土 壤 学 报
Acta Pedologica Sinica

DOI: 10.11766/trxb202502150061 CSTR: 32215.14.trxb202502150061

李欣烨，石璞，刘航，杨永. 蒙东黑土区有效土层厚度空间制图及其地表基质制约机理研究[J]. 土壤学报, 2026,

LI Xinye, SHI Pu, LIU Hang, YANG Yong. Spatial Mapping of Effective Soil Thickness and Its Surface Substrate Constraint Mechanism in the Black Soil Area of Eastern Mongolia[J]. Acta Pedologica Sinica, 2026,

蒙东黑土区有效土层厚度空间制图及其地表基质制约机理研究*

李欣烨[1]，石 璞[1]，刘 航[2†]，杨 永[2]

（1.吉林大学地球科学学院,长春 130061; 2.中国地质调查局廊坊自然资源综合调查中心,河北廊坊 065000）

摘 要：有效土层厚度是评价土壤健康和生产力的一个决定性指标，精确描绘有效土层厚度的空间分布格局及其对土地利用变化及地表基质类型的响应机制对于土壤资源的可持续保护具有重要意义。本文选取内蒙古东部黑土区为研究区，基于地表基质调查数据和土壤-景观建模，开展该地区有效土层厚度数字化制图与空间格局分析，通过 SHAP 解释模型（Shapley additive explanation）识别有效土层厚度空间变异的主控因子，查明不同土地利用类型和地表基质分区下有效土层厚度的差异性分布规律。结果表明：基于 Cubist 的有效土层厚度回归模型性能良好（R^2=0.5,RMSE=43.8），所生成的空间分布图能够准确揭示其空间格局特征。SHAP 分析揭示了地形和气候因子是决定有效土层厚度空间变异的主控因子，具体表现在同一流域内高海拔区域受侵蚀作用影响，土层较薄；而月均气温极值对有效土层厚度的影响为正向。地表基质分区和土地利用类型均对有效土层厚度的空间特征有重要制约作用，坡洪积土和坡积土区土层厚度整体大于残坡土区；林地由于多分布于山区或坡度较大的区域，有效土层厚度最低。本文为有效土层厚度的空间建模与表征提供了方法参考，研究结果可为查明区域自然资源本底条件及其对人类互动的响应机制提供数据基础。

关键词：有效土层厚度；数字土壤制图；SHAP；地表基质；土地利用类型

中图分类号：S159 **文献标志码**：A

Spatial Mapping of Effective Soil Thickness and Its Surface Substrate Constraint Mechanism in the Black Soil Area of Eastern Mongolia

LI Xinye[1]，SHI Pu[1]，LIU Hang[2†]，YANG Yong[2]

(1. *College of Earth Sciences, Jilin University, Changchun* 130061, *China; 2.Langfang Center for General Survey of Natural Resources, China Geological Survey, Langfang, Hebei* 065000, *China*)

Abstract:【Objective】Effective soil thickness is a decisive indicator for evaluating soil health and productivity, it is therefore of great significance to accurately depict the spatial distribution pattern of effective soil thickness and its response mechanism to land use change and surface substrate type for the sustainable protection of soil resources.【Method】In this study, leveraging on the surface substrate survey data and soil-landscape modeling, we carried out predictive modelling and mapping of effective soil thickness in the black soil area of eastern Inner Mongolia. Based on the modelling results, the spatial variability of effective soil thickness was analyzed among land use types and surface substrate zoning. SHAP

* 国家重点研发计划项目（2024YFD1500300）、中国地质调查局项目（DD20242037）资助 Supported by the National Key Research and Development Program of China（No. 2024YFD1500300）and the Project from China Geological Survey（No. DD20242037）

† 通讯作者 Corresponding author，E-mail：lh950511@qq.com

作者简介：李欣烨（2002—），女，博士，主要从事自然资源调查与遥感监测研究。E-mail: lixy2220@mails.jlu.edu.cn

收稿日期：2025-02-15；收到修改稿日期：2025-08-03；网络首发日期（www.cnki.net）：2025- -

analysis was employed to identify the main controlling factors underlying the spatial distribution pattern of effective soil thickness.【Result】The results show that the Cubist-based regression model had a good performance (R^2=0.5,RMSE=43.8) for effective soil thickness prediction, and the generated spatial distribution map could accurately reveal its spatial pattern. SHAP analysis revealed that topographic and climatic factors were the main controlling factors determining the spatial variability of effective soil thickness, which was specifically reflected in the fact that highly eroded areas with high elevation had thinner soils, while monthly average temperature extremes had a positive effect.【Conclusion】Both surface substrate zoning and land use types exerted important constraints on the spatial characteristics of the effective soil layer thickness, with the overall soil layer thickness in the floodplain and sloping deposit areas being greater than that in the residual slope deposit area. Forestlands, which were mostly distributed in mountainous areas or regions with steep slopes, had the thinnest effective soil layer. This study provides a methodological reference for the spatial modeling and characterization of effective soil thickness, and the results can provide a data basis for identifying the background conditions of regional natural resources and their response mechanisms to human interactions.

Key words: Effective soil thickness; Digital soil mapping; SHAP; Surface substrate; Landuse

有效土层厚度作为土壤的重要属性之一[1]，直接制约植物生长[2]、植被覆盖、土壤养分和水分迁移[3]等一系列地表生物和水土过程，对土壤碳库的固存、流失、迁移与转化具有重要的控制作用。农业农村部《耕地质量等级》GB/T33469—2016 将有效土层厚度定义为作物能够利用的母质层以上的土地总厚度，当有障碍层时，为障碍层以上的土层厚度。在我国东北黑土区，长期高强度的耕作和漫川漫岗的地形条件导致黑土层变薄，严重影响土壤质量和农作物产量[4-5]，近年的研究报道典型坡耕地受土壤侵蚀等退化过程影响，出现以土层变薄和土壤碳库流失为典型表现的“破皮黄”现象[6]。因此，亟需开展黑土区有效土层厚间分布特征研究，探索其空间变异特征对成土因素、土地利用变化等因素的差异性响应特征，促进生态保育与修复、降低土壤退化及地质灾害的潜在风险，为区域土地资源的可持续利用与保护提供科学支撑[7]。

有效土层厚度空间特征的定量评价方法首先依赖于野外调查与实测，通过实地挖掘土壤剖面和利用钻探设备进行定点取样获取有效土层厚度数据。例如，刘凯等[8]对东北地区 61 处典型土壤剖面进行实地考察，研究表明东北地区 40 年来土层厚度减少了 12cm。由于野外实地调查难以进行大尺度、高密度的采样和数据获取，基于野外点数据的有效土层厚度调查通常难以详细揭示其空间异质性及驱动机制。为克服这一瓶颈，土壤-景观模型模拟、遥感大数据集成应用、机器学习建模等多元化手段逐渐应用于有效土层厚度数字化制图领域。根据 McBratney 提出的 SCORPAN 范式[9]，数字土壤制图的工作原理是在有限的实测数据的支撑下，建立气候、地形、植被、成土母质、时间、空间等环境协变量与待预测土壤指标之间的映射关系，通过统计模型或机器学习算法实现空间显式制图[10]。伴随着人工智能技术的快速发展，机器学习算法由于能够表征环境变量与土壤属性之间的非共线关系而被广泛采用。Yamashita[11]利用随机森林预测了日本山区 A 层厚度超过 15 cm、A 层和 B 层厚度超过 75 cm 的概率图；Wang 等[12]利用快速均值聚类和地理加权回归将研究区聚类为若干子区域，建立回归模型，通过分层抽样对土层厚度进行了区域性的制图；Chen 等[13]利用分位数随机森林方法测试黄土沉积物的厚度，结果表明航空伽马射线放射测量数据和 Sentinel 2 产品在法国地区黄土厚度预测中的重要性。目前，国内外学者对土层厚度的研究多聚焦于利用机器学习方法提

