

DOI: 10.11766/trxb202406180242

CSTR: 32215.14.trxb202406180242

王晓盼, 王昌昆, 孙海军, 郭志英, 刘杰, 高磊, 马海艺, 袁自然, 姚成硕, 潘贤章. 基于随机森林的黑土地尺度耕地土壤容重传递函数构建[J]. 土壤学报, 2025, 62 (4): 970–982.

WANG Xiaopan, WANG Changkun, SUN Haijun, GUO Zhiying, LIU Jie, GAO Lei, MA Haiyi, YUAN Ziran, YAO Chengshuo, PAN Xianzhang. Construction of Pedotransfer Function for Predicting Soil Bulk Density in Cultivated Land of Northeast China Using Random Forest[J]. Acta Pedologica Sinica, 2025, 62 (4): 970–982.

## 基于随机森林的黑土地尺度耕地土壤容重传递函数构建<sup>\*</sup>

王晓盼<sup>1, 2</sup>, 王昌昆<sup>2, 3†</sup>, 孙海军<sup>1</sup>, 郭志英<sup>2, 3</sup>, 刘杰<sup>2, 3</sup>, 高磊<sup>2, 3</sup>,  
马海艺<sup>2, 3</sup>, 袁自然<sup>2, 3</sup>, 姚成硕<sup>2, 3</sup>, 潘贤章<sup>2, 3</sup>

(1. 南京林业大学 南方现代林业协同创新中心, 南京 210037; 2. 土壤与农业可持续发展国家重点实验室(中国科学院南京土壤研究所),  
南京 210008; 3. 中国科学院大学 现代农业科学学院, 北京 100049)

**摘要:** 土壤容重 (Bulk Density, BD) 信息对了解东北黑土地耕地土壤物理状况及推进黑土地利用与保护至关重要。传统环刀方法在容重采样环节耗时耗力, 难以获得大空间尺度容重数据。土壤传递函数 (Pedotransfer function, PTF) 可以利用容易获取的土壤变量实现容重信息估算, 然而当前缺少针对东北黑土地区域尺度的 PTF 模型研究, 而且用于 PTF 模型构建的潜在土壤属性变量的重要性尚待揭示。本研究针对东北黑土地区域尺度, 基于实测容重数据, 使用随机森林 (Random forest, RF) 机器学习方法, 以土壤有机质 (Soil organic matter, SOM)、含水量 (Moisture content, MC) 和土壤质地相关变量作为输入变量, 构建面向大尺度耕地土壤容重预测的 PTF 模型, 并分析土壤属性变量的重要性, 同时评估国内外已发表 PTF 模型在东北黑土地耕地容重预测中的适用性。结果显示: 已发表 PTF 模型在表层、亚表层以及全部土壤容重预测中的最优  $R^2$  分别为 0.17、0.22 和 0.26, 均方根误差 (Root Mean Squared Errors, RMSE) 分别为 0.16、0.13 和  $0.15 \text{ g} \cdot \text{cm}^{-3}$ ; 本研究中基于 RF 方法的 PTF 模型在表层、亚表层以及全部土壤中的最优预测  $R^2$  分别为 0.22、0.45 和 0.37, RMSE 分别为 0.16、0.11 和  $0.14 \text{ g} \cdot \text{cm}^{-3}$ 。研究表明: 已发表 PTF 模型的容重预测精度较低, 难以用于东北黑土地区域尺度的容重预测; 基于 RF 开发的模型具有预测东北黑土地容重的潜力; 有机质是预测东北黑土地容重的最重要变量, 其次是含水量, 土壤质地相关变量影响较小。

**关键词:** 东北黑土地; 土壤容重; 土壤传递函数; 随机森林; 变量重要性

中图分类号: S152.4 文献标志码: A

## Construction of Pedotransfer Function for Predicting Soil Bulk Density in Cultivated Land of Northeast China Using Random Forest

WANG Xiaopan<sup>1, 2</sup>, WANG Changkun<sup>2, 3†</sup>, SUN Haijun<sup>1</sup>, GUO Zhiying<sup>2, 3</sup>, LIU Jie<sup>2, 3</sup>, GAO Lei<sup>2, 3</sup>, MA Haiyi<sup>2, 3</sup>,

\* 国家重点研发计划项目 (2021YFD1500102)、中国科学院战略性先导科技专项 (XDA28010102) 和国家自然科学基金项目 (42107145) 资助  
Supported by the National Key R&D Program of China (No. 2021YFD1500102), Strategic Priority Research Program of the Chinese Academy of Sciences (No. XDA28010102), and the National Natural Science Foundation of China (No. 42107145)

† 通讯作者 Corresponding author, E-mail: ckwang@issas.ac.cn

作者简介: 王晓盼 (1999—), 女, 甘肃武山人, 硕士研究生, 主要从事土壤物理特性光谱分析研究。E-mail: wangxp@njfu.edu.cn

收稿日期: 2024-06-18; 收到修改稿日期: 2024-10-22; 网络首发日期 (www.cnki.net): 2024-11-22

YUAN Ziran<sup>2,3</sup>, YAO Chengshuo<sup>2,3</sup>, PAN Xianzhang<sup>2,3</sup>

(1. Co-Innovation Center for Sustainable Forestry in Southern China, Nanjing Forestry University, Nanjing 210037, China; 2. State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China; 3. College of Advanced Agricultural Sciences, University of Chinese Academy of Sciences, Beijing 100049, China)

**Abstract:** 【Objective】Soil bulk density (BD) is crucial for understanding the physical condition of black soil in cultivated land of Northeast China and advancing its utilization and protection. The traditional cutting ring method for determining BD is time-consuming and laborious, making the evaluation of BD on a large spatial scale difficult. The pedo-transfer function (PTF) can estimate BD information using readily available soil variables. However, there is currently a lack of research on PTF models specifically targeting the whole of Northeast China, and the importance of potential soil attribute variables for PTF model construction remains to be elucidated. 【Method】By incorporating soil organic matter (SOM), moisture content (MC), and soil texture-related variables as input features, we constructed PTF models capable of predicting BD on a large scale. Furthermore, we delved into the significance of these soil attribute variables in the constructed PTF models. Additionally, we assessed the suitability of existing published PTF models for BD prediction in the black soil of Northeast China. 【Result】The optimal predicted  $R^2$  values of published PTF were 0.17, 0.22, and 0.26, respectively, for the topsoil, subsoil, and all soil samples, and Root Mean Squared Errors (RMSE) were 0.16, 0.13, and 0.15 g·cm<sup>-3</sup>, respectively. Also, the optimal predicted  $R^2$  values of PTF for the topsoil, subsoil, and all soil samples based on the proposed RF method were 0.22, 0.45, and 0.37, respectively, while the RMSE values were 0.16, 0.11, and 0.14 g·cm<sup>-3</sup>, respectively. 【Conclusion】The published PTF models had low BD prediction accuracy and were difficult to use for BD prediction on the scale of black soil in Northeast China whereas the PTF model constructed in this study has the potential to predict BD on the scale of black soil in Northeast China. Among the variables, SOM was the most important variable for predicting BD in the black soil of Northeast China, followed by MC, while soil texture-related variables had a relatively small impact.

**Key words:** Black soil in Northeast China; Soil bulk density; Pedotransfer function; Random forest; Variable importance

土壤容重 ( Bulk density, BD ) 是指未扰动状态下单位体积的干土质量, 是土壤最重要的物理属性之一。它不仅影响土壤水分和养分的有效性, 还间接影响土壤质量和生产力<sup>[1]</sup>。容重反映了土壤的结构功能, 包括作物的机械支撑、土壤的水分循环和通气性等, 在计算土壤水力特性以及评估土壤压实性能等方面发挥着重要作用<sup>[2-3]</sup>。此外, 容重还是估算土壤碳储量的关键参数, 对分析土壤在全球碳循环中的作用和土壤储存碳的潜力有很大影响<sup>[4-5]</sup>。东北黑土地以高有机质和高肥力而著称, 是中国的粮仓, 在保障国家粮食安全中具有举足轻重的地位<sup>[6]</sup>。近年来, 由于过渡垦殖和长期高强度利用, 东北黑土地耕地质量不断下降, 存在黑土层变薄、有机质含量锐减、以及耕层变浅变硬等问题, 严重影响了粮食生产的可持续性<sup>[7-8]</sup>。获取容重数据对于了解东北黑土地耕地土壤物理特性、评估区域土壤碳储量, 推进黑土地利用与保护至关重要。

