

DOI: 10.11766/trxb202311270498

CSTR: 32215.14.trxb202311270498

李艾雯, 李文丹, 宋靛颖, 冉敏, 陈丹, 成金礼, 齐浩然, 郭聪慧, 李启权. 四川盆地耕地表层土壤容重缺失数据填补方法[J]. 土壤学报, 2025, 62(1): 40–53.

LI Aiwen, LI Wendan, SONG Liangying, RAN Min, CHEN Dan, CHENG Jinli, QI Haoran, GUO Conghui, LI Qiquan. Methods of Filling in Bulk Density Gaps of Cropland Topsoil in The Sichuan Basin[J]. Acta Pedologica Sinica, 2025, 62(1): 40–53.

## 四川盆地耕地表层土壤容重缺失数据填补方法\*

李艾雯, 李文丹, 宋靛颖, 冉敏, 陈丹, 成金礼, 齐浩然, 郭聪慧,  
李启权<sup>†</sup>

(四川农业大学资源学院, 成都 611130)

**摘要:** 构建土壤容重高精度预测方法是准确补全区域土壤属性数据库的需要。本研究基于全国第二次土壤普查时获得的四川盆地(含四川省和重庆市)2883个典型耕地样点数据,运用相关分析、方差分析和回归分析等方法揭示表层土壤容重的统计特征及其主控因素,采用传统土壤传递函数(PTFs)、多元线性回归(MLR)模型、径向基函数神经网络(RBFNN)模型和随机森林(RF)模型通过不分区、分流域以及分土类3种建模方式建立土壤容重预测模型,以期实现对该区域土壤容重缺失值的填补。结果表明:研究区耕地表层土壤容重处于 $0.60\sim 1.71\text{ g}\cdot\text{cm}^{-3}$ 之间,均值为 $1.29\text{ g}\cdot\text{cm}^{-3}$ 。土壤有机质、土壤亚类和夏季降雨量是土壤容重最重要的影响因素。分流域构建的RBFNN预测模型能较好地捕捉土壤容重与各影响因素的非线性关系以及这种关系的空间非平稳性,432个独立验证样点预测结果的决定系数( $R^2$ )和均方根误差(RMSE)分别为0.519和 $0.095\text{ g}\cdot\text{cm}^{-3}$ ,明显优于其他方法,其不仅有助于提高四川盆地耕地表层土壤容重缺失值的填补精度,也为其他区域土壤性质缺失值的填补提供了方法参考。

**关键词:** 土壤容重; 传递函数; 四川盆地; 多元线性回归模型; 径向基函数神经网络模型; 随机森林模型

中图分类号: S153.6 文献标志码: A

## Methods of Filling in Bulk Density Gaps of Cropland Topsoil in the Sichuan Basin

LI Aiwen, LI Wendan, SONG Liangying, RAN Min, CHEN Dan, CHENG Jinli, QI Haoran, GUO Conghui, LI Qiquan<sup>†</sup>

(College of Resources, Sichuan Agricultural University, Chengdu 611130, China)

**Abstract:** [Objective] This study aimed to construct a high precision prediction method for soil bulk density to accurately complete the regional soil attribute database. [Method] Based on the data of 2883 typical cropland samples in the Sichuan Basin (including Sichuan Province and Chongqing Municipality) obtained during the second national soil census, this study used

\* 四川省自然科学基金项目(2022NSFSC0104)资助 Supported by the Natural Science Foundation of Sichuan Province, China (No. 2022NSFSC0104)

<sup>†</sup> 通讯作者 Corresponding author, E-mail: liqq@lreis.ac.cn

作者简介: 李艾雯(1995—),女,四川西昌人,博士研究生,主要从事土壤质量演变与环境变化研究。E-mail: ivy\_laww@163.com

收稿日期: 2023-11-27; 收到修改稿日期: 2024-03-21; 网络首发日期(www.cnki.net): 2024-06-06

correlation analysis, variance analysis, and regression analysis to reveal the statistical characteristics and main controlling factors of the cropland topsoil bulk density in the Sichuan Basin. The traditional pedotransfer functions (PTFs), multiple linear regression (MLR) models, radial basis function neural network (RBFNN) model, and random forest (RF) models were used to establish a soil bulk density prediction model through three modeling methods: whole region, by river basin and by soil type, to fill the missing value of soil bulk density. 【 Result 】 The results show that the cropland topsoil bulk density in the study area ranged from 0.60 to 1.71  $\text{g}\cdot\text{cm}^{-3}$ , with a mean value of 1.29  $\text{g}\cdot\text{cm}^{-3}$ . Soil organic matter, soil subgroup, and rainfall in summer were the most important factors influencing bulk density. The RBFNN model constructed by the river basin can better capture the nonlinear relationship between soil bulk density and the influencing factors and the spatial non-stationarity of this relationship. The coefficient of determination ( $R^2$ ) and root mean square error (RMSE) of the 432 independent validation samples were 0.519 and 0.095  $\text{g}\cdot\text{cm}^{-3}$ , respectively, which were significantly better than those of other methods. 【 Conclusion 】 Therefore, the RBFNN prediction model constructed in sub-basin is helpful to improve the imputation accuracy of the missing values of topsoil bulk density in the Sichuan Basin, and also provides a method reference for the imputation of missing values of soil properties in other regions.

**Key words:** Soil bulk density; Pedotransfer functions; Sichuan Basin; Multiple linear regression model; Radial basis function neural network model; Random forest model

作为衡量土壤紧实度的重要指标，容重不仅可以基于土壤孔隙度表征土壤的物理性质，还可以基于土壤有机质等要素的储量反映土壤的肥力水平<sup>[1-2]</sup>。了解土壤容重状况是研究土壤理化性质的基础。传统的土壤容重测定方法为环刀法，该方法原理简单、测定精度高，但存在野外采样环节多、某些情况下操作过程难度大以及人力物力成本高等问题<sup>[3-4]</sup>，导致全球范围内土壤属性数据库均存在容重数据大面积缺失的情况<sup>[5-6]</sup>。填补土壤容重缺失数据是健全区域土壤属性数据库和全面掌握土壤理化性质演变信息所必需的环节。因此，通过已有样点数据实现对土壤容重低成本且高精度的预测成为了土壤科学的一个研究重点。

国内外学者根据容重与土壤属性间的关系构建数学模型以实现土壤容重的预测，该方法被称为土壤传递函数 (Pedotransfer Functions, PTFs)<sup>[7-8]</sup>。我国多地使用 PTFs 对土壤容重进行预测并取得了不错的效果。如门明新等<sup>[9]</sup>、刘继红等<sup>[10]</sup>和李颖等<sup>[11]</sup>基于全国第二次土壤普查资料分别构建了河北省、河南省和南方喀斯特地区的土壤容重传递函数模型，补全了地方土壤属性数据库中缺失的容重数据。韩光中等<sup>[12]</sup>将 2 441 个土壤样点划分为 12 个土纲，归纳出我国主要土壤类型的土壤容重传递函数。这些研究发现，PTFs 虽然操作简单快速，但存在较差的区域适用性。这是因为成土因素、土壤类型以及土壤性质等因素对土壤容重均有影响，且在不同区域或同一区域不同亚区的影响程度明显不同<sup>[13-16]</sup>。

而基于回归模型构建的土壤容重预测模型很难准确揭示土壤容重与各影响因素间的复杂关系，导致现有模型的预测精度普遍不高<sup>[17-19]</sup>。因此，准确表达土壤容重与各影响因素间的复杂映射关系进而构建适合特定区域的高精度土壤容重预测模型亟待解决。

近年来，人工神经网络 (Artificial Neural Network, ANN) 模型和随机森林 (Random Forest, RF) 模型等机器学习方法凭借计算快速准确、学习能力强且对多元共线性不敏感等特点在预测土壤容重上受到了研究人员的青睐<sup>[20-22]</sup>。如高如泰等<sup>[23]</sup>构建了地处黄淮海平原的曲周县基于 BP 神经网络的土壤转换函数模型，预测效果明显优于多元线性回归 (Multiple Linear Regression, MLR) 模型。杨玮等<sup>[24]</sup>利用 100 个土壤样本构建的支持向量机回归土壤容重预测模型，也大幅度提高了土壤容重的预测精度。然而，基于 140 个土壤表层调查数据构建的安徽省土壤容重随机森林预测模型结果表明，机器学习方法可以有效降低模型过拟合问题，但样点密度较低和考虑的环境因子较少会导致其预测精度不高<sup>[25]</sup>。这些研究表明，运用机器学习方法构建的土壤容重预测模型在小区域范围内预测精度优于传统的 PTFs，但在较大的研究尺度下其预测精度还有待进一步提高。

作为我国四大盆地之一，四川盆地耕地面积广阔，农业历史悠久，是我国重要的粮食产区之一<sup>[26-28]</sup>。全国第二次土壤普查资料中，该区域近 30%耕地土

壤容重数据缺失, 制约了对该区域耕地土壤理化性质演变规律的准确把握。因此, 本研究在收集整理全国第二次土壤普查四川盆地典型土壤样点数据的基础上, 运用方差分析、回归分析和随机森林模型分析了成土因素、土壤类型以及土壤性质(质地类型、pH 和有机质)等因素对土壤容重的影响, 采用径向基函数神经网络模型构建该区域土壤容重预测模型, 并与传统土壤传递函数、多元线性回归模型和随机森林模型进行对比, 以期提高较大研究尺度下耕地土壤容重缺失值的预测精度, 为更好地补全区域土壤属性数据库和准确掌握耕地土壤质量演变提供方法参考。

