

DOI: 10.11766/trxb202211020602

周琪清, 赵小敏, 郭熙, 周洋. 基于物候与极端气候信息的耕地土壤有机碳空间分布预测研究[J]. 土壤学报, 2024, 61(3): 648–661.
ZHOU Qiqing, ZHAO Xiaomin, GUO Xi, ZHOU Yang. Prediction of Spatial Distribution of Soil Organic Carbon in Cultivated Land Based on Phenology and Extreme Climate Information[J]. Acta Pedologica Sinica, 2024, 61(3): 648–661.

基于物候与极端气候信息的耕地土壤有机碳空间分布预测研究*

周琪清^{1, 2}, 赵小敏^{1, 2†}, 郭熙^{1, 2}, 周洋^{1, 2}

(1. 江西农业大学, 国土资源与环境学院, 南昌 330045; 2. 农业农村部鄱阳湖流域农业资源与生态重点实验室(江西农业大学), 南昌 330045)

摘要: 土壤有机碳 (Soil Organic Carbon, SOC) 作为陆地生态系统中最大的碳库, 在农田土壤质量和作物产量方面发挥着重要作用。准确预测耕地 SOC 的空间分布对于制定农业管理措施至关重要。在数字土壤制图 (Digital Soil Mapping, DSM) 框架下, 选择有效的环境协变量是提高 SOC 空间预测精度的重要方法。以往遥感指数和气候变量通常使用某个时段或时点的 (平均) 值作为输入变量, 而很少有研究将时间特性和事件用于土壤有机碳预测。因此, 引入物候变量、极端气候变量弥补部分损失的地物信息和气候特征, 探讨其对研究区耕地 SOC 空间变异的响应特性及预测 SOC 空间分布的可行性。以江西省上高县为研究区域, 采用随机森林模型, 选取遥感数据、DEM 衍生变量、物候参数、气候特征因子等作为环境协变量引入模型中, 并用普通克里格 (Ordinary Kriging, OK) 对模型结果进行残差修正, 最后对比不同类型变量组合下模型的预测效果及预测精度。结果表明, 时序变量、物候变量及极端气候变量能够改善模型的预测性能, 并且残差作为误差项还能进一步提升模型的精度。结合时序变量、物候变量、极端气候变量、地形变量和残差的组合拥有最高的预测精度, 相较于地形变量、遥感变量和气候变量的组合, 将 R^2 、MAE 和 RMSE 提升了 90.00%、58.95% 和 57.14%。变量贡献率分析显示, SU、a3 和 TXx 是影响研究区耕地 SOC 分布的重要变量。因此, 物候变量和极端气候变量具有较好的应用前景, 未来还需验证极端气候变量作为环境变量在不同土地利用、大尺度研究区下预测土壤属性的有效性。

关键词: 土壤有机碳; 数字土壤制图; 随机森林残差克里格; 物候; 极端气候

中图分类号: S158.9 **文献标志码:** A

Prediction of Spatial Distribution of Soil Organic Carbon in Cultivated Land Based on Phenology and Extreme Climate Information

ZHOU Qiqing^{1, 2}, ZHAO Xiaomin^{1, 2†}, GUO Xi^{1, 2}, ZHOU Yang^{1, 2}

(1. Academy of Land Resource and Environment, Jiangxi Agricultural University, Nanchang 330045, China; 2. Key Laboratory of Agricultural Resources and Ecology in Poyang Lake Watershed of Ministry of Agriculture and Rural Affairs in China, Jiangxi Agricultural University, Nanchang 330045, China)

* 国家重点研发计划项目 (2020YFD1100603) 和国家自然科学基金项目 (41361049) 资助。Supported by the National Key R&D Program of China (No. 2020YFD1100603) and the National Natural Science Foundation of China (No. 41361049)

† 通讯作者 Corresponding author, E-mail: zhaoxm889@126.com

作者简介: 周琪清, 男, 硕士研究生, 主要从事数字土壤制图研究。E-mail: zhouqq0423@126.com

收稿日期: 2022-11-02; 收到修改稿日期: 2023-09-24; 网络首发日期 (www.cnki.net): 2023-12-12

Abstract: 【Objective】As the largest Carbon pool in the terrestrial ecosystem, Soil Organic Carbon(SOC)plays an important role in Soil quality and crop yield. Accurate prediction of the spatial distribution of SOC on cropland is essential for the development of agricultural management measures. In the framework of Digital Soil Mapping(DSM), an important method to improve the precision of SOC spatial prediction is to select an effective environmental covariate. In previous studies, the mean values of remote sensing indices and climate variables for a certain period or point in time were usually used as input variables, while temporal characteristics and events were rarely used for SOC prediction. Therefore, in order to reduce the impact of the lack of part of physical information and climate characteristics, phenological variables and extreme climate variables were added in this study. The response characteristics to the spatial variability of SOC of cultivated land and the feasibility of predicting the spatial distribution of SOC were discussed. 【Method】The research area of this paper is Shanggao County, Jiangxi province. A random forest model was used, in which remote sensing data, DEM-derived variables, phenological parameters and climatic characteristics were selected as environmental covariates, and the model results were corrected for residuals using Ordinary Kriging(OK). The prediction effect and prediction accuracy of the model under different types of variable combinations were compared. 【Result】The results show that chronological variables, phenological variables, and extreme climate variables can improve the prediction performance of the model, and the residual error as an error item can further improve the accuracy of the model. The combination of chronological variables, phenological variables, extreme climate variables, topographic variables, and residuals had the highest prediction accuracy, improving R^2 , MAE, and RMSE by 90.00%, 58.95%, and 57.14% compared to the combination of topographic variables, remote sensing variables, and climate variables. The analysis of variable contribution rates shows that SU, a3 and TXx were important variables affecting the distribution of cultivated land SOC in the study area. 【Conclusion】Phenological variables and extreme climate variables have good application prospects. In the future, it is necessary to verify the validity of extreme climate variables as environmental variables in predicting soil properties under different land use and large-scale study areas.

Key words: Soil organic carbon; Digital soil mapping; Random forest with residual kriging; Phenology; Extreme climate

土壤是地球表面最大的有机碳储库，在陆地生态系统功能中发挥着重要作用^[1]。土壤有机碳（Soil Organic Carbon, SOC）能够影响土壤中的生物、化学和物理过程^[2]，决定着耕地土壤质量和作物产量。鉴于对精准制定农业管理措施及评估耕地土壤健康的需求，详细了解耕地 SOC 的含量及空间分布至关重要。

数字土壤制图（Digital Soil Mapping, DSM）已被学者证明是一种效率高、成本低的制图方法^[3]。DSM 主要通过地理统计模型或机器学习模型建立。机器学习模型由于包含算法较多且自适应性较强，被广泛应用于不同尺度的土壤属性预测制图中^[4-6]，如多元线性回归，人工神经网络、随机森林、支持向量机等。其中，随机森林算法（RF）具有稳定性强、超参数少、处理多元非线性数据能力好等特点^[5]，逐渐成为研究土壤属性制图的主要方法^[7-8]。然而机器学习在预测过程中，只考虑了土壤环境之间的关系，忽略了相邻观测数据的影响（空间自相关）。所以有研究^[9]将机器学习所建立的趋势项与含有空

间自相关信息的预测残差项相结合，来提升空间分布的预测精度，但在国内 DSM 应用的较少。

提高空间分布的预测精度，模型选择仅一方面，从 scorpan 模型可知，引入更多有效的环境协变量也能提升预测精度^[3-5]。在以往 SOC 空间分布预测中，通常将遥感指数和气候变量视为相对静止，即使用某个时段或时点的（平均）值作为输入变量^[4-8]，而地物信息和气候特征是动态变化的过程，如果未考虑时间特性和事件对 SOC 空间变异的影响，可能无法充分体现气候、生物等因素在土壤形成过程中的作用与强度。为此引入物候参数描述植被长期生长的特征^[10]，探究其与土壤功能的关系，归一化植被指数（NDVI）是植物生长状态以及植被空间分布密度的最佳指示因子，与植被分布密度呈线性相关^[11]，因此采用 NDVI 时间序列提取物候参数。朱丹瑶^[12]研究表明植被物候变化与气候模式变化密切相关，并可能对生态系统的生态过程产生重要影响，如陆地表面-大气的水和碳交换等。然而气候变化不仅包括平均值的变化，还包括极端气候的变化^[13]。目前，