高土层厚度制图的预测精度，然而，对土层厚度空间分布规律的分析及其与地质、地理、地貌等因素的响应关系探讨尚显不足，相关研究亟待进一步深入。

“地表基质”这一概念由自然资源部于 2020 年 1 月 18 日印发的《自然资源调查监测体系构建总体方案》首次提出，是指“当前出露于地球陆域地表浅部或水域水体底部，主要由天然物质经自然作用形成，正在或可以孕育和支撑森林、草原、水等各类自然资源的基础物质”。作为国内首次提出的概念，地表基质调查旨在深入理解土地资源的分布、评估和保护生态环境，查清高强度人类活动制约下地表基底条件对地球关键物质循环和能量流动过程的协同影响机理。国际上与地表基质相关的研究范畴包括生态地质学、地球关键带等[14]，但这些研究范式均面临跨学科整合困难、观测数据不足等瓶颈，亟需建立“地质调查-数据科学-自然资源系统评价”有机融合的研究新范式。中国地质调查局于 2020 年在全国范围内开展了地表基质调查试点工作，通过地表基质分区和构型的划分，厘清地表基质的物质成分、空间结构、理化性质、景观属性和生态功能等。其中，地表基质调查工作包括对有效土层厚度的调查，这为分析有效土层厚度对地表基质结构和功能的响应机制，揭示地表基质空间异质性对土壤生产和生态功能的控制机理提供了重要的数据基础。

蒙东地区（内蒙古自治区东部）是我国重要的农业生产区域，但该地区地形和气候恶劣，坡耕地集中，风蚀和水蚀现象严重，耕作层逐渐变薄，土地肥力下降，严重限制了农业的可持续发展。土地退化现象的加剧还影响了该区域草地资源的承载力，草原沙化日趋突出[15]。在此背景下，揭示蒙东地区有效土层厚度与区域地表基质本底条件和人类活动的空间异质关系，有助于明晰该区域的自然资源概况和未来重点发展方向。为此，本文以内蒙古东部阿荣旗和莫力达瓦达斡尔族自治旗（以下简称“莫旗”）为研究区，基于区内地表基质调查的 518 个采样点和 34 个环境变量，开展基于 Cubist 模型的有效土层厚度预测建模和空间制图，通过 SAHP 分析法测度不同环境变量对模型预测的重要性，探究有效土层厚度与地表基质分区、土地利用类型的空间关联关系，为查明该区域自然资源本底条件、制定区域黑土资源可持续利用与保护策略提供参考依据。

1 材料与方法

1.1 研究区概况

1.1.1 自然地理概况　研究区为内蒙古自治区呼伦贝尔市阿荣旗和莫旗（纬度：47°56'54"—49°50′50″N，经度：122°02'30"—125°16′14″E），占地约 2.4×10^4 km^2，地理位置位于内蒙古自治区东北部，大兴安岭向松嫩平原过渡的黑土带，全境地貌呈中低山—丘陵漫岗地形，海拔范围为 106~1152 m，地势从西北向东南呈阶梯式下降[16]（图 1）。该区域土壤类型主要包括黑土、暗棕壤、草甸土、沼泽土等，其中黑土占耕地面积的 90%以上，农业资源优越；气候类型属于温带大陆性半湿润气候，年平均气温为 1.3℃至 1.7℃，平均降水量为 400~500 mm，主要集中在夏季 6—8 月，占全年降水量的 70%。

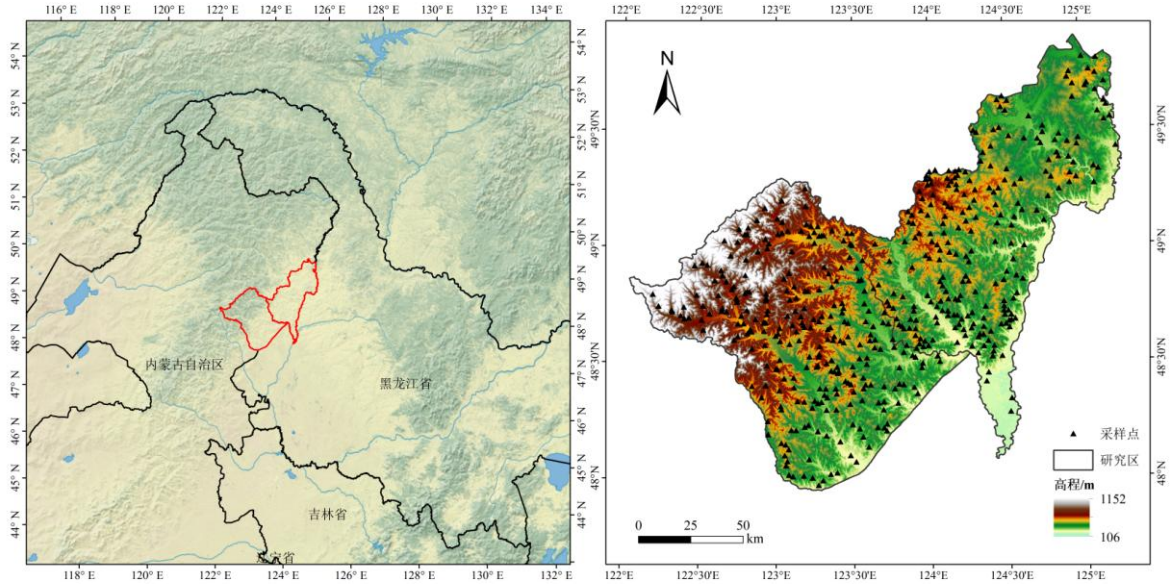

图 1 研究区区位及采样点分布

Fig. 1 Location of the study area and distribution of sampling sites

1.1.2 区域地表基质概况　　自然资源部 2020 年发布的《地表基质分类方案（试行）》旨在对地表基质的主要成分进行细致分类，采用包含四个一级类别和三级细分的体系。一级类别涵盖了岩石、砾质、土质和泥质四种基本类型；在二级类别进一步扩展至 14 种不同类别；三级类别在二级类别的基础上进行更详尽的划分。

如图 2 所示，研究区地表基质分区表现出显著的空间分异特征。大兴安岭隆起带中低山残破积地表基质区主要分布在西部和南部的山区等海拔较高的区域，是面积最大的地表基质分区，占研究区面积的 44%；其次，大兴安岭隆起带丘陵残破积地表基质区，分布在研究区的东部和南部一小部分，通常位于海拔中等或较低的平原地带，是山地向平原或河谷过渡的中间地形，占研究区面积的 22.8%；大兴安岭隆起带山间谷地冲洪积地表基质区呈脉状分布，沿河谷平原扩展，占研究区的 11.8%；大兴安岭隆起带丘陵坡积地表基质区集中在研究区东北呈块状分布；大兴安岭隆起带丘陵坡洪积地表基质区在东北部沿着山间谷地冲洪积地表基质区呈脉状分布；松嫩平原沉降带阶地漫滩冲洪积地表基质区分布在研究区地势最低的区域，沿着水体分布。

