Key Laboratory of Water Engineering Ecology and Environment in Arid Area, Inner Mongolia Agricultural University, Hohhot 010018, China
3. Water Resources Research Institute of Inner Mongolia Autonomous Region, Hohhot 010010, China
4. Institute of Water Resources Science for Pastoral Area, Ministry of Water Resources, Hohhot 010020, China)

Abstract: Accurate estimation of soil moisture, salinity and pH value is essential for soil and water environment monitoring and sustainable land use in arid and semi-arid agricultural irrigation districts. Taking farmland and other vegetation-covered areas in the Hetao Irrigation District as the research objects, the first-order derivative, second-order derivative, continuum removal, standard normal variate

收稿日期: 2026-06-03 修回日期: 2026-07-04

基金项目: 国家自然科学基金黄河水科学研究联合基金重点支持项目(U2443210)、内蒙古自治区科技计划项目(2025YFHH0166)和内蒙古自治区自然科学基金项目(2026MS0804)

作者简介: 李瑞平(1973—),男,教授,博士生导师,主要从事节水灌溉与农业水利遥感技术研究, E-mail: nmglrp@163.com

and continuous wavelet transform (CWT) were adopted to preprocess raw spectra and enhance features. Five base learners, including PLSR, SVM, RF, XGBoost and LightGBM were selected, and three ensemble strategies of Stacking, Voting and Bagging were combined to construct comprehensive estimation models for soil moisture, salinity and pH value. The results showed that continuous wavelet transform performed the best in removing high-frequency noise and highlighting local features, with an applicability score of 4.7, which greatly improved the sensitivity of spectral data to soil moisture, salinity and pH value. Compared with single models, ensemble learning strategies effectively reduced the risk of overfitting and enhanced the generalization ability of models. Among all strategies, Stacking achieved the best overall performance, and the coefficients of determination R^2 of the training sets for soil moisture, salinity and pH value were increased by 4.9%, 4.7% and 5.1%, respectively. The optimal model constructed by CWT-Stacking realized high-precision estimation of soil moisture, salinity and pH value. The R^2 of the validation sets reached 0.85, 0.89 and 0.82, and the corresponding root mean square errors (RMSE) were 2.26%, 0.61g/kg and 0.30, respectively. The combination of advanced spectral transformation technology and complex ensemble learning algorithms can provide a high-precision and feasible technical means for the dynamic monitoring of large-scale soil health in arid and semi-arid regions.

Key words: soil water, salt content, and pH value estimation; ground object spectral; ensemble learning; spectral transformation; Hetao Irrigation District

0 引言

土壤是维系农业生产、生态系统稳定与区域可持续发展的核心基础,其中含水率、含盐量与 pH 值是表征干旱半干旱灌区土壤质量的关键指标,直接调控作物生长发育、养分有效性及土壤生态过程^[1]。河套灌区作为我国西北干旱半干旱区典型农业灌区,气候干旱、蒸发强烈,土壤盐渍化风险高、空间异质性显著。因此快速、精准地获取土壤水盐含量与 pH 值信息,对灌区盐渍化防控、精准灌溉及土地资源高效利用具有重要意义^[2-3]。与传统土壤理化监测方法相比,高光谱数据具有光谱分辨率高、波段密集且连续优势,利用高光谱数据可充分挖掘土壤的光谱特征,构建高精度估算模型,实现土壤水盐含量、pH 值快速准确估算^[4-5]。

高光谱定量分析通过建立光谱与物质含量之间的数学关系,利用光谱数据估算物质含量,因此省去了繁琐的化学分析步骤,在土壤含水率、含盐量及 pH 值估算中展现出巨大应用潜力^[6-10]。可见光-近红外波段能有效捕捉土壤组分光谱吸收特征,为定量建模提供可靠数据支撑^[10-14]。一阶导数、二阶导数、包络线去除、标准正态变换等预处理方法被广泛用于光谱特征增强以削弱仪器噪声、基线漂移、背景干扰等影响^[15-17]。连续小波变换 (CWT) 可在时频域内实现信号精细解析,在降噪与微弱特征提取方面优于常规变换,更适用于复杂土壤光谱信号处理^[18-22]。但高光谱数据存在维度高、冗余度大、非线性关系强等问题,单一预处理方法难以充分挖

掘有效特征,制约模型精度提升^[23-25]。此外在干旱半干旱灌区,土壤含水率、含盐量及 pH 值并非相互独立的随机变量,盐分累积与运移直接影响土壤溶液的离子组成及酸碱缓冲体系,而水分蒸发-盐分结晶的周期性循环则是驱动盐分表聚和 pH 值升高的核心水文-地球化学过程。三者的时空演变受控于共同的气候、水文与人为灌溉管理因素,存在内在的耦合关系^[26]。因此,在灌区复杂环境下开展三者的同步估算,不仅在科学上更具合理性,在工程应用中也更有实际价值。

机器学习为解决土壤光谱非线性建模问题提供了有效途径,PLSR、SVM、RF、XGBoost、LightGBM 等模型已广泛应用于土壤水盐含量及 pH 值估算^[27-30]。相较于单一模型,集成学习通过多模型优势互补,可有效降低过拟合风险、提升泛化能力,成为提升反演精度的重要手段^[31-33]。Stacking、Voting、Bagging 等集成策略在土壤水盐含量、pH 值预测中表现突出,预测能力更强、表现出更好的稳定性和适用性^[34-36]。尽管相关研究已取得诸多进展,但多数成果仍以单一模型参数优化或浅层集成策略为主,系统地将数学变换与多种小波变换相结合增强光谱特征,并与集成策略协同估算土壤属性的研究是土壤水盐遥感反演亟待解决的关键技术问题,特别是面向灌区复杂环境下多指标同步反演的相关方法体系与应用亟需完善^[37-40]。

基于此,本文以河套灌区为研究区,系统对比分析多种光谱变换方法的特征增强效果;结合优选后的机器学习基模型,构建 Stacking、Voting、Bagging 集成模型;建立 WT-Stacking 土壤水盐含量、pH 值同

步估算模型,以期为干旱半干旱灌区土壤大尺度动态监测与精准管理提供技术支撑。

1 材料和方法

1.1 研究区概况

河套灌区位于内蒙古自治区西部巴彦淖尔市(图 1),地理坐标为 $40^{\circ}19' \sim 41^{\circ}18'N$, $106^{\circ}20' \sim 109^{\circ}19'E$,总面积约为 $1.19 \times 10^4 \text{ km}^2$ 。北与蒙古国接壤,东北与包头市相邻,东南与鄂尔多斯市毗连,南与宁夏回族自治区相接,西南与阿拉善盟相连。地势由南向北倾斜,南部为河套平原,中部横亘阴山

山脉(含乌拉山),北部为乌拉特草原,西部零散分布乌兰布和沙漠等戈壁沙漠。巴彦淖尔市属典型的中温带大陆性干旱半干旱气候,年平均气温 $3.7 \sim 7.6^{\circ}\text{C}$,最热月(7 月)平均气温 $22 \sim 25^{\circ}\text{C}$,最冷月(1 月)平均气温 $-15 \sim -10^{\circ}\text{C}$;年平均降水量 $130 \sim 215 \text{ mm}$,主要集中 6—9 月,占全年降水量的 70% 以上;年蒸发量 $2\,032 \sim 2\,881 \text{ mm}$,为降水量的 10~20 倍,干旱指数高达 $9 \sim 22$;年日照时数 $3\,100 \sim 3\,300 \text{ h}$,太阳辐射强度大;春季多风沙天气,年均风速 $2.5 \sim 3.5 \text{ m/s}$ 。上述地理与气候特征共同塑造了河套灌区独特的土壤环境。

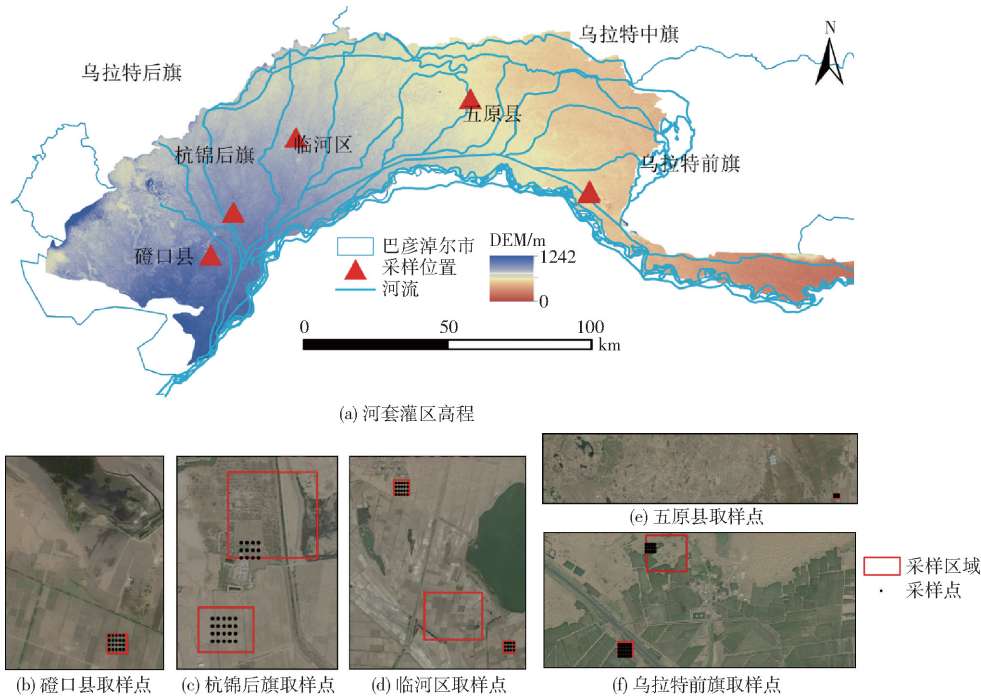


图 1 研究区概况
Fig.1 Overview of study area

1.2 数据来源与预处理

1.2.1 土壤理化指标

综合考虑土地利用类型、土壤类型、地形地貌等因素,在研究区内布设 375 个采样点(图 1)。采样点分布:农田 240 个采样点,其他植被覆盖区 135 个采样点,不同区域内采样点间距保持 $30 \sim 40 \text{ m}$,区域之间保持 $50 \sim 100 \text{ km}$,确保空间代表性。采样时间为 2025 年 4 月。在每个采样点采集深度 $0 \sim 20 \text{ cm}$ 土壤样品,混合后取 1 kg 土样,土壤样品经风干、研磨、过筛处理后,采用电导率仪测定土水质量比 1:5 浸提液电导率(EC),通过标定回归方程将 EC 换算为土壤含盐量(g/kg)。采用烘干法测定土壤含水率。土壤 pH 值采用电位法测定,浸提剂为去除 CO_2 的去离子水,土液质量比为 1:2.5,浸提液经振荡平衡后,用酸度计测定 pH 值。