全球持续变暖,极端气候事件发生的频率和幅度逐渐增强^[14],极端气候预计会增加未来对生态系统的非生物胁迫,进而影响土壤生态,但至今关于极端天气对土壤属性的响应机制仍然知之甚少。世界气象组织委员会(WMO)为有效推动世界对极端气候事件的研究,共定义 27 个极端气候指数,大致可分为:特征量的极值、特征量超某阈值的天数、特征量的持续天数等。该指标综合考虑了极端气候事件的强度和持续时间,弥补部分气候特征的缺失。因此,将极端气候指数和物候参数纳入环境协变量,深入研究气候事件、植被生长周期对耕地 SOC 空间变异的响应特性及预测研究区耕地 SOC 空间分布的可行性。

为验证该假设,本文以江西省上高县为研究区域,采用随机森林模型,选取遥感数据、DEM 衍生变量、物候参数、气候特征因子等作为环境协变量引入模型中,并用普通克里格(Ordinary Kriging, OK)对模型结果进行残差修正,对比不同类型变量组合下模型的预测效果及预测精度。

1 材料与方法

1.1 研究区概况

上高县位于江西省西北部,位置处于赣江支流的锦江流域中段,东部与高安相毗邻,南邻新余,

西接宜春市万载县,北连宜春市宜丰县,如图 1 所示。地理坐标介于 28°02'N~28°25'N 和 114°28'E~115°10'E 之间,东西长 68 km,南北宽 45 km,总面积 1350.25 km²。气候温暖湿润,雨量充沛,日照充足,无霜期长,全年平均气温为 17.6℃,年平均降水量为 1 718mm。境内地势平缓,以低丘区为主,由西南向东北倾斜,具有西南高、中部平、东北部低的特征,地貌以山地、丘陵、平原为主,分别占全县总面积的 1.76%、65.79%、32.45%。上高县土壤类型多样,共分为 6 个土类,14 个亚类,40 个土属,72 个土种,适宜各种农作物的栽种,主要有水稻土、红壤、草甸土、紫色土、石灰石土和山地黄壤六种土壤类型。研究区主要的土地利用类型为林地、耕地和建设用地,其中主要的农业作物包括水稻、油菜和花生,农作物种植制度为一年一季和两季种植。

1.2 样品采集与制备

本研究采用随机布点法布设了 108 个耕地土壤采样点,样点分布如图 1 所示,于 2021 年 7 月完成实地采样。每个采样点使用五点采样法,收集分布在采样位置周围五个点的土样,将其混合,获得 1kg 左右的土壤样品。土壤样品经剔除杂物、自然风干、研磨和过筛后采用氧化还原体积法^[15]测定土壤样品的有机碳含量。

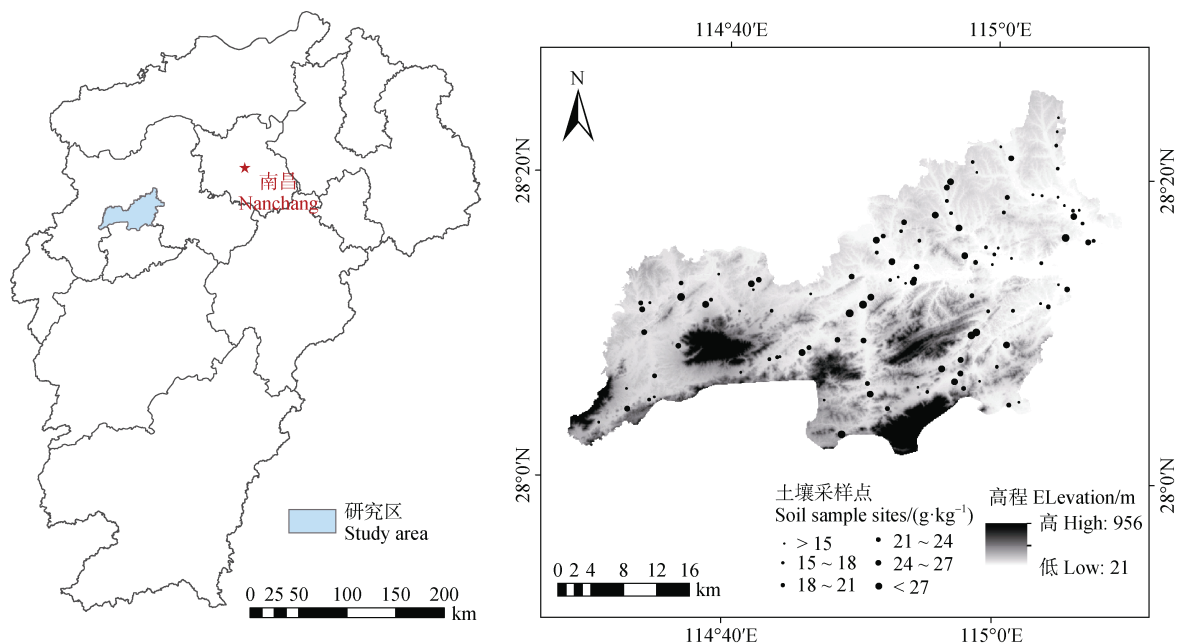


图 1 研究区地理位置及采样点分布

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

1.3 环境协变量及数据来源

将不同类别的变量分为四组,以验证物候变量和极端气候变量在耕地 SOC 预测中的有效性。(1)地形变量、气候变量和遥感变量(模型 1);(2)地形变量、气候变量和时序遥感变量(模型 2);(3)地形变量、气候变量、时序遥感变量和物候变量(模型 3);(4)地形变量、气候变量、时序遥感变量、物候变量和极端气候变量(模型 4)。本研究所采用的耕地范围数据为 2019 年哥白尼全球土地覆盖层 CGLS-LC100 系列(分辨率为 100 m),由于不同分辨率尺度可能会给制图带来不确定性,所以选取空间分辨率最低的 100 m 作为标准进行重采样。

地形变量的数据源为 ALOS PALSAR 数据,分辨率为 12.5 m。使用 ArcGIS 软件得到数字高程模型(DEM)、地形湿度指数(TWI)、坡度(Slope)、坡向(Aspect)、平面曲率(Plan)、剖面曲率(Profile)。

遥感数据为基于 GEE 平台下载的 2020 年 3 月—2021 年 2 月 Sentinel-2 L2A 产品数据。由于研究区属亚热带季风气候型,降水充沛,云覆盖率较大,致使部分数据无法使用,采用 QA60 质量波段去云方法,对去云的影像进行月度中值合成。由于 6 月份数据缺失严重,所以舍弃 6 月份遥感数据。随后通过波段计算提取归一化植被指数(NDVI)、归一化绿度指数(GNDVI)、修正土壤调整植被指数(MSAVI)、增强型植被指数(EVI)。遥感变量为 8

月份遥感指数,时序遥感变量为 1—12 月份遥感指数。遥感数据受到多种因素影响会出现异常点等现象,物候参数又易受异常值干扰,因此提取前需对植被指数时间序列进行平滑处理。Savitzky-Golay 滤波法是一种基于局域多项式最小二乘法拟合的多项式平滑滤波算法,在消除噪声和异常值的同时,可以保证时间序列的形状和宽度不变,能够较好地保持原始数据的真实性^[16],因此采用 Savitzky-Golay 滤波法重新构建植被指数时间序列,为避免过度平滑,将移动窗口宽度设置为 1,多项式拟合阶数设置为 2。通过参考相关文献研究^[10, 17-18]并反复试验后,采用动态阈值法(阈值为 0.2)提取物候变量。物候变量通过基于 Matlab 开发的 Timesat3.3 实现,具体定义见表 1。

气象数据来源于国家气象中心网站提供的全国地面气候资料日值数据集。极端气候变量为 2010—2019 年均极端气候指数(表 2),为探究极端气候变量与耕地土壤有机碳的关系,对极端气候变量与研究区耕地 SOC 进行 Pearson 相关性分析。气候变量为 2010—2019 年均最高气温(Maxt)、年最低气温(Mint)、年均气温(Meant)、年降水(Rainfall)。气候数据均采用反距离权重法插值获得^[19]。极端气候变量通过加拿大气象研究中心基于 R 开发的 RclimDex 计算得到,反距离权重法通过 ArcGIS 实现, Pearson 相关性分析通过 SPSS 实现。

表 1 物候变量定义

Table 1 Definition of phenological variables

| 标识 Identification | 定义 Definition | 单位 Unit | 标识 Identification | 定义 Definition | 单位 Unit |
|----------------------|----------------------------|------------|----------------------|----------------------------|------------|
| a1 | 季起期 | d | a2 | 季末期 | d |
| a3 | 季节长度 | d | a4 | 曲线左右两侧最小值的平均值 | value |
| a5 | 季中期 | d | a6 | 生长季曲线中的最大值 | value |
| a7 | 季节期间拟合函数的最大数据值与基准水平之间的差异 | value | a8 | 曲线左侧 80%和 20%的数据差与相应时间差的比率 | Value/d |
| a9 | 曲线右侧 80%和 20%的数据差与相应时间差的比率 | Value/d | a10 | 曲线函数数据在季节长度中的积分 | Value × d |
| a11 | 季节长度中曲线函数数据与基准面值之差的积分 | Value × d | a12 | 季节开始时对应于曲线函数的值 | value |
| a13 | 季节结束时对应于曲线函数的值 | value | | | |

