

目次

中国城市碳排放强度的时空演变、动态跃迁及收敛趋势 杨清可, 王磊, 朱高立, 李颖, 范业婷, 王雅竹 (1869)

交通运输业碳排放效率时空演变及趋势预测 郑琰, 蒋雪梅, 肖玉杰 (1879)

基于扩展STIRPAT模型LMDI分解的碳排放脱钩因素 张江艳 (1888)

基于LEAP模型的工业园区碳达峰路径:以南京某国家级开发区为例 李慧鹏, 李荔, 殷茵, 何文大, 宿杰, 赵秋月 (1898)

高校碳排放核算与分析:以北京A高校为例 曹睿, 封莉, 张立秋 (1907)

北京市制造业减排降碳协同效应分析和驱动因素 俞珊, 韩玉花, 牟洁, 张双, 张增杰 (1917)

碳排放权交易价格与全要素生产率:来自中国的证据 吴雪萍 (1926)

中国城市PM_{2.5}和PM₁₀时空分布特征和影响因素分析 李江苏, 段良荣, 张天娇 (1938)

2017~2021年苏皖鲁豫交界区域PM_{2.5}和O₃时空变化特征及影响因素 陈伟, 徐学哲, 刘文清 (1950)

疫情管控期西安PM_{2.5}和O₃污染特征及成因分析 原晓红, 张强, 李琦, 谢文豪, 刘跃廷, 樊亭亭, 姜旭朋 (1963)

苏南五市秋冬季PM_{2.5}化学组成特征和空间差异 冯蔚, 丁峰, 尚玥, 谢鸣捷 (1975)

湖北咸宁细颗粒物PM_{2.5}来源 罗怡, 朱宽广, 陈璞琬, 田军, 谢旻, 战杨志豪, 赵润琪 (1983)

邢台氨减排对京津冀PM_{2.5}改善的溢出效益 边泽君, 闻超玉, 郎建垒, 范晓茜, 夏祥晨, 周颖 (1994)

聊城市冬季PM_{2.5}载带金属元素污染特征、风险评价及来源分析 张敬巧, 朱瑶, 曹婷, 燕丽, 王淑兰, 刘铮 (2003)

高原城市拉萨典型VOCs排放源成分谱特征 郭淑政, 叶春翔, 林伟立, 陈熠, 曾立民, 尹晴晴, 刘雪莉 (2011)

北京市工业园区VOCs污染特征及健康风险评估案例:高新技术产业的环境影响 王洁, 姚震, 王敏燕, 陈速敏, 龙腾, 王海滨, 李红, 郭秀锐, 郝江虹, 聂磊 (2019)

高温极端天气影响下的成都平原一次典型臭氧污染过程分析 雷丽娟, 张懿, 罗伊娜, 张潇, 冯森 (2028)

水环境中抗病毒毒物的存在、行为与风险 葛林科, 李璇艳, 曹胜凯, 郑金帅, 张蓬, 朱超, 马宏瑞 (2039)

黄河小浪底水库地表水中重金属的时空变化与概率健康风险 王亮, 邓雪娇, 王潇磊, 李明, 刘奕尧, 姜亚敏, 涂响, 张坤锋 (2054)

藏东多曲河流域总富集水化学特征及控制因素 李敬杰, 连晟, 王明国, 张智印, 张涛 (2067)

西北内陆区降水稳定同位素时空分布特征及其水汽来源 张炎炎, 辛存林, 郭小燕, 张博, 陈宁, 史延飞 (2080)

基于水化学和氢氧同位素的泰安城区岩溶地下水补给来源及演化过程 孟令华 (2096)

丹江流域山区地表水-地下水水化学特征及其影响因素 张子燕, 伏永朋, 王宁涛, 谭建民, 刘亚磊 (2107)

金塔盆地鸳鸯池灌区地下水水化学特征及控制因素 王晓燕, 韩双宝, 张梦南, 尹德超, 吴呈, 安永会 (2118)

德阳市平原区浅层地下水水化学特征与健康风险评价 刘楠, 陈盟, 高东东, 吴勇, 王橹橹 (2129)

典型城市河网沉积物微塑料时空分布特征 许万璐, 范一凡, 钱新 (2142)

抚仙湖流域尺度氮排放清单构建及关键源解析 王延杰, 梁启斌, 王艳霞, 侯磊, 陈奇伯, 王伟, 李晓琳, 高俊淑 (2150)

过氧化钙/海泡石海藻酸钠缓释凝胶复合材料的制备及其对内源磷的控制性能 曲思彤, 单苏洁, 王崇铭, 吴玲子, 李大鹏, 黄勇 (2160)

矿物超细颗粒的形成机制、结构特征及其环境行为和效应 刘振海, 张展华, 袁语欣, 朱盼盼, 陈威, 张彤 (2171)

生物炭固定化菌复合材料在环境修复中的应用研究进展 孙淑玉, 黄梦鑫, 孔强, 张焕新, 刘继伟 (2185)

微塑料对沸石吸附水体氨氮的影响及其机制 练建军, 谢诗婷, 吴培, 孟冠军, 陈波 (2195)

紫外老化作用对纳米生物炭吸附环丙沙星的影响机制 马锋锋, 薛之一, 赵保卫 (2203)

土地利用影响下永定河流域浮游植物群落与环境因子响应 郭善嵩, 胡恩, 丁一桐, 张嘉渭, 孙长顺, 卢悦, 潘保柱 (2211)

鄱阳湖湿地细菌群落多样性和可培养细菌功能基因丰度 喻江, 王淳, 龙永, 刘贵花, 李春杰, 范国权, 于镇华 (2223)

典农河沉积物细菌群落结构特征及其与重金属的关系 刘双羽, 蒙俊杰, 邱小琼, 周瑞娟, 李霖 (2233)

污水处理厂尾水排放对受纳河流细菌和真菌微生物群落的影响 郭有顺, 余仲, 郝文彬, 孟凡刚 (2246)

宏基因组学分析深度处理阶段污水中细菌的赋存特征及其功能 胡健双, 王燕, 周政, 汪雅琴, 王秉政, 李激 (2259)

大别山区生境质量时空特征及自然-人为因素驱动机制 郑亚平, 张俊华, 田惠文, 朱航成, 刘舒, 丁亚鹏 (2268)

基于连续小波变换、SHAP和XGBoost的土壤有机质含量高光谱反演 叶森, 朱琳, 刘旭东, 黄勇, 陈蓓蓓, 李欢 (2280)

秸秆还田、覆膜和施氮对旱地麦田土壤质量的影响 叶子壮, 王松燕, 陆潇, 史多鹏, 吕慎强, 李嘉, 杨泽宇, 王林科 (2292)

长期施用有机肥对土壤微塑料赋存及迁移特征的影响 王长远, 马啸驰, 郭德杰, 刘新红, 马艳, 罗佳 (2304)

土壤盐分变化对N₂O排放影响:基于Meta分析 黄艺华, 余冬立, 史祯琦, 胡磊, 潘永春 (2313)

基于土地利用/覆被动态变化的粤港澳大湾区碳储量评价与预测 郑慧玲, 郑辉峰 (2321)

基于InVEST模型的伏牛山地区生态系统碳储量时空变化模拟 张哲, 时振钦, 朱文博, 孙梓欣, 赵体侠, 邓文萍, 刘志强 (2332)

防护林建设过程中土壤微生物养分限制与有机碳组分之间的关系 徐凤璟, 黄懿梅, 黄倩, 申继凯 (2342)

长期秸秆还田褐土有机碳矿化特征及其驱动力 赵宇航, 殷浩凯, 胡雪纯, 解文艳, 刘志平, 周怀平, 杨振兴 (2353)

冬绿肥覆盖对土壤团聚体及有机碳和AMF多样性的影响 鲁泽让, 陈佳钰, 李智贤, 李永梅, 罗志章, 杨锐, 田明洋, 赵吉霞, 范茂攀 (2363)

辽河流域氮素时空分布及其对土地利用和降雨的响应 周波, 李晓光, 童思陈, 吕旭波, 郭朝臣, 雷坤 (2373)

高强度农业种植区不同景观池塘氧化亚氮排放特征 张欣悦, 肖启涛, 谢晖, 刘臻婧, 邱银国, 罗菊花, 徐向华, 段洪涛 (2385)

生物炭与不同类型氮肥施肥对菜地土壤反硝化细菌群落的影响 柳晓婉, 刘杏认, 高尚洁, 李贵春 (2394)

大岳山不同林龄人工油松林土壤微生物特征 马义淑, 曹亚鑫, 牛敏, 张明昱, 程曼, 文永莉 (2406)

基于稀疏样点的土壤重金属含量模拟方法 张佳琦, 潘瑜春, 高世臣, 赵亚楠, 景胜强, 周艳兵, 郇允兵 (2417)

基于特定源-风险评估模型的兰州黄河风情线绿地土壤重金属污染优先控制源分析 李军, 李旭, 李开明, 焦亮, 台喜生, 臧飞, 陈伟, 脱新颖 (2428)

西南不同类型紫色土pH变化、重金属累积与潜在生态风险评估 张海琳, 张雨, 王顶, 谢军, 张跃强, 张宇亭, 王洁, 石孝均 (2440)

重庆某铁矿周边耕地土壤重金属污染评价及来源解析 廖泽源, 李杰芹, 沈智杰, 李彩霞, 罗程钟, 梅楠, 张成, 王定勇 (2450)

金属矿区周边农田土壤与农作物重金属健康风险评估 魏洪斌, 罗明, 向奎, 查理思 (2461)

赤泥基纳米零价铁对多金属污染土壤修复效果 刘龙宇, 杨世利, 赵黄诗雨, 常凯威, 余江 (2473)

施锌对碱性土壤-小麦幼苗体系累积镉的影响 张瑶, 王天齐, 牛硕, 杨阳, 陈卫平 (2479)

基于连续小波变换、SHAP 和 XGBoost 的土壤有机质含量高光谱反演

叶淼^{1,2,3}, 朱琳^{1,2,3*}, 刘旭东^{1,2,3}, 黄勇⁴, 陈蓓蓓^{1,2,3}, 李欢⁴

(1. 首都师范大学资源环境与旅游学院, 北京 100048; 2. 首都师范大学城市环境过程和数字模拟国家重点实验室培育基地, 北京 100048; 3. 首都师范大学水资源安全北京实验室, 北京 100048; 4. 北京市生态地质研究所, 北京 100120)

摘要: 针对土壤有机质含量高光谱反演中存在的光谱有效信号薄弱和光谱信息冗余问题, 提出结合连续小波变换、SHAP 和 XGBoost 的土壤有机质含量高光谱反演框架. 以北京市延庆区和房山区永久基本农田土壤为研究对象, 首先, 基于连续小波变换处理的土壤光谱反射率数据构建初始 XGBoost 模型; 然后, 利用 SHAP 方法分析模型中各波段的贡献度以筛选特征波段; 最后, 基于特征波段重新构建和优化 XGBoost 模型, 实现土壤有机质含量高光谱反演. 发现连续小波变换尺度为 2^5 时, 利用 SHAP 方法选取的 40 个特征波段构建的 XGBoost 模型准确性最高, 有机质含量反演值和实测值之间的 R^2 为 0.80, RMSE 为 $3.60 \text{ g}\cdot\text{kg}^{-1}$; 随着连续小波变换尺度的增大, R^2 呈现先升高后降低的趋势, 2^5 尺度下的 R^2 比 2^1 尺度的高 0.37; SHAP 方法选取的特征波段比 Pearson 相关分析法少 682 个, RMSE 低 $0.69 \text{ g}\cdot\text{kg}^{-1}$; XGBoost 模型的 R^2 分别比随机森林和支持向量机模型高 4% 和 8%. 验证了结合连续小波变换、SHAP 和 XGBoost 在土壤有机质含量高光谱反演中的有效性, 为快速、准确监测土壤有机质含量提供技术支撑.

