

DOI:10.13870/j.cnki.stbcbx.2026.02.014 CSTR:32310.14.stbcbx.2026.02.014

赵彤彤,李华朋.基于多时相高分遥感数据与机器学习的黑土层厚度反演[J].水土保持学报,2026,40(2):121-130.

ZHAO Tongtong, LI Huapeng. Inversion of black soil layer thickness based on multi-temporal high-resolution remote sensing data and machine learning[J]. Journal of Soil and Water Conservation, 2026, 40(2): 121-130.

## 基于多时相高分遥感数据与机器学习的黑土层厚度反演

赵彤彤<sup>1</sup>, 李华朋<sup>2</sup>

(1. 吉林建筑大学测绘与勘查工程学院, 长春 130118; 2. 中国科学院东北地理与农业生态研究所, 长春 130102)

**摘要:** [目的] 为评估作物生长期遥感信息与地形因子对提升黑土层厚度反演制图精度的影响。[方法] 在海伦农场小流域, 综合应用裸土期及生长期 8 m 高分遥感数据、SRTM 数字高程模型 (digital elevation model, DEM) 地形数据和实地土壤采样点数据, 采用随机森林 (random forest, RF)、梯度提升决策树 (gradient boosting decision tree, GBDT) 和极端梯度提升 (extreme gradient boosting, XGBoost) 模型进行预测与制图, 并以决定系数 ( $R^2$ ) 和均方根误差 (RMSE) 评估性能。[结果] 1) 引入作物生长期遥感特征后, 所有模型反演精度显著优于仅使用裸土期数据的模型, 其中 RF、GBDT 和 XGBoost 模型的  $R^2$  分别提高 0.09、0.11 和 0.10, RMSE 分别降低 1.15、1.34 和 1.35; 2) 在联合模型中引入地形因子后精度进一步提高, 上述 3 种模型的  $R^2$  分别增加 0.03、0.12 和 0.05, RMSE 分别下降 0.38、1.49 和 0.81; 3) XGBoost 模型在所有特征组合中表现最优 ( $R^2=0.60$ , RMSE=12.31), 显著优于 GBDT ( $R^2=0.53$ , RMSE=13.27) 和 RF ( $R^2=0.46$ , RMSE=14.33), 结合地形因子后解释力更强。[结论] 成功构建 8 m 高分辨率的黑土层厚度分布图, 证实“多时相遥感+地形因子+机器学习”技术范式在提升土壤属性遥感制图精度方面的有效性, 为精准农业实践、黑土资源可持续利用、国家黑土保护战略及粮食安全保障提供高精度的空间决策依据。

**关键词:** 小流域; 作物生长期; 地形因子; 机器学习; 黑土层厚度

中图分类号: S152.2 文献标识码: A 文章编号: 1009-2242(2026)02-0121-10

## Inversion of Black Soil Layer Thickness Based on Multi-Temporal High-Resolution Remote Sensing Data and Machine Learning

ZHAO Tongtong<sup>1</sup>, LI Huapeng<sup>2</sup>

(1. School of Geomatics and Prospecting Engineering, Jilin Jianzhu University, Changchun 130118, China; 2. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China)

**Abstract:** [Objective] This study aims to evaluate the influence of remote sensing data from crop growth periods and topographic factors on improving the accuracy of black soil layer thickness inversion and mapping. [Methods] The study was conducted in a small watershed in Hailun Farm. 8 m high-resolution remote sensing data from both the bare soil and crop growth periods, SRTM digital elevation model (DEM) data, and in situ soil sampling data were comprehensively used. Three machine learning models, such as random forest (RF), gradient boosting decision tree (GBDT), and extreme gradient boosting (XGBoost) were applied for spatial prediction and mapping. Additionally, model performance was evaluated using the coefficient of determination ( $R^2$ ) and root mean square error (RMSE). [Results] 1) After incorporating remote sensing features from the crop growth period, the inversion accuracy of all models was significantly higher than that of models using data from the bare soil period alone. Specifically, the  $R^2$  values of the RF, GBDT, and XGBoost models increased by 0.09, 0.11, and 0.10, respectively, while RMSE decreased by 1.15, 1.34, and 1.35, respectively. 2) By adding topographic factors to the integrated models, the accuracy was further improved. The  $R^2$  values of the three models increased by 0.03, 0.12, and 0.05, respectively,

收稿日期: 2025-07-31 修回日期: 2025-09-22 录用日期: 2025-09-28 网络首发日期 (www.cnki.net): 2025-11-24

资助项目: 国家重点研发计划项目 (2021YFD1500100); 中国科学院战略性先导专项项目 (XDA28070500)

第一作者: 赵彤彤 (1997—), 男, 硕士研究生, 主要从事农业遥感、土壤理化性质的定量遥感反演研究。E-mail: zhaotongtong@student.jlzu.edu.cn

通信作者: 李华朋 (1984—), 男, 副研究员, 主要从事农业遥感、遥感信息智能化提取与分析、时空大数据挖掘研究。E-mail: lihuapeng@iga.ac.cn

http://stbcbx.alljournal.com.cn

while RMSE decreased by 0.38, 1.49, and 0.81, respectively. 3) The XGBoost model delivered the best performance across all feature combinations ( $R^2=0.60$ , RMSE=12.31), significantly outperforming both GBDT ( $R^2=0.53$ , RMSE=13.27) and RF ( $R^2=0.46$ , RMSE=14.33), and exhibited stronger explanatory power when combined with topographic factors. [Conclusion] This study successfully constructs an 8 m high-resolution map of black soil layer thickness. The results demonstrate the effectiveness of the “multi-temporal remote sensing + topographic factors + machine learning” technical paradigm in improving the accuracy of soil property mapping. This provides high-precision spatial decision-making support for precision agriculture practices, sustainable utilization of black soil resources, national black soil conservation strategies, and food security assurance.

**Keywords:** small watershed; crop growth period; topographic factors; machine learning; black soil layer thickness

Received: 2025-07-31

Revised: 2025-09-22

Accepted: 2025-09-28

Online(www.cnki.net): 2025-11-24

黑土是一种具有黑色或暗黑色腐殖质表土层的优质土壤资源,以其优良的性状、较高的肥力和突出的农耕适宜性而著称。作为珍贵的稀缺资源,黑土被誉为耕地中的“大熊猫”,是大自然赋予人类的宝贵资源。我国东北地区凭借其肥沃的黑土资源,已成为国家重要的粮食生产基地,保障着国家粮食安全。然而,长期的高强度连续耕作且缺乏有效保护措施,导致该地区面临严峻的土壤侵蚀问题。由于土地从原生植被转变为农业用地,使得该地区的土壤侵蚀在过去 300 年中几乎呈指数级增长<sup>[1]</sup>。严重的土壤侵蚀会显著降低黑土层厚度。据估计,黑土层厚度的年损失量在 0.3~1.0 cm<sup>[2]</sup>。更严重的是,黑土区部分地区的黑土已经消失,严重威胁我国粮食安全。因此,迫切需要对黑土区黑土层厚度的空间变化进行估算。然而,黑土区的土壤侵蚀与黄土高原等地区有所不同,黑土区土壤侵蚀具有隐蔽性:其主要表现形式为面蚀和缓坡长坡条件下的细沟侵蚀,侵蚀物质大多在坡麓或河道附近沉积,并未长距离输移至河流中。因此,尽管坡面侵蚀严重,但河流输沙量却偏低<sup>[3]</sup>。这一特性导致黑土层厚度的空间变异主要集中于小流域内部。因此,精确掌握流域尺度的黑土层厚度现状,对于建立黑土层厚度动态监测体系、实施以小流域为单元的水土保持精准治理具有重要科学意义和实践价值<sup>[4]</sup>。