表 2 极端气候变量定义

Table 2 Definition of extreme climate variables

| 分类 Classification | 指数 Index | 定义 Definition | 单位 Unit |
|-----------------------------|--|--|------------|
| 降水指数 Precipitation index | 年最大连续无降雨天数 CDD | 年最大连续几天日降雨量均小于 1 mm 的天数 | d |
| | 年最大连续降雨天数 CWD | 年最大连续几天日降雨量均大于等于 1 mm 的天数 | d |
| | 强降雨日数 R10 | 年日降雨量大于 10 mm 的天数 | d |
| | 极强降雨日数 R20 | 年日降雨量大于 20 mm 的天数 | d |
| | 年平均日降雨强度 SDII | 年雨日降雨总量除以雨日天数 | mm/d |
| | 年强降雨量 R95p | 年日降雨大于 2010—2019 年雨日降雨量第 95 分位数的降雨总量 | mm |
| | 年极强降雨量 R99p | 年日降雨大于 2010—2019 年雨日降雨量第 99 分位数的降雨总量 | mm |
| | 年总雨日降水量 PRCPTOT | 年日降雨大于 1 mm 的降雨总量 | mm |
| | 年最大 1 日降雨量 Rx1day | 年最大连续 1 日降雨量 | mm |
| | 年最大连续 5 日降雨量 Rx5day | 年最大连续 5 日降雨量 | mm |
| 温度指数 Temperature index | 霜冻日数 FD0 | 年日最低温小于 0℃ 的天数 | d |
| | 夏日日数 SU | 日最高温 > 25℃ 的日数 | d |
| | 结冰日数 ID0 | 日最高温 < 0℃ 的日数 | d |
| | 炎热夜数 TR | 日最低温 > 20℃ 的日数 | d |
| | 生长期长度 GSL | 连续 6 日 > 5℃ 或 < 5℃ 的时间跨度 | d |
| | 平均温差 DTR | 日温差的平均值 | ℃ |
| | 年日最高温的最大值 TXx | 年日最高温的最大值 | ℃ |
| | 年日最低温的最大值 TNx | 年日最低温的最大值 | ℃ |
| | 年日最高温的最小值 TXn | 年日最高温的最小值 | ℃ |
| | 年日最低温的最小值 TNn | 年日最低温的最小值 | ℃ |
| | 热持续指数 WSDI | 连续 6 日最高温在 2010—2019 年 90% 分位值的日数 | d |
| | 冷持续指数 CSDI | 连续 6 日最低温在 2010—2019 年 10% 分位值的日数 | d |
| | 冷夜日数 TN10p | 年日最低温小于 2010—2019 年同期前后 5 天日最低温第 10 分位数的日数 | d |
| | 冷昼日数 TX10p | 年日最高温小于 2010—2019 年同期前后 5 天日最高温第 10 分位数的日数 | d |
| | 暖夜日数 TN90p | 年日最低温大于 2010—2019 年同期前后 5 天日最低温第 90 分位数的日数 | d |
| 暖昼日数 TX90p | 年日最高温大于 2010—2019 年同期前后 5 天日最高温第 10 分位数的日数 | d | |

1.4 特征选择

衍生变量易存在多重共线性，增加过拟合的风险。采用递归特征消除法（RFE）寻找四组预测模型的最佳变量组合。该算法的优点是能够指定某个模型进行训练的，例如，使用 RF 进行预测可基于 RFE-RF 指定变量。为验证剩余变量间是否存在多重共线性，计算方差膨胀因子（VIF）。流程为：采用 RFE 方法运行 10 次，将选择次数最多的环境变量作为最佳的变量数据集，随后计算各变量间的方差膨胀因子，忽略 VIF 值大于 10 的变量。

RFE 算法通过 Python scikit-learn 库中的 RFE 包实现，计算方差膨胀因子通过 SPSS 软件实现。

1.5 预测模型

RF 是一种基于树状结构的集成学习方法^[3]。利用训练集构建决策树，然后组成多棵决策树进行预测，最终预测结果基于所有决策树投票确定。参数直接影响预测模型的精度，RF 模型中 $n_estimators$ 和 max_depth 两个参数最为重要^[20]，分别代表决策树的数量和决策树的最大深度。采用网格搜索法选择最佳参数，得到最优参数用来建立模型。最终筛

选出最优参数模型 1 为 ($n_estimators=500$, $max_depth=7$)、模型 2 为 ($n_estimators=200$, $max_depth=9$)、模型 3 为 ($n_estimators=200$, $max_depth=10$) 和模型 4 为 ($n_estimators=100$, $max_depth=11$)。

采用普通克里格 (OK) 残差修正。根据已知点位数据进行半方差函数分析拟合预测残差的最优插值模型, 利用最优插值模型生成残差分布图并与预测结果进行叠加求和, 以达到提高模型精度的目的。

RF 模型通过 Python scikit-learn 库中 Random Forest Regressor 包实现, OK 插值通过 ArcGIS 实现, 半方差函数分析通过 GS+ 实现。

1.6 模型精度评估

为衡量不同类型变量组合下模型的表现, 采用留一交叉验证评估模型性能。采用决定系数 (R^2)、平均绝对误差 (MAE) 和均方根误差 (RMSE) 3 个模型精度指标评价模型预测结果。其中, MAE、RMSE 越小, R^2 越大, 预测精确度越高。

为评估预测模型的稳定性, 利用标准差 (Standard deviation, SD) 的空间变化来表示预测的不确定性^[3]。基于环境变量的最佳组合预测生成 100 幅 SOC 图, 计算每个栅格像素的标准差。

2 结果

2.1 土壤有机碳的统计特征

上高县耕地土壤 SOC 的描述性统计结果表明, 土壤 SOC 含量最小值为 $6.11 \text{ g}\cdot\text{kg}^{-1}$, 最大值为 $30.65 \text{ g}\cdot\text{kg}^{-1}$, 平均值为 $18.79 \text{ g}\cdot\text{kg}^{-1}$, 标准差为 $5.28 \text{ g}\cdot\text{kg}^{-1}$, 偏度系数为 0.30, 峰度系数为 -0.78, 变异系数为 28.10%, 属于中等程度的空间变异。

2.2 极端气候变量与耕地有机碳的相关性

按照特征量的极值、特征量超某阈值的天数、特征量的持续天数分类, 挑选部分变量进行 Pearson 相关性分析。年尺度视角下, 研究区耕地 SOC 含量与极端降雨事件的极值、天数及持续天数 (CWD 0.233^* 、R95p 0.288^{**} 、Rx5day 0.094), 极端高温事件的极值、持续天数 (TXx 0.342^{**} 、WSDI 0.130) 呈正相关; 与极端低温事件的极值、持续天数 (TNx -0.142 、CSDI -0.077) 呈负相关; 与极端高、低温事件的天数 (SU 0.275^{**} 、TR -0.087 、ID0 -0.192^* 、FD0 0.119) 呈不确定性。

将气候事件分类后, 可以发现热、冷事件的天数并不总为正、负相关, 说明极端气候指数综合考虑了极端气候事件的强度和持续时间, 当气候超过某阈值时, 研究区耕地 SOC 呈反向相关。总体而言, 极端气候事件与研究区耕地 SOC 具有显著的相关性, 说明极端气候指数能够很好地反映气候对研究区耕地 SOC 的响应机制, 极端气候的响应方式及强度直接或间接 (例如植被覆盖) 地影响耕地 SOC 含量。当研究区气候多热湿润时, 耕地 SOC 含量越高。

2.3 模型精度与不确定性评价

为消除变量间的多重共线性, 采用 RFE 筛选出模型所使用的变量。通过共线性分析, 四组模型的方差膨胀因子均小于 10, 说明 RFE 方法能够有效剔除变量间的多重共线性。基于不同变量组合的 RF 模型在预测 SOC 含量方面的性能。模型 1 中 R^2 为 0.30, MAE 为 $4.58 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $5.32 \text{ g}\cdot\text{kg}^{-1}$; 模型 2 中 R^2 为 0.35, MAE 为 $4.24 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $4.98 \text{ g}\cdot\text{kg}^{-1}$; 模型 3 中 R^2 为 0.36, MAE 为 $4.21 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $4.95 \text{ g}\cdot\text{kg}^{-1}$; 模型 4 中 R^2 为 0.44, MAE 为 $4.03 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $4.87 \text{ g}\cdot\text{kg}^{-1}$ 。结果表明, 基于不同环境变量类型构建的模型, 预测精度存在差异。例如: 以模型 1 作为基线, 其余三组模型精度得到显著提升。说明时序变量、物候变量和极端气候变量含有未利用信息, 对于 SOC 的有效建模和预测具有一定的价值。