关键词: 土壤有机质(SOM); 高光谱反演; 连续小波变换; SHAP 方法; XGBoost 模型

中图分类号: X53 文献标识码: A 文章编号: 0250-3301(2024)04-2280-12 DOI: 10.13227/j.hjkk.202304100

Hyperspectral Inversion of Soil Organic Matter Content Based on Continuous Wavelet Transform, SHAP, and XGBoost

YE Miao^{1,2,3}, ZHU Lin^{1,2,3*}, LIU Xu-dong^{1,2,3}, HUANG Yong⁴, CHEN Bei-bei^{1,2,3}, LI Huan⁴

(1. College of Resources Environment and Tourism, Capital Normal University, Beijing 100048, China; 2. Laboratory Cultivation Base of Environment Process and Digital Simulation, Capital Normal University, Beijing 100048, China; 3. Beijing Laboratory of Water Resources Security, Capital Normal University, Beijing 100048, China; 4. Beijing Institute of Ecological Geology, Beijing 100120, China)

Abstract: Aiming to address the problems of weak spectral signals and redundant spectral information existing in hyperspectral inversion of soil organic matter content, a hyperspectral inversion framework combining continuous wavelet transform, SHAP, and XGBoost was proposed. Taking the permanent basic farmland soil in Yanqing District and Fangshan District of Beijing as the research object, an initial XGBoost model was first constructed based on the spectral reflectance data of soil processed by continuous wavelet transform. Then, the SHAP method was used to analyze the contribution of each band in the model to select the characteristic bands. Finally, the XGBoost model was reconstructed and optimized based on the characteristic bands to realize the hyperspectral inversion of soil organic matter content. It was found that the XGBoost model based on the 40 characteristic bands of continuous wavelet transform at the 2^5 scale selected by the SHAP method showed the highest accuracy, with the R^2 and RMSE between the inverted and measured value of the organic matter content being 0.80 and $3.60 \text{ g}\cdot\text{kg}^{-1}$, respectively. The R^2 first increased and then decreased with the increase in the scale of continuous wavelet transform, and the R^2 at the 2^5 scale was 0.37 higher than that at the 2^1 scale. The number of characteristic bands selected by the SHAP method was 682 less than that by the Pearson correlation analysis method, and the RMSE was $0.69 \text{ g}\cdot\text{kg}^{-1}$ lower. The R^2 of the XGBoost model was 4% and 8% higher than that of the random forest and support vector machine models, respectively. The results demonstrated the effectiveness of the combination of continuous wavelet transform, SHAP, and XGBoost in the hyperspectral inversion of soil organic matter content, which provides technical support for rapid and accurate monitoring of soil organic matter content.

Key words: soil organic matter(SOM); hyperspectral inversion; continuous wavelet transform; SHAP method; XGBoost model

土壤有机质(soil organic matter, SOM)是土壤中含碳有机物的总称,影响着全球碳循环^[1]、土壤碳固定^[2]、肥力^[3]、微生物群落^[4]及其生态功能^[5]等.传统土壤有机质含量的检测主要通过土壤样品进行化学分析来实现,检测方法包括容量法、比色法和灼烧法等^[6],这些方法准确性高,但操作复杂、耗时且成本高,不适合在较大尺度开展^[7,8].高光谱遥感具有光谱分辨率高、数据信息丰富和实时高效等特点,已逐渐成为获取土壤有机质含量的重要手段之一^[9,10].然而,由于受到光谱测量环境、光源、

仪器以及土壤样本质量等多种因素的干扰,以及高光谱数据自身存在的相邻波段共线性较强和光谱信息冗余度较高等问题^[11],土壤有机质含量高光谱反演模型的性能和精度较低.突出高光谱数据中的有效信号,滤除冗余和共线的光谱变量,筛选出与有

收稿日期: 2023-04-12; 修订日期: 2023-06-25

基金项目: 国家自然科学基金项目(42271082); 北京卓越青年科学家计划项目(BJJWZYJH01201910028032)

作者简介: 叶淼(1995~),女,博士研究生,主要研究方向为生态环境遥感, E-mail: 2200901018@cnu.edu.cn

* 通信作者, E-mail: lin.zhu@cnu.edu.cn

机质含量相关性较强的特征波段, 对降低模型复杂度和提升模型精度具有重要意义。

为了增强光谱中的有效信号, 通常对土壤原始光谱数据进行倒数、倒数对数以及一阶微分和二阶微分等多种数学变换^[12,13]。聂哲等^[14]通过对原始光谱开展一阶微分、倒数对数和连续统去除变换, 建立东北典型黑土区表层土壤有机质含量高光谱反演模型, 结果表明基于一阶微分变换处理的光谱数据构建的模型精度最高, 有机质含量反演值和实测值之间的 R^2 比基于原始光谱构建的模型高0.33。随着研究的深入, 有学者发现光谱数据经连续小波变换(continuous wavelet transform, CWT)后与土壤元素具有更强的相关性, 模型反演精度更高^[15,16]。玉米提·买明等^[17]对原始光谱分别进行倒数对数、一阶微分、倒数对数一阶微分以及连续小波变换, 结果表明光谱经过连续小波变换处理后与有机质含量的相关性相较于其他变换方法得到显著提高, 相关系数最大提升0.15。Tan等^[18]研究了连续小波变换不同尺度对黑土有机质含量高光谱反演的影响, 结果表明不同尺度小波系数可以在不同程度上提高特征波段与有机质含量的相关性, 2^7 尺度的相关系数比 2^1 尺度大0.2。

上述研究均采用相关分析法在连续小波变换的基础上选取特征波段, 该方法也是土壤有机质含量高光谱反演研究中使用最多的特征波段筛选方法^[19,20], 通过计算有机质含量和小波系数之间的相关系数, 选择达到极显著水平且相关系数较大的波段作为特征波段^[21]。然而, 相关系数表征变量之间的线性相关程度, 由于光谱反射率与有机质含量之间的关系复杂, 仅考虑线性关系存在一定的局限性, 并且通过相关系数筛选的特征波段之间仍然存在多重共线性问题^[22]。SHAP (Shapley additive explanation)是由Lundberg等^[23]于2017年在合作博弈论启发下构建的一种可加性解释方法, 通过计算多个变量协同作用下单个变量的贡献度来确定该变量的重要程度, 不仅考虑单个变量的影响, 还考虑变量之间的协同效应, 解决了变量之间的多重共线性问题, 有学者尝试将该方法用于模型特征变量筛选。赵健等^[24]利用SHAP方法筛选与路面类型相关性大的特征参与路面识别模型构建, 在多种路面类型识别中模型的准确率均在96%以上。Liu等^[25]针对帕金森病诊断中医疗数据特征维数高的问题, 引入SHAP方法对医疗数据进行特征筛选, 结果表明, 基于SHAP特征筛选的医疗数据分类模型在帕金森病诊断中准确率达到91.78%。据检索, SHAP方法暂未应用于高光谱数据特征波段选择相关

研究。

除光谱预处理和特征波段选择外, 反演模型对提升土壤有机质含量高光谱反演性能也至关重要, 常用的反演模型主要包括统计模型和机器学习模型。统计模型有多元线性回归^[26]和偏最小二乘回归^[27]等, 机器学习模型有支持向量机(support vector machine, SVM)^[28]、随机森林(random forest, RF)^[29]和卷积神经网络^[30]等。已有较多研究证明了机器学习模型比统计模型在土壤有机质含量高光谱反演中具有更好的精度^[31,32]。周伟等^[33]构建偏最小二乘回归、SVM和RF模型来反演三江源区土壤有机质含量, 结果表明, RF和SVM模型精度高于偏最小二乘回归模型, R^2 分别为0.78、0.78和0.71。极限梯度提升树(extreme gradient boosting, XGBoost)是由Chen等^[34]于2016年提出的一种集成学习算法, 在土壤组分高光谱反演中已得到广泛运用^[35,36]。Sun等^[37]对比了XGBoost、BP神经网络、RF和SVM等5种机器学习模型在土壤镍含量高光谱反演中的性能, 结果表明, XGBoost表现最好, 相对分析误差(ratio of prediction to deviation, RPD)为2.08。葛翔宇等^[38]利用XGBoost、RF和梯度提升回归树共3种集成学习算法, 建立土壤含水量高光谱反演模型, 结果表明, 3种模型精度排序为: XGBoost > RF > 梯度提升回归树, R^2 分别为0.93、0.91和0.89。

本研究提出结合连续小波变换、SHAP和XGBoost的土壤有机质含量高光谱反演框架, 以北京市延庆区和房山区永久基本农田土壤为例, 基于193个土壤样品的实验室光谱和有机质含量数据, 对土壤光谱反射率进行连续小波变换, 利用SHAP方法筛选与土壤有机质含量密切相关的特征波段, 构建基于XGBoost算法的土壤有机质含量高光谱反演模型, 以期对土壤有机质含量高光谱监测提供新思路。

1 材料与方法

1.1 研究区概况

为增加土壤样本丰富性, 分别选择北京市延庆区中部($115^{\circ}57' \sim 116^{\circ}12' E$, $40^{\circ}21' \sim 40^{\circ}34' N$)与房山区东南部($115^{\circ}54' \sim 116^{\circ}13' E$, $39^{\circ}41' \sim 39^{\circ}33' N$)作为研究区(图1), 面积共422 km²。区内属大陆性季风气候, 夏季炎热多雨、冬季寒冷干燥。2021年, 延庆区平均气温为10.5℃, 降水量为605.3 mm; 房山区平均气温为12.7℃, 降水量为865.7 mm。区内永久基本农田广泛分布, 延庆区中部永久基本农田分布范围内海拔为482~1 124 m, 房山区东南部永久基本农田分布范围内海拔为22~127 m。