目前, 国内外土壤容重测定的最常用方法为环刀法<sup>[9]</sup>。但对于整个东北黑土地这一类较大尺度

耕地土壤容重数据的获取, 环刀法在样品采集环节十分耗时、耗力, 导致当前全球许多土壤数据库中缺乏直接测量获得的容重数据。以我国为例, 在第二次全国土壤普查数据中, 容重数据的缺失问题尤为突出, 这在一定程度上影响了对土壤资源全面、准确评估的能力<sup>[10]</sup>。因此, 为满足较大尺度容重数据需求, 亟需寻找一种高效精确的容重获取方法。

20世纪90年代末以来, 土壤传递函数 ( Pedotransfer function, PTF ) 的发展为容重获取提供了新思路。PTF 是一种利用已知土壤信息来估算难以直接测量或测量成本较高的土壤属性的建模方法<sup>[11]</sup>。近年来, 基于其他土壤属性建立了众多容重预测的 PTF 模型<sup>[12-13]</sup>。这些模型中使用的变量主要包括土壤有机碳 ( Soil organic carbon, SOC ) 或有机质 ( Soil organic matter, SOM ) 和土壤质地<sup>[14-15]</sup>, 部分研究还包括土壤深度<sup>[16]</sup>、含水量 ( Moisture content, MC ) 、阳离子交换量<sup>[17]</sup>、pH<sup>[18]</sup>等。但基于 PTF 的容重预测精度在不同土壤类型及区域间存在较大差异<sup>[19-20]</sup>, 因此

需要针对不同区域构建专有的容重预测 PTF 模型。当前在东北黑土地区域土壤容重预测 PTF 模型研究还相对较少。王思楚等<sup>[21]</sup>在松嫩平原黑土区, 使用有机碳、土壤质地和深度作为预测变量, 建立了容重预测 PTF 模型; 韩光中和李秀芝<sup>[22]</sup>在黑龙江省使用有机质和土壤深度, 通过校正已发表的容重预测 PTF 模型开发了针对黑龙江省域尺度的 PTF 模型。然而, 这些研究并未针对整个东北黑土地区域尺度开发有效的容重预测 PTF 模型, 而且并未充分揭示容重预测 PTF 模型中土壤变量的重要性。

多元线性回归方法在构建 PTF 模型中应用最早<sup>[23-24]</sup>。通常, 使用线性回归方法构建的 PTF 模型结构简单, 且变量重要性易于量化, 因此被广泛使用<sup>[25]</sup>。但由于容重与其他土壤属性之间有可能存在非线性关系, 因此其预测精度难以得到有效保障。为提高容重预测 PTF 模型的精度, 指数、对数、幂函数等模型也被大量采用, 这些模型对于容重与预测土壤属性之间具有非线性关系的情形时效果较好<sup>[14]</sup>。近年来, 随着土壤数据规模的增加以及对容重预测精度的更高要求, 机器学习方法, 如随机森林 (Random forest, RF)<sup>[26]</sup>、人工神经网络<sup>[27]</sup>等被广泛用于 PTF 建模。与简单多元线性回归模型或者指数等模型方法相比, 机器学习方法可以更好处理自变量和因变量之间的非线性和复杂关系, 从而改善 PTF 模型的预测效果<sup>[28]</sup>。RF 是一种基于决策树的集成学习方法, 能够处理非线性问题, 而且能够量化模型中各因子的相对重要性, 从而能够实现较高精度容重预测 PTF 模型的构建以及揭示预测变量的重要性<sup>[29]</sup>。然而, 基于 RF 针对东北黑土地区域尺度 PTF 模型的土壤容重预测潜力尚不清楚。

因此, 基于当前针对东北黑土地区域尺度耕地土壤容重预测 PTF 研究的不足以及机器学习方法在 PTF 模型构建中的优势, 本研究使用 RF 方法, 针对东北黑土地区域尺度的耕地土壤, 利用大规模实际采样、测量的容重数据, 结合多种潜在土壤属性, 构建容重预测 PTF 模型。目的在于: (1) 建立适用于东北黑土地区域尺度耕地土壤容重预测的高精度 PTF 模型; (2) 揭示东北黑土地区域尺度耕地土壤容重预测 PTF 模型中的土壤属性变量重要性。为准确估算中国东北黑土地耕地土壤容重提供方法支持, 支撑区域碳储量估算等研究。

## 1 材料与方法

### 1.1 研究区概况

中国东北黑土地涉及黑龙江省、吉林省、辽宁省全境、以及内蒙古自治区东部 ( $38^{\circ}43' \text{--} 53^{\circ}33' \text{N}$ ,  $118^{\circ}35' \text{--} 135^{\circ}05' \text{E}$ )<sup>[30]</sup> (图 1), 共 246 个县, 总面积 109 万  $\text{km}^2$ <sup>[31]</sup>。东北黑土地地形三面环山, 西有大兴安岭, 西北有小兴安岭, 东有长白山; 松嫩平原、松辽平原及辽河平原分布在群山之中<sup>[31]</sup>。东北黑土地属温带大陆性季风气候, 夏季炎热多雨, 冬季寒冷干燥<sup>[32]</sup>。整个地区的年平均气温为  $-7$  至  $11^{\circ}\text{C}$ , 大部分地区的年降水量为 450 至 850  $\text{mm}$ <sup>[33]</sup>。东北黑土地主要有 6 种土壤类型: 黑土、黑钙土、暗棕壤、棕壤、白浆土和草甸土<sup>[34]</sup>, 主要种植玉米、大豆和水稻<sup>[8]</sup>。

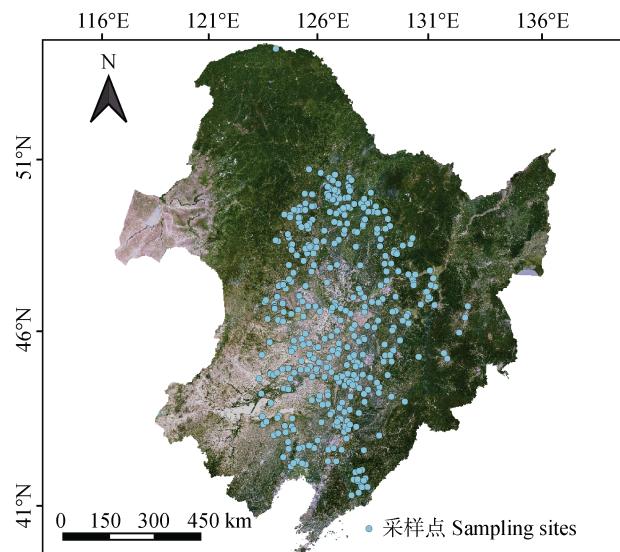


图 1 研究区及耕地土壤采样点分布

Fig. 1 Study area and distribution of sampling points in cultivated land

### 1.2 耕地土壤样品采集与实验室分析

土壤样品采集于 2022 年 7 月至 9 月, 包括表层 ( $0 \text{--} 20 \text{ cm}$ ) 和亚表层 ( $20 \text{--} 40 \text{ cm}$ ) 样品。其中用于有机质和土壤质地测量的土壤样品由 5 m 范围内 3~5 个子样混合而成, 用于容重和含水量测量的土壤样品采用体积为  $100 \text{ cm}^3$  的环刀采集。样品采集完成后, 贴上标签, 密封并送到实验室进行后续分析和测试。

其中容重和含水量采用烘干法测定, 有机质采

用重铬酸钾氧化—外加热法测定<sup>[35]</sup>，土壤质地采用激光粒度仪测定<sup>[36]</sup>。土壤质地参考美国农业部(USDA)的等级划分标准分级，即砂粒(Sand: 2000~50 μm)、粉粒(Silt: 50~2 μm)、黏粒(Clay: <2 μm)。需要说明的是，由于运输和实验室分析过程中出现了样品损坏等情况，经过异常值去除后，本研究共有容重、有机质、含水量、砂粒、粉粒和黏粒6种土壤属性均完整存在的数据993组，其中表层497组，亚表层496组，包括549个点位(如图1所示)。

### 1.3 数据集的划分

使用Kennard-Stone(KS)算法<sup>[37]</sup>选出60%的土壤样本作为建模集来构建模型，剩下40%的土壤样本作为测试集。KS算法是一种迭代样本选择算法，能够在数据空间中选择有代表性的子集，保证建模集和验证集样本之间数据分布的一致性。首先，

该算法依据输入样本的变量主成分在多元空间中选取彼此距离最远的一对样本点，然后依次选择使已选样本点之间的距离度量最大的样本点，重复该过程，直至达到样本数量要求。本研究中KS算法使用R语言“prospectr”包执行，所使用的距离度量为马氏距离。

### 1.4 已发表土壤传递函数

为验证国内外已发表的PTF模型在东北黑土地耕地容重预测中的适宜性，本研究从已发表的容重预测PTF模型中，选择了7个代表性PTF模型(表1)，包括2个黑土地范围内的模型，2个中国国家尺度的模型，以及美国加利福尼亚州、巴西亚马逊河流域2个区域尺度和1个英国国家尺度的模型。这些PTF模型的开发区域不同，但均具有较好的容重预测性能。

