

DOI: 10.11766/trxb202508100388

CSTR: 32215.14.trxb202508100388

覃建祥, 孔超, 石镇南, 张海涛, 陈家赢, 王天巍, 宋效东, 李德成, 赵玉国, 张甘霖, 郭龙. 基于土壤成土理论与机器学习的历史土壤图分类逐级更新研究[J]. 土壤学报, 2026,

QIN Jianxiang, KONG Chao, SHI Zhennan, ZHANG Haitao, CHEN Jiaying, WANG Tianwei, SONG Xiaodong, LI Decheng, ZHAO Yuguo, ZHANG Ganlin, GUO Long. Hierarchical Classification Update of Historical Soil Maps Based on Pedogenic Theory and Machine Learning[J]. Acta Pedologica Sinica, 2026,

基于土壤成土理论与机器学习的历史土壤图分类逐级更新研究*

覃建祥¹, 孔超¹, 石镇南¹, 张海涛¹, 陈家赢¹, 王天巍¹, 宋效东², 李德成², 赵玉国², 张甘霖^{2,3}, 郭龙^{1,2†}

(1.华中农业大学资源与环境学院, 武汉 430070; 2.土壤与农业可持续发展全国重点实验室(中国科学院南京土壤研究所), 南京 211135; 3.中国科学院南京地理与湖泊研究所, 南京 211135)

摘要: 土壤图是土地资源管理、农业规划与生态保护的重要基础数据, 其精度直接影响耕地质量评价、作物适种区划及土壤碳储量估算等。受限于传统土壤调查的高成本与低更新频率, 历史土壤图存在更新滞后、分类系统不完善及数据精度不足等问题, 难以满足当前土壤资源动态监测的精度与时效性需求, 需要进行有效的更新。如何从历史土壤图中挖掘有效的土壤-环境关系知识, 并结合当前最新的土壤信息进行土壤图的更新, 是一个值得探讨的问题。为此, 本研究以湖北省京山市第二次全国土壤普查形成的历史土壤图为基础, 以土地利用、母质、遥感和地形等多源异构数据为支撑, 从土壤形成理论出发, 借助于机器学习算法和基础数据提取土壤类型关键划分指标, 构建从高级(土类)到基层(土种)的土壤分类逐级更新策略。研究结果表明: (1) 分类逐级更新策略能有效完成研究区历史土壤图的更新, 经 30 个野外样点独立验证, 土类与亚类的更新精度达到 90%, 土属与土种的识别准确率达到 83%和 73%; (2) 更新后土壤分类结果有所调整, 土类数维持 5 个不变, 亚类从 13 个减少至 12 个(剔除沼泽型水稻土), 土属从 43 个调整为 38 个(剔除 14 个, 新增 9 个), 土种从 283 个增加至 323 个; (3) 从空间格局看, 约 22%的区域土类发生了显著变化, 其中潮土面积较第二次全国土壤普查增加约 172%, 水稻土面积减少约 32%, 亚类与土属的变化面积分别为 32%和 36%, 土种命名规则重构后, 大部分图斑中的土种类型发生明显改变。本研究验证了基于成土理论与机器学习的分类逐级更新框架在提升历史土壤图精度、完善分类体系及揭示空间格局演变方面的有效性, 也为历史土壤图的高效更新提供了参考。

关键词: 历史土壤图更新; 土壤成土理论; 逐级土壤分类; 数字土壤制图; 机器学习算法

中图分类号: S159

文献标志码: A

Hierarchical Classification Update of Historical Soil Maps Based on Pedogenic Theory and Machine Learning

QIN Jianxiang¹, KONG Chao¹, SHI Zhennan¹, ZHANG Haitao¹, CHEN Jiaying¹, WANG Tianwei¹, SONG

* 国家重点研发计划项目(2023YFD1500102, 2023YFC3709801)和国家自然科学基金项目资助(42571065) Supported by the National Key Research and Development Program of China (Nos.2023YFD1500102, 2023YFC3709801), and National Natural Science Foundation of China (No.42571065)

† 通讯作者 Corresponding author, E-mail: guolong@mail.hzau.edu.cn

作者简介: 覃建祥(1997-), 男, 广西贵港人, 硕士研究生, 主要从事数字土壤制图。E-mail:1450297013@qq.com
收稿日期: 2025-08-10; 收到修改稿日期: 2025-12-14; 网络首发日期(www.cnki.net):

Xiaodong², LI Decheng², ZHAO Yuguo², ZHANG Ganlin^{2,3}, GUO Long^{1,2†}

(1.College of Resources and Environment, Huazhong Agricultural University, Wuhan 430070, China; 2.State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 211135, China; 3. Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 211135, China)

Abstract: 【Objective】 Soil maps are essential for land resource management, agricultural planning, and ecological conservation. Their accuracy directly affects applications such as arable land quality assessment, crop suitability zoning, and soil carbon stock estimation. However, limited by the high cost and low frequency of traditional surveys, historical soil maps often suffer from outdated classifications, incomplete taxonomy, and low precision, making them inadequate for current monitoring needs. This study explores how to extract effective soil–environment relationships from historical maps and combine them with new sample data to support efficient updates. 【Method】 Using the soil map from the second national soil survey of Jingshan City (Hubei Province) as a baseline, this study integrates multi-source data—including land use, parent material, remote sensing, and topography—and applies pedogenic theory to guide a hierarchical soil classification updating strategy. Key classification indicators were extracted via machine learning algorithms to iteratively update soil types from soil order to soil species. 【Result】 (1) The proposed strategy effectively updated the historical soil map. Validation using 30 field samples showed classification accuracies of 90% for soil order and suborder, and 83% and 73% for soil group and species. (2) Classification structure was adjusted: soil orders remained at five; suborders decreased from 13 to 12; soil groups from 43 to 38 (14 removed, 9 added); and soil species increased from 283 to 323. (3) Spatially, about 24% of the area underwent soil order changes, with fluvisols increasing by 172% and paddy soils decreasing by 32%. Suborder and group-level changes covered 32% and 36% of the area, respectively. Most soil species categories were revised due to naming logic reconstruction. 【Conclusion】 This study verified the effectiveness of the classification hierarchical update framework based on soil formation theory and machine learning in improving the accuracy of historical soil maps, improving the classification system, and revealing the evolution of spatial patterns. It also provided feasibility and reference for the efficient updating of historical soil maps.

Key words: Historical soil map updating; Pedogenic theory; Hierarchical soil classification; Digital soil mapping; Machine learning algorithms

土壤图是开展耕地质量评价、作物适宜性评价、生态保护与碳储量评估等工作的重要基础性资料^[1]。目前我国仍广泛使用第二次全国土壤普查时形成的土壤图,然而受自然环境演变与人类活动的双重影响,其所反映的土壤类型名称、边界及理化性质等已发生显著变化^[2]。因此,推进历史土壤图的高精度更新,提升土壤类型空间信息的现势性与表达能力,已成为当前土壤资源调查与管理的关键任务^[3-4]。

土壤发生分类系统是在成土过程理论基础上建立的层级化分类框架,依据土壤形态特征、诊断层属性及形成机制,对土壤类型进行系统划分^[5]。作为连接土壤形成机制与空间分布表达的重要桥梁,该系统不仅体现了气候、母质、地形、生物与时间等成土因子的协同作用路径,也为土壤图的分级更新和命名规范提供了科学依据^[6-7]。在地貌复杂或土壤类型多样的区域,该系统分层清晰、逻辑严密的结构为机器学习建模提供了分类依据与逻辑约束,有助于提升模型在空间预测中的结构准确性与对成土过程的解释力^[8]。

传统土壤图更新方法以野外调查与专家判读为主,具有可操作性强、精度高等特点,但也存在周期长、成本高等缺点,且在地形复杂、交通不便地区效率显著受限^[9]。随着遥感、地理信息系统与机器学习技术的发展,基于环境协变量的数字土壤制图逐步替代人工方法,成为现代土壤学制图与更新的主要路线^[10-12]。早期制图方法多采用线性统计模型^[13],但面对土壤与环境因子间高度非线性和空间异质性的关系,存在较大的局限性和不确定性^[14-16]。近年来,随机森林(Random Forest, RF)、极限梯度提升(Extreme Gradient Boosting, XGBoost)与轻量级梯度提升机(Light Gradient Boosting Machine, LightGBM)等非线性模型在土壤类型预测中取得突破,显著提升了模型对高维、多源异构