目前,国家虽已在东北黑土区开展多次土壤调查,但这些调查大部分并不以黑土层厚度作为主要调查目的;且由于黑土层厚度调查困难、成本高,目前仅有较少研究开展黑土层厚度空间估算工作<sup>[5]</sup>。传统的黑土厚度估算方法是基于有限的土壤样本剖面进行插值,但由于黑土层厚度空间异质性较高,插值结果不具有代表性且存在一定偏差<sup>[6]</sup>。为此,部分学者应用基于物理的方法和环境相关方法开展研究<sup>[7]</sup>,然而这些方法往往需要大量模型参数,模型建

立困难。此外,传统模型普遍难以实现实时或准实时大范围监测,给实际应用带来极大困难。

面对传统方法在大范围实时监测的困境,近年来随着高分辨率遥感技术的飞速发展,多光谱卫星数据在土壤调查中的应用日益广泛<sup>[8]</sup>。已有研究<sup>[9]</sup>表明,裸土期多光谱和高光谱遥感影像在数字土壤制图中具有重要的应用价值,可以提供有关土壤表面覆盖状况的信息。裸土期影像结合地面实测土壤数据,可建立土壤性质预测模型。然而,裸地数据的获取受云量、大气条件等因素显著制约,其数据质量与可用性难以保障。因此,探索利用其他时期(特别是作物生长期)遥感影像进行土壤制图,成为克服裸土期局限的关键研究方向。有研究<sup>[10]</sup>证实,生长期遥感提取的光谱信息和植被指数中蕴含丰富的作物生长状态信息。由于作物生长状况在很大程度上受到其根系所处土壤环境(如有机质含量、水分状况、养分供应等)的影响和制约,因此这些遥感提取的信息与底层土壤属性之间存在着显著的空间协同关系。这种关联性使得生长期数据能够有效替代或补充裸土期信息,成为构建土壤性质预测模型的关键输入。为了充分挖掘遥感数据的潜力并构建稳健的预测模型,机器学习算法已成为该领域的核心工具<sup>[11]</sup>。其中随机森林(RF)凭借其优异的稳定性和易用性,被广泛应用于土壤属性的空间预测研究。目前,基于裸土期遥感影像反演黑土层厚度时,RF 仍被视为首选方法;尤其在东北地区的流域尺度研究中,其预测效果优于经典贝叶斯克里格、普通克里格和简单克里格等多种克里格方法。与此同时,梯度提升决策树(GBDT)和极端梯度提升(XGBoost)等算法在农业遥感中的应用也逐渐增多,并在土壤有机质及有机碳含量预测等任务中展现出更高的精度潜力。因此,系统比较上述先进机器学习算法在黑土层厚度反演中的性能,对优化模型选择与提升预测精度具有重要意义。

近年来,黑土层厚度空间制图的分辨率不断提高,体现了遥感与空间分析技术在土壤属性精细化识别方面的显著进展。这一趋势在多项研究中得到明确体现。例如,在东北典型黑土区小流域生成30 m分辨率的耕地黑土层厚度分级图<sup>[12]</sup>;通过对Landsat8影像进行重采样处理,进一步实现10 m精度的黑土层厚度分布制图<sup>[13]</sup>。从30 m到10 m的空间分辨率提升,不仅大幅增强黑土资源的精细刻画能力,也表明高分辨率土壤属性反演技术日趋成熟。然而,面对小流域尺度水土保持精准治理的实际需求,当前制图精度仍存在进一步提升的空间。发展更高分辨率的黑土层厚度监测技术,对于实现土壤退化精准评价与治理具有重要意义。

本研究通过获取并处理国产高分系列卫星(GF-1B/D和GF-6)多时相遥感数据,利用其8 m高分辨率的特性,通过多时相观测有效捕捉作物冠层光谱变化,从而构建集成作物生长信息和地形特征的机器学习预测框架,系统比较不同算法在黑土层厚度反演中的表现。研究主要目标包括:1)制作8 m分辨率小流域黑土层厚度空间分布图;2)评估在可获取条件下,裸地信息与作物生长信息协同反演黑土层厚度的潜力;3)分析地形因子对黑土层厚度空间分异的影响;4)对比RF、GBDT和XGBoost 3种先进机器学习模型的反演性能,为黑土厚度精准预测提供方法优化依据。这一研究框架不仅拓展了农业遥感的应用维度,也为黑土资源精准管理提供新的技术支撑。

## 1 材料与方法

### 1.1 研究区概况

研究区选择黑龙江省海伦市东北部的海伦农场(47°39′~47°48′N, 127°04′~127°24′E),地处小兴安岭南麓,通肯河畔的漫川漫岗黑土区。研究区总面积约4 333.3 hm<sup>2</sup>,地势平缓,东高西低,属中温带大陆性季风气候,冬季寒冷干燥,夏季温热多雨,雨热同期,平均降水量550~560 mm,年蒸发量1 200 mm,年平均气温1.5℃。主要土壤类型为黑钙壤土,其性状是土层深、肥力高,且土性温和、保水供水能力强。玉米-大豆轮作是该地区主要种植制度,主要土地利用类型为农田和林地。

### 1.2 土壤取样与分析

于2024年4月中旬至5月初采集土壤样品,以农田土壤为研究对象,划定小流域研究区(图1),综合考虑典型区域土地利用现状、地形特征和交通可达性等关键因素,采用动态采样策略进行土壤剖面调查。采样过程中,通过高精度手持GPS实时定位采