## 1 材料与方法

### 1.1 研究区概况

四川盆地(含四川省和重庆市)位于我国西南腹地(27°38'—32°54' N、102°48'—109°16' E, 图 1a),

面积约为 26 万 km<sup>2</sup>; 耕地面积约占盆地总面积的 76%<sup>[28]</sup> (中国科学院资源环境科学与数据中心 (<https://www.resdc.cn/>) 2020 年中国土地利用遥感监测数据集), 是我国长江中上游地区重要的粮食产区。盆地整体地势中间低, 四周高, 按照地形和水系可进一步细分为岷江流域、沱江流域、涪江流域、嘉陵江流域、渠江流域以及长江干流区 6 个亚流域(图 1a)<sup>[29]</sup>。四川盆地属亚热带季风性气候, 气象站点数据统计显示年平均气温为 14.9~18.6℃, 年均降水量为 700~1 700 mm, 年日照时数为 900~1 300 h, 无霜期约为 280~350 d。成土母质主要有侏罗系的蓬莱镇组、沙溪庙组和遂宁组、白垩系的夹关组和灌口组紫色砂泥岩风化物, 上更新统黄色黏土和中下更新统老冲积物以及近代河流冲积物。土壤类型以紫色土和水稻土为主, 分别约占盆地总面积的 45%和 31%, 其次为地带性土壤黄壤。耕地利用类型主要包括旱地、水旱轮作地以及水田三类。

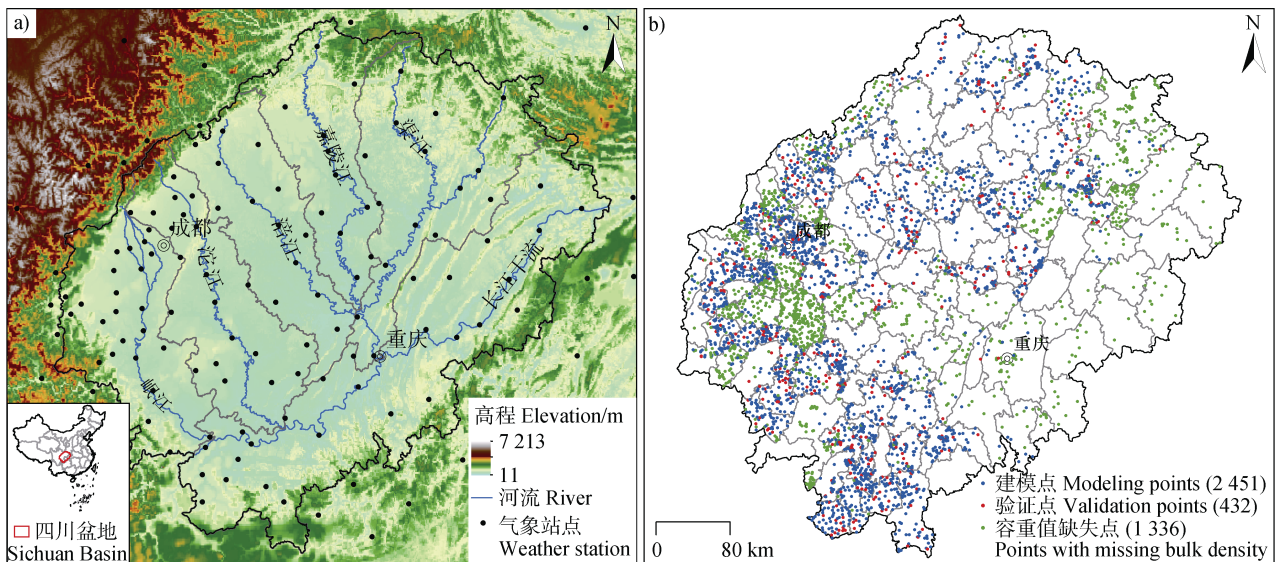


图 1 四川盆地位置和土壤样点分布图

Fig. 1 Distribution of the Sichuan Basin and soil sample sites

### 1.2 数据来源与处理

(1) 土壤数据。通过收集整理四川盆地范围内各区县全国第二次土壤普查土种志, 获得了可代表整个区域主要土种的 4219 个典型土壤样点数据, 包括采样地点、成土母岩、地形部位、土壤类型、土壤 pH 以及有机质等环境信息和土壤理化指标。其中, 2883 个土壤样点记录了表层土壤容重值, 1336

个土壤样点的表层土壤容重值缺失(图 1b)。此外, 近 60%土壤样点没有详细的土壤颗粒组成信息, 只记录了土壤质地类型(卡庆斯基制)。根据记录资料, 土壤样点所在母岩主要有第四系全新统和更新统, 白垩系城墙岩群、嘉定群、灌口组和夹关组, 侏罗系蓬莱镇组、沙溪庙组、遂宁组 and 自流井组以及三叠系须家河组等 10 余个群组。土壤类型主要为紫色

土、水稻土、黄壤、新积土、灰潮土以及石灰(岩)土。根据土壤样点所在地块的种植制度,将耕地利用类型划分为旱地、水旱轮作地以及水田 3 种。为避免受到异常值的影响,分别在土壤类型和耕地利用类型中用均值加减 3 倍标准差判断异常值<sup>[29]</sup>,并以两种分类方式下的正常最大值或最小值替换被同时标记为异常值的数值。为后续评价模型预测精度,利用 ArcGIS 10.6 软件的地统计分析模块将已有容重数值的 2883 个土壤样点划分为两个部分,即随机抽出 15% (432 个) 的土壤样点作为独立样点验证集,其余 85% (2451 个) 的土壤样点作为建模样点数据集(图 1b)。

(2) 影响因素数据。参照现有研究结果,选择气候(年、四季气温和降雨量)、地形(高程、坡度、坡向、平面曲率、剖面曲率以及地形湿度指数)、成土母岩(地层群组类型)、耕地利用类型以及土壤属性(土壤亚类类型、质地类型、土壤 pH 和有机质) 5 方面因素,共 22 个因子作为预测模型的辅助变量。其中,气象数据为研究区及其周边 144 个县级气象站点 1961—1980 年的逐日气温数据和降雨量数据;首先利用气象数据计算得到各气象站点年和四个季节的气温均值和降雨量均值,然后在 ArcGIS 10.6 软件中采用普通克里格法分别对其进行插值,从而得到四川盆地年和四个季节的气温均值和降雨量均值空间分布栅格数据。地形数据为 ALOS 12.5 m DEM,来源于中国科学院资源环境科学与数据中心(<https://www.resdc.cn/>);基于该数据在 ArcGIS 10.6 软件中计算得到研究区高程、坡度、坡向、平面曲率、剖面曲率以及地形湿度指数等地形因子。各土壤样点的气象因子及地形因子信息通过其位置提取得到。对地层群组类型、耕地利用类型、土壤亚类类型及质地类型等定性分类变量采用类别均值代替类别名称,作为变量值用于模型构建。

### 1.3 研究方法

(1) 径向基函数神经网络模型。径向基函数神经网络(Radial Basis Function Neural Network, RBFNN)模型是一种以函数逼近为基础,具有单隐层的 3 层前馈人工神经网络,结构简单、参数调节方便且学习速度快<sup>[30]</sup>。该方法可以准确揭示多元环境变量与目标变量间的非线性映射关系,被广泛运用到土壤性质的预测研究中<sup>[31-33]</sup>。本研究利用 RBFNN 模型建立了影响因子与土壤容重间的映射关系预测模型,具体过程如下:

(a) 将土壤样点按其所在的亚流域分为 6 组,采用相关分析确定各亚流域中与土壤容重具有显著相关关系( $P < 0.05$ )的影响因子。

(b) 在不同亚流域中,将与土壤容重具有显著相关性的影响因子作为 RBFNN 模型的输入变量,土壤容重作为输出变量,建立土壤容重与影响因素间关系的预测模型。

(c) 通过编程搜索,获得各亚流域 RBFNN 预测模型的最优扩展常数和隐含层节点数。即在 0.01 至 3.00 范围内以 0.01 作为间隔,隐含层节点数以 1 为间隔,逐步搜索将验证点预测结果均方根误差最小的最优参数组合,进而确定各亚流域土壤容重的最优 RBFNN 预测模型,记为 RBFNN\_分流域。

为对比分流域建模与不分区建模和分土类建模的预测效果,根据上述过程依次建立不分区的 RBFNN 预测模型(记为 RBFNN\_不分区)和分土类的 RBFNN 预测模型(记为 RBFNN\_分土类)。其中,潮土和新积土合并后能更好找到影响这两个土类容重的影响因子,故在分土类建模时将这两个土类进行合并。上述方法建模过程均在 MATLAB 2022 软件中完成。

(2) 随机森林模型。随机森林(Random Forest, RF)模型是一种基于决策树延伸出的机器学习算法,对多元共线性不敏感,不仅可以通过减少模型过拟合提高预测精度,更能对缺失数据和非平衡数据实现稳健估计<sup>[24]</sup>。本研究中采用与 RBFNN 模型相同的建模过程,按不分区、分流域和分土类 3 种方式建立基于随机森林模型的土壤容重预测模型,分别记为 RF\_不分区、RF\_分流域和 RF\_分土类。为获得不同方式下构建的 RF 模型最优参数,本研究以验证点预测结果均方根误差最小作为评价指标,采用迭代法编程搜索决策树数(ntree)和最佳分裂次数(mtry)的最优组合。RF 模型的建模过程在软件 R 中完成。