采样点土壤水盐含量、pH 值统计分布特征如

图 2 所示。土壤含盐量中位数和均值均在 3.5 g/kg 附近,分布较为集中,大部分样本处于中低水平,但也存在少数超过 8.0 g/kg 高值点。土壤含水率分布范围较广,均值与中位数均接近 15%,主要分布在 $5\% \sim 27\%$ 之间,反映了研究区土壤湿度的空间异质性。土壤 pH 值整体呈碱性,中位数与均值高度一致(约 8.5),分布高度集中,绝大多数样本 pH 值介于 $8.0 \sim 9.5$ 之间。

1.2.2 光谱数据获取和预处理

使用莱森光学(LiSen Optics)宽光谱微型光谱仪获取地面光谱数据,光谱范围为 $200 \sim 2\,500 \text{ nm}$,光谱分辨率为 1 nm ,测量条件为晴朗天气、太阳天顶角小于 60° ,每个采样点测量 5 次。取平均值采用 Savitzky-Golay 平滑滤波去除光谱噪声,使用 3σ 准则剔除异常光谱数据,使用标准白板校正,计算相对反射率,将光谱数据重采样到 10 nm 间隔,减少数

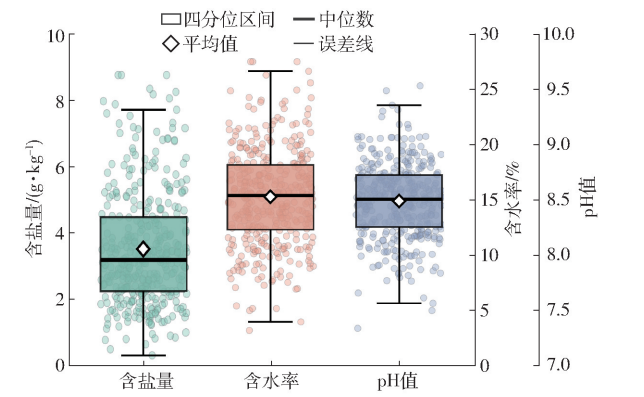


图 2 土壤水盐含量、pH 值箱线图

Fig.2 Box plot of soil water and salt content as well as pH value

据维度,剔除质量差的光谱数据,最终保留 332 个有效样本。

1.3 研究方法

高光谱数据由于维度高且极易受到背景环境及仪器噪声的干扰,直接建模易导致模型过拟合或精度受限。因此,采用一阶导数变换、二阶导数变换、包络线去除变换(Continuum removal, CR)、标准正态变换(Standard normal variate, SNV)、连续小波变换(Continuous wavelet ransform CWT)5 种经典光谱变换算法对原始反射率数据进行预处理及特征增强。

小波变换能够同时在频域和空间域(波长域)分析信号,尤其适用于处理高光谱这些包含复杂环境干扰的非稳定信号。为在感知光谱高频噪声的同时最大程度特征保留吸收峰信息,引入连续小波变换对原始光谱进行多信号重构。

由于不同小波基函数对光谱微弱特征的响应敏感度存在显著差异,为系统对比并优选出最适合提取土壤水盐含量、pH 值光谱特征的变换方法,分别采用高斯一阶导数小波(Gaussian first-derivative wavelet, Gaus1)、墨西哥帽小波(Mexican hat wavelet, Mexh)和莫雷特小波(Morlet wavelet, Morl)3 种非经典解析小波基,对光谱信号进行特征增强。

其中 Gaus1 基函数呈单峰形态,能有效放大光谱吸收谷、反射峰的边缘变化特征,用于增强土壤光谱中与水盐组分相关的弱吸收信号,提升特征波段辨识度。

Mexh 可精准定位吸收峰与反射峰的中心位置,同时对高频噪声具有一定的平滑作用,用于提取光谱的多尺度形态特征,区分土壤水盐组分的不同吸收响应。

Morl 基函数兼具高斯窗的局部化特性与复指数函数的频率分析能力,能在不同尺度下捕捉光谱

信号的细节变化与整体趋势,用于挖掘土壤水盐光谱的多尺度时频特征。

1.4 机器学习方法与集成策略

1.4.1 机器学习模型

分别采用偏最小二乘回归(PLSR)、支持向量机(SVM)、随机森林(RF)、极端梯度提升(XGBoost)与轻量级梯度提升(LightGBM)5 种模型,基于高光谱反射率数据进行土壤水盐含量、pH 值估算。

所有样本按比例 7:3 随机划分为训练集与验证集,训练集用于模型拟合与参数调优,验证集用于模型泛化能力评估。模型预测性能采用 4 项核心指标综合评价:决定系数(R^2)、均方根误差(RMSE)、平均绝对误差(MAE)及性能偏差比(RPD)。

1.4.2 集成策略

为克服单一模型在高维光谱数据下易陷入局部最优及泛化能力受限的问题,引入 3 种集成框架(Stacking(堆叠泛化)、Voting(投票集成)和 Bagging(装袋集成))以进一步优化预测性能(图 3)。

其中 Stacking 集成采用层级级联架构。第 1 层(Base-layer)并行训练上述多种异构基学习器;第 2 层训练元学习器(Meta-learner),以第 1 层模型预测输出作为输入特征进行融合重构,实现多模型的优势互补。Voting 集成包含简单投票与加权投票机制。加权投票基于各模型在验证集上性能表现动态分配权重,赋予高精度模型更大的决策比例,从而提升最终估算结果的可靠性。Bagging 集成利用 Bootstrap 重采样技术生成多个相互独立的训练子集,分别训练同构模型后对预测结果取算术平均,以此有效抑制模型方差,增强应对数据波动的稳健性。

2 结果与分析

2.1 光谱变换特性

2.1.1 不同土地利用类型光谱特征

研究区内农田与林草覆盖区土壤的光谱反射率曲线整体形态存在差异,但在特征吸收带上表现出高度共性(图 4)。在可见光波段(400~700 nm),农田土壤反射率极低(自 0.03 左右起步),随后随波长增加迅速攀升,在 800~1 800 nm 区间形成 0.14~0.16 高反射平台;而林草等其他植被覆盖区在该波段起点更低(接近于 0),并在 600~700 nm 区间呈现陡峭上升趋势,整体最高反射率(约 0.125)略低于农田。在短波红外波段(1 300~2 500 nm),2 类用地光谱均在 1 400、1 900、2 200 nm 附近出现显著吸收特征。其中,1 900 nm 处吸收最为剧烈,主要对应水分子和羟基的强吸收带,农田在此波段反射率跌至