数据的处理能力^[17-18]。如 Aydin 等^[19]采用 XGBoost 与 LightGBM 模型,获得大于 90%的分类准确率,展现了该模型在捕捉土壤与环境因子复杂非线性关系中的优势。同样, Liu 等^[20]应用 RF 模型,在高维数据和类别不平衡问题中展现了良好的准确性和鲁棒性。然而,此类模型高度依赖数据驱动,缺乏成土过程约束,易出现“黑箱”问题,在样本稀疏或异质性显著区域存在失真风险^[21-22]。

为此,越来越多的研究尝试将成土过程理论引入模型构建与分类逻辑中,通过引入母质约束、分层分类与属性特征耦合等策略,提升预测结果的科学性与可解释性^[23-24]。这一融合路径既保障了土壤类型划分的发生学合理性和科学性^[25],又增强了制图结果的空间表达能力,逐步成为复杂地形与多样成土环境下历史土壤图更新的研究趋势^[26]。

本研究以湖北省京山市为例,融合土壤成土理论与机器学习方法,构建了基于逐级分层约束的 4 级土壤名称(土类-亚类-土属-土种)更新框架,系统整合历史土壤图、多源环境变量与野外踏勘数据,依托土壤属性数据、命名规则与空间叠置逻辑,实现了县域尺度下土壤图的分层逐级更新。研究旨在突破历史土壤图更新的效率瓶颈,探索面向复杂地貌区的土壤图更新机制,并为区域土壤资源管理、农业可持续利用与生态调控提供科学依据和技术支撑。

1 材料与方法

1.1 研究区概况

本研究区位于湖北省荆门市京山市,地处 $30^{\circ}42' - 31^{\circ}29' N$ 、 $112^{\circ}43' - 113^{\circ}29' E$ 之间,总面积约 3520 km^2 ,属鄂中低山丘陵向江汉平原的过渡地带。如图 1 所示,地势总体西北高、东南低,区域内地貌类型多样,主要有低山、丘陵、岗地和平原,地形起伏明显。京山市属中亚热带向北亚热带过渡区的典型季风气候,年均气温 16.3°C ,年降水量 1067.1 mm ,气候四季分明。水系方面,区域内河流密布,北接大富水水系,南临汉江流域,地表水资源相对丰富。据《京山县土壤志》^[27]等资料,京山市土壤类型丰富,主要包括水稻土、潮土、黄棕壤、石灰岩土和紫色土 5 个土类,细分为 13 个亚类、43 个土属及 283 个土种。研究区成土母质多样,包括砂页岩、第四纪更新世黄土、富钙黏土、河流冲积物等多种类型。

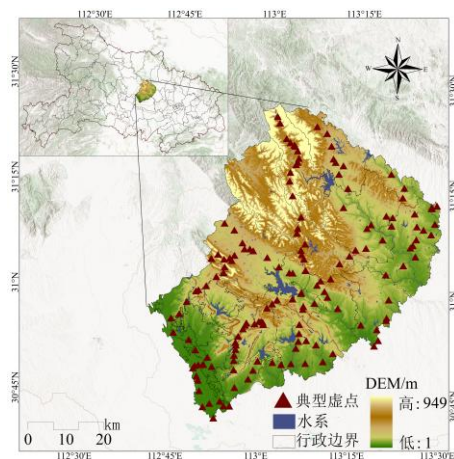


图 1 研究区概况图

Fig. 1 Overview map of the study area

1.2 基础数据来源及预处理

本研究所用数据包括历史土壤图、土地利用现状、土壤关键属性、地形、遥感和气候等 6 大类,构建了由 338 个环境因子组成的多源异构数据集。在建模与分析前,完成数据的格式转换、掩膜提取、坐标投影、空间配准及空间重采样等处理,确保数据在空间分辨率、坐标系及分析区域上的一

致性。

(1) 基础土壤数据。基础土壤数据包括土壤志、历史土壤图、土壤关键属性数据、母质数据、典型虚设点样本及野外实测样点数据。其中历史土壤图源自第二次全国土壤普查 1:5 万纸质图件,经数字化矢量处理与分类标准规范化转换后,依据《第三次全国土壤普查土壤分类系统(试行版)》(2023)

(以下简称《土壤分类系统(试行版)》)重新识别为 5 个土类、12 个亚类、39 个土属和 121 个土种,土壤图中土种数量与《土壤志》记载存在差异,不仅与历史图分类精度相关,也受制图比例尺、最小图斑面积及制图方法等因素共同影响。土壤关键属性数据主要来源于国家科技资源共享服务平台—国家地球系统科学数据中心—土壤分中心(<http://soil.geodata.cn>),涵盖了研究区的土体厚度、质地、砾石含量及 pH 等关键属性。母质数据基于 1:20 万地质图和历史土壤图提取岩性、岩相、岩色与地层年代等属性。通过图幅拼接与坐标统一,构建水稻土和非水稻土的母岩—母质映射表,生成分类明确、空间精度一致的母质图。典型虚设点样本选取自第二次全国土壤普查以来土壤类型保持稳定、未发生显著变化的区域,结合成土环境、遥感、地形和土壤属性等多维因子确定土壤类型的合理性和可靠性。通过分层均匀抽样与概率分布相结合的方式布设采样点,以确保样本在成土环境与空间分布上的代表性,最终获得 1 019 个典型虚设点。此外,为提高模型准确性与样本代表性,本研究依据空间均匀性、地貌格局和成土环境差异,在研究区内布设采样点,重点覆盖淹育水稻土、潜育水稻土等小面积类型,并确保平原、丘陵与冲积平原等典型地貌均得到有效覆盖。最终共采集野外实测样点 157 个(如图 1 所示),记录土壤类型、关键土壤属性数据(质地、砾石含量、土体厚度、母质、障碍层/特征层与土地利用)等,并和典型虚设点共同构成模型训练与验证的基础样本集。

(2) 基础环境变量。基础环境变量包括地形、土地利用、气候、遥感及植被指数。地形数据来自国家地球系统科学数据共享平台—地理空间数据云,分辨率为 30 m,本研究利用数字地形分析方法提取坡度、坡向、地形湿度指数、相对高程等,最终获得包括 DEM 等 19 类地形因子。土地利用现状数据引自中国科学院地理资源所提供的 1980—2020 年中国土地利用数据库,数据将研究区划分为耕地(旱地和水田)、林地、草地、水域等 14 类土地利用类型,全面反映了不同时期人类活动对土地格局的影响。为分析人类活动对土壤类型演变的驱动效应,研究选取 2020 年土地利用数据作为更新依据,并依据土壤分类的应用需求,将其重新归类为旱地、水田、林地、园地、草地、水系、建筑用地等 7 类。遥感数据基于 Sentinel-2 影像与 GEE 平台,计算 2020—2023 年期间每月中值,剔除因云雨遮挡无全域有效观测的月份后,提取归一化植被指数(Normalized Difference Vegetation Index, NDVI)等 10 类植被与水文指数,形成 264 个 30 m 分辨率的遥感变量。气候数据来自中国科学院成都山地灾害与环境研究所,涵盖气温、降雨量、风速、干燥度、径流量和水分盈亏量各月平均、年平均和年极值等数据,共 89 个气候因子。

1.3 土壤类型与关键土壤属性建模方法

本研究采用 RF、XGBoost 与 LightGBM 三种基于树结构的集成学习算法,构建土壤类型分类与关键土壤属性预测的统一建模体系。这 3 类模型均具备非线性建模能力,对噪声与异常值的鲁棒性强,适用于地形复杂、样本不均匀与数据异构性强的区域土壤制图任务。

其中,RF 模型通过 Bootstrap 重采样与多数投票机制集成多棵决策树,具备良好的泛化性能和抗过拟合能力,适用于土壤类型复杂、多样的非线性识别任务^[28]。XGBoost 在传统梯度提升树的基础上引入正则化项与剪枝策略,在处理高维数据与类别不平衡样本时表现优越,适合应用于障碍层识别、亚类与土属的细分类任务^[29]。LightGBM 通过直方图优化与叶子优先生长策略,提升训练效率并强化特征交互能力,尤其适用于大规模样本条件下的精细分类与属性预测^[30]。

在土壤类型分类建模中,3 种模型分别应用于亚类与灰潮土石灰性识别的分类任务,结合分层分类策略与命名规则约束,确保分类结果符合《土壤分类系统》的层级逻辑;在关键土壤属性建模中,模型用于预测影响土属与土种命名的判别性属性数据,如腐殖质层厚度、障碍层/特征层埋深等,以实现土壤内部结构的定量刻画与成土过程约束的间接表达。