(3) 多元线性回归模型。多元线性回归(Multiple Linear Regression, MLR)模型是一种通过多个自变量的最优组合对因变量进行预测的数学方法,建模过程简单快捷。本研究中,与 RBFNN 模型的建模过程相同,采用多元逐步回归方法按不分区、分流域和分土类建立土壤容重的回归预测模型,分别记为 MLR\_不分区、MLR\_分流域和 MLR\_分土类。MLR 模型的建模过程在软件 IBM Statistics SPSS 27.0 中完成。

(4) 土壤传递函数。土壤传递函数 (Pedotransfer Functions, PTFs) 是一种根据已有土壤理化性质间的关系对未知理化性质进行模拟的数学方法。由于土壤有机质与土壤容重的相关性最为密切, 且历史土壤数据库中的土壤有机质数据一般均被较完整地记录, 因此本研究选用的传统土壤传递函数为土壤容重与土壤有机质间的非线性模型<sup>[12]</sup>, 具体表达式如下:

$$BD = \frac{a \times b}{b \times SOM + a \times (1 - SOM)} \quad (1)$$

式中,  $a$  和  $b$  为常数项,  $BD$  为土壤容重值 ( $\text{g} \cdot \text{cm}^{-3}$ ),  $SOM$  为土壤有机质含量 ( $\text{g} \cdot \text{kg}^{-1}$ )。本研究中, 根据上式按不分区、分流域和分土类建立土壤容重与土壤有机质间的非线性模型, 分别记为 PTFs\_不分区、PTFs\_分流域和 PTFs\_分土类。模型中的参数  $a$  和  $b$  通过 MATLAB 2022 软件拟合获得。

(5) 精度评价方法。本研究采用决定系数 (Coefficient of Determination,  $R^2$ )、平均绝对误差 (Mean Absolute Error, MAE)、平均相对误差 (Mean

Relative Error, MRE) 和均方根误差 (Root Mean Square Error, RMSE) 作为模型的精度评价指标<sup>[25, 31-32]</sup>, 通过对对比验证样点的实测值和各方法的预测值确定最优预测方法。一般认为,  $R^2$  越接近 1, MRE、MAE 和 RMSE 越接近 0, 预测值越接近实测值, 即模型的预测精度越高。

## 2 结果

### 2.1 土壤容重统计特征

统计结果表明 (表 1), 全国第二次土壤普查时四川盆地耕地表层土壤容重变化范围在  $0.60 \sim 1.71 \text{ g} \cdot \text{cm}^{-3}$  之间, 均值为  $1.29 \text{ g} \cdot \text{cm}^{-3}$ , 变异系数为 11.90%, 属中度变异。不同流域的耕地表层土壤容重均值存在显著差异 ( $P < 0.05$ )。其中, 嘉陵江流域土壤容重最高, 平均值为  $1.36 \text{ g} \cdot \text{cm}^{-3}$ ; 岷江流域土壤容重最低, 平均值为  $1.23 \text{ g} \cdot \text{cm}^{-3}$ ; 涪江流域和渠江流域的土壤容重显著低于嘉陵江流域 ( $P < 0.05$ ), 但显著高于沱江流域和长江干流区 ( $P < 0.05$ )。涪江流域耕地表层土壤容重为弱变异, 变异系数为 8.38%; 其余流域为中度变异, 变异系数处于 10.25%~13.25% 之间。

表 1 四川盆地耕地表层土壤容重统计特征

Table 1 Statistical characteristics of cropland topsoil bulk density in the Sichuan Basin

流域 Basin	样本数 Number	均值 Mean/ ( $\text{g} \cdot \text{cm}^{-3}$ )	最小值 Min/ ( $\text{g} \cdot \text{cm}^{-3}$ )	最大值 Max/ ( $\text{g} \cdot \text{cm}^{-3}$ )	标准差 Standard deviation/ ( $\text{g} \cdot \text{cm}^{-3}$ )	变异系数 Coefficient of variation/ %	偏度 Skewness	峰度 Kurtosis
涪江流域 <sup>①</sup>	298	1.33±0.01b	1.04	1.68	0.11	8.38	0.31	-0.22
嘉陵江流域 <sup>②</sup>	303	1.36±0.01a	0.93	1.71	0.14	10.04	-0.47	0.58
岷江流域 <sup>③</sup>	674	1.23±0.01d	0.60	1.58	0.16	13.25	-0.73	0.89
渠江流域 <sup>④</sup>	440	1.31±0.01b	0.91	1.64	0.13	10.25	-0.29	0.23
沱江流域 <sup>⑤</sup>	585	1.29±0.01c	0.78	1.66	0.14	10.98	-0.15	0.07
长江干流区 <sup>⑥</sup>	583	1.27±0.01c	0.78	1.64	0.16	12.98	-0.26	-0.27
四川盆地 <sup>⑦</sup>	2 883	1.29±0.00	0.60	1.71	0.15	11.90	-0.48	0.64

注: 不同小写字母表示不同流域之间表层土壤容重均值存在显著差异 ( $P < 0.05$ )。Note: Different lowercase letters indicate significant differences in the mean bulk density among different basins ( $P < 0.05$ ). ①Fujiang basin, ②Jialingjiang basin, ③Minjiang basin, ④Qujiang basin, ⑤Tuojiang basin, ⑥Yangtze River mainstem area, ⑦Sichuan basin

### 2.2 土壤容重的影响因素

(1) 成土母岩。方差分析结果表明 (表 2), 研究区不同成土母岩发育的土壤容重存在显著差异

( $P < 0.05$ )。蓬莱镇组、城墙岩群、沙溪庙组和遂宁组的土壤容重均值最高, 但与嘉定群和自流井组不存在显著差异 ( $P > 0.05$ )。该系统的土壤容重均值低

表 2 不同成土母岩发育土壤的容重统计特征

Table 2 Statistical characteristics of cropland topsoil bulk density in different soil parent rocks

母岩 (群或组) Soil parent rock	样本数 Number	均值 Mean/ (g·cm <sup>-3</sup> )	最小值 Min/ (g·cm <sup>-3</sup> )	最大值 Max/ (g·cm <sup>-3</sup> )	标准差 Standard deviation/ (g·cm <sup>-3</sup> )	变异系数 Coefficient of variation/ %	偏度 Skewness	峰度 Kurtosis
沙溪庙组 <sup>①</sup>	586	1.31±0.01a	0.75	1.64	0.15	11.22	-0.48	0.65
蓬莱镇组 <sup>②</sup>	562	1.32±0.01a	0.60	1.68	0.14	10.56	-0.57	1.93
全新统 <sup>③</sup>	359	1.25±0.01bc	0.69	1.71	0.15	11.68	-0.33	0.70
遂宁组 <sup>④</sup>	284	1.31±0.01a	0.66	1.64	0.16	12.21	-0.63	0.72
更新统 <sup>⑤</sup>	164	1.27±0.01b	0.65	1.57	0.16	12.50	-0.66	1.13
城墙岩群 <sup>⑥</sup>	137	1.32±0.01a	0.73	1.60	0.14	11.00	-1.00	2.74
自流井组 <sup>⑦</sup>	121	1.28±0.01ab	0.87	1.62	0.15	11.92	-0.02	-0.22
嘉定群 <sup>⑧</sup>	96	1.29±0.02ab	0.85	1.60	0.15	11.87	-0.44	-0.04
须家河组 <sup>⑨</sup>	74	1.22±0.02c	0.80	1.54	0.16	13.02	-0.26	-0.11
夹关组 <sup>⑩</sup>	54	1.27±0.02bc	0.72	1.58	0.17	13.31	-0.75	0.95
灌口组 <sup>⑪</sup>	50	1.22±0.02c	0.84	1.53	0.15	12.06	-0.32	-0.23
名山组 <sup>⑫</sup>	26	1.25±0.02bc	0.93	1.44	0.12	9.86	-1.00	1.11
其他 <sup>⑬</sup>	370	1.24±0.01bc	0.78	1.64	0.16	12.82	-0.32	-0.08

注：小写字母表示不同成土母岩之间的表层土壤容重均值存在显著差异 ( $P < 0.05$ )。Note: Different lowercase letters indicate significant differences in the mean bulk density among different soil parent rocks ( $P < 0.05$ )。①Shaximiao group, ②Penglaizhen group, ③Holocene, ④Suining group, ⑤Pleistocene, ⑥Chengqiangyan formation, ⑦Ziliujing group, ⑧Jiading formation, ⑨Xujiahe group, ⑩Jiaguan group, ⑪Guankou group, ⑫Minshan group, ⑬Others

于嘉定群和自流井组，高于夹关组，名山组、全新统和其他群组，但与这 6 个地层间均不存在显著差异 ( $P > 0.05$ )。须家河组和灌口组的土壤容重均值最低，为  $1.22 \text{ g}\cdot\text{cm}^{-3}$ ，较其他母岩低  $0.02 \sim 0.10 \text{ g}\cdot\text{cm}^{-3}$ 。名山组的土壤容重呈弱变异，变异系数为 9.86%；其余成土母岩的土壤容重为中度变异，变异系数在  $10.56\% \sim 13.31\%$  之间。