表1 已发表土壤传递函数的基本信息

Table 1 Basic information on published pedo-transfer functions (PTFs)

编号 ID	研究区域 Study area	样本量 Samples	PTF	参考文献 Reference
PTF1	松嫩黑土区	265	BD=exp ( 0.341-0.084SOC+0.002Sand )	[21]
PTF2	黑龙江省	408	BD=0.167×1.526/ ( 1.526SOM+0.167 ( 1-SOM ) )	[22]
PTF3	中国	1566	BD=exp ( 0.5379-0.0653SOM <sup>0.5</sup> )	[38]
PTF4	中国	4765	BD=1.3770exp ( -0.0048SOC )	[39]
PTF5	巴西亚马逊河流域	323	BD=1.398-0.042SOC-0.0047Clay	[40]
PTF6	美国加利福尼亚州	386	BD=1.66-0.308SOC <sup>0.5</sup>	[41]
PTF7	英国	333	BD=0.80806+0.823844exp ( -0.27993SOC ) + 0.0014065Sand-0.0010299Clay	[42]

### 1.5 随机森林模型构建

本研究采用RF构建东北黑土地土壤容重预测PTF模型。RF是一种基于决策树的机器学习算法<sup>[43]</sup>，被广泛用于回归问题，它通过使用多个决策树对数据进行建模，并将它们的预测结果进行集成，从而能够提高模型性能和稳定性。决策树个数、随机选择变量数以及节点数对RF模型精度具有重要影响，本研究针对这3个参数，采用5-fold交叉验证，通过R语言“caret”包进行参数调优。为验证不同输入土壤属性对容重PTF模型预测精度的影响，本研究以不同土壤属性为自变量，包括含水量、有机质、土壤质地(砂粒+粉粒+黏粒)、含水量+土

壤质地、有机质+含水量、有机质+土壤质地以及有机质+含水量+土壤质地共计7种土壤属性输入组合。

### 1.6 评价指标

利用验证数据集，采用决定系数(Coefficient of determination,  $R^2$ )、均方根误差(Root mean square error, RMSE)和平均误差(Mean error, ME)3个指标来评价已发表PTF模型和RF算法构建PTF模型的容重预测精度。其中，RMSE越小， $R^2$ 越接近1，模型的预测精度越高。ME可以量化系统误差，并指出模型高估或低估的趋势。各评价参数计算公式如下：

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

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

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

式中,  $n$  为样本数量,  $y_i$  和  $\hat{y}_i$  分别为第  $i$  个样本的真实值和预测值;  $\bar{y}_i$  为真实值的平均值。

### 1.7 变量重要性

变量重要性可以通过 RF 模型计算得到的均方误差增量 (IncMSE) 进行评价<sup>[44]</sup>。通过对每一个预测变量随机赋值, 如果该预测变量更为重要, 那么其值被随机替换后模型预测的误差会增大。因此, 该值越大表示变量越重要。本研究中 IncMSE 值使用 R 语言 “randomForest” 包进行计算。需要说明的是, 为了最大程度利用采集土壤样本, 准确揭示不同土壤属性在容重预测中的重要程度, 计算变量重要性时的 RF 模型是基于完整数据集构建的, 同样采用 5-fold 交叉验证进行参数调优, 避免模型过拟合对变量间相对重要程度的影响。

## 2 结 果

### 2.1 土壤属性基本特征

土壤容重范围介于  $0.70\sim1.81\text{ g}\cdot\text{cm}^{-3}$  之间, 均值为  $1.36\text{ g}\cdot\text{cm}^{-3}$ , 其中表层为  $1.31\text{ g}\cdot\text{cm}^{-3}$ , 亚表层为  $1.41\text{ g}\cdot\text{cm}^{-3}$ , 亚表层土壤容重均值显著大于表层 ( $P<0.05$ )。土壤含水量均值为  $21.32\%$ , 表层和亚表层均值间无显著性差异, 分别为  $21.62\%$  和  $21.03\%$ 。因采样区域覆盖整个东北黑土地, 有机质含量范围较大, 土壤有机质含量介于  $1.20\sim139.40\text{ g}\cdot\text{kg}^{-1}$  之间, 均值为  $31.93\text{ g}\cdot\text{kg}^{-1}$ , 亚表层均值含量显著低于表层含量, 分别为  $28.56\text{ g}\cdot\text{kg}^{-1}$  和  $35.29\text{ g}\cdot\text{kg}^{-1}$ 。砂粒、粉粒和黏粒的均值分别为  $31.17\%$ 、 $62.75\%$  和  $6.08\%$ , 粉粒在表层、亚表层土壤中均占比最大, 其次是砂粒, 黏粒占比最小, 且砂粒和粉粒在表层和亚表层土壤间无显著性差异, 但黏粒表层均值显著大于亚表层 ( $P<0.05$ )。从各属性的变异程度来看<sup>[45]</sup>, 容重

的变异程度最低, 变异系数为  $12.64\%$ , 粉粒属于中等变异, 变异系数为  $27.69\%$ , 而砂粒、有机质、黏粒和含水量的变异程度较大, 变异系数分别为  $62.46\%$ 、 $61.19\%$ 、 $41.67\%$  和  $40.20\%$ , 表明这 4 种属性在不同土壤样本间存在较大差异, 具有较大的空间变异性。

### 2.2 土壤属性间的相关性特征

图 2 展示了各土壤属性之间的 Pearson 相关性分析结果, 所有土壤属性间均呈现出极显著相关关系 ( $P<0.001$ )。具体而言, 容重与含水量、有机质、粉粒以及黏粒之间呈负相关关系, 而与砂粒呈正相关关系。对于表层土壤, 容重与有机质的相关性最为显著, 相关系数为  $-0.50$ ; 与含水量的相关性次之, 相关系数为  $-0.44$ ; 而与砂粒、粉粒和黏粒的相关系数绝对值均接近  $0.30$ 。相较于表层土壤, 亚表层土壤中容重与其他土壤属性的相关性更为显著。其中, 与有机质的相关系数达到  $-0.67$ , 相较表层增加了  $17\%$ ; 与含水量的相关系数为  $-0.65$ , 增加了  $21\%$ ; 与砂粒、粉粒和黏粒的相关系数绝对值均在  $0.40$  左右, 增加了约  $10\%$ 。有机质与含水量之间呈现显著正相关关系, 且表层、亚表层以及所有土壤中, 两者的相关系数均大于  $0.6$ , 在亚表层土壤中, 有机质与含水量的相关系数达到  $0.68$ , 高于表层的  $0.64$ 。此外, 有机质和含水量与砂粒、粉粒和黏粒之间的相关系数绝对值均在  $0.5$  以上。

### 2.3 已发表土壤传递函数的适用性

本研究验证数据集在国内外已发表 PTF 模型中的容重预测结果如表 2 所示。对于表层土壤, 容重验证  $R^2$  介于  $-0.17\sim0.17$  之间, 其中 PTF1 (松嫩黑土区)、PTF4 (中国)、PTF5 (巴西亚马逊河流域) 和 PTF6 (美国加利福尼亚州) 的  $R^2$  值大于  $0$ , 而且四种 PTF 模型的 RMSE 值等于或小于  $0.17\text{ g}\cdot\text{cm}^{-3}$ ; PTF2 (黑龙江省)、PTF3 (中国) 和 PTF7 (英国) 的  $R^2$  值小于  $0$ , RMSE 分别为  $0.18\text{ g}\cdot\text{cm}^{-3}$ 、 $0.19\text{ g}\cdot\text{cm}^{-3}$  和  $0.18\text{ g}\cdot\text{cm}^{-3}$ 。ME 预测结果表明, 7 种已发表模型中只有 PTF7 (英国) 的容重预测 ME 为正值 ( $0.05$ ), 其他 PTF 的 ME 值均为负值。对于亚表层, PTF7 (英国) 的容重预测精度最高,  $R^2$  为  $0.22$ , RMSE 为  $0.13\text{ g}\cdot\text{cm}^{-3}$ 。但其他 PTF 模型的预测效果均较差,  $R^2$  均为负值, 而且这些 PTF 模型均低估了容重值, 表现为 ME 均为负值。对于所有土壤样本 (表层+

亚表层),  $R^2$  的范围在  $-0.24 \sim 0.26$  之间, 其中 PTF7 (英国) 的  $R^2$  为 0.26, RMSE 为  $0.15 \text{ g} \cdot \text{cm}^{-3}$ , 预测

