

文慧, 杨晨, 王天巍, 李朝霞. 历史数据和专家知识融合的土壤矿物学类型制图研究 [J]. 土壤学报, 2026,
WEN Hui, YANG Chen, WANG Tianwei, LI Zhaoxia. Mapping Research of Soil Mineralogical Classes Integrating Expert
Knowledge and Legacy Data[J]. Acta Pedologica Sinica, 2026,

历史数据和专家知识融合的土壤矿物学类型制图研究*

文 慧, 杨 晨, 王天巍[†], 李朝霞

(华中农业大学资源与环境学院, 武汉 430070)

摘 要: 土壤矿物学类型的空间分布特征反映了区域成土环境与发生过程, 对于理解土壤发生发育具有重要意义。以亚热带气候区湖北省为研究对象, 整合历史土壤矿物学数据, 提出基于专家知识决策树的概率制图框架, 并结合 3 种不同的模型对土壤矿物学类型的空间概率分布及环境驱动进行研究。结果表明: CatBoost 模型具有良好的适用性, 在所有模型中表现最优。研究区土壤矿物学类型以混合型、伊利石混合型和硅质混合型为主, 蒙脱石型、高岭石型与蛭石型呈零星斑块状分布。验证结果表明, 不同矿物学类型的验证概率存在明显差异, 其中硅质混合型、蒙脱石混合型等混合矿物类型具有较高的预测可靠性, 而部分单一矿物类型的预测不确定性相对较高。不同类型预测效果的差异与其形成过程及环境控制因素有关, 水热条件通过影响风化淋溶过程参与调节土壤矿物组成演变, 母质则对原生矿物类型的形成与空间分布具有重要影响。研究成果有助于深入理解亚热带过渡区土壤发生演变过程, 并可为第三次土壤普查系统分类土族命名提供参考依据。

关键词: 土壤系统分类; 土族; 数字土壤制图; 专家知识

中图分类号: S155.3

文献标志码: A

Mapping Research of Soil Mineralogical Classes Integrating Expert Knowledge and Legacy Data

WEN Hui, YANG Chen, WANG Tianwei[†], LI Zhaoxia

(College of Resource and Environment, Huazhong Agricultural University, Wuhan 430070, China)

Abstract: 【Objective】 The spatial distribution of mineralogical classes in mineral soils reflects the regional pedogenic environment and soil-forming processes, and this information is essential for understanding soil genesis and development. **【Method】** This study focuses on Hubei Province, a region with a subtropical climate in China. By integrating legacy soil mineralogical classes data, a fuzzy mapping framework was proposed based on an expert knowledge decision tree. The study further investigated the spatial probability distribution and driving factors of soil mineralogical classes using three different models. **【Result】** The results showed that the CatBoost model demonstrated the best performance among the three models. The soil mineralogical classes in the study area were

* 国家自然科学基金项目 (42471059, 42377328) 资助 Supported by the National Natural Science Foundation of China (Nos. 42471059, 42377328)

[†] 通讯作者 Corresponding author, E-mail: wangtianwei@webmail.hzau.edu.cn

作者简介: 文 慧 (1995—), 女, 博士研究生, 主要研究领域为土壤地理。E-mail: 734951507@qq.com

收稿日期: 2026-03-01; 收到修改稿日期: 2026-06-29; 网络首发日期 (www.cnki.net): 2026- -

predominantly Mixed, Illite mixed, and Siliceous mixed, while Smectite, Kaolinite, and Vermiculite classes were distributed in scattered patches. Validation results revealed pronounced differences in mapping accuracy among mineralogical classes. Mixed mineral assemblages, including Siliceous mixed and Smectite mixed, were predicted with relatively high reliability, whereas several single-mineral types showed greater uncertainty. These class-dependent differences may reflect variation in mineral formation pathways and environmental controls. Moisture and thermal regimes may influence soil mineralogical evolution by regulating weathering and leaching, whereas parent material exerts an important control on the occurrence and spatial distribution of primary minerals.

【Conclusion】 These research findings contribute to a deeper understanding of soil formation and evolution processes in the subtropical transition zone and provide a reference basis for the systematic classification and naming of soil families in the 3rd National Soil Census.

Key words: Soil taxonomy; Soil family; Digital soil mapping; Expert knowledge

土壤矿物学类型是指控制层段内特定颗粒大小组分的土壤矿物组成,按成因可分为原生矿物与次生黏土矿物。原生矿物与母岩关系密切,其种类及分布在较大程度上受母岩矿物组成制约。相比之下,次生黏土矿物形成于风化和成土过程中,主要富集在粒径小于 $2\ \mu\text{m}$ 的土壤颗粒中,对环境变化响应敏感,其结构和类型的变化能够反映土壤演化过程及环境变迁特征,是判定土壤矿物学类型的重要依据^[1-3]。

黏土矿物的形成与分布受母质、气候、地形、生物及水文条件等多种因素共同影响,空间分异规律复杂。这种环境控制在垂直尺度上有所体现,研究表明,随着海拔升高,黏粒矿物组成有从 1:1 型为主向 2:1 型矿物为主过渡的趋势^[4-5]。然而,该序列仅反映一般规律,具体演变仍受气候、母质等条件制约^[6-7]。例如,华中地区部分黄褐土和黄棕壤主要黏土矿物为 1.4 nm 过渡矿物、水云母和高岭石,而华北地区的部分黄褐土和黄棕壤则以伊利石为主,其次为高岭石^[3,8]。华中以南区域,总体呈现随着雨量的增高,高岭石的含量呈显著增加的趋势,但在富含云母类矿物的母质发育的土壤中,伊利石却占据主要地位^[9]。因此,探究土壤矿物学类型的空间分布特性,对目前开展的土壤普查工作具有重要意义。

我国目前正在开展第三次土壤普查,土壤命名与类型判定主要采用土壤发生分类与土壤系统分类两套框架同步推进^[10]。其中,土族作为土壤系统分类的基层单元,可为土地利用和评价服务提供重要依据^[11-13]。土壤矿物学类型是土族不可或缺的划分指标,反映了该类土壤的发生发育过程,有较强的区域性。然而,已有矿物学类型数据覆盖范围有限,且缺少颗粒大小级别分布信息,难以满足当前土壤系统分类定名的需求。传统的实验室测定方法虽然精确,但成本高、周期长^[1,14-15],一定程度上限制了区域尺度土壤矿物学组成的空间信息获取。因此,如何在数据不足的情况下,表征区域尺度上土壤矿物学的类型,是亟待解决的问题。

随着高精度环境变量数据的日益丰富与技术算法的迭代,数字土壤制图方法被广泛应用于土壤属性空间预测。其核心在于构建合适的预测模型,常用预测模型包括分类和回归树、支持向量机、深度学习、极致梯度提升树和随机森林等^[16-17]。其中,随机森林模型(RF)能够捕捉和表达土壤属性空间变化与环境变量之间复杂的非线性关系,并基于这种关系推测土壤属性的空间分布^[18]。优化的梯度提升算法(CatBoost)在处理非线性关系、复杂特征组合及类别差异方面具有更强优势。而支持向量机(SVM)则可以解决小样本情况下的机器学习问题。