经过残差修正后精度获得进一步提升。模型 1_OK 中 R^2 为 0.42, MAE 为 $2.04 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $2.52 \text{ g}\cdot\text{kg}^{-1}$; 模型 2_OK 中 R^2 为 0.50, MAE 为 $1.99 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $2.43 \text{ g}\cdot\text{kg}^{-1}$; 模型 3_OK 中 R^2 为 0.52, MAE 为 $1.93 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $2.39 \text{ g}\cdot\text{kg}^{-1}$; 模型 4_OK 中 R^2 为 0.57, MAE 为 $1.88 \text{ g}\cdot\text{kg}^{-1}$, RMSE 为 $2.28 \text{ g}\cdot\text{kg}^{-1}$ 。总体来看, 残差修正将 R^2 提升至 0.4 以上, 并且均改善了最小值偏大、最大值偏小的预测区间, 与实测值更为接近。其中, 模型 4 仍可以获得最高的精度, 模型 4_OK, 相较于模型 4 将 R^2 、MAE 和 RMSE 提升了 29.54%、53.35% 和 53.18%; 相较于模型 1 将 R^2 、MAE 和 RMSE 提升了 90.00%、58.95% 和 57.14%。

计算 100 幅 SOC 预测值标准差评估模型预测的不确定性, 如图 2 所示。模型 1 和模型 4 具有最高和最低的不确定性, 模型 2 和模型 3 的精度与不确定性呈负相关, 说明拥有更好精度的模型在制图过

程中并不总是最稳定的模型。不确定性主要出现在研究区的中部、东北部、西南部,可能是因为该地区的 SOC 空间变异性较强,采样点数量较少且缺乏

代表性,所以预测此区域 SOC 含量出现明显波动。因此量化预测的不确定性对未来指导采样具有重要意义。

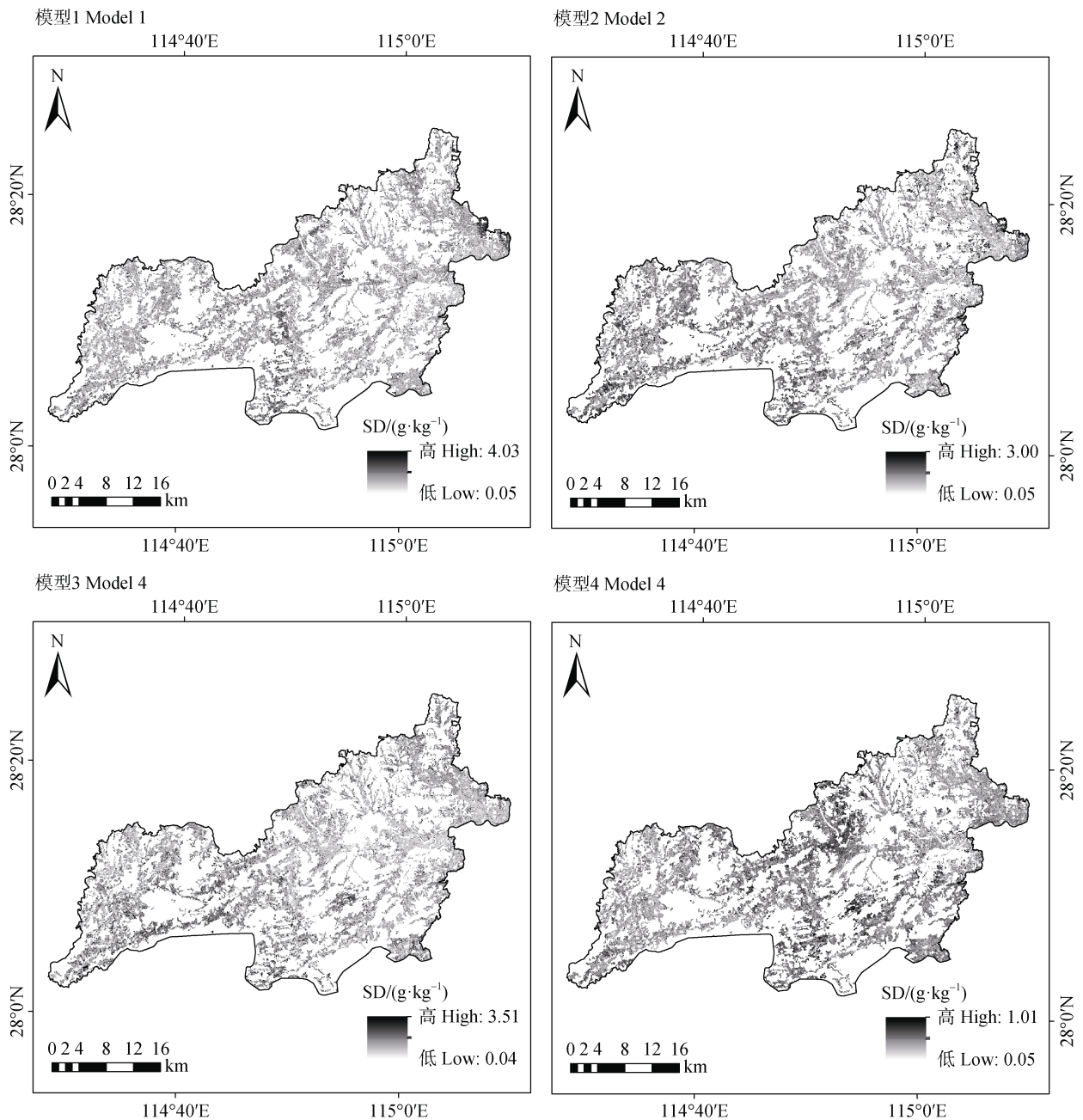


图2 模型预测的不确定性

Fig. 2 Uncertainty in model predictions

2.4 RF 模型环境协变量的重要性

四组模型变量的相对重要性排序,如图3所示。环境协变量的重要性会在不同变量组合下发生变化,此外每组模型的主要协同因子均有不同。以地形变量为标准,在模型1中综合重要性最高,而在

其余模型中重要性基本处于末位。模型2中,时序变量部分变量重要性稍弱,但累积贡献率最高。模型3中,虽NDVI在其他模型中仅一次纳入,但由NDVI时序数据提取的物候变量a3在模型中贡献第二,说明物候变量能够捕捉到植被的特定信息。模

