

基于智能学习的光滑土壤表面光谱反射率模型

马静文^{1,2}, 李祥东³, 郑兴明^{1,2}

¹吉林建筑大学测绘与勘查工程学院, 吉林 长春

²中国科学院东北地理与农业生态研究所, 吉林 长春

³吉林农业大学信息技术学院, 吉林 长春

收稿日期: 2026年4月7日; 录用日期: 2026年5月14日; 发布日期: 2026年7月21日

摘要

土壤光谱反射率是遥感反演的关键参数。本研究采集东北地区700个干土样品的实验室光谱数据, 基于SHAP方法分析土壤理化性质及成土母质等协变量的贡献, 构建了基于梯度提升回归(GBR)的干土光谱反射率模型(400~2500 nm), 并结合比尔-朗伯定律预测湿土反射率。结果表明: 1) 干土模型预测精度较高($R^2 = 0.93$, RMSE = 0.018), 引入成土母质使精度提升了4.5%; 2) 模型预测的反射率空间分布与卫星观测结果具有良好一致性; 3) 湿土反射率模拟精度在8%、16%、20%、24%和30%五个质量含水量梯度下, R^2 分别达到0.78、0.85、0.85、0.89和0.85。综上, 本研究提出了一种高效且具有良好可解释性的土壤光谱反射率建模方法, 可为土壤反射率预测及基于遥感的土壤性质反演提供可靠支撑。

关键词

土壤光谱, 机器学习, SHAP, 比尔-朗伯定律

An Intelligent Learning-Based Model of Spectral Reflectance for Smooth Soil Surfaces

Jingwen Ma^{1,2}, Xiangdong Li³, Xingming Zheng^{1,2}

¹School of Geomatics and Prospecting Engineering, Jilin Jianzhu University, Changchun Jilin

²Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun Jilin

³College of Information Technology, Jilin Agricultural University, Changchun Jilin

Received: April 7, 2026; accepted: May 14, 2026; published: July 21, 2026

Abstract

Soil spectral reflectance is a critical parameter for the remote sensing retrieval of land surface

文章引用: 马静文, 李祥东, 郑兴明. 基于智能学习的光滑土壤表面光谱反射率模型[J]. 土壤科学, 2026, 14(3): 168-180. DOI: 10.12677/hjss.2026.143017

properties. In this study, laboratory spectral data were collected from 700 dry soil samples across Northeast China. The contributions of environmental covariates, including soil physicochemical properties and parent material, were analyzed using the SHAP method. Subsequently, a dry soil spectral reflectance model spanning 400~2500 nm was developed based on Gradient Boosting Regression (GBR), and wet soil reflectance was predicted by coupling this model with the Beer-Lambert law. The results indicate that: 1) The dry soil model exhibited high predictive accuracy ($R^2 = 0.93$, RMSE = 0.018), and the incorporation of parent material improved the model's accuracy by 4.5%. 2) The spatial distribution of the model-simulated reflectance demonstrated good consistency with satellite observations. 3) Under five gravimetric water content gradients (8%, 16%, 20%, 24%, and 30%), the simulation accuracy (R^2) for wet soil reflectance reached 0.78, 0.85, 0.85, 0.89, and 0.85, respectively. In conclusion, this study proposes an efficient and highly interpretable modeling methodology for soil spectral reflectance, providing reliable support for reflectance prediction and the remote sensing-based retrieval of soil properties.

Keywords

Soil Spectra, Machine Learning, SHAP, Beer-Lambert Law

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

土壤光谱反射率表征土壤表面对入射辐射的响应能力,是揭示土壤组成及表面性质的重要参数[1]。其不仅反映矿物组成、质地结构、有机质含量及盐分状况等信息,还为土壤理化性质反演、土地退化评估及盐渍化监测提供关键依据[2]。因此,构建高精度土壤反射率模型具有重要意义。

土壤反射率受多因素耦合控制,主要包括环境因素、矿物组成以及土壤物理和化学性质[3]。其中,成土母质决定土壤物质来源及演化过程,对光谱特征具有基础性影响,尤其在短波红外(SWIR)波段表现显著[4]。观测几何、地形及光照等条件虽不改变土壤本征属性,但通过影响光与地表的相互作用,导致光谱响应差异[5]。矿物组成主导近红外(NIR)和 SWIR 吸收特征,而粗糙度和水分通过改变散射与吸收过程影响反射率水平[6];土壤质地则通过调节结构与孔隙间接影响可见光(VIS)波段反射特性[7]。此外,有机质、pH 值及电导率等化学性质也对光谱响应具有重要调控作用[8]。

现有土壤反射率建模方法主要包括经验模型、物理模型和智能学习模型三类。经验模型(如 Price [9]、GSV [10])形式简单但物理解释能力有限;物理模型(如 Hapke 模型[11]-[14]、MARIT 模型[15]及 BSM 模型[16])基于辐射传输理论,具有较强机理性但参数复杂;智能学习模型 SOGM 模型则通过数据驱动方式刻画非线性关系,具备较强的拟合能力[17]。总体来看,土壤反射率建模正由经验方法向机理与数据驱动融合方向发展,实现土壤反射率的高精度模拟已成为土壤光谱研究的重要基础。

2. 数据与方法

2.1. 研究区

研究区位于东北地区(117°46'~130°67'E, 40°94'~49°25'N),属温带季风型大陆性气候,年平均气温 0℃~10℃,年降水量约 400~1000 mm,夏季降水占全年 60%~80%。土壤类型丰富,包括黑土、草甸土、白浆土、暗棕壤等;土壤母质多样,有深浅色结晶岩风化物、砂页岩、石灰岩风化物、洪积物、冲积物、湖

积物及粘质黄土等, 为研究提供了充分的基础。

2.2. 土壤数据

2.2.1. 土壤样品采集与处理

本研究采用五点法采集 0~5 cm 表层土壤, 子样混合后去杂质, 共计 700 份土壤样品, 将样品烘干 (85℃) 并筛分 (2 mm)。样品分两部分, 用于理化性质测定 (pH、电位法; 有机质, 重铬酸钾外加热氧化法; 粒径组成, 比重计法) 和干土光谱测量, 其分布特征见图 1。