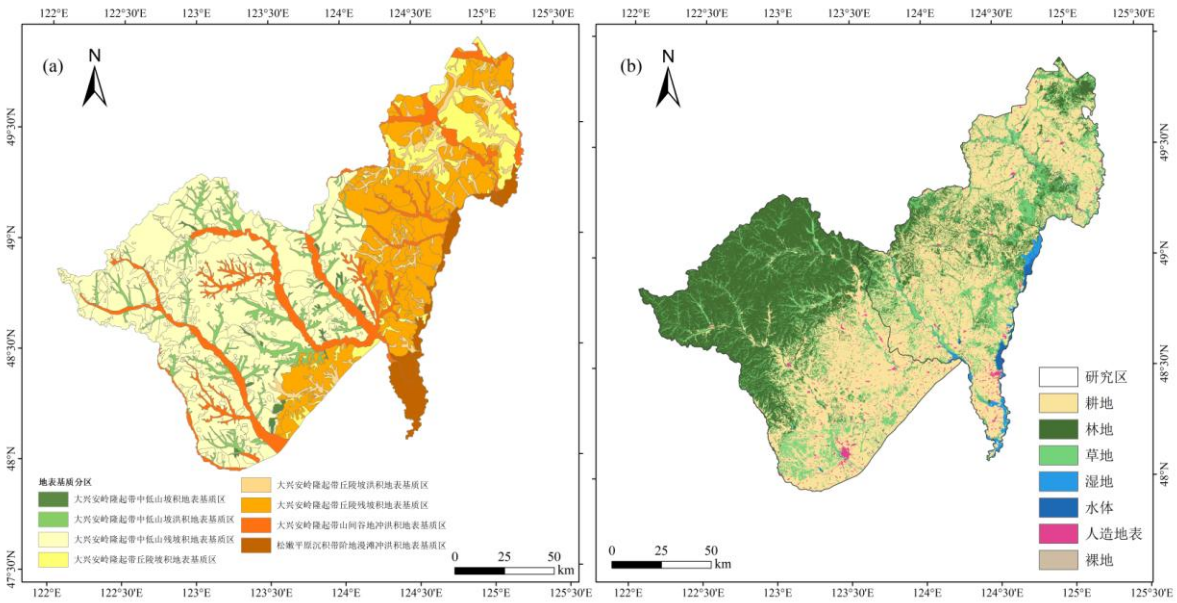

图 2 研究区地表基质分区空间分布图（a）和土地利用类型空间分布图（b）

Fig.2 Spatial distribution of surface substrate types (a) and spatial distribution of land use types (b)

有效土层厚度的野外测定工作于 2022 年进行。根据研究区土壤类型、成土母质、地质成因、地形地貌以及土地利用类型等环境变量，遵循分区、分类、分层次的原则，利用背包钻和洛阳铲进行有效土层厚度测量。基于《黑土地地表基质调查技术要求（试行）》：有效土层厚度定义为作物能够利用的母质层以上的土体总厚度，当有障碍层或非土体时，为障碍层或非土体以上的土层厚度。常见障碍层包括粘盘层、铁盘层、砂姜层、砂砾层、盐积层、石膏层、白土层、白浆层、灰化层、潜育层、冻土层等。有效土层厚度最大值定为 2 m，即对于有效土层厚度超过 2 m 且无法识别有效土层厚度的调查点位，未进行 2 m 以深的测量。为保证所测数据的可信度，项目组在野外工作过程中进行了 100%的自检，100%的互检以及 30%的项目组抽检。最终，研究区内可用于有效土层厚度空间预测建模的样本量为 518 个。

1.1.3 土地利用概况　　从土地利用类型空间分布特征来看（图 2），农业生产在该区域占据主导地位，超过 46%的面积为耕地，主要分布在中部和东部等地势较为平坦的地区；其次为林地，占研究区面积的 31.4%，主要分布在西部的山地、丘陵地区；草地资源的分布相对较为分散，多处于耕地和林地的交界地带，占研究区面积的 19.3%；湿地则主要沿水体集中分布在东部和中部地区；人造地表和裸地仅占研究区的一小部分，不作为后续研究对象。

1.2 研究方法

本研究采取的技术流程如图 3 所示。首先，基于地形、气候、遥感和土壤四个方面共计 34 个环境变量作为模式解释因子，对比 Cubist、随机森林（Random forest，RF）、极端梯度提升（Extreme gradient boosting，XGBoost）三种建模算法对有效土层厚度的预测精度。在模型评估方面，利用 Kennard-Stone 算法划分 80%的训练集和 20%的验证集，以此来评估模型的准确性并筛选出性能最优的算法。其次，基于最佳建模算法，通过 50 次 Bootstrap 重复抽样分析模型不确定性。最后，根

据模型预测和空间制图结果，计算各因子 SHAP 值进行变量重要性分析，并对比不同土地利用类型与地表基质分区下有效土层厚度的空间差异。

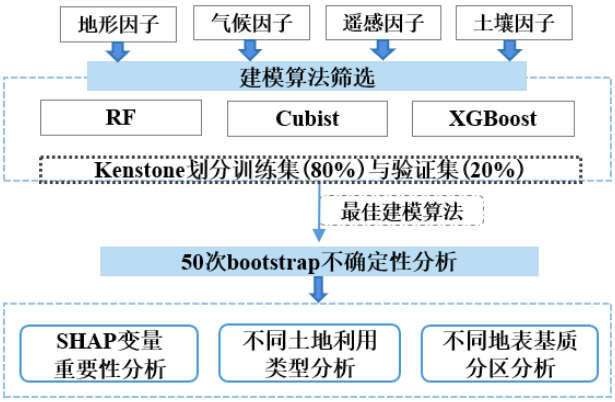

图 3 研究流程图

Fig.3 Research workflow diagram

1.2.1 Cubist 回归预测　Cubist 模型是一种基于规则和实例的回归预测模型，它是对 Quinlan 的 M5 模型树的扩展[17]。其基于回归树，这种树状结构将输入空间划分为不同的区域，并在每个区域中拟合一个简单的线性模型，从而捕获数据中的复杂非线性关系。其主要参数是模型复杂度（commities），用于指定要创建的回归树的数量，默认是 1。本文基于 Cubist 模型利用环境协变量对有效土层厚度进行预测。本文在建模时使用 Kennard-Stone 算法[18]选取样本作为模型的训练集和验证集。

1.2.2 随机森林　随机森林算法是一种集成学习方法，它通过构建多棵决策树并结合它们的预测结果来提高整体模型的准确性。构建过程包括从原始训练集中有放回地抽取样本形成新的训练子集，对每个节点的分裂随机选择特征，并使用最佳分裂方式进行分裂[19]。重复这一过程构建多棵决策树，形成随机森林。随机森林的关键参数包括森林中决策树的数量（n_estimators）、每次分裂时考虑的特征数量（max_depth）、决策树的最大深度（max_features）。

1.2.3 XGBoost　XGBoost[20]的基本原理是梯度提升，通过逐步构建多个弱模型并将它们组合成一个强模型。每个弱模型都试图纠正前一个模型的错误，最终形成一个高度准确的预测模型。XGBoost 使用二阶泰勒展开近似损失函数，并在每一轮迭代中通过最小化目标函数来选择最佳的树分裂点。重要参数包括决策树的数量（n_estimators）、决策树的最大深度（max_features）、学习率（learning_rate）。

1.2.4 评价指标　本文采取均方根误差（Root Mean Square Error，RMSE）、平均绝对误差（Mean Absolute Error，MAE）和决定系数（R^2）三种评估指标[21]对 cubist 模型的性能进行评价，其公式为：

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(p_i - \hat{p}_i)^2} \quad (1)$$

$$\mathrm{MAE} = \frac{\sum_{i=1}^{N}|(p_i - \hat{p}_i)|}{n} \quad (2)$$

$$R^2 = 1 - \frac{\sum_{i=0}^{N}(p_i - \hat{p}_i)^2}{\sum_{i=1}^{N}(p_i - \bar{p}_i)^2} \quad (3)$$

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

1.2.5 SHAP 解释模型　　SHAP 解释模型是 Lundberg 和 Lee[22]提出通过计算一个特征加入到模型时的边际贡献，然后考虑到该特征在所有的特征序列的情况下不同的边际贡献，取均值，其表达式如下：

$$\varphi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|s|!\,(|N| - |S| - 1)!}{|N|!} [vS \cup \{i\} - v(S)] \quad (4)$$

式中，φ_i表示特征i的 shapley 值，S表示特征子集，N表示所有特征的集合，$v(S)$表示子集S的值函数。

在 SHAP 框架下，该值可理解为特征i对模型预测的期望边际贡献：

$$\mathrm{SHAP}_i = \mathrm{Expected}[f(S \cup \{i\} - f(S))] \quad (5)$$

本文通过计算 SHAP 值对 Cubist 模型进行解释，分析环境变量的重要性。

1.3 数据处理

基于 Scorpan 土壤景观模型[9]，本研究选取 34 个变量作为预测有效土层厚度的环境因子（表 1），将其空间分辨率重采样为 30 m，并统一掩膜至与研究区空间范围一致。