性能优于其他 PTF 模型。从 ME 值来看, 除了 PTF7 的 ME 大于 0, 其他 PTFs 的 ME 值均小于 0。

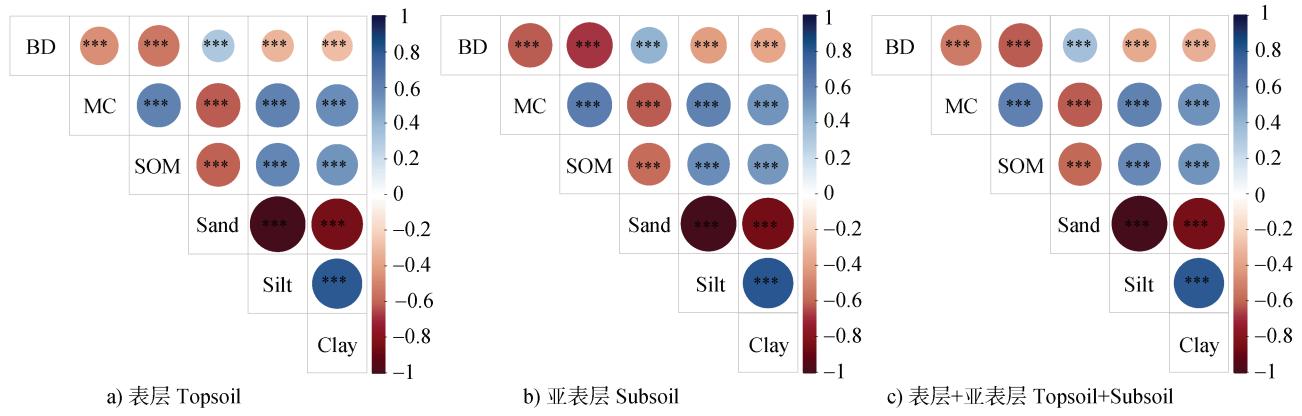


图 2 土壤属性间的相关性

Fig. 2 The correlation between soil properties

表 2 基于已发表土壤传递函数的容重预测精度

Table 2 Prediction accuracy of bulk density values based on published pedo-transfer functions (PTFs)

模型编号 ID	输入变量 Input variables	表层 Topsoil			亚表层 Subsoil			表层+亚表层 Topsoil + subsoil		
		$R^2$	RMSE	ME	$R^2$	RMSE	ME	$R^2$	RMSE	ME
PTF1 (松嫩黑土区)	SOC, Sand	0.07	0.17	-0.04	-0.22	0.17	-0.11	0.14	0.16	-0.08
PTF2 (黑龙江省)	SOM	-0.08	0.18	-0.08	-0.65	0.19	-0.15	-0.10	0.18	-0.12
PTF3 (中国)	SOM	-0.17	0.19	-0.10	-0.92	0.21	-0.17	-0.24	0.19	-0.13
PTF4 (中国)	SOC	0.12	0.17	-0.06	-0.62	0.19	-0.15	-0.01	0.17	-0.10
PTF5 (巴西亚马逊河流域)	SOC, Clay	0.17	0.16	-0.03	-0.35	0.17	-0.12	0.11	0.16	-0.07
PTF6 (美国加利福尼亚州)	SOC	0.10	0.17	-0.05	-0.30	0.17	-0.12	0.12	0.16	-0.09
PTF7 (英国)	SOC, Sand, Clay	-0.07	0.18	0.05	0.22	0.13	-0.01	0.26	0.15	0.02

从散点图可以看出, 整体上, 除了 PTF7 (图 3g: 表层, 图 3n: 亚表层, 图 3u: 所有土壤样本) 的容重真实值和预测值较为均匀地分布于  $1:1$  线两侧外, 其他基于表层、亚表层以及所有土壤样本构建的 PTF 模型的容重预测结果均低于真实值。因而这些 PTF 模型的容重预测 ME 均为负值, 这一结果可能反映了不同模型在预测过程中的系统偏差。因此, 综合  $R^2$ 、RMSE 和 ME 结果, 已发表 PTF 模型的容重预测精度较低, 估计误差较大, 难以满足东北黑土地区域尺度土壤容重预测精度的要求。

#### 2.4 随机森林模型容重预测结果

以不同土壤属性变量结合 RF 算法构建的 PTF 模型的容重预测结果如表 3 所示。对于表层土壤,  $R^2$  介于  $0.05 \sim 0.22$  之间, RMSE 最大值为

$0.17 \text{ g} \cdot \text{cm}^{-3}$ , 最小值为  $0.16 \text{ g} \cdot \text{cm}^{-3}$ , ME 介于  $0.00 \sim 0.01$  之间。当单独使用含水量、有机质或土壤质地来建模时, 基于有机质的 PTF 模型容重预测结果最优,  $R^2$  为 0.21, RMSE 为  $0.16 \text{ g} \cdot \text{cm}^{-3}$ ; 单独使用含水量或土壤质地的容重预测结果较差。使用两类土壤属性组合建模时, 有机质+含水量和有机质+土壤质地组合构建的 PTF 模型容重预测效果相近,  $R^2$  均为 0.21, RMSE 均为  $0.16 \text{ g} \cdot \text{cm}^{-3}$ , 但相较于单独使用有机质的 PTF 模型, 容重预测精度并未提升。此外, 可以发现在有机质+含水量组合中再加入土壤质地时, 对模型精度的提升效果也不显著,  $R^2$  和 RMSE 分别为 0.22 和  $0.16 \text{ g} \cdot \text{cm}^{-3}$ 。

使用含水量、有机质和土壤质地的不同组合建模时, 亚表层土壤容重的预测精度明显高于表层。

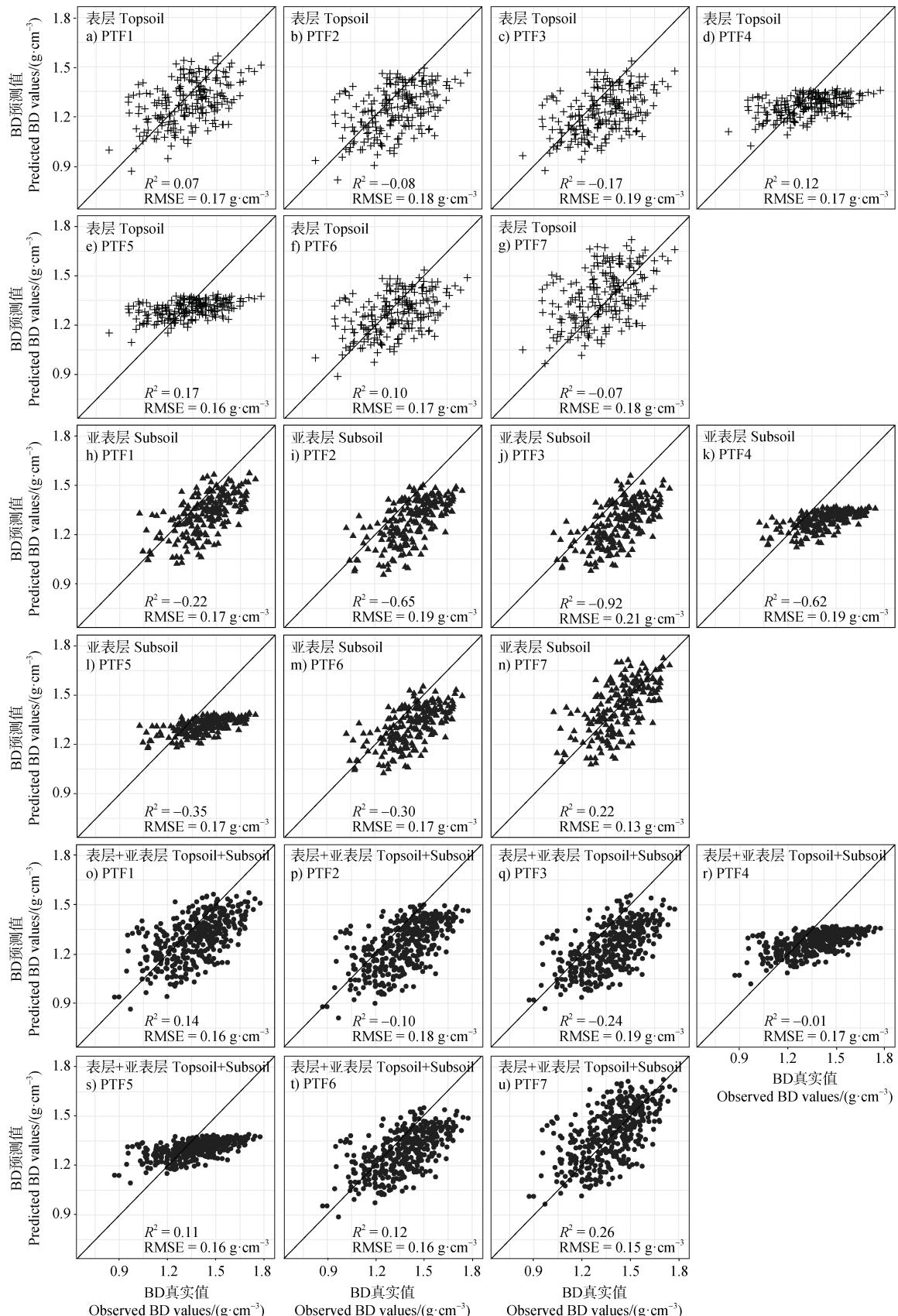


图3 已发表 PTF 模型的容重预测值与真实值散点图

Fig. 3 Scatter plots of observed vs. predicted bulk density (BD) values based on published PTFs