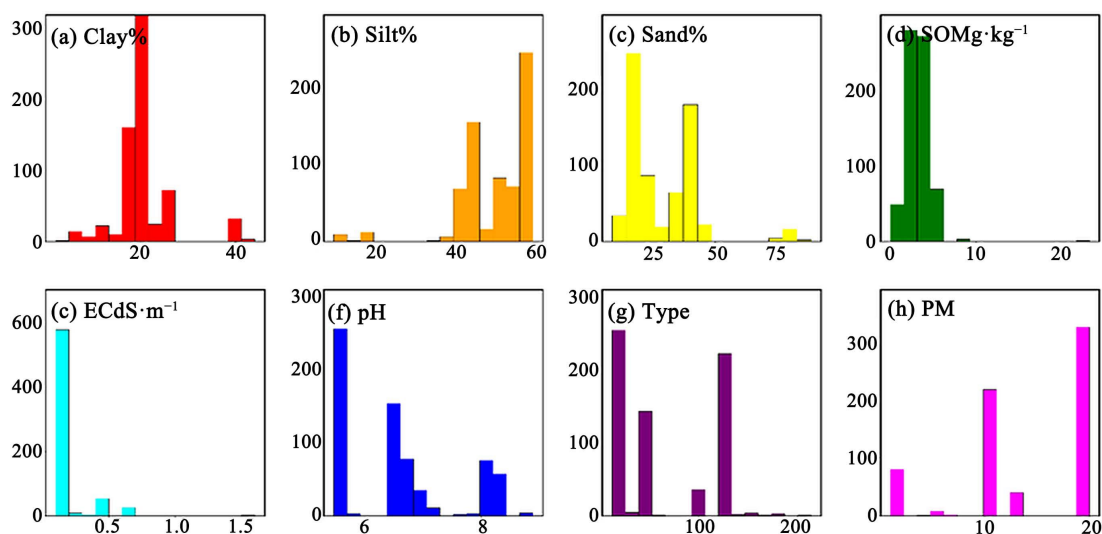


Figure 1. Distribution of soil property ranges
图 1. 属性范围分布图

研究区土壤中黏粒、粉粒和砂粒的平均含量分别为 20.93%、50.40% 和 28.67%, 其变化范围分别为 2.17%~43.98%、9.01%~59.46% 和 8.30%~88.82%。有机质 (SOM)、电导率 (EC) 及 pH 值的平均值分别为 3.40 g·kg⁻¹、0.15 dS·m⁻¹ 和 6.52, 其对应范围分别为 0.12~23.08 g·kg⁻¹、0.01~1.60 dS·m⁻¹ 和 5.46~8.89。

2.2.2. 暗室土壤光谱测量

光谱测量采用美国 ASD 公司 FieldSpec®4 地物光谱仪在暗室条件下进行, 波段范围覆盖 350~2500 nm [18]。干土样品经 105℃ 烘干并过 2 mm 筛后直接测量, 以表征土壤固有光谱特性; 湿土样品在干土基础上按质量含水量法分别配制 8%、16%、20%、24% 和 30% 五个梯度, 加水后以保鲜膜密封并于室温静置 48 h, 待水分充分均衡后测量。测量时, 样品置于直径 60 mm、深度 10 mm 的培养皿中, 光纤探头垂直固定于样品表面上方 3 cm 处, 视场完全覆盖样品以规避边缘效应。仪器预热 30 min 后以标准白板定标, 测量过程中每 30 min 复校一次以抑制信号漂移。每份样品连续测量 10 次 (间隔 1 s), 取均值作为最终光谱反射率 (图 2)。

2.2.3. Sentinel-2 卫星影像数据

本研究基于 Google Earth Engine 平台的 Sentinel-2 S2_HARMONIZED 地表反射率数据, 合成东北地区裸土最大反射率, 用于与模型模拟干土光谱进行对比验证 [19]。该数据集经辐射、几何及大气校正处理后可直接提供表面反射率, 覆盖可见光、近红外及短波红外波段。其中, B10 波段 (1375 nm) 主要用于卷云检测, 不纳入地表反射率分析。



Figure 2. Laboratory spectral measurements of dry and wet soils
图 2. 干土与湿土实验室光谱测量

为获取裸土信息并降低植被影响,选取 2015~2025 年春耕前后(4 月 1 日至 5 月 20 日)期间 Sentinel-2 影像。该时段农田多处于播种前或作物尚未出苗阶段,裸土暴露充分,有利于提取真实土壤反射率。影像筛选过程中,剔除云量大于 10%的数据,并结合场景分类层(Scene Classification Layer, SCL)去除不透明云、卷云及阴影像元,以确保像元反射率尽可能不受云层干扰[20]。

在时序反射率合成中,为减少残余云雾、大气扰动、秸秆覆盖及仪器观测不确定性对结果的影响,采用第 90 百分位数(90th percentile)统计方法。该方法提取时间序列反射率的高值部分,有效剔除异常低值干扰,同时能较好反映裸土在最佳观测条件下的反射率特征[21]。表示为 R_{p90} :

$$R_{p90}(x, b) = \text{percentile}\left(\left\{R_t(x, b)\right\}_{t=1}^T, 90\right) \quad (2.1)$$

其中, $R_t(x, b)$ 表示像元 x 在波段 b 上第 t 时刻的反射率, T 为研究时段内该像元的有效影像总数。将反射率合成结果统一重采样至 1000 m 空间分辨率。在重采样过程中,仅保留掩膜后的有效像元,重采样过程中仅保留有效像元,空值区域(云掩膜及非裸土像元)不作插值填补。为去除非耕地和植被的影响,基于 ESA WorldCover 2020 数据集的耕地范围和 NDVI 阈值($\text{NDVI} < 0.2$)进行掩膜。

2.3. 土壤反射率建模方法

干土反射率模拟采用梯度提升回归(Gradient Boosting Regression, GBR)构建环境与土壤因子驱动的干土光谱反射率模型。覆盖 350~2500 nm 波段,每隔 10 nm 构建独立子模型,共 215 个单波长模型。700 份样本按 7:3 划分为训练集(490 份)与验证集(210 份)。为提高泛化能力并防止过拟合,训练集采用 10 折交叉验证结合 GridSearchCV 优化超参数,搜索空间包括:决策树数量($n_estimators$) {900, 1100, 1300}、学习率($learning_rate$) {0.01, 0.05, 0.1}、最大树深度(max_depth) {5, 6, 7}、最小分裂样本数($min_samples_split$) {2, 5, 10} 及最小叶节点样本数($min_samples_leaf$) {1, 2, 4}。最终基于 Sentinel-2 中心波长选择交叉验证最佳参数构建子模型,并以决定系数(R^2)和均方根误差(RMSE)评估预测精度与稳定性。

湿土反射率基于比尔-朗伯定律构建,实现不同含水量条件下土壤光谱反射率的定量模拟[22]。在模型构建中,以干土光谱反射率作为基准特征,结合等效水厚度表征土壤水分对反射率的调控作用。通过耦合干土反射率与等效水厚度,可建立土壤含水量变化对各波段反射率的定量关系,从而精确模拟湿土光谱。