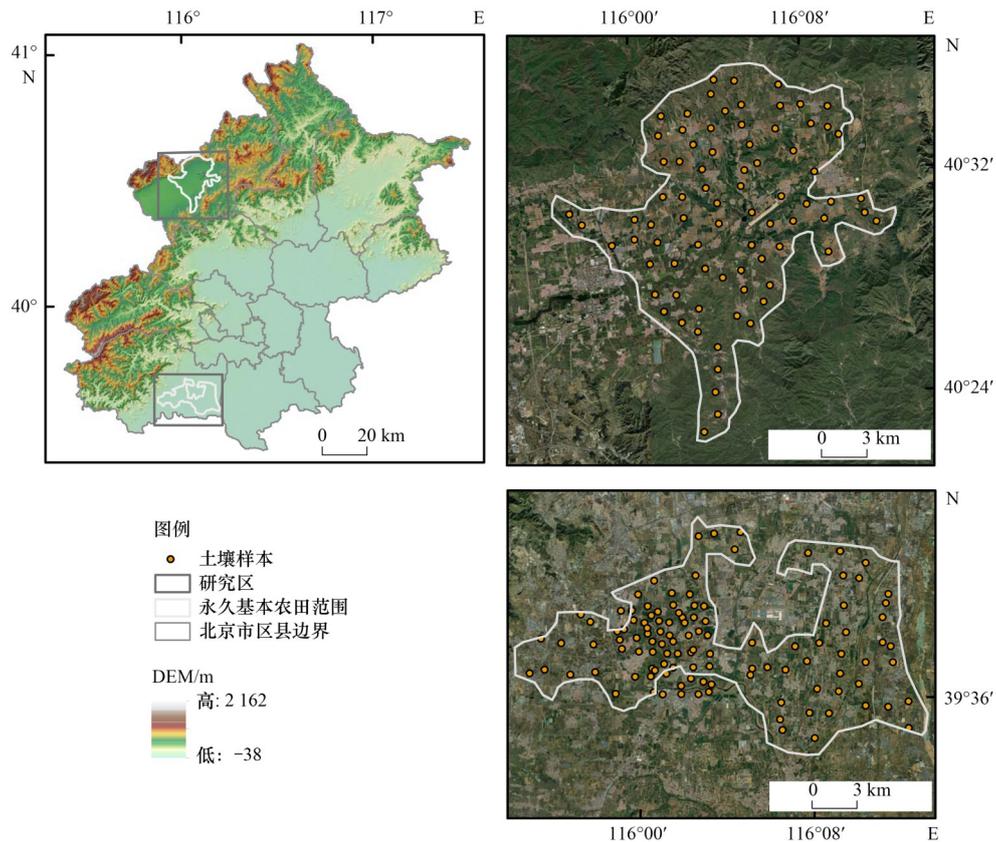


图1 研究区地理位置及土壤样本空间分布示意

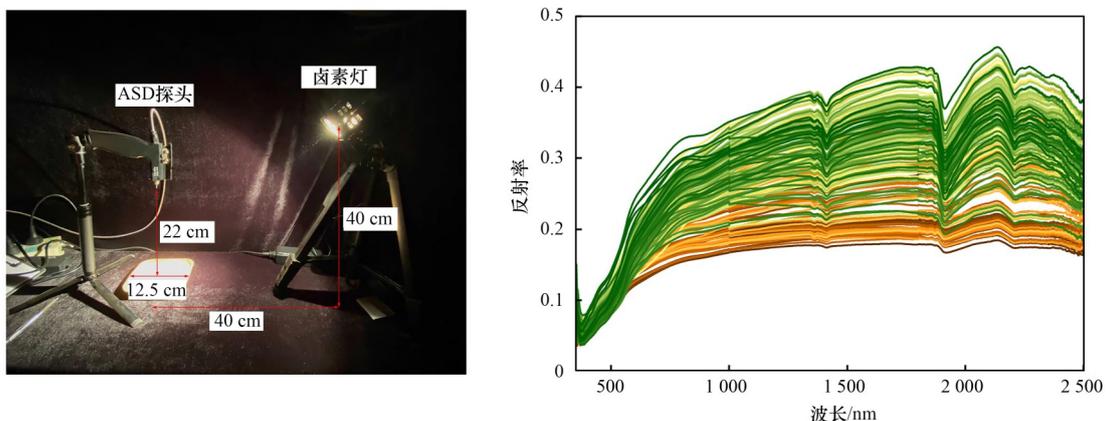
Fig. 1 Geographical location of study area and spatial distribution of soil samples

1.2 数据获取

2021年,结合遥感影像和永久基本农田分布数据,在研究区共采集193个土壤样品,采集深度为0~20 cm,空间位置如图1所示.将土壤样品自然风干后用孔径为2 mm的10目尼龙筛研磨和筛分,以分离大碎片、岩石和植物残留物.每个样品被平均分为2份,1份用于检测有机质含量,1份用于测量光谱反射率.

使用地物光谱仪FieldSpec 4采集土壤光谱反射率数据,光谱波段范围为350~2 500 nm,采样间隔

为1 nm.将土壤样品放置在直径为12.5 cm、深度为2.3 cm的黑色哑光器皿内,轻轻震动器皿使得土壤表面呈自然状态,以40 W卤素灯作为唯一光源,采用45°光源照射角度、22 cm探头距离及40 cm光源距离,从垂直于土壤表面的方向进行光谱测量,并利用40 cm×40 cm的白板进行定标,获取绝对反射率.为消除测量过程中的不稳定性,每个样品采集10条光谱曲线,利用ViewSpec Pro软件剔除异常曲线后取反射率平均值作为样品的原始光谱反射率(图2).



颜色由深绿至黄棕表示有机质含量由少至多

图2 土壤光谱实验室测量示意及土壤样品原始光谱反射率曲线

Fig. 2 Experimental apparatuses for hyperspectral reflectance measurement and raw spectral reflectance of soil samples

1.3 研究方法

结合连续小波变换、SHAP 和 XGBoost 的土壤有机质高光谱反演框架主要包括土壤光谱变换、特征波段筛选和有机质含量反演这 3 个步骤。首先, 利用连续小波变换方法对土壤原始光谱反射率数据进行多尺度分解, 得到各尺度下各波段的小波系数; 然后, 根据十折交叉验证方法将小波系数和有机质含量数据划分为训练集和验证集, 基于训练集中全部波段小波系数和有机质含量数据, 构建初始 XGBoost 模型, 并应用 SHAP 方法分析模型中各波段小波系数对有机质含量的贡献度, 以此为依据筛选特征波段; 最后, 基于验证集中特征波段小波系数和有机质含量数据, 重新构建和优化 XGBoost 模型, 实现有机质含量高光谱反演和精度验证。

其中十折交叉验证方法步骤如下: 首先, 在无先验知识的条件下将数据集随机分为 10 个大小一致的互斥子集, 随机划分 10 次; 然后, 对于每次划分的数据集, 轮流将其中 9 个子集的并集作为训练集, 剩余 1 个子集作为验证集; 最后, 求取 10 次模型输出值的平均值作为最终输出值^[39,40]。

1.3.1 连续小波变换

连续小波变换是一种时频分析方法, 能通过“伸缩”和“平移”自动适应时频信号分析的要求, 可聚焦到信号的任意细节^[41]。连续小波变换通过小波基函数, 将土壤光谱反射率在不同尺度下进行分解, 生成一系列小波系数, 计算公式如下:

$$w_{ij} = \int_{-\infty}^{+\infty} v_{ij} \Psi_{a,b}(j) dj \quad (1)$$

式中, w_{ij} 和 v_{ij} 分别表示第 i 个土壤样本的第 j 个波段的小波系数和反射率, a 表示尺度因子, b 表示平移因子, $\Psi_{a,b}(j)$ 表示小波基函数, 计算公式如下:

$$\Psi_{a,b}(j) = \frac{1}{\sqrt{a}} \Psi\left(\frac{j-b}{a}\right) \quad (2)$$

由于土壤光谱曲线特征与 Gaussian 函数相似^[42], 因此本研究采用 Gaus4 函数作为小波基函数。

1.3.2 SHAP

SHAP 的核心是计算每个特征的 Shaply 值, 以此来反映特征对模型输出的贡献度^[43]。运用 SHAP 进行特征波段筛选的步骤如下:

(1) 计算各土壤样本各波段小波系数对有机质含量的贡献度, 即 Shaply 值, 计算公式如下:

$$S_{ij} = \sum_{Q \in N/\{j\}} \frac{q!(n-q-1)!}{n!} [\hat{y}_i(w_i^{Q \cup \{j\}}) - \hat{y}_i(w_i^Q)] \quad (3)$$

式中, S_{ij} 表示第 i 个土壤样本的第 j 个波段的小波系数 w_{ij} 和第 i 个土壤样本的有机质含量 y_i 之间的 Shaply 值, N 表示所有波段的集合, 共 n 个波段, Q 表示集

合 N 中部分波段(至少 1 个波段)组成的子集, q 表示子集 Q 中包含的波段数量, $Q \in N$, $\hat{y}_i(w_i^Q)$ 表示在不加入波段 j 时, 使用子集 Q 中的波段构建模型得到的第 i 个土壤样本的有机质含量反演值, $\hat{y}_i(w_i^{Q \cup \{j\}})$ 表示在加入波段 j 后, 对应模型得到的反演值。

由于保留波段 j 的结果取决于模型之前已输入的其他波段, 因此需遍历所有可能的子集 $Q \in N/\{j\}$, 分别计算波段 j 加入前后的模型反演值之差 $\hat{y}_i(w_i^{Q \cup \{j\}}) - \hat{y}_i(w_i^Q)$, 加权求和后作为第 i 个土壤样本第 j 个波段的 Shaply 值。

(2) 对所有土壤样本的第 j 个波段的 Shaply 值取均值得到该波段对有机质含量的贡献度, 作为筛选特征波段的依据。

1.3.3 XGBoost

XGBoost 算法是由 Chen 等^[34]于 2016 年提出的一种集成学习算法, 涉及加法模型(强学习器由一系列弱学习器线性相加而成)和前向分布算法(下一轮迭代产生的新学习器是在上一轮的基础上训练得到的)两个关键部分。XGBoost 算法的基学习器是决策树, 将所有决策树的结果相加作为模型最终输出, 计算公式如下:

$$\hat{y}_i = \sum_{p=1}^P f_p(W_i), f_p \in F \quad (4)$$

式中, \hat{y}_i 表示第 i 个土壤样本的有机质含量反演值, f_p 表示第 p 棵决策树, P 表示决策树的总数, F 表示所有决策树的集合, W_i 表示第 i 个土壤样本的小波系数数据集。

XGBoost 算法目标函数计算公式如下:

$$L^{(t)} = \sum_{i=1}^n l[y_i, \hat{y}_i^{(t-1)} + f_i(W_i)] + \Omega(f_i) \quad (5)$$

式中, $\hat{y}_i^{(t-1)}$ 表示第 $t-1$ 次迭代时第 i 个土壤样本的有机质含量反演值, l 表示损失函数, 衡量有机质含量反演值 \hat{y}_i 和真实值 y_i 之间的误差, Ω 表示正则化函数, 防止模型过拟合^[44]。

2 结果与分析

2.1 土壤有机质含量统计及光谱特征分析

研究区 193 个土壤样本中 ω (有机质)最大值为 $45.86 \text{ g} \cdot \text{kg}^{-1}$, 最小值为 $5.17 \text{ g} \cdot \text{kg}^{-1}$, 平均值为 $20.64 \text{ g} \cdot \text{kg}^{-1}$ 。根据全国第二次土壤普查养分分级标准, 研究区有机质含量大多处于中等水平, 其中, 达到一级 ($> 40 \text{ g} \cdot \text{kg}^{-1}$) 的样本数为 2 个, 二级 ($30 \sim 40 \text{ g} \cdot \text{kg}^{-1}$) 的为 25 个, 三级 ($20 \sim 30 \text{ g} \cdot \text{kg}^{-1}$) 的为 62 个, 四级 ($10 \sim 20 \text{ g} \cdot \text{kg}^{-1}$) 的为 93 个, 五级 ($6 \sim 10 \text{ g} \cdot \text{kg}^{-1}$) 的为 9 个, 六级 ($< 6 \text{ g} \cdot \text{kg}^{-1}$) 的为 2 个。有机质含量空间分布

如图3所示,可以看到,房山区永久基本农田范围内有机质含量整体高于延庆区,房山区有机质含量

空间分布严重不均,呈现西部高东部低的特征,延庆区则相对均匀.

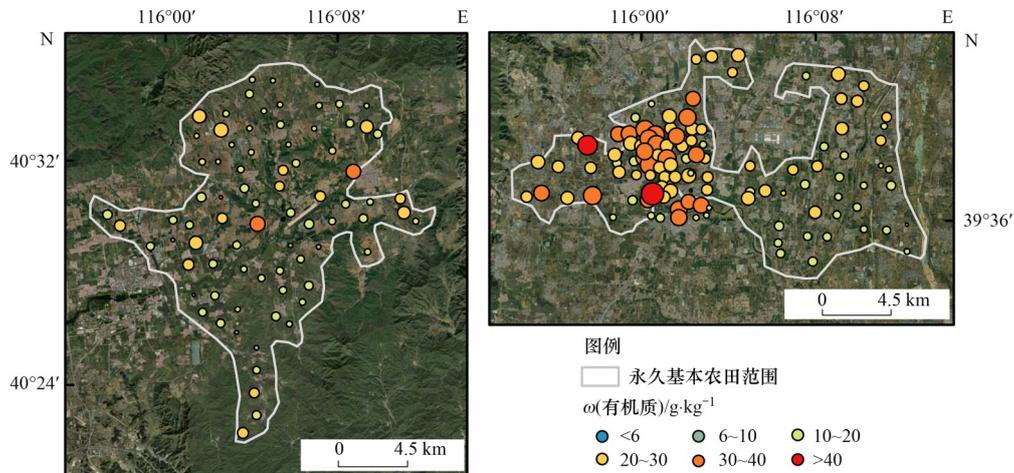


图3 研究区土壤有机质含量空间分布

Fig. 3 Spatial distribution of soil organic matter content in study area

为探究不同级别土壤有机质含量光谱反射率特征,求取每个级别土壤样本光谱反射率平均值,绘制曲线如图4所示.可以发现,不同级别的光谱反射率曲线大致走向类似,在400~1 000 nm区间,随着波长的增加,反射率呈现上升趋势;在1 000 nm以后,除水分吸收谷(1 400、1 900和2 200 nm附近)外,曲线较为平稳.整体上,较高的有机质含量呈现出较低的反射率,但从一级和二级、五级和六级的有机质含量反射率曲线可以看出,二者之间并非绝对的负相关关系,有机质含量较低的情况下也可能呈现相对较高的反射率.