模型构建过程中,各模型在相同的多源异构环境变量数据集上独立训练,但训练样本则按分类级别单独分配。为提升分类性能,本研究采用分级变量筛选与样本均衡策略:前者依据不同亚类模型对地形、气候和遥感因子的判别力差异,筛选最优变量组合;后者通过对样本量差异较大的类别实施均衡抽样,缓解类别不平衡问题,从而增强模型在稀疏样本区域和类别异质性显著区域的鲁棒性与预测精度。

1.4 模型评价与验证方法

为确保模型在不同类型任务中具备良好的识别精度与泛化能力,针对土壤类型分类任务和关键属性的数值预测任务分别建立了对应的评估体系。

(1) 分类模型评估指标。针对土壤类型的分类任务,采用总体精度(OA)与 Kappa 系数对模型性能进行评估。OA 衡量模型正确分类的样本在总样本中的比例,反映整体识别准确率。Kappa 系数用于校正预测一致性中随机因素的影响,适用于类别不平衡或样本数量差异较大的情形。公式分别为:

$$OA = \frac{\sum_{i=1}^N Tp_i}{\text{总样本数}} \times 100\% \quad (1)$$

式中, Tp_i 表示第 i 类土壤的正确预测数。

$$k = \frac{OA - P_e}{1 - P_e}, \quad P_e = \sum_{i=1}^N P_{\text{obs},i} \times P_{\text{pred},i} \quad (2)$$

式中, $P_{\text{obs},i}$ 和 $P_{\text{pred},i}$ 分别为第 i 类真实和预测样本占比, P_e 为随机预期一致率。

(2) 属性预测模型评估指标。对于腐殖质层厚度、障碍层/特征层埋深等关键属性的连续数值预测任务,选取均方根误差(RMSE)和决定系数(R^2)作为评估指标。RMSE 衡量模型预测值与真实值的平均偏离程度,数值越小表示预测误差越小。 R^2 衡量模型对数据变异的解释能力,越接近 1 表示模型拟合得越好,公式分别为:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2} \quad (4)$$

式中, y_i 为真实值, $\hat{y}_i$ 为模型预测值, $\bar{y}$ 为真实值的平均值。

(3) 验证策略与方法。模型训练采用训练集与测试集 7:3 的分层抽样划分,确保各类别样本在两个子集中的比例一致,从而提高评估结果的代表性与稳定性。为增强模型的空间泛化能力,在典型景观区域选取虚设点,结合采样数据进行外部独立验证,并将其用于后续图斑修正与模型反馈优化。

1.5 历史土壤图更新方法

本研究融合土壤成土理论与机器学习方法,构建基于分类逐级的更新框架体系(土类-亚类-土属-土种),先系统整合历史土壤图、多源环境变量与野外踏勘数据,通过历史土壤数据与土地利用的空间叠置判定类型变化区;继而完成 USDA 质地向国标 3/5 级转换与质地构型判定,生成土体厚度与砾质度栅格,并以 RF、XGBoost、LightGBM 两阶段建模估测灰潮土石灰性、腐殖质层(有无分类与厚度预测)及水稻土潜育/漂洗层(空间分布与埋深预测);随后按母质与成土理论判定土类,在土类先验下预测亚类,依据命名规则融合关键属性更新土属与土种,并结合野外样点独立校验与土壤成土逻辑的修正,最终完成分类完整、结构清晰的 4 级土壤图的联动式更新(如图 2)。

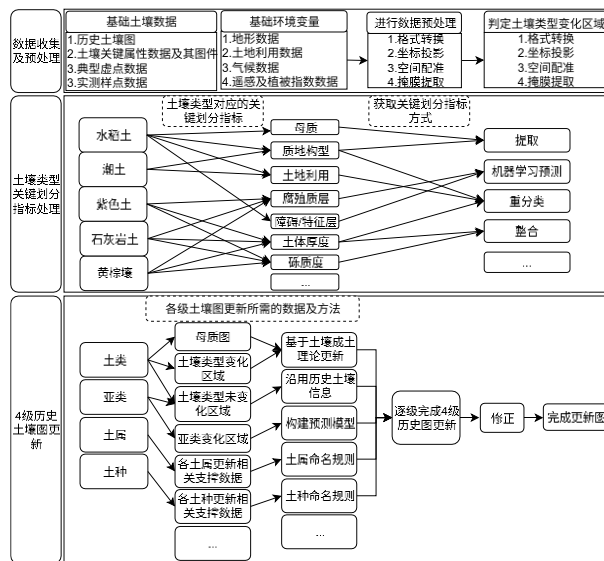


图 2 技术路线图
Fig. 2 Technology roadmap

2 结果

2.1 土壤类型关键划分指标

为支撑 4 级分类，本研究基于土壤关键属性数据与《土壤分类系统（试行版）》命名规则，提取并标准化质地、土体厚度与砾质度 3 类指标，进一步构建石灰性灰潮土、腐殖质层与障碍层/特征层的空间分布及厚度（深度）预测模型，完成分类逐级更新所需的关键划分指标图层，用于支撑不同级别土壤类型的判定与命名（如表 1）。

表 1 不同土壤分类级别的关键划分指标

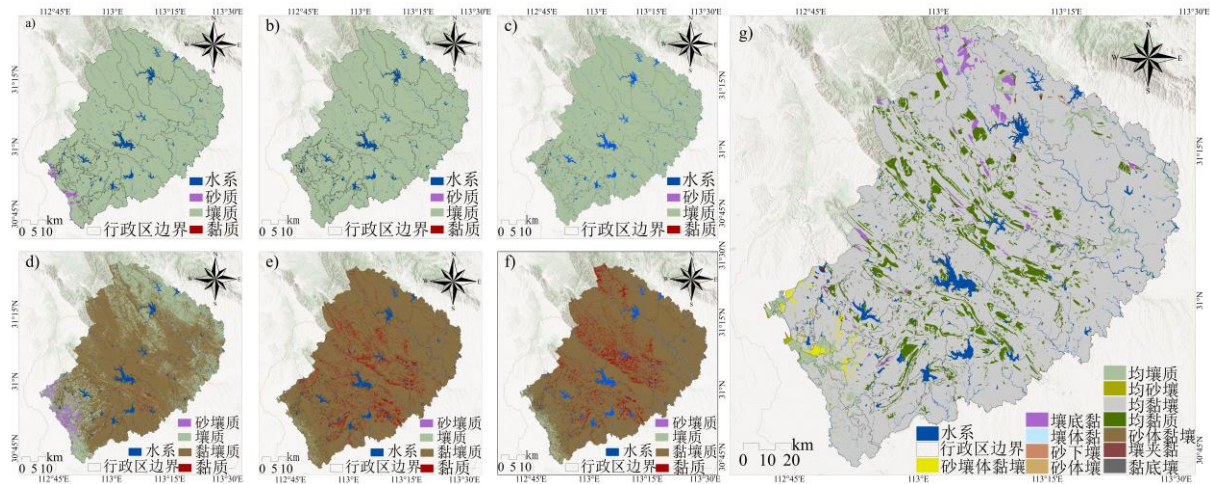
Table 1 Key diagnostic indicators for different taxonomic levels in soil classification

土壤分类级别 Soil classification level	水稻土 Paddy soil	潮土 Fluvo-aquic soil	黄棕壤 Yellow-brown soil	石灰岩土 Limestone soil	紫色土 Purple soil
土类 Soil order	土地利用类型	土地利用类型	母质	母质	母岩、母质
亚类 Soil suborder	模型预测	/	模型预测	模型预测	模型预测
土属 Soil group	亚类定性词、母质	石灰性	母质	/	质地、砾质度
土种 Soil species	障碍层/特征层、表层质地、质地构型	表层质地、质地构型	土地利用类型、腐殖质层、土体厚度、表层质地	土地利用类型、腐殖质层、土体厚度、表层质地	土地利用类型、腐殖质层、土体厚度

（1）土壤质地、质地构型、土体厚度和砾质度数据。土壤质地数据为 0~100 cm 剖面质地的美国农业部（United States Department of Agriculture, USDA）12 级标准。本研究在依据国标 3 级、5 级质地与美国制 12 级质地转换规则的基础上，将其分别转换为 3 级（砂质、壤质、黏质）和 5 级（砂质、砂壤质、壤质、黏壤质、黏质）质地类型。在 5 级质地基础上，结合均质型、夹层型、体型、底型与上下型的深度区间要求，结合质地变化规律，并以质地等级差异大于 2 级为判据，判定土体质

地构型类型,通过对质地数据编码与剖面叠置解析获得构型分布图。如图3所示,研究区表层与深层土壤在3级质地中以壤质为主,而在5级质地中,表层以砂壤和壤土为主,质地构型则呈现出11种类型,以均黏壤质为主,为水稻土与潮土等命名的判定提供依据。