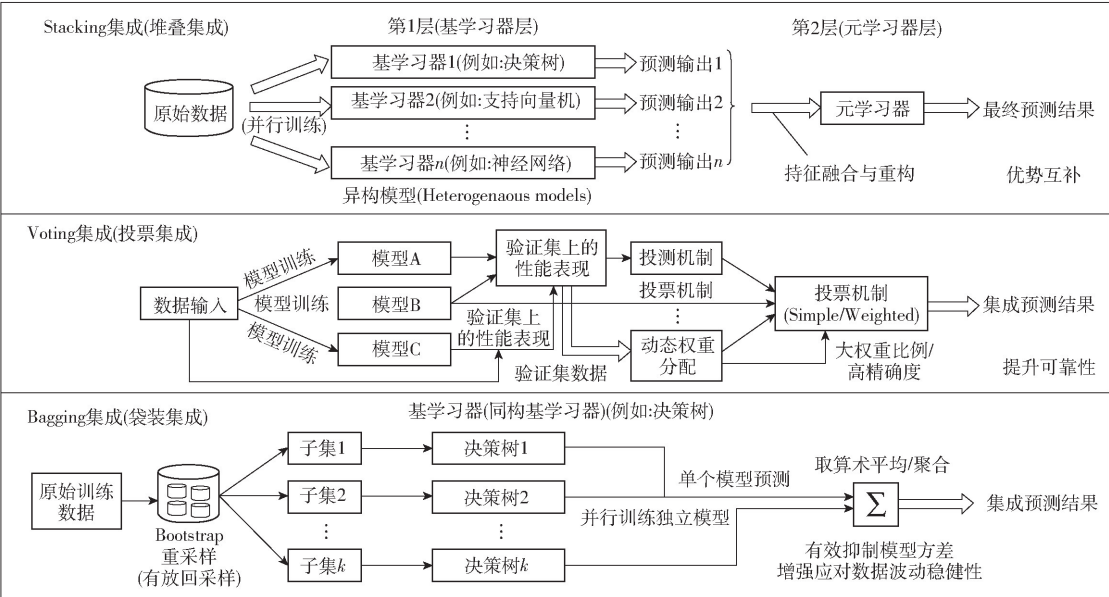


图 3 集成策略示意图

Fig. 3 Integration strategy diagram

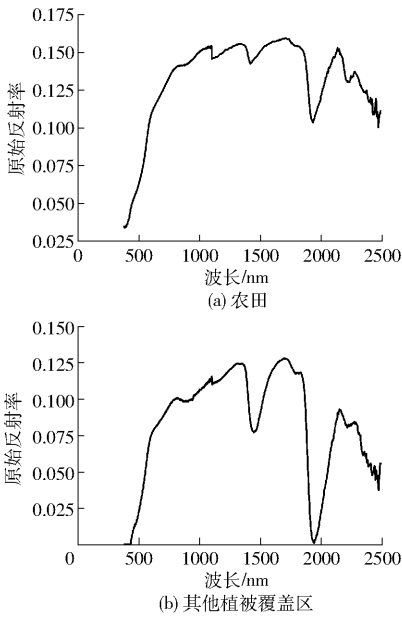


图 4 原始光谱反射率曲线

Fig. 4 Original spectral reflectivity curves

0.10 左右,而植被覆盖区受更高含水率或植被冠层影响,该处反射率几乎骤降至 0;相比之下,1 400 nm

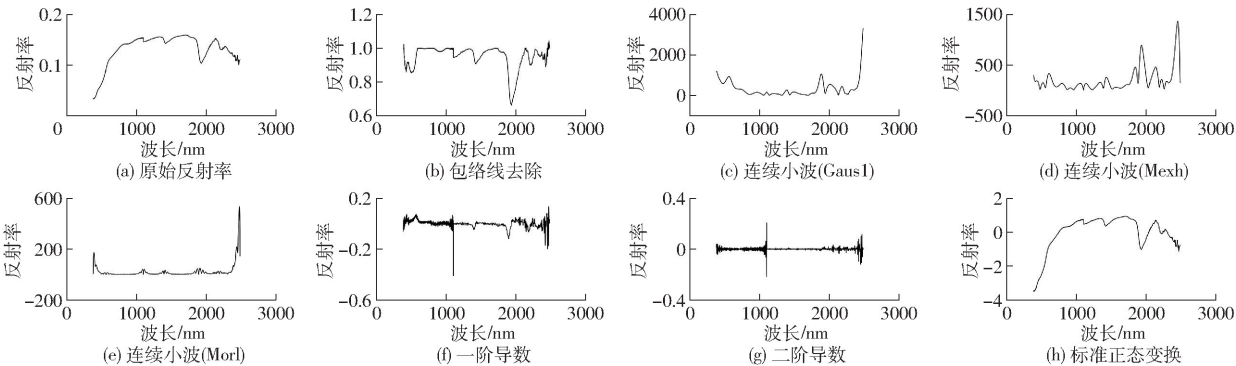


图 5 不同光谱变换方法特征曲线(农田)

Fig. 5 Characteristic curves of different spectral transformation methods (farmland)

处存在次级吸收谷,而 2 200 nm 处(指示黏土矿物含量)吸收特征则相对微弱。

2.1.2 光谱变换前后特征

原始光谱经不同数学与小波变换处理后,特征增强效果呈现显著差异(图 5、6)。由图 5f、6f 可以看出,一阶导数有效消除了光谱基线漂移,使 1 400、1 900 nm 区域隐藏的细微吸收特征得以显著显现。二阶导数虽进一步放大局部曲率特征,但在全波段特别是高频区域引入强烈的振荡噪声,降低了有效信噪比(图 5g、6g)。标准正态变换对样本整体的散射效应和基线偏移具有一定校正作用,使光谱整体分布更集中、趋势更稳定,但该方法仅能实现光谱整体的标准化处理,对局部吸收特征的增强效果有限,1 400、1 900 nm 等关键吸收谷的辨识度并未得到明显提升(图 5h、6h)。包络线去除在农田土壤样本中表现优异,将 1 900 nm 等特征波段的吸收深度显著拓宽至 0.35 左右;但该方法对局部奇异值较为敏感,在其他植被覆盖区样本中易受短波端异常值干扰而导致基线严重失真(图 6b)。相较于常规数学

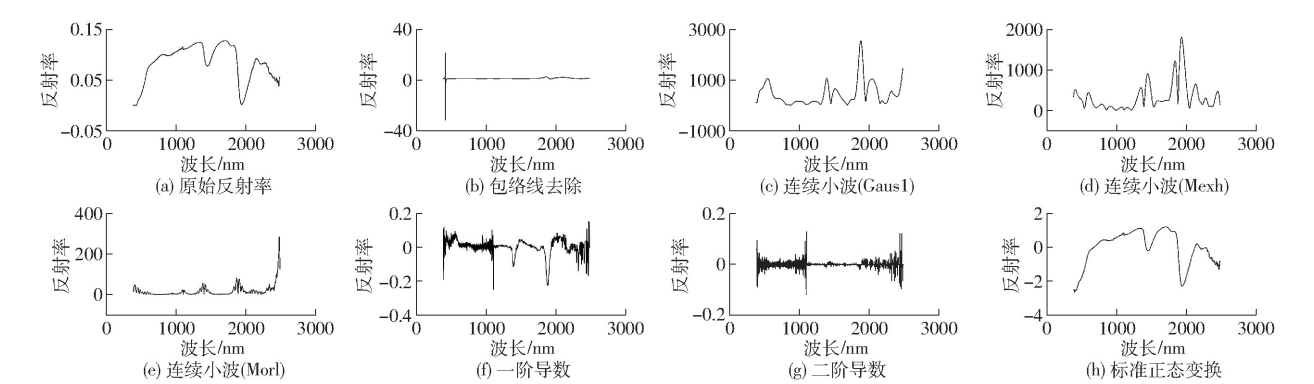


图 6 不同光谱变换方法特征曲线(其他植被覆盖区)

Fig. 6 Characteristic curves of different spectral transformation methods (other vegetation)

变换,连续小波变换在实现光谱降噪与特征提取方面综合优势最为突出,但不同解析小波基的特征响应各异。莫雷特小波(Morl)虽能捕捉周期性微弱信号,但在光谱端点($>2\,400\text{ nm}$)产生了极其明显的边界效应,易造成假性特征干扰(图 5e、6e);墨西哥帽小波(Mexh)能够准确定位吸收峰拐点,但平滑能力有限,导致处理后的光谱呈锯齿状,包含较多非中心的最高尖峰(图 5d、6d)。相比之下,高斯一阶导数小波(Gaus1)表现出最优化的综合时频局部化特性,不仅有效滤除了环境高频干扰噪声,而且精准放大了 $1\,400$ 、 $1\,900$ 、 $2\,200\text{ nm}$ 附近与土壤水盐组分密切相关的吸收信号(图 5c、6c)。Gaus1 分解得到的小波系数精准捕捉局部突变特征,适用性和稳定性最高,为高精度估测模型构建提供了最优光谱输入变量。

2.1.3 光谱特征与土壤水盐含量、pH 值相关系数

通过相关性分析,识别不同土地利用类型下土壤水盐含量、pH 值光谱响应特征存在显著差别(图 7、8)。整体而言,单一光谱波段与土壤水盐含量、pH 值的全局线性相关性以中低强度为主。

对于土壤含盐量,光谱响应方向在 2 类用地中发生显著反转。在农田区域(图 7),含盐量与全波段反射率多呈弱负相关(相关系数主要在 $-0.2 \sim -0.35$ 之间波动);而在其他植被覆盖区(图 8),可能受植被冠层及盐分胁迫的综合影响,转变为中等强度正相关(相关系数多集中在 $0.3 \sim 0.5$ 之间)。这种全局偏低的线性相关性表明,盐分信息在光谱序列中存在严重的非线性隐蔽性。

对于土壤含水率高的农田区域(图 7)常规连续光谱(如原始反射率及小波变换)与其表现出较平缓的相关趋势。然而,经过包络线去除(CR)变换后,在其他植被覆盖区(图 8),位于 970 、 $1\,200$ 、 $1\,400$ 、 $1\,900\text{ nm}$ 附近的水分子基频与倍频吸收波段被精准剥离,暴露出局部高相关特征的窄带。说明特定数学变换对挖掘隐藏水分敏感波段具有关键

作用。

对于土壤 pH 值,在农田区(图 7)几乎不表现出明显的线性光谱响应(全波段相关系数趋近于 0);但在其他植被覆盖区(图 8),pH 值与可见光及近红外波段呈显著的广域负相关(相关系数达 $-0.4 \sim -0.6$)。进一步印证了 pH 值本身缺乏直接的物理吸收特征,其光谱响应本质上高度依赖于由土地利用类型和地表覆盖所调控的土壤伴生矿物与有机质的协变关系。

综合来看,土壤水盐含量、pH 值与单一波段光谱的线性相关系数未能突破 0.6。从数据底层逻辑上证实:在干旱半干旱灌区,面对空间异质性极强的复杂土壤组分,仅依赖简单的线性关系难以实现高精度反演,须引入非线性特征挖掘能力极强的集成机器学习算法,方能充分解析高维光谱数据与土壤属性间的复杂映射关系。

为评估 5 种光谱变换方法的综合性能,从信噪比提升率、特征增强度、计算复杂度和场景适用性 4 个关键维度构建评估体系,各评估维度定义与量化方式如下:信噪比提升率以变换前后光谱曲线的平滑程度、噪声抑制效果为依据,通过曲线光滑度与噪声幅值变化进行量化打分;特征增强度以变换后光谱与土壤含水率、含盐量、pH 值的平均相关系数提升幅度为评价依据,相关系数提升越显著得分越高;计算复杂度依据算法运算量、迭代次数与运行耗时进行分级评定,复杂度越低得分越高;场景适用性综合农田与其他植被覆盖区 2 种地表类型下模型拟合优度(R^2)、稳定性与鲁棒性进行打分。各维度采用 5 分制(1 为最差,5 为最优)进行独立评分,再通过等权重加权得到综合适用性评分。结果如表 1 所示,小波变换虽计算复杂度最高,但在特征增强度与信噪比提升方面均表现最优,能有效滤除高频噪声并放大土壤含水率、含盐量、pH 值相关的微弱吸收特征,综合适用性评分最高(4.7/5.0);包络线去除变换次之,特征增强度与小波变换相当,