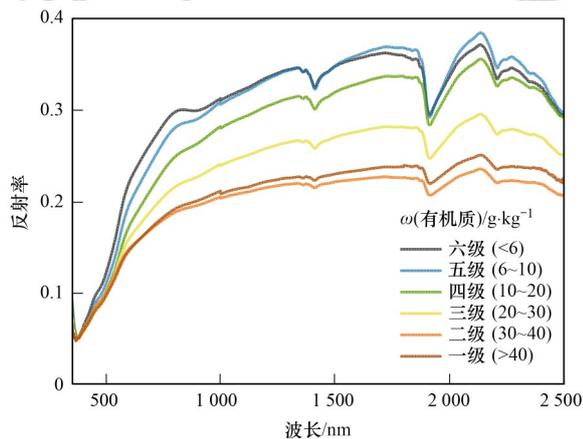


图4 不同级别土壤有机质含量光谱曲线

Fig. 4 Spectral curves of soil organic matter content at different levels

2.2 小波系数与土壤有机质含量相关性分析

选取 2^n ($n = 1, 2, 3, \dots, 6$)共6个尺度,对400~2 400 nm原始光谱数据进行连续小波变换,生成的小波系数如图5所示. 2^1 和 2^2 尺度光谱响应相似且较弱,随着尺度不断增大,光谱响应差异愈加明显,但中高尺度例如 2^5 和 2^6 尺度光谱弧度出现越来越

平滑的趋势,细节信息逐渐消失.

利用SHAP方法分别计算6种尺度的小波系数和有机质含量之间的Shaply值,结果如图6所示. $2^1 \sim 2^3$ 尺度Shaply值较大的波段主要分布在1 380~1 440 nm附近, 2^4 尺度主要分布在644、1 469和1 924 nm附近, 2^5 尺度主要分布在789 nm和1 938 nm附近,而 2^6 尺度则主要分布在713 nm和1 639 nm附近.各尺度Shaply值较大的波段在可见光和近红外波段中均有分布,大多在近红外波段,这是由于北京地区旱作农田有机质主要源于农作物残体,其由糖类化合物、含氮化合物和纤维素等组成,这些成分中C—H键、C—O键、N—O键和N—H键等的光谱响应波段位于近红外区域^[45].

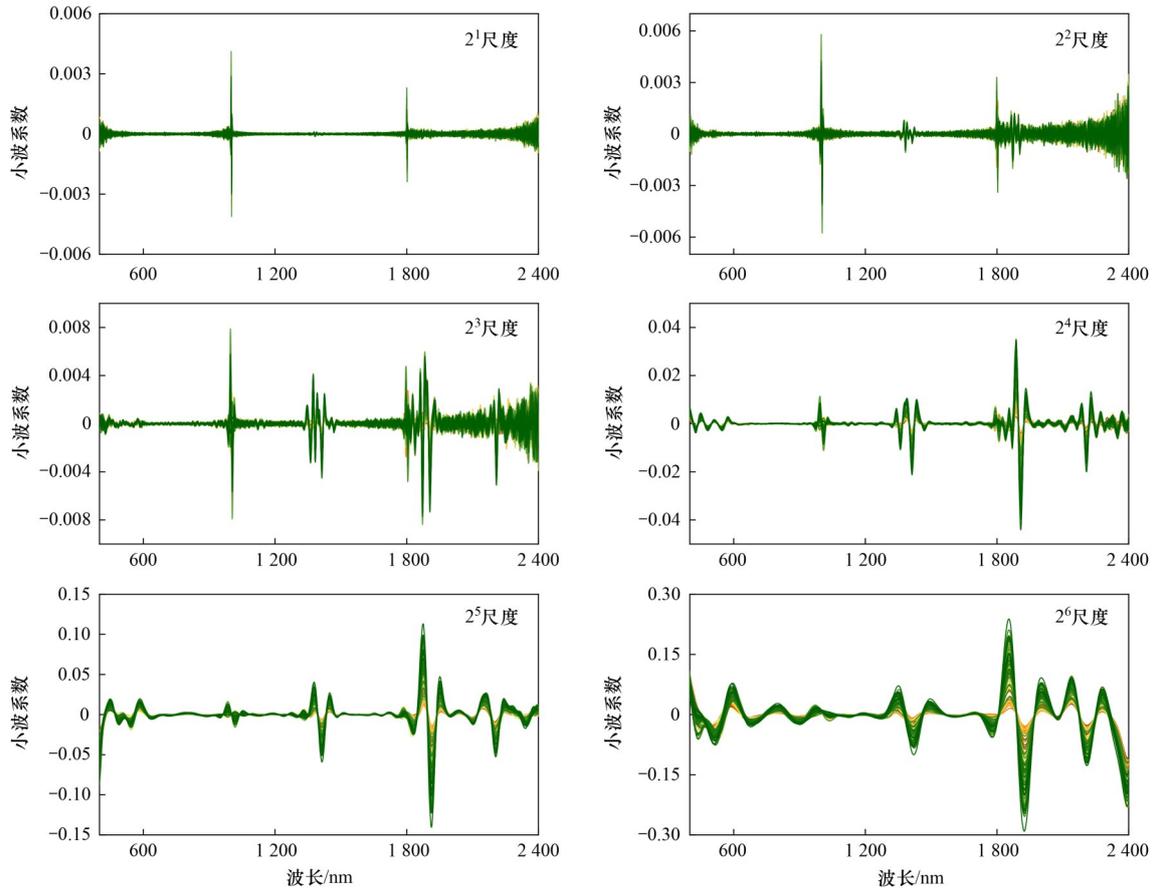
2.3 土壤有机质含量高光谱反演与分析

2.3.1 特征波段选择

基于训练集,将各尺度下全部波段的小波系数作为自变量,有机质含量作为因变量,构建初始XGBoost模型.从Shaply绝对值最大的波段开始,依次加入绝对值大于0的波段参与模型构建,选取均方根误差(root mean square error, RMSE)作为损失函数,以RMSE值达到最小为标准,获取各尺度下特征波段的最佳数量.从图7中可以看出,随着波段数量的增加, RMSE值迅速下降,其中除 2^5 尺度在RMSE值达到最小后趋于稳定外,所有尺度在RMSE值达到最小后呈现明显上升趋势. $2^1 \sim 2^6$ 尺度下分别选取23、100、45、73、40和23个波段.

2.3.2 模型参数优化与训练过程

基于各尺度下选取的特征波段小波系数对XGBoost模型进行优化.XGBoost模型的重要参数如



颜色由深绿至黄棕表示有机质含量由少至多

图5 2¹~2⁶尺度下的小波系数

Fig. 5 Wavelet coefficients at scales ranging from 2¹ to 2⁶

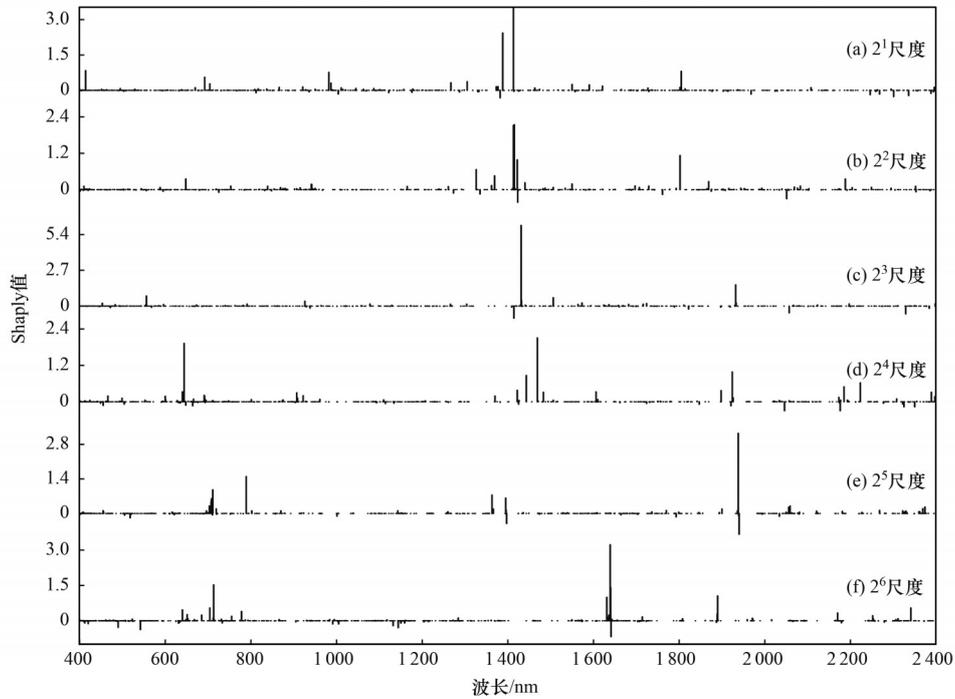


图6 土壤有机质含量与小波系数之间的 Shaply 值

Fig. 6 Shaply value between soil organic matter content and wavelet coefficients

表1所示，以最小化RMSE值为模型训练目标，确定最佳参数组合，训练过程如图8所示。

2.3.3 模型结果与精度验证

基于验证集，采用决定系数(R^2)、平均绝对误

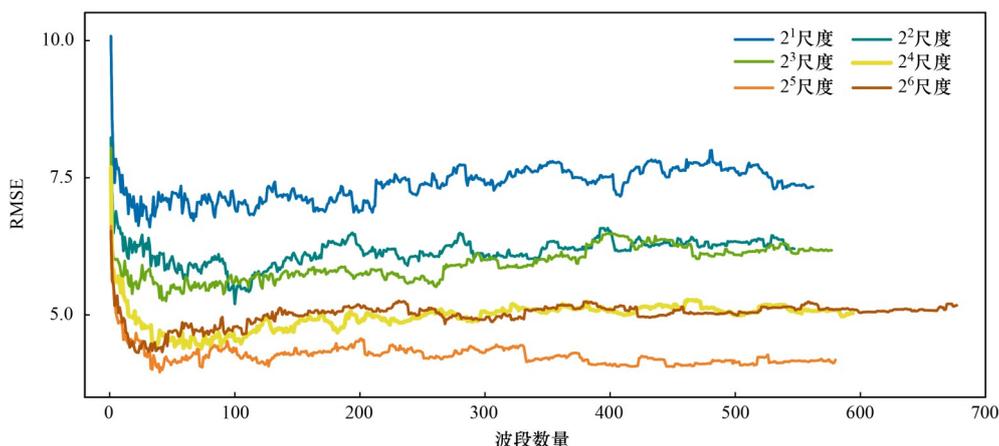


图 7 SHAP方法筛选特征波段的损失函数曲线

Fig. 7 Loss function curve of characteristic bands selected by SHAP method

表 1 XGBoost模型参数

Table 1 Parameters of XGBoost model

参数	释义	取值范围	优选值
max_depth	树的最大深度	[1,20]	3
min_child_weight	叶节点权重的最小总和	[1,20]	7
gamma	树的叶节点进一步划分所需的最小损失减少	[0,1]	0
subsample	对样本随机采样的比例	(0,1]	0.8
colsample_bytree	对特征随机采样的比例	(0,1]	0.6
alpha	L1正则项的权重	[0,1]	0.3
eta	学习率	(0,0.2]	0.1
n_estimators	决策树的数量	(0,200]	140