鉴于此,以湖北省为研究对象,耦合多源环境因子及样点数据,旨在实现以下目标:(1)构建融合专家知识与历史数据的土壤矿物学类型先验概率框架;(2)比较不同模型在概率制图中的适用性,筛选最优模型;(3)量化表征土壤矿物学类型的空间概率分布。研究结果旨在为区域土壤系统分类提供关键的诊断层定量依据,并为农业、环境等领域的管理提供数据参考。

1 材料与方法

1.1 研究区概况

湖北省(29°01'53"—33°6'47"N, 108°21'42"—116°07'50"E)地处我国中部,坐落于长江中上游地区。具有典型的盆地式地貌格局,其东、西、北三面被武陵山、大巴山、桐柏山等山系环绕,中部为地势低平、河网密布的江汉平原,整体地势向南敞开。全省地貌类型多样,以山地(56%)和丘陵(24%)为主,平原湖区(20%)集中于中南部。土地利用类型以林地和耕地为主。据湖北第三次国土调查主要数据成果,林地面积占主要土地利用类型总面积的63.19%,耕地面积占32.47%,园地面积占3.32%。区域属于亚热带季风气候,多年平均气温16.7~18℃,年均降雨量1200~1400 mm。受复杂的地形、气候和成土母质等因素的综合影响,湖北省土壤类型繁多,空间分异显著。根据第二次土壤普查的结果,主要土壤类型占比约为黄棕壤面积占37.93%,水稻土面积占24.97%,石灰岩土面积占9.86%,潮土面积占8.53%等。

1.2 专家知识和历史数据获取

基于系统文献计量检索与数据挖掘相结合的方法获取专家知识和历史数据。首先通过中国知网、Web of Science等学术资源平台,围绕湖北省、土壤类型、黏土矿物等关键词构建检索式,进行系统性筛选。随后对入选文献中的原始监测数据、附表数据及补充材料进行逐条提取,并建立统一的数据录入模板。在数据整理过程中,对不同年代、不同测定标准的数据进行指标统一、单位换算和方法校正,剔除重复样点及异常值,确保数据的可比性与一致性。图1为历史数据点位分布信息、发表年份、发表数量及数据来源统计图谱结果。

历史数据共来源于37篇公开发表的湖北省相关学术成果及系统性土壤调查资料,发表年份跨度为1987-2018年,以反映第二次全国土壤普查成果形成和应用之后、第三次全国土壤普查启动之前湖北省土壤研究资料的历史积累。研究团队主要来自湖北省农业科学院、华中农业大学、中国科学院南京土壤研究所等科研机构及高校。



图1 土壤黏土矿物专家知识数据统计图

Fig. 1 Statistical atlas of expert knowledge data on soil clay mineralogy

依据历史研究资料中已有的矿物学类型记录、判定规则及区域土壤资料,对研究区在现有资料精度和制图分辨率条件下可识别的主要土壤矿物学类型进行归纳、统一命名和编号。共整理形成 12 种类型,分别为云母混合型(Micaceous mixed, Mic-M)、硅质型(Siliceous, Sil)、硅质混合型(Siliceous mixed, Sil-M)、高岭石型(Kaolinite, Kao)、高岭石混合型(Kaolinite mixed, Kao-M)、伊利石型(Illite, Ill)、伊利石混合型(Illite mixed, Ill-M)、蛭石型(Vermiculite, Ver)、蛭石混合型(Vermiculite mixed, Ver-M)、蒙脱石型(Smectite, Sme)、蒙脱石混合型(Smectite mixed, Sme-M)和混合型(Mixed)。其中,混合型包括 0.02~2 mm 和 ≤ 0.002 mm 两个颗粒大小级别的土壤矿物学类型。长石型、长石混合型和云母型等虽在部分资料中有记录,但在湖北省多呈零散、局部块状分布,空间连续性较弱,未将其纳入后续预测。

1.3 先验概率制图框架

收集历史数据中所有土壤矿物学类型及其对应的环境因子信息作为数据集。以土壤矿物学类型为目标变量,将所有数据信息转化为层级化决策规则,应用 1-SE 准则选择最优树模型,构建反映环境因子与矿物类型逻辑关系的决策树模型^[19]。最后通过检索虚点对应的特定土壤亚类及其对应环境组合下的矿物类型,量化该土壤类型中不同矿物的分布概率,将其定义为模型的先验概率,结果如图 2 所示。同时,为防止模型过拟合,应用十折交叉验证与复杂性参数的智能修剪,对重建后的树模型进行优化,得到最后的概率范围。

以土壤亚类空间分布为基础,根据各亚类面积占比采用自然断点法进行分层随机抽样,并将各亚类样点数量按 10 的倍数进行取整,最终生成 9 450 个虚拟点位(图 3)。在此基础上,采用上述决策树判定不同点位土壤亚类对应土壤矿物学类型的概率。若决策树中判定某亚类中该矿物类型概率为 $0 < P < 1$,则基于退火算法和环境因子进行聚类,从而将其样点分配为 0 或 1^[20-21]。在上述概率的基础上,采用重复 N ($N=100$) 次的方法计算概率。针对研究区每个栅格单元,统计虚拟采样点在 N 次随机模拟中某类矿物的出现频次 S ,据此计算该矿物在对应位置的出现概率 $P(P=S/N \times 100\%)$ 。以此虚拟样点概率数据集为训练样本,利用 RF、CatBoost 和 SVM 进行空间建模比较,选取最优结果生成土壤矿物学类型的空间分布概率。最后,采用逐像元最大值合成法,识别每个栅格单元中概率最高的矿物类别,从而得到研究区土壤矿物学类型最大概率空间分布图。