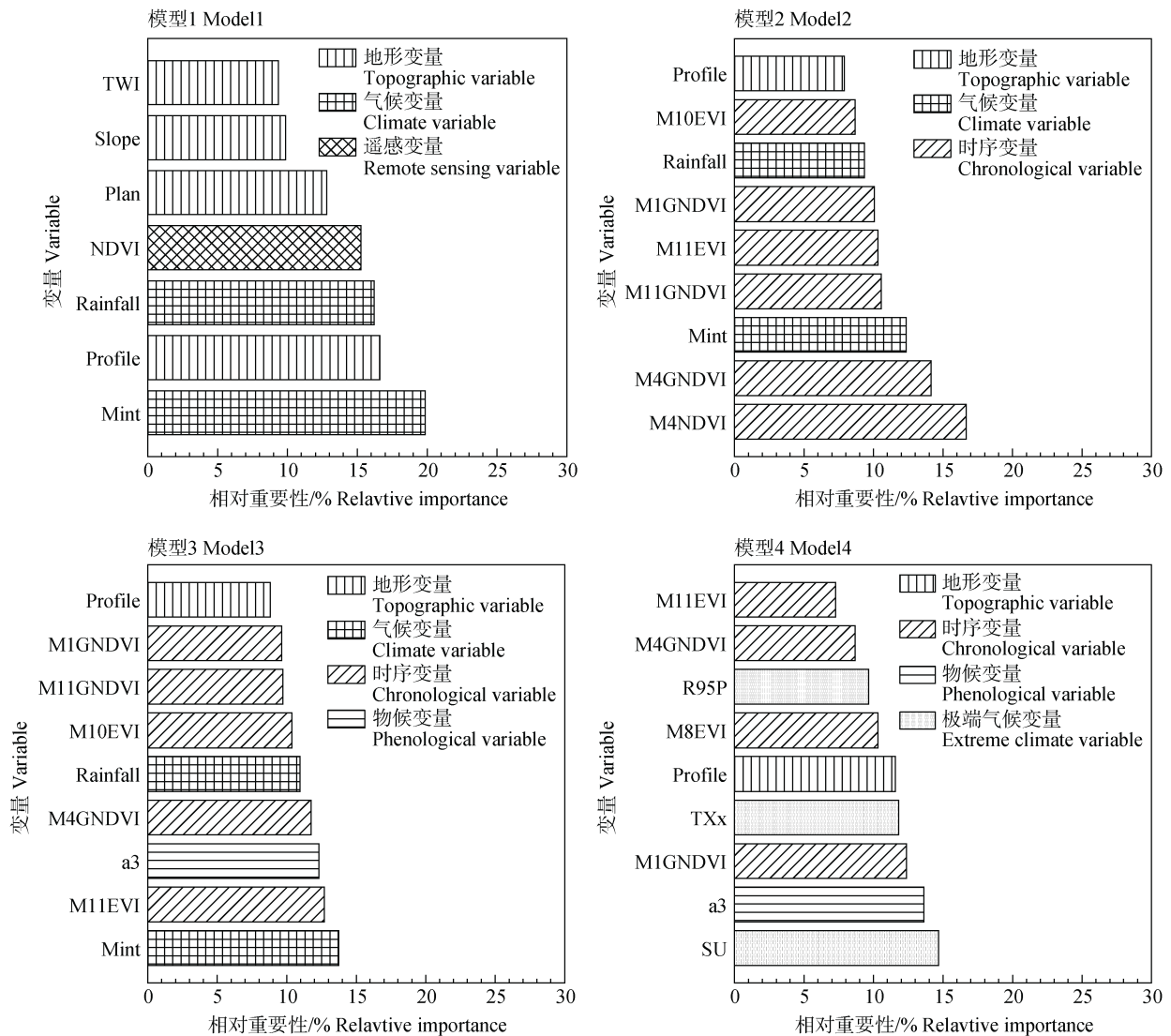


图3 RF模型各变量相对重要性

Fig. 3 Relative importance of variables in the RF model

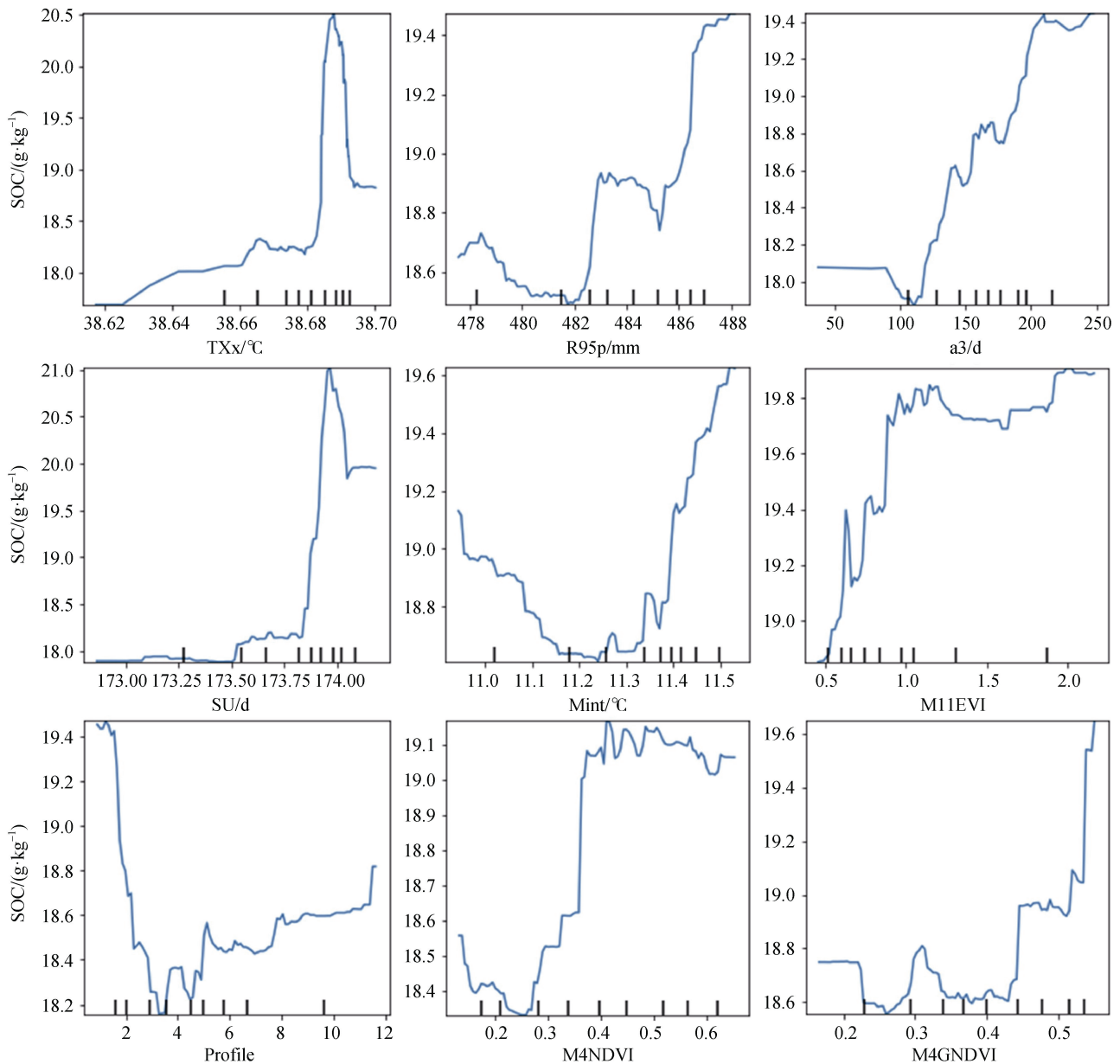
型4中,重要协同因子Mint和Rainfall被极端气候变量所代替,并解释了SOC变化的36.12%,其中SU为单个贡献率最高的因子,是模型的主导因子。表明物候变量和极端气候变量具有预测研究区耕地SOC空间分布的应用潜力。

为探究环境变量对SOC的响应变化关系,绘制SOC与环境变量的局部依赖性图,该曲线代表变量与SOC的变化关系,如图4所示。SOC含量与植被指数、a3、R95p总体呈正相关关系;与Mint、Profile、SU和TXx表现为波动变化趋势,部分变量当达到临界值时,变量与SOC含量之间显著响应,例如TXx超过38.68, SOC含量急剧下降。但Pearson相关性分析显示, SU、TXx与SOC含量呈正相关关

系,说明变量与SOC之间并非只是简单的线性关系,而是存在复杂的非线性关系,且响应关系差异较大。

2.5 土壤有机碳空间分布

将筛选出来的最优协变量组合及最优参数运行100次,绘制出100张SOC结果进行平均,生成最终的SOC空间分布图,如图5所示。模型1至模型4预测研究区SOC的变化范围分别为10.00~26.71、12.90~28.34、12.92~28.34和11.54~28.68 $\text{g}\cdot\text{kg}^{-1}$, 均值分别为 19.00 ± 0.71 、 21.44 ± 0.67 、 20.74 ± 0.69 和 $20.64\pm 0.65 \text{ g}\cdot\text{kg}^{-1}$ 。四组模型预测的整体趋势相似,均呈现中部高值区,东北部、中西部低值区。研究区中部地形复杂,耕地景观格局存在多样化,因此



注: SOC: 土壤有机碳; TXx: 年日最高温的最大值; R95p: 年强降雨量; a3: 季节长度; SU: 夏日日数; Mint: 年最低气温; M11EVI: 11 月份增强型植被指数; Profile: 剖面曲率; M4NDVI: 4 月份归一化植被指数; M4GNDVI: 4 月份归一化绿度指数。Note: SOC: Soil Organic Carbon; TXx: Annual maximum daily maximum temperature; R95p: Intense annual rainfall; a3: Season length; SU: Summer days; Mint: Annual minimum temperature; M11EVI: Enhanced vegetation index in November; Profile: Section curvature; M4NDVI: Normalized vegetation index for April; M4GNDVI: Normalized Greenness Index for April.