根据 William (2010),土壤光谱反射率随含水量变化并非单调递减:当土壤水分低于临界值(田间持水量)时,传感器接收的主要是土壤散射与吸收后的反射光;高于临界值时,反射逐渐受土壤颗粒表面水膜的镜面反射影响,甚至出现镜面反射现象。当土壤含水量明显低于田间持水量时,湿土反射率可近似服从比尔-朗伯定律(Beer-Lambert law):

$$r_w = r_0 * \exp(-\alpha \cdot \zeta) \quad (2.2)$$

式中, α 为土壤水分吸收系数; ζ 为土壤水分的等效水厚度; r_w 为湿土双半球反射率; r_0 为干土反射率。其中, 水吸收系数 α 的取值是参考邓孺孺水吸收系数的测量方法, 该方法采用控制水层厚度的方法, 测量了范围为 400~2500 nm 的水吸收系数固定值[23] [24]。水吸收系数 α 表示为:

$$\alpha = k(\lambda) - b_w(\lambda) - K_s(\lambda) \quad (2.3)$$

式中, $b_w = 0.0001451\lambda - 4.32$ 为水分子散射系数, $K_s(\lambda) = 0.07581\lambda - 0.2718$ 为水中悬浮粒子的消光系数, 其光学厚度和消光系数可以通过 Ångstrom 公式拟合得到。最终得到的 400~2400 nm 波段的水吸收系数测量结果如图 3 所示。

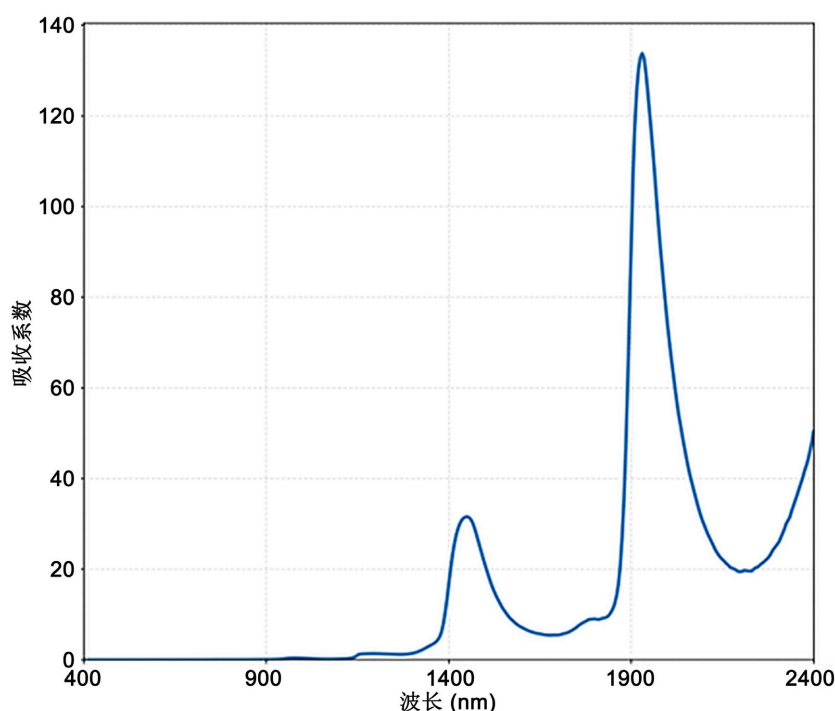


Figure 3. Water absorption coefficients at different spectral bands

图 3. 不同波段水吸收系数

3. 结果

3.1. 干土反射率模型特征选择

在构建干土反射率预测模型前, 首先通过 Pearson 相关性分析和 SHAP 特征重要性评估对候选特征进行筛选, 以确保输入参数选取的合理性。

图 5 展示了 Sentinel-2 所有波段下的干土反射率预测模型的 SHAP 变量重要性分布。图中采用环形图形式, 内环表示可见光波段(443 nm)下各变量的重要性, 外环表示短波红外波段(2190 nm)下的变量贡献。颜色代表不同类别的土壤属性, 包括地理位置(Lon、Lat)、颗粒组成(Clay、Silt、Sand)、土壤有机质(SOM)、电导率(EC)、酸碱度(pH)、母质(PM)和土壤类型(Type), 扇形的角度大小表示对应变量在该波段下对模型输出的平均贡献程该环形图直观呈现了各特征变量在不同波段下的重要性排序。明显看出, 不同波段对变量响应存在显著差异。

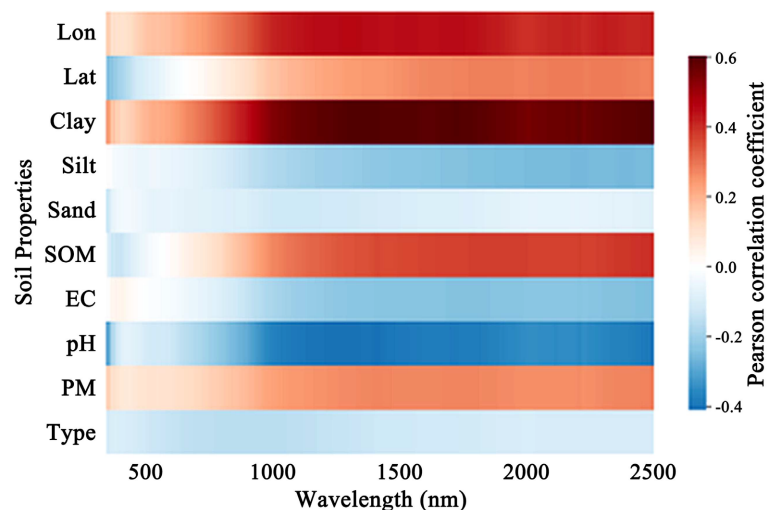


Figure 4. Pearson correlation between soil properties and dry-soil reflectance in the 350~2500 nm spectral range

图 4. 土壤属性与 350 nm~2500 nm 波段干土反射率的皮尔逊相关性

图 4 是土壤属性与 350~2500 nm 波段干土反射率的皮尔逊相关性结果, 整体来看, Clay、Lon 和 PM 在 350~2500 nm 波段与反射率呈正相关, Silt、Sand、pH 和 Type 整体呈负相关, SOM 和 Lat 在可见光 (VIS) 负相关, 在 NIR 和 SWIR 呈弱正相关, EC 在 VIS 正相关, 在 NIR 和 SWIR 弱负相关。

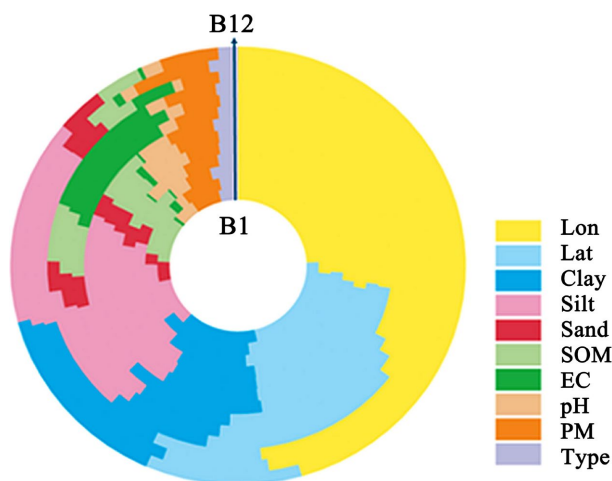


Figure 5. SHAP-Based feature importance of soil properties influencing dry-soil reflectance