表 3 基于随机森林模型的容重预测精度

Table 3 Prediction accuracy of bulk density values based on random forest models

输入变量	表层 Topsoil			亚表层 Subsoil			表层+亚表层 Topsoil + subsoil		
	$R^2$	RMSE	ME	$R^2$	RMSE	ME	$R^2$	RMSE	ME
MC	0.13	0.17	0.00	0.30	0.12	0.00	0.23	0.15	0.00
SOM	0.21	0.16	0.01	0.39	0.12	0.00	0.36	0.14	0.01
SSC	0.05	0.17	0.00	0.11	0.14	-0.01	0.12	0.16	-0.01
MC + SSC	0.13	0.17	0.00	0.30	0.13	0.00	0.23	0.15	0.00
SOM + MC	0.21	0.16	0.01	0.44	0.11	0.00	0.37	0.14	0.01
SOM + SSC	0.21	0.16	0.01	0.38	0.12	0.00	0.37	0.14	0.01
SOM + MC + SSC	0.22	0.16	0.01	0.45	0.11	0.00	0.37	0.14	0.01

注: SSC 为 Sand、Silt 和 Clay 的缩写。下同。Note: SSC stands for Sand, Silt, and Clay. The same as below.

对于亚表层土壤, 除基于土壤质地构建的 PTF 模型容重预测精度较低外 ( $R^2$  和 RMSE 分别为 0.11 和  $0.14 \text{ g}\cdot\text{cm}^{-3}$ ), 其余 PTF 模型的容重预测精度均相对较高,  $R^2$  介于 0.30~0.45 之间, RMSE 介于  $0.11\sim0.13 \text{ g}\cdot\text{cm}^{-3}$  之间。当单独使用含水量、有机质或者土壤质地预测容重时, 基于有机质的 PTF 模型的容重预测效果精度最高,  $R^2$  达到 0.39, RMSE 为  $0.12 \text{ g}\cdot\text{cm}^{-3}$ ; 此外, 单独使用含水量的预测效果也较好,  $R^2$  为 0.30, RMSE 为  $0.12 \text{ g}\cdot\text{cm}^{-3}$ 。当使用两类土壤属性组合建模时, 有机质+含水量的组合预测精度最高,  $R^2$  达到了 0.44, RMSE 为  $0.11 \text{ g}\cdot\text{cm}^{-3}$ ; 相较于单独使用有机质时, 有机质+土壤质地的容重预测精度稍有降低,  $R^2$  和 RMSE 分别为 0.38 和  $0.12 \text{ g}\cdot\text{cm}^{-3}$ 。与表层土壤类似, 在有机质+含水量组合中再加入土壤质地时, 容重预测精度无明显提升。

对于所有土壤样本, PTF 模型的容重预测精度高于表层土壤, 但低于亚表层土壤。与表层和亚表层土壤类似, 使用两类土壤属性时的容重预测精度优于使用一类土壤属性; 而与两类土壤属性的结果相比, 三类土壤属性同时使用时容重预测精度并无明显提升。整体上, 容重预测  $R^2$  介于 0.12~0.37 之间, RMSE 介于  $0.14\sim0.16 \text{ g}\cdot\text{cm}^{-3}$  之间。最优 PTF 使用的土壤输入变量组合为有机质+含水量、有机质+土壤质地和有机质+含水量+土壤质地,  $R^2$  均为 0.37, RMSE 均为  $0.14 \text{ g}\cdot\text{cm}^{-3}$ 。

基于一类、两类和三类土壤属性时的最优容重预测散点图如图 5 所示。整体来看, 亚表层土壤的

容重预测值与真实值在 1:1 线两侧分布较为均匀, 预测结果较好。表层土壤和所有土壤样本时的容重预测值与真实值差距相对较大, 而且当容重含量较高时 (容重大于  $1.5 \text{ g}\cdot\text{cm}^{-3}$ ) 预测结果存在一定程度的饱和现象。与已发表 PTF 模型的容重预测精度相比 (表 2), 无论表层、亚表层还是所有土壤样本共同分析时, 基于 RF 构建的 PTF 模型的预测精度均较高 (表 3), 突出表现在已发表 PTF 模型的容重预测值与真实值差距较大 (偏离 1:1 线更多, ME 绝对值相对较大, 图 3), 而基于 RF 构建的 PTF 模型的容重预测值与真实值差距较小 (偏离 1:1 线较小, ME 接近零, 图 4)。

## 2.5 土壤属性变量重要性

为确定东北黑土地耕地土壤容重预测模型中土壤属性变量的重要性, 分别针对表层、亚表层以及全部土壤数据集 (表层+亚表层), 计算了 5 种输入土壤属性变量在 RF 模型中的 IncMSE 值, 结果如图 5 所示。结果表明, 无论是表层、亚表层还是全部土壤数据集, 有机质均为容重预测的最重要变量, 含水量次之; 相较于有机质和含水量, 土壤质地的重要程度较低; 有机质、含水量和土壤质地三者在表层、亚表层以及全部土壤数据集间的重要性排序区别不大。此外, 表层以及全部土壤数据集共同分析时, 有机质和含水量之间的重要性差异较大; 而亚表层土壤有机质和含水量之间的重要性差异较小, 而且有机质和含水量与土壤质地之间的重要性差异更大。

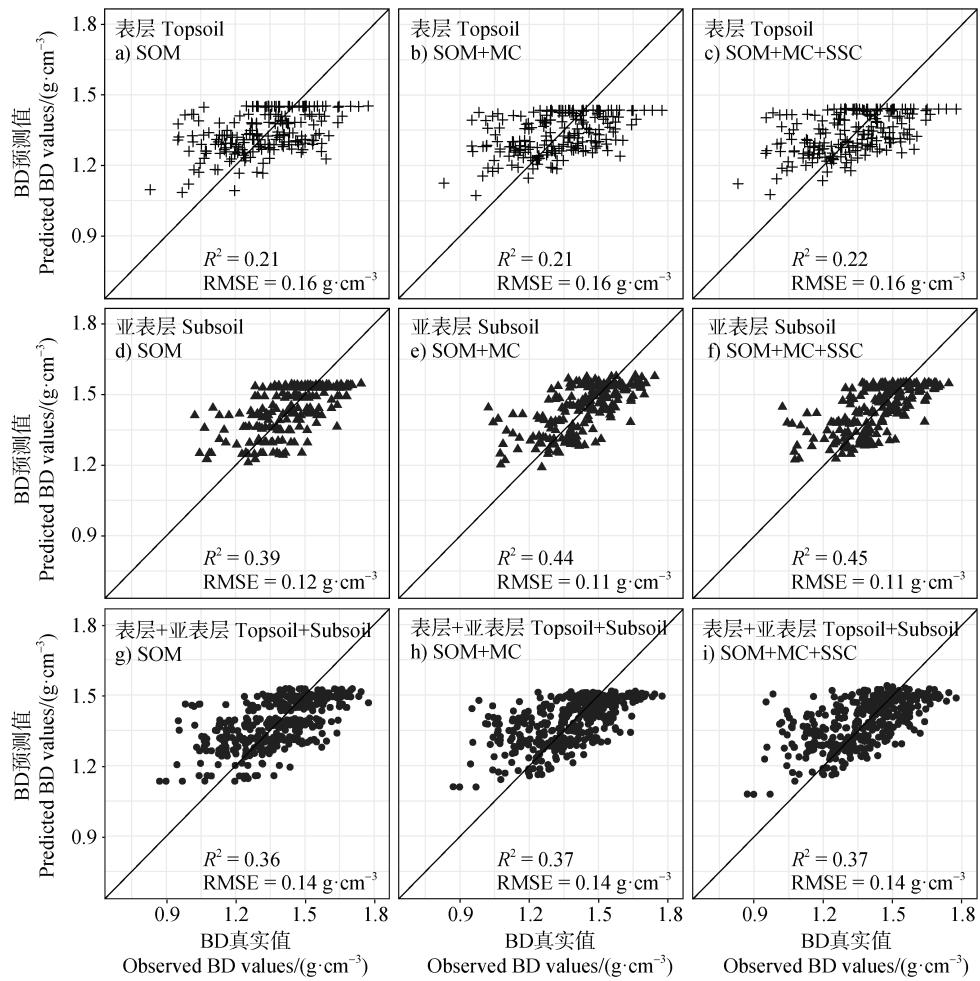


图4 基于随机森林模型的土壤容重真实值和预测值散点图

Fig. 4 Scatter plots of observed vs. predicted bulk density values based on random forest models