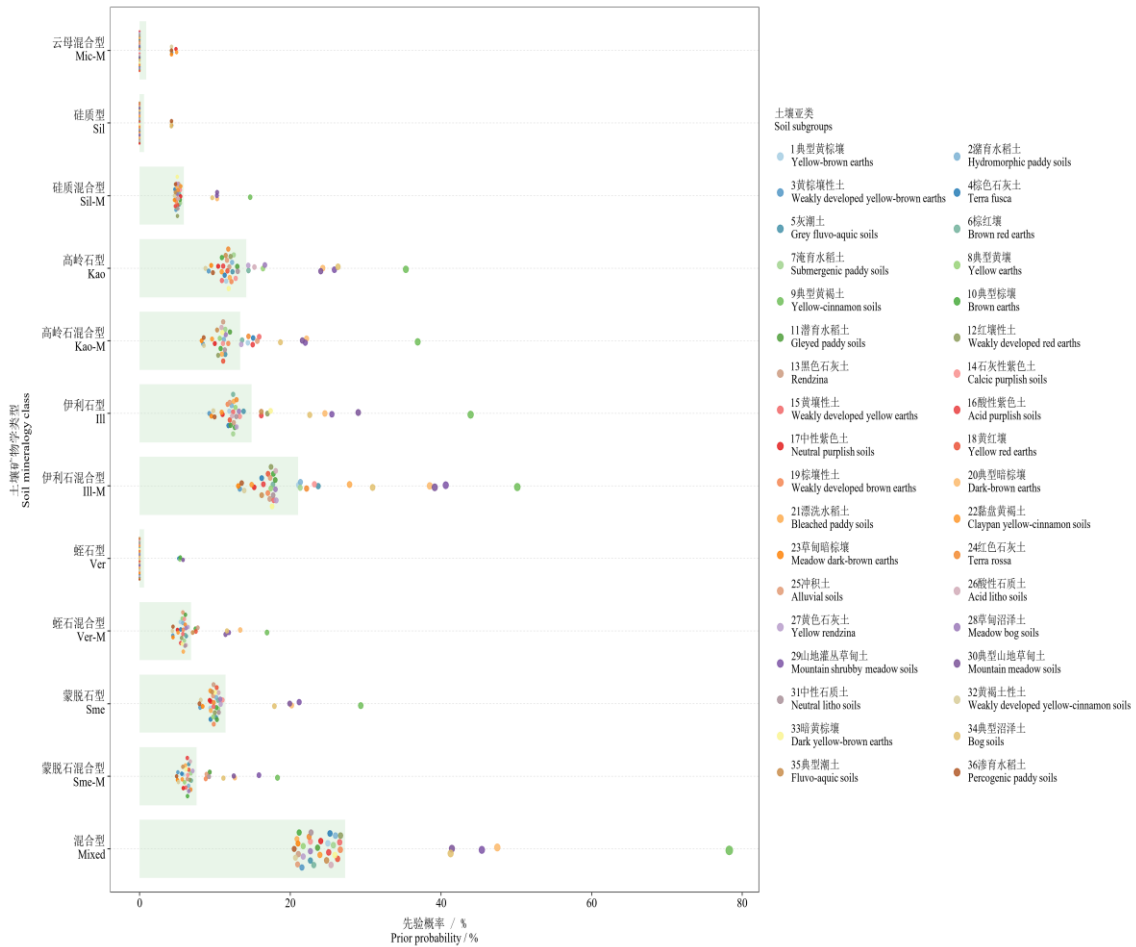


图 2 土壤矿物学类型专家知识先验概率分布图

Fig. 2 Prior probability distribution map of soil mineralogical classes based on expert knowledge

1.4 环境变量数据来源及处理

1.4.1 数据来源 选取与土壤矿物学类型形成和空间分异密切相关的环境变量数据,主要包括气候、植被、土地利用和土壤属性等 17 个类型因子。土壤类型与黏土矿物息息相关,采用第二次全国土壤普查的矢量化土壤类型图亚类层级。亚类级别既能够反映不同土壤之间成土过程的差异,又与现有资料支撑条件较为匹配,适用于省级尺度的矿物学类型预测。土壤母质是原生矿物风化的基础,从 1:50 万湖北省地质图中数字化提取,并转换为栅格数据。气候数据来源于中国气象数据网获取 2018—2023 年的气象站点数据,通过空间插值生成了年均气温、年均降雨和水分盈余指数 3 个栅格图层,用以表征区域水热条件差异。生物因素由植被和人类活动共同表征,植被因子能够通过调节有机质输入、根系活动、养分循环及土壤水热条件,影响土壤风化过程和矿物转化特征。归一化植被指数通过 Sentinel-2 影像产品 2023 年 12 月至 2024 年 3 月第四季度 (NDVI_4) 的 NDVI 平均值处理得到,第四季度数据捕捉了植被从生长末期到休眠期的衰退过程,能够反映土壤在植物生命周期末尾阶段对其持续供养能力。人类活动的影响则通过武汉大学黄昕教授团队制作的中国年度土地覆盖数据集 (annual China Land Cover Dataset) 2023 年土地利用数据来量化^[23]。人类活动通过改变土地利用方式、耕作措施和蓄排条件,影响土壤水分状况及物质迁移过程,从而影响土壤矿物风化过程。同时,为进一步表征水文条件对土壤发育的影响,基于土地利用数据计算各栅格单元至最近河流的欧氏距离。基于航天飞机雷达地形任务 (SRTM) 30 m 数字高程模型,提取研究区高程和地形湿度指数 2 个地形因子。为反映物质组成、离子交换特征和风化程度等方

面对土壤矿物的形成与转化环境的影响,选取的土壤属性变量包括 pH、有机质含量、阳离子交换量、土壤质地、氧化铁含量和土内砾石丰度 7 个因子。其中,土壤质地由黏粒、粉粒和砂粒含量表征。交换性离子包括交换性钙、铝、镁、钠和钾。相关数据数据来源于中国科学院南京土壤研究所制作的中国高分辨率国家土壤信息格网基本属性数据集^[24]。

1.4.2 数据处理 在土壤系统分类中,矿物学特征的判定通常依据标准化的“控制层段”^[2]。然而,由于土壤类型多样,在区域尺度上不同土体的控制层段在深度和厚度上存在显著异质性,难以保证采样层位的绝对一致^[24-25]。尤其是在跨类型模型训练中,层位差异可能导致系统误差。为确保本研究中所有样本在发生学意义上的均一性和各属性数据间的可比性,统一选取成土过程最活跃且性质最稳定的淀积层(B 层)作为研究对象。B 层能较好地反映黏土矿物的发生演变特征及环境响应过程。在土层深度处理上,基于等面积二次样条函数来拟合连续的深度函数,反映垂直方向上土壤属性的变化。最终,从拟合曲线中提取深度层(25~50 cm)的平均值,用作后续相关属性及土壤矿物学类型预测的标准化样本数据。

所有空间数据的坐标系统均统一为 CGCS2000 GK Zone 19,空间分辨率重采样为 30 m。在对数据进行掩膜提取与检查后,对缺失值进行填充,连续变量采用局部均值(3×3)与全局均值相结合的方式填充。分类型变量则采用众数进行填充,并且对所有分类型变量进行二进制格式编码。对于包含 K 个类别的变量,采用 $\lceil \log_2(K) \rceil$ 个二进制位表示(K=该列类别数)。

1.5 精度评价

环境变量的重要性是基于平均不纯度减少方法计算得到的。变量的不纯度减少量总和越大,表明该变量对降低模型预测误差的贡献越大,其重要性越高。

$$i(t) = \frac{1}{N_t} \sum_{k \in D_t} (y_k - \bar{y}_t)^2 \quad (1)$$

式中,节点 t 的不纯度 $i(t)$ 定义为该节点内样本标签值的均方误差, N_t 表示节点 t 中的样本总数, D_t 表示节点 t 中的样本集合, y_k 表示第 k 个样本的真实值。

验证样点数据涵盖了从文献数据中挖掘的 144 个土壤矿物学类型以及 164 个野外实测数据,共 308 个实测样点数据(图 3),采用基于概率的制图精度指标(Mapping accuracy, MA)作为衡量预测结果和实际样点相符合的程度,数值越大表示制图效果越好,越贴近实际。MA 公式如下:

$$MA = \frac{1}{N} \sum_{i=1}^N P_i I[M_i = A_i] \times 100\% \quad (2)$$

式中, N 表示样点数, i 表示调查样点序号, P_i 表示第 i 个点获取的黏土矿物类型概率。 $I[M_i = A_i]$ 是一个内指示函数,当信息类型或最大概率黏土矿物类型 M_i 与实际调查 A_i 相等时 $I=1$,否则 $I=0$ 。