(2) 土壤类型和耕地利用类型。不同土类的土壤容重存在显著差异 (表 3,  $P < 0.05$ )。紫色土的容重均值最高，为  $1.34 \text{ g}\cdot\text{cm}^{-3}$ ，但与灰潮土和石灰(岩)土间差异不显著 ( $P > 0.05$ )。黄壤的容重均值低于灰潮土和石灰(岩)土，高于新积土，但与这 3 个土类差异均不显著 ( $P > 0.05$ )。水稻土的容重均值最低，较紫色土低  $0.08 \text{ g}\cdot\text{cm}^{-3}$ ，但与新积土无显著差异 ( $P > 0.05$ )。四川盆地灰潮土、新积土和紫色土容重的变异系数介于  $7.73\% \sim 8.86\%$  之间，具有较弱程度的变异性；而水稻土、石灰(岩)土和黄壤容重的变异系数介于  $11.36\% \sim 13.16\%$  之间，

均为中等变异。

对不同土壤亚类而言，石灰性紫色土、中性紫色土、灰潮土、石灰(岩)土和淹育水稻土的容重均值最高，但与漂洗水稻土差异不显著 (表 3,  $P > 0.05$ )。潜育水稻土、酸性紫色土和黄壤的容重均值略低于漂洗水稻土，高出新积土  $0.03 \sim 0.05 \text{ g}\cdot\text{cm}^{-3}$ ，但与漂洗水稻土和新积土均无显著差异 ( $P > 0.05$ )。渗育水稻土表层容重低于新积土且无显著差异 ( $P > 0.05$ )，但较潜育水稻土显著高出  $0.14 \text{ g}\cdot\text{cm}^{-3}$  ( $P > 0.05$ )。从变异程度看，灰潮土、新积土、石灰性紫色土和中性紫色土的容重变异系数处于  $7.73\% \sim 9.46\%$  之间，变异性较弱；其余土壤亚类的容重具有中等程度的变异性，潜育水稻土的变异系数最高，为 15.67%。

3 种耕地利用类型中，旱地的土壤容重均值最高，为  $1.32 \text{ g}\cdot\text{cm}^{-3}$ ，较水旱轮作地和水田分别显著高出  $0.05 \text{ g}\cdot\text{cm}^{-3}$  和  $0.08 \text{ g}\cdot\text{cm}^{-3}$ ；水田的土壤容重均值最低，显著低于另外 2 种利用类型 (表 4,  $P < 0.05$ )。

表 3 不同土类表层土壤容重统计特征

Table 3 Statistical characteristics of cropland topsoil bulk density in different soil groups

土类 Soil group	亚类 Soil subgroup	样本数 Number (个)	均值 Mean/ (g·cm <sup>-3</sup> )	最小值 Min/ (g·cm <sup>-3</sup> )	最大值 Max/ (g·cm <sup>-3</sup> )	标准差 Standard deviation/ (g·cm <sup>-3</sup> )	变异系数 Coefficient of variation/ %	偏度 Skewness	峰度 Kurtosis
紫色土 Purpli-Cambosols	石灰性紫色土 <sup>①</sup>	383	1.35±0.01a	1.12	1.65	0.11	7.98	0.21	-0.71
	中性紫色土 <sup>②</sup>	199	1.34±0.01a	1.07	1.63	0.13	9.46	0.13	-0.58
	酸性紫色土 <sup>③</sup>	121	1.30±0.01b	0.97	1.64	0.13	10.12	-0.14	-0.50
	合计 <sup>④</sup>	703	1.34±0.00A	0.97	1.65	0.12	8.86	0.04	-0.43
水稻土 Stagnic Anthrosols	淹育水稻土 <sup>⑤</sup>	104	1.32±0.01a	0.96	1.66	0.13	10.02	-0.21	0.22
	漂洗水稻土 <sup>⑥</sup>	51	1.32±0.02ab	0.89	1.63	0.15	11.70	-0.41	0.59
	潴育水稻土 <sup>⑦</sup>	644	1.30±0.01b	0.80	1.68	0.15	11.64	-0.34	0.13
	渗育水稻土 <sup>⑧</sup>	541	1.25±0.01c	0.81	1.60	0.14	11.45	-0.21	-0.02
	潜育水稻土 <sup>⑨</sup>	253	1.11±0.01d	0.60	1.53	0.17	15.67	-0.30	-0.15
	合计 <sup>④</sup>	1593	1.26±0.00C	0.60	1.68	0.17	13.16	-0.44	0.32
黄壤 <sup>⑩</sup>	黄壤 <sup>⑩</sup>	340	1.30±0.01Bb	0.81	1.71	0.15	11.36	-0.20	0.37
新积土 <sup>⑪</sup>	新积土 <sup>⑪</sup>	102	1.27±0.01BCbc	1.05	1.56	0.10	7.84	0.14	-0.23
灰潮土 <sup>⑫</sup>	灰潮土 <sup>⑫</sup>	99	1.33±0.01ABa	1.10	1.61	0.10	7.73	0.50	0.43
石灰(岩)土 <sup>⑬</sup>	黄色石灰土 <sup>⑭</sup>	46	1.33±0.02ABa	0.96	1.58	0.16	11.89	-0.42	-0.57

注: 不同大写字母表示不同土类之间的容重均值存在显著差异 ( $P<0.05$ ); 不同小写字母表示不同亚类之间的容重均值存在显著差异 ( $P<0.05$ )。Note: Different uppercase letters indicate significant differences in the mean of bulk density among different soil groups ( $P<0.05$ ), and different lowercase letters indicate significant differences in the mean of bulk density among different soil subgroups ( $P<0.05$ ). ①Calcaric Purpli-Cambosols, ②Typic Purpli-Cambosols, ③Dystric Purpli-Cambosols, ④Total, ⑤Hapli-Stagnic Anthrosols, ⑥Albic Fe-accumuli-Stagnic Anthrosols, ⑦Fe-accumuli-Stagnic Anthrosols, ⑧Fe-leachi-Stagnic Anthrosols, ⑨Gleyi-Stagnic Anthrosols, ⑩Ali-Perudic Argosols, ⑪Orthic Primosols, ⑫Ochri-Aquic Cambosols, ⑬Perudic Cambosols, ⑭Carbonati-Perudic Cambosols

表 4 不同耕地利用类型的表层土壤容重统计特征

Table 4 Statistical characteristics of cropland topsoil bulk density in different cropland use types

耕地利用类型 Cropland use type	样本数 Number (个)	均值 Mean/ (g·cm <sup>-3</sup> )	最小值 Min/ (g·cm <sup>-3</sup> )	最大值 Max/ (g·cm <sup>-3</sup> )	标准差 Standard Deviation/ (g·cm <sup>-3</sup> )	变异系数 Coefficient of variation/%	偏度 Skewness	峰度 Kurtosis
旱地 Dryland	1 290	1.32±0.00a	0.81	1.71	0.13	9.66	-0.12	0.20
水旱轮作地 Paddy-dryland rotation cropland	807	1.27±0.01b	0.69	1.68	0.15	11.68	-0.30	0.56
水田 Paddy fields	786	1.24±0.01c	0.6	1.61	0.18	14.46	-0.43	-0.04

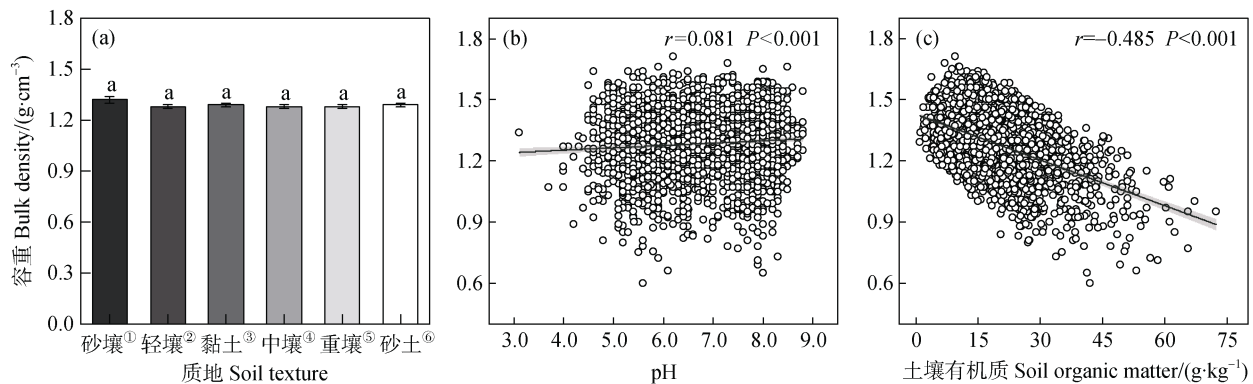
注: 不同小写字母表示不同耕地利用类型之间的表层土壤容重均值存在显著差异 ( $P<0.05$ )。Note: Different lowercase letters indicate significant differences in the mean bulk density among different cropland use types ( $P<0.05$ ).