图 4 土壤 SOC 含量与各环境协变量的响应曲线

Fig. 4 Response of soil total nitrogen content to the variation of each environmental predictive factor

微生物群体的多样性可能很高, 多物种能够有效地将 C 从地表输送至土壤当中, 并且该区域人为扰动较弱, 具有较好的连通性, 连通的景观有利于物种和属性的移动, 因而 SOC 含量较高^[21]。研究区东北部和西部 SOC 含量偏低, 可能是该区域地势平坦且

耕地集中连片, 农业生产强度高, 频繁耕作加剧了对农业生态系统的扰动, 导致土壤的生物种群遭受破坏, 致使 SOC 含量较低。四组模型虽然预测的整体趋势相似, 但值域区间范围和空间变异存在明显差异。模型 4 相较其他模型, 在研究区中部高值区、

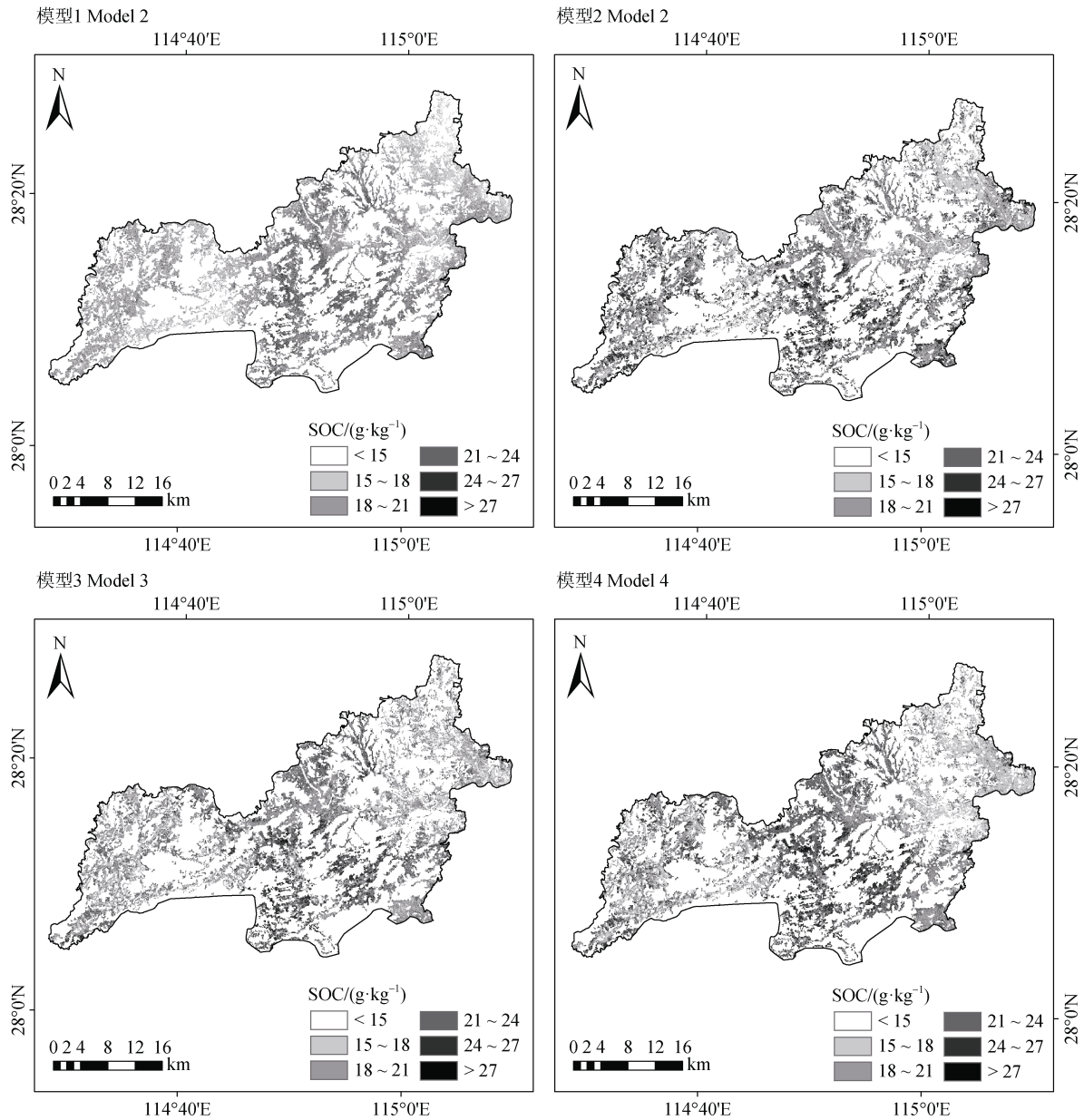


图5 不同模型下研究区土壤 SOC 空间分布

Fig. 5 Spatial distribution of SOC by different model

中西部低值区，分级区间与实测值更为接近，且拥有最宽的预测区间。说明模型 4 无论在模型精度还是空间变异性表达上均为最优模型，证明 RF 模型预测研究区耕地 SOC 空间分布格局的可行性和合理性。

3 讨论

3.1 影响土壤有机碳分布的主要环境变量

通过变量贡献率分析显示，遥感变量 GNDVI、EVI 在模型中具有较好的贡献。因为作物地上部及

根系凋落物是土壤碳输入的重要部分，当在高生物量条件下或生物物理参数达到极值时，EVI 和 GNDVI 能够有效监测作物状态及其覆盖度^[22-23]，并且计算两个指数中的近红外波段也能有效反映土壤水分和作物健康^[24]，因此植被指数通过反映植被覆盖度、土壤水分等差异间接解释 SOC 空间变化。通常植被覆盖度较高、土壤水分较多处，SOC 含量较高^[25]，所以植被指数与 SOC 呈正相关关系。值得注意的是，所选多时相遥感变量中春季、秋季占绝大部分，春季可能是此时气温回暖，降雨集中，使得

土壤温度、含水量提升,影响净初级生产力以及微生物代谢^[26-27],进而影响 SOC 含量,此外正值农耕时节,施肥量等因素影响植被生长速率,通常 SOC 含量累积的越多生长速率越快^[18]。秋季可能是间接反映秸秆还田状况,影响土壤碳输入,从而反映 SOC 含量。总体而言,相较单时相,多时相数据可以更好地反馈地表动态变化^[28-29],为 scorpan 模型提供潜在信息。

地形变量中, Profile 虽在四组模型中贡献率较低,但对预测研究区 SOC 含量是不可或缺的变量。从响应曲线中可以看出,当 Profile < 8 时与 SOC 呈负相关关系,可能是因为地形会对区域小气候特征产生一定影响,海拔随曲率提升逐渐降低,气温、降雨也随之增加或减少,致使微生物活性提升,土壤碳矿化率变快, SOC 含量下降;当 Profile > 9 时呈正相关关系,是因为 Profile 能够影响排水的流速并影响侵蚀和沉积^[30],通常地势低洼处土壤水分含量较多,植物生长茂盛,归还量高,因而 SOC 含量较高。由此也解释了气候变量 Mint 与 SOC 间先负后正的响应特性。

物候变量中, a3 描述生长季结束日期和生长季开始日期的差值。朱丹瑶^[12]研究表示温度与降水量的增加可为植被提供有效的辐射能、热量和水分,致使植被物候期提前或推迟。此时作物的生长特性体现水热条件的差异,水热条件的改变可以影响农田微生物结构功能进而影响 SOC 组分特征^[31]。这与以往研究得出相似结果^[32],说明植被生长周期特征的生物学机制包含 SOC 含量的重要信息。

气候变量在土壤属性制图中的应用已有探索,但关于极端气候变量与土壤属性反映特性的研究较为缺失。通过贡献率发现,极端气候的贡献率显著大于年均气候,这表明 SOC 含量受极端气候的变化影响更大。其中极端高温事件 SU、TXx 在模型中均存在积极作用,因为研究区降水充沛地区,高温能够促进水分蒸发,防止植被水分过多,避免因土壤水分过多造成缺氧,根系呼吸作用减弱,产生多种有毒物质,造成土壤养分流失^[12,33-34],进而影响 SOC 含量。但由响应曲线可知,当极端高温事件的强度、持续时间达到临界值, SOC 含量急剧下降,可能是水分蒸发,致使真菌以及其他微生物活性提高,繁殖加快,矿化率加速,导致 SOC 含量降低。极端降水事件 R95p 与 SOC 含量呈正相关关系,有研究表

示^[35]在植被稀疏等地区,短时强降水极易加剧水土流失,不利于植被生长,而研究区强降雨通常发生在春夏两季,此时作物覆盖度较高,能够削减雨滴动能,减少了雨滴对地面的打击,降低对表土层的冲击侵蚀作用^[36]。与 SOC 呈正相关关系,可能是因为短时强降雨在土壤的中上层形成滞水,导致土壤中微生物厌氧类菌种和数量增加,使 SOC 的腐殖化程度降低,难氧化有机碳的含量增加, SOC 稳定性增大,所以 SOC 的含量较高^[37]。值得注意的是,以往研究^[5]发现部分年均气候与 SOC 含量呈负相关,而某些极端气候事件对 SOC 含量却有促进作用,例如极端高温、降雨事件。说明基于年平均气温进行预测,会损失强度、持续时间等气候特征信息,当气候特征不断完善,年均气候在模型中的贡献率就会逐渐下降。