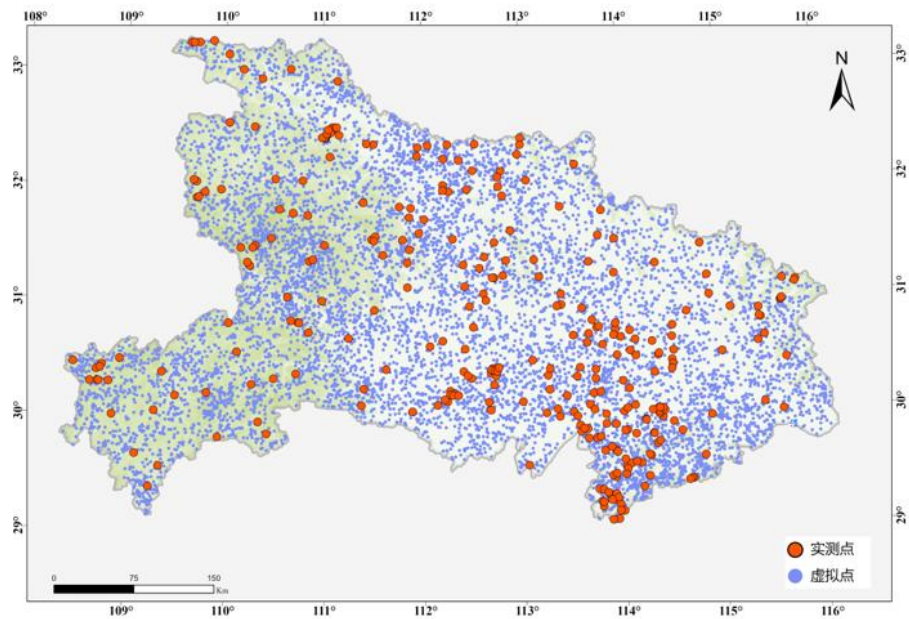


图3 虚拟点位和实测点空间分布
Fig. 3 Spatial distribution of virtual and validation points

2 结果与讨论

2.1 不同模型在土壤矿物学类型中的预测精度

基于不同模型对土壤矿物学类型的分类结果比较可知，CatBoost 模型整体表现最优，在多数矿物学类型上具有更高的拟合度和稳定性，因此选取其概率结果进行最大像元合成。从表 1 可以看出，CatBoost 模型在所有矿物类型上的拟合度均高于 RF 和 SVM，表现出更好的稳定性。其中，CatBoost 在蛭石混合型、蒙脱石混合型、高岭石型、伊利石混合型、伊利石型等类型上的拟合度较高，均大于 0.65。说明 CatBoost 对大部分类型的训练样本具有较好的拟合能力，但其空间预测可靠性仍需结合独立验证点进一步判断。蛭石混合型的拟合度最高，表明该类型能更好的受环境变量响应。相比之下，云母混合型在所有模型中最低，拟合度为 0.02，显著低于其他矿物类型。表明这类矿物类型可能存在样本数量不足，边界模糊或空间分布零散等问题，导致模型难以识别。基于此，认为云母混合型的预测结果难以准确反映其在研究区内的空间分布关系，故在后续制图分析中将其剔除，以避免对结果解释造成干扰。

表 1 不同模型在土壤矿物学类型分类中的拟合度

Table 1 Evaluation metrics of different models in soil mineralogical classes (R^2)

土壤黏土矿物类型 Soil mineralogy classes	随机森林 RF	优化的梯度提升算法 CatBoost	支持向量机 SVM
硅质混合型 Sil-M	0.52	0.60	0.51
硅质型 Sil	0.46	0.54	0.44
混合型 Mixed	0.58	0.61	0.53
蒙脱石混合型 Sme-M	0.68	0.71	0.70
蒙脱石型 Sme	0.61	0.67	0.59
高岭石混合型 Kao-M	0.63	0.67	0.59

高岭石型 Kao	0.52	0.68	0.66
伊利石混合型 Ill-M	0.62	0.67	0.64
伊利石型 Ill	0.63	0.67	0.59
云母混合型 Mic-M	0.01	0.02	0.02
蛭石混合型 Ver-M	0.67	0.72	0.65
蛭石型 Ver	0.47	0.53	0.47

根据最优模型的概率图结果显示（图 4），多数矿物学类型在研究区范围内表现为局部高概率分布。不同类型在区域上存在明显的过渡，尤其在鄂中丘陵到平原的过渡区域、鄂东低山丘陵以及鄂西山地。平均概率排序前 5 位的土壤矿物学类型依次为：Mixed>Ill-M>Kao-M>Ill>Kao，表明混合型、伊利石混合型及高岭石混合型在研究区内占比较高。其中，混合型的平均概率最高，高概率区分布范围最广，在鄂西南、鄂东以及中南部部分区域，呈大面积带状连续分布，而在鄂中北部整体概率相对较低。伊利石混合型高概率区集中在鄂东北、鄂西南部分地区，在江汉平原整体偏低，表现出东西分异格局。高岭石混合型与伊利石混合型分布相似，但整体概率偏低。伊利石型的高概率区域主要集中在鄂东和鄂西南地区，并在中北部部分丘陵区形成局部聚集。高岭石型则主要分布于鄂北局部区域，高值区范围小于混合型 and 伊利石混合型。蒙脱石型、蒙脱石混合型和蛭石混合型分布范围有限，在鄂北形成了局部高概率区，在鄂东和鄂西南边缘零星分布。这表明这几类矿物类型受局部环境影响强烈。硅质混合型在鄂东北地区 and 鄂东南局部区域存在明显高概率区。其他类型在研究区内总体概率较低，仅在局部地区出现零星高值斑块，并非优势矿物学类型，区域代表性弱。