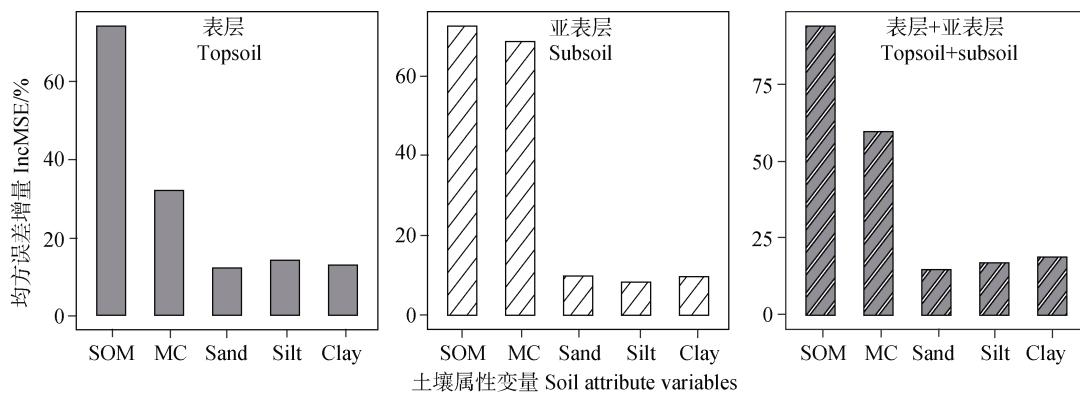


图5 基于随机森林模型预测东北黑土地土壤容重的变量重要性

Fig. 5 The variable importance for predicting bulk density values in black soil in Northeast China based on a random forest model

### 3 讨 论

本研究变量重要性计算结果表明, 土壤有机质是预测东北黑土地耕地土壤容重的最重要变量, 其

次为含水量, 土壤质地的重要性相对较小。有机质通过改变土壤颗粒的胶结状况, 从而促进土壤团聚体的形成、增加土壤孔隙度进而降低容重<sup>[5, 46]</sup>。因此, 有机质与容重具有极显著的负相关关系(图2)。

在已有基于 PTF 模型预测容重的研究中也发现有机质或有机碳是容重预测的最重要变量<sup>[13, 47]</sup>。有机质含量越高, 土壤越肥沃, 对容重的影响也越大<sup>[48]</sup>。东北黑土地作为我国的重要粮食产地, 土壤肥沃, 有机质含量较高, 因此, 在本研究中, 无论是表层、亚表层还是全部土壤数据, 使用有机质作为预测变量均能够获得相对较好的容重预测效果。含水量是东北黑土地耕地土壤容重预测时的第二重要变量。然而, 相较于有机质和土壤质地, 已有研究中含水量被用来预测容重的频率并不高。东北黑土地耕地有机质含量高, 土壤团粒结构多且稳定性好, 土壤孔隙分布比较均匀, 且毛管孔隙数量较多, 可以贮存大量的水分, 因此含水量与有机质具有极显著的正相关关系(图 2)。这也导致了含水量与容重间呈极显著的负相关关系, 且在亚表层中含水量与有机质和容重的相关性要强于表层。土壤质地在预测东北黑土地耕地土壤容重时的贡献较小。然而, 一些研究显示, 土壤质地对容重的预测起着重要作用<sup>[40, 49]</sup>。但在本研究中, 土壤质地的重要性不及有机质和含水量, 且单独使用土壤质地预测容重的效果较差。这可能是因为东北黑土地土壤有机质含量较高, 容重主要受有机质的影响。因此, 虽然土壤质地与容重关系密切, 在一些预测土壤容重的研究中发挥了重要作用, 但在本研究针对东北黑土地区域尺度土壤容重预测时并不是主要变量。

从模型的预测精度来看, 已发表的 7 种 PTF 模型在本研究—东北黑土地耕地土壤容重中的预测精度较低(表 2 和图 3)。已有基于 PTF 模型的容重预测研究也表明, PTF 模型在超出其构建区域或所在土壤类型时, 容重的预测精度通常较低<sup>[19, 50]</sup>。例如, Chen 等<sup>[18]</sup>研究表明, 当利用构建的容重预测 PTF 模型进行区域外推时, 只有当外推土壤样本与 PTF 模型构建所用土壤样本相似时才能具有一定的预测精度。因此, 利用 PTF 模型进行容重预测时有必要构建不同区域和尺度的专有模型。与已发表 PTF 模型在本研究中的容重预测精度相比, 本研究基于 RF 构建的 PTF 模型预测精度更高, 尤其是在亚表层和全部土壤预测中表现更出色(表 2 和表 3, 图 3 和图 4)。

本研究基于 RF 构建的东北黑土地 PTF 模型的土壤容重预测精度, 与 Wang 等<sup>[51]</sup>和 Schillaci 等<sup>[52]</sup>的研究结果较为接近, 但低于王思楚等<sup>[21]</sup>、韩光中

和李秀芝<sup>[22]</sup>的研究结果。这可能是由于不同研究区域土壤类型及土地利用类型等的差异以及研究尺度的差异导致的。本研究针对东北黑土地区域尺度开展容重预测, 研究区域涉及多种土壤类型和气候区域, Wang 等<sup>[51]</sup>在黄土高原大区域构建的容重预测 PTF 模型精度与本研究类似( $R^2 = 0.398$ ), 该区域也涉及多种土壤类型, 区域土壤变异较大。王思楚等<sup>[21]</sup>建立的 PTF 模型容重预测精度较高( $R^2 = 0.71$ ), 但与本研究相比, 该研究区域为松嫩黑土区, 范围较小。但韩光中和李秀芝<sup>[22]</sup>在黑龙江省域构建的 PTF 模型预测精度较高( $R^2 = 0.93$ )。因此结合已有研究可以发现 PTF 模型的容重预测精度并未随着尺度增大而降低, 但可能需要针对研究区构建专有的 PTF 模型。此外, 与基于耕地土壤的研究不同, Nanko 等<sup>[53]</sup>和 Rodríguez-Lado 等<sup>[54]</sup>在针对森林土壤的研究中容重平均值分别为 0.60 和  $0.87 \text{ g} \cdot \text{cm}^{-3}$ , 远小于本研究, 但这些研究中容重预测精度相对较高, 多元线性回归模型的  $R^2$  分别为 0.678 和 0.79。这可能是由于森林土壤受人为扰动较少, 而耕地土壤受翻耕、机械压实等人为活动影响较大导致的。De Vos 等<sup>[23]</sup>认为可以利用 RMSE 作为容重预测精度的判定标准, 其推荐的范围为  $0.12 \sim 0.25 \text{ g} \cdot \text{cm}^{-3}$ (表层土壤) 和  $0.14 \sim 0.29 \text{ g} \cdot \text{cm}^{-3}$ (全部土壤)。本研究基于 RF 构建的 PTF 模型的容重预测 RMSE 值相对较小, 在其推荐标准范围内。因此, 本研究基于东北黑土地耕地土壤构建的 PTF 模型具有预测土壤容重的潜力。但与国内外研究结果对比分析发现, 本研究容重预测  $R^2$  处于中等水平。

虽然本研究构建的 PTF 模型具有预测东北黑土地耕地土壤容重的潜力, 但其在表层土壤的预测精度仍然较低。因此, 未来的研究需要进一步优化模型, 以提升预测精度。有研究表明, 除本研究所选变量之外, 其他与容重密切相关的变量也有预测容重的潜力。例如, Quraishi 和 Mouazen 在建立 PTF 模型时除利用有机质, 含水量和黏粒外, 还选择了穿透阻力作为输入变量, 获得了较高的预测精度( $R^2 = 0.79$ ); Chen 等<sup>[55]</sup>加入 pH 来预测容重也取得了较好的预测效果,  $R^2$  达到了 0.646。因此, 将 pH 和土壤穿透阻力等相对较容易获取的土壤属性加入到 PTF 模型中, 可能有助于提高容重预测精度。其次, 深度、土壤类型、地形和气候等环境协变量也能在一定程度上提高容重预测精度。Reidy 等<sup>[20]</sup>的研究