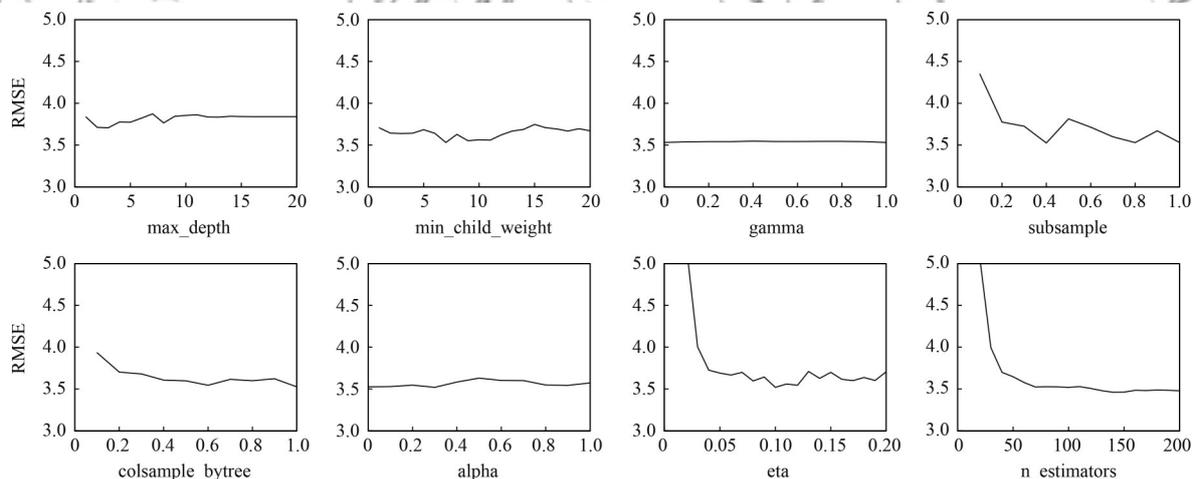


图 8 XGBoost模型参数优化过程

Fig. 8 Optimization process of XGBoost model parameters

差(mean absolute error, MAE)、RMSE 和 RPD 对各尺度下的模型性能进行评估, R^2 越接近 1, MAE 和 RMSE 越小, RPD 越大, 模型性能越好. 模型精度统计结果以及有机质含量实测值和模型反演值散点图如图 9 所示. 6种尺度模型精度由高到低依次为: $2^5 > 2^6 > 2^4 > 2^3 > 2^2 > 2^1$. 其中 2^5 尺度下模型精度最高, R^2 达到 0.80, MAE、RMSE 和 RPD 分别为

$2.72 \text{ g} \cdot \text{kg}^{-1}$ 、 $3.60 \text{ g} \cdot \text{kg}^{-1}$ 和 1.94, R^2 和 RPD 相较于 2^1 尺度分别提高了 0.37 和 1.11, MAE 和 RMSE 分别降低 $2.10 \text{ g} \cdot \text{kg}^{-1}$ 和 $2.49 \text{ g} \cdot \text{kg}^{-1}$. 整体看, 随着尺度的增加, 模型性能也逐步提升, 但尺度超过 2^5 后, 模型性能呈现减弱趋势, 这是由于有机质成分复杂多样, 各成分均有不同的光谱相应范围, 如果尺度过高则不能充分利用有机质各成分的光谱信息,

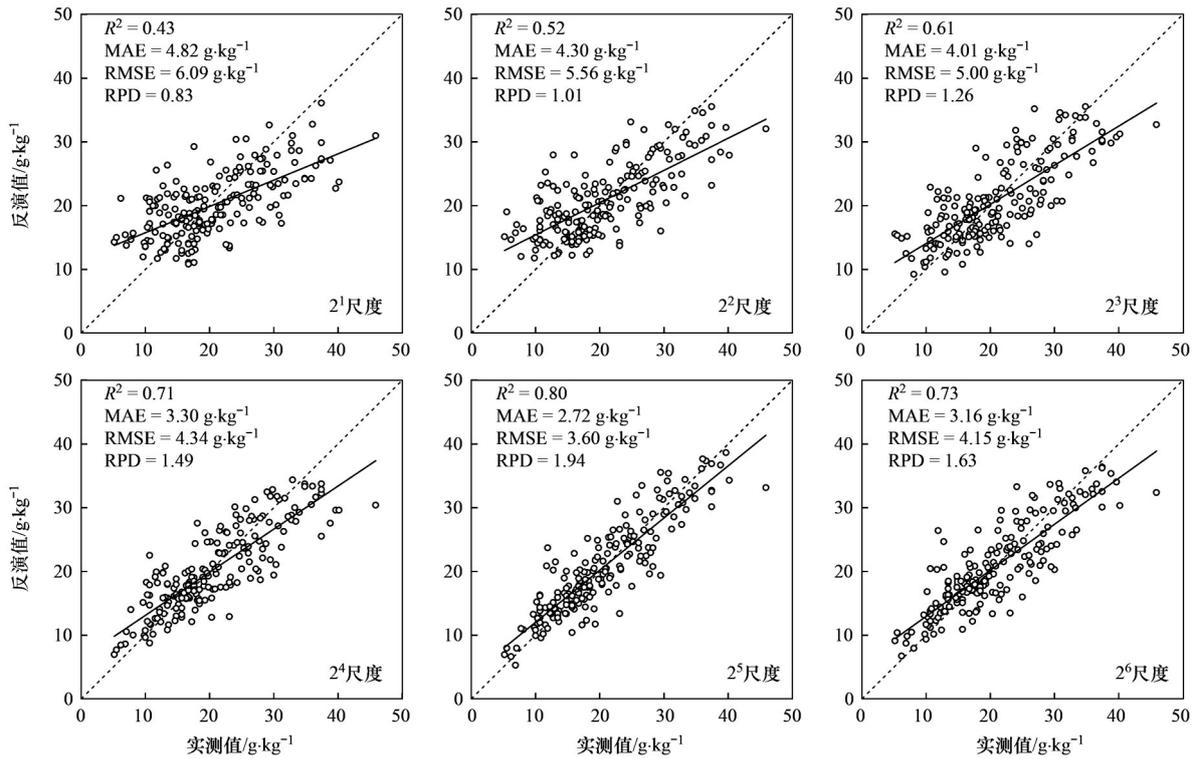


图9 2^1 ~ 2^6 尺度下有机质含量反演值-实测值散点图

Fig. 9 Scatter plot of inverted and measured values of organic matter content at scales ranging from 2^1 to 2^6

而尺度较低时会降低信噪比，进而对模型产生负面影响。

3 讨论

3.1 SHAP与Pearson相关分析法对比

为验证SHAP方法在特征波段筛选方面的优势，与应用最多的Pearson相关分析法进行对比，图10为 2^5 尺度下各波段对应的Pearson相关系数和Shaply值。可以看到，Pearson相关系数较大的波段数目众多且连续，而Shaply值较大的波段数量较少，彼此间隔较远。

以RMSE值达到最小为标准，确定Pearson相关分析法选取的特征波段数量，将达到极显著水平

($P < 0.01$)且Pearson相关系数绝对值大于0.5的波段按照从大至小的顺序依次加入初始XGBoost模型中，RMSE随波段数量的变化如图11所示。当波段数量达到722时，RMSE达到最小，与SHAP方法相比，波段数量多682个，RMSE却高0.69 $\text{g}\cdot\text{kg}^{-1}$ 。图12为采用这两种方法选取的特征波段以及不筛选特征波段的建模结果，可以发现，通过SHAP和Pearson方法筛选的特征波段和不筛选特征波段相比均对模型性能有提升作用。由于Pearson选取的波段数量众多，相邻波段之间信息冗余严重，增加了模型复杂度， R^2 仅从0.67提升至0.72。而SHAP方法有效选取了对有机质含量高光谱反演有重要作用的特征波段，在波段数量较

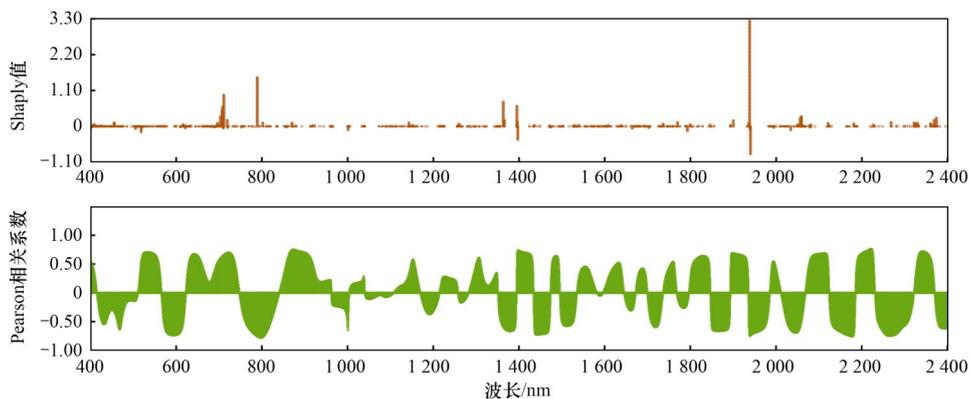


图10 2^5 尺度小波系数与有机质含量之间的Shaply值和Pearson相关系数

Fig. 10 Shaply value and Pearson correlation coefficient between wavelet coefficients at 2^5 scale and organic matter content

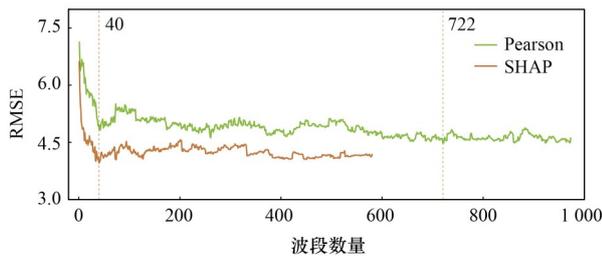


图 11 SHAP 和 Pearson 方法筛选特征波段的损失函数曲线

Fig. 11 Loss function curve of characteristic bands selected by SHAP and Pearson methods

少的情况下, 还提升了模型精度, R^2 从 0.67 提升至 0.80.

3.2 XGBoost 与 SVM 和 RF 模型对比

为验证 XGBoost 模型在土壤有机质含量高光谱反演中的性能, 与应用较多的 RF 和 SVM 模型进行对比. 图 13 和图 14 分别为基于 2^5 尺度下 SHAP 方法选取的特征波段构建的 SVM、RF 和 XGBoost 模型的

有机质含量反演值-实测值散点图 and 对比. 可以发现, XGBoost 模型精度最高, RF 次之, SVM 最低, XGBoost 模型的 R^2 比 RF 和 SVM 模型分别提升 4% 和 8%. 图 14 显示部分样本的有机质含量模型反演值和实测值相差较大, 以样本 50 为例, 分析其原因. 样本 50 与样本 124、68 和 182 的有机质含量相似, 但从图 15 中可以看出, 其光谱反射率和小波系数曲线却与样本 61、5 和 89 接近, 两组样本的有机质含量相差约 $11 \text{ g} \cdot \text{kg}^{-1}$. 研究区内有机质含量较高的土壤通常呈现出较低的反射率(图 4), 但也存在不符合上述规律的样本(例如: 样本 50), 当这种样本数量较少时, 模型将不能很好地学习到二者之间的关系. 因此, 在未来的研究中, 应当适当增加此类复杂样本的数量. 此外, 在今后的光谱采集工作中应考虑土壤二向反射特性, 通过旋转样本实现从多个角度采集光谱, 从而更好地反映土壤光谱特征.