表 1 环境变量基本信息

Table 1 Basic information of environmental variables

	环境变量 Environmental variables	简称 Abbreviation	分辨率 Resolution	数据来源 Data Sources
地形因子 Topographic factor	高程 Elevation	DEM	30	DEM
	坡度 Slope	Slope	30	DEM
	坡向 Aspect	Aspect	30	DEM
	流量 Flow accumulation	FA	30	DEM
	流向 Flow direction	FD	30	DEM
	坡长 Flow length	FL	30	DEM
	坡度之坡度 Slope of slope	SOS	30	DEM
	单位汇水面积 Specific catchment area	SCA	30	DEM
	水流强度指数 Stream power index	SPI	30	DEM
	地形湿度指数 Topographic wetness index	TWI	30	DEM
	地形位置指数 Topographic position index	TPI	30	DEM
	地形粗糙度指数 Topographic roughness index	TRI	30	DEM
	泥沙输送指数 Sediment transport index	STI	30	DEM
气候因子 Climate factor	年平均降水量 Mean annual precipitation	MAP	1 000	中国 1km 分辨率逐月平均降水数据集
	年平均气温 Mean annual	MAT	1 000	中国 1km 分辨率逐

	temperature			月平均气温数据集
	月均气温最大值 Mean month temperature maximun	MMTmax	1 000	中国 1km 分辨率逐月平均气温数据集
	月均气温最小值 Mean month temperature minimun	MMTmin	1 000	中国 1km 分辨率逐月平均气温数据集
遥感因子 Remote sensing factor	归一化植被指数 Normalized difference vegetation index	NDVI	30	Landsat8
	增强植被指数 Enhanced vegetation index	EVI	30	Landsat8
	地表温度 Land surface temperature	LST	30	Landsat8
	地表温度最大值 Land surface temperature maximun	LSTm	30	Landsat8
	光合有效辐射 Fraction of photosynthetically active radiation	FPAR	500	MODIS
	光合有效辐射最大值 Fraction of photosynthetically active radiation maximum	FPARm	500	MODIS
土壤因子 Soil factor	砂粒含量 Sand	Sand	90	中国高分辨率国家土壤信息格网基本属性数据集
	粉粒含量 Silt	Silt	90	中国高分辨率国家土壤信息格网基本属性数据集
	黏粒含量 Clay	Clay	90	中国高分辨率国家土壤信息格网基本属性数据集
	土壤 pH Soil pH	pH	90	中国高分辨率国家土壤信息格网基本属性数据集
	土壤有机碳 Soil organic carbon	SOC	90	中国高分辨率国家土壤信息格网基本属性数据集
	阳离子交换量 Cationic exchange capacity	CEC	90	中国高分辨率国家土壤信息格网基本属性数据集
	全氮含量 Total nitrogen	TN	90	中国高分辨率国家土壤信息格网基本属性数据集
	全磷含量 Total phosphorus	TP	90	中国高分辨率国家土壤信息格网基本属性数据集

土壤质地类型 Soil texture classification	TEXCL	90	中国高分辨率国家土壤信息格网基本属性数据集
砾石含量	Cf	90	中国高分辨率国家土壤信息格网基本属性数据集

数字高程模型（Digital Elevation Model, DEM）数据采用 ASTER GDEM 数据（国家基础数据中心），其余地形衍生因子利用 Arcmap 和 SAGA 8.4.1 计算；气候因子采用中国 1 km 分辨率 2021 年逐月平均降水和气温数据集；遥感因子基于谷歌地球引擎（Google Earth Engine, GEE）平台获取，在 ArcMap10.4 版本中进行预处理。其中 NDVI、EVI、LST、LSTm 来源于 Landsat8 数据，FPAR、FPARm 来源于 MODIS 数据，时间范围是 2023 年；土壤因子采用中国高分辨率国家土壤信息网格基本属性数据集；以上数据均用 R4.4.2 进行计算和裁剪。

研究中土地利用数据来自 2020 年 Globe30（http://www.globallandcover.com），分辨率 30 m，研究区土地利用分类为耕地、林地、草地、湿地、水体、人造地表和裸地。其中水体与人造地表不在后续研究范围。

2 结果与讨论

2.1 描述性统计

根据有效土层厚度的调查结果，研究区平均土层厚度为 94.55 cm，但呈现出显著的空间变异性。研究区面积为 2.4×10^4 km^2，采样点为 518 个，相较于已有部分研究，采样密度较高。对于自然资源进行研究时，采样密度是一个非常重要的因素，他直接影响到数据收集的全面和预测结果的准确性。如表 2 所示，有效土层厚度的变异系数为 60.95，数据分布离散，偏度值为 0.66（＞0.5），数据呈轻微的正偏态，峰度值为 2.08（＞3），数据的分布相对正态分布较为平坦，尾部更重。使用对数变换，虽然可以转化为正态分布，但对模型精度没有提升。因此，本文使用原始数据进行预测模型的构建。

表 2 有效土层厚度描述性统计

Table 2 Descriptive statistical results of the effective soil thickness

样本数量 Number	中位数 Median/cm	平均数 Mean/cm	标准差 SD/cm	偏度 Skewness	峰度 Kurtosis	变异系数 CV/%
518	75	94.55	57.63	0.66	2.08	60.95

注：Standard Deviation，SD

2.2 有效土层厚度空间建模与制图

不同模型 Cubist、RF 和 XGBoost 对有效土层厚度进行建模精度评估，R^2、MAE 和 RMSE 变化如表 3 所示。结果表明，Cubist 模型预测精度最好（R^2=0.50，MAE=33.05 cm，RMSE=43.79 cm）。

建模时采用了被广泛使用的 Kennard-Stone 算法，该算法是 Kennard 和 Stone[23]提出的一种基于样本对之间欧氏距离比较的算法，其可以选出数据空间中最能代表整体分布信息的样本。

表 3 不同模型的预测性能

Table 3 Predictive performance of different models

	决定系数 R^2	平均绝对误差 MAE/cm	均方根误差 RMSE/cm
Cubist	0.50	33.05	43.79
随机森林 RF	0.45	35.63	45.64
极端梯度提升 XGBoost	0.45	35.62	46.06

基于 Cubist 算法，训练集占比为 80%的有效土层厚度预测模型训练集与验证集的精度评估结果如图 4 所示。可以看出，利用 80%调查点训练的模型可以较好的捕捉有效土层厚度的变异，预测值与实测值之间存在显著的线性关系（R^2=0.59）。由直方图可知，目标值与预测值均为轻微的正偏态，散点图中的点大致沿着蓝色的拟合线分布，即模型存在对部分有效土层厚度 150 cm 以上点位的低估现象，可能与采样时未采集 200 cm 以深的有效土层厚度有关。从 20%的独立验证点来看，校准模型的表现力有较好的迁移性，验证集预测模型的 R^2 为 0.50。

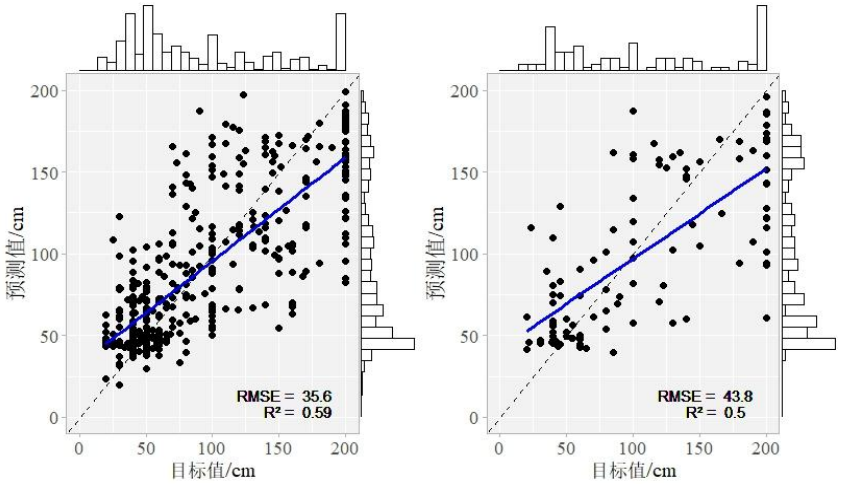

图 4 基于 Cubist 的有效土层厚度模型训练与预测结果

Fig. 4 Model calibration and validation results for effective soil layer thickness prediction based on Cubist regression