图 5. 基于 SHAP 的土壤属性对干土反射率影响的重要性

结果表明, 所有变量均对干土反射率的预测具有一定作用, 其中空间因子贡献最为突出, 经度(Lon, 0.330)与纬度(Lat, 0.176)在多数波段上表现出较高的解释力。其次为土壤粒级组成, 粉粒(Silt, 0.149)和黏粒(Clay, 0.125)均对反射率预测有显著贡献。有机质(SOM, 0.054)在可见光区的作用相对突出, 而电导率(EC, 0.036)、砂粒(Sand, 0.030)、pH (0.028)及土壤类型(Type, 0.025)贡献相对较弱。整体而言, SHAP 分析验证了所选变量在光谱预测中的科学性, 并揭示了干土光谱反射对多因子的依赖性及各变量的重要性差异(图 5)。

综上, 本研究所选取的 10 个土壤属性(Clay、Silt、Sand、SOM、pH、EC、PM、Type、Lon、Lat)具有合理性, 在相关性与模型解释层面均得到验证。这不仅保证了变量选择的科学性, 也揭示了干土反射率的波段依赖性与多因子驱动特征, 为模型构建提供了坚实的理论与实证基础。

3.2. 基于比尔定律得光滑湿土反射率模型构建

在进行湿土反射率模拟前, 本研究首先分析了等效水厚度与土壤理化属性(如粒径组成、有机质含量、电导率、含盐量、酸碱度)及土壤含水量之间的相关性关系。

在进行等效水厚度参数拟合之前, 本研究首先分析了 EWT 与土壤理化属性(如粒径组成、有机质含量、电导率、含盐量、酸碱度)及土壤含水量之间的相关性关系。结果表明, EWT 不仅与体积含水量密切相关, 同时受土壤结构特性的调制影响。因此, 在模型构建过程中, 需要综合考虑含水量与土壤属性的协同作用, 以提高湿土反射率模拟的精度与泛化能力(图 6)。

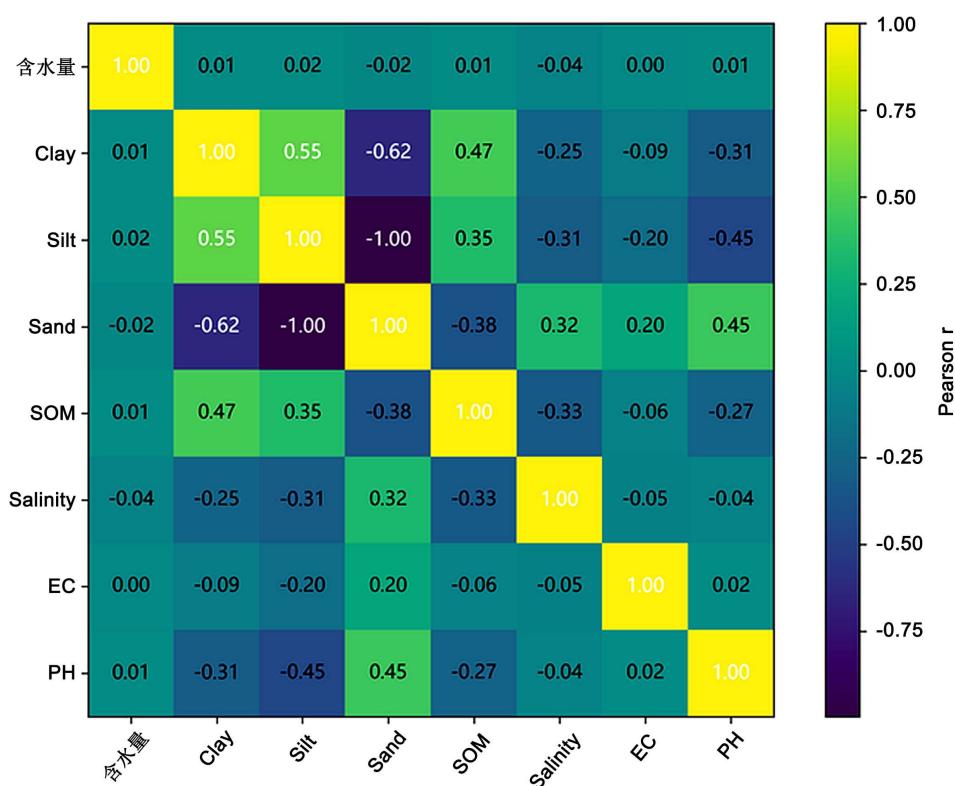


Figure 6. Correlation between soil moisture content, soil properties, and equivalent water thickness

图 6. 土壤含水量及土壤属性与等效水厚度之间的相关性

为避免自变量间强相关性对模型稳定性与解释能力的影响, 本研究对土壤理化属性进行了多重共线性诊断。通过计算容忍度(Tolerance)和方差膨胀因子(VIF)评估变量间相关性。通常, $VIF > 10$ 或容忍度 < 0.1 表示存在严重多重共线性。

结果显示, Silt、Sand 和 pH 的 VIF 值分别为 21.5、12.4 和 31.4, 容忍度分别为 0.05、0.08 和 0.03, 存在明显共线性; Clay 和含水量 VIF 接近阈值(9.1、8.0), SOM、Salinity 和 EC VIF 较低(2.3、1.8、1.1), 共线性较弱。

在模型构建中, 初始使用全部 8 个变量的模型 R^2 为 0.75。逐步剔除共线性较强的 Silt、Sand 和 pH

后， R^2 略降至 0.73；进一步去除相关性较弱的 Clay 或 EC 后， R^2 维持在 0.69~0.72。综合考虑共线性与解释能力，最终选择去除 Clay 或 EC 后的变量组合构建模型，最优模型 R^2 为 0.72，表明在降低共线性影响的同时，模型仍保持较高拟合精度与稳定性(表 1)。

Table 1. Multicollinearity of variables
表 1. 变量多重共线性

参数	含水量	Clay	Silt	Sand	SOM	Salinity	EC	pH
VIF	8.0	9.1	21.5	12.4	2.3	1.8	1.1	31.4
容忍度	0.13	0.11	0.05	0.08	0.38	0.56	0.89	0.03

经多重共线及相关性分析最终选择(含水量、Clay、SOM、Salinity、pH)作为拟合等效水厚度的特征。经多重共线及相关性分析最终选择(含水量、Clay、SOM、Salinity、pH)作为拟合等效水厚度的特征，等效水厚度拟合均值 $R^2 = 0.72$ (图 7)。