指出,不同土层深度的容重预测精度存在显著差异; Palladino 等<sup>[19]</sup>发现,土壤类型对模型的预测精度影响很大; Schillaci 等<sup>[52]</sup>的研究则表明,气候和地形数据的加入可以增强模型的可靠性。后续研究可以考虑将环境协变量如深度、土壤类型等作为预测容重的输入变量。此外,本研究仅采用了 RF 方法,其他机器学习方法如人工神经网络和支持向量机也是可行的技术。尽管这些方法尚未广泛应用于容重预测,但它们具有独特的优势。例如,人工神经网络具有较强的自学习能力和处理非线性问题能力<sup>[56]</sup>,而支持向量机鲁棒性强,具有良好的泛化能力<sup>[49]</sup>。因此,在将来研究中可以进一步探索这些机器学习方法在东北黑土地土壤容重预测中的应用潜力。

## 4 结 论

本研究针对东北黑土地区域尺度耕地土壤,以有机质、含水量和土壤质地作为输入变量,采用 RF 机器学习方法构建了面向容重预测的 PTF 模型,并揭示了土壤属性变量的重要性。此外,本研究还验证了国内外已发表 PTF 模型在东北黑土地耕地土壤容重预测中的适用性。所得结论包括以下几点:(1)已发表 PTF 模型的容重预测精度较低,难以用于东北黑土地区域尺度土壤容重的预测;(2)采用 RF 方法建立的 PTF 模型具有容重预测的潜力,表层、亚表层以及全部土壤均是有机质+含水量和有机质+含水量+土壤质地变量组合的预测效果最好,其中亚表层土壤容重的预测精度最高;(3)有机质是东北黑土地耕地土壤容重预测的最重要变量,其次为含水量,土壤质地相关变量的重要性相对较低。

**致 谢** 感谢中国科学院南京土壤研究所刘峰、赵玉国、于东升、赵永存、胡文友、姜军、高璐璐和马利霞在土壤样品采集过程中所做的贡献。

## 参考文献 (References)

- [ 1 ] Sequeira C H, Wills S A, Seybold C A, et al. Predicting soil bulk density for incomplete databases[J]. *Geoderma*, 2014, 213: 64—73.
- [ 2 ] Abbaspour-Gilandeh Y, Abbaspour-Gilandeh M, Babaie H A, et al. Modeling agricultural soil bulk density using artificial neural network and adaptive neuro-fuzzy inference system[J]. *Earth Science Informatics*, 2023, 16 ( 1 ): 57—65.
- [ 3 ] Bruand A, Pérez Fernández P, Duval O. Use of class pedotransfer functions based on texture and bulk density of clods to generate water retention curves[J]. *Soil Use and Management*, 2003, 19 ( 3 ): 232—242.
- [ 4 ] Yu D X, Jia X X, Huang L M, et al. Spatial variation of soil bulk density in different soil layers in the loess area and simulation[J]. *Acta Pedologica Sinica*, 2019, 56 ( 1 ): 55—64. [于冬雪, 贾小旭, 黄来明, 等. 黄土区不同土层土壤容重空间变异与模拟[J]. 土壤学报, 2019, 56 ( 1 ): 55—64.]
- [ 5 ] Chai H, He N P. Evaluation of soil bulk density in Chinese terrestrial ecosystems for determination of soil carbon storage on a regional scale[J]. *Acta Ecologica Sinica*, 2016, 36 ( 13 ): 3903—3910. [柴华, 何念鹏. 中国土壤容重特征及其对区域碳贮量估算的意义[J]. 生态学报, 2016, 36 ( 13 ): 3903—3910.]
- [ 6 ] Han X Z, Li N. Research progress of black soil in Northeast China[J]. *Scientia Geographica Sinica*, 2018, 38 ( 7 ): 1032—1041. [韩晓增, 李娜. 中国东北黑土地研究进展与展望[J]. 地理科学, 2018, 38 ( 7 ): 1032—1041.]
- [ 7 ] Wang J K, Xu X R, Pei J B, et al. Current situations of black soil quality and facing opportunities and challenges in Northeast China[J]. *Chinese Journal of Soil Science*, 2021, 52 ( 3 ): 695—701. [汪景宽, 徐香茹, 裴久渤, 等. 东北黑土地区耕地质量现状与面临的机遇和挑战[J]. 土壤通报, 2021, 52 ( 3 ): 695—701.]
- [ 8 ] Xu Y D, Pei J B, Li S Y, et al. Main characteristics and utilization countermeasures for black soils in different regions of Northeast China[J]. *Chinese Journal of Soil Science*, 2023, 54 ( 2 ): 495—504. [徐英德, 裴久渤, 李双异, 等. 东北黑土地不同类型区主要特征及保护利用对策[J]. 土壤通报, 2023, 54 ( 2 ): 495—504.]
- [ 9 ] Abed Gatea Al-Shammary A, Kouzani A Z, Kaynak A, et al. Soil bulk density estimation methods: A review[J]. *Pedosphere*, 2018, 28 ( 4 ): 581—596.
- [ 10 ] Liu J H, Lan C B, Chen J. Fitting and precision evaluation of regional PTFs for soil bulk density—A case study from Fengqiu County, Henan Province[J]. *Chinese Journal of Soil Science*, 2013, 44 ( 1 ): 77—82. [刘继红, 兰传宾, 陈杰. 区域土壤容重转换函数构建与预测结果评价——以河南省封丘县为例[J]. 土壤通报, 2013, 44 ( 1 ): 77—82.]
- [ 11 ] Padarian J, Morris J, Minasny B, et al. Pedotransfer functions and soil inference systems[M]//*Progress in Soil Science*. Cham: Springer International Publishing, 2018: 195—220.
- [ 12 ] Brahim N, Bernoux M, Gallali T. Pedotransfer functions to estimate soil bulk density for Northern Africa: Tunisia case[J]. *Journal of Arid Environments*, 2012, 81: 77—83.
- [ 13 ] Heuscher S A, Brandt C C, Jardine P M. Using soil