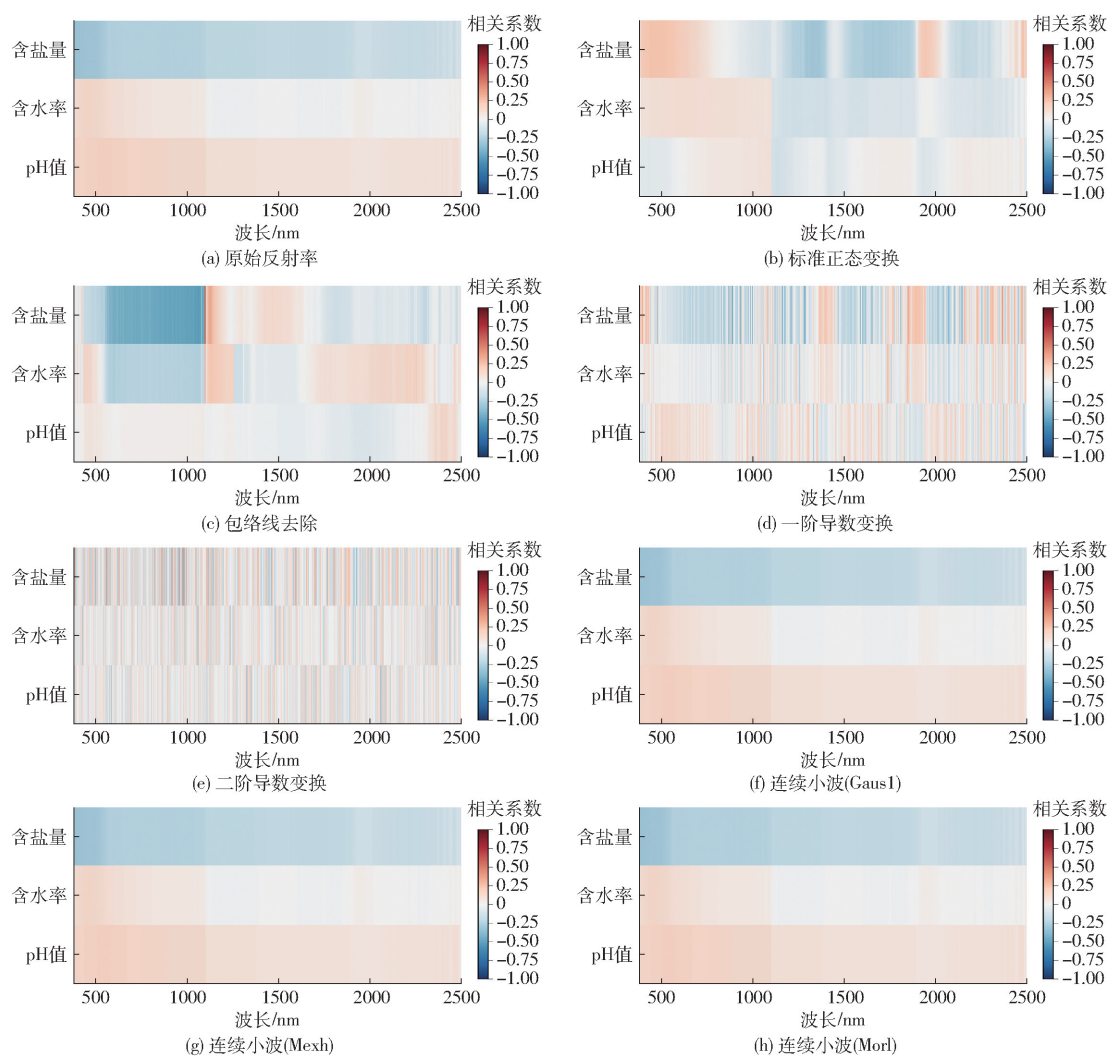


图 7 光谱反射率与土壤水盐含量、pH 值相关系数矩阵(农田)

Fig. 7 Correlation coefficient matrix of spectral reflectance with soil moisture, salinity, and pH value (farmland)

但在信噪比提升方面略低;一阶导数与标准正态变换综合表现中等,可满足常规建模需求;二阶导数因易放大高频噪声、降低曲线平滑度,导致适用性评分最低。

2.2 模型性能评估

2.2.1 单一模型性能评估

为定量评估不同算法对农田和其他植被覆盖区域土壤含水率、含盐量、pH 值的光谱预测能力,构建并比较 5 种单一模型:偏最小二乘回归(PLSR)、支持向量机(SVM)、随机森林(RF)、XGBoost 和 LightGBM。为直观对比各模型在多维度指标上综合表现,绘制了模型综合性能雷达图(图 9、10)。综合分析雷达图各项维度指标可知,不同算法在构建土壤属性估测模型时表现出显著的精度差异。整体而言,基于树结构和集成思想的非线性模型(SVM、RF、XGBoost、LightGBM)建模效果全面优于传统线性模型(PLSR)。表明在面对土壤组分复杂、光谱存在严重多重共线性和冗余信息时,线性回归难以充分拟合光谱与土壤含水率、含盐量、pH

值(尤其是含水率和 pH 值)之间的复杂非线性映射关系。

在非线性模型中,XGBoost 表现出最优综合预测能力与稳健性。具体而言,在农田土壤含盐量估算中(图 9b),XGBoost 获得了最高预测精度,验证集决定系数(R^2)达到 0.84,均方根误差(RMSE)为 0.65 g/kg,且相对分析误差(RPD)达 2.60(RPD 大于等于 2.0 具备极强的定量预测能力)。相比之下,PLSR 模型 R^2 仅为 0.72,拟合效果相对较差。RF 模型紧随 XGBoost 之后,表现出较好的抗过拟合能力。在土壤含水率与 pH 值的估算中,XGBoost 依然保持领先, R^2 分别为 0.80 和 0.82。主要得益于 XGBoost 采用的梯度提升框架能够有效捕捉高维光谱序列中的微弱敏感信息。基于上述分析,在后续的集成模型构建中,优先选择表现优异的模型作为基学习器。

2.2.2 集成模型性能评估

尽管 XGBoost 等单一模型表现优异,但面对干旱半干旱区土壤强烈的空间异质性,单一模型仍存

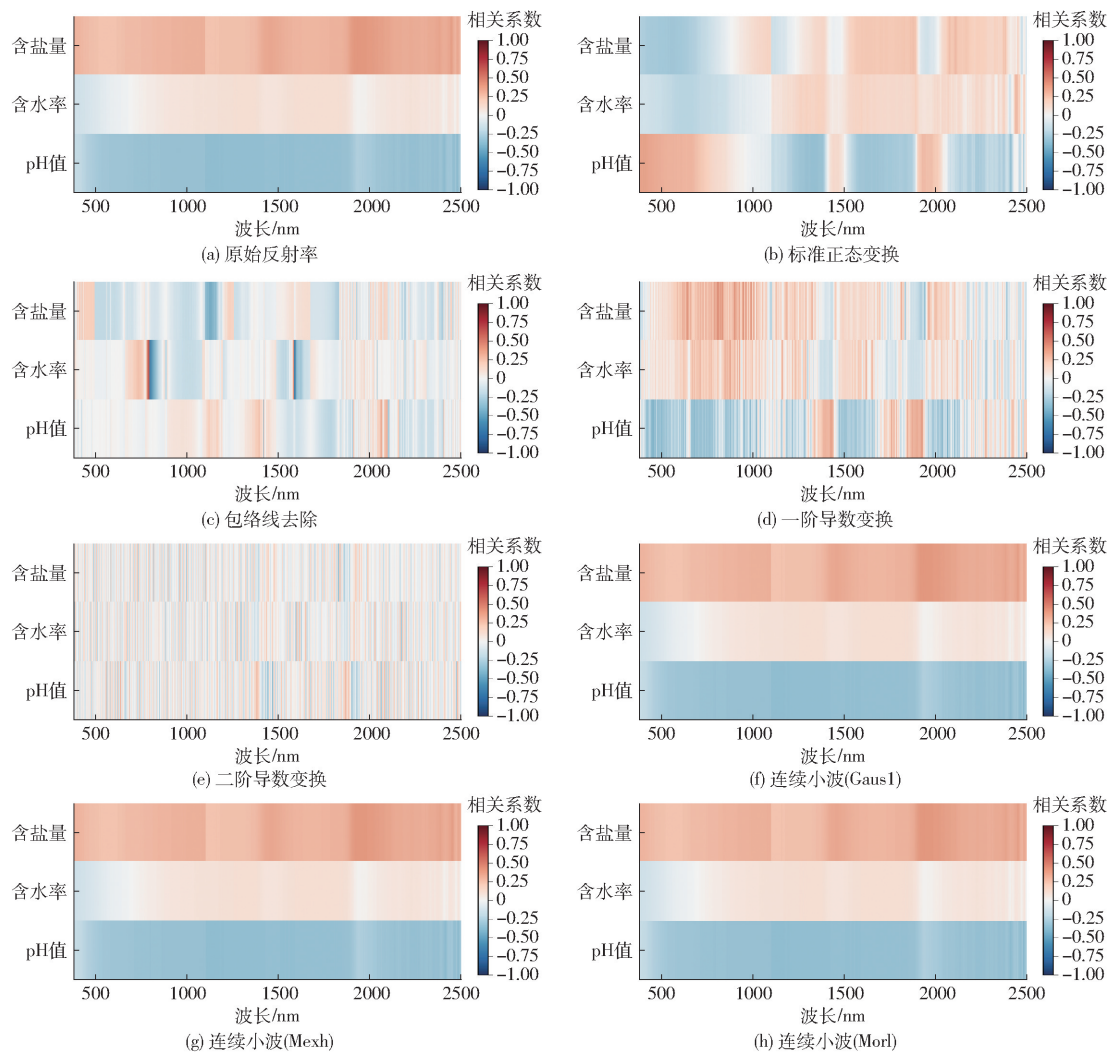


图 8 光谱反射率与土壤水盐含量、pH 值相关系数矩阵 (其他植被覆盖区)

Fig. 8 Correlation coefficient matrix of spectral reflectance with soil moisture, salinity, and pH value (other vegetation)