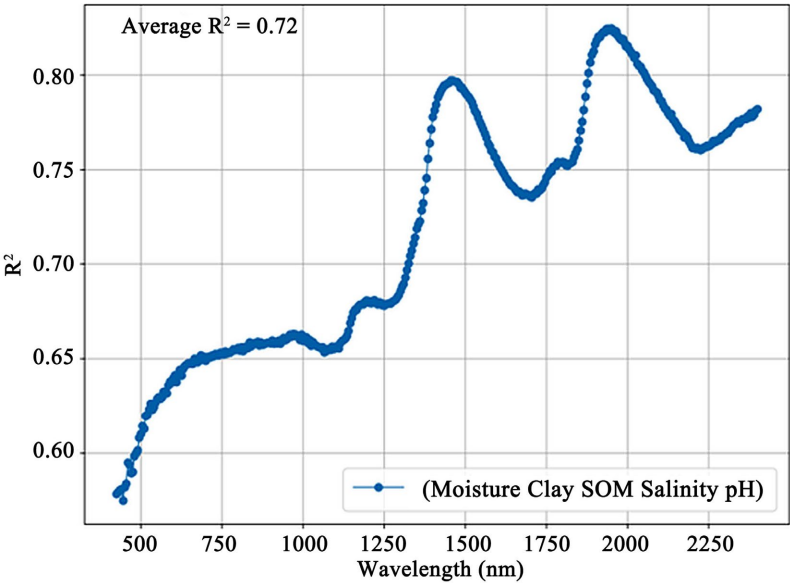


Figure 7. Mean R^2 of equivalent water thickness fitting using soil moisture content, clay, SOM, salinity, and pH
图 7. (含水量、Clay、SOM、Salinity、pH)拟合等效水厚度的平均 R^2

3.3. 光滑表面土壤反射率模型精度

3.3.1. 干土模拟精度

十个变量(Group1)与去掉母质的 9 个变量(Group2)在(350~2500 nm)间隔 10 nm 的 R^2 和 RMSE。如图 8(a)所示，Group1 整体 R^2 高于 Group2，Group1、Group2 在 VIS 至 SWIR 波段 R^2 随波长升高呈现先低后高的趋势；如图 8(b)所示，Group1 整体 RMSE 低于 Group2，Group1、Group2 在 VIS 至 SWIR 波段 RMSE 随波长升高呈先高后底的趋势，干土反射率模型整体表现出较高的拟合度和稳定性，Group1 平均 R^2 为 0.93、RMSE 为 0.018，Group2 平均 R^2 为 0.89、RMSR 为 0.02。结果表明，模型在不同波段均表现出良好的预测能力。

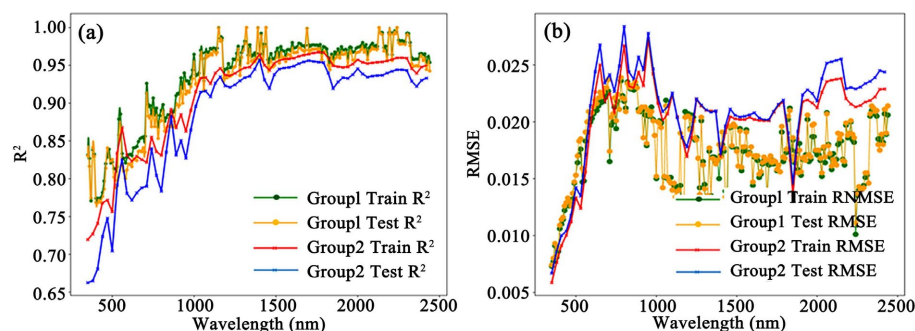


Figure 8. Model accuracy using 10 dry-soil reflectance features with parent material excluded
图 8. 干土反射率 10 个特征与去掉母质的模型精度

3.3.2. 干土反射率模型预测与 Sentinel-2 P90 反射率空间分布

为验证干土反射率模型在区域尺度上的适用性，本研究选取 Sentinel-2 卫星影像数据进行空间对比分析。对于被云或非耕地掩膜覆盖的像元，Google Earth Engine 在统计处理过程中自动忽略 NoData 值，因此未进行插值或缺失值填补处理。

图 9 和图 10 为东北耕地区域在 Sentinel-2 中心波长下的干土反射率预测值空间分布和 Sentinel-2 P90 反射率的空间分布。干土反射率呈现明显的波长依赖特征：在可见光(VIS, 400~700 nm)范围内反射率较低，进入近红外(NIR, 700~1300 nm)后明显升高；在短波红外(SWIR, 1300~2500 nm)区间，反射率曲线表现出显著波动，这主要源于羟基、黏土矿物及碳酸盐的特征吸收作用，其中在约 1400 nm、1900 nm 和 2200 nm 附近出现明显吸收峰，导致整体呈下降趋势。空间分布上，干土反射率存在显著异质性，西部地区反射率普遍高于东部地区。这种空间差异主要由土壤理化性质分布引起。东北平原西北部土壤砂粒含量高、保水能力低、盐分含量高，因此在 NIR 和 SWIR 波段表现出较高反射率；而东南部土壤有机质含量高、颜色较暗，可增强 VIS 和 NIR 波段吸收，从而降低反射率。

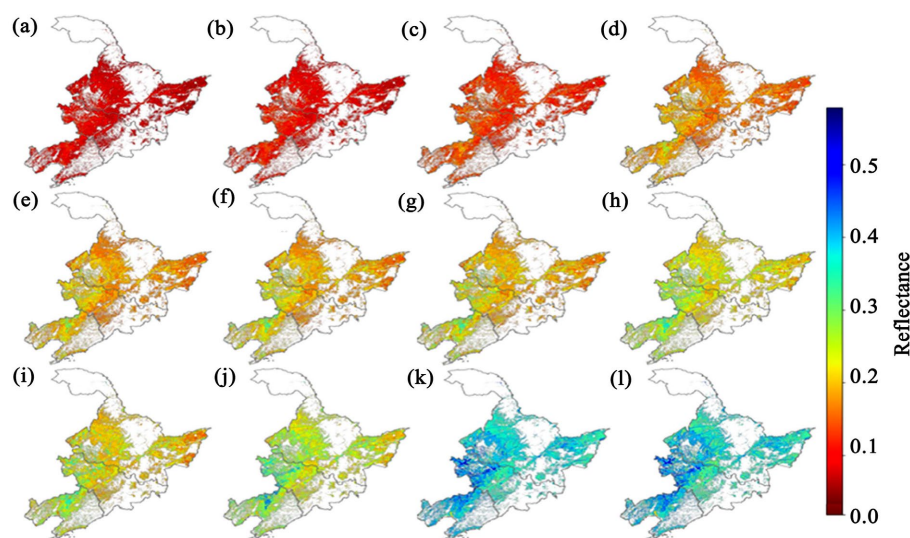


Figure 9. Spatial distribution of predicted dry-soil reflectance. (a)~(l) correspond to the central wavelengths of Sentinel-2 bands: (a) 443 nm, (b) 490 nm, (c) 560 nm, (d) 665 nm, (e) 705 nm, (f) 740 nm, (g) 783 nm, (h) 842 nm, (i) 865 nm, (j) 945 nm, (k) 1610 nm, and (l) 2190 nm