3.2 不同类型环境变量组合下的模型表现

精度分析结果表明, SOC 空间分布的预测精度受模型及环境变量的组合影响较大。此前已有研究证明,随机森林残差克里格(RFRK)要优于部分数字土壤制图方法。Szatmári 和 Pásztor^[38]预测匈牙利的 SOC 储量,发现 RFRK 的准确度最高,优于 OK、序列高斯模拟法和分位数回归森林。Guo 等^[9]预测橡胶种植园 SOM 的空间分布发现, RFRK 相较逐步线性回归、RF 拥有更低的预测误差(ME、MAE 和 RMSE)和更高的 R^2 , 并表明 RFRK 对捕捉土壤属性与环境因素之间的非线性和层次关系具有优势。但采用 OK 插值可能会出现不确定性因素(比如平滑效应等)导致精度降低,所以未来还需深入探讨残差克里格法在特定情况(比如随机因素为主要作用的研究区等)是否会产生更大的误差。

此前研究已经证明了时序变量和物候变量预测土壤属性的有效性。Zhou 等^[29]以中国西北部黑河中上游流域地区为研究区;刘焕军等^[39]以黑龙江省农垦总局胜利农场耕地范围为研究区,表明多时相遥感数据能够提升预测模型的精度。He 等^[18]在安徽省宣州市利用物候参数和遥感因子通过 RF 模型进行农田 SOC 空间分布预测研究,结果显示,基于物候参数和遥感因子的模型(R^2 为 0.38)提高了 SOC 的可预测性(R^2 提升了 171%, RMSE 降低了 14%),值得注意的是,其预测变量对 SOC 空间变异的解释程度为 38%低于以往 DSM 研究精度^[40],因为耕地常处农业生产较便利的低海拔区域,致使所选变量

差异不显著，导致耕地较其他土地利用更难预测，精度更低，但文中表示在同一地区的农田 SOC 预测，预测精度略高于类似研究。由于缺少针对极端气候变量作为输入集的土壤属性制图研究案例，无法得到有效的精度评价结果。本研究显示，极端气候变量在 RF 模型中累积解释度最高，且随极端气候变量的加入能够显著提升预测精度。研究表明^[41]小尺度内大气候对土壤形成基本为均质，更多受小气候影响，因此极端气候在更大尺度下应该能够得到更好的解释度。

3.3 预测的不确定性及其影响因素

结果的不确定性可归因于数据、模型和环境协变量等因素^[42]。用于模型训练的样本应代表地表生物物理特征的异质性^[43]，但外业时可能会遇到无法采样的情况，这会降低样本的代表性和数量，影响模型在预测时的稳定性。此外，不同的模型也会对不确定性造成显著影响，Fatholouloumi 等^[43]和 Zeraatpisheh 等^[44]对比 RF、Cubist 发现，RF 显示出更低的不确定性，而 Cubist 可能存在过拟合的情况，所以拥有更好的模型精度。本研究显示，相同预测模型下，包含较高预测潜力的环境协变量可以降低部分结果的不确定性。

影响 SOC 制图精度主要有四个部分。第一，未考虑人为活动因素的影响，有研究表明耕地 SOC 受人为影响因素较大，比如耕作制度、施肥方式等因素^[45]。第二，未考虑灌排条件的影响，极端气候作用强度可能会受到灌排条件的部分影响，比如极端降雨事件。第三，研究区春夏季降水充沛，导致部分影像缺失，对提取物候变量和时序变量造成不确定影响。第四，所用的气候数据是来自国家气象中心网站提供的全国地面气候资料日值数据集，其站点密度较稀疏，导致数据精细化较低。

4 结 论

本文以江西省上高县为研究区域，采用随机森林模型，选取遥感数据、DEM 衍生变量、物候参数、气候特征因子等作为环境协变量引入模型中，并用普通克里格对模型结果进行残差修正，最后对比不同类型变量组合下模型的预测效果及预测精度。结果表明，时序变量、物候变量及极端气候变量能够

改善 SOC 预测的性能，结合时序变量、物候变量、极端气候变量、地形变量和残差的组合拥有最高的预测精度，相较于地形变量、遥感变量和气候变量的组合，将 R^2 、MAE 和 RMSE 提升了 90.00%、58.95% 和 57.14%，且预测 SOC 分布更接近实测值。其中，SU、a3 和 TXx 是影响研究区耕地 SOC 分布的重要变量。因此，本文认为物候变量和极端气候变量具有较好的应用前景，未来还需验证极端气候变量作为环境变量在不同土地利用、大尺度研究区下预测土壤属性的有效性。

参考文献 (References)

- [1] Lal R. Soil carbon sequestration impacts on global climate change and food security[J]. *Science*, 2004, 304 (5677): 1623—1627.
- [2] McBratney A, Field D J, Koch A. The dimensions of soil security[J]. *Geoderma*, 2014, 213: 203—213.
- [3] Zhou Y, Zhao X M, Guo X. Prediction of total nitrogen distribution in surface soil based on multi-source auxiliary variables and random forest approach[J]. *Acta Pedologica Sinica*, 2022, 59 (2): 451—460. [周洋, 赵小敏, 郭熙. 基于多源辅助变量和随机森林模型的表层土壤全氮分布预测[J]. *土壤学报*, 2022, 59 (2): 451—460.]
- [4] Wei Y C, Lu X L, Zhu C D, et al. High-resolution digital mapping of soil organic carbon at small watershed scale using landform element classification and assisted remote sensing information[J]. *Acta Pedologica Sinica*, 2023, 60 (1): 63—76. [魏宇宸, 卢晓丽, 朱昌达, 等. 基于地形与遥感辅助信息的小流域尺度高分辨率有机碳空间分布预测研究[J]. *土壤学报*, 2023, 60 (1): 63—76.]
- [5] Yuan Y Q, Chen H Y, Zhang L M, et al. Prediction of spatial distribution of soil organic carbon in farmland based on multi-variables and random forest algorithm—A case study of a subtropical complex geomorphic region in Fujian as an example[J]. *Acta Pedologica Sinica*, 2021, 58 (4): 887—899. [袁玉琦, 陈瀚阅, 张黎明, 等. 基于多变量与 RF 算法的耕地土壤有机碳空间预测研究——以福建亚热带复杂地貌区为例[J]. *土壤学报*, 2021, 58 (4): 887—899.]
- [6] Hao C K, Sun X L, Wang H L. Application of generalized linear geostatistical model for digital soil mapping in a typical subtropical hilly area[J]. *Acta Pedologica Sinica*, 2023, 60 (4): 993—1006. [郝辰恺, 孙孝林, 王会利. 广义线性地统计模型在典型亚热带丘陵区数字土壤制图中的应用[J]. *土壤学报*, 2023, 60 (4): 993—1006.]
- [7] Jiang S P, Zhang H Z, Zhang R L, et al. Research on spatial distribution of soil organic matter in Hainan Island