表 1 不同光谱变换方法效果评估

Tab. 1 Evaluation of different spectral transformation methods				
变换方法	信噪比提升率/%	特征增强度	计算复杂度	综合适用性评分
一阶导数	15 ~ 25	中等	低	4. 2/5. 0
二阶导数	10 ~ 20	高	低	3. 8/5. 0
包络线去除	20 ~ 30	高	中等	4. 5/5. 0
标准正态	25 ~ 35	中等	低	4. 0/5. 0
小波变换	35 ~ 50	高	高	4. 7/5. 0

在陷入局部最优及泛化能力受限的风险。为此,引入 Stacking、Voting (包含简单投票与加权投票) 和 Bagging 3 种集成策略,对优选的基学习器进行级联与重构,性能评估结果如图 11、12 所示。

整体来看,集成策略提升效果呈由大到小依次为 Stacking、Voting、Bagging 趋势。Bagging 平抑了极端值波动,但整体精度提升有限;Voting 通过多模型加权使预测值更加平滑;而 Stacking 展现出最优综合估算能力。

基于 WT – Stacking 模型在农田含盐量反演中取得重大突破 (图 11), 验证集 R^2 跃升至 0. 89, RMSE 降至 0. 61 g/kg, RPD 高达 3. 44。散点图中高值区置信带显著收窄,单线斜率高度贴合 1: 1 基准线,有效克服了单一模型对高盐分样本的“低估”缺陷。土壤含水率的集成估测性能同样优异,农田区 R^2 达到 0. 85, RMSE 为 2. 26%。虽然在高含水率区间由于光谱特征吸收趋于饱和存在极轻微偏移,但整体预测值与实测值保持高度一致。堆叠策略显著提升了弱光谱响应土壤水盐含量、pH 值的预测能力。对于缺乏直接物理吸收特征的 pH 值,WT – Stacking 模型验证集 R^2 从单一模型的 0. 78 显著提升至 0. 82, RMSE 控制在 0. 30。核心聚集区 (pH 值 8. 0 ~ 9. 5) 数据点高度贴合基准线,克服了物理特征缺失的劣势,实现了复杂环境下偏碱性土壤水盐含量、pH 值高精度间接预测。

不同土地利用类型下预测精度存在显著差异,农田、其他植被覆盖区平均 R^2 分别为 0. 88、0. 84。

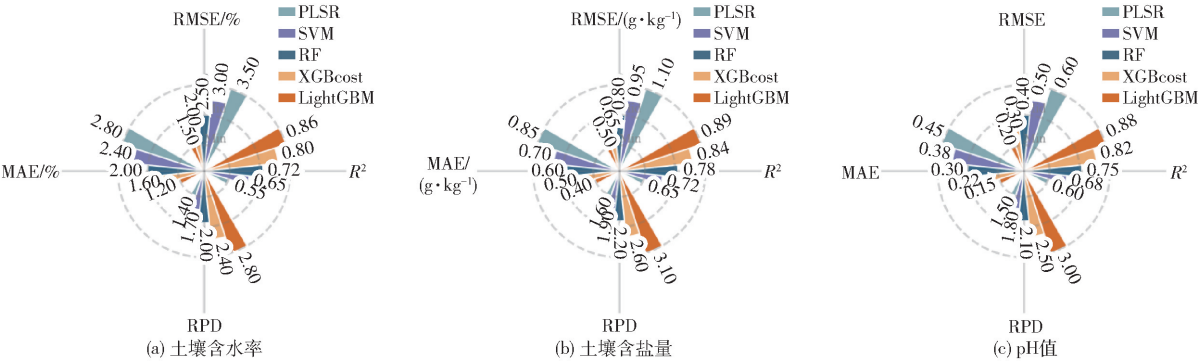


图 9 单一模型土壤水盐含量、pH 值精度(农田)

Fig. 9 Accuracy of soil water – salt content and pH value measurements using single model (farmland)

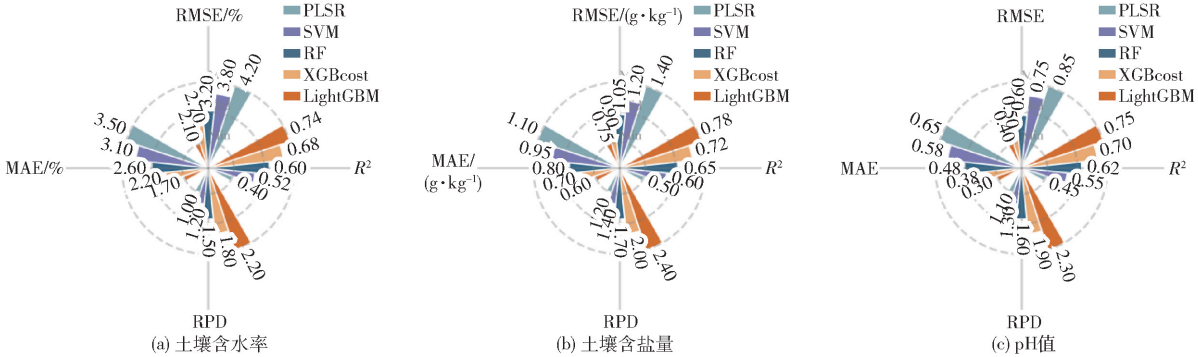


图 10 单一模型土壤水盐含量、pH 值精度(其他植被覆盖区)

Fig. 10 Accuracy of soil water – salt content and pH value measurements using single model (other vegetation)

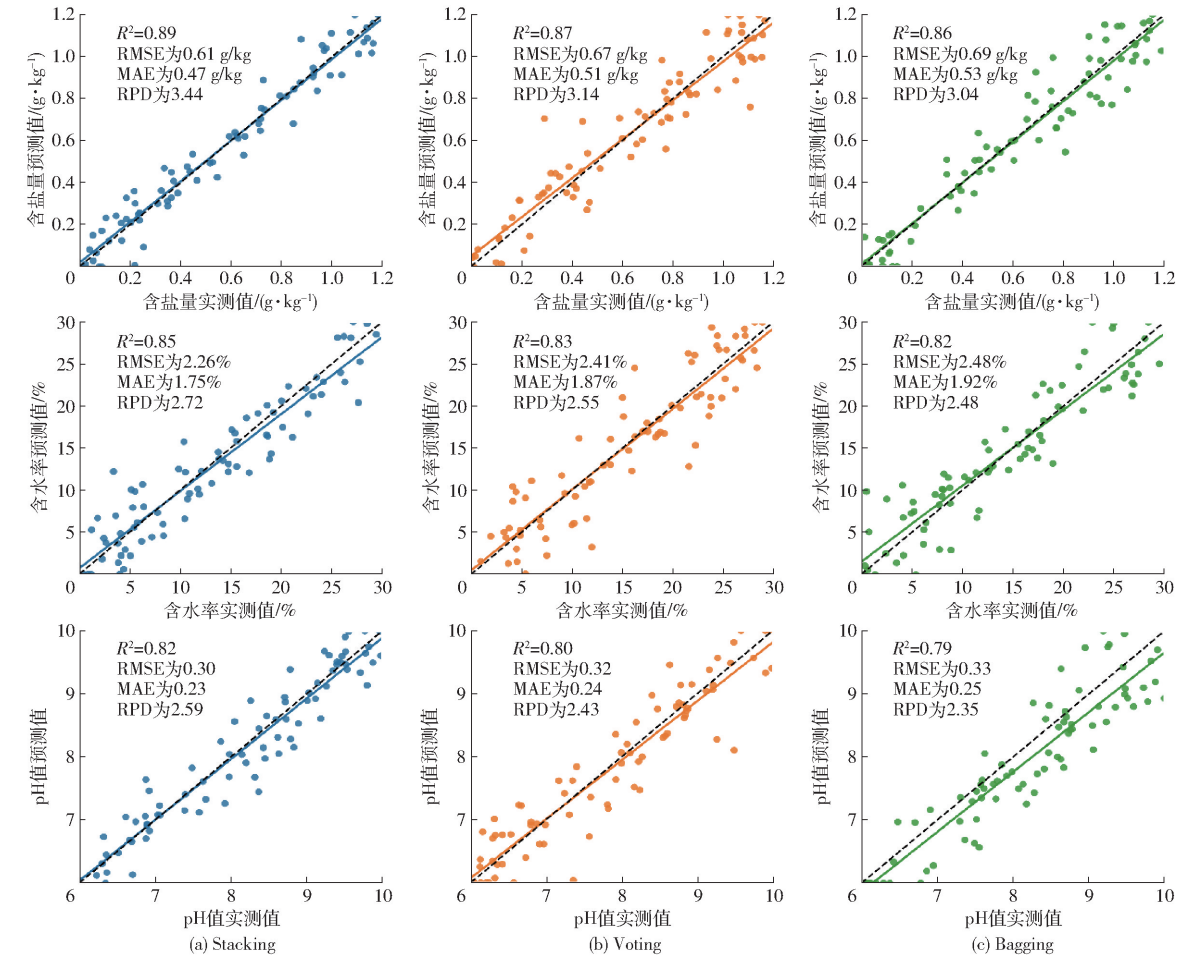


图 11 集成模型土壤水盐含量、pH 值精度(农田)

Fig. 11 Accuracy of integrated model predictions for soil water – salt content and pH values (farmland)