图 9. 干土反射率模型预测空间分布。(a)~(l)分别表示 Sentinel-2 影像各波段的中心波长：(a) 443 nm, (b) 490 nm, (c) 560 nm, (d) 665 nm, (e) 705 nm, (f) 740 nm, (g) 783 nm, (h) 842 nm, (i) 865 nm, (j) 945 nm, (k) 1610 nm, (l) 2190 nm

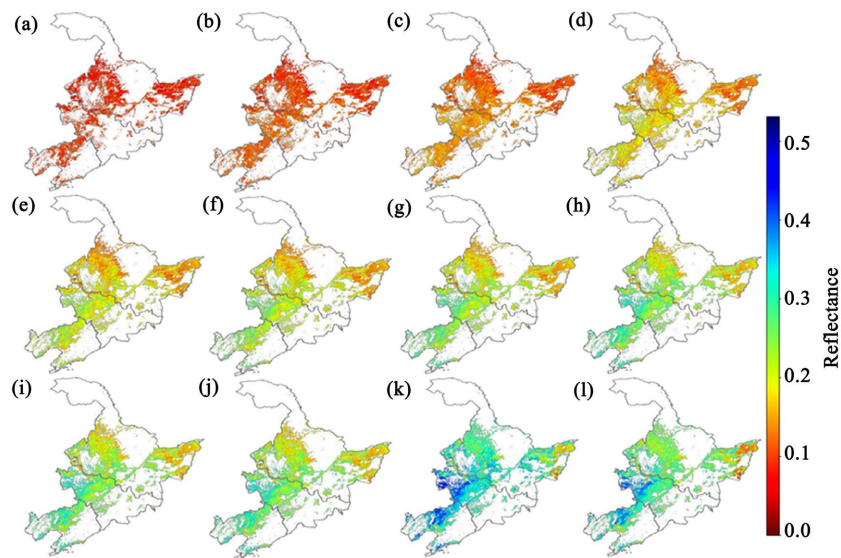


Figure 10. Spatial distribution of Sentinel-2 P90 reflectance. (a)~(l) correspond to the central wavelengths of Sentinel-2 bands: (a) 443 nm, (b) 490 nm, (c) 560 nm, (d) 665 nm, (e) 705 nm, (f) 740 nm, (g) 783 nm, (h) 842 nm, (i) 865 nm, (j) 945 nm, (k) 1610 nm, and (l) 2190 nm

图 10. Sentinel-2 P90 反射率的空间分布。(a)~(l)分别表示 Sentinel-2 各波段的中心波长: (a) 443 nm, (b) 490 nm, (c) 560 nm, (d) 665 nm, (e) 705 nm, (f) 740 nm, (g) 783 nm, (h) 842 nm, (i) 865 nm, (j) 945 nm, (k) 1610 nm, (l) 2190 nm

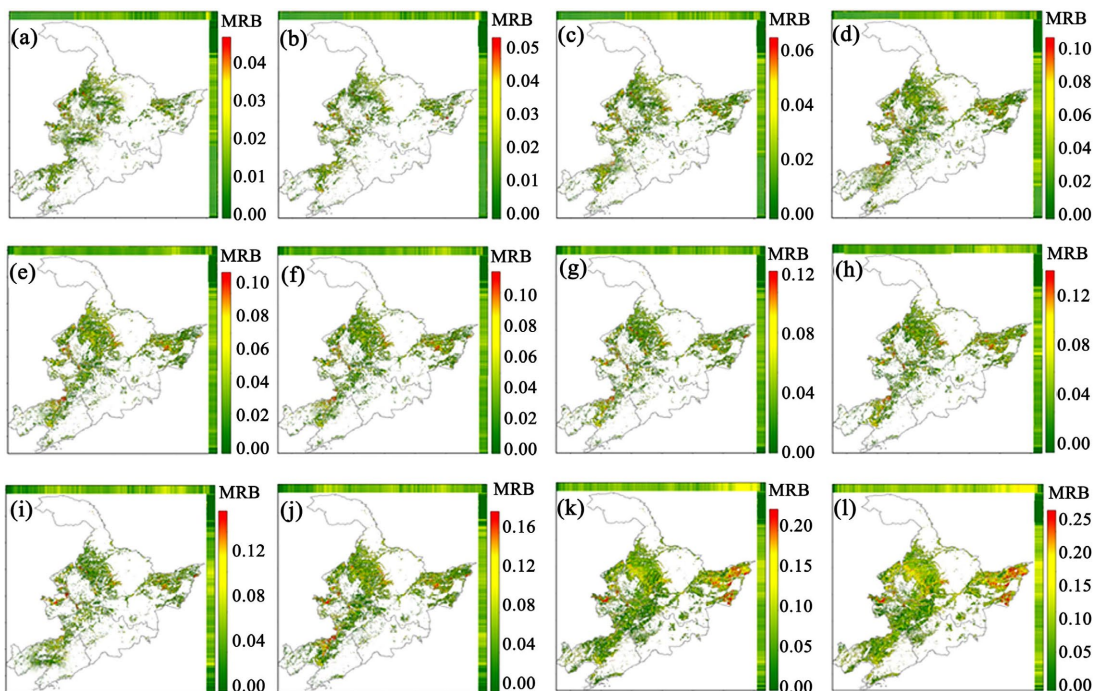


Figure 11. Spatial distribution of differences between modeled dry-soil reflectance and Sentinel-2 P90 reflectance, where the strip gradient represents the mean bias. MRB: mean reflectance bias. (a)~(l) correspond to the central wavelengths of Sentinel-2 bands: (a) 443 nm, (b) 490 nm, (c) 560 nm, (d) 665 nm, (e) 705 nm, (f) 740 nm, (g) 783 nm, (h) 842 nm, (i) 865 nm, (j) 945 nm, (k) 1610 nm, and (l) 2190 nm

图 11. 模型预测干土反射率与 Sentinel-2 P90 反射率差值的空间分布, 其中条带梯度表示平均偏差。MRB: 平均反射率偏差。(a)~(l)分别表示 Sentinel-2 各波段的中心波长: (a) 443 nm, (b) 490 nm, (c) 560 nm, (d) 665 nm, (e) 705 nm, (f) 740 nm, (g) 783 nm, (h) 842 nm, (i) 865 nm, (j) 945 nm, (k) 1610 nm, (l) 2190 nm