研究区水田和水旱轮作地的土壤容重变异系数分别为 14.46% 和 11.68%，变异程度均为中等；而旱地的土壤容重变异系数为 9.66%，变异性较弱。

(3) 质地类型、土壤 pH 和有有机质。方差分析结果表明 (图 2a)，不同质地类型的土壤容重略有不同，但差异并不显著 ( $P>0.05$ )。其中，砂壤的土壤容重均值最高，为  $1.32 \text{ g}\cdot\text{cm}^{-3}$ ，较黏土和砂土高  $0.03 \text{ g}\cdot\text{cm}^{-3}$ 。轻壤、中壤和重壤的土壤容重均值最低，较砂壤低  $0.04 \text{ g}\cdot\text{cm}^{-3}$ 。研究区轻壤和砂壤的土壤

容重变异系数分别为 8.46% 和 8.90%，变异性均较弱；其余各质地类型的土壤容重变异系数在 10.45%~13.17% 之间变化，表现为中等程度的变异性。

相关分析表明 (图 2b, 图 2c)，研究区耕地表层土壤容重与土壤 pH 呈极显著正相关关系 ( $P<0.001$ )，与有机质呈极显著负相关关系 ( $P<0.001$ )。土壤容重与土壤 pH 和有有机质的相关系数分别为 0.081 和 0.485，表明有机质对该区耕地表层土壤容重的影响程度远高于土壤 pH。



注：不同小写字母表示不同质地之间的表层土壤容重均值存在显著差异 ( $P<0.05$ )。Note: Different lowercase letters indicate significant differences in the mean bulk density among different soil texture types ( $P<0.05$ )。①Sandy loam, ②Light loam, ③Clay soil, ④Medium loam, ⑤Heavy loam, ⑥Sandy soil

图 2 耕地表层土壤容重与质地、土壤 pH 和有有机质的相关关系

Fig. 2 Relationships between cropland topsoil bulk density and soil texture, soil pH and soil organic matter

(4) 地形和气候。地形因子中，坡度、坡向、平面曲率和剖面曲率与土壤容重间的相关关系不显著 ( $P>0.05$ )，但高程和地形湿度指数与土壤容重之间均呈显著负相关，相关系数分别为  $-0.158$  和  $-0.049$  ( $P<0.01$ )。气象因子中，研究区耕地表层土壤容重与气温均呈极显著正相关关系 ( $P<0.001$ )，与年降雨量和夏季降雨量呈极显著负相关关系 ( $P<0.001$ )，与春季降雨量和秋季降雨量呈极显著正相关关系 ( $P<0.001$ )，与冬季降雨量不具有相关性 ( $P>0.05$ )。土壤容重与夏季气温的相关系数最大，为  $0.281$  ( $P<0.001$ )；与冬季降雨量均值的相关系数最小，为  $-0.017$  ( $P>0.05$ )。

(5) 主控因素。采用 RF 模型揭示不同影响因素对四川盆地耕地表层土壤容重的相对重要性，结果表明不同影响因子在整个四川盆地、不同亚流域和不同土类中对土壤容重的相对重要性存在明显差异 (图 3)。对整个四川盆地而言，对土壤容重影响

最大的因素为土壤有机质、土壤亚类和降雨量，其次为气温、高程和地质地层，而耕地利用类型、土壤 pH、土壤质地和其他地形因子的影响程度较低。

不同土类中，气温和降雨量对潮土和新积土的容重影响程度最大，超过土壤有机质和高程；土壤有机质、降雨量和耕地利用类型则是石灰 (岩) 土容重的最重要影响因素。水稻土容重最重要的影响因素为土壤有机质、土壤亚类和降雨量，与整个四川盆地的情况基本一致。相较于其他土类，高程分别是紫色土和黄壤容重第 1 和第 2 重要因素。尽管土壤有机质是其他四类土壤容重最重要的影响因素，但对紫色土的影响程度相对较低。

不同亚流域上，岷江流域、渠江流域和沱江流域耕地土壤容重最重要的影响因素为土壤有机质、土壤亚类和降雨量，与整个四川盆地一致，但其他 3 个亚流域则有所不同。涪江流域耕地土壤容重最重要的因素为土壤有机质、高程、气温和降雨量



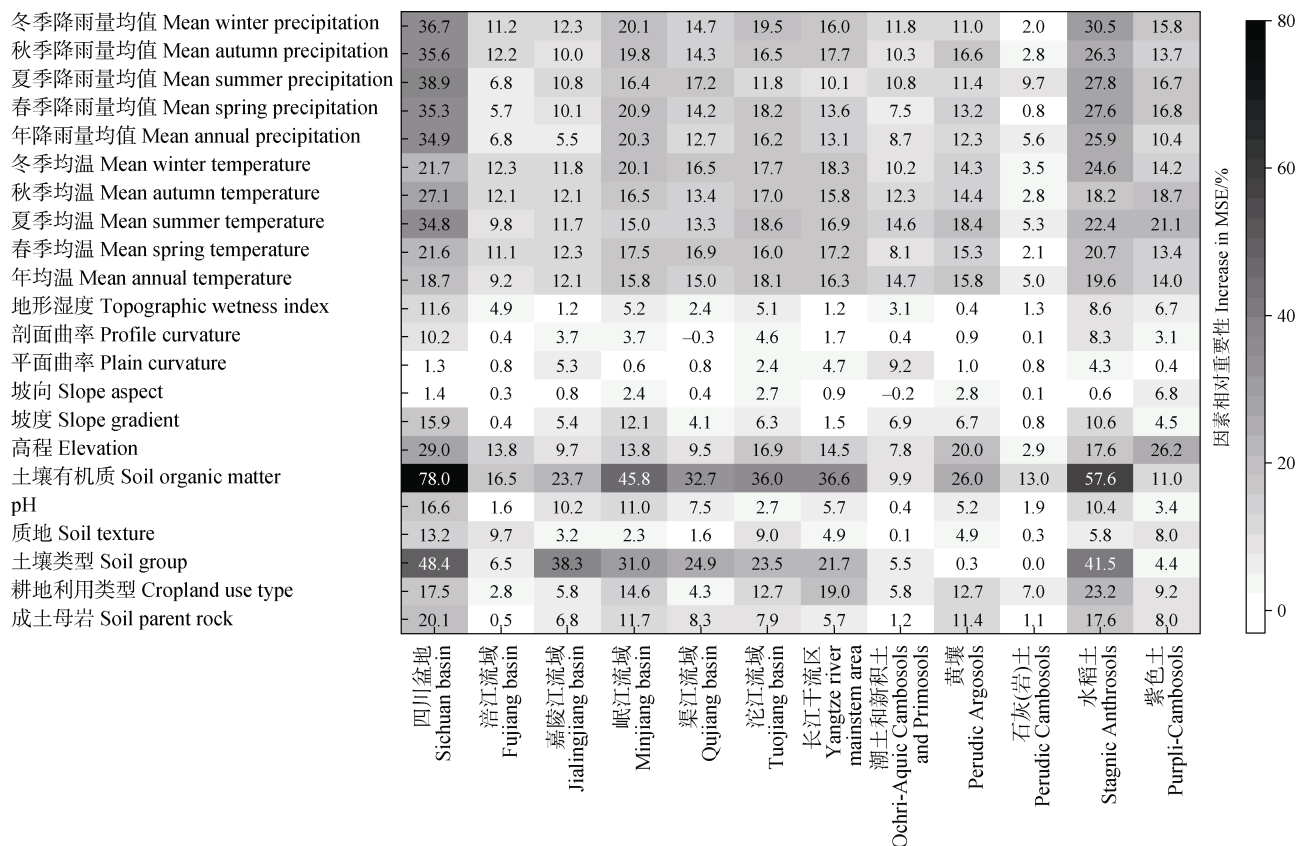


图3 不同亚流域和土类表层土壤容重影响因素的相对重要性

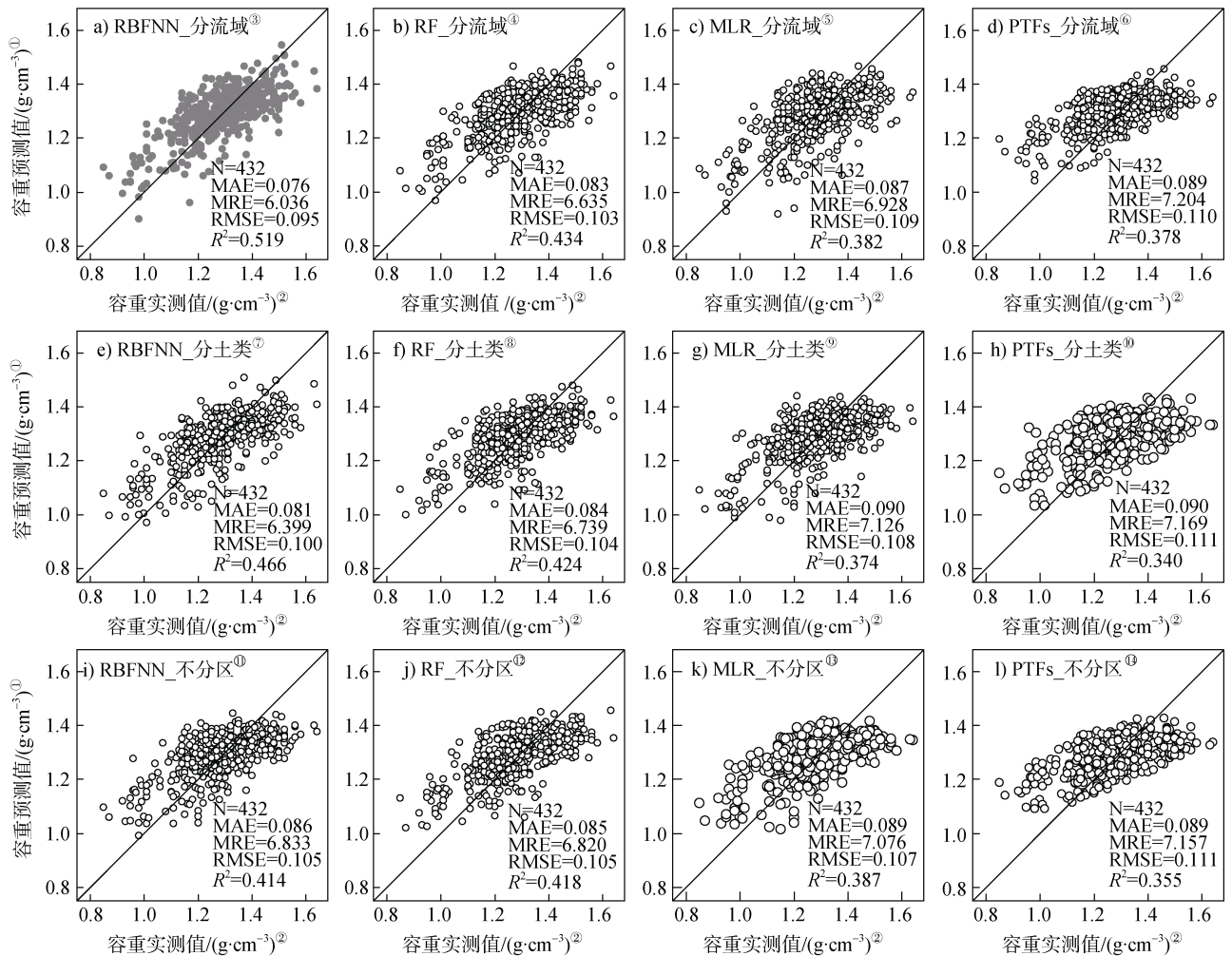
Fig. 3 The relative importance of factors affecting topsoil bulk density in different sub-basins and soil groups