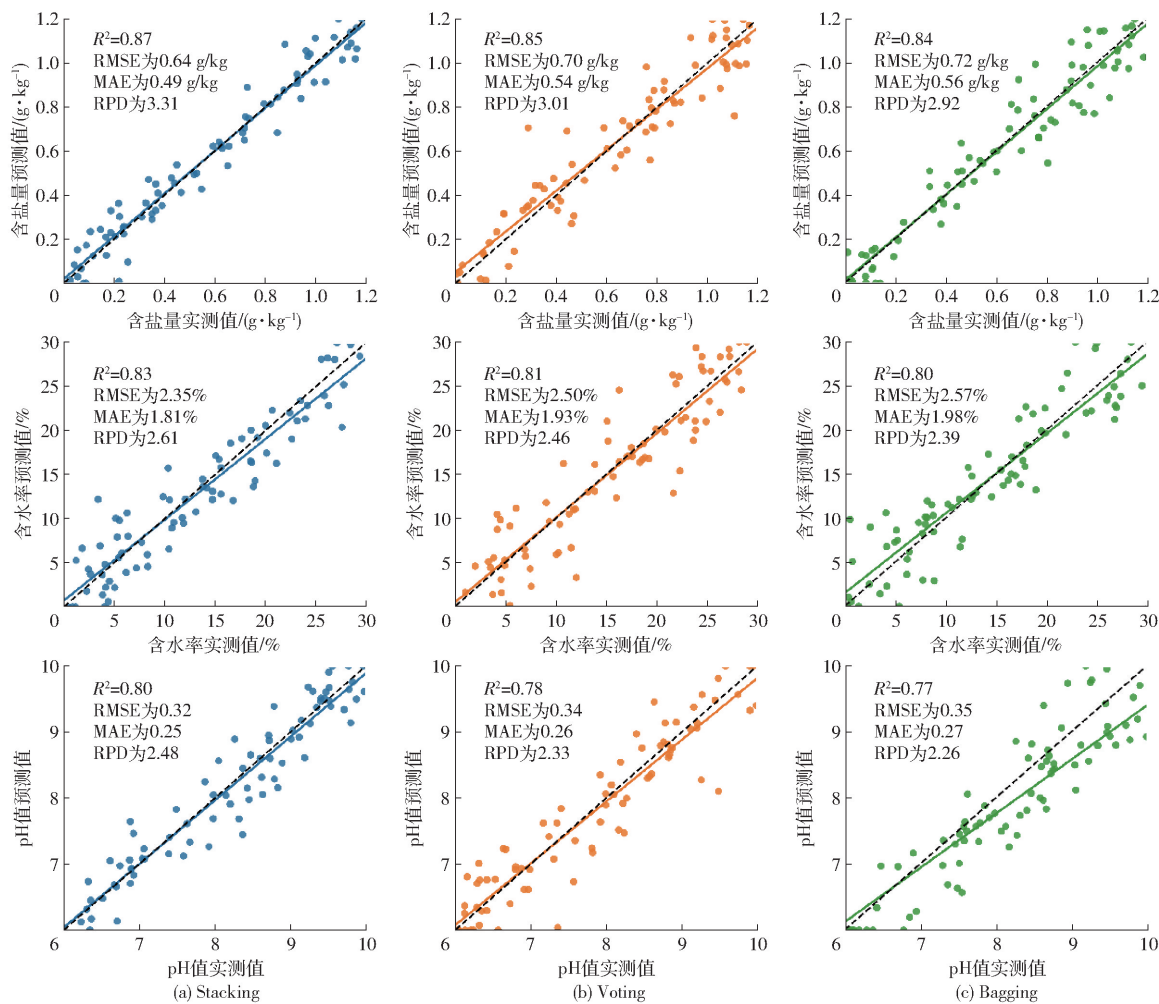


图 12 集成模型土壤水盐含量、pH 值精度(其他植被覆盖区)

Fig. 12 Accuracy of integrated model predictions for soil water and salt content as well as pH value (other vegetation)

3 讨论

3.1 光谱变换效果

野外实测原始光谱受光照、仪器噪声、土壤散射与基线漂移干扰严重,土壤含水率、含盐量、pH 值的光谱响应多为弱隐性特征,直接建模难以挖掘有效关联。一阶导数、二阶导数、包络线去除(CR)、标准正态变换(SNV)等传统方法各有短板,这些传统光谱增强方法虽各有适用场景,但在抑制噪声与保持特征之间往往难以兼顾,存在一定局限性^[41]。连续小波变换(CWT)能在频率域与波长域(空间域)同时对光谱信号进行多分辨率解析,具备优异的时频局部化特性。本研究经系统对比后优选高斯一阶导数小波(Gaus1),信噪比提升率达 35%~50%,在有效滤除环境与仪器噪声的同时,可精准凸显水分、盐分及黏土矿物对应的微弱吸收特征,显著增强了有效光谱信息的可辨识度。从物理光谱学与分子振动理论视角来看,高光谱数据中水分与盐分的特征吸收本质上是分子内部特定化学键(如 O—H、Al—OH)在吸收特定频率电磁波后产生的泛频与合频振

动信号,同土壤背景散射信号相互叠加的结果;而 Gaus1 小波多尺度分解正是利用优异的时频局域化特性,在不同分解尺度下实现了对复杂叠加物理信号的有效解耦与剥离^[42]。在小尺度下,由于小波基函数的支撑区间极窄,频域能量高度集中于高频段,能够与光谱信号中具有高频突变特征的精细吸收结构,比如在 1 400、1 900 nm 处水分子 O—H 键伸缩振动的局部微小尖峰产生强烈共振,同时将仪器记录的高频随机白噪声有效隔离并限制在极小尺度内。随着尺度逐步扩展至中等尺度,小波基有效宽度发生平滑展宽,形态学特征与土壤光谱中主要特征吸收谷的半高宽达到最佳空间几何匹配。在此尺度区间,Gaus1 小波表现出最优的带通滤波特性,有效克服了光谱信噪比衰减,精准捕获了 1 900 nm 处水分子强吸收带的宽化包络形态,以及 2 200 nm 附近因盐渍化伴生硅酸盐黏土矿物(含 Al—OH/Si—OH 键)晶格结构变形而引起的特征吸收谷轮廓^[43]。当分解尺度进一步进入大尺度时,宽幅的小波基函数转为响应光谱信号中的低频成分,即由土壤颗粒粒径分布、表面粗糙度及整体反照率差异引

起的宏观连续统计基线漂移。这种大尺度分解不仅拟合了光谱全局的基线变化趋势,更对中、小尺度提取的局部高频和中频吸收特征起到了平滑去势与消除低频背景混叠的关键作用^[44]。综上,Gaus1 小波多尺度分解并非单纯的数学平滑工具,而是严密契合土壤光谱从微观分子键振动(高频/小尺度)、宏观矿物结构吸收带(中频/中尺度)到整体物理散射背景(低频/大尺度)的物理响应层级,从而为复杂基底环境下水分与盐分特征波段提取提供物理与数学基础。

3.2 模型性能分析

构建的基于连续小波变换(CWT)特征增强的Stacking集成学习模型,在土壤含水率、含盐量及pH值3项指标估算中均表现优异。与单一模型相比,集成学习策略显著提升了预测精度,这一结论与已有研究成果相一致^[45]。集成学习优势主要体现在3方面:①不同基学习器对光谱特征的响应敏感度各异,集成策略能够充分发挥各模型的互补优势。②通过多模型预测结果的融合,有效降低了单一模型过拟合的风险。③集成模型在验证集等未知数据上的泛化能力更为稳健。试验结果显示,相较于最优单一模型,Stacking集成策略使土壤含水率、含盐量和pH值训练集决定系数(R^2)分别提升4.9%、4.7%和5.1%,表现出显著的协同增益效应。此外pH值本身并不具备直接的光谱吸收特征,光谱响应并非源于氢离子浓度对辐射的直接作用,而是依赖于与盐分、水分及伴生黏土矿物之间的内在协变关系^[46]。在干旱半干旱灌区,盐渍化过程的钠离子富集与碳酸盐/重碳酸盐累积,同时驱动土壤含盐量升高和pH值向碱性方向偏移^[47]。换句话说,pH值并非光谱可直接探测的物理量,其变化总是伴随着盐分和水分的变动,模型正是通过捕捉这种伴随关系实现间接估算^[48]。本研究中pH值预测精度(验证集 $R^2=0.82$)系统性地低于含盐量(0.89)和含水率(0.85),正是这种间接光谱响应的客观定量反映。Stacking集成策略之所以能够在pH值分布范围较窄(8.0~9.5)条件下实现相对有效的估算,关键在于通过多个异构基学习器从不同角度捕捉了含水率、含盐量、pH值三者之间的协方差结构,而非简单建立pH值与光谱之间的直接映射关系^[49]。此外,pH值实测值在8.0~9.5之间的窄幅分布确实存在一定过拟合统计隐患。为最大程度降低这一风险,在Stacking集成框架中采用五折交叉验证生成元特征(即基学习器预测输出),有效避免信息泄漏问题;同时,独立验证集(占总样本30%)与训练集严格分离,验证集

$R^2(0.82)$ 与训练集 $R^2(0.84)$ 之间差值仅为0.02,表明模型未出现显著过拟合迹象。尽管如此,若要进一步提升pH值估算模型外部泛化能力,未来亟需引入覆盖更广泛pH值梯度(如酸性或中性土壤样本)独立外部验证数据集,这也是本课题组后续研究的重点方向之一。

3.3 不同土地利用类型对预测精度的影响

不同土地利用类型误差差异,反映了“土壤植被管理”复合生态系统复杂性。本研究中农田土壤预测精度最高,主要由于农田受人为规范管理,空间变异相对较小、均一,与已有的高光谱反演研究结果相一致^[50];在受控或均质化程度较高的农田生态系统中(如对春玉米、水稻等作物的精准监测),高光谱技术往往能发挥出极高的优势预测。

相比之下,自然植被覆盖区预测准确性居中,主要受2方面因素的制约:①植被冠层与地表污染物会掩盖真实的土壤光谱信号,根系分泌物及内陆水体背景等因素会进一步改变光谱反射特征。②其他复杂的野外环境降低了光谱特征稳定性^[51]。例如植物覆盖度会显著影响土壤光谱特征,从而降低预测精度^[52],为理解跨生态系统下保证精度的误差机制提供了重要参考。

基于本研究结果,未来研究可从以下方面深入:结合多时相数据,开展土壤含水率、含盐量及pH值时间序列分析,揭示时间变化规律。将方法扩展到更大的空间尺度,结合卫星遥感数据实现区域尺度的土壤含水率、含盐量及pH值监测。探索深度学习在高光谱土壤含水率、含盐量及pH值估算中的应用,进一步提升模型性能,为环境保护和可持续发展提供有力支撑。