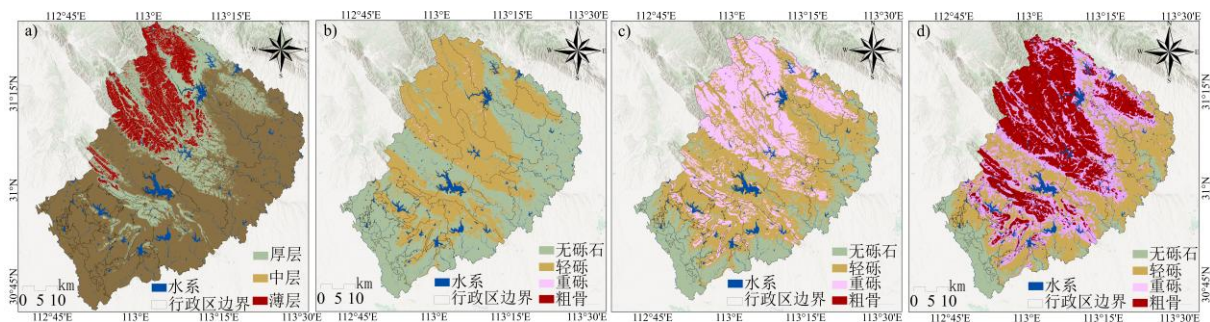
土体厚度数据根据《土壤分类系统(试行版)》标准,结合母质分布与地形特征,将土体厚度重分类为薄层、中层和厚层,结果显示(图4a),北部山区以薄层为主,其余区域多为中厚层。砾质度同样依据标准并结合母质差异设定阈值,划分为无砾石、轻砾、重砾和粗骨(图4b~图4d),砾质度总体呈东南向西北增强的趋势,随海拔升高由轻砾逐渐过渡至重砾和粗骨类型。所有指标经标准化处理后生成栅格图层,用于黄棕壤、石灰岩土和紫色土等类型的属性识别与基础类型判定。



注: a)~c): 0~30 cm、30~60 cm、60~100 cm 的3级质地, d)~f): 0~30 cm、30~60 cm、60~100 cm 的5级质地, g): 0~100 cm 的质地构型。Note: a)~c): 3-level textures at 0~30 cm, 30~60 cm, and 60~100 cm; d)~f): 5-level textures at 0~30 cm, 30~60 cm, and 60~100 cm; g): texture configuration at 0~100 cm.

图3 土壤3级、5级质地及质地构型分布图

Fig. 3 Spatial distribution of soil textures (3-level and 5-level) and texture configuration



注: a): 土体厚度, b)~d): 0~20 cm、20~60 cm、60~100 cm 的砾质度。Note: a): soil thickness; b)~d): gravel content at 0~20 cm, 20~60 cm, and 60~100 cm, respectively.

图4 土体厚度及砾质度空间分布图

Fig. 4 Spatial distribution of soil thickness and gravel content

(2) 石灰性灰潮土预测、腐殖质层、障碍层/特征层预测。为实现石灰性灰潮土、腐殖质层厚度以及水稻土障碍层/特征层的空间表达,本研究基于典型虚设点与采样数据,构建了RF、XGBoost与LightGBM 3种模型进行独立建模,输入变量均为与相应成土过程密切相关的环境因子,并采用分层建模策略,先分类判别空间分布区域,再进行属性回归预测。

在石灰性灰潮土识别中,针对石灰性与非石灰性样本不平衡问题,采用均衡抽样构建训练集,结果显示(如表2)RF模型表现最佳(OA=0.94, Kappa=0.93),成功区分两类灰潮土,并生成其空

间分布图（图 5 a），其中石灰性灰潮土集中分布于南部汉江冲积平原，非石灰性灰潮土沿内陆河流呈带状分布。

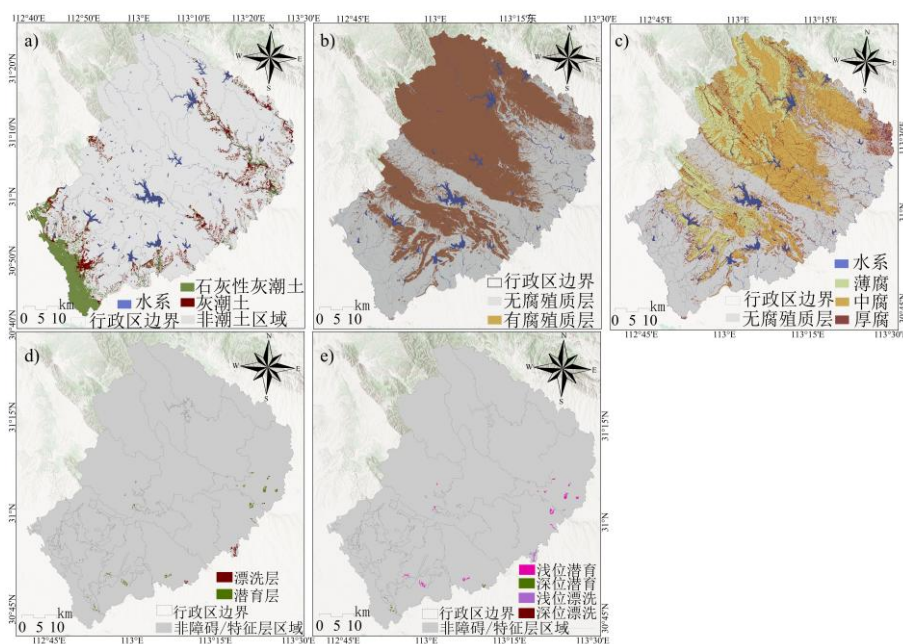
在腐殖质层建模中，首先通过分类模型确定其分布范围，再在区域内利用回归模型预测厚度。空间分布预测中 RF 模型表现最佳（OA=0.93，Kappa=0.92），而 XGBoost 在厚度预测中精度最高（RMSE=10.83， $R^2=0.60$ ）。结果表明，腐殖质层主要分布于中北部山区，以“薄腐”为主，为土属和土种中腐殖质等级划分提供了定量依据（图 5 b、图 5c）。

在水稻土障碍层/特征层预测中，采用相同建模策略识别潜育层与漂洗层的空间-垂向分布。XGBoost 在分类与回归任务中均表现最优（OA=0.86，Kappa=0.86；RMSE=9.81， $R^2=0.66$ ），最终生成空间分布与埋深图层。结果显示（图 5 d、图 5e），浅位潜育层主要分布于东部低洼区，漂洗层集中于东南边缘，为水稻土种划分提供了关键支撑。

表 2 石灰性灰潮土空间分布、腐殖质层空间分布预测及厚度、障碍层/特征层空间分布及埋深预测模型精度

Table 2 Accuracy of prediction models for calcareous grey fluvo-aquic soil , spatial distribution and thickness of humus layer, and spatial distribution and burial depth of obstacle/characteristic layers

模型 Model	OA	Kappa 系数	RMSE	R^2
石灰性灰潮土空间分布预测最优模型（RF） Optimal model for predicting spatial distribution of calcareous soil (RF)	0.94	0.93	/	/
腐殖质空间分布及其厚度预测最优模型（RF/XGBoost） Optimal model for predicting humus spatial distribution and thickness (RF/XGBoost)	0.93	0.92	10.83	0.60
障碍层/特征层空间分布及其埋深位置预测最优模型（XGBoost/XGBoost） Optimal model for predicting the spatial distribution of obstacles/feature layers and their buried depth (XGBoost/XGBoost)	0.86	0.86	9.81	0.66



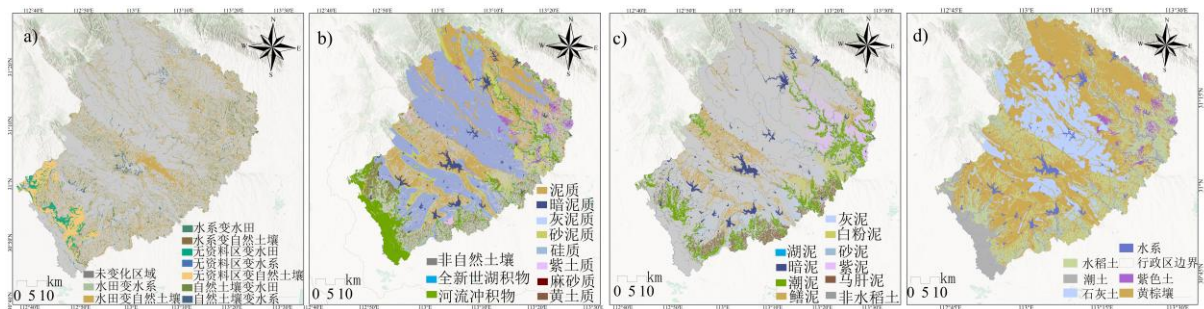
注：a)为石灰性灰潮土空间分布预测图，b)~c)为腐殖质层空间分布及厚度预测图，d)~e)为障碍层/特征层空间分布及埋深预测图。Note:a): predicted spatial distribution of calcareous gray soil; b)~c): predicted spatial distribution and thickness of the humus layer; d)~e): predicted spatial distribution and burial depth of the barrier/characteristic layer.