样点坐标,并由至少3名具备专业土壤学背景的研究人员组成鉴定小组,共同完成黑土层厚度的识别工作。

黑土层厚度的判定基于暗沃表层的诊断指标。暗沃表层作为土壤系统分类中的核心诊断表层之一,其在中美两国分类体系中的判定标准存在差异:中国土壤系统分类主要依据厚度、颜色、有机质含量、盐基饱和度和结构5项指标<sup>[14]</sup>;而美国体系则扩展至包含岩石结构、磷含量、湿润天数及n值在内的9项指标。本次调查对黑土层的判定综合借鉴中美土壤分类体系中关于“暗沃表层”的核心指标,并重点聚焦于颜色与结构2项野外可快速评估的标准。为最大限度减少机械扰动对土壤原始结构和真实厚度的影响,采取分层连续采样策略:自地表开始,以每10 cm为1个采样单元逐层向下采集,直至完整穿越黑土层并获得全剖面序列样品。该方法显著降低单次深挖可能引起的土壤压缩效应,从而更准确地还原黑土层的实际厚度。所依据的具体判定标准为<sup>[15]</sup>:颜色(润态明度 $\leq 3.5$ 、干态明度 $\leq 5.5$ 、彩度 $\leq 3.5$ )和结构(主要呈粒状结构和小角块状结构,干态时需同时满足不形成大块状或整块状结构,抗破裂强度属于中等硬度或更软级别)。基于野外操作的可行性与效率考量,本研究选择以颜色(现场比色)、结构(形态观测)2项可快速评估的指标作为核心标准,其余指标未纳入本次调查范围。颜色判读采用双重比对方法以提高准确性:首先观察湿润状态下土壤剖面的颜色,随后将剖面样品置于田间日光下自然稍作风干后再次进行颜色比对,综合判定其是否达到黑土颜色的判定阈值。当土壤样本同时满足上述颜色和结构的要求时,即认定为黑土层<sup>[16]</sup>。在研究区内系统采集的58个代表性黑土层样品的空间分布见图1。

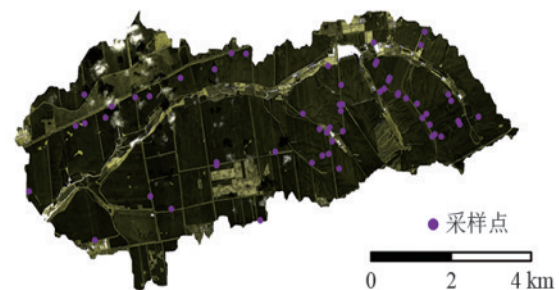


图1 海伦农场黑土层厚度采样点

Fig. 1 Sampling points of black soil layer thickness in Hailun Farm

### 1.3 数据来源与预处理

本研究采用的遥感数据为高分一号B星(GF-1B)和高分一号D星(GF-1D)及高分六号(GF-6)1A级8 m多光谱影像,通过中国资源卫星应用中心(<https://data.cresda.cn/#/home>)获取,其影像空间

分布与云量信息见图2。其中,GF-1B/1D卫星搭载的全色多光谱传感器(PMS)提供空间分辨率为2 m(全色)和8 m(多光谱)的影像数据,GF-6卫星搭载的全色多光谱传感器(PMS)在8 m分辨率模式下提供与GF-1B/1D相似的多光谱数据,包含蓝光(0.45~0.52  $\mu\text{m}$ )、绿光(0.52~0.59  $\mu\text{m}$ )、红光(0.63~0.69  $\mu\text{m}$ )和近红外(0.77~0.89  $\mu\text{m}$ )4个波段。

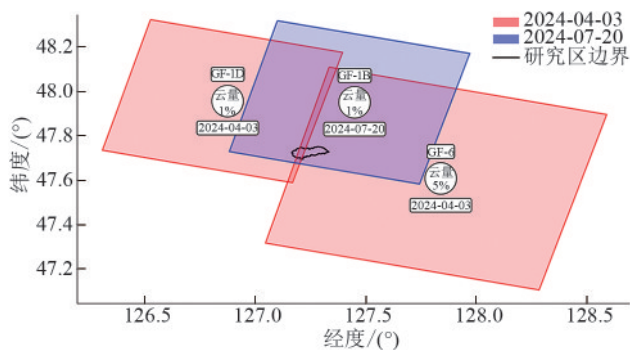


图2 高分卫星影像空间分布与云量信息

Fig. 2 Spatial distribution and cloud coverage information of GF satellite imagery

在数据预处理方面,采用ENVI5.3软件对原始影像进行系统处理,包括辐射定标、大气校正和正射校正等关键步骤。为解决单景影像覆盖范围有限的问题,将2024年4月3日同期获取的GF-1D和GF-6多光谱影像进行空间融合处理,通过影像镶嵌和匀色技术生成覆盖完整研究区的高质量裸土期基准影像,同时选用2024年7月20日获取的GF-1B多光谱影像作为作物生长期的特征数据。影像融合过程中,采用基于地理坐标的严格配准方法,确保不同数据源间的空间一致性,同时通过直方图匹配消除影像间的辐射差异,保证数据质量的一致性。

研究所用到的12.5 m分辨率数字高程模型(DEM)数据来源于美国宇航局(NASA)的航天飞机雷达地形测绘任务(SRTM)v3.0数据集。该数据由2000年奋进号航天飞机搭载的合成孔径雷达(SAR)系统采集,采用C波段(5.6 cm波长)雷达干涉测量技术获取全球80%陆地表面的高精度地形数据。SRTM DEM数据采用WGS84地理坐标系和EGM96高程基准,其标称垂直精度为 $\pm 16$  m(绝对高程误差)和 $\pm 6$  m(相对高程误差)。本研究通过NASA Earthdata平台(<https://earthdata.nasa.gov>)下载研究区DEM数据,通过重采样得到8 m分辨率的DEM影像,为后续黑土层厚度反演提供基础地理数据支持。

本研究中的农田掩膜空间分布数据是通过多时相高分辨率遥感影像进行人工目视解译,采用数字化、矢量化方法构建完成的,最终生成研究区农田掩膜分布的矢量数据集。

#### 1.4 黑土层厚度反演方法

本研究构建考虑加入作物生长信息和地形因子的黑土层厚度反演模型,技术路线见图3。首先基于GF-1B/D和GF-6卫星影像,分别提取裸土期和作物生长期的光谱特征及植被指数,同时结合DEM数据获取地形因子,随后构建3类特征组合:1)裸土期单时相特征;2)裸土期与生长期联合特征;3)联合含DEM的多源特征。基于这些特征组合,研究对比评估随机森林(RF)、梯度提升决策树(GBDT)和XGBoost 3种机器学习算法的预测性能,最终确定最优模型并实现研究区黑土层厚度的高精度空间制图。

#### 1.5 黑土层厚度预测的解释协变量

黑土层厚度的空间分布特征受多种环境生态因子及其复杂相互作用的影响。基于区域环境特征,本研究系统选取3类解释变量构建预测指标体系:1)光谱特征变量,包括各波段反射率及其变换形式;2)植被光谱指数,如NDVI、EVI等反映植被生长状况的指标;3)地形因子,包括高程、坡度、剖面曲率等地形属性参数(表1)。这些协变量为建立高精度预测模型提供全面的特征支撑。

#### 1.6 模型构建

1.6.1 随机森林(RF) 随机森林是一种用于分类和回归的机器学习方法,是随机分类树和回归树的集合<sup>[17]</sup>。随机森林中的树是基于原始数据的自举样本构建的。