4 结论

(1)小波变换在5种光谱变换方法中表现最优,可同步滤除原始光谱随机噪声、放大土壤水盐矿物弱吸收特征,显著提升光谱信噪比与土壤指标相关性。

(2)Stacking集成策略显著提升了模型性能,相比最佳单一模型,在土壤含水率、含盐量及pH值估计中 R^2 分别提升4.9%、4.7%和5.1%。基于高斯一阶导数小波变换特征增强的Stacking集成(WT-Stacking)学习模型在土壤含水率、含盐量及pH值预测中 R^2 分别达到0.85、0.89和0.82, RMSE分别为2.26%、0.61 g/kg和0.30。

(3)不同土地利用类型下预测精度存在显著差异,农田、其他植被覆盖区平均 R^2 分别为0.88、0.84。

参 考 文 献

- [1] 赵晓阳,胡月明,杨颢,等. 基于多源数据与多模型集成的多种土壤属性制图方法对比[J/OL]. 农业资源与环境学报, 2026-01-28. <https://link.cnki.net/urlid/12.1437.S.20260128.0932.001>.
- [2] 史海滨,杨树青,李瑞平,等. 内蒙古河套灌区水盐运动与盐渍化防治研究展望[J]. 灌溉排水学报,2020,39(8):1-17. Shi Haibin, Yang Shuqing, Li Ruiping, et al. Soil water and salt movement and soil salinization control in Hetao Irrigation District: current state and future prospect[J]. Journal of Irrigation and Drainage,2020,39(8):1-17. (in Chinese)
- [3] 郭文川,马恒,龚城圩,等. 基于可见/近红外光谱的中国北方苹果园土壤有机质含量检测方法[J]. 农业机械学报,2026, 57(9):350-357. Guo Wenchuan, Ma Heng, Gong Chengxu, et al. Methods for detecting soil organic matter content of apple orchards in Northern China based on visible/near-infrared spectra[J]. Transactions of the Chinese Society for Agricultural Machinery,2026,57(9): 350-357. (in Chinese)
- [4] 骆夏杨,李娜,张运林,等. 基于便携式高光谱近感快速监测内陆水体叶绿素 a 浓度[J]. 湖泊科学,2026,38(3):980-995. Luo Xiyang, Li Na, Zhang Yunlin, et al. Rapid monitoring of chlorophyll-a concentration in inland water bodies using a portable proximal sensing technology[J]. Journal of Lake Sciences,2026,38(3):980-995. (in Chinese)
- [5] 邓天民,彭扬洋,赵西崇. 高光谱与气象数据融合的物理引导 XCO₂ 反演方法[J]. 中国环境科学,2026,46(5):2824-2836. Deng Tianming, Peng Yangyang, Zhao Xichong. Physics-guided XCO₂ retrieval method based on the fusion of hyperspectral and meteorological data[J]. China Environmenta Science,2026,46(5):2824-2836. (in Chinese)
- [6] 吴翠玲,李静泰,闫丹丹,等. 基于高光谱数据的互花米草磷素含量反演[J]. 草业科学,2024,41(12):2847-2858. Wu Cuiling, Li Jingtai, Yan Dandan, et al. Phosphate content prediction models of *Spartina alterniflora* based on hyperspectral data[J]. Pratacultural Science,2024,41(12):2847-2858. (in Chinese)
- [7] 杨玲,陈林,吕鹏远,等. 宁夏农田土壤镉含量高光谱反演模型研究[J]. 光谱学与光谱分析,2025,45(增刊1):252-259. Yang Ling, Chen Lin, Lü Pengyuan, et al. Hyperspectral inversion model study for cadmium content in farmland soil in Ningxia [J]. Spectroscopy and Spectral Analysis,2025,45(S1):252-259. (in Chinese)
- [8] 付波霖,黄柯越,杨艳丽,等. 基于实测全谱段高光谱数据的多场景红树林土壤光谱响应特性解析及土壤有机碳含量反演[J]. 测绘学报,2026,55(4):604-617. Fu Bolin, Huang Keyue, Yang Yanli, et al. Multi-scene analysis of mangrove soil spectral response characteristics and inversion of soil organic carbon content based on measured full-spectrum hyperspectral data[J]. Acta Geodaetica et Cartographica Sinica, 2026,55(4):604-617. (in Chinese)
- [9] Melikishvili S, Dizon M, Hianik T. Application of high-resolution ultrasonic spectroscopy for real-time monitoring of trypsin activity in B-casein solution[J]. Food Chemistry,2021,337:127759.
- [10] Eltaweil A, Ayoup S M, Alkahtani M F, et al. Irntegrating ensemble machine learning with mechanistic adsorption analysis for Cr(VI) removal using architecture-engineered LCys-UiO66/h-BN hybrid[J]. Inorganic Chemistry Communications,2026, 189(P2):116786.
- [11] 李笑芳,王金杲,霍建鸿,等. 利用离散小波耦合算法的土壤有机质含量高光谱反演[J]. 光谱学与光谱分析,2025, 45(12):3488-3497. Li Xiaofang, Wang Jingao, Huo Jianhong, et al. Hyperspectral inversion of soil organic matter content using a discrete wavelet coupling algorithm[J]. Spectroscopy and Spectral Analysis,2025,45(12):3488-3497. (in Chinese)
- [12] 魏薇,王丹,王博韬,等. 融合高光谱图谱特征的油菜角果含水率反演研究[J]. 光谱学与光谱分析,2025,45(10):2863-2874. Wei Wei, Wang Dan, Wang Botao, et al. Research on the inversion of moisture content in rapeseed silique peel based on hyperspectral fusion imaging[J]. Spectroscopy and Spectral Analysis,2025,45(10):2863-2874. (in Chinese)
- [13] 于雷,洪永胜,耿雷,等. 基于偏最小二乘回归的土壤有机质含量高光谱估算[J]. 农业工程学报,2015,31(14):103-109. Yu Lei, Hong Yongsheng, Geng Lei, et al. Hyperspectral estimation of soil organic matter content based on partial least squares regression[J]. Transactions of the CSAE,2015,31(14):103-109. (in Chinese)
- [14] 汪杰,孙孝林,伍丹华,等. 高密度植被覆盖区基于无人机高光谱数据的农田土壤有机质反演[J]. 光谱学与光谱分析, 2025,45(10):2885-2896. Wang Jie, Sun Xiaolin, Wu Danhua, et al. Prediction of soil organic matter for farmlands covered with high density of vegetation based on UAV hyperspectral data[J]. Spectroscopy and Spectral Analysis, 2025, 45(10):2885-2896. (in Chinese)
- [15] 毛亚纯,夏安妮,曹旺,等. 基于高光谱数据和 Stacking 集成学习算法的金矿品位快速反演[J]. 光谱学与光谱分析, 2025,45(7):2061-2067. Mao Yachun, Xia Anni, Cao Wang, et al. Rapid inversion of gold ore grades based on hyperspectral data and Stacking ensemble learning algorithm[J]. Spectroscopy and Spectral Analysis,2025,45(7):2061-2067. (in Chinese)
- [16] 郭辉,韩紫薇,吴斗庆. 基于光谱指数的采煤拉张裂隙区土壤有机质含量估测[J]. 光谱学与光谱分析,2025,45(9): 2569-2577.

Guo Hui, Han Ziwei, Wu Douqing. Estimation of soil organic matter content in coal mining tensile fracture area based on spectral indices[J]. Spectroscopy and Spectral Analysis,2025,45(9):2569–2577. (in Chinese)

[17] 古甜甜,刘严松,胡薇,等. 基于机器学习的铅锌铜矿区土壤重金属含量高光谱反演分析[J]. 环境科学,2026,47(6):4211–4222.

Gu Tiantian, Liu Yansong, Hu Wei, et al. Hyperspectral inversion analysis of soil heavy metal content in lead-zinc-copper mining area based on machine learning[J]. Environmental Science,2026,47(6):4211–4222. (in Chinese)

[18] 王涵颖,张美玲,巩彦君,等. 基于集成机器学习的祁连山草地全球变暖潜力预测[J/OL]. 环境科学,2025–11–05. <https://link.cnki.net/urlid/11.1895.x.20251104.1548.032>.

[19] 丁朋朋,王苏桐,郭政,等. 基于 Stacking 集成机器学习的银川平原地下水位动态预测[J/OL]. 灌溉排水学报,2026–05–09. <https://link.cnki.net/urlid/41.1337.S.20260509.1050.001>.

[20] 汪立,樊后宝,张晏维,等. 基于可解释集成机器学习的土壤表层 pH 值空间分布预测及影响因素[J]. 环境科学,2026,47(3):1941–1953.

Wang Li, Fan Houbao, Zhang Yanwei, et al. Spatial distribution prediction and influencing factors of soil surface pH based on interpretable ensemble machine learning[J]. Environmental Science,2026,47(3):1941–1953. (in Chinese)

[21] Kevin S, Kalifa G, Mickaël G. A semi-empirical approach to quantify and handle the effect of moisture on spectral unmixing [J]. International Journal of Applied Earth Observations and Geoinformation,2021,96:102259.

[22] Li Kuangyu, Yan Weijie, Huang Xianliang, et al. In-situ measurement of temperature and potassium concentration during the combustion of biomass pellets based on the emission spectrum[J]. Fuel,2021,289:119863.

[23] 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. RF hyperspectral inversion of oasis plough layer soil salinity based on optimal mathematics and wavelet transform[J]. Transactions of the CSAE,2025,41(3):83–93. (in Chinese)

[24] 朱则东,张清河,刘含. 堆栈集成学习联合多源数据反演草地土壤水分[J]. 遥感信息,2024,39(3):104–112.

Zhu Zedong, Zhang Qinghe, Liu Han. Inversion of soil moisture in grassland based on Stacking ensemble learning with multi-source data[J]. Remote Sensing Information,2024,39(3):104–112. (in Chinese)

[25] 甘甜,李雷,李红叶,等. 基于多源遥感数据和机器学习算法的冬小麦产量预测研究[J]. 麦类作物学报,2022,42(11):1419–1428.

