

# 基于 SHAP 方法的东北黑土区土壤质量评价指标最小数据集构建<sup>①</sup>

张超<sup>1,2</sup>, 樊亚男<sup>2</sup>, 刘汉强<sup>2,3</sup>, 曾文彪<sup>2,4</sup>, 姜军<sup>2</sup>, 张凤<sup>2</sup>, 胡文友<sup>2,3\*</sup>, 刘刚<sup>1</sup>

(1 南京信息工程大学环境科学与工程学院, 南京 210044; 2 土壤与农业可持续发展全国重点实验室(中国科学院南京土壤研究所), 南京 211135; 3 中国科学院大学, 北京 100049; 4 河海大学环境学院, 南京 210098)

**摘要:** 针对传统耕地土壤质量评价指标最小数据集(MDS)筛选方法在区域尺度应用中存在的局限性问题, 本文以我国东北黑土区耕地表层(0–20 cm) 土壤为研究对象, 基于随机森林模型和 Shapely 加性解释(SHAP)模型, 提出了适用于区域尺度的耕地土壤质量评价指标 MDS 筛选方法。结果表明: 基于传统的主成分分析法(PCA)筛选出的 MDS 包括有机质、pH、砂粒、有效磷和速效钾; 基于 SHAP 法筛选出的 MDS 包括有机质、pH、黏粒、有效磷和容重。对比结果显示, 基于 SHAP 法构建的 MDS 的土壤质量指数(SQI)与基于全量数据集的 SQI 之间具有更强的相关性( $R^2=0.82$ ), 相比之下, 基于 PCA 法构建的 MDS 的 SQI 与之对应的相关性仅为  $R^2=0.63$ , 表明相较于传统的 PCA 法, SHAP 法在精简土壤指标数量的同时还能够保留更丰富的信息。此外, 由 SHAP 法获得的 SQI 与作物产量之间的相关性更强( $R^2=0.40$ ), 远高于由 PCA 法得出的结果( $R^2=0.15$ ), 表明由 SHAP 法构建的 MDS 更能准确反映研究区域内土壤质量的实际状况。综上所述, 本研究基于 SHAP 法构建的 MDS 和 SQI 模型不仅实用且稳健, 可为东北黑土区耕地土壤质量评估与管理提供强有力的工具。

**关键词:** 黑土; 土壤质量; 土壤肥力; 评价指标; 机器学习

**中图分类号:** S158 **文献标志码:** A

## Construction of Minimum Data Set of Soil Quality Assessment Indicators in Black Soil Region of Northeast China Based on SHAP Method

ZHANG Chao<sup>1,2</sup>, FAN Ya'nan<sup>2</sup>, LIU Hanqiang<sup>2,3</sup>, ZENG Wenbiao<sup>2,4</sup>, JIANG Jun<sup>2</sup>, ZHANG Feng<sup>2</sup>, HU Wenyu<sup>2,3\*</sup>, LIU Gang<sup>1</sup>

(1 School of Environmental Science and Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China; 2 State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 211135, China; 3 University of Chinese Academy of Sciences, Beijing 100049, China; 4 College of Environment, Hohai University, Nanjing 210098, China)

**Abstract:** In response to the limitations of the traditional minimum data set (MDS) screening method for cultivated soil quality assessment indicators in the regional scale application, this study focused on the cultivated surface soil (0–20 cm) in the black soil area of northeastern China, and proposed a MDS screening method for cultivated soil quality assessment applicable to the regional scale based on the random forest and Shapley additive explanations (SHAP) models. The results indicated that the MDS screened by the traditional principal component analysis (PCA) method included organic matter, pH, sand, available phosphorus, and available potassium, whereas MDS screened by the SHAP method included organic matter, pH, clay, available phosphorus, and bulk density. The comparative results showed that the soil quality index (SQI) derived from the MDS constructed using the SHAP method exhibited a stronger correlation with the SQI based on the total dataset ( $R^2=0.82$ ). In contrast, the correlation for the SQI derived from the PCA-based MDS was only  $R^2=0.63$ . This indicates that, compared to the traditional PCA method, the SHAP method can retain more comprehensive information while effectively reducing the number of soil indicators. Furthermore, SQI obtained by SHAP method showed a stronger correlation with crop yield ( $R^2=0.40$ ), significantly higher than the result from