有效土层厚度受到多种环境因素的影响，包括气候条件、地形特征、植被覆盖以及人类活动等，这些因素之间存在着复杂的相互作用和反馈机制。因此目前，研究对于有效土层厚度的预测相比其他土壤性质如土壤有机碳、pH 等具有更大的难度：Zhang 等[24]用 RF、支持向量机（Support vector regression，SVR）、神经网络算法（Artificial neural networks，ANN）对有效土层厚度进行预测，结果显示 SVR 能更好预测有效土层厚度（R^2>0.63）；郭俊辉等[25]通过比较不同数字制图方法，结果表明堆叠泛化模型（Stacking）模型表现最佳（R^2=0.47）；Zhang 等[?]采用多元线性回归（Multiple linear regression，MLR）预测有效土层厚度，其 R^2仅为 0.28；Li 等[26]对比 MLR、SVR、RF、XGBoost 预测土层厚度分布图，其 R^2 为 0.43~0.64。

基于 Cubist 模型，将所有变量重新采样至最小分辨率 30 m，并作为出图的基础。采用 Bootstrap

重复计算 50 次得到均值，并计算 90%预测区间，得到如图 5 所示的空间预测结果及其不确定性分布。有效土层厚度的分布在空间上呈现明显异质性，与 DEM 的分布大体相同，在同一流域内，地势较高、容易侵蚀的地方有效土层厚度较薄，而地势低洼的地方土层较厚。有效土层厚度较高值在研究区内分布较广泛，空间上东南普遍高于西北。西北部有效土层厚度分布普遍较低，可能是在山地或岩石裸露地区，山地地势高，土壤容易受到侵蚀，导致土层变薄且不连续；中间和东南区域土层分布较厚，适合农业发展，可能是平原、河谷或沉积区，地势较低，泥沙沉积作用明显，形成厚土层。橙色区域有效土层厚度呈脉状分布，河流带来的沉积物堆积，尤其是沿着主河道和支流的交汇处，促进了土层的加厚；水系分布与有效土层厚度关系紧密。此外，通过 90%预测区间表达的不确定性图能够从空间角度对有效土层厚度建模进行评估。由图 5b 可知，大部分地区的预测区间范围都较小，不确定性较高的区域主要集中在地势较低的平原地区，可能是因为平原地区土壤类型分布较为复杂且空间异质性较高。相比之下，高程较高的山地地区预测结果的不确定性相对较低。这些区域通常土地利用相对单一，自然干扰较少，土壤发育过程较为稳定，从而使得模型预测更加一致。

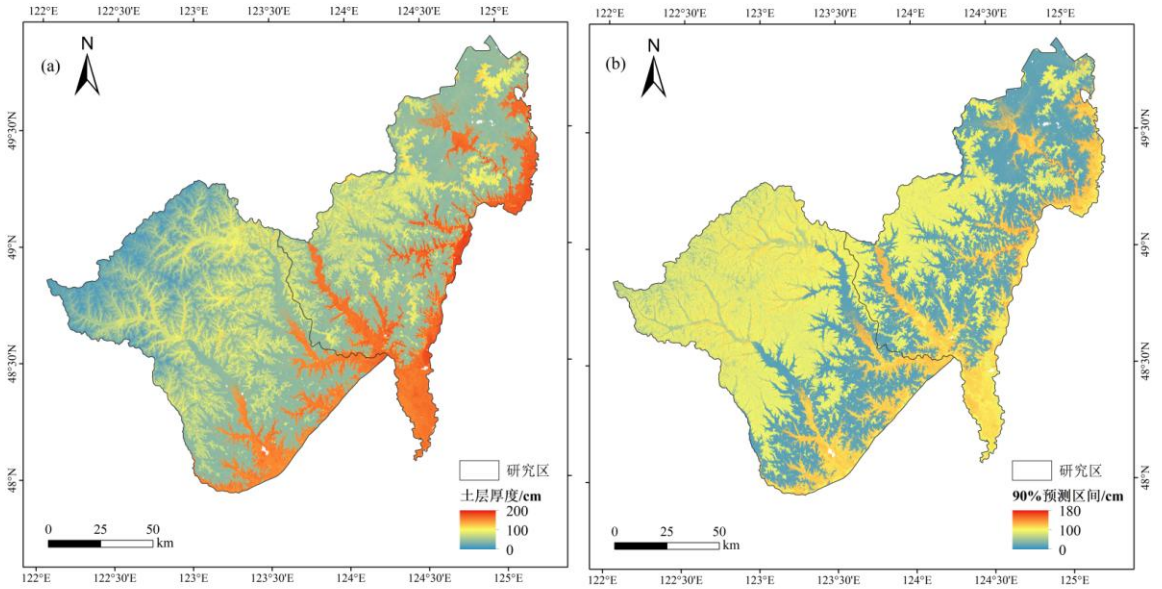

图 5 有效土层厚度空间分布图（a）和有效土层厚度 90%预测区间空间分布图（b）

Fig. 5 Spatial distribution of effective soil thickness (a) and 90% prediction interval spatial distribution of effective soil thickness (b)

计算 34 个环境变量的 SHAP 值和相对重要性排名，结果如图 6 所示。使用 Cubist 建模时，对有效土层厚度建模预测时比较重要的前 5 个环境变量分别为：DEM，MMTmax，MAT，MMTmin 和 TRI。其中 DEM、TRI 为地形因子，其余为气候因子。DEM 与 TRI 对有效土层厚度的影响均为负向。气温在模型中占据主导地位，可能是由于 1 000 m 分辨率的气温数据无法捕捉局部变化，导致重要性被高估。

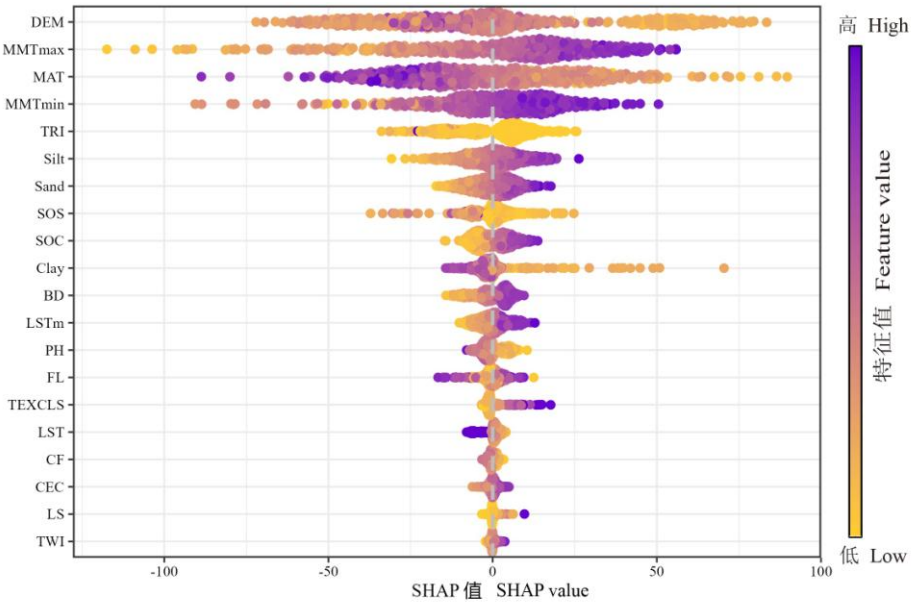

图 6 环境变量 SHAP 因素重要性分析结果

Fig.6 SHAP value importance results of environmental covariates

为了更直观显示各因子对有效土层厚度的影响，将变量重要性排名前九的环境变量特征值与 SHAP 值做部分依赖分析，如图 7 所示。每个因子对有效土层厚度的影响不同。变量重要性最高的 DEM 在建立模型时，Cubist 将其进行分区间预测，SHAP 值大部分时候为负值，与有效土层厚度呈负相关性。MMTmax 和 MMTmin 的分布趋势相同，SHAP 值在较高水平时为正值，较低水平时为负值；MAT 与此相反，在较高水平为负值，较低水平为负值。月均气温极值与有效土层厚度呈正相关性，原因是较高的温度会促进有机质的产生与积累，有利于植被的根系生长[27]，从而增加有效土层厚度。当月均气温在 0℃上下浮动时，反复的冻融作用会使细小颗粒向下迁移，形成更厚的有效土层[28]。