此外, 质地类型也具有较重要的影响。对嘉陵江流域而言, 土壤亚类是该区土壤容重最重要的影响因素, 其余依次为土壤有机质、气温、降雨量和土壤 pH。长江干流区则是以土壤有机质、土壤亚类和耕地利用类型为土壤容重最重要的 3 个影响因素, 其次为气温、降雨量和高程等因素。

### 2.3 预测精度评价

对 432 个验证样点的预测结果表明 (图 4), 分流域建立的 RBFNN 预测模型 (RBFNN\_分流域) 各项预测误差明显小于其他建模方式和预测方法, 预测值与实测值间的  $R^2$  为 0.519, 也远高于其他建模方式和预测方法。具体而言, 4 种方法中, RBFNN 模型和 RF 模型对验证样点的预测精度远高于 MLR 模型和 PTFs 模型; 仅以土壤有机质作为自变量的 PTFs 模型在 3 种建模方式 (不分区、分流域和分土类) 下预测精度均为最低, 对验证样点预测结果的 MRE 和 RMSE 分别在 7.157%~7.204%和 0.340~0.378 之间; MLR 模型对验证样点预测结果的精度虽略高于 PTFs 模型, 但其预测

值与实测值间的  $R^2$  也未超过 0.4; 相比 MLR 模型和 PTFs 模型, RBFNN 模型和 RF 模型预测值与实测值间的  $R^2$  均在 0.4 以上, 各项预测误差也明显降低。对比 3 种建模方式可以发现, 总体上分流域建模的预测精度效果最好, 分土类和不分区建模的预测精度相对较低。如, RF\_分流域的预测值与实测值间的  $R^2$  较分土类和不分区分别提高了 0.01 和 0.016, MRE 从 6.82%和 6.739%降低至 6.635%; RBFNN 模型分流域的预测结果  $R^2$  从不分区和分土类建模方式的 0.414 和 0.466 提升至 0.519, MRE 从 6.833%和 6.399%降至 6.036%, RMSE 则低于 0.1。此外, 对比 RBFNN 模型和 RF 模型的不同建模方式可以发现, 尽管 RBFNN\_分流域模型精度达到最高, 但 RBFNN\_不分区模型的各项精度评价指标远不如 RF\_分流域模型和 RF\_分土类模型。这说明, RBFNN\_分流域模型之所以具有较高的预测精度, 不仅与模型本身有关, 还与分流域的建模方式和将更多的影响因素作为模型的输入变量有关。



Note: ①Estimated soil bulk density, ②Observed soil bulk density, ③RBFNN\_Sub-basin, ④RF\_Sub-basin, ⑤MLR\_Sub-basin, ⑥PTFs\_Sichuan basin, ⑦RBFNN\_Soil group, ⑧RF\_Soil group, ⑨MLR\_Soil group, ⑩PTFs\_Soil group, ⑪RBFNN\_Sichuan basin, ⑫RF\_Sichuan basin, ⑬MLR\_Sichuan basin, ⑭PTFs\_Sichuan basin

图4 不同预测方法下表层土壤容重预测结果的比较

Fig. 4 The prediction results of topsoil bulk density under different prediction methods

### 3 讨论

填补土壤属性数据库中缺失的土壤容重值一直备受关注<sup>[9-12, 17-19, 21, 25]</sup>, 已有研究对预测模型和精度评价方法的选取各不相同。在精度评价方法上, 少部分研究从样点数据中划分出 20%~30%的样点作为独立验证数据集对预测方法进行验证<sup>[17, 25]</sup>, 但大部分研究则未划分出独立的验证数据集, 而是直接将全部的样点用于建模, 然后以建模样点的拟合结果作为精度评价依据<sup>[9-12, 18-19, 21]</sup>。已有研究中, 对我国人为土土纲的土壤容重拟合  $R^2$  达到 0.551, RMSE 为  $0.16 \text{ g}\cdot\text{cm}^{-3}$  (建模样本数  $N=878$ , 土壤

容重均值为  $1.35 \text{ g}\cdot\text{cm}^{-3}$ )<sup>[12]</sup>。利用多种方法拟合的河北省土壤容重传递函数模型 RMSE 在  $0.16 \sim 0.46 \text{ g}\cdot\text{cm}^{-3}$  之间 (建模样本数  $N=695$ , 土壤容重均值为  $1.37 \text{ g}\cdot\text{cm}^{-3}$ )<sup>[9]</sup>。东北松嫩黑土区土壤容重预测研究中, 各方法的拟合  $R^2$  在  $0.544 \sim 0.717$  之间, RMSE 在  $0.112 \sim 0.145 \text{ g}\cdot\text{cm}^{-3}$  之间 (建模样本数  $N=146$ , 土壤容重均值为  $1.26 \text{ g}\cdot\text{cm}^{-3}$ )<sup>[18]</sup>。而采用独立验证数据集的土壤容重预测研究中, 新疆主要草地的土壤容重最优预测结果  $R^2$  为 0.29, RMSE 为  $0.238 \text{ g}\cdot\text{cm}^{-3}$  (验证样本数  $N=672$ , 土壤容重均值为  $1.24 \text{ g}\cdot\text{cm}^{-3}$ )<sup>[17]</sup>。在安徽省土壤容重的预测研究中, 验证样本的预测结果  $R^2$  为 0.22, RMSE 为  $0.12 \text{ g}\cdot\text{cm}^{-3}$

(验证样本数  $N=28$ , 土壤容重均值为  $1.24 \text{ g}\cdot\text{cm}^{-3}$ )<sup>[25]</sup>。本研究采用 4 种预测方法分 3 种建模方式对四川盆地耕地土壤容重进行预测, 所选方法既包括了土壤容重填补常用方法 (PTFs 和 MLR 模型), 也有当前在土壤性质预测中性能较好而备受关注的机器学习算法 (RBFNN 模型和 RF 模型)。由于 RBFNN 模型和 RF 模型对建模样点的拟合精度较高, RBFNN 模型甚至可以零误差逼近建模样点实测值, 因此本研究未列出两种方法对建模样点的拟合结果, 而是选用了独立验证数据集对模型精度进行评价。研究结果发现, 以多种影响因素作为模型输入变量构建的 RBFNN\_分流域模型获得了最高的预测精度, 对 432 个验证样点的预测结果  $R^2$  达到了 0.519, RMSE 小于  $0.1 \text{ g}\cdot\text{cm}^{-3}$ , 相对误差为 6.036%, 非常接近实验室实测过程中 5% 的平行误差要求。同时, RBFNN\_分流域模型对验证样点的预测能力达到甚至超过了部分研究中各方法对建模样点的拟合能力。

对比 3 种建模方式下的 4 种方法可以发现, RBFNN\_分流域模型具有较高的预测能力主要源于将多种影响因素作为模型输入变量、分流域建模方式以及 RBFNN 的模型性能 3 个方面。首先, 土壤是多种因素共同作用下的产物, 其性质受各种因素的共同影响。研究表明, 土壤容重与成土母岩、气象因子和土地利用等因素密切相关<sup>[4, 12-16, 34]</sup>。如, 母岩决定着土壤中矿物组成, 进而影响着土壤密度和土壤容重<sup>[34]</sup>。气候和地形则控制着母岩的风化与物质的运移, 从而与土壤容重呈显著的相关关系<sup>[4, 12-13]</sup>。不同土地利用类型的耕作管理措施差异明显, 影响着土壤孔隙度和压实度等土壤理化性质, 使得不同利用方式下土壤容重存在显著差异<sup>[15, 34]</sup>。因此, 仅以一两个因子作为自变量很难揭示土壤容重的变异性<sup>[30]</sup>。本研究中仅以土壤有机质作为输入变量的 PTFs 模型预测效果最差也证明了这一点。类似的, 新疆主要草地土壤容重预测模型中也仅以土壤有机碳作为自变量, 未考虑地形、气候和母质等因素, 使得其预测精度提高有限<sup>[17]</sup>。引入多种因素作为模型输入变量可以获得更好的预测效果, 这与已有研究观点一致<sup>[12, 17, 25]</sup>。其次, 土壤性质与各影响因素间的关系存在着空间非平稳性, 即不同土壤类型、不同地区或同一区域的不同亚区各因素对土壤性质的影响程度不同<sup>[31-32]</sup>, 大区域尺