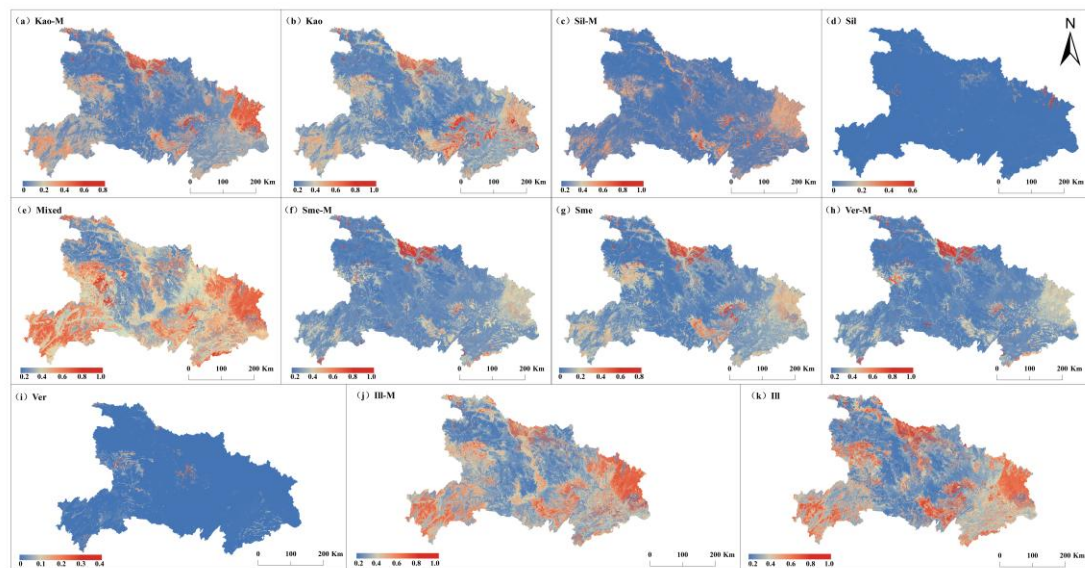


图 4 土壤矿物学类型概率结果图

Fig. 4 Probability result diagram of soil mineralogical classes

2.2 土壤矿物学类型的空间分布

土壤矿物学类型发生发育过程是多因素相互作用的复杂性结果。基于模型中平均不纯度减少的方法，来量化土壤矿物学类型环境变量的相对重要性。如图 5 所示，研究筛选了对上述 11 种土壤矿物学类型预测贡献度最高的关键变量。水热条件、土壤类型被识别为大多数类型组别的主导驱动因子，重要性值较高。对于高岭石组（Kao、Kao-M），降雨量是关键驱动因子，这与土壤发生学特性一致，高岭石主要富集在经历强烈淋溶和酸性环境的区域，水热淋溶强度在驱动原生层状硅酸盐向过渡性风化产物中具有重要作用。正如黄巧云等^[26]研究指出，区域水热条件通过控制化学风化强度，主导了原生矿物向次生黏土矿物的定向转化

与演替。

蒙脱石组 (Sme、Sme-M) 与蛭石组 (Ver、Ver-M) 主要受地形和降雨的影响。相关研究表明, 这两类 2:1 型矿物的保存主要与淋溶作用相对较弱、排水受限及局部母质具备较丰富硅铝来源等条件有关^[27]。在这种环境下, 淋溶作用相对较弱, 有利于硅和盐基离子的原位滞留, 从而维持 2:1 型膨胀晶格的稳定性。此外, 该组别与土壤类型的关联受到母质特性的影响, 其分布可能与富含钙镁的湖相沉积物或基性岩母质在特定地貌单元的堆积密切相关。

伊利石组 (Ill、Ill-M) 受水分盈余指数、年均降雨的影响。伊利石是多钾原生铝硅酸盐矿物在强物理风化、弱碱性及弱淋溶条件下风化的产物^[28]。然而, 在长期淹水条件下, 铝硅酸盐矿物的水解反应会受到抑制, 减缓了伊利石向其他矿物类型的转变。硅质组 (Sil、Sil-M) 为原生矿物类别, 受土壤类型、降雨和 pH 的影响。混合型作为成土过程中的过渡性类型, 包含了原生矿物以及次生黏土矿物两种类别。其含量是成土母质与水热条件等协同作用的产物, 其中降雨占据重要位置。

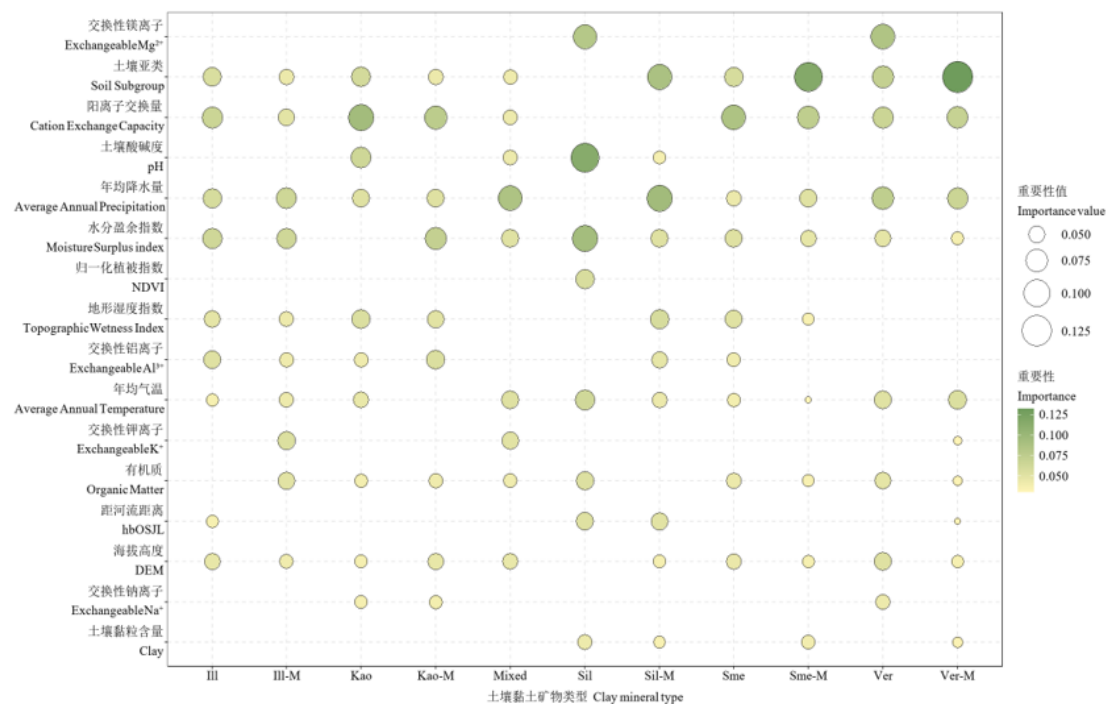


图 5 土壤矿物学类型环境因子相对重要性排序

Fig. 5 Relative importance ranking of environmental factors for soil mineralogical classes

研究区土壤矿物学类型最大概率空间分布如图 6 所示, 与相关环境因子分布密切相关。空间上, 土壤矿物学类型以混合型和伊利石组为主, 蒙脱石型呈现局部富集特征。伊利石组 (Ill、Ill-M) 沿流域分布, 从鄂西山地源头出发, 沿着汉江流域直至中下游的河谷平原, 整体呈典型的“树枝状”延伸结构^[8]。在流域上游山区, 伊利石主要为原生矿物的风化产物。而在中下游平原区, 长期水耕环境导致土壤处于频繁的“干湿交替”, 氧化还原过程呈周期性波动。一部分伊利石脱钾, 当环境中镁离子较多时, 会促使矿物结构向伊利石-蛭石混层体演变^[29-30]。另一部分受长期淹水或滞水环境, 抑制了持续氧化作用和脱硅过程, 使得原生伊利石的保存。

高岭石组 (Kao、Kao-M) 零星分布, 主要分布于鄂东的大别山南麓, 在荆山区域也有少量分布。该区域地势起伏明显, 坡面排水条件良好, 长期的湿热环境及淋溶作用, 使得土壤中的盐基离子大量淋失, 层间钾和镁基本被移除, 硅酸盐矿物强烈分解, 最终演变为以高岭石为主的稳定矿物组合。有研究表明, 酸性介质环境易于高岭石形成, 因此在气候炎热潮