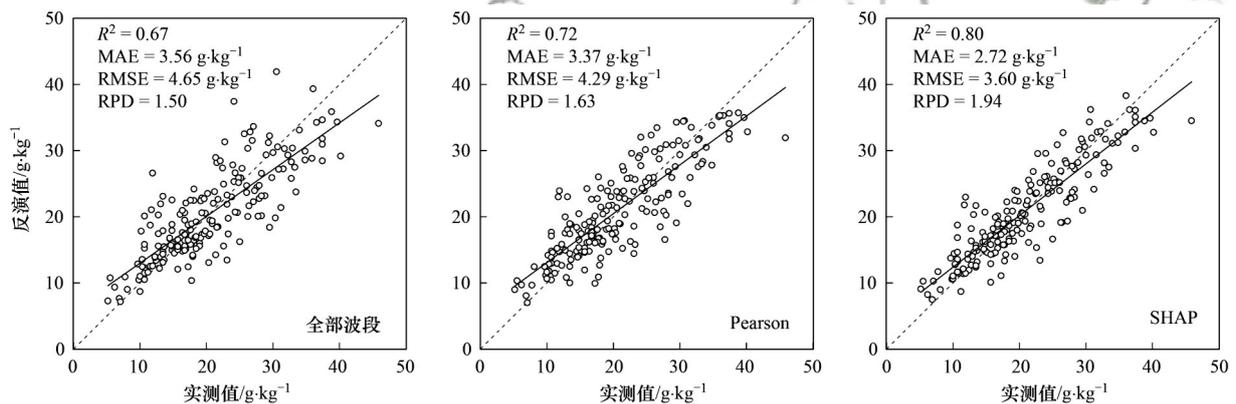


图 12 2^5 尺度下特征波段和全部波段的有机质含量反演值-实测值散点图

Fig. 12 Scatter diagram of inverted and measured values of organic matter content for characteristic bands and all bands at 2^5 scale

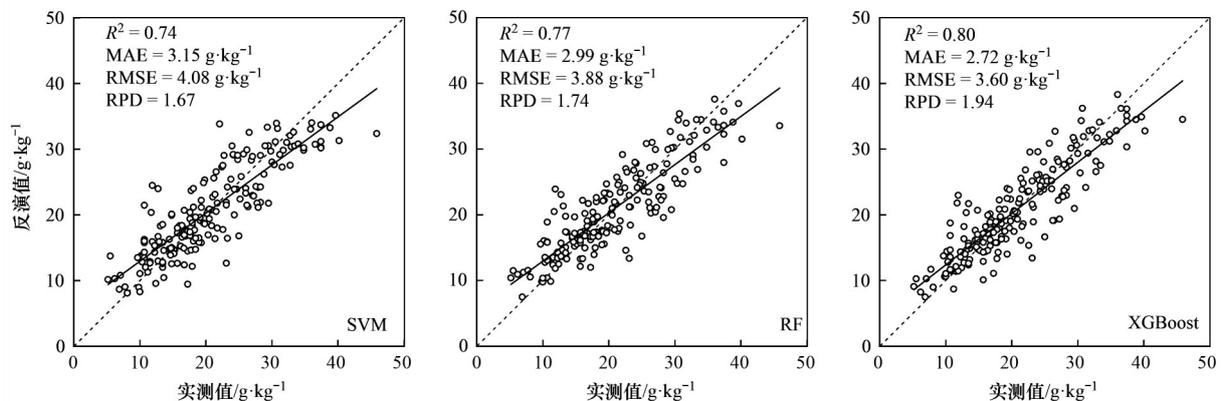


图 13 2^5 尺度下不同模型的有机质含量反演值-实测值散点图

Fig. 13 Scatter diagram of inverted and measured values of organic matter content for models at 2^5 scale

4 结论

(1) 研究区土壤有机质含量大多处于中等水平

($10\sim 30 \text{ g} \cdot \text{kg}^{-1}$), 整体上延庆区有机质含量低于房山区; 延庆区有机质含量空间分布相对均匀, 房山区则呈现西部高东部低的特征.

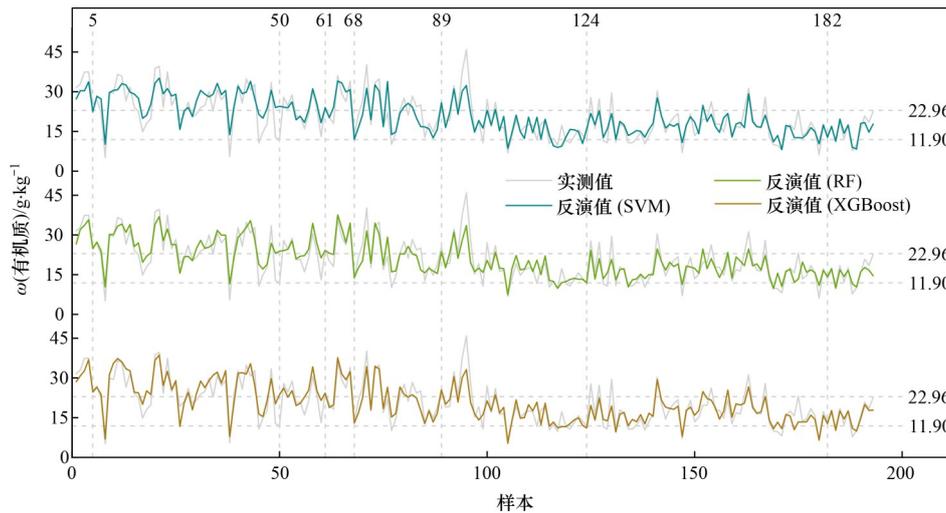


图 14 2⁵尺度下不同模型的有机质含量反演值-实测值对比

Fig. 14 Comparison of inversed and measured values of organic matter content for models at 2⁵ scale

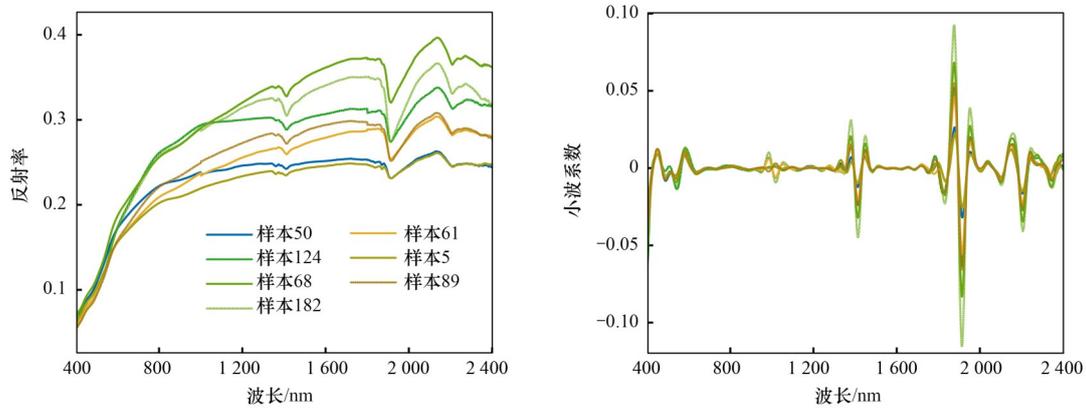


图 15 样本 50 的有机质含量反演值与实测值对应样本的原始光谱反射率及 2⁵ 尺度下的小波系数
Fig. 15 Raw spectral reflectance and wavelet coefficients at 2⁵ scale for the samples corresponding to the inversed and the measured value of sample 50

(2)有机质含量反演精度随着连续小波变换尺度的增大呈现先升高再降低的趋势，其中 2⁵ 尺度精度最高，有机质含量反演值和实测值之间的 R² 比 2¹ 尺度大 0.37，表明中高尺度连续小波变换能有效去除光谱噪声并放大光谱细节特征。

(3)基于 SHAP 方法选取的 40 个特征波段构建的反演模型 R² 达到 0.80，而基于 Pearson 相关分析法选取的 722 个波段构建的模型 R² 仅达到 0.72，表明 SHAP 方法比 Pearson 相关分析法更能有效去除冗余光谱信息，在特征波段筛选中具有可行性和优越性。

(4)XGBoost 相较于 RF 和 SVM 模型在土壤有机质含量高光谱反演中具有最好的性能。表明结合连续小波变换、SHAP 和 XGBoost 的土壤有机质含量高光谱反演框架可以为土壤有机质含量快速监测提供科学依据。

参考文献：

[1] 汪景宽, 徐英德, 丁凡, 等. 植物残体向土壤有机质转化过程及其稳定机制的研究进展[J]. 土壤学报, 2019, 56(3):

528-540.

Wang J K, Xu Y D, Ding F, *et al.* Process of plant residue transforming into soil organic matter and mechanism of its stabilization: a review[J]. Acta Pedologica Sinica, 2019, 56(3): 528-540.

[2] Yazdanshenas H, Tavili A, Jafari M, *et al.* Evidence for relationship between carbon storage and surface cover characteristics of soil in rangelands [J]. CATENA, 2018, 167: 139-146.

[3] Zhang Z S, Wang J J, Lyu X, *et al.* Impacts of land use change on soil organic matter chemistry in the Everglades, Florida—a characterization with pyrolysis-gas chromatography - mass spectrometry[J]. Geoderma, 2019, 338: 393-400.

[4] Picariello E, Baldantoni D, Izzo F, *et al.* Soil organic matter stability and microbial community in relation to different plant cover: A focus on forests characterizing Mediterranean area [J]. Applied Soil Ecology, 2021, 162, doi: 10.1016/j.apsoil.2021.103897.

[5] 储成, 吴赵越, 黄欠如, 等. 有机质提升对酸性红壤氮循环功能基因及功能微生物的影响[J]. 环境科学, 2020, 41(5): 2468-2475.

Chu C, Wu Z Y, Huang Q R, *et al.* Effect of organic matter promotion on nitrogen-cycling genes and functional microorganisms