Gan Tian, Li Lei, Li Hongye, et al. Winter wheat yield prediction based on multi-source remote sensing data and machine learning algorithms[J]. Journal of Triticeae Crops,2022,42(11):1419–1428. (in Chinese)

[26] 刘凯,王玉峰,彭志青,等. 基于机器学习的高光谱土壤养分特征波段提取方法[J]. 光学学报,2025,45(6):230–239.

Liu Kai, Wang Yufeng, Peng Zhiqing, et al. Machine learning for hyperspectral characteristic band extraction in soil nutrient analysis[J]. Acta Optica Sinica,2025,45(6):230–239. (in Chinese)

[27] 魏雪馨,刘洋,闵庆文,等. 基于地物光谱季节曲线特征的毛竹林分布提取[J]. 遥感技术与应用,2021,36(5):1178–1188.

Wei Xuexin, Liu Yang, Min Qingwen, et al. Extraction of moso bamboo forest distribution based on characteristics of vegetation spectral seasonal curves[J]. Remote Sensing Technology and Application,2021,36(5):1178–1188. (in Chinese)

[28] 王薪宇. 基于高光谱数据的土壤速氮含量反演模型比较[J]. 测绘与空间地理信息,2023,46(3):201–204.

Wang Xinyu. Comparison on inversion models of soil rapid nitrogen content based on hyperspectral data[J]. Geomatics & Spatial Information Technology,2023,46(3):201–204. (in Chinese)

[29] Zhang Chunxiao, Shen Chao, Wei Shen, et al. Flexible management of heat/electricity of novel PV/T systems with spectrum regulation by Ag nanofluids[J]. Energy,2021,221:119903.

[30] Mohammed N H, Arafa B, Rana M, et al. Spectral characterization, DFT, docking and cytotoxicity of N-benzyl-4,5-dihydro-3-methyl-5-oxo-1H-pyrazole-4-carbothioamide and its metal complexes[J]. Journal of Molecular Structure,2021,1232:130020.

[31] Li Jiaqi, Xia Xinlin, Yang Youwei, et al. Pore scale simulation of spectral radiative properties of Voronoi open-cell carbon foams at high temperatures[J]. Infrared Physics and Technology,2021,114:103655.

[32] 贺洋,徐韬,宋云. ASD FieldSpec 3 地物光谱数据在水磨沟-大河地区岩性分类中的应用[J]. 地质力学学报,2015,21(1):21–29.

He Yang, Xu Tao, Song Yun. Application of ASD FieldSpec3 ground object spectral data in lithology classification of Shuimo–Dahe area[J]. Journal of Geomechanics,2015,21(1):21–29. (in Chinese)

[33] 洪琴,胡清华,陈文惠,等. 基于高光谱数据藻类水华重要参数:叶绿素 a 的定量反演模型研究[J]. 环境生态学,2023,5(2):23–31.

Hong Qin, Hu Qinghua, Chen Wenhui, et al. Quantitative inversion model of chlorophyll a, an important parameter of algal blooms, based on hyperspectral data[J]. Environmental Ecology,2023,5(2):23–31. (in Chinese)

[34] 热依拉·艾合买提,吾木提·艾山江,阿不都艾尼·阿不里,等. 基于机器学习的春小麦叶片水分含量高光谱估算[J]. 麦类作物学报,2022,42(5):640–648.

Rahila Ahmat, Umut Hasan, Abugheni Abliz, et al. Hyperspectral estimation of spring wheat leaf water content based on machine learning[J]. Journal of Triticeae Crops,2022,42(5):640–648. (in Chinese)

[35] Abakar R, Isah M, Shafi I Q, et al. Data-driven rational design of functionalized COFs for PFAS remediation via DFT and

- machine learning integration[J]. Microporous and Mesoporous Materials, 2026, 410: 114213.
- [36] Haemers H W. Cospectral pairs of regular graphs with different connectivity[J]. Discussiones Mathematicae Graph Theory, 2020, 40(2): 577–584.
- [37] Zhang H, Chen A, Huang L, et al. Transcriptomic analysis unravels the modulating mechanisms of the biomass and value-added bioproducts accumulation by light spectrum in *Eustigmatos* cf. *Polyphem* (Eustigmatophyceae) [J]. Bioresource Technology, 2021, 338: 125523.
- [38] Tsogtbilegt B, Minjung L, Honghyun C. Enhancing the solar energy conversion and harvesting characteristics of multiwalled carbon nanotubes-modified 1-hexyl-3-methyl-imidazolium cation ionic liquids[J]. International Journal of Energy Research, 2022, 46(7): 8891–8907.
- [39] Rajiv D, Wu Yupeng. Review of methods to account for the solar spectral influence on photovoltaic device performance[J]. Energy, 2024, 286: 129461.
- [40] Zhang Y, Chai X, Hu J, et al. Enhancing rapeseed biomass and yield estimation with ensemble learning and synergistic multidimensional features. [J]. Journal of Zhejiang University (Science B), 2026, 27(5): 499–516.
- [41] 周夕, 夏鸿斌, 王晓锋. 基于小波变换增强引导的扩散序列推荐模型[J]. 模式识别与人工智能, 2025, 38(10): 893–910. Zhou Xi, Xia Hongbin, Wang Xiaofeng, et al. Wavelet-enhanced guidance for diffusion-based sequential recommendation[J]. Pattern Recognition and Artificial Intelligence, 2025, 38(10): 893–910. (in Chinese)
- [42] 辜姣, 陈国庆, 张笑河, 等. 基于小波分解和因子分析的白酒香型和年份鉴定的研究[J]. 光谱学与光谱分析, 2018, 38(8): 2511–2515. Gu Jiao, Chen Guoqing, Zhang Xiaohe, et al. Classification and year prediction of Chinese liquors based on wavelet decomposition and factor analysis[J]. Spectroscopy and Spectral Analysis, 2018, 38(8): 2511–2515. (in Chinese)
- [43] 周锦程, 杨清亮, 张伟. 基于高斯小波变换的测井曲线自动分层模型[J]. 黔南民族师范学院学报, 2017, 37(4): 31–39, 48. Zhou Jincheng, Yang Qingliang, Zhang Wei. Automatic stratification model of well logging curves based on Gaussian wavelet transform[J]. Journal of Qiannan Normal University for Nationalities, 2017, 37(4): 31–39, 48. (in Chinese)
- [44] 陈中杰, 蔡勇, 蒋刚. 复高斯小波核函数的支持向量机研究[J]. 计算机应用研究, 2012, 29(9): 3263–3265. Chen Zhongjie, Cai Yong, Jiang Gang. Study on SVM of complex Gaussian wavelet kernel function[J]. Application Research of Computers, 2012, 29(9): 3263–3266. (in Chinese)
- [45] 王鹏新, 王静怡, 郭丰玮, 等. 基于遥感多参数和 Stacking 集成学习的冬小麦单产估测[J]. 农业机械学报, 2025, 56(11): 369–377. Wang Pengxin, Wang Jingyi, Guo Fengwei, et al. Yield estimation of winter wheat based on multiple remotely sensed parameters and Stacking ensemble learning[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(11): 369–377. (in Chinese)
- [46] 张恒, 梁太波, 宋效东, 等. 基于高光谱成像的烟田土壤 pH 估测[J]. 西南农业学报, 2023, 36(12): 2771–2779. Zhang Heng, Liang Taiibo, Song Xiaodong, et al. Estimation of soil pH in tobacco field based on hyperspectral imaging[J]. Southwest China Journal of Agricultural Sciences, 2023, 36(12): 2771–2779. (in Chinese)
- [47] 崔瑾, 刘凯, 冯欣玉, 等. 基于机器学习与 SHAP 算法的胶东半岛土壤酸化影响因素分析[J]. 土壤与作物, 2026, 15(2): 191–201. Cui Jin, Liu Kai, Feng Xinyu, et al. Analysis of factors influencing soil acidification in the Jiaodong Peninsula based on machine learning and the SHAP algorithm[J]. Soils and Crops, 2026, 15(2): 191–201. (in Chinese)
- [48] 胡亚男, 高小红, 申振宇, 等. 基于野外实测 Vis-NIR 光谱的土壤肥力估算研究: 以湟水流域为例[J]. 土壤通报, 2021, 52(3): 575–584. Hu Ya'nan, Gao Xiaohong, Shen Zhenyu, et al. Estimating fertility index by using field-measured Vis-NIR spectroscopy in the Huanghui River Basin[J]. Chinese Journal of Soil Science, 2021, 52(3): 575–584. (in Chinese)
- [49] Liu X, Pang H, Kou X, et al. Synergistic interaction of ternary Pd-Cu-Ni confined in nanoparticles as pH-universal catalysts for enhanced hydrogen evolution reaction[J]. Chinese Chemical Letters, 2026, 37(4): 367–372.
- [50] Long T, Wang H. Super-hybrid transition metal sulfide nanoarrays of NiS nanoparticle/WS₂ nanosheet/Ni₃S₄ nanoparticle with abundant plane and edge-type active interfaces for robust all-pH hydrogen evolution [J]. Chinese Chemical Letters, 2026, 37(3): 783–788.
- [51] 刘彦随, 金淑贞, 王伟, 等. 基于多源数据城市土地利用类型精细化识别与功能演替分析[J]. 人文地理, 2026, 41(1): 1–11, 33. Liu Yansui, Jin Shuzhen, Wang Wei, et al. Refined identification and functional succession analysis of urban land use types based on multi-source data[J]. Human Geography, 2026, 41(1): 1–11, 33. (in Chinese)
- [52] 苏鑫森, 祁迷, 滑永春, 等. 科尔沁沙地不同土地利用变化的生态系统服务功能[J/OL]. 水土保持学报, 2026-05-22. <https://link.cnki.net/urlid/61.1362.TV.20260521.1325.003>.