湿的红壤区域，高岭石占主导^[14,31]，这与本研究 pH 空间分布规律一致。

值得注意的是，在鄂东北、鄂东南部分丘陵岗地，出现了面积集中的硅质混合型（Sil-M），这一类型与成土母质、砾石含量及土壤质地密切相关。根据该区域地质图分析，该地区广泛分布着石英砂岩、花岗岩等富含石英的岩系。也有相关研究表明，该地区广泛分布着由花岗岩发育而成的崩岗地貌，水土流失严重。在水动力分选作用影响下，细粒物质被优先搬运流失，而石英砂砾等粗骨颗粒则大量残留并富集^[32-33]。

蒙脱石组（Sme-M、Sme）类型在全省范围内总体概率较低，但在鄂北漫岗洼地（随州、襄阳一带）及局部湖沼低洼地带呈现斑块状分布。其分布范围与该地区砂姜黑土一致，砂姜黑土多位于排水不畅的低洼地或地下水位较高的区域，淋溶水中的钙、镁、硅等离子容易汇聚，形成相对富碱性的地球化学环境，有利于蒙脱石等 2:1 型膨胀性矿物的合成^[27,34]。此外，该类矿物多通过流水悬浮搬运，并在静水环境中沉淀富集。该地区土壤多发育于黄土性古河沉积物，母质本身富含二价阳离子，为其形成提供了基础。此外，蛭石组（Ver、Ver-M）作为风化序列中的过渡性或残留矿物，在研究区内整体含量较低且呈零星分布，仅在局部岩性或特定地形部位分布。混合型在全省主要区域均有分布，集中在江汉平原、鄂北岗地以及鄂西山地有较高丰度的分布。这可能是因为，研究区处于亚热带湿润季风气候，降水充沛但又未达到热带剧烈淋溶的程度，导致大部分土壤处于“中度风化”阶段，从而存在大量的过渡性混层矿物。

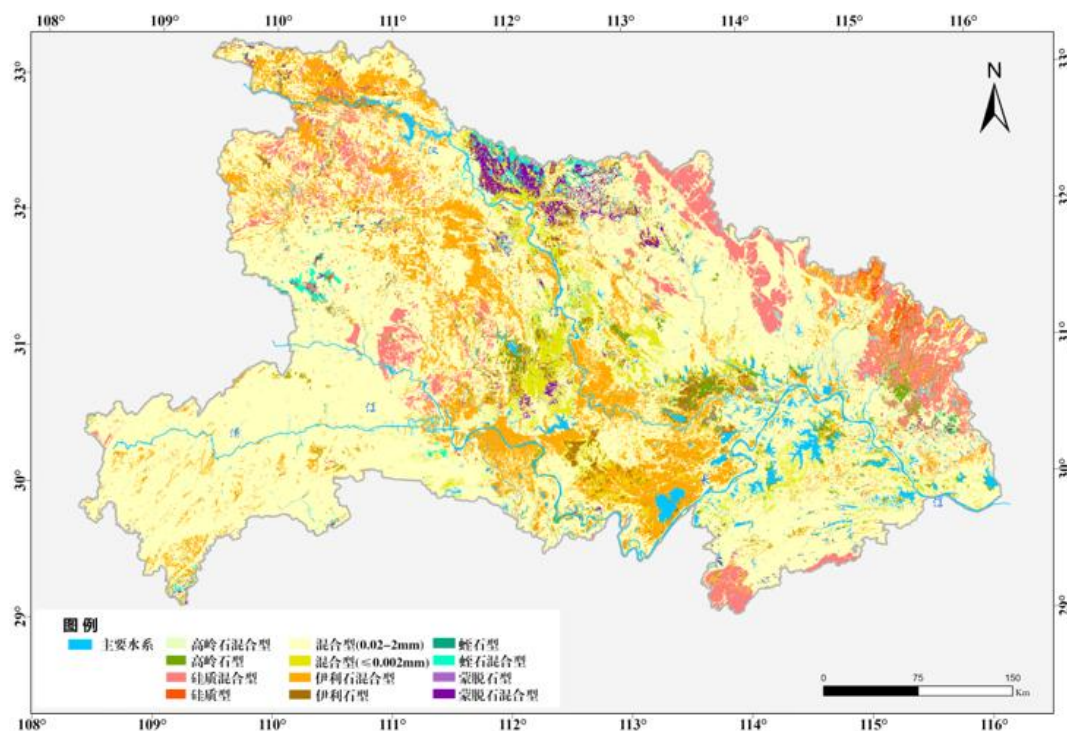


图 6 土壤矿物学类型空间分布

Fig. 6 Spatial distribution of soil mineralogical classes

2.3 制图结果验证

利用实测点对制图结果进行验证，结果表明，不同矿物组分在研究区内的预测概率大小有差异，整体上呈现出混层矿物为主，单一矿物为辅的趋势（图 7）。从各矿物学类型的验证概率均值来看，硅质混合型的均值最高，达到 0.62。混合型、蒙脱石混合型和蛭石混合型，验证概率均值均大于 0.5。其中，混合型类别内最大值达到 0.89，且精度值超过 0.5 的样点占总样点数的 65%。表现出相对较好的可识别性。

相比之下,高岭石混合型和高岭石型的验证概率均值较低,在 0.2 左右,表明模型在判定这些样点时存在较大的不确定性。然而,从类别内最大值来看,高岭石混合型的最大值达到 0.55,高岭石型为 0.54,说明模型并非完全无法识别这些类型,在部分特定环境条件下仍可形成一定程度的有效判别。其均值偏低,更多反映的是该类矿物学类型在区域尺度上的空间分异较强、稳定识别难度较大,而不意味着其空间响应关系完全缺失。其原因可能是,高岭石混合型处于不同矿物组合的过渡状态,矿物组成更为复杂,进一步增加了模型判别难度。与此同时,高岭石通常形成于强风化、强淋溶条件下。而研究区内易形成高岭石的区域同时受到母质条件的显著控制,其土体中砾石含量相对较高,颗粒组成多以粗骨颗粒为主。

此外,伊利石型、伊利石混合型、蒙脱石型和蛭石型的验证概率均值均在 0.32 左右,整体处于中等水平。说明这些类型虽然能够被一定程度识别,但模型判定仍存在较大不确定性。从单个类别内最大值来看,伊利石型为 0.65,伊利石混合型为 0.74,这些类型在局部区域仍表现出较强的可识别性。这表明模型在某些特定环境背景下能够较好地识别这些类型,只是在区域整体尺度上,矿物转化是一个连续过程。各类混合矿物类型往往处于过渡矿物的中间态,不同类型之间在光谱特征、土壤属性以及环境条件上较为相似,导致其平均预测水平较低。