图 5 土壤类型更新关键划分指标分布图

Fig. 5 Spatial distribution of key indicators for updating soil types

2.2 历史土壤图分类逐级更新

(1) 土类图更新。历史土类图更新以成土过程理论为指导,融合历史土壤图、土地利用现状与母岩母质信息,通过空间叠置分析识别土壤类型变化区域(如图 6 a),并整合提取完善母质信息(如图 6 b、图 6 c),结合分类规则与环境因子完成土类更新。对未发生显著变化区域,判定历史土壤类型正误后沿用其信息;对变化区域,按照以下规则更新:水稻土依据水田和水耕熟化特征进行判别;潮土以发育于河流冲积母质上的非水田耕作土壤进行判定;紫色土依据其发育于紫色母岩区域的特征确定。针对石灰岩土与黄棕壤母岩交叉区域,则依据土壤发生过程,优先判定典型碳酸盐岩环境中受淋溶与黏化作用主导形成的石灰岩土,其余区域依据亚热带气候主导的水热成土特征归为黄棕壤。最终,通过分类规则与空间判别逻辑,完成土类图的系统更新(如图 6 d),为后续各级分类建模提供基础支撑。



注: a 为土壤类型变化区域空间分布图, b、c 为分别为自然土壤和水稻土母质空间分布图, d 为更新后土类空间分布图。Note: a: spatial distribution of soil order change areas; b~c: spatial distribution of parent material for natural soils and paddy soils, respectively; d: spatial distribution of soil order after update.

图 6 土类更新关键信息空间分布图

Fig. 6 Spatial distribution of key information for soil order update

(2) 亚类图更新。亚类更新基于土类更新结果,采用分层建模策略,分别利用 RF、XGBoost 与 LightGBM 构建黄棕壤、水稻土、紫色土与石灰岩土的亚类预测模型。模型以多源异构的环境变量数据集作为输入,以 1019 个典型虚设点作为样本数据(如图 7 a),训练集与测试集按 7:3 划分。亚类划分延续历史分类体系,其中潮土仅对应 1 类,可直接继承其历史亚类信息,黄棕壤 2 类,水稻土 4 类,紫色土 3 类,石灰岩土 2 类,为避免类别干扰,这 4 个土类采用分层建模策略针对不同土壤类型进行单独建模。训练中采用样本均衡策略控制类别比例,模型性能以 OA 与 Kappa 系数进行评估。预测过程中引入土类-亚类层级约束,确保亚类输出逻辑一致。最终输出为:类型变化区域采用模型预测结果,未变化区域继承历史亚类信息,二者合并形成亚类更新图,并结合野外采样数据作为外部独立验证以修正亚类更新结果(如表 3),最终获得亚类更新图(如图 7 b)。

表 3 各亚类最优预测模型精度

Table 3 Accuracy of the optimal prediction model for each suborder

亚类预测模型 Soil suborder prediction model	OA	Kappa 系数
黄棕壤亚类最优预测模型 (RF) Optimal Prediction Model for Suborders of Yellow-Brown Soils (RF)	0.92	0.92
水稻土亚类最优预测模型 (XGBoost) Optimal Prediction Model for Suborders of Paddy Soils (XGBoost)	0.92	0.91
石灰岩土亚类最优预测模型 (RF) Optimal Prediction Model for Suborders of Limestone Soils (RF)	0.93	0.92
紫色土亚类最优预测模型 (LightGBM) Optimal Prediction Model for Suborders of Purple Soils (LightGBM)	0.91	0.90

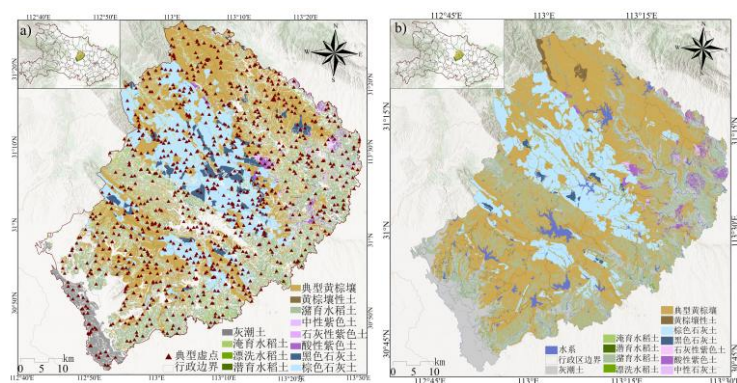


图 7 典型虚设点分布图及更新后亚类空间分布图

Fig. 7 Distribution of typical virtual points and spatial distribution of suborders after update

(3) 土属图更新。土属更新依托亚类更新结果，融合母质类型、砾质度等关键土壤类型划分指标，结合《土壤分类系统（试行版）》命名规则，采用空间叠置与分类判别方法构建研究区土属更新图。其中，黄棕壤与水稻土土属以“母质定性词+亚类名称”进行命名，亚类更新图与母质图进行空间叠置后，提取泥质、黄土质、灰泥质等母质类型，作为定性词用于命名黄棕壤土属类型。水稻土在提取对应水稻土母质的基础上，进一步根据淹育、漂洗、潜育亚类前缀，即“浅”、“漂”、“青”，与母质定性词组合形成水稻土土属名称。潮土土属依据是否具有石灰性划分，通过上述灰潮土石灰性预测结果，即灰潮土与石灰性灰潮土的空间分布，完成潮土土属更新。紫色土土属更新依据其表层质地及砾质度，与亚类叠置形成命名结果。石灰岩土土属类型直接沿用其亚类结果，分为棕色石灰土与黑色石灰土 2 类。通过多源属性图层的空间叠置与土属规则约束，完成土属的更新，结合野外采样数据作为外部独立验证以修正土属更新结果（如图 8 所示）。