随机森林的主要优点有:1)对于原始数据,即使存在缺失值,随机森林也可以分析连续变量和分类变量;2)对于训练过程,随机森林节省计算时间,因为每棵树都可以并行构建,不需要修剪;3)对于模型性能而言,随机森林总是收敛的,因此过拟合风险较低,其误差估计基于袋外样本,能够提供无偏且稳定的性能评估。

1.6.2 梯度提升决策树(GBDT) 梯度提升决策树模型是一种集成学习模型,是决策树和增强算法的结合<sup>[18]</sup>。该模型是计算真实值和预测值之间残差的综合树模型。该集成算法模型采用梯度、提升和决策树等方法来解决分类问题并进行回归预测。其关键机制包括:1)通过计算损失函数的负梯度(近似残差)来指导每一轮新树的构建,从而逐步逼近真实目标;2)将多个弱模型以线性叠加的方式组合,并以阶段性的方式逐步减小训练误差;3)支持多种损失函数,可灵活应用于回归与分类任务。

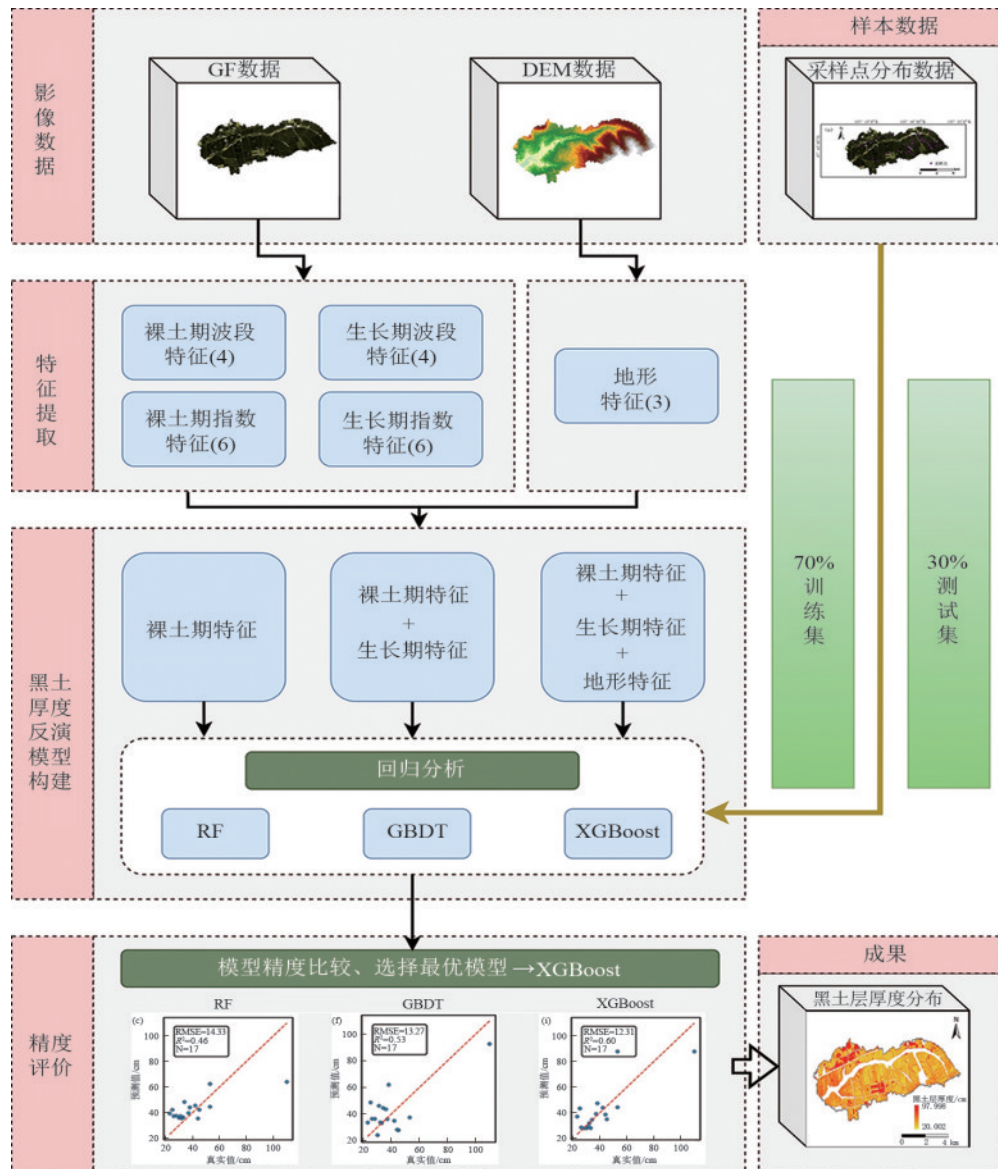


图 3 技术路线

Fig.3 Technical framework

1.6.3 极端梯度提升(XGBoost) XGBoost是一种高效、灵活的梯度提升决策树(GBDT)框架,由CHEN等<sup>[19]</sup>于2016年提出,因其卓越的性能和计算效率,在许多国际数据挖掘竞赛中取得优异的成绩,其性能超过深度学习算法。XGBoost算法是对GBDT算法的改进。其核心机制包括:1)在损失函数中引入正则化项(L1/L2正则化),用于控制模型复杂度,从而直接抑制过拟合;2)利用损失函数的二阶泰勒展开,更精确地计算梯度方向,加快收敛速度;3)支持特征抽样和并行处理,进一步提升训练效率。正因为这些优化,XGBoost通常比传统GBDT具有更高的准确性和鲁棒性。

1.7 模型评估与超参数调优

将全部采样点按7:3比例随机划分,即41个点用于模型训练,剩余17个点用于模型测试。采用决定系数( $R^2$ )和均方根误差(RMSE)来评价模型的准

确性和稳定性,即 $R^2$ 值越接近1, RMSE值越小,表明模型性能越强。使用K折交叉验证对模型的训练集进行验证,并使用网格搜索对超参数进行调整。K折交叉验证将训练集分成K个相等的部分,其中一部分作为测试集,其余部分作为训练集。交叉验证重复K次,每个子样本验证1次,取K次结果的平均值作为模型结果。本研究中,由于样本点数量较少,所以K值取5,即应用5折交叉验证。

计算公式为:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{1}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{2}$$

式中: $n$ 为样本总数; $y_i$ 为样本*i*的真实值; $\hat{y}_i$ 为样本*i*的预测值; $\bar{y}$ 为真实值的对应平均值。

表 1 黑土层厚度预测的解释协变量  
Table 1 Explanatory variables for black soil layer thickness prediction

类别	变量	来源	分辨率/m
波段	蓝 B1(0.45~0.52 μm)	GF-1B/D	8
	绿 B2(0.52~0.59 μm)		
	红 B3(0.63~0.69 μm)		
	近红外 B4(0.77~0.89 μm)		
光谱指数	归一化差值植被指数(NDVI)	GF-1B/D	8
	归一化差值水体指数(NDWI)		
	增强型植被指数(EVI)		
	土壤调节植被指数(SAVI)		
	绿光归一化差值植被指数(GNDVI)		
地形因子	比值植被指数(RVI)	DEM	8
	高程(E)		
	坡度(S)		
	剖面曲率(C)		