度下只建立一个模型很难准确捕捉这种空间非平稳性。已有研究中, 在不同土壤类型下建立土壤容重与有机质和土壤质地间的预测模型<sup>[9, 12]</sup>, 正是利用了这种空间非平稳性。四川盆地面积广阔, 区内地形、气候和成土母岩等因素差异较大, 导致不同亚流域和不同土类的土壤容重主要控制因素与整个盆地存在差异 (图 3)。因此, 分流域和分土类建模的方式总体上优于不分区 (图 4)。本研究中, 分流域建模的效果优于分土类建模, 这是因为盆地内几种最主要的土壤类型, 包括紫色土、水稻土、黄壤、潮土和新积土, 均广泛地分布于整个盆地, 分土类建模仍不能较好地揭示某些因素在盆地内不同区域对土壤容重影响程度的差异。而相同亚流域范围缩小, 土壤容重与各影响因素间关系的空间非平稳性降低, 各预测模型能更好地描述各因素在不同流域的影响差异。第三, 土壤性质与各种影响因素间的关系多表现为复杂的非线性关系<sup>[30-31]</sup>。ANN 模型具有较强的自学习能力和处理非线性问题能力<sup>[22-23]</sup>, 特别是 RBFNN 模型是一种能以任意精度逼近任意连续函数、需要调节的参数较少的理想非线性计算工具, 通过选择合适的扩展常数和隐含层节点数据, 即可在土壤性质预测中取得较好的预测效果<sup>[30-32]</sup>。虽然近年来 RF 模型在土壤性质预测中表现出了较好的预测性能<sup>[33, 35]</sup>, 但本研究中 RF 模型的预测能力略逊于 RBFNN 模型。两种方法在各个亚流域的预测结果, 也是 RBFNN 模型的预测效果优于 RF 模型 (由于需要制作的图表较多, 文中没有一一列出不同方法对各个亚流域的预测结果)。相似的, 在采用 RF 模型对安徽省耕地土壤容重的预测研究中, 其预测结果  $R^2$  仅为 0.22<sup>[25]</sup>。另一项填补我国第二次土壤普查土壤容重缺失值的研究也发现<sup>[36]</sup>, 人工神经网络模型对验证数据集预测结果的  $R^2$  为 0.47, RMSE 为  $0.14 \text{ g}\cdot\text{cm}^{-3}$  (样本数  $N=315$ , 容重均值为  $1.32 \text{ g}\cdot\text{cm}^{-3}$ ), 预测能力优于包括梯度增强回归树模型在内的其他方法。

本研究中 RBFNN\_分流域模型对验证样点集的预测结果  $R^2$  为 0.519, 精度仍需要进一步提升。原因如下, 一是全国第二次土壤普查资料中未提供采样点的准确坐标信息, 当前使用的样点坐标是依据文献[37]和[38]的方法对所有样点进行逐一定位获得的, 即根据相似的成土环境和具有相近的土壤性质这一规律, 充分利用各种成土因素的图件数据,

在资料记录的采样地点寻找与样点具有一致的环境信息（相同的地形部位、成土母质、土壤类型、土地利用方式等）的位置作为样点的采样位置。虽然该方法具有科学性，但所获得的采样位置可能与实际采样点存在一定的偏差，使得以此位置提取的样点地形因子、气象因子等信息也会与实际采样点有一定的差异。二对耕地土壤而言，一些耕作施肥管理措施也会影响土壤容重，合理地引入并量化这些因素将有助于预测精度的进一步提高。三是辅助变量的精度有待进一步提高，如可利用更高精度的地形图生成地形因子空间分布图，以及在引入相关辅助因素后采用更优的空间预测方法获得气象因子的空间分布格局。本研究中，由于大部分样点未给出详细的土壤颗粒组成数据，建模过程中采用土壤质地类型代替，同时选择对土壤颗粒组成有较大影响或相关性较强的地形<sup>[15-16, 25]</sup>、气候<sup>[2, 4]</sup>以及其他土壤性质（如 pH<sup>[3, 39-40]</sup>）作为替代。因素分析结果表明所选指标与土壤容重均存在不同程度的相关性，RBFNN\_分流域模型也表现出了相对较优的预测性能，这为一些关键指标（如颗粒组成）缺失情况下进行容重值插补提供了参考。RBFNN\_分流域模型对验证数据集的预测结果  $R^2$  和 RMSE 表明，采用该方法对四川盆地耕地土壤容重缺失值进行填补，将极大地降低现有各方法预测结果的不确定性，从而更好地揭示该区域耕地土壤碳、氮、磷等元素自全国第二次土壤普查以来的储量变化。当前，包括深度学习等在内的一系列先进方法日益成熟，今后可在不同亚流域内采用多种先进方法进行预测，分流域择优选择不同的预测模型，进而在整个区域尺度上达到最佳的预测效果，也将是一个理想可行的建模策略。

## 4 结 论

四川盆地耕地表层土壤容重变化范围介于  $0.60 \sim 1.71 \text{ g}\cdot\text{cm}^{-3}$  之间，均值为  $1.29 \text{ g}\cdot\text{cm}^{-3}$ 。土壤有机质、土壤亚类和降雨量是影响该区域土壤容重最重要的因素，不同亚流域和土类容重的主控因素存在明显差异。分流域构建 RBFNN 模型能准确捕捉土壤容重与各影响因素的非线性关系以及这种关系的空间非平稳性，获得了优于其他方法的预测能力。

以该方法填补四川盆地耕地表层土壤容重的缺失值，有助于极大地降低现有方法预测结果的不确定性。同时，本研究提出的建模策略也可为其他区域土壤性质缺失值的填补提供方法参考。

## 参考文献 (References)

- [ 1 ] Martín M Á, Reyes M, Taguas F J. Estimating soil bulk density with information metrics of soil texture[J]. *Geoderma*, 2017, 287: 66—70.
- [ 2 ] Zhao Y L, Wang Y Q, Zhang X C. Distribution characteristics of bulk density and saturated hydraulic conductivity in intensive land restoration project areas on the Loess Plateau[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2020, 36( 10): 83—89. [赵亚丽, 王云强, 张兴昌. 黄土高原生态工程区土壤容重及饱和导水率的分布特征[J]. *农业工程学报*, 2020, 36( 10): 83—89.]
- [ 3 ] Benites V M, Machado P L O A, Fidalgo E C C, et al. Pedotransfer functions for estimating soil bulk density from existing soil survey reports in Brazil[J]. *Geoderma*, 2007, 139 ( 1/2): 90—97.
- [ 4 ] Yi X B, Shao M A, Zhao C L, et al. Variation analysis and simulation of soil bulk density within different soil depths along south to north transect of Loess Plateau[J]. *Transactions of the Chinese Society for Agricultural Machinery*, 2017, 48 ( 4): 198—205. [易小波, 邵明安, 赵春雷, 等. 黄土高原南北样带不同土层土壤容重变异分析与模拟[J]. *农业机械学报*, 2017, 48 ( 4): 198—205.]
- [ 5 ] 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.
- [ 6 ] 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.
- [ 7 ] Feng Z J, Nie W B, Yu M, et al. Multiple scale variability of soil infiltration characteristics and establishment of pedo-transfer function[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2022, 38 ( 13): 64—75. [冯正江, 聂卫波, 余淼, 等. 多尺度土壤入渗特性的变异特征和传递函数构建[J]. *农业工程学报*, 2022, 38 ( 13): 64—75.]
- [ 8 ] Xu L, He N P, Yu G R. Methods of evaluating soil bulk density: Impact on estimating large scale soil organic carbon storage[J]. *Catena*, 2016, 144: 94—101.
- [ 9 ] Men M X, Peng Z P, Xu H, et al. Investigation on pedo-transfer function for estimating soil bulk density in Hebei Province[J]. *Chinese Journal of Soil Science*, 2008, 39( 1): 33—37. [门明新, 彭正萍, 许皞, 等. 河北省土壤容重的传递函数研究[J]. *土壤通报*, 2008, 39