整体来看, Sentinel-2 P90 的反射率略低于模型预测的干土反射率, 但两者在光谱模式和空间分布上保持一致, 表明两类数据在光谱响应特征和地理差异性上高度对应。同时, 绘制了干土反射率预测模型与 Sentinel-2 P90 在上述三个波段的偏差空间分布(图 11), 结果显示, 偏差总体上由西向东逐步增大, 东部耕地土壤区域的平均偏差高于西部地区。验证了干土反射率模型的预测准确性。

3.3.3. 湿土模拟精度

在干土反射率研究基础上, 进行了湿土反射率预测。湿土反射率模拟精度在 8%、16%、20%、24% 和 30% 五个质量含水量梯度下, R^2 分别达到 0.78、0.85、0.85、0.89 和 0.85。本研究构建的湿土光谱反射率模型能够较好地反映不同水分条件下土壤光谱响应的变化规律, 具有较好的稳定性和适用性(图 12)。

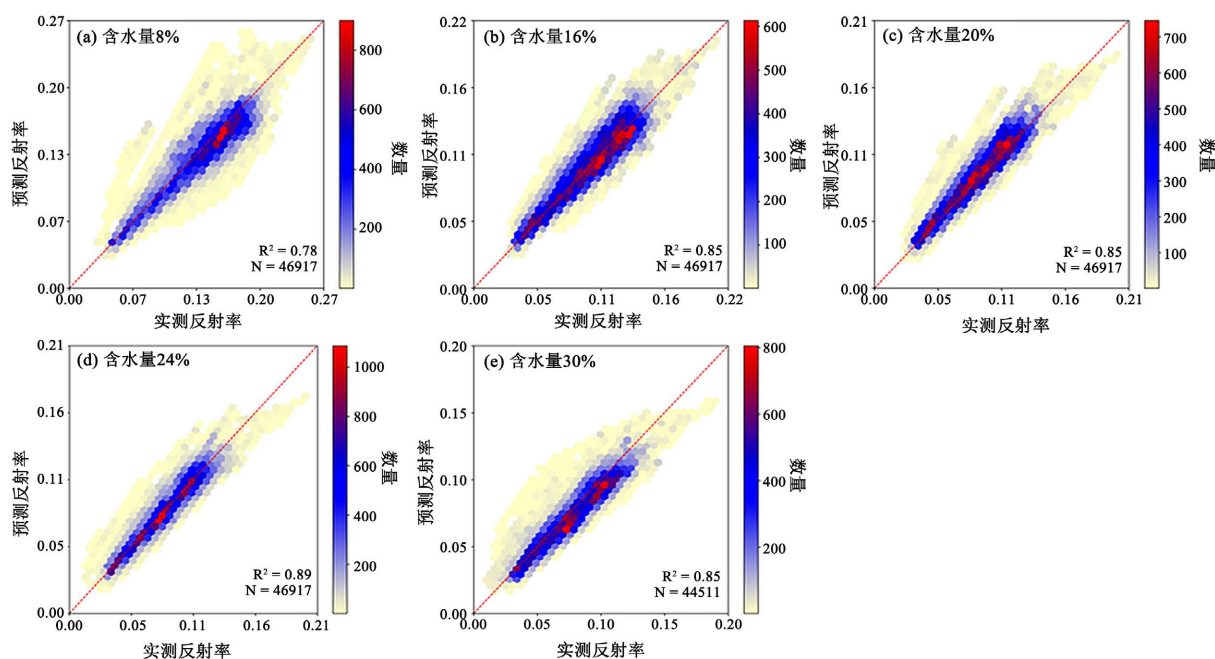


Figure 12. Model accuracy of wet soil (8%, 16%, 20%, 24%, and 30%) based on the Beer-Lambert law

图 12. 基于比尔-郎伯定律的 8%、16%、20%、24%、30% 的湿土模型精度

在(8%~30%)含水量范围内, 本研究构建的湿土光谱反射率模型能够较好地反映不同水分条件下土壤光谱响应的变化规律, 具有较好的稳定性和适用性, 为后续区域尺度土壤湿度遥感反演及相关研究提供了可靠的理论基础。

4. 讨论

母质作为决定土壤形成初始条件的关键因子, 在干土光谱响应机制中发挥着显著作用。为进一步量化母质特征对干土反射率贡献, 构建了不包含母质特征的梯度提升回归训练模型, 与本研究结果进行对比分析。在全特征建模条件下, 各波段平均验证集 R^2 达到 0.93, RMSE 为 0.018 图 8(a), 去除母质特征后的训练模型平均 $R^2 = 0.89$, RMSE 为 0.02 图 8(b), 表明母质对土壤光谱预测干土反射率具有显著增强作用。

5. 结论

针对干土反射率的建模精度不足问题, 本研究以中国东北黑土区为研究区域, 构建了一种综合考虑

多种土壤性质(土壤质地、有机质、酸碱度、电导率、土壤类型、母质)以及地理位置的 400~2500 nm 干土光谱反射率智能学习模型,并在干土反射率模型基础上构建湿土反射率模型,模型精度与评价结果如下:

(1) EEDSR 模型在 400~2500 nm 的模拟干土反射率的平均 R^2 和 RMSE 分别为 0.93 和 0.018,在 VIS 至 SWIR 波段, R^2 随波长升高呈现先低后高的趋势, RMSE 随波长升高呈先高后低的趋势;

(2) 将母质(PM)作为土壤反射率的预测因子,加入母质后干土反射率预测平均 R^2 由 0.89 提升至 0.93,相对提升幅度达 4.5%;

(3) EEDSR 模型预测的干土反射率与 Sentinel-2 P90 裸土最大反射率具有相似的空间分布格局,二者偏差自西向东逐渐增大,符合地表湿度的空间分布规律。这也侧面证实了本研究提出 EEDSR 模型的合理性;

(4) 湿土反射率模拟精度在 8%、16%、20%、24%和 30%五个质量含水量梯度下, R^2 分别达到 0.78、0.85、0.85、0.89 和 0.85;

模型在可见光至短波红外波段实现了高精度土壤反射率预测,具有良好的可解释性与泛化能力,可为土壤光谱建模与土壤属性反演提供技术支撑,后期可考虑在此基础上加入土壤水分、地表粗糙度和观测几何的影响,从而提供与卫星观测更为一致的土壤反射率模拟能力。