从离散程度来看,蛭石混合型的概率标准差最大,为 0.27。其次为硅质混合型,概率标准差为 0.22。表明这两类矿物学类型具有较强的空间异质性。特别是蛭石混合型,其预测值范围跨度极大。推测在某些环境因子处于特定阈值区间内时,蛭石混合型的特征非常显著,能给出接近 0.95 的高概率值,而一旦环境条件稍有偏离或处于过渡带,概率迅速降低。相比之下,蒙脱石混合型、硅质型和蛭石型的概率标准差均小于 0.05,离散程度低,可能是由于该类型在研究区内分布范围较小且较为集中,空间异质性较弱所致。

验证结果显示出较大的离散度和相对较低的预测置信度,这在很大程度上可归因于验证数据与协变量之间的时空差异以及历史纸质记录造成的定位误差。然而,鉴于土壤矿物的形成和转化通常跨越数千甚至数百万年的地质时间尺度,而非短期的土地覆盖变化。因此,历史测定的矿物类型,在绝大多数样点上,依然代表了当前的矿物学背景。历史数据在“相对稳定性”的假设下,依然是有效的参照基准。

本研究提出方法在区域尺度土壤矿物学类型制图中具有较好的适用性,能够有效揭示研究区主要矿物学类型的空间分布特征。但同时,受模型拟合能力、样本条件、历史数据精度以及类型本身复杂性等因素影响,部分矿物学类型的空间预测仍存在较大不确定性。后续研究仍需进一步加强对拟合类型样本补充,并结合更精确母质信息、砾石含量等指标,以提高土壤矿物学类型制图的精度和可靠性。

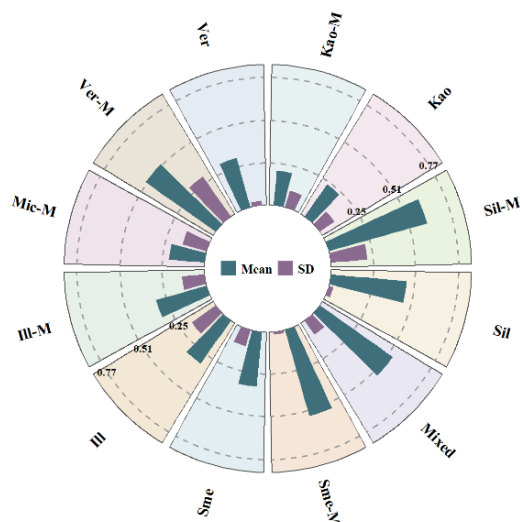


图 7 土壤矿物学类型制图结果准确性平均值和标准差

Fig. 7 Mean and standard deviation of accuracy for soil mineralogical classes results

3 结 论

CatBoost 模型能够较好表征研究区主要土壤矿物学类型空间分布特征,整体预测性能优于其他模型。空间上,研究区土壤矿物学类型以混合型及各类混合矿物类型为主,高岭石相关类型主要集中于鄂东和鄂东南地区,伊利石和蒙脱石相关类型则主要见于鄂北、鄂中及部分沉积环境区域。水热条件、土壤类型、地形水文、母质及沉积环境相关因素共同驱动了土壤矿物学类型的空间分异。不同矿物学类型的验证概率均值存在明显差异,其中混合型、硅质混合型、蒙脱石混合型和蛭石混合型具有较高的概率均值,而高岭石型、高岭石混合型及伊利石相关类型仍具有较高不确定性。总体而言,基于机器学习的数字制图方法在一定程度上能够有效反映研究区主要土壤矿物学类型的基本空间格局,但不同类型之间的预测稳定性和置信度仍存在差异。这种差异不仅与模型性能有关,也与土壤矿物学类型本身的形成机制和判定基础有关。仍需结合更多高精度样点数据和环境协变量提高预测可靠性。

参考文献 (References)

- [1] Belghazdis M, Hachem E K. Clay and clay minerals: A detailed review[J]. *International Journal of Recent Technology and Applied Science*, 2022, 4(2): 54-75.
- [2] Zhang G L, Wang Q B, Zhang F R, et al. Criteria for establishment of soil family and soil series in Chinese soil taxonomy[J]. *Acta Pedologica Sinica*, 2013, 50(4): 826-834. [张甘霖, 王秋兵, 张凤荣, 等. 中国土壤系统分类土族和土系划分标准[J]. *土壤学报*, 2013, 50(4): 826-834.]
- [3] Zhang Z Y, Huang L, Liu F, et al. The properties of clay minerals in soil particles from two ultisols, China[J]. *Clays and Clay Minerals*, 2017, 65(4): 273-285.
- [4] Liu F, Xu F L, Li X Y. Studies on clay minerals of mountain soils in Hubei and Hunan Provinces—III. Clay minerals and surface chemical characteristics of soils on north slope of Shennongjia natural conservation area[J]. *Acta Pedologica Sinica*, 1996, 33(1): 59-69. [刘凡, 徐凤琳, 李学垣. 鄂湘两省山地土壤粘粒矿物的研究—III. 神农架自然保护区北坡土壤的粘粒矿物与表面化学特性[J]. *土壤学报*, 1996, 33(1): 59-69.]
- [5] He J Z, Li X Y, Liu F, et al. Clay mineral composition and surface charge properties of soils in Wudang Mountain[J]. *Tropical and Subtropical Soil Science*, 1995, 4(4): 228-232. [贺纪正, 李学垣, 刘凡, 等. 武当山土壤的粘粒矿物与表面电荷性质[J]. *热带亚热带土壤科学*, 1995, 4(4): 228-232.]
- [6] Li X Y, Xu F L. Surface chemical properties, chemical and mineral composition of clay for main soils distributed in alluvial plain and hills mound in Hubei Province[J]. *Journal of Huazhong Agricultural University*, 1999, 18(5): 420-426. [李学垣, 徐凤琳. 湖北省丘岗、平原主要土壤的表面化学性质与粘粒的矿物、化学组成[J]. *华中农业大学学报*, 1999, 18(5): 420-426.]
- [7] Huang L, Liu F, Tan W F, et al. The difference and significance of 1.4 nm intergrade minerals between matrix and cutan in several subtropical alfisol[J]. *Acta Mineralogica Sinica*, 2002, 22(4): 315-320. [黄丽, 刘凡, 谭文峰, 等. 几种亚热带淋溶土基质和胶膜中 1.4 nm 过渡矿物组成的差异及其意义[J]. *矿物学报*, 2002, 22(4): 315-320.]
- [8] Hu H Q, He J Z, Wu C B. Clay mineralogy and surface properties of Hubei Province soils derived from Quaternary period deposit[J]. *Journal of Huazhong Agricultural University*, 1998, 17(4): 335-340. [胡红青, 贺纪正, 武从斌. 湖北省几种土壤的粘土矿物组成与表面化学性质[J]. *华中农业大学学报*, 1998, 17(4): 335-340.]
- [9] Hseung Y, Hsu C C. Frequency distribution of clay minerals in the soils of China[J]. *Acta Pedologica Sinica*,