①基金项目: 中国科学院战略性先导科技专项子课题(XDA28010403, XDA28010101)、江苏省科技计划“一带一路”创新合作项目(BZ2023003)和中国科学院南京土壤研究所“十四五”自主部署项目(ISSAS2418)资助。

\* 通信作者(wyhu@issas.ac.cn)

作者简介: 张超(2001—), 男, 安徽宣城人, 硕士研究生, 主要从事区域土壤质量评价研究。E-mail: 202212480132@nuist.edu.cn

PCA method ( $R^2=0.15$ ). This indicated that MDS established by the SHAP method could more accurately reflect the actual soil quality conditions in the study area. In conclusion, MDS and SQI models constructed in this study based on SHAP method are not only practical and robust but also can provide a powerful tool for soil quality assessment and management of cultivated land in the black soil region of Northeast China.

**Key words:** Black soil; Soil quality; Soil fertility; Assessment indicators; Machine learning

土壤质量作为评估土壤条件变化的一个重要概念通常被定义为土壤在生态系统边界内维持生物生产力、保持环境质量和促进动植物健康的能力<sup>[1]</sup>。作为土壤质量定量评估的有效工具,土壤质量指数(SQI)已在不同尺度和区域得到广泛应用<sup>[2]</sup>。然而,构建 SQI 的关键挑战是评价指标的筛选:过多的指标会导致共线性问题并增加评估成本,而过少的指标则难以有效反映土壤质量<sup>[3]</sup>。为此, Larson 和 Pierce<sup>[4]</sup>提出了最小数据集(MDS)的概念,即通过筛选关键指标来简化评估过程。这一方法已成为土壤质量评估的重要环节<sup>[5]</sup>。然而,如何从众多可能的土壤质量指标中挑选出最具代表性的最小数据集来生成能够准确反映特定区域土壤状况的 SQI,仍然是土壤质量评价研究面临的一项重大挑战<sup>[6]</sup>。

近年来,国内外学者在 MDS 筛选方法上进行了大量探索,其中主成分分析(PCA)法因其能够有效识别和保留数据集中的关键信息而得到广泛应用<sup>[7-8]</sup>。然而该方法主要依赖指标间的统计特征进行筛选,难以充分反映指标与土壤功能之间的实际关联。随着机器学习技术的发展,以随机森林为代表的功能导向型方法为 MDS 筛选提供了新的思路。Li 等<sup>[9]</sup>基于作物产量,利用随机森林模型中的特征变量重要性来筛选 MDS,并通过与 PCA 法和相关性分析法的对比研究,证实了该方法在指标筛选中的有效性。然而,机器学习方法虽能捕捉复杂的非线性关系,但其“黑箱”特性限制了结果的可解释性,难以直观揭示输入特征与输出结果之间的内在联系,这在一定程度上制约了其在区域尺度土壤质量评价 MDS 筛选中的应用。针对这些问题,Shapely 加性解释(SHAP)模型作为一种强大的机器学习可解释性工具被引入<sup>[10]</sup>。该模型能够精确量化每个特征对模型输出的贡献,从而精准识别对土壤功能具有显著影响的关键指标。这一特性为构建科学合理的 MDS 提供了新的解决方案。

东北黑土区作为我国重要的商品粮生产基地,在保障国家粮食安全方面发挥着不可替代的作用<sup>[11]</sup>。然而,黑土地长期重用轻养,导致土壤有机质下降、生产性能和生态服务功能降低等土壤退化问题,严重威胁东北黑土区农业生产的可持续性<sup>[12]</sup>。因此,开

展以生产功能为核心的耕地质量评价研究,对科学指导黑土地保护与可持续利用具有重要的理论与实践意义。

基于以上,本研究以生产功能为导向,选取我国东北黑土区耕地作为研究对象,创新性地将随机森林模型与 SHAP 方法相结合,基于作物产量筛选关键土壤属性并构建 MDS,同时通过与传统 PCA 法对比,验证了该方法的有效性和区域适用性,以为黑土地耕地土壤质量评估提供方法学支撑,为制定黑土地保护与可持续利用策略提供科学依据。