TRI 值分布较集中，在大多情况下 SHAP 值其负值，与有效土层厚度呈负相关性。粉粒、砂粒和 SOC 的分布趋势相同，且呈较好的线性，在较高水平时为正值，促进有效土层厚度的积累，较低水平时为负值，不利于有效土层厚度的积累。SOS 总体 SHAP 值较小且为负值，与有效土层厚度为负相关性。

地形因子 DEM 被识别为影响土壤厚度预测的首要关键因子，这一发现与先前研究的结论相吻合：陈玉蓝等[29]通过对比不同机器学习方法，发现控制四川省土层厚度变化的主要因子为地形因子；Zhang 等[24]发现影响喀斯特流域土层厚度的最重要因素是坡位；Zahedi 等[30]研究也发现坡度和坡向是预测土壤厚度最重要的解释变量之一。除地形因子外，气候因子在预测有效土层厚度方面也占据了相对显著的位置，这与刘凯等[31]研究发现一致。气候作为五大成土母质之一，温度通过影响土壤有机碳的分解、土壤结构的紧实度以及土壤温度的垂直分布，对有效土层厚度也会产生重要影响。

而在目前已有的研究中，其不作为主要的解释变量，可能是由于尺度导致研究区气候相对均一。

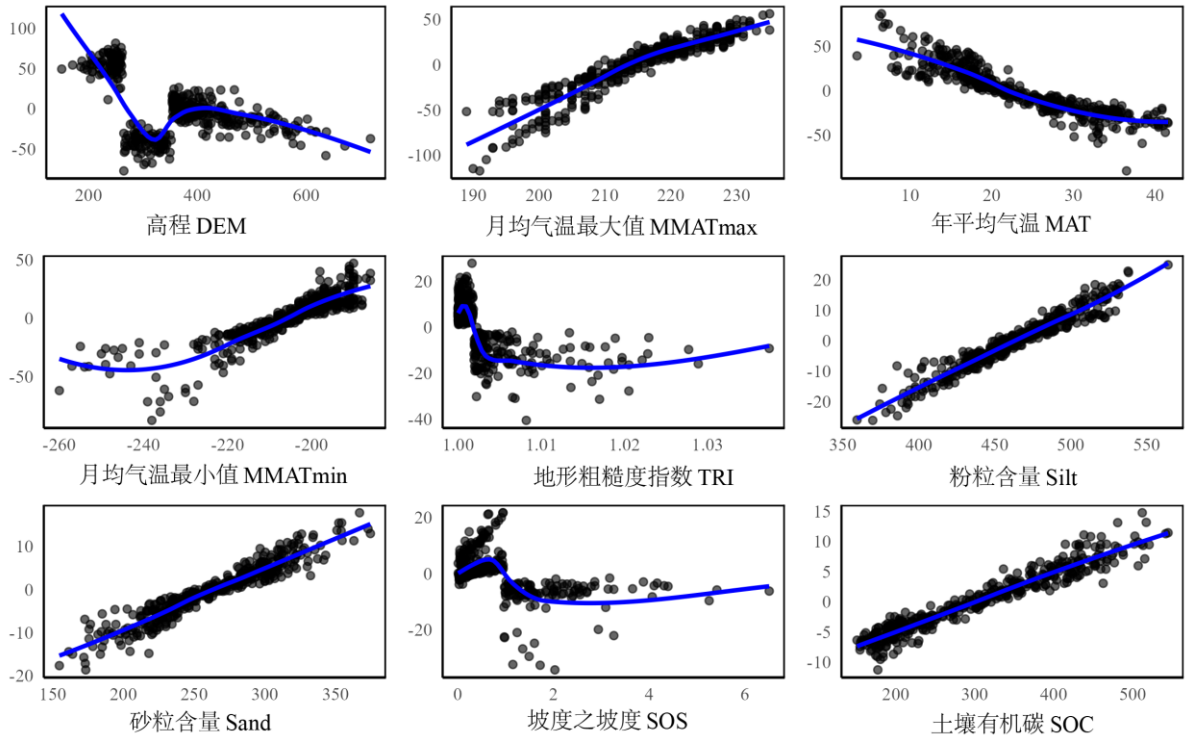

图 7 环境变量部分依赖图

Fig.7 Partial dependence plots of environmental covariates

2.3 有效土层厚度与土地利用、地表基质的关系

为更准确地评估各类土地的生产力和适宜性，进而为土地的合理利用和规划提供科学依据，根据土地利用类型，制作不同土地利用类型的有效土层厚度箱线图。由图 8 可知，不同土地利用类型有效土层厚度平均值顺序为湿地＞裸地＞耕地＞草地＞林地，有效土层厚度分布范围顺序为草地＞耕地＞林地＞裸地＞湿地。湿地分布集中，有效土层厚度最高，湿地环境稳定，影响因子单一，受到外界干扰小；裸地大部分集中在 100~150 cm 之间，其因为缺少人类活动干扰，有效土层厚度也较高；耕地分布范围较广，集中于中低有效土层厚度，耕地受人为耕作与干扰较多；草地有效土层厚度分布范围最广，植被覆盖能力相对较弱，有效土层厚度也较低；林地有效土层厚度最低，主要是其多分布于山区或坡度较大的区域，地形陡峭，流失严重。

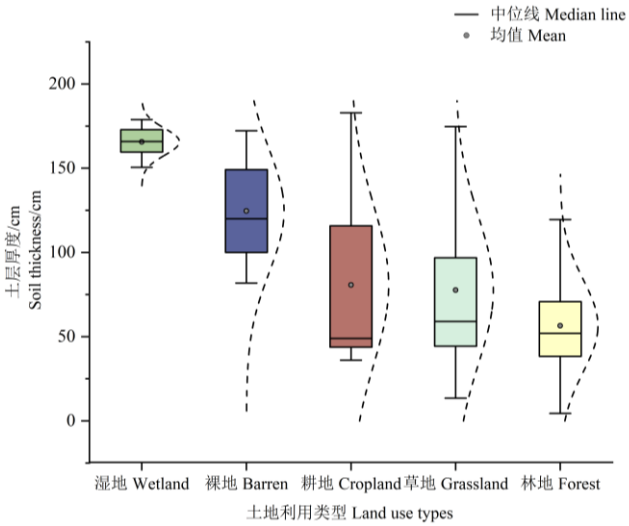

图 8 不同土地利用类型的有效土层厚度箱线图

Fig.8 Box plots of effective soil thickness for different land use types

不同土地利用类型有效土层厚度平均值顺序为湿地＞裸地＞耕地，这与艾晓军等[32]研究结果一致。湿地分布较均一，这可能与湿地环境较稳定有关，其均值最高，长期水分饱和可能不利于土地的侵蚀作用；裸地通常位于平坦或缓坡区域，植被覆盖少且侵蚀较少，沉积物易堆积；耕地有效土层厚度较低，反映了其受到人为耕作导致的土壤侵蚀与压实作用[33]，低于自然沉积的裸地；一些坡耕地的水土流失也是导致耕土层变薄的重要原因；草地微生物及有机质积累较少，且缺乏树木保护，导致其更容易受到风蚀和水蚀的影响，且人为的放牧也会影响土壤退化；林地多分布于山区，土壤侵蚀严重，且由于较少人工干预，恢复速度较低。