- 1964, 12(3): 266-274. [熊毅, 许冀泉. 中国土壤中粘粒矿物的分布规律[J]. 土壤学报, 1964, 12(3): 266-274.]
- [10] Long H Y, Lu C A, Ji H J, et al. A retrieval system for great soil groups from China's Provisional Soil Classification System for the 3rd National Soil Census[J]. Scientia Agricultura Sinica, 2024, 57(21): 4264-4275. [龙怀玉, 卢昌艾, 冀宏杰, 等. 《第三次全国土壤普查暂行土壤分类系统》土类检索系统构建[J]. 中国农业科学, 2024, 57(21): 4264-4275.]
- [11] Li J W, Chen Y M, Zheng L Y, et al. Development of soil classification and research avenue[J]. Soils and Crops, 2014, 3(4): 146-150. [李建维, 陈一民, 郑利远, 等. 土壤分类的发展与方向[J]. 土壤与作物, 2014, 3(4): 146-150.]
- [12] Zhang G L, Zhu A X, Shi Z, et al. Progress and future prospect of soil geography[J]. Progress in Geography, 2018, 37(1): 57-65. [张甘霖, 朱阿兴, 史舟, 等. 土壤地理学的进展与展望[J]. 地理科学进展, 2018, 37(1): 57-65.]
- [13] Yang J W, Guan X F, Luo M Y, et al. Cross-system legacy data applied to digital soil mapping: A case study of Second National Soil Survey data in China[J]. Geoderma Regional, 2022, 28: e00489.
- [14] Ndzana G M, Huang L, Wang J B, et al. Characteristics of clay minerals in soil particles from an argillic horizon of Alfisol in Central China[J]. Applied Clay Science, 2018, 151: 148-156.
- [15] Guan Z Z. The significance of clay minerals characteristics & the environmental changes in lower reaches of the Yangtze River[D]. Shanghai: East China Normal University, 2007. [管章志. 长江下游流域环境变化的粘土矿物表征及其意义[D]. 上海: 华东师范大学, 2007.]
- [16] Guo L, Sun X R, Fu P, et al. Mapping soil organic carbon stock by hyperspectral and time-series multispectral remote sensing images in low-relief agricultural areas[J]. Geoderma, 2021, 398: 115118.
- [17] Bao Y L, Yao F M, Meng X T, et al. A fine digital soil mapping by integrating remote sensing-based process model and deep learning method in Northeast China[J]. Soil & Tillage Research, 2024, 238: 106010.
- [18] Pouladi N, Gholizadeh A, Khosravi V, et al. Digital mapping of soil organic carbon using remote sensing data: A systematic review[J]. Catena, 2023, 232: 107409.
- [19] Chen Y C, Yang Y H. The one standard error rule for model selection: Does it work?[J]. Stats, 2021, 4(4): 868-892.
- [20] Yang J W, Wang T W, Que S X, et al. Effect of colour calibration on the prediction of soil organic matter content based on original soil images obtained from smartphones under different lighting conditions[J]. Soil & Tillage Research, 2024, 238: 106018.
- [21] Yang J W, Wang T W, Bi Y H, et al. Harnessing expert knowledge and legacy data for digital soil mapping with no new field surveys[J]. Geoderma Regional, 2025, 42: e00998.
- [22] Zhang Q H, Cao B G, Jiang L. Development and construction of the 1: 500000 geological map database[J]. Geology in China, 2002, 29(2): 208-212. [张庆合, 曹邦功, 姜兰. 1: 50 万地质图数据库的研建[J]. 中国地质, 2002, 29(2): 208-212.]
- [23] Yang J, Huang X. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019[J]. Earth System Science Data, 2021, 13(8): 3907-3925.
- [24] Liu F, Wu H Y, Zhao Y G, et al. Mapping high resolution national soil information grids of China[J]. Science Bulletin, 2022, 67(3): 328-340.
- [25] Liu F, Zhang G L, Song X D, et al. High-resolution and three-dimensional mapping of soil texture of China[J]. Geoderma, 2020, 361: 114061.
- [26] Huang Q Y, Xu F L, Li X Y. The clay mineral association and 14 Å mineral of the zonal soils in central and Southern China[J]. Journal of Huazhong Agricultural University, 1989, 8(1): 45-52. [黄巧云, 徐凤琳, 李学垣. 中南地区几种地带性土壤的粘粒矿物组合与 14 埃矿物[J]. 华中农业大学学报, 1989, 8(1): 45-52.]
- [27] Jin P P, Ou C H, Ma Z G, et al. Evolution of montmorillonite and its related clay minerals and their effects on

- shale gas development[J]. *Geophysical Prospecting for Petroleum*, 2018, 57(3): 344-355. [靳平平, 欧成华, 马中高, 等. 蒙脱石与相关黏土矿物的演变规律及其对页岩气开发的影响[J]. *石油物探*, 2018, 57(3): 344-355.]
- [28] Zhang Z Y, Huang L, Lu S, et al. Characteristics of clay minerals in the zonal soil from two kinds of parent rock[J]. *Journal of Huazhong Agricultural University*, 2015, 34(3): 51-58. [张志毅, 黄丽, 卢胜, 等. 2 种母岩发育的地带性土壤中粘土矿物的组成特点[J]. *华中农业大学学报*, 2015, 34(3): 51-58.]
- [29] Xu F L, Li X Y, Liu F, et al. The 1.4 nm intergrade mineral of acid soils in the Southern China[J]. *Journal of Huazhong Agricultural University*, 1991, 10(3): 274-280. [徐凤琳, 李学垣, 刘凡, 等. 我国南方酸性土壤中的 1.4 nm 过渡矿物(英文) [J]. *华中农业大学学报*, 1991, 10(3): 274-280.]
- [30] Ndzana G M. The transformation of clay minerals from macro zonal soils in China to nanoscale in subsoil fractions of several[D]. Wuhan: Huazhong Agricultural University, 2018. [乔治. 中国几种地带性土壤颗粒中黏粒矿物从宏观到纳米尺度的转化[D]. 武汉: 华中农业大学, 2018.]
- [31] Sun Q F, Colin C, Chen F H, et al. A discussion on the factors affecting formation and quantity of clay minerals in climatic and environmental researches[J]. *Acta Petrologica et Mineralogica*, 2011, 30(2): 291-300. [孙庆峰, Christophe Colin, 陈发虎, 等. 气候环境变化研究中影响粘土矿物形成及其丰度因素的讨论[J]. *岩石矿物学杂志*, 2011, 30(2): 291-300.]
- [32] Ni S M, Zhou Y, Cai C F, et al. Insights into linkage between water erosion and Benggang formation: Trajectory tracing from slope to gully[J]. *Catena*, 2025, 254: 109001.
- [33] Li K K, Yang J Y, Wang J Y, et al. Human-altered soil loss dominates nearly half of water erosion in China but surges in agriculture-intensive areas[J]. *One Earth*, 2024, 7(11): 2008-2018.
- [34] Li W D, Wang Q Y. Composition of clay minerals and discussion of the origin of smectite in Shajiang black soils[J]. *Acta Agriculturae Universitatis Pekinensis*, 1994, 20(2): 192-196. [李卫东, 王庆云. 砂姜黑土的粘粒矿物组成和蒙皂石来源的探讨[J]. *北京农业大学学报*, 1994, 20(2): 192-196.]

(责任编辑: 卢 萍)