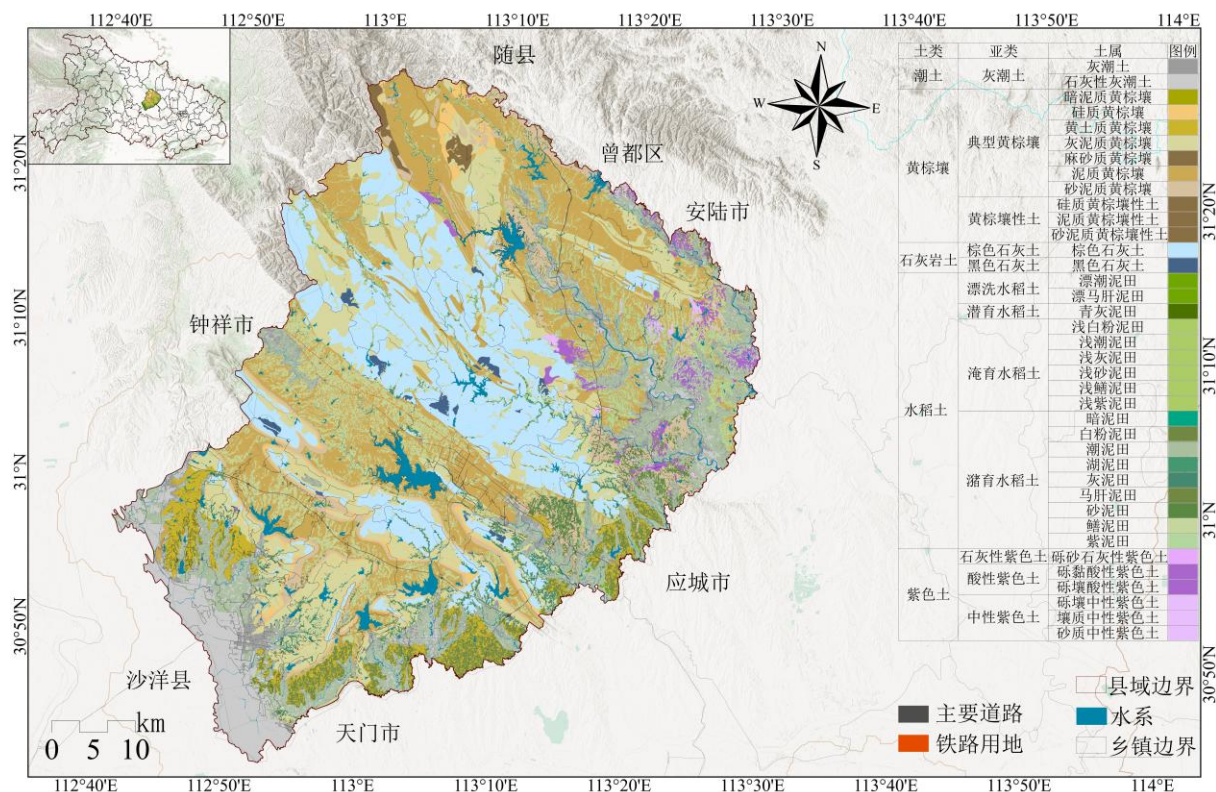


图 8 更新后各土属类型的空间分布

Fig. 8 Spatial distribution of soil groups after update

(4) 土种图更新。土种图更新以土属更新图为基础,融合土体厚度、5级质地、质地构型、腐殖质层及障碍层/特征层等关键属性数据,结合《土壤分类系统(试行版)》命名规则,构建多层嵌套的命名体系。其中,关键土壤属性数据均使用经过前期处理的数据结果,土种命名体系按土类差异构建。黄棕壤的自然土壤依据表层砾质度、腐殖质层与土体厚度组合命名,其耕作土壤依据表层质地与土体厚度;水稻土是综合障碍/特征层、表层质地与质地构型,构建“(障碍层/特征层)+表层质地+质地构型+土属”命名结构;石灰岩土的自然土壤以腐殖质层厚度与土体厚度为依据,耕作土壤结合表层质地与砾质度进行划分;潮土是根据表层质地与质地构型组合命名;紫色土的自然土壤依据腐殖质层与土体厚度分类,其耕作土壤则仅依据土体厚度分类。然后采用多图层空间叠置与土种的命名逻辑,完成研究区历史土种的系统更新。通过野外踏勘数据、土壤成土理论逻辑校验,对土种图进行修正,并通过最小上图面积(0.36 hm²)控制,合并细碎土种类型,进一步优化土种图。最终共判别土种 323 种,包括水稻土 77 种、黄棕壤 157 种、石灰岩土 45 种、潮土 15 种、紫色土 29 种(如图 9)。

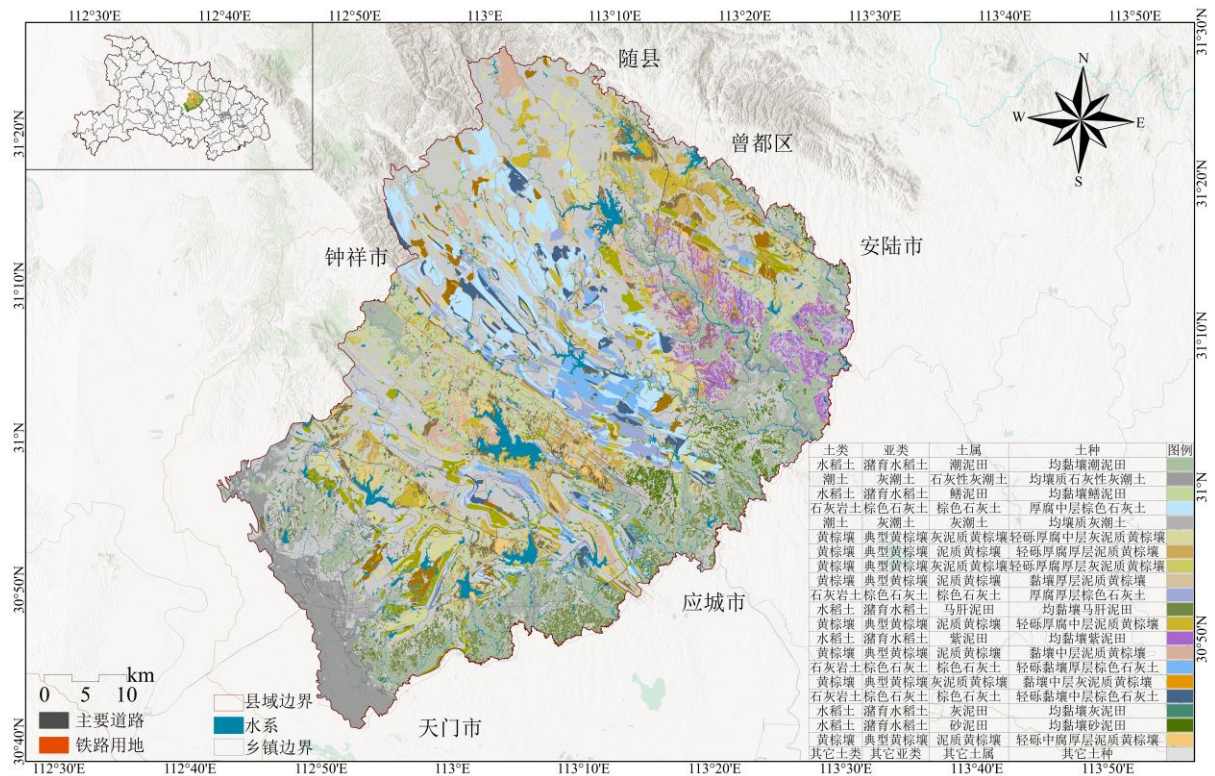


图 9 更新后各土种类型的空间分布

Fig. 9 Spatial distribution of soil species after update

2.3 历史土壤图更新结果对比与精度验证

本研究基于分类逐级更新策略,在融合历史土壤图、多源环境变量与成土逻辑的基础上,在京山市共识别出 5 个土类、12 个亚类、38 个土属和 323 种土种。分类结构相较历史图有所调整:土类数量保持不变,亚类由 13 个减少至 12 个(剔除沼泽型水稻土),土属由 43 个减少至 38 个(剔除 14 个,新增 9 个),土种由 283 种增加至 323 种,分类体系自高层单元(土类)向基层单元(土种)递进,其结构调整的幅度逐级增强。

更新后总图斑面积为 342580.77 hm²,其中在土类级别上土壤类型发生变化区域为 73898.97 hm²,占研究区总面积的 21.57%(如表 4、图 10)。黄棕壤、水稻土和潮土为主要土类。对比历史图可见,黄棕壤主要在中北部丘陵和岗地地带大幅扩展,成为分布最广的主导土类;水稻土面积则减少约 32%,

集中分布于东部和东南部灌溉区，受农田整治和退耕还林等政策影响明显；潮土扩展最为显著，面积较第二次土壤普查增长约 172%，主要归因于历史无资料区的识别补全和分类归属修正；石灰岩土主要发育于北部和中部的石灰岩地区，其分布区在空间上邻近黄棕壤；紫色土则因紫色砂页岩分布的局限性，主要集中于东北部地区。整体来看，土类空间格局的演变受气候演替、母质差异及人类活动共同驱动，揭示了区域成土过程的阶段性特征，并体现出土壤分类更新对环境变化的敏感性。

在下一层级上，亚类和土属的变化区域分别占总面积的约 32%和 36%，土种层级新增比例超过 40%，命名体系重构与属性分类标准的更新，促使图斑中土种类型发生系统性重构，结构更细化、表达更规范。

表 4 更新后土类面积变化
Table 4 Changes in land type areas after update

土类 Soil order	更新后面积 Updated area / hm ²	更新后占比 Updated percentage / %	变化量 Change / hm ²	变化率 Rate of change / %
黄棕壤 Yellow brown soil	160877.45	46.96	39599.24	31.22
水稻土 Paddy soil	78994.69	23.06	-34143.85	-32.07
石灰岩土 Limestone soil	61605.62	17.98	-8714.60	-14.23
潮土 Fluvo-aquic soil	34951.95	10.20	22437.24	171.70
紫色土 Purple soil	6151.06	1.80	257.16	4.53