参考文献

- [1] 夏学齐, 季峻峰, 陈骏, 廖启林, 杨忠芳, 等. 土壤理化参数的反射光谱分析[J]. 地学前缘, 2009, 16(4): 356-364.
- [2] Bartholomeus, H.M., Schaepman, M.E., Kooistra, L., Stevens, A., Hoogmoed, W.B. and Spaargaren, O.S.P. (2008) Spectral Reflectance Based Indices for Soil Organic Carbon Quantification. *Geoderma*, **145**, 28-36. <https://doi.org/10.1016/j.geoderma.2008.01.010>
- [3] Ribeiro, S.G., Oliveira, M.R.R.D., Lopes, L.M., Costa, M.C.G., Toma, R.S., Araújo, I.C.D.S., et al. (2023) Reflectance Spectroscopy in the Prediction of Soil Organic Carbon Associated with Humic Substances. *Revista Brasileira de Ciência do Solo*, **47**, e0230051. <https://doi.org/10.36783/18069657rbcs20220143>
- [4] Cierniewski, J. (1987) A Model for Soil Surface Roughness Influence on Soil Reflectance in the Visible and Near-Infrared Range. *Remote Sensing of Environment*, **23**, 79-115. [https://doi.org/10.1016/0034-4257\(87\)90073-3](https://doi.org/10.1016/0034-4257(87)90073-3)
- [5] Tiruneh, G.A., Meshesha, D.T., Adgo, E., Tsunekawa, A., Haregeweyn, N., Fenta, A.A., et al. (2022) Use of Soil Spectral Reflectance to Estimate Texture and Fertility Affected by Land Management Practices in Ethiopian Tropical Highland. *PLOS ONE*, **17**, e0270629. <https://doi.org/10.1371/journal.pone.0270629>
- [6] Stoner, E.R. and Baumgardner, M.F. (1981) Characteristic Variations in Reflectance of Surface Soils. *Soil Science Society of America Journal*, **45**, 1161-1165. <https://doi.org/10.2136/sssaj1981.03615995004500060031x>
- [7] Clark, R.N. (1999) Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy. In: Rencz, A.N., Ed., *Remote Sensing for the Earth Sciences*, John Wiley & Sons, 3-58.
- [8] Gomez, C., Viscarra Rossel, R.A. and McBratney, A.B. (2008) Soil Organic Carbon Prediction by Hyperspectral Remote Sensing and Field Vis-Nir Spectroscopy: An Australian Case Study. *Geoderma*, **146**, 403-411. <https://doi.org/10.1016/j.geoderma.2008.06.011>
- [9] Price, J.C. (1990) On the Information Content of Soil Reflectance Spectra. *Remote Sensing of Environment*, **33**, 113-121. [https://doi.org/10.1016/0034-4257\(90\)90037-m](https://doi.org/10.1016/0034-4257(90)90037-m)
- [10] Jiang, C. and Fang, H. (2019) GSV: A General Model for Hyperspectral Soil Reflectance Simulation. *International Journal of Applied Earth Observation and Geoinformation*, **83**, Article ID: 101932. <https://doi.org/10.1016/j.jag.2019.101932>
- [11] Hapke, B. (1981) Bidirectional Reflectance Spectroscopy: 1. Theory. *Journal of Geophysical Research: Solid Earth*, **86**, 3039-3054. <https://doi.org/10.1029/jb086ib04p03039>
- [12] Hapke, B. (1984) Bidirectional Reflectance Spectroscopy: 3. Correction for Macroscopic Roughness. *Icarus*, **59**, 41-59. [https://doi.org/10.1016/0019-1035\(84\)90054-x](https://doi.org/10.1016/0019-1035(84)90054-x)
- [13] Hapke, B. (2002) Bidirectional Reflectance Spectroscopy: 5. The Coherent Backscatter Opposition Effect and Anisotropic Scattering. *Icarus*, **157**, 523-534. <https://doi.org/10.1006/icar.2002.6853>
- [14] Hapke, B. (2012) Bidirectional Reflectance Spectroscopy 7: The Single Particle Phase Function Hockey Stick Relation. *Icarus*, **221**, 1079-1083. <https://doi.org/10.1016/j.icarus.2012.10.022>

-
- [15] Babet, A., Vu, P.V.H., Jacquemoud, S., Viallefont-Robinet, F., Fabre, S., Briottet, X., *et al.* (2018) MARMIT: A Multilayer Radiative Transfer Model of Soil Reflectance to Estimate Surface Soil Moisture Content in the Solar Domain (400-2500 nm). *Remote Sensing of Environment*, **217**, 1-17. <https://doi.org/10.1016/j.rse.2018.07.031>
- [16] Yang, P., van der Tol, C., Yin, T. and Verhoef, W. (2020) The SPART Model: A Soil-Plant-Atmosphere Radiative Transfer Model for Satellite Measurements in the Solar Spectrum. *Remote Sensing of Environment*, **247**, Article ID: 111870. <https://doi.org/10.1016/j.rse.2020.111870>
- [17] Lei, T. and Bailey, B.N. (2025) A Text-Based, Generative Deep Learning Model for Soil Reflectance Spectrum Simulation in the Solar Range (400-2499 nm). *Remote Sensing of Environment*, **318**, Article ID: 114527. <https://doi.org/10.1016/j.rse.2024.114527>
- [18] Roda, F. and Nemiña, F. (2025) From Spectra to Semantics: An Ontology-Based Model of Spectral Observations Results. 2025 *Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)*, Foz do Iguaçu, 10-13 November 2025, 1-7. <https://doi.org/10.1109/lagirs68367.2025.11414769>
- [19] Gascon, F., Bouzinac, C., Thépaut, O., *et al.* (2017) Copernicus Sentinel-2A Calibration and Products Validation Status. *Remote Sensing*, **9**, Article 584. <https://doi.org/10.3390/rs9060584>
- [20] Ibrahim, E. and Gobin, A. (2021) Sentinel-2 Recognition of Uncovered and Plastic-Covered Agricultural Soil. *Remote Sensing*, **13**, Article 4195. <https://doi.org/10.3390/rs13214195>
- [21] Farzamian, M., Castanheira, N., Gonçalves, M.C., Freitas, P., Saberioon, M., Ramos, T.B., *et al.* (2025) Predicting Soil Organic Carbon from Sentinel-2 Imagery and Regional Calibration Approach in Salt-Affected Agricultural Lands: Feasibility and Influence of Soil Properties. *Remote Sensing*, **17**, Article 2877. <https://doi.org/10.3390/rs17162877>
- [22] 邓孺孺, 田国良, 柳钦火, 辛晓洲. 粗糙地表土壤含水量遥感模型研究[J]. 遥感学报, 2004, 8(1): 75-80.
- [23] 邓孺孺, 何颖清, 秦雁, 等. 分离悬浮质影响的光学波段(400-900 nm)水吸收系数测量[J]. 遥感学报, 2012, 16(1): 174-191.
- [24] 邓孺孺, 何颖清, 秦雁, 等. 近红外波段(900-2500 nm)水吸收系数测量[J]. 遥感学报, 2012, 16(1): 192-206.