为更好地理解地表基质对自然资源的支撑作用，根据地表基质分区，制作不同地表基质分区的有效土层厚度箱线图。由图 9 可知，在大兴安岭隆起带中低山地形地貌分区内，由成因类型所划分的地表基质区有效土层厚度均值顺序为：坡积地表基质区＞坡洪积地表基质区＞残坡积基质区；在同大兴安岭隆起带丘陵地形地貌分区内，由成因类型所划分的地表基质区有效土层厚度均值顺序为：坡洪积地表基质区＞坡积地表基质区＞残坡积基质区；总体而言，有效土层厚度均值顺序为；山间谷地地表基质区 > 中低山地表基质区 > 丘陵地表基质区；针对冲洪积地表基质区而言，松嫩平原沉降带阶地漫滩地表基质区＞大兴安岭隆起带山间谷地地表基质区。

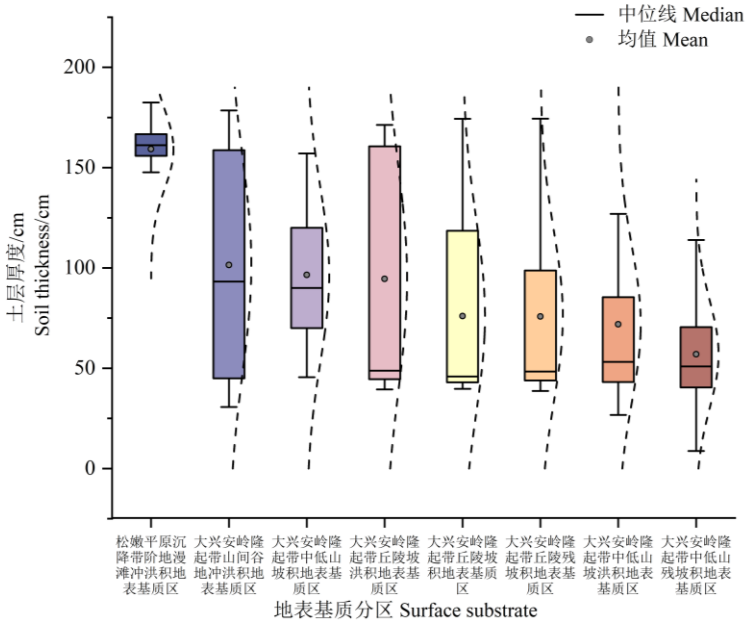

图 9 不同地表基质分区的有效土层厚度箱线图

Fig.9 Box plots of effective soil thickness for different surface substrate

不同地表基质分区对有效土层厚度的影响具有显著的规律性，地表基质作为五大成土因素之一，对有效土层厚度的形成与发展起着绝对性作用[34]。坡积土则是由于重力作用和雨水冲刷，土粒在坡面上移动并在坡脚堆积形成；由于搬运距离短，层理不明显，且可能持续从坡上冲刷下来，有效土层厚度较厚。坡洪积土多由大雨或洪水冲刷山坡时沉积形成的，通常在山前平原堆积；但由于洪水的冲刷和搬运，坡洪积土的层理和分选性相对较好[35]；多分布于洪水冲刷后形成的扇形地貌[36]，厚度因水流强度和沉积物类型而异，一般较残坡积土厚。残坡积土由于其在岩石风化后未被搬运[36]，以残积物为主，搬运较少，导致土层发育不够厚；且通常处于较高坡度，侵蚀作用较强。

本研究系统整合地形、气候、遥感及土壤多源数据，建立了典型黑土区有效土层厚度的空间预测模型，并从地表基质分区构型的视角分析了不同地表基底条件对土壤演育的影响。受限于高质量的耕作方式、土壤机械压实状况、土壤侵蚀空间显式数据的缺失，本研究难以实现多形式人类活动对土层厚度的深入刻画，这也一定程度上构成了预测模型的误差来源。此外，基于 SHAP 值的影响因子重要性分析解析了地形、气候、遥感等关键因子的独立贡献度，其全局特征重要性排序为土壤-景观模型提供了清晰的解释框架，然而对于解释因子彼此之间的非线性交互效应，传统 SHAP 方法仍存在理论局限性。根据本研究提供的有效土层厚度预测结果，发现有效土层厚度分布与土地利用类型和地表基质分区关系密切，可持续土地资源管理与保护措施应充分考虑地形和成土母质的不同，采取差异化的土地利用策略。根据宜林则林、宜耕则耕的原则，推广适宜的农业技术。在对地观测卫星大数据集成应用蓬勃发展的背景下，未来研究可开展面向保护性耕作、土壤物理结构定量评价、土壤侵蚀遥感反演等方面的研究，生成高质量土壤基准信息和管理措施数据，深入评估多形

式人类干扰下有效土层厚度的时空变异特征。

3 结 论

本文利用蒙东黑土区地表基质调查基础数据，构建了以地形、气候、遥感和土壤因子为自变量的有效土层厚度 Cubist 预测模型，预测性能良好，可有效捕捉有效土层厚度的空间分布特征。SHAP 因素重要性分析显示地形特征是决定有效土层厚度的首要因子，同一流域内海拔较高的区域受土壤侵蚀等地表外动力过程影响而土层较薄，而地势低洼地带土层较厚；气候因子也对有效土层厚度的空间分异具有显著影响。由于土壤形成与演育过程的控制作用，不同地表基质分区有效土层厚度平均值坡积、坡洪积均大于残坡积土；林地由于多分布于山区或坡度较大的区域，有效土层厚度最低。

参考文献（References）

[1] Yi C, Li D C, Zhang G L, et al. Criteria for partition of soil thickness and case studies [J]. Acta Pedologica Sinica, 2015,52(1): 220-227. [易晨,李德成,张甘霖,等.土壤厚度的划分标准与案例研究[J].土壤学报,2015,52(1):220-227.]

[2] Li Z, Zhao Y J, Song H Y, et al. Effects of karst soil thickness heterogeneity on the leaf anatomical structure and photosynthetic traits of two grasses under different water treatments[J].Acta Ecologica Sinica,2018,38(2):721-732.[李周,赵雅洁,宋海燕,等.不同水分处理下喀斯特土层厚度异质性对两种草本叶片解剖结构和光合特性的影响[J].生态学报,2018,38(2):721-732.]

[3] Li L, Fu X, Cui J, et al. Soil layer identification based on envelope detector and STFT spectrum analysis of ground penetrating radar signals[J]. Journal of Geo-information Science, 2020,22(2):316-327. [李俐,付雪,崔佳,等.基于包络检波和 STFT 谱分析的探地雷达土壤分层信息识别[J].地球信息科学学报,2020,22(2):316-327.]

[4] Liu H, Wei Y X. Influence of soil erosion thickness on soil productivity of black soil and its evaluation[J]. Transactions of the Chinese Society of Agricultural Engineering, 2014,30(20):288-296. [刘慧,魏永霞.黑土区土壤侵蚀厚度对土地生产力的影响及其评价[J].农业工程学报,2014,30(20): 288-296.]

[5] Xu X Z, Xu Y, Chen S C, et al. Soil loss and conservation in the black soil region of Northeast China: A retrospective study[J]. Environmental Science, 2010.13(8):793-800.

[6] Lei G P, Dai L, Song G. Evaluation of soil ecological environment quality of typical black soils in Heilongjiang Province[J]. Transactions of the Chinese Society of Agricultural Engineering, 2009,25(7):243-248. [雷国平,代路,宋戈.黑龙江省典型黑土区土壤生态环境质量评价[J].农业工程学报,2009,25(7): 243-248.]