- (1): 33—37.]
- [ 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 ] Li Y, Liu X M, Wang S J, et al. A study on the pedo-transfer functions and influencing factors for prediction of soil bulk density for limestone soil in Karst area of South China[J]. Journal of Earth Environment, 2018, 9 ( 3 ): 245—256. [李颖, 刘秀明, 王世杰, 等. 中国南方喀斯特地区石灰土容重传递函数模型及影响因素研究[J]. 地球环境学报, 2018, 9 ( 3 ): 245—256.]
- [ 12 ] Han G Z, Wang D C, Xie X J. Pedotransfer functions for prediction of soil bulk density for major types of soils in China[J]. Acta Pedologica Sinica, 2016, 53 ( 1 ): 93—102. [韩光中, 王德彩, 谢贤健. 中国主要土壤类型的土壤容重传递函数研究[J]. 土壤学报, 2016, 53 ( 1 ): 93—102.]
- [ 13 ] 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.]
- [ 14 ] Zhu L T, Wang X Z, Jiao F. Research of soil bulk density distribution in Loess Hilly Region based on TPS interpolation methods[J]. Chinese Journal of Soil Science, 2012, 43 ( 5 ): 1043—1048. [朱乐天, 王信增, 焦峰. 基于 TPS 插值的黄土丘陵区土壤容重空间分布研究[J]. 土壤通报, 2012, 43 ( 5 ): 1043—1048.]
- [ 15 ] Wei J B, Xiao D N, Zhang X Y, et al. Spatial variation of soil bulk density in relation to terrain and land use in eroded of black soil region[J]. Journal of Soil and Water Conservation, 2006, 20 ( 3 ): 118—122. [魏建兵, 肖笃宁, 张兴义, 等. 侵蚀黑土容重空间分异与地形和土地利用的关系[J]. 水土保持学报, 2006, 20 ( 3 ): 118—122.]
- [ 16 ] Zhang L, Liu Y X, Peng H W, et al. Spatial characteristics of bulk density and related influencing factors of Bijie tobacco-planting soils[J]. Soils, 2022, 54 ( 1 ): 145—151. [张龙, 刘燕翔, 彭华伟, 等. 毕节市植烟土壤耕层容重的空间特征及影响因素[J]. 土壤, 2022, 54 ( 1 ): 145—151.]
- [ 17 ] Di X S, Wu H Q, Jia H T, et al. Construction of relationship model between soil bulk density and soil organic carbon content of main grasslands in Xinjiang[J]. Chinese Journal of Soil Science, 2021, 52 ( 6 ): 1323—1329. [狄晓双, 武红旗, 贾宏涛, 等. 新疆主要草地土壤容重与有机碳含量关系模型构建[J]. 土壤通报, 2021, 52 ( 6 ): 1323—1329.]
- [ 18 ] 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.]
- [ 19 ] Fang B, Wu S C, Chen H Y, et al. Comparison of modeling methods for region soil bulk density in different soil layers[J]. Southwest China Journal of Agricultural Sciences, 2021, 34 ( 10 ): 2258—2268. [方兵, 吴思聪, 陈弘扬, 等. 区域不同土层土壤容重建模方法比较[J]. 西南农业学报, 2021, 34 ( 10 ): 2258—2268.]
- [ 20 ] Nie C Y, Hu K L, Shao Y H, et al. Comparison of predicting results of soil hydraulic parameters by SVR and Rosetta models[J]. Journal of China Agricultural University, 2010, 15 ( 6 ): 102—107. [聂春燕, 胡克林, 邵元海, 等. 基于支持向量机和神经网络的土壤水力学参数预测效果比较[J]. 中国农业大学学报, 2010, 15 ( 6 ): 102—107.]
- [ 21 ] Yang W, Lan H, Li M Z, et al. Prediction of top soil layer bulk density based on image processing and gradient boosting regression tree model[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51 ( 9 ): 193—200. [杨玮, 兰红, 李民赞, 等. 基于图像处理 and GBRT 模型的表土层土壤容重预测[J]. 农业机械学报, 2020, 51 ( 9 ): 193—200.]
- [ 22 ] Guo L N, Fan G S. Prediction of surface soil bulk density based on grey theory and BP Neural Network[J]. Water Saving Irrigation, 2018 ( 2 ): 93—97. [郭李娜, 樊贵盛. 基于灰色理论—BP 神经网络方法的表层土壤容重预测[J]. 节水灌溉, 2018 ( 2 ): 93—97.]
- [ 23 ] Gao R T, Chen H W, Li B G, et al. Prediction of soil hydraulic characteristic parameters based on BP Neural Network[J]. Chinese Journal of Soil Science, 2005, 36 ( 5 ): 641—646. [高如泰, 陈焕伟, 李保国, 等. 基于 BP 神经网络的土壤水力学参数预测[J]. 土壤通报, 2005, 36 ( 5 ): 641—646.]
- [ 24 ] Yang W, Lan H, Li M Z, et al. Predicting bulk density and porosity of soil using image processing and support vector regression[J]. Transactions of the Chinese Society of Agricultural Engineering, 2021, 37 ( 12 ): 144—151. [杨玮, 兰红, 李民赞, 等. 基于图像处理和 SVR 的土壤容重与土壤孔隙度预测[J]. 农业工程学报, 2021, 37 ( 12 ): 144—151.]
- [ 25 ] Lu H L, Zhao M S, Liu B Y, et al. Spatial prediction of soil properties based on Random Forest Model in Anhui Province[J]. Soils, 2019, 51 ( 3 ): 602—608. [卢宏亮, 赵明松, 刘斌寅, 等. 基于随机森林模型的安徽省土壤属性空间分布预测[J]. 土壤, 2019, 51 ( 3 ): 602—608.]

- [ 26 ] Li A W, Ran M, Song L Y, et al. Spatial distribution characteristics and influencing factors of cropland topsoil organic carbon content in the Sichuan Basin[J]. Resources and Environment in the Yangtze Basin, 2023, 32 ( 5 ): 1102—1112. [李艾雯, 冉敏, 宋靛颖, 等. 四川盆地耕地表层土壤有机碳含量空间分布特征及其影响因素[J]. 长江流域资源与环境, 2023, 32 ( 5 ): 1102—1112.]
- [ 27 ] Hu S J, Deng O P, Zhang S R, et al. Spatial-temporal variability of soil organic carbon and total nitrogen in paddy soils in Sichuan Basin[J]. Soils, 2016, 48 ( 2 ): 401—408. [胡嗣佳, 邓欧平, 张世熔, 等. 四川盆地水稻土有机碳与全氮的时空变异及影响因素研究[J]. 土壤, 2016, 48 ( 2 ): 401—408.]
- [ 28 ] Li A W, Zhang Y Y, Li C J, et al. Divergent responses of cropland soil organic carbon to warming across the Sichuan Basin of China[J]. Science of the Total Environment, 2022, 851: 158323.
- [ 29 ] Li A W, Song L Y, Ran M, et al. Impact of climate warming on paddy soil organic carbon change in the Sichuan Basin of China[J]. Environmental Science, 2023, 44 ( 8 ): 4679—4688. [李艾雯, 宋靛颖, 冉敏, 等. 气候变暖对四川盆地水稻土有机碳含量变化的影响[J]. 环境科学, 2023, 44 ( 8 ): 4679—4688.]
- [ 30 ] Li Q Q, Wang C Q, Yue T X, et al. Method for spatial variety of soil organic matter based on radial basis function neural network[J]. Transactions of the Chinese Society of Agricultural Engineering, 2010, 26 ( 1 ): 87—93, 6. [李启权, 王昌全, 岳天祥, 等. 基于 RBF 神经网络的土壤有机质空间变异研究方法[J]. 农业工程学报, 2010, 26 ( 1 ): 87—93, 6.]
- [ 31 ] Li Q Q, Yue T X, Wang C Q, et al. Spatially distributed modeling of soil organic matter across China: An application of artificial neural network approach[J]. Catena, 2013, 104: 210—218.
- [ 32 ] Li Q Q, Zhang H, Jiang X Y, et al. Spatially distributed modeling of soil organic carbon across China with improved accuracy[J]. Journal of Advances in Modeling Earth Systems, 2017, 9 ( 2 ): 1167—1185.
- [ 33 ] Zhang X H, Meng X T, Tang H T, et al. Random forest prediction model for the soil organic matter with optimized spectral inputs[J]. Transactions of the Chinese Society of Agricultural Engineering, 2023, 39 ( 2 ): 90—99. [张笑寒, 孟祥添, 唐海涛, 等. 优化光谱输入量的土壤有机质随机森林预测模型[J]. 农业工程学报, 2023, 39 ( 2 ): 90—99.]
- [ 34 ] Li S, Li Q Q, Wang C Q, et al. Spatial variability of soil bulk density and its controlling factors in an agricultural intensive area of Chengdu Plain, Southwest China[J]. Journal of Integrative Agriculture, 2019, 18 ( 2 ): 290—300.
- [ 35 ] Sun Y Q, Sun X M, Wu Z F, et al. Impact of sample size and sampling method on accuracy of topsoil pH prediction on a regional scale[J]. Acta Pedologica Sinica, 2023, 60 ( 6 ): 1595—1609. [孙越琦, 孙笑梅, 巫振富, 等. 样点规模与采样方法对表层土壤 pH 空间预测精度的影响[J]. 土壤学报, 2023, 60 ( 6 ): 1595—1609.]
- [ 36 ] Liang Z Z. Research and application of high-resolution digital mapping of key properties of soil at national scale[D]. Hangzhou: Zhejiang University, 2019. [梁宗正. 国家尺度土壤关键属性高分辨率数字制图研究及应用[D]. 杭州: 浙江大学, 2019.]
- [ 37 ] Li Q Q, Yue T X, Fan Z M, et al. Spatial simulation of topsoil TN at the national scale in China[J]. Geographical Research, 2010, 29 ( 11 ): 1981—1992. [李启权, 岳天祥, 范泽孟, 等. 中国表层土壤全氮的空间模拟分析[J]. 地理研究, 2010, 29 ( 11 ): 1981—1992.]
- [ 38 ] [38] Li Q Q, Yue T X, Fan Z M, et al. Study on method for spatial simulation of topsoil SOM at national scale in China[J]. Journal of Natural Resources, 2010, 25 ( 8 ): 1385—1399. [李启权, 岳天祥, 范泽孟, 等. 中国表层土壤有机质空间分布模拟分析方法研究[J]. 自然资源学报, 2010, 25 ( 8 ): 1385—1399.]
- [ 39 ] 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.
- [ 40 ] 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.

(责任编辑: 檀满枝)