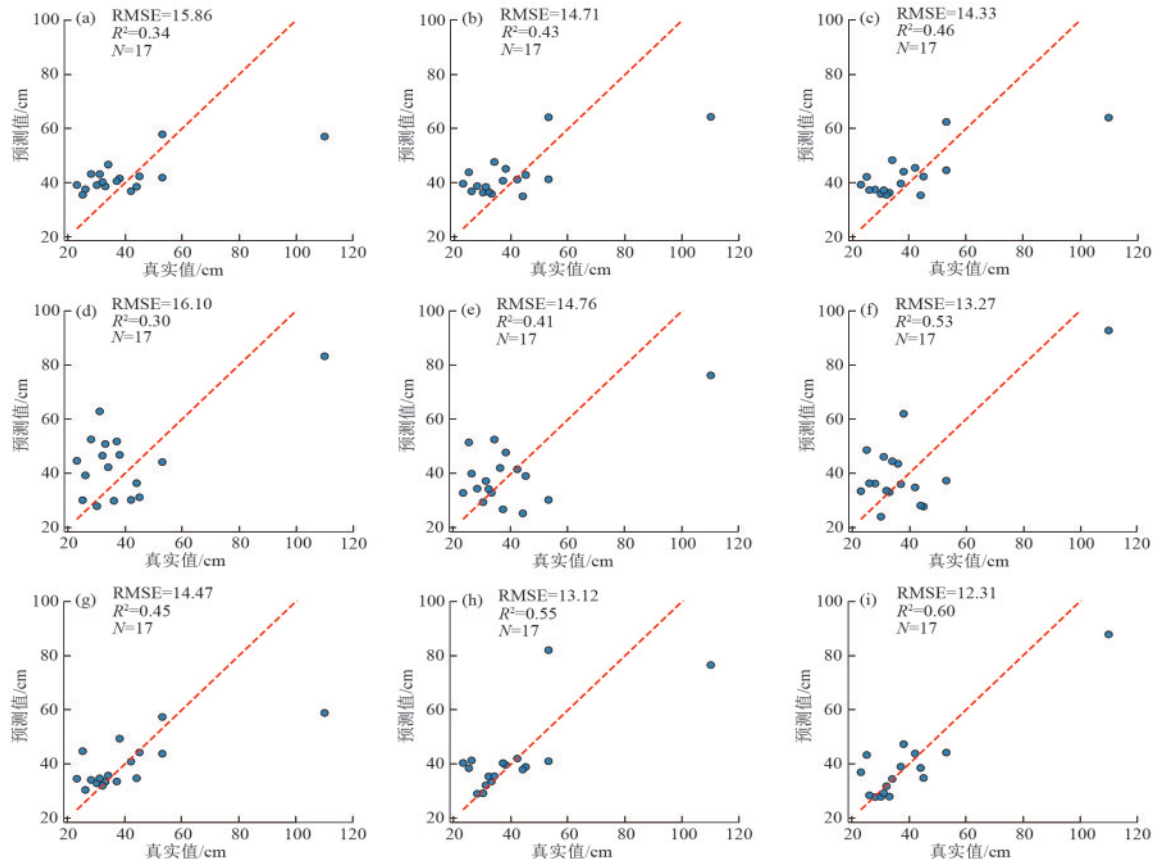
## 2 结果与分析

### 2.1 不同模型下对不同特征组合的精度比较

本研究采用随机森林(RF)、梯度提升决策树(GBDT)和极端梯度提升(XGBoost)3种算法分别对裸土期特征、裸土期和生长期联合特征及包含地形

因子的联合特征(联合含DEM)建立预测模型,并通过计算训练集和测试集的决定系数( $R^2$ )和均方根误差(RMSE)来评估模型的准确性(图4)。

首先,在裸土期特征基础上引入作物生长信息(裸土期+生长期)后,所有模型的预测性能均显著提升,RF模型的 $R^2$ 提高0.09,GBDT模型提升0.11,XGBoost模型提升0.1;相应RMSE分别降低1.15、1.34、1.35。其次,进一步引入地形因子(联合含DEM)后,模型精度得到更明显的改善,与仅联合生长期特征相比,RF、GBDT和XGBoost模型的 $R^2$ 分别进一步提升0.03、0.12和0.05,RMSE分别降低0.38、1.49和0.81。通过XGBoost模型对不同特征组合下的特征重要性分析(图5)显示,作物生长信息和地形因子均具有较高的贡献度,说明二者对黑土层厚度反演均起到重要作用,且存在协同增效效应。横向对比3种算法表现发现,在所有特征组合下,XGBoost模型均取得最优的预测效果,其 $R^2$ 最高、RMSE最低,表明该算法在黑土层厚度反演任务中具有明显优势,推荐作为该类建模的首选方法。本研究结果为土壤属性遥感反演研究提供有益的方法借鉴和参考。



注:图(a)~(c)、(d)~(f)、(g)~(i)分别为RF、GBDT和XGBoost模型在裸土期、裸土期联合生长期及联合含DEM特征组合。

图 4 不同特征组合在不同模型中的黑土层厚度预测散点图

Fig. 4 Scatter plots of black soil layer thickness predictions under different feature combinations across models

### 2.2 黑土层厚度的空间分布

由图 6 可知,裸土期和生长期联合含 DEM 特征(联合含 DEM)组合在使用 XGBoost 算法下有最高精度。因此,采用该特征组合在 XGBoost 模型中预

测研究区 2024 年黑土层厚度的空间分布。预测结果见图 6。研究区黑土层厚度空间分布呈现明显的区域差异,即西部区域黑土层最为深厚,中部偏北区域最为浅薄,东部区域黑土层厚度显著低于西部。