[7] Liu Y J, Wu H Q, Li R, et al. Study on the distribution characteristics of effective soil thickness of cultivated land in Tacheng City based on GIS[J]. Geospatial Information, 2023,21(6): 58-62. [刘宇杰,武红旗,李润,等.基于 GIS 的塔城市耕地有效土层厚度分布特征研究[J].地理空间信息, 2023, 21(6): 58-62.]

[8] Liu K, Wei M H, Dai H M, et al. Spatiotemporal variation of black soil layer thickness in black soil region of northeast China[J]. Geology & Resources, 2022,31(3): 434-442, 394. [刘凯,魏明辉,戴慧敏,等.东北黑土区黑土层厚度的时空变化[J].地质与资源, 2022, 31(3): 434-442, 394.]

[9] Mcbratney A B, Mendonca Santos M L, Minasny B. On digital soil mapping[J]. Geoderma, 2003, 117(1/2): 3-52.

[10] Wang S Q, Zhu S L, Zhou C H. Characteristics of spatial variability of soil thickness in China[J], Geographical Research,2001,20(2):161-169. [王绍强,朱松丽,周成虎. 中国土壤土层厚度的空间变异性特征[J]. 地理研究, 2001, 20(2): 161-169.]

[11] Yamashita N. Yasuhiro O, Junko I, et al. National-scale mapping of soil-thickness probability in hilly and

mountainous areas of Japan using legacy and modern soil survey[J]. Geoderma, 2024, 446:116896.

[12] Wang W, Zhao Y, Zhang T, et al. Regional soil thickness mapping based on stratified sampling of optimally selected covariates[J]. Geoderma, 2021, 400: 115092.

[13] Chen S, Richer-de-Forges A C, Leatitia M V, et al. Digital mapping of the soil thickness of loess deposits over a calcareous bedrock in central France[J]. Catena, 2021, 198: 105062.

[14] Hou H X, Li R H, Li J H, et al. Current situation, progress and prospect of ground substrate survey[J]. Natural Resources Information, 2024(11): 1-8. [侯红星, 李瑞红, 李俊华, 等. 地表基质调查：现状、进展与展望[J]. 自然资源情报, 2024(11): 1-8.]

[15] Su R H, Guo E L, Wang Y F, et al. Extreme climate changes in the Inner Mongolia and their impacts on vegetation dynamics during 1982—2020[J]. Acta Ecologica Sinica, 2023,43(1):419-431. [苏日罕,郭恩亮,王永芳,等.1982—2020 年内蒙古地区极端气候变化及其对植被的影响[J].生态学报,2023,43(1):419-431.]

[16] Cui W F, Huang Z G. Analysis on productive potential of climate resources in dry-farming land in Hulunbeier[J]. Agricultural Research in the Arid Areas, 2005,23(3): 122-127. [崔文芳,黄振刚. 呼伦贝尔岭东南旱作丘陵区气候资源潜力分析[J]. 干旱地区农业研究, 2005,23(3): 122-127.]

[17] Quinlan J R. Learning with continuous classes[C]//Proceedings of Australian Joint Conference on Artificial Intelligence. Hobart, 1992: 343-348.

[18] Cortes C, Vapnik V. Support-vector networks[J]. Machine Learning, 1995,20:273-297.

[19] Breiman L. Random forests[J]. Machine Learning, 2001, 45(1): 5-32.

[20] Chen T, 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.

[21] Armstrong J S, Collopy F. Error measures for generalizing about forecasting methods: Empirical comparisons[J]. International Journal of Forecasting, 1992, 8(1): 69-80.

[22] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[J]. Proceedings of the 31st International Conference on Neural Information Processing Systems,2017,10:4768-4777.

[23] Kennard R W, Stone L A. Computer aided design of experiments[J]. Technometrics, 1969, 11(1): 137-148.

[24] Zhang Y, Xu X, Li Z, et al. Modelling soil thickness using environmental attributes in karst watersheds[J]. Catena, 2022, 212: 106053.

[25] Guo J H, Liu F, Xu S X, et al. Comparison of digital mapping methods for the thickness of black soil layer of cultivated land in typical blacksoil area of Songnen Plain[J]. Journal of Geo-information Science, 2024,26(6):1452-1468. [郭俊辉,刘峰,徐胜祥,等.松嫩典型黑土区耕地黑土层厚度数字制图方法比较[J]. 地球信息科学学报, 2024, 26(6): 1452-1468.]

[26] Li X, Luo J, Jin X, et al. Improving soil thickness estimations based on multiple environmental variables with stacking ensemble methods[J]. Remote Sensing, 2020, 12(21): 3609.

[27] Zhang Y, Hartemink A E, Vanwalleghem T, et al. Climate and land use changes explain variation in the A horizon and soil thickness in the United States[J]. Communications Earth & Environment, 2024, 5(1): 129.

[28] Fan G S, Jia H J. Experimental study on the freezing and thawing features and reduction infiltration characteristics of seasonal freezjng and thawing soils[J]. Acta Pedologica Sinica, 2000,37(1): 24-32. [樊贵盛,贾宏骥. 季节性冻融土壤的冻融特点和减渗特性的研究[J]. 土壤学报, 2000,37(1): 24-32.]

[29] Chen Y L, Liang T B, Zhang Y L, et al. Spatial prediction of soil thicknesses in Sichuan Province based on feature-ensemble learning[J]. Soils, 2023, 55(4): 894-902. [陈玉蓝,梁太波,张艳玲,等.基于特征集成学习的四川省土壤厚度预测[J]. 土壤, 2023, 55(4): 894-902.]

[30] Zahedi S, Shahedi K, Rawshan M H, et al. Soil depth modelling using terrain analysis and satellite imagery: The case study of Qeshlaq mountainous watershed (Kurdistan, Iran)[J]. Journal of Agricultural Engineering, 2017,48(3):167.

[31] Liu K, Dai H M, Liu G D, et al. Organic carbon content-baesd prediction and influencing factors of black soil layer thicknesses[J]. Geophysical and Geochemical Exploration, 2024, 48(5): 1368-1376. [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.]

[32] Ai X J, Chen Z S, Geng G S, et al. Distribution patterns and influencing factors of effective soil layers in the surface matrix of black soil in Liaoyang-Dandong Area—Taking Fengcheng City as an example[J]. Journal of Hebei Agricultural Sciences, 2023, 27(3): 54-59,65. [艾晓军,陈占生,耿国帅,等.辽阳—丹东地区黑土地地表基质有效土层分布规律及影响因素——以凤城市为例[J].河北农业科学,2023,27(3):54-59,65.]

[33] Shi D M, Jin H F, Jiang G Y. Degradation effect of soil erosion on tillage-layer quality of slope farmland and its evaluation trend[J]. Transactions of the Chinese Society of Agricultural Engineering, 2019,35(18):118-126. [史东梅,金慧芳,蒋光毅.土壤侵蚀对坡耕地耕层质量退化作用及其评价趋势展望[J].农业工程学报, 2019,35(18):118-126.]

[34] Liu J T, Zhao W, Liu Y. Modelling soil thickness evolution: Advancements and challenges[J]. Acta Pedologica Sinica,2024,61(2):319-330. [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.]

[35] Alexander D, Coppola L. Structural geology and the dissection of alluvial fan sediments by mass movement: An example from the southern Italian apennines[J]. Geomorphology, 1989, 2(4): 341-361.

[36] Dong H, Li Z F, Fu H L, et al. A study of the spatial variability of physical characteristics and water-holding characteristics of the residual diluvia gravel[J].Hydrogeology & Engineering Geology, 2015, 42(3): 71-78. [董辉, 李智飞, 傅鹤林,等. 残坡积碎石物理特征与持水性的空间变异性研究[J]. 水文地质工程地质, 2015, 42(3): 71-78.]

（责任编辑：陈德明）