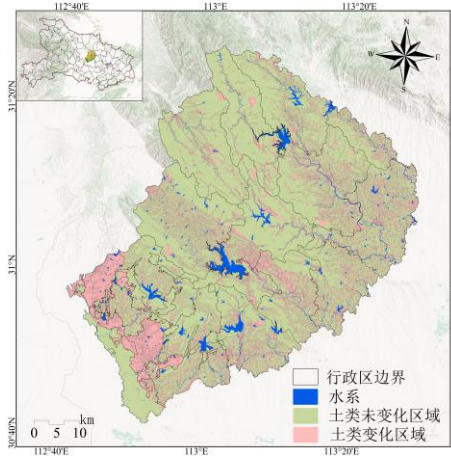


图 10 土类变化区域空间分布图

Fig. 10 Regional spatial distribution map of soil order changes

为验证模型更新结果的精度与可靠性，本研究选取京山市 30 个野外样点开展独立校核。结果表明，土类与亚类更新精度达 90%，土属与土种识别准确率分别为 83.33%和 73.33%，整体表现出良好的空间识别能力与分类一致性。依据成土理论与各级土壤命名规则，对误判图斑逐一修正：土类与亚类的 3 处误判均与近 5 年水改旱的土地利用变更有关，导致水稻土转化为潮土或黄棕壤。土属层级中的 5 处偏差主要包括：2 处“壤质”中性紫色土因深层砾石含量超过 10%，应更名为“砾壤”；2 处“灰潮土”厚度>50 cm 的土层发生石灰反应，应更名为“石灰性灰潮土”；1 处“厚层”黑色石灰土因厚度不足，应调整为“薄层”。土种层级 8 处误判主要由关键属性判定偏差导致：2 处潴育水稻土障碍层埋深误判为“深位”，实测为“浅位”；4 处腐殖质厚度分类误差，因山脊与谷地腐殖质积累差异，均被误判为“薄腐”或“厚腐”；2 处质地构型分类错误，一是“均壤质”应为“均黏壤”，二是“壤底黏”应为“均黏质”，均与地下水位高、水分淤积有关。经修正后，形成经实测优化的 4 级土壤空间分布图（图 9），显著提升了分类精度与成果图的应用可靠性。

3 讨论

构建具备分类逻辑一致性、可持续更新性和高精度表达能力的多层级土壤图是土壤资源调查与管理的重要内容^[31-32]。本研究立足京山市复杂地形与母质类型多样的特征,提出了一套“土类-亚类-土属-土种”4级联动的分类逐级更新机制,探索了土壤成土过程与机器学习推理的融合路径,为县域尺度的高精度土壤图更新提供了新范式。

本研究揭示了近40年来研究区不同类型土壤在空间格局上的系统演变。北部低山丘陵区以黄棕壤为主,具有薄层、砾质和酸性反应特征;中部碳酸盐岩区主要发育石灰岩土,pH较高,表现出碳酸盐富集;东南部冲积平原及谷地为潮土和水稻土集中分布区,受地下水位及灌排措施影响,土壤长期处于还原-氧化交替环境;东北部则以紫色母岩发育的紫色土为主。沿典型地貌梯度(山脊-谷地-冲积平原),土壤类型呈现由黄棕壤-水稻土-潮土的有序过渡,伴随质地由砂壤向黏壤转变,pH由酸性逐渐过渡至中性-碱性,腐殖质层加厚,障碍层/特征层逐渐显现。与历史土壤图相比,潜育水稻土和石灰性灰潮土分布范围显著扩大,其中淹育水稻土在长期水耕条件下,逐渐向潜育型转化,凸显人为活动对稻田土壤演化的深远影响。同时,黄棕壤在丘陵与岗地的扩展更能反映水热自然条件综合作用下的演变趋势。此外,在土属和土种层级上也表现出明显差异,黄棕壤区以“厚腐殖质中层灰泥质黄棕壤”为主导类型,而潮土区则以石灰性类型的显著扩张为主要特征。

相较于传统更新方法,本研究在方法路径和机制构建上均有所提升。首次提出了“成土规则嵌套融合”的分类逐级更新机制,通过地形湿度指数、年均水分盈亏量与母质类型等关键环境因子,增强了分类模型的逻辑性和成因合理性。同时,结合质地构型判断逻辑,在壤底黏、砂壤体黏壤等深层构型类型中,通过上下层质地差异来反映剖面发育特征与成土过程。这种模型推理路径兼顾模型数据驱动的效率与成土过程机制的解释力,为构建“结构-过程-算法”耦合的分类模型体系提供了范式参考。

在模型构建策略方面,本研究基于土壤发生学理论,通过嵌入母质约束、分层分类建模和属性耦合策略,构建了多级分类模型体系。这一方法路径与Zhong等^[33]提出的“母质优先”预测模型高度一致,即优先限定母质空间范围可显著提升分类准确率。在紫色土更新中,通过引入紫色母岩的分布掩膜,使紫色土亚类与土属的误判率显著降低,印证了母质作为“首要诊断因子”的有效性。

在既有研究中,常采用的方法是直接利用单层模型预测土种,或先预测土种再反推高级分类单元^[34-35],此类方法虽然简便,但往往忽略土类-亚类-土属-土种的层级依赖,容易导致跨层级逻辑不一致,特别是在受人为强干预的区域,表层与深层诊断特征不匹配,从而影响分类的可靠性。此外,小样本类别和过渡带区域常出现误判,空间边界破碎,泛化能力有限。相比之下,本文提出的“母质约束-分层建模-关键属性嵌套”的逐级更新框架,从高层级到低层级依次约束,显著提高了分类的逻辑一致性、边界连贯性与整体精度,更适用于复杂地貌区的历史土壤图更新。

尽管本研究在县域尺度土壤图更新方面取得了良好效果,但仍存在一定局限。首先,在少样本土种(如潜育水稻土、石灰性紫色土)识别中,即使采用了均衡采样策略,数据稀疏区域仍存在误判,说明当前样本增强方式在复杂异质环境下尚显不足。其次,模型输入仍依赖静态土地利用、气候和地貌数据,未能充分考虑土地覆被的动态变化,因而在快速城市化背景下的泛化能力受限,与Hengl等^[32]指出的“静态模型对时序变化不敏感”的问题相一致。此外,模型精度亦受到特征选择与数据质量的制约,当土壤类型多样且受人为强干扰时,预测不确定性更为突出。其中,不确定性主要源于少数类样本的不足、环境因子的时空变异以及不同地理单元间的空间异质性,使环境变量与土壤类型之间的关系更为复杂和非线性,从而增加了模型预测的偏差与不稳定性。最后,本研究的制图结果尚未考虑土壤类型之间的合并以及制图综合等技术问题,未来需结合实际需求进一步完善。

未来研究可从以下几方面展开：一是结合土壤空间分布的一致性及成土环境的地学约束条件，生成更具代表性的少数类样本，从而缓解类型不平衡问题，提升模型在小面积土壤类型上的识别能力^[36-37]。二是构建具备时间动态感知能力的土壤分类体系，从而提升模型对土地利用变化与季节性干扰的响应能力，更好地反映土壤属性的时空演化过程。三是构建“过程驱动+数据驱动”的混合建模框架，通过模拟生成土壤属性的动态演化序列作为训练输入，提升模型的物理一致性与过程可解释性^[38]。值得强调的是，更新后的土壤图不仅优化了分类体系和制图方法，还在应用层面为耕地质量评价、作物适宜性区划及碳储量估算等提供了更可靠的数据支撑，凸显其在粮食安全、精准农业与“双碳”战略中的重要价值。此外，未来可依托第三次全国土壤普查成果，推动省域乃至全国尺度的土壤图智能化更新，为耕地保护与土地整治提供更加全面、准确的基础数据支撑。

4 结 论

本研究以湖北省京山市为研究区域，构建融合土壤成土理论与机器学习算法的历史土壤图更新策略，实现对历史土壤图中土类、亚类、土属与土种的分类逐级更新。结果表明，土类演变显著受地貌格局与土地利用变迁的共同驱动，黄棕壤在自然风化与人类扰动的交互作用下明显扩展，水稻土则因耕作技术优化、农田整治与退耕还林等措施而面积大幅缩减。在亚类更新中，分别构建了 RF、XGBoost 和 LightGBM 模型，结合母质约束与分层分类建模策略，有效提升了分类精度，最高 OA 达 0.92，Kappa 系数为 0.91，验证了将成土理论嵌入数据驱动模型的可行性与有效性。土属更新以灰潮土的石灰性预测、质地等级转换和砾质度重分类为基础，识别出 38 种土属类型。在土种更新阶段，进一步整合土体厚度、5 级质地、质地构型、腐殖质层和障碍层/特征层等多源信息，建立多层级逻辑约束机制。最终共识别出 323 种土种，全面揭示了研究区土壤类型的空间异质性和多样性。研究表明，将成土机理与数据驱动方法有机融合，可显著提升土壤图更新的系统性、科学性与分辨率，为高精度土壤制图、耕地资源管理与农业空间优化布局提供坚实支撑与技术路径。

参考文献 (References)