- in acidic red soils [J]. *Environmental Science*, 2020, **41**(5): 2468-2475.
- [6] 鲍士旦. 土壤农化分析[M]. (第三版). 北京: 中国农业出版社, 2000.
- [7] Horta A, Azevedo L, Neves J, *et al.* Integrating portable X-ray fluorescence (pXRF) measurement uncertainty for accurate soil contamination mapping[J]. *Geoderma*, 2021, **382**, doi: 10.1016/j.geoderma.2020.114712.
- [8] Nawar S, Buddenbaum H, Hill J, *et al.* Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy [J]. *Soil and Tillage Research*, 2016, **155**: 510-522.
- [9] Conforti M, Castrignanò A, Robustelli G, *et al.* Laboratory-based Vis-NIR spectroscopy and partial least square regression with spatially correlated errors for predicting spatial variation of soil organic matter content[J]. *CATENA*, 2015, **124**: 60-67.
- [10] Sun W C, Liu S, Zhang X, *et al.* Estimation of soil organic matter content using selected spectral subset of hyperspectral data [J]. *Geoderma*, 2022, **409**, doi: 10.1016/j.geoderma.2021.115653.
- [11] Nowkandeh S M, Noroozi A A, Homaei M. Estimating soil organic matter content from Hyperion reflectance images using PLSR, PCR, MinR and SWR models in semi-arid regions of Iran [J]. *Environmental Development*, 2018, **25**: 23-32.
- [12] 于雷, 洪永胜, 周勇, 等. 连续小波变换高光谱数据的土壤有机质含量反演模型构建[J]. *光谱学与光谱分析*, 2016, **36**(5): 1428-1433.
- Yu L, Hong Y S, Zhou Y, *et al.* Inversion of soil organic matter content using hyperspectral data based on continuous wavelet transformation [J]. *Spectroscopy and Spectral Analysis*, 2016, **36**(5): 1428-1433.
- [13] 钟浩, 李西灿, 翟浩然, 等. 耕层土壤有机质高光谱间接估测模型[J]. *测绘科学技术学报*, 2019, **36**(1): 74-78, 85.
- Zhong H, Li X C, Zhai H R, *et al.* Hyperspectral indirect estimation model of soil organic matter content in plough layer [J]. *Journal of Geomatics Science and Technology*, 2019, **36**(1): 74-78, 85.
- [14] 裴哲, 李秀芬, 吕家欣, 等. 东北典型黑土区表层土壤有机质含量高光谱反演研究[J]. *土壤通报*, 2019, **50**(6): 1285-1293.
- Nie Z, Li X F, Lv J X, *et al.* Hyperspectral retrieval of surface soil organic matter content in a typical black soil region of northeast China [J]. *Chinese Journal of Soil Science*, 2019, **50**(6): 1285-1293.
- [15] 勾宇轩, 赵云泽, 李勇, 等. 基于CWT-sCARS的东北旱作农田土壤有机质高光谱反演[J]. *农业机械学报*, 2022, **53**(3): 331-337.
- Gou Y X, Zhao Y Z, Li Y, *et al.* Soil organic matter content in dryland farmland in northeast China with hyperspectral reflectance based on CWT-sCARS [J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2022, **53**(3): 331-337.
- [16] 赵海龙, 甘淑, 袁希平, 等. 基于多尺度连续小波分解的土壤氧化铁反演[J]. *光学学报*, 2022, **42**(22), doi: 10.3788/AOS202242.2230002.
- Zhao H L, Gan S, Yuan X P, *et al.* Inversion of soil iron oxide based on multi-scale continuous wavelet decomposition [J]. *Acta Optica Sinica*, 2022, **42**(22), doi: 10.3788/AOS202242.2230002.
- [17] 玉米提·买明, 王雪梅. 连续小波变换的土壤有机质含量高光谱估测[J]. *光谱学与光谱分析*, 2022, **42**(4): 1278-1284.
- Yumiti M M, Wang X M. Hyperspectral estimation of soil organic matter content based on continuous wavelet transformation [J]. *Spectroscopy and Spectral Analysis*, 2022, **42**(4): 1278-1284.
- [18] Tan Y, Jiang Q G, Yu L F, *et al.* Reducing the moisture effect and improving the prediction of soil organic matter with VIS-NIR spectroscopy in black soil area [J]. *IEEE Access*, 2021, **9**: 5895-5905.
- [19] 王荐一, 杨雯, 王玉川, 等. 辽宁省黄土状母质发育土壤有机质含量高光谱预测模型的构建[J]. *土壤通报*, 2022, **53**(6): 1320-1330.
- Wang J Y, Yang W, Wang Y C, *et al.* A hyperspectral prediction model for organic matter content in soil developed from loess-like parent material in Liaoning Province [J]. *Chinese Journal of Soil Science*, 2022, **53**(6): 1320-1330.
- [20] 赵瑞, 崔希民, 刘超. GF-5 高光谱遥感影像的土壤有机质含量反演估算研究[J]. *中国环境科学*, 2020, **40**(8): 3539-3545.
- Zhao R, Cui X M, Liu C. Inversion estimation of soil organic matter content based on GF-5 hyperspectral remote sensing image [J]. *China Environmental Science*, 2020, **40**(8): 3539-3545.
- [21] 章涛, 于雷, 易军, 等. 高光谱小波能量特征估测土壤有机质含量[J]. *光谱学与光谱分析*, 2019, **39**(10): 3217-3222.
- Zhang T, Yu L, Yi J, *et al.* Determination of soil organic matter content based on hyperspectral wavelet energy features [J]. *Spectroscopy and Spectral Analysis*, 2019, **39**(10): 3217-3222.
- [22] 尚天浩, 毛鸿欣, 张俊华, 等. 基于PCA敏感波段筛选与SVM建模的银川平原土壤有机质高光谱估算[J]. *生态学杂志*, 2021, **40**(12): 4128-4136.
- Shang T H, Mao H X, Zhang J H, *et al.* Hyperspectral estimation of soil organic matter content in Yinchuan plain, China based on PCA sensitive band screening and SVM modeling [J]. *Chinese Journal of Ecology*, 2021, **40**(12): 4128-4136.
- [23] Lundberg S M, Lee S I. A unified approach to interpreting model predictions [A]. *Proceedings of the 31st International Conference on Neural Information Processing Systems* [C]. Long Beach: Curran Associates Inc., 2017. 4768-4777.
- [24] 赵健, 刘彦辰, 朱冰, 等. 基于SHAP-RF框架的越野车辆路面识别算法研究[J]. *力学学报*, 2022, **54**(10): 2922-2935.
- Zhao J, Liu Y C, Zhu B, *et al.* Research on road recognition algorithm of off-road vehicle based on SHAP-RF framework [J]. *Chinese Journal of Theoretical and Applied Mechanics*, 2022, **54**(10): 2922-2935.
- [25] Liu Y C, Liu Z H, Luo X, *et al.* Diagnosis of Parkinson's disease based on SHAP value feature selection [J]. *Biocybernetics and Biomedical Engineering*, 2022, **42**(3): 856-869.
- [26] 孙浩然, 赵志根, 赵佳星, 等. 珠海一号高光谱遥感的表层土壤有机质含量反演方法[J]. *遥感信息*, 2020, **35**(4): 40-46.
- Sun H R, Zhao Z G, Zhao J X, *et al.* Inversion of topsoil organic matter content by hyperspectral remote sensing of Zhuhai-1 [J]. *Remote Sensing Information*, 2020, **35**(4): 40-46.
- [27] 颜祥照, 姚艳敏, 张霄羽, 等. 星载高分五号高光谱耕地主要土壤类型土壤有机质含量估测——以黑龙江省建三江农垦区为例[J]. *中国土壤与肥料*, 2021, (5): 10-20.
- Yan X Z, Yao Y M, Zhang X Y, *et al.* Estimation of soil organic matter content in different soil types of cultivated land based on hyperspectral data of GF-5 satellite-A case study of Jiansanjiang reclamation area in Heilongjiang province [J]. *Soils and Fertilizers Sciences in China*, 2021, (5): 10-20.
- [28] 谭琨, 张倩倩, 曹茜, 等. 基于粒子群优化支持向量机的矿区土壤有机质含量高光谱反演[J]. *地球科学—中国地质大学学报*, 2015, **40**(8): 1339-1345.
- Tan K, Zhang Q Q, Cao Q, *et al.* Hyperspectral retrieval model of soil organic matter content based on particle swarm optimization-support vector machines [J]. *Earth Science—Journal of China*

- University of Geosciences, 2015, **40**(8): 1339-1345.
- [29] Meng X T, Bao Y L, Ye Q, *et al.* Soil organic matter prediction model with satellite hyperspectral image based on optimized denoising method[J]. *Remote Sensing*, 2021, **13**(12), doi: 10.3390/rs13122273.
- [30] 钟亮, 郭熙, 国佳欣, 等. 基于不同卷积神经网络模型的红壤有机质高光谱估算[J]. *农业工程学报*, 2021, **37**(1): 203-212.
- Zhong L, Guo X, Guo J X, *et al.* Hyperspectral estimation of organic matter in red soil using different convolutional neural network models [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2021, **37**(1): 203-212.
- [31] 陈玮, 徐占军, 郭琦. 煤炭矿区耕地土壤有机质无人机高光谱遥感估测[J]. *农业工程学报*, 2022, **38**(8): 98-106.
- Chen W, Xu Z J, Guo Q. Estimation of soil organic matter by UAV hyperspectral remote sensing in coal mining areas[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, **38**(8): 98-106.
- [32] 侯化刚, 王丹阳, 马斯琦, 等. 黄河三角洲不同盐渍度土壤有机质含量的高光谱预测研究[J]. *中国农业科学*, 2023, **56**(10): 1905-1919.
- Hou H G, Wang D Y, Ma S Q, *et al.* Hyperspectral prediction of organic matter in soils of different salinity levels in the Yellow River Delta[J]. *Scientia Agricultura Sinica*, 2023, **56**(10): 1905-1919.
- [33] 周伟, 谢利娟, 杨晗, 等. 基于高光谱的三江源区土壤有机质含量反演[J]. *土壤通报*, 2021, **52**(3): 564-574.
- Zhou W, Xie L J, Yang H, *et al.* Hyperspectral inversion of soil organic matter content in the Three-Rivers Source region [J]. *Chinese Journal of Soil Science*, 2021, **52**(3): 564-574.
- [34] Chen T Q, Guestrin C. XGBoost: A scalable tree boosting system [A]. *Proceedings of the 22nd ACM Sigkdd International Conference on Knowledge Discovery and Data Mining [C]*. San Francisco: ACM, 2016. 785-794.
- [35] Ye M, Zhu L, Li X J, *et al.* Estimation of the soil arsenic concentration using a geographically weighted XGBoost model based on hyperspectral data[J]. *Science of The Total Environment*, 2023, **858**, doi: 10.1016/j.scitotenv.2022.159798.
- [36] Ge X Y, Ding J L, Jin X L, *et al.* Estimating agricultural soil moisture content through UAV-based hyperspectral images in the arid region [J]. *Remote Sensing*, 2021, **13**(8), doi: 10.3390/rs13081562.
- [37] Sun Y S, Chen S S, Dai X M, *et al.* Coupled retrieval of heavy metal nickel concentration in agricultural soil from spaceborne hyperspectral imagery [J]. *Journal of Hazardous Materials*, 2023, **446**, doi: 10.1016/j.jhazmat.2023.130722.
- [38] 葛翔宇, 丁建丽, 王敬哲, 等. 一种基于无人机高光谱影像的土壤墒情检测新方法[J]. *光谱学与光谱分析*, 2020, **40**(2): 602-609.
- Ge X Y, Ding J L, Wang J Z, *et al.* A new method for predicting soil moisture based on UAV hyperspectral image [J]. *Spectroscopy and Spectral Analysis*, 2020, **40**(2): 602-609.
- [39] Geisser S. A predictive approach to the random effect model [J]. *Biometrika*, 1974, **61**(1): 101-107.
- [40] 孙成, 王卫, 刘方田, 等. 基于线性混合效应模型的河北省PM_{2.5}浓度时空变化模型研究[J]. *环境科学研究*, 2019, **32**(9): 1500-1509.
- Sun C, Wang W, Liu F T, *et al.* Spatial-temporal simulation of PM_{2.5} concentration in Hebei Province based on linear mixed effects model [J]. *Research of Environmental Sciences*, 2019, **32**(9): 1500-1509.
- [41] Farge M. Wavelet transforms and their applications to turbulence [J]. *Annual Review of Fluid Mechanics*, 1992, **24**: 395-458.
- [42] Zhang S W, Shen Q, Nie C J, *et al.* Hyperspectral inversion of heavy metal content in reclaimed soil from a mining wasteland based on different spectral transformation and modeling methods [J]. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2019, **211**: 393-400.
- [43] 董佳奇, 胡冬梅, 闫雨龙, 等. 基于可解释性机器学习的城市O₃驱动因素挖掘[J]. *环境科学*, 2023, **44**(7): 3660-3668.
- Dong J Q, Hu D M, Yan Y L, *et al.* Revealing driving factors of urban O₃ based on explainable machine learning [J]. *Environmental Science*, 2023, **44**(7): 3660-3668.
- [44] 吴迪, 杜宁, 王莉, 等. 基于GTWR-XGBoost模型的四川省PM_{2.5}小时浓度估算[J]. *环境科学*, 2023, **44**(7): 3738-3748.
- Wu D, Du N, Wang L, *et al.* Estimation of PM_{2.5} hourly concentration in Sichuan province based on GTWR-XGBoost model [J]. *Environmental Science*, 2023, **44**(7): 3738-3748.
- [45] 王延仓, 杨贵军, 朱金山, 等. 基于小波变换与偏最小二乘耦合模型估测北方潮土有机质含量[J]. *光谱学与光谱分析*, 2014, **34**(7): 1922-1926.
- Wang Y C, Yang G J, Zhu J S, *et al.* Estimation of organic matter content of north fluvo-aquic Soil based on the coupling model of wavelet transform and partial least squares [J]. *Spectroscopy and Spectral Analysis*, 2014, **34**(7): 1922-1926.