- physical and chemical properties to estimate bulk density[J]. *Soil Science Society of America Journal*, 2005, 69 ( 1 ): 51—56.
- [ 14 ] Yi X S, Li G S, Yin Y Y. Pedotransfer functions for estimating soil bulk density: A case study in the three-river headwater region of Qinghai Province, China[J]. *Pedosphere*, 2016, 26 ( 3 ): 362—373.
- [ 15 ] Zheng G H, Jiao C X, Xie X L, et al. Pedotransfer functions for predicting bulk density of coastal soils in East China[J]. *Pedosphere*, 2023, 33 ( 6 ): 849—856.
- [ 16 ] Qiao J B, Zhu Y J, Jia X X, et al. Development of pedotransfer functions for predicting the bulk density in the critical zone on the Loess Plateau, China[J]. *Journal of Soils and Sediments*, 2019, 19 ( 1 ): 366—372.
- [ 17 ] Ghehi N G, Nemes A, Verdoodt A, et al. Nonparametric techniques for predicting soil bulk density of tropical rainforest topsoils in Rwanda[J]. *Soil Science Society of America Journal*, 2012, 76 ( 4 ): 1172—1183.
- [ 18 ] Chen S C, Richer-de-Forges A C, Saby N P A, et al. Building a pedotransfer function for soil bulk density on regional dataset and testing its validity over a larger area[J]. *Geoderma*, 2018, 312: 52—63.
- [ 19 ] Palladino M, Romano N, Pasolli E, et al. Developing pedotransfer functions for predicting soil bulk density in Campania[J]. *Geoderma*, 2022, 412: 115726.
- [ 20 ] Reidy B, Simo I, Sills P, et al. Pedotransfer functions for Irish soils—estimation of bulk density ( $\rho_b$ ) per horizon type[J]. *Soil*, 2016, 2 ( 1 ): 25—39.
- [ 21 ] Wang S C, Wang Z Q, Cheng C C, et al. Pedotransfer functions to estimate soil bulk density in Song Nen black soil region of Northeast China[J]. *Journal of Beijing Normal University ( Natural Science )*, 2018, 54 ( 3 ): 381—390. [王思楚, 王志强, 成聪聪, 等. 东北松嫩黑土区土壤密度传递函数及其适用性[J]. 北京师范大学学报(自然科学版), 2018, 54 ( 3 ): 381—390.]
- [ 22 ] Han G Z, Li X Z. Pedotransfer functions for estimating soil bulk density in Heilongjiang Province[J]. *Journal of Neijiang Normal University*, 2014, 29 ( 2 ): 53—55. [韩光中, 李秀芝. 黑龙江省土壤容重传递函数研究[J]. 内江师范学院学报, 2014, 29 ( 2 ): 53—55.]
- [ 23 ] De Vos B, van Meirvenne M, Quataert P, et al. Predictive quality of pedotransfer functions for estimating bulk density of forest soils[J]. *Soil Science Society of America Journal*, 2005, 69 ( 2 ): 500—510.
- [ 24 ] Kaur R, Kumar S, Gurung H P. A pedo-transfer function (PTF) for estimating soil bulk density from basic soil data and its comparison with existing PTFs[J]. *Soil Research*, 2002, 40 ( 5 ): 847.
- [ 25 ] Zhang H X, Lin C, Cheng H, et al. Variation of soil organic carbon content of Moso Bamboo forest along altitudinal gradient in Wuyi Mountain in China[J]. *Soils*, 2019, 51 ( 4 ): 821—828. [张厚喜, 林丛, 程浩, 等. 武夷山不同海拔梯度毛竹林土壤有机碳特征及影响因素[J]. 土壤, 2019, 51 ( 4 ): 821—828.]
- [ 26 ] Ramcharan A, Hengl T, Beaudette D, et al. A soil bulk density pedotransfer function based on machine learning: A case study with the NCSS soil characterization database[J]. *Soil Science Society of America Journal*, 2017, 81 ( 6 ): 1279—1287.
- [ 27 ] Quraishi M Z, Mouazen A M. Development of a methodology for *in situ* assessment of topsoil dry bulk density[J]. *Soil and Tillage Research*, 2013, 126: 229—237.
- [ 28 ] Katuwal S, Knadel M, Norgaard T, et al. Predicting the dry bulk density of soils across Denmark: Comparison of single-parameter, multi-parameter, and vis-NIR based models[J]. *Geoderma*, 2020, 361: 114080.
- [ 29 ] Song S F, He R Y, Shi Z Y, et al. Variable importance measure system based on advanced random forest[J]. *Computer Modeling in Engineering & Sciences*, 2021, 128 ( 1 ): 65—85.
- [ 30 ] Wang M Q, Lei G P. Relative and cumulative effects of climate and land use change on hydrological ecosystem services in Northeast China[J]. *Land*, 2023, 12 ( 7 ): 1298.
- [ 31 ] Liu B Y, Zhang G L, Xie Y, et al. Delineating the black soil region and typical black soil region of Northeastern China[J]. *Chinese Science Bulletin*, 2021, 66 ( 1 ): 96—106. [刘宝元, 张甘霖, 谢云, 等. 东北黑土区和东北典型黑土区的范围与划界[J]. 科学通报, 2021, 66 ( 1 ): 96—106.]
- [ 32 ] Gao J B, Liu L L, Guo L H, et al. The effects of climate change and phenological variation on agricultural production and its risk pattern in the black soil area of Northeast China[J]. *Journal of Geographical Sciences*, 2023, 33 ( 1 ): 37—58.
- [ 33 ] Liu Z W, Wang M C, Liu X N, et al. Ecological security assessment and warning of cultivated land quality in the black soil region of Northeast China[J]. *Land*, 2023, 12 ( 5 ): 1005.
- [ 34 ] Zou W X, Han X Z, Lu X C, et al. Effects of the construction of fertile and cultivated upland soil layer on soil fertility and maize yield in black soil region in Northeast China[J]. *Chinese Journal of Applied Ecology*, 2020, 31 ( 12 ): 4134—4146. [邹文秀, 韩晓增, 陆欣春, 等. 肥沃耕层构建对东北黑土区旱地土壤肥力和玉米产量的影响[J]. 应用生态学报, 2020, 31 ( 12 ): 4134—4146.]
- [ 35 ] Walkley A, Black I A. An examination of the degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method[J]. *Soil Science*, 1934, 37 ( 1 ): 29—38.
- [ 36 ] Sperazza M, Moore J N, Hendrix M S. High-resolution particle size analysis of naturally occurring very fine-grained sediment through laser diffractometry[J].

- Journal of Sedimentary Research, 2004, 74 ( 5 ): 736—743.
- [ 37 ] Kennard R W, Stone L A. Computer aided design of experiments[J]. *Technometrics*, 1969, 11 ( 1 ): 137—148.
- [ 38 ] Han G Z, Zhang G L, Gong Z T, et al. Pedotransfer functions for estimating soil bulk density in China[J]. *Soil Science*, 2012, 177 ( 3 ): 158—164.
- [ 39 ] Song G H, Li L Q, Pan G X, et al. Topsoil organic carbon storage of China and its loss by cultivation[J]. *Biogeochemistry*, 2005, 74 ( 1 ): 47—62.
- [ 40 ] Bernoux M, Cerri C, Arrouays D, et al. Bulk densities of Brazilian Amazon soils related to other soil properties[J]. *Soil Science Society of America Journal*, 1998, 62 ( 3 ): 743—749.
- [ 41 ] Alexander E B. Bulk densities of California soils in relation to other soil properties[J]. *Soil Science Society of America Journal*, 1980, 44 ( 4 ): 689—692.
- [ 42 ] Hollis J M, Hannam J, Bellamy P H. Empirically-derived pedotransfer functions for predicting bulk density in European soils[J]. *European Journal of Soil Science*, 2012, 63 ( 1 ): 96—109.
- [ 43 ] [43]Breiman L. Random forests[J]. *Machine Learning*, 2001, 45 ( 1 ): 5—32.
- [ 44 ] Feng N, Liu D D, She D L. Effects of vegetation restoration on carbonate-derived laterite erodibility in Karst mountain areas[J]. *Land Degradation & Development*, 2022, 33 ( 9 ): 1347—1365.
- [ 45 ] Song F F, Xu M G, Duan Y H, et al. Spatial variability of soil properties in red soil and its implications for site-specific fertilizer management[J]. *Journal of Integrative Agriculture*, 2020, 19 ( 9 ): 2313—2325.
- [ 46 ] Li D C, Huang J, Ma C B, et al. Soil organic matter content and its relationship with pH and bulk density in agricultural areas of China[J]. *Journal of Soil and Water Conservation*, 2020, 34 ( 6 ): 252—258. [李冬初, 黄晶, 马常宝, 等. 中国农耕区土壤有机质含量及其与酸碱度和容重关系[J]. 水土保持学报, 2020, 34 ( 6 ): 252—258.]
- [ 47 ] Martin M P, Seen D L, Boulonne L, et al. Optimizing pedotransfer functions for estimating soil bulk density using boosted regression trees[J]. *Soil Science Society of America Journal*, 2009, 73 ( 2 ): 485—493.
- [ 48 ] Liu L Y, Xu Y, Zhu S H, et al. Meta-analysis on the responses of soil bulk density to supplementation of organic fertilizers in croplands in China[J]. *Journal of Agricultural Resources and Environment*, 2021, 38 ( 5 ): 867—873. [刘丽媛, 徐艳, 朱书豪, 等. 有机肥配施对中国农田土壤容重影响的整合分析[J]. 农业资源与环境学报, 2021, 38 ( 5 ): 867—873.]
- [ 49 ] Guo L N, Fan G S, Zhang Y B, et al. Estimating the bulk density in 0—20 cm of tilled soils in China's Loess Plateau using support vector machine modeling[J]. *Communications in Soil Science and Plant Analysis*, 2019, 50 ( 14 ): 1753—1763.
- [ 50 ] Schjonning P, McBride R A, Keller T, et al. Predicting soil particle density from clay and soil organic matter contents[J]. *Geoderma*, 2017, 286: 83—87.
- [ 51 ] Wang Y Q, Shao M A, Liu Z P, et al. Prediction of bulk density of soils in the Loess Plateau Region of China[J]. *Surveys in Geophysics*, 2014, 35 ( 2 ): 395—413.
- [ 52 ] Schillaci C, Perego A, Valkama E, et al. New pedotransfer approaches to predict soil bulk density using WoSIS soil data and environmental covariates in Mediterranean agro-ecosystems[J]. *Science of the Total Environment*, 2021, 780: 146609.
- [ 53 ] Nanko K, Ugawa S, Hashimoto S, et al. A pedotransfer function for estimating bulk density of forest soil in Japan affected by volcanic ash[J]. *Geoderma*, 2014, 213: 36—45.
- [ 54 ] Rodríguez-Lado L, Rial M, Taboada T, et al. A pedotransfer function to map soil bulk density from limited data[J]. *Procedia Environmental Sciences*, 2015, 27: 45—48.
- [ 55 ] Chen Y, Huang Y, Sun W J. Using organic matter and pH to estimate the bulk density of afforested/reforested soils in northwest and Northeast China[J]. *Pedosphere*, 2017, 27 ( 5 ): 890—900.
- [ 56 ] Gao Z W, Wu D M, Chen X, et al. Machine learning in nitrogen cycle research: A review[J]. *Soils*, 2023, 55 ( 4 ): 689—698. [高志炜, 吴电明, 陈曦, 等. 机器学习在氮循环领域的应用研究进展[J]. 土壤, 2023, 55 ( 4 ): 689—698.]

(责任编辑: 檀满枝)