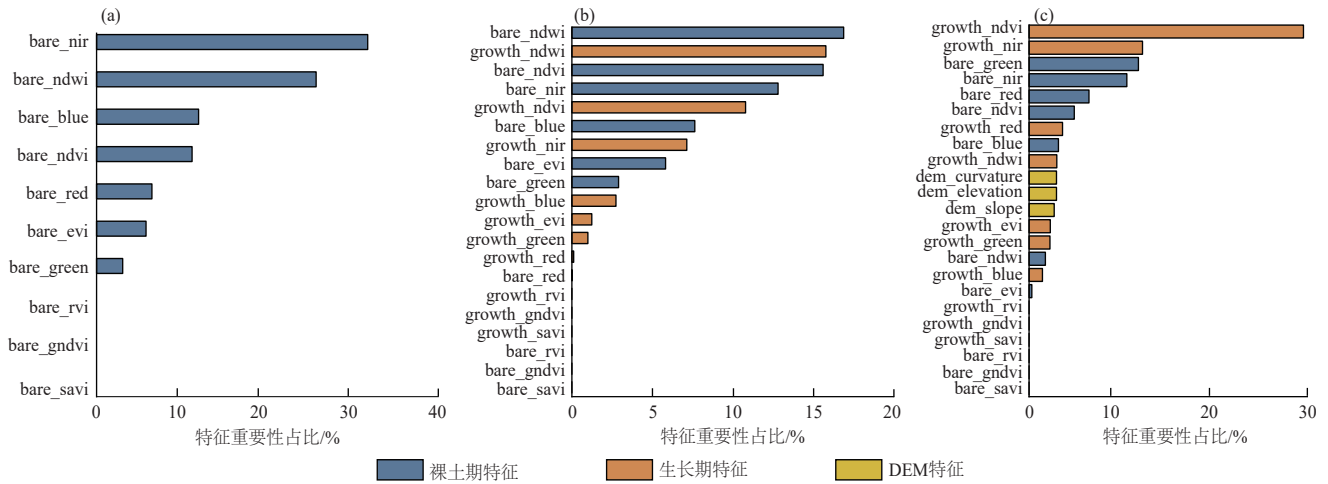


图 5 XGBoost 模型在不同特征组合下的特征重要性排序

Fig. 5 Feature importance ranking of XGBoost model under different feature combinations

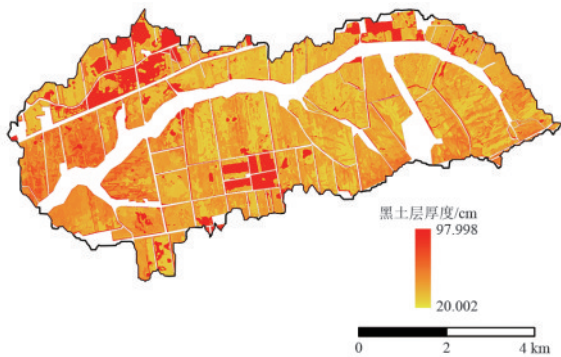


图 6 研究区黑土层厚度分布

Fig 6 Distribution map of black soil layer thickness in the study area

从图 7 可以看出,通过百分比条形图展示高程(E)、坡度(S)和曲率(C)在不同的分类情况下黑土层厚度的分布占比,对研究区的黑土层厚度进行分级统计,以黑土层厚度为划分标准进行划分,分为 0~20 cm 的“破皮黄”区域、20~30 cm 的薄层黑土区、30~60 cm 的中层黑土区及 >60 cm 的厚层黑土区<sup>[4,20]</sup>,不同厚度级别所占比例差别较大。通过对地形因子和黑土层厚度的分级处理显示,随着高程的增加,厚层黑土占比逐渐减少。当高程 < 265 m 时,厚层黑土区占 37.6%,而当高程 > 275 m 时,厚层黑土区仅占 19.2%;坡度和曲率分类中以中层黑土区为主,占比接近 65%,其他各黑土层厚度占比接近。从高程、坡度和曲率整体来看,研究区以中层黑土区为主,厚层黑土区次之,薄层黑土区最少,未出现破皮黄区域,表明研究区整体黑土资源保存状况良好,但存在明显的空间异质性。

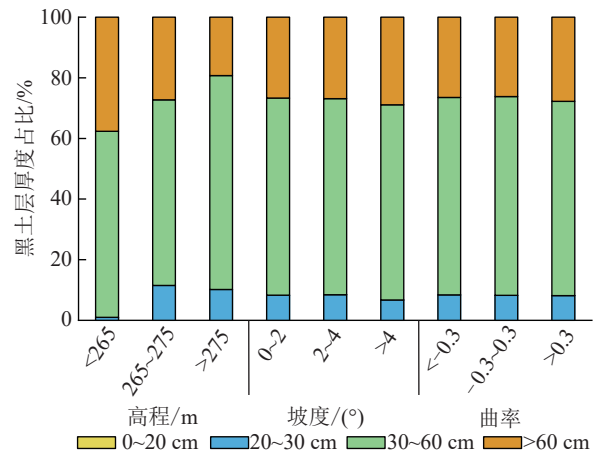


图 7 黑土层厚度与地形因子累计分布

Fig 7 Cumulative distribution of black soil layer thickness and topographic factors

## 3 讨论

### 3.1 结合生长季节影像对黑土层厚度预测的影响分析

现有关于黑土层厚度的遥感反演研究主要依赖于裸土期影像数据,主要是裸土期的地面作物残渣覆盖较少,能够更直接地反映土壤特性<sup>[21]</sup>。然而,裸土期遥感数据的获取往往受到云层覆盖和大气条件等因素的限制,导致数据质量难以保证<sup>[22]</sup>。与以往研究相比,本研究通过比较不同模型下裸土期与裸土期和生长期联合特征(裸土期+生长期)的黑土层厚度制图精度(图 4),可以看出,添加作物生长信息可以有效提高黑土层厚度的预测性能。

图 5 中的 XGBoost 模型在不同特征组合下的特征重要性显示,在裸土期模型中,近红外波段(bare\_nir)的特征重要性占比高达 33.1%,主要是因为近红外对植被和土壤具有高透射率特性,能够有效反映土壤表层的有机质质量分数<sup>[8]</sup>,裸土条件下土壤反射光谱占据主导地位,使得 bare\_nir 成为最具判别力的特征。

当引入生长期特征后,模型的特征重要性分布呈现出明显变化,其中生长期特征表现出显著的重要性占比,排序靠前的 2 个生长期特征 NDWI (growth\_ndwi)和 NDVI(growth\_ndvi)分别占 15.7% 和 10.7%,显示出植被生长状况对模型的重要贡献。值得注意的是,植被覆盖引入强烈信号(如生长期的 ndwi 对叶片水分的响应),导致模型注意力从裸土特征转向植被相关特征,使得特征重要性占比发生改变。尽管生长期特征影响显著,但裸土期特征在加入生长信息后依旧重要,其裸土期 NDWI(bare\_ndwi)的重要性上升至第 1 位(16.8%),可能是因为该特征不仅反映裸土期的土壤湿度,还与生长期植被水分状况存在密切联系,共同刻画“土壤-植被”水分传递的动态过程。

当地形因子(DEM)被引入模型后,生长期特征的重要性进一步凸显并占据绝对主导地位,其生长期的 NDVI(growth\_ndvi)的重要性占比大幅提升至 29%,成为最具影响力的特征。此现象可以从地形因素对生态环境的调控作用来解释,即高程、坡度等地形特征显著影响水分分布、温度和日照条件,从而放大 NDVI 的空间分异性<sup>[23-24]</sup>。将作物生长信息纳入黑土层厚度制图具有科学依据,由于黑土层厚度与土壤有机质含量等理化参数密切相关<sup>[25]</sup>,以及土壤有机质含量直接影响作物生长状况,作物生长状况可以间接反映黑土层的厚度特征。