- based on three spatial prediction models[J]. *Acta Pedologica Sinica*, 2018, 55 (4): 1007—1017. [姜赛平, 张怀志, 张认连, 等. 基于三种空间预测模型的海南岛土壤有机质空间分布研究[J]. *土壤学报*, 2018, 55 (4): 1007—1017.]
- [8] Zhang Y, Sui B, Shen H O, et al. Mapping stocks of soil total nitrogen using remote sensing data: A comparison of random forest models with different predictors[J]. *Computers and Electronics in Agriculture*, 2019, 160: 23—30.
- [9] Guo P T, Li M F, Luo W, et al. Digital mapping of soil organic matter for rubber plantation at regional scale: An application of random forest plus residuals kriging approach[J]. *Geoderma*, 2015, 237: 49—59.
- [10] Araya S, Lyle G, Lewis M, et al. Phenologic metrics derived from MODIS NDVI as indicators for plant available water-holding capacity[J]. *Ecological Indicators*, 2016, 60: 1263—1272.
- [11] Testa S, Soudani K, Boschetti L, et al. MODIS-derived EVI, NDVI and WDRVI time series to estimate phenological metrics in French deciduous forests[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2018, 64: 132—144.
- [12] Zhu D Y. Land surface phenology characteristics and its response to extreme climate in Xiaoxing'an mountains[D]. Harbin: Harbin Normal University, 2021. [朱丹瑶. 小兴安岭地表物候特征及其对极端气候的响应[D]. 哈尔滨: 哈尔滨师范大学, 2021.]
- [13] Donat M G, Alexander L V, Yang H, et al. Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset[J]. *Journal of Geophysical Research: Atmospheres*, 2013, 118 (5): 2098—2118.
- [14] Lin E D, Xu Y L, Jiang J H, et al. National assessment report of climate change (II): Climate change impacts and adaptation[J]. *Advances in Climate Change Research*, 2006, 2 (2): 51—56. [林而达, 许吟隆, 蒋金荷, 等. 气候变化国家评估报告 (II): 气候变化的影响与适应[J]. *气候变化研究进展*, 2006, 2 (2): 51—56.]
- [15] Bao S D. Soil and agricultural chemistry analysis[M]. Beijing: China Agriculture Press, 1999. [鲍士旦. 土壤农化分析[M]. 北京: 中国农业出版社, 1999.]
- [16] Bian J H, Li A N, Song M Q, et al. Reconstruction of NDVI time-series datasets of MODIS based on Savitzky-Golay filter[J]. *Journal of Remote Sensing*, 2010, 14 (4): 725—741. [边金虎, 李爱农, 宋孟强, 等. MODIS 植被指数时间序列 Savitzky-Golay 滤波算法重构[J]. *遥感学报*, 2010, 14 (4): 725—741.]
- [17] Zu J X, Yang J. Temporal variation of vegetation phenology in northeastern China[J]. *Acta Ecologica Sinica*, 2016, 36 (7): 2015—2023. [祖佳星, 杨健. 东北地区植被物候时序变化[J]. *生态学报*, 2016, 36 (7): 2015—2023.]
- [18] He X, Yang L, Li A, et al. Soil organic carbon prediction using phenological parameters and remote sensing variables generated from Sentinel-2 images[J]. *Catena*, 2021, 205: 105442.
- [19] Cai F, Yu G R, Zhu Q L, et al. Comparison of precisions between spatial methods of climatic factors: A case study on mean air temperature[J]. *Resources Science*, 2005, 27 (5): 173—179. [蔡福, 于贵瑞, 祝青林, 等. 气象要素空间化方法精度的比较研究——以平均气温为例[J]. *资源科学*, 2005, 27 (5): 173—179.]
- [20] Ließ M, Glaser B, Huwe B. Uncertainty in the spatial prediction of soil texture: Comparison of regression tree and Random Forest models [J]. *Geoderma*, 2012, 170: 70—79.
- [21] Zhang J. Effects of farmland landscape pattern on soil organic carbon in lower Liaohe plain[D]. Shenyang: Shenyang Agricultural University, 2020. [张俊. 下辽河平原农田景观格局对土壤有机碳的影响[D]. 沈阳: 沈阳农业大学, 2020.]
- [22] Lillesand T, Kiefer R W, Chipman J. Remote sensing and image interpretation[M]. John Wiley & Sons, 2015.
- [23] Sabetizade M, Gorji M, Roudier P, et al. Combination of MIR spectroscopy and environmental covariates to predict soil organic carbon in a semi-arid region[J]. *Catena*, 2021, 196: 104844.
- [24] Aldana-Jague E, Heckrath G, MacDonald A, et al. UAS-based soil carbon mapping using VIS-NIR (480-1000 nm) multi-spectral imaging: Potential and limitations[J]. *Geoderma*, 2016, 275: 55—66.
- [25] Wen L N, Zhou B, Wang L, et al. Analysis of vegetation cover, land utilization and soil moisture in the controlled area of check dam in loess gully region[J]. *Grassland and Turf*, 2021, 41 (3): 41—47. [温丽娜, 周波, 王立, 等. 淤地坝沟道植被覆盖/土地利用和土壤水分的分析[J]. *草原与草坪*, 2021, 41 (3): 41—47.]
- [26] Hu Y G, Wang Z R, Wang Q, et al. Climate change affects soil labile organic carbon fractions in a Tibetan alpine meadow[J]. *Journal of Soils and Sediments*, 2017, 17(2): 326—339.
- [27] Qi R, Li J, Lin Z, et al. Temperature effects on soil organic carbon, soil labile organic carbon fractions, and soil enzyme activities under long-term fertilization regimes[J]. *Applied Soil Ecology*, 2016, 102: 36—45.
- [28] Yang R M, Guo W W. Modelling of soil organic carbon and bulk density in invaded coastal wetlands using Sentinel-1 imagery[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2019, 82: 101906.
- [29] Zhou T, Geng Y, Chen J, et al. Mapping soil organic carbon content using multi-source remote sensing

- variables in the Heihe River Basin in China[J]. *Ecological Indicators*, 2020, 114: 106288.
- [30] Kimerling A J, Buckley A R, Muehrcke P C, et al. *Map use: Reading, analysis, interpretation*[M]. 7th ed. Redlands, California: Esri Press Academic, 2016.
- [31] Zhang P. Spatial mapping of soil organic matter in Jiangsu Province based on hydrothermal conditions[D]. Huainan, Anhui: Anhui University of Science & Technology, 2019. [张平. 基于水热条件的江苏省土壤有机质空间制图[D]. 安徽淮南: 安徽理工大学, 2019.]
- [32] Yang L, Cai Y, Zhang L, et al. A deep learning method to predict soil organic carbon content at a regional scale using satellite-based phenology variables[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2021, 102: 102428.
- [33] Ni M, Zhang X Y, Jiang C, et al. Responses of vegetation to extreme climate events in southwestern China[J]. *Chinese Journal of Plant Ecology*, 2021, 45 (6): 626—640. [倪铭, 张曦月, 姜超, 等. 中国西南部地区植被对极端气候事件的响应[J]. *植物生态学报*, 2021, 45 (6): 626—640.]
- [34] Chen W Z, Huang C J, Wang L C, et al. Climate extremes and their impacts on interannual vegetation variabilities: A case study in Hubei Province of central China[J]. *Remote Sensing*, 2018, 10 (3): 477.
- [35] Wang X, Wang M T, Feng Y, et al. Variation characteristics of normalized difference vegetation index in Northwestern Sichuan Plateau and its response to extreme climate during 2001-2020[J]. *Chinese Journal of Applied Ecology*, 2022, 33 (7): 1957—1965. [王鑫, 王明田, 冯勇, 等. 2001—2020 年川西北高原归一化植被指数演变特征及其对极端气候的响应[J]. *应用生态学报*, 2022, 33 (7): 1957—1965.]
- [36] Hua D W. Investigation and study on soil erosion under extreme sub-rainfall events[J]. *Agriculture and Technology*, 2017, 37 (19): 50—52. [花东文. 极端次降雨事件下的土壤侵蚀调查研究[J]. *农业与技术*, 2017, 37 (19): 50—52.]
- [37] Xu H J. Effects of hydrothermal conditions on soil organic carbon distribution in northern slope of middle Tianshan Mountains[J]. *Journal of Xinjiang University: Natural Science Edition*, 2020, 37 (4): 534—542, 550.
- [38] Szatmári G, Pásztor L. Comparison of various uncertainty modelling approaches based on geostatistics and machine learning algorithms[J]. *Geoderma*, 2019, 337: 1329—1340.
- [39] Liu H J, Zhang M W, Yang H X, et al. Inversion of cultivated soil organic matter content combining multi-spectral remote sensing and random forest algorithm[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2020, 36 (10): 134—140. [刘焕军, 张美薇, 杨昊轩, 等. 多光谱遥感结合随机森林算法反演耕作土壤有机质含量[J]. *农业工程学报*, 2020, 36 (10): 134—140.]
- [40] Odebiri O, Mutanga O, Odindi J. Deep learning-based national scale soil organic carbon mapping with Sentinel-3 data[J]. *Geoderma*, 2022, 411: 115695.
- [41] Liang Z Z. Research and application of high-resolution digital mapping of soil key attributes 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.]
- [42] Oduor C O, Karanja N K, Onwonga R N, et al. Enhancing soil organic carbon, particulate organic carbon and microbial biomass in semi-arid rangeland using pasture enclosures[J]. *BMC Ecology*, 2018, 18 (1): 1—9.
- [43] Fatholoulumi S, Vaezi A R, Alavipanah S K, et al. Improved digital soil mapping with multitemporal remotely sensed satellite data fusion: A case study in Iran[J]. *Science of the Total Environment*, 2020, 721: 137703.
- [44] Zeraatpisheh M, Garosi Y, Owliaie H R, et al. Improving the spatial prediction of soil organic carbon using environmental covariates selection: A comparison of a group of environmental covariates[J]. *Catena*, 2022, 208: 105723.
- [45] Wang Y, Yang D L, Wang L L, et al. Effects of farmland management measures on soil organic carbon turnover and microorganisms[J]. *Journal of Agricultural Resources and Environment*, 2020, 37 (3): 340—352. [汪洋, 杨殿林, 王丽丽, 等. 农田管理措施对土壤有机碳周转及微生物的影响[J]. *农业资源与环境学报*, 2020, 37 (3): 340—352.]

(责任编辑: 檀满枝)