CONTENTS

Spatio-temporal Evolution, Dynamic Transition, and Convergence Trend of Urban Carbon Emission Intensity in China	YANG Qing-ke, WANG Lei, ZHU Gao-li, <i>et al.</i> (1869)
Spatio-temporal Evolution and Trend Prediction of Transport Carbon Emission Efficiency	ZHENG Yan, JIANG Xue-mei, XIAO Yu-jie (1879)
Research on Carbon Emission Decoupling Factors Based on STIRPAT Model and LMDI Decomposition	ZHANG Jiang-yan (1888)
Carbon Peak Pathways of Industrial Parks Based on the LEAP Model: A Case Study of a National Development Zone in Nanjing	LI Hui-peng, LI Li, YIN Yin, <i>et al.</i> (1898)
Accounting and Analysis of Carbon Emissions in Universities: A Case Study of Beijing A University	CAO Rui, FENG Li, ZHANG Li-qiu (1907)
Synergy Effects and Driving Factors of Pollution and Carbon Emission Reduction in Manufacturing Industry in Beijing	YU Shan, HAN Yu-hua, MU Jie, <i>et al.</i> (1917)
Carbon Emissions Trading Prices and Total Factor Productivity: Evidence from China	WU Xue-ping (1926)
Analysis of Spatio-temporal Distribution Characteristics and Influencing Factors of PM _{2.5} and PM ₁₀ in Chinese Cities	LI Jiang-su, DUAN Liang-rong, ZHANG Tian-jiao (1938)
Spatial-temporal Characteristics and Influencing Factors of PM _{2.5} and Ozone in the Border Area of Jiangsu, Anhui, Shandong, and Henan from 2017 to 2021	CHEN Wei, XU Xue-zhe, LIU Wen-qing (1950)
Analysis of the Characteristics and Causes of PM _{2.5} and O ₃ Pollution in Xi'an During the Epidemic Lockdown Period	YUAN Xiao-hong, ZHANG Qiang, LI Qi, <i>et al.</i> (1963)
PM _{2.5} Chemical Composition and Spatial Variability in the Five Cities of Southern Jiangsu During Fall and Winter	FENG Wei, DING Feng, SHANG Yue, <i>et al.</i> (1975)
Source Apportionment of Fine Particles in Xianning, Hubei	LUO Yi, ZHU Kuan-guang, CHEN Pu-long, <i>et al.</i> (1983)
Spillover Benefits of Ammonia Emission Reduction in Xingtai Area to the Improvement of Atmospheric PM _{2.5} Concentration in the Beijing-Tianjin-Hebei Region	BIAN Ze-jun, WEN Chao-yu, LANG Jian-lei, <i>et al.</i> (1994)
Characteristics, Risk Assessment, and Sources of the Polluted Metallic Elements in PM _{2.5} During Winter in Liaocheng City	ZHANG Jing-qiao, ZHU Yao, CAO Ting, <i>et al.</i> (2003)
Composition Characteristics of Typical VOCs Sources in the Highland City of Lhasa	GUO Shu-zheng, YE Chun-xiang, LIN Wei-li, <i>et al.</i> (2011)
VOCs Pollution Characteristics and Health Risk Assessment in Typical Industrial Parks in Beijing: Environmental Impact of High and New Technology Industries	WANG Jie, YAO Zhen, WANG Min-yan, <i>et al.</i> (2019)
Analysis of a Typical Ozone Pollution Process in the Chengdu Plain Under the Influence of High Temperature Extremes	LEI Li-juan, ZHANG Yi, LUO Yi-na, <i>et al.</i> (2028)
Presence, Behavior, and Risk of Antiviral Drugs in the Aqueous Environment	GE Lin-ke, LI Xuan-yan, CAO Sheng-kai, <i>et al.</i> (2039)
Spatio-temporal Variation and Probability Health Risk of Heavy Metals in Surface Water of Xiaolangdi Reservoir in the Yellow River	WANG Liang, DENG Xue-jiao, WANG Xiao-lei, <i>et al.</i> (2054)
Hydrochemical Characteristics and Control Factors of Strontium Enrichment of the Duoqu River Basin in Eastern Xizang	LI Jing-jie, LIAN Sheng, WANG Ming-guo, <i>et al.</i> (2067)
Characteristics of Stable Isotopes in Precipitation and Its Moisture Sources in the Inland Regions of Northwest China	ZHANG Yan-yan, XIN Cun-lin, GUO Xiao-yan, <i>et al.</i> (2080)
Recharge Source and Evolution Process of Karst Groundwater in Tai'an Urban Area Based on Hydrochemistry and Hydrogen and Oxygen Isotopes	MENG Ling-hua (2096)
Hydrochemical Characteristics and Influencing Factors of Surface Water and Groundwater in the Mountainous Area of Danjiang River Basin	ZHANG Zi-yan, FU Yong-peng, WANG Ning-tao, <i>et al.</i> (2107)
Hydrochemical Characteristics and Control Factors of Groundwater in Yuanyangchi Irrigation Area, Jinta Basin	WANG Xiao-yan, HAN Shuang-bao, ZHANG Meng-nan, <i>et al.</i> (2118)
Hydrochemical Characteristics and Health Risk Assessment of Shallow Groundwater in the Plain Zone of Deyang City	LIU Nan, CHEN Meng, GAO Dong-dong, <i>et al.</i> (2129)
Spatial and Temporal Distribution of Microplastics in the Sediments of Typical Urban River Network	XU Wan-lu, FAN Yi-fan, QIAN Xin (2142)
Construction of Nitrogen Emission Inventory at Sub-basin Scale and Analysis of Key Sources in Fuxian Lake Watershed	WANG Yan-jie, LIANG Qi-bin, WANG Yan-xia, <i>et al.</i> (2150)
Preparation of Lanthanum Crosslinked Calcium Peroxide/Sepiolite/Sodium Alginate Composite Hydrogels and Their Elimination Performance for Endogenous Phosphorus	QU Si-tong, SHAN Su-jie, WANG Chong-ming, <i>et al.</i> (2160)
Formation Mechanism, Structural Characteristics of Ultrafine Mineral Particles, and Their Environmental Effects	LIU Zhen-hai, ZHANG Zhan-hua, YUAN Yu-xin, <i>et al.</i> (2171)
Research Progress in Application of Biochar-immobilized Bacteria Composites in Environmental Remediation	SUN Shu-yu, HUANG Meng-xin, KONG Qiang, <i>et al.</i> (2185)
Effect of Microplastics on Ammonia Nitrogen Adsorption by Zeolite in Water and Its Mechanism	LIAN Jian-jun, XIE Shi-ting, WU Pei, <i>et al.</i> (2195)
Mechanism of Ultraviolet Aging Effect on the Adsorption of Ciprofloxacin by Nano-biochar	MA Feng-feng, XUE Zhi-yi, ZHAO Bao-wei (2203)
Response of Phytoplankton Communities and Environmental Factors Under the Influence of Land Use in the Wuding River Basin	GUO Shan-song, HU En, DING Yi-tong, <i>et al.</i> (2211)
Bacterial Community Diversity and Functional Gene Abundance of Culturable Bacteria in the Wetland of Poyang Lake	YU Jiang, WANG Chun, LONG Yong, <i>et al.</i> (2223)
Bacterial Community Structure and Its Relationship with Heavy Metals in Sediments of Diannong River	LIU Shuang-yu, MENG Jun-jie, QIU Xiao-cong, <i>et al.</i> (2233)
Impacts of Treated Wastewater on Bacterial and Fungal Microbial Communities in Receiving Rivers	GUO You-shun, YU Zhong, HAO Wen-bin, <i>et al.</i> (2246)
Metagenomics Reveals the Characteristics and Functions of Bacterial Community in the Advanced Wastewater Treatment Process	HU Jian-shuang, WANG Yan, ZHOU Zheng, <i>et al.</i> (2259)
Spatio-temporal Characteristics of Habitat Quality and Natural-human Driven Mechanism in Dabie Mountain Area	ZHENG Ya-ping, ZHANG Jun-hua, TIAN Hui-wen, <i>et al.</i> (2268)
Hyperspectral Inversion of Soil Organic Matter Content Based on Continuous Wavelet Transform, SHAP, and XGBoost	YE Miao, ZHU Lin, LIU Xu-dong, <i>et al.</i> (2280)
Effects of Straw Retention, Film Mulching, and Nitrogen Input on Soil Quality in Dryland Wheat Field	YE Zi-zhuang, WANG Song-yan, LU Xiao, <i>et al.</i> (2292)
Effects of Long-term Application of Organic Fertilizer on the Occurrence and Migration Characteristics of Soil Microplastics	WANG Chang-yuan, MA Xiao-chi, GUO De-jie, <i>et al.</i> (2304)
Effect of Different Soil Salinities on N ₂ O Emission: A Meta-analysis	HUANG Yi-hua, SHE Dong-li, SHI Zhen-qi, <i>et al.</i> (2313)
Assessment and Prediction of Carbon Storage Based on Land Use/Land Cover Dynamics in the Guangdong-Hong Kong-Macao Greater Bay Area	ZHENG Hui-ling, ZHENG Hui-feng (2321)
Simulation of Temporal and Spatial Changes in Ecosystem Carbon Storage in Funiu Mountains Based on InVEST Model	ZHANG Zhe, SHI Zhen-qin, ZHU Wen-bo, <i>et al.</i> (2332)
Relationship Between Microbial Nutrient Limitation and Soil Organic Carbon Fraction During Shelterbelts Construction	XU Feng-jing, HUANG Yi-mei, HUANG Qian, <i>et al.</i> (2342)
Characteristics and Driving Forces of Organic Carbon Mineralization in Brown Soil with Long-term Straw Returning	ZHAO Yu-hang, YIN Hao-kai, HU Xue-chun, <i>et al.</i> (2353)
Effects of Winter Green Manure Mulching on Soil Aggregates, Organic Carbon, and AMF Diversity	LU Ze-rang, CHEN Jia-yan, LI Zhi-xian, <i>et al.</i> (2363)
Spatial and Temporal Distribution of Nitrogen in the Liaohu River Basin and Its Responses to Land Use and Rainfall	ZHOU Bo, LI Xiao-guang, TONG Si-chang, <i>et al.</i> (2373)
Nitrous Oxide Emissions from Ponds in Typical Agricultural Catchment with Intensive Agricultural Activity	ZHANG Xin-yue, XIAO Qi-tao, XIE Hui, <i>et al.</i> (2385)
Effects of Biochar Combined with Different Types of Nitrogen Fertilizers on Denitrification Bacteria Community in Vegetable Soil	LIU Xiao-wan, LIU Xing-ren, GAO Shang-jie, <i>et al.</i> (2394)
Investigation of Soil Microbial Characteristics During Stand Development in <i>Pinus tabulaeformis</i> Forest in Taiyue Mountain	MA Yi-shu, CAO Ya-xin, NIU Min, <i>et al.</i> (2406)
Simulation of Heavy Metal Content in Soil Based on Sparse Sample Sites	ZHANG Jia-qi, PAN Yu-chun, GAO Shi-chen, <i>et al.</i> (2417)
Identification Priority Source of Heavy Metal Pollution in Greenspace Soils Based on Source-specific Ecological and Human Health Risk Analysis in the Yellow River Custom Tourist Line of Lanzhou	LI Jun, LI Xu, LI Kai-ming, <i>et al.</i> (2428)
Heavy Metal Accumulation and Assessment of Potential Ecological Risk Caused by Soil pH Changes in Different Types of Purple Soils in Southwest China	ZHANG Hai-lin, ZHANG Yu, WANG Ding, <i>et al.</i> (2440)
Assessment and Source Analysis of Heavy Metal Pollution in Arable Land Around an Iron Ore Mining Area in Chongqing	LIAO Ze-yuan, LI Jie-qin, SHEN Zhi-jie, <i>et al.</i> (2450)
Health Risk Assessment of Heavy Metals in Farmland Soils and Crops Around Metal Mines	WEI Hong-bin, LUO Ming, XIANG Lei, <i>et al.</i> (2461)
Effect of Red Mud-based Nano Zero-valent Iron on Remediation of Polymetallic Contaminated Soil	LIU Long-yu, YANG Shi-li, ZHAO Huang-shi-yu, <i>et al.</i> (2474)
Effects of Zinc Application on Cadmium Accumulation in Alkaline Soil-Wheat Seedling Systems	ZHANG Yao, WANG Tian-qi, NIU Shuo, <i>et al.</i> (2480)