- [1] Jouquet P, Harit A, Hervé V, et al. The impact of termites on soil sheeting properties is better explained by environmental factors than by their feeding and building strategies[J]. *Geoderma*, 2022, 412: 115706.
- [2] Wu S X, Zhang W L, Xu A G, et al. Development of model for integrating and mapping of soil nutrient maps with different grading systems[J]. *China Agricultural Informatics*, 2019, 31(5): 110-120. [武淑霞, 张维理, 徐爱国, 等. 不同分级标准下土壤养分图的整合模型构建[J]. *中国农业信息*, 2019, 31(5): 110-120.]
- [3] Gong Z T, Zhang G L, Chen H Z, et al. Responses and feedbacks of soils to enviromental change[J]. *Quaternary Sciences*, 2004, 24: 140-148.[龚子同, 张甘霖, 陈鸿昭, 等. 环境变化中土壤的响应与反馈[J]. *第四纪研究*, 2004, 24: 140-148.]
- [4] Ma Y X, Minasny B, Malone B P, et al. Pedology and digital soil mapping (DSM)[J]. *European Journal of Soil Science*, 2019, 70(2): 216-235.
- [5] IUSS and FAO. World Reference Base for Soil Resources. International soil classification system for naming soils and creating legends for soil maps[M]. 4th ed. Vienna: International Union of Soil Sciences (IUSS), 2023: 1-17.
- [6] Bouma J, Bonfante A, Basile A, et al. How can pedology and soil classification contribute towards sustainable development as a data source and information carrier?[J]. *Geoderma*, 2022, 424: 115988.
- [7] Zhu A X, Yang L, Fan N Q, et al. The review and outlook of digital soil mapping[J]. *Progress in Geography*, 2018, 37(1): 66-78.[朱阿兴, 杨琳, 樊乃卿, 等. 数字土壤制图研究综述与展望[J]. *地理科学进展*, 2018, 37(1): 66-78.]
- [8] Wadoux A M J, Molnar C. Beyond prediction: Methods for interpreting complex models of soil variation[J]. *Geoderma*, 2022, 422: 115953.

- [9] Hou K, Qian H, Zhang Y T, et al. Evolutionary process of soil grain-size fractals in irrigation areas of northern China: A pedological analysis of ancient irrigation wisdom[J]. Catena, 2025, 257: 109160.
- [10] Wadoux A M J C, Minasny B, McBratney A B. Machine learning for digital soil mapping: Applications, challenges and suggested solutions[J]. Earth-Science Reviews, 2020, 210: 103359.
- [11] Xu M P, Gao D X, Fu S Y, et al. Long-term effects of vegetation and soil on the microbial communities following afforestation of farmland with *Robinia pseudoacacia* plantations[J]. Geoderma, 2020, 367: 114263.
- [12] De Caires S A, St Martin C, Atwell M A, et al. Advancing soil mapping and management using geostatistics and integrated machine learning and remote sensing techniques: A synoptic review[J]. Discover Soil, 2025, 2(1): 53.
- [13] Heung B, Ho H C, Zhang J, et al. An overview and comparison of machine-learning techniques for classification purposes in digital soil mapping[J]. Geoderma, 2016, 265: 62-77.
- [14] Maino A, Alberi M, Anceschi E, et al. Airborne radiometric surveys and machine learning algorithms for revealing soil texture[J]. Remote Sensing, 2022, 14(15): 3814.
- [15] Cheng L W, Yan M Z, Zhang W H, et al. Interpretable digital soil organic matter mapping based on geographical Gaussian process-generalized additive model (GGP-GAM)[J]. Agriculture, 2024, 14(9): 1578.
- [16] Rohmer J, Belbeze S, Guyonnet D. Insights into the prediction uncertainty of machine-learning-based digital soil mapping through a local attribution approach[J]. Soil, 2024, 10(2): 679-697.
- [17] Gayen A, Mafizul Haque S. Soil erodibility assessment of laterite dominant sub-basin watersheds in the humid tropical region of India[J]. Catena, 2022, 213: 106161.
- [18] Tian Y C, Zheng D L, Zhang Q, et al. Inversion of coastal soil salinity in Qinzhou Bay based on domestic resources-1 02D satellite and machine learning algorithm[J]. China Environmental Science, 2024, 44(1): 371-385. [田义超, 郑丹琳, 张强, 等. 基于国产资源一号02D卫星和机器学习算法的钦州湾滨海土壤盐分反演[J]. 中国环境科学, 2024, 44(1): 371-385.]
- [19] Aydın Y, Işıkdag Ü, Bekdaş G, et al. Use of machine learning techniques in soil classification[J]. Sustainability, 2023, 15(3): 2374.
- [20] Liu C Y, Ku C Y, Wu T Y, et al. An advanced soil classification method employing the random forest technique in machine learning[J]. Applied Sciences, 2024, 14(16): 7202.
- [21] M S, Lee S I. A unified approach to interpreting model predictions[C]//Neural Information Processing Systems., 2017.
- [22] Zhang L, Chu Q, Zhou J W, et al. Soil phosphorus availability determines the preference for direct or mycorrhizal phosphorus uptake pathway in maize[J]. Geoderma, 2021, 403: 115261.
- [23] Mello F A O, Bellinaso H, Mello D C, et al. Soil parent material prediction through satellite multispectral analysis on a regional scale at the Western Paulista Plateau, Brazil[J]. Geoderma Regional, 2021, 26: e00412.
- [24] Zhu C D, Wei Y C, Zhu F B, et al. Digital mapping of soil organic carbon based on machine learning and regression Kriging[J]. Sensors, 2022, 22(22): 8997.
- [25] Owens P R, Mancini M, Winzeler E H, et al. Simulating water dynamics related to pedogenesis across space and time: Implications for four-dimensional digital soil mapping[J]. Geoderma, 2024, 447: 116911.
- [26] Qi J T, Cheng P T, Gao F F, et al. Research progress and prospect of rapid acquisition technology of physical and chemical parameters of farmland soil[J]. Smart Agriculture, 2024, 6(3): 17-33. [齐江涛, 程盼婷, 高芳芳, 等. 农田土壤理化参数快速获取技术研究进展与展望[J]. 智慧农业, 2024, 6(3): 17-33.]
- [27] Soil Survey Office of Jingshan County. Soil Chronicles of Jingshan County[M]. Jingshan, Hubei: Soil Survey Office of Jingshan County, 1984. [京山县土壤普查办公室. 京山县土壤志. 湖北京山: 京山县土壤普查办公室, 1984.]
- [28] Bo Q, Lü P H, Wang Z G, et al. Predication of the post mining land use based on random forest and DBSCAN[J]. PLoS One, 2024, 19(1): e0287079.
- [29] 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.
- [30] Zhang H, Si S, Hsieh C. GPU-acceleration for large-scale Tree Boosting[J]. arXiv preprint arXiv:1706.08359. 2017, <https://arxiv.org/abs/1706.08359>.
- [31] Minasny B, McBratney A B. Digital soil mapping: A brief history and some lessons[J]. Geoderma, 2016, 264: 301-311.
- [32] Hengl T, Mendes de Jesus J, Heuvelink G B M, et al. SoilGrids250m: Global gridded soil information based on machine learning[J]. PLoS One, 2017, 12(2): e0169748.
- [33] Zhong S Q, Han Z, Du J, et al. Relationships between the lithology of purple rocks and the pedogenesis of purple soils in the Sichuan Basin, China[J]. Scientific Reports, 2019, 9: 13272.
- [34] Liang Z P, Zou T X, Gong J L, et al. Evaluation of soil nutrient status based on LightGBM model: An example of tobacco planting soil in Debao County, Guangxi[J]. Applied Sciences, 2022, 12(23): 12354.
- [35] Liu C. Study on soil type prediction model and digital mapping in hilly area[D]. Taian, Shandong: Shandong Agricultural University, 2024. [刘成.山丘区土壤类型预测模型与数字制图研究[D]. 山东泰安: 山东农业大学, 2024.]
- [36] Taghizadeh-Mehrjardi R, Schmidt K, Eftekhari K, et al. Synthetic resampling strategies and machine learning for digital soil mapping in Iran[J]. European Journal of Soil Science, 2020, 71(3): 352-368.
- [37] Wang L P, Wang X, Kooch Y, et al. Improvement of data imbalance for digital soil class mapping in Eastern China[J]. Computers and Electronics in Agriculture, 2023, 214: 108322.
- [38] Vereecken H, Herbst M, Nemes A, et al. Modeling soil structure and function[J]. Annual Review of Earth and Planetary Sciences, 2016, 44: 73-99.

(责任编辑: 檀满枝)