考虑到裸土期影像获取存在时间窗口限制的现实困难,本研究建议在保证裸土期数据的基础上,合理引入作物生长期信息,从而有效提高黑土层厚度制图的精度和可靠性。

### 3.2 地形因子对黑土层厚度预测的影响分析

通过引入地形因子数据显著提升黑土层厚度预测模型的精度。特征重要性分析(图 5)表明,在裸土期和生长期联合含 DEM 特征组合中,地形因子具有重要的贡献度。相关性分析(图 8)进一步揭示,高程与黑土层厚度呈显著负相关( $r=-0.26$ ),而坡度和曲率与黑土层厚度的相关性相对较弱( $r=0.048$ ),可能与研究区整体地形较为平缓的特征有

关。地形因子对黑土层厚度的影响机制主要体现在水土流失过程方面,在重力、水力等外营力作用下,坡耕地土壤发生迁移,导致地面高程变化,进而改变土壤厚度分布格局<sup>[26]</sup>。近年来的土壤扩散模型研究表明,土壤扩散过程是塑造土壤厚度空间格局的主导机制<sup>[27]</sup>。具体来说,土壤颗粒在重力作用下发生顺坡移动现象,包括土壤蠕动和浅层滑坡过程。在缓坡、地形发散的凸坡区,在重力驱动下,土壤输移以缓慢的蠕动为主,而在陡坡、地形收敛的凹坡、直坡处,浅层滑坡则成为影响土壤厚度分布的主要过程<sup>[28]</sup>。因此,海拔和坡度对黑土层厚度具有较大影响。在构建土壤厚度空间分布的物理-随机混合预测模型时发现,土壤厚度与地形曲率参数表现出显著的相关性<sup>[29]</sup>,为利用地形特征预测黑土层厚度提供理论依据。

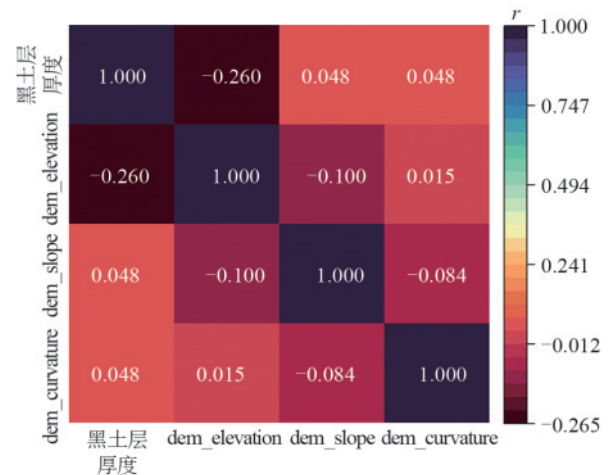


图 8 黑土层厚度与地形因子相关性矩阵

Fig 8 Correlation matrix between black soil layer thickness and topographic factors

在研究区内部,黑土层厚度存在显著的空间异质性,一方面是自然成土过程的结果;另一方面,受到土壤侵蚀和沉积的影响<sup>[30]</sup>。调查小流域黑土层厚度发现,小流域内黑土层厚度呈明显的空间分异规律,高海拔陡峭处浅薄,低洼平缓处相对深厚,从侧面印证土壤侵蚀过程对黑土层厚度空间格局的重要塑造作用。基于上述分析,本研究选取的地形因子与黑土层厚度变化之间存在显著相关性,因此,在预测模型中引入地形协变量将有望显著提升黑土层厚度的空间预测精度。

### 3.3 研究优势、不足与未来方向

本研究通过引入作物生长期遥感信息,有效提升黑土层厚度反演的精度,突破传统方法仅依赖裸土期影像的局限,体现多时相遥感数据在黑土属性建模中的重要作用。同时,本研究构建“多时相遥

感+地形因子+机器学习”的融合技术框架,为区域土壤属性高分辨率制图提供一种可复制、可推广的新路径。所生成的8 m分辨率黑土层厚度分布图,相比以往基于低分辨率遥感或传统地统计方法的研究,在空间细节表征能力方面具有明显优势,不仅为黑土区精准农业提供更可靠的空间决策依据,也为景观条件相似的其他黑土区域提供方法参考。

尽管如此,本研究仍存在一定的局限性。首先,所采用的SRTM DEM数据分辨率与精度有限,对微地形特征的表达能力不足,可能影响坡度、曲率等派生地因子的提取准确性,进而弱化其与黑土层厚度之间关系的揭示能力。未来研究可引入无人机遥感或激光雷达技术,生成高精度DEM数据,以更精确地刻画微地形特征,深化侵蚀-沉积过程对黑土厚度空间分异的影响机制理解。其次,本研究的实地样本数量相对有限,虽已通过交叉验证和正则化等手段增强模型稳健性,但仍在一定程度上制约了模型对复杂空间异质性的捕捉能力,存在过拟合风险。后续工作需开展更系统、规模化的土壤采样调查,构建更具代表性的样本集,进一步提升模型的泛化能力和预测精度。未来将通过引入更高精度DEM和扩大样本库,持续优化黑土层厚度制图方法的精度与鲁棒性,为黑土资源监测与可持续管理提供更坚实的科学支撑。

## 4 结论

1)作物生长期遥感信息的引入显著提升了黑土层厚度的反演精度。相比仅依赖裸土期遥感数据的模型,联合裸土期与生长期遥感特征的XGBoost模型在 $R^2$ 上提高0.1,表明多时相遥感数据对黑土层厚度的空间预测具有重要增强作用。

2)地形因子的引入有助于进一步优化模型性能。相关分析显示,高程与黑土层厚度呈显著负相关( $r=-0.26$ ),而受研究区地形平缓影响,坡度和曲率与黑土层厚度的相关性较弱( $r=0.048$ ),但仍对模型优化具有一定贡献。

3)在多种机器学习模型中,XGBoost算法表现最优( $R^2=0.60$ , $RMSE=12.31$ ),其预测精度显著高于GBDT( $R^2=0.53$ , $RMSE=13.27$ )和随机森林模型( $R^2=0.46$ , $RMSE=14.33$ )。结合地形因子后,模型解释能力得到进一步提升。本研究提出的“多时相遥感+地形因子+机器学习”技术框架,不仅拓展土壤属性遥感制图的方法路径,也为精准农业实践与黑土资源可持续管理提供高精度空间决策依据,对国家黑土保护与粮食安全战略具有重要支撑意义。

## 参考文献:

- [1] XIE Y, LIN H H, YE Y, et al. Changes in soil erosion in cropland in northeastern China over the past 300 years[J].*Catena*, 2019, 176: 410-418.
- [2] 范昊明,蔡强国,陈光,等.世界三大黑土区水土流失与防治比较分析[J].*自然资源学报*, 2005, 20(3): 387-393.  
FAN H M, CAI Q G, CHEN G, et al. Comparative study of the soil erosion and control in the three major black soil regions in the world [J]. *Journal of Natural Resources*, 2005, 20(3): 387-393.
- [3] 张晓平,梁爱珍,申艳,等.东北黑土水土流失特点[J].*地理科学*, 2006, 26(6): 687-692.  
ZHANG X P, LIANG A Z, SHEN Y, et al. Erosion characteristics of black soils in northeast China [J]. *Scientia Geographica Sinica*, 2006, 26(6): 687-692.
- [4] 刘若轩,王志强,谭玉萍.东北典型黑土区小流域黑土层厚度及影响因素[J].*水土保持学报*, 2024, 38(4): 346-353.  
LIU R X, WANG Z Q, TAN Y P. The thickness of black soil layer and its influencing factors in a small watershed of typical black soil region in northeast China [J]. *Journal of Soil and Water Conservation*, 2024, 38(4): 346-353.
- [5] TESFA T K, TARBOTON D G, CHANDLER D G, et al. Modeling soil depth from topographic and land cover attributes [J]. *Water Resources Research*, 2009, 45(10): W10438.
- [6] GAO Q Z, KANG M Y, XU H M, et al. Optimization of land use structure and spatial pattern for the semi-arid loess hilly-gully region in China [J]. *Catena*, 2010, 81(3): 196-202.
- [7] KURIAKOSE S L, DEVKOTA S, ROSSITER D G, et al. Prediction of soil depth using environmental variables in an anthropogenic landscape, a case study in the western ghats of Kerala, India [J]. *Catena*, 2009, 79(1): 27-38.
- [8] ZHANG Y, GUO L, CHEN Y Y, et al. Prediction of soil organic carbon based on landsat 8 monthly NDVI data for the Jiangnan Plain in Hubei Province, China [J]. *Remote Sensing*, 2019, 11(14): e1683.
- [9] DE SOUSA MENDES W, DEMATTÊ J A M, SILVERO N E Q, et al. Integration of multispectral and hyperspectral data to map magnetic susceptibility and soil attributes at depth: A novel framework [J]. *Geoderma*, 2021, 385: e114885.
- [10] YANG C, YANG L, ZHANG L, et al. Soil organic matter mapping using INLA-SPDE with remote sensing

- based soil moisture indices and Fourier transforms decomposed variables[J].*Geoderma*,2023,437:e116571.
- [11] KOBAYASHI Y, YOSHIDA K. Quantitative structure-property relationships for the calculation of the soil adsorption coefficient using machine learning algorithms with calculated chemical properties from open-source software [J]. *Environmental Research*, 2021, 196:e110363.
- [12] XU K K, DAI H M, ZHANG X J, et al. Black soil layer thickness prediction and soil erosion risk assessment in a small watershed in northeast China[J].*PLoS One*, 2025, 20(6):e0324368.
- [13] ZHANG S, LIU G, CHEN S L, et al. Assessing soil thickness in a black soil watershed in northeast China using random forest and field observations [J].*International Soil and Water Conservation Research*, 2021, 9(1):49-57.
- [14] GONG Z T, CHEN H Z, WANG H L. The distribution regularity of higher categories in Chinese soil taxonomy [J].*Chinese Geographical Science*,1997,7(1):1-10.
- [15] Soil Survey Division Staff. Soil survey manual [M]. Washington D C: US Department of Agriculture,1993.
- [16] SCHOENEBERGER P J, WYSOCKI D A, BENHAM E C, et al. Field book for describing and sampling soils [M]. USDA-NRCS: US Government Printing Office,2012.
- [17] BREIMAN L. Random forests [J].*Machine Learning*, 2001,45(1):5-32.
- [18] FRIEDMAN J H. Greedy function approximation: A gradient boosting machine [J].*Annals of Statistics*, 2001, 29: 1189-1232.
- [19] CHEN T Q, GUESTRIN C. XGBoost: A scalable tree boosting system [C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. San Francisco California USA. ACM,2016:785-794.
- [20] 张之一. 黑土开垦后黑土层厚度的变化[J]. 黑龙江八一农垦大学学报, 2010,22(5):1-3.  
ZHANG Z Y. The thickness changes of ah horizon after the phaeozems cultivated[J].*Journal of Heilongjiang Bayi Agricultural University*,2010,22(5):1-3.
- [21] BASTIDA F, ZSOLNAY A, HERNÁNDEZ T, et al. Past, present and future of soil quality indices: A biological perspective[J].*Geoderma*,2008,147(3/4):159-171.
- [22] ZHANG Y, LUO C, ZHANG Y H, et al. Integration of bare soil and crop growth remote sensing data to improve the accuracy of soil organic matter mapping in black soil areas[J].*Soil and Tillage Research*,2024,244:e106269.
- [23] XIONG Y L, LI Y R, XIONG S Y, et al. Multi-scale spatial correlation between vegetation index and terrain attributes in a small watershed of the upper Minjiang River[J].*Ecological Indicators*,2021,126:e107610.
- [24] 刘焕军, 鲍依临, 孟祥添, 等. 黑土区田块尺度下地形影响作物长势机理分析[J]. 农业机械学报, 2020, 51(9): 274-283.  
LIU H J, BAO Y L, MENG X T, et al. Effect of topography on crop growth in black soil area[J].*Transactions of the Chinese Society for Agricultural Machinery*, 2020, 51(9):274-283.
- [25] 刘凯, 戴慧敏, 刘国栋, 等. 基于土壤有机碳含量的黑土层厚度预测及影响因素分析[J]. 物探与化探, 2024, 48(5):1368-1376.  
LIU K, DAI H M, LIU G D, et al. Organic carbon content-baed prediction and influencing factors of black soil layer thicknesses [J]. *Geophysical and Geochemical Exploration*,2024,48(5):1368-1376.
- [26] 刘金涛, 赵薇, 刘彧. 土壤厚度演化模型理论方法研究进展[J]. 土壤学报, 2024, 61(2):319-330.  
LIU J T, ZHAO W, LIU Y. Modelling soil thickness evolution: Advancements and challenges [J].*Acta Pedologica Sinica*,2024,61(2):319-330.
- [27] GABET E J, MUDD S M, WOOD R W, et al. Hilltop curvature increases with the square root of erosion rate [J]. *Journal of Geophysical Research: Earth Surface*, 2021,126(5):e2020JF005858.
- [28] ROERING J J. How well can hillslope evolution models explain topography? Simulating soil transport and production with high-resolution topographic data [J].*Geological Society of America Bulletin*,2008,120(9/10):1248-1262.
- [29] PATTON N R, LOHSE K A, GODSEY S E, et al. Predicting soil thickness on soil mantled hillslopes [J]. *Nature Communications*,2018,9:e3329.
- [30] FENG Z Z, ZHENG F L, HU W, et al. Impacts of mollic epipedon thickness and overloaded sediment deposition on corn yield in the Chinese mollisol region [J]. *Agriculture, Ecosystems and Environment*, 2018, 257: 175-182.