## 1 材料与方法

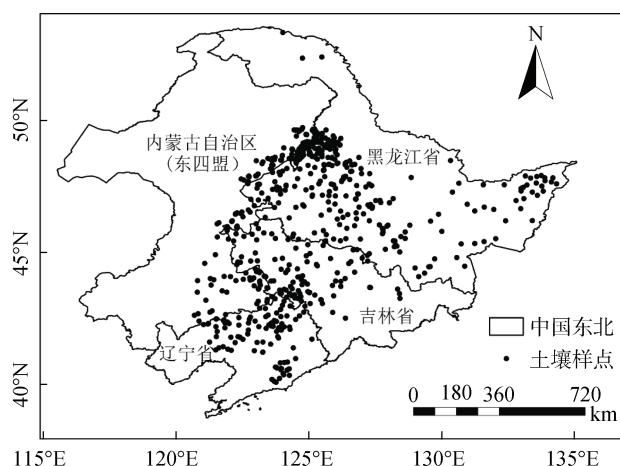
### 1.1 研究区概况

研究区域位于我国东北地区,研究对象为耕地。东北地区(115°30'E~135°20'E, 38°43'N~53°30'N)涵盖黑龙江省、吉林省、辽宁省以及内蒙古东部的呼伦贝尔市、兴安盟、通辽市和赤峰市。该地区被公认为世界四大黑土区之一,区域总面积约为 124.86 万 km<sup>2</sup>,约占我国国土面积的 1/7,其中耕地面积约为 35.84 万 km<sup>2</sup>,约占全国耕地面积的 1/4。作为中国的战略粮仓,东北黑土区贡献了全国粮食总产量的 1/4,主要作物为玉米、大豆和水稻。区内属温带大陆性季风气候,夏季温暖短暂,冬季寒冷干燥,年平均气温介于 -2~8 °C,年平均降水量在 400~700 mm。地形以平原和丘陵为主,地势平缓辽阔,根据中国土壤发生分类系统,土壤类型主要包括黑土、黑钙土、草甸土、白浆土、暗棕壤和棕壤。

### 1.2 土壤样品采集

2022 年 7—9 月在中国东北地区共采集 813 个土壤样本(图 1)。采样前,首先在 ArcGIS 10.6 软件中通过研究区域的卫星照片人工识别出采样田块。然后,在实地调查时使用便携式全球定位系统(GPS)记录土壤采样的精确位置信息。土壤采样点主要分布在耕地,其布设充分考虑了土壤类型、土地利用方式以及空间分布的均匀性和随机性。为了确保样本的可靠性,采样时已避开最近施过肥的田块以及降雨或灌溉后不久的田块;同时选择作物生长阶段相近的时期开展土壤样品采集。采样过程中同步记录环境数据如海

拔高度、地形特征、作物类型、土壤深度等。采用五点采样法和四分法获得代表性田块 0~20 cm 表土土样，并置于塑料自封袋中运回室内。土样自然风干后，进一步除杂、研磨，按分析项目需求过筛、保存、待测。



(该图基于自然资源部标准地图服务网站下载的审图号为 GS(2024)0650 号的标准地图制作，底图无修改)

图 1 研究区域的位置和样点分布

Fig. 1 Location of study area and sampling sites

### 1.3 土壤理化性质测定

本研究对每个土壤样本进行了 11 项土壤性质分析，包括 8 项化学性质(pH、速效氮、有效磷、速效钾、有机质、全氮、全磷和全钾)和 3 项物理性质(颗粒组成、容重和含水量)<sup>[13]</sup>。其中，pH 采用电位法定；有机质采用  $K_2Cr_2O_7-H_2SO_4$  氧化容量法定；全氮采用凯氏定氮法定；全磷采用高氯酸-氢氟酸-盐酸消解，钼锑抗比色法定；全钾采用酸溶-火焰光度法定；速效氮采用碱解扩散法定；有效磷采用  $NaHCO_3$  浸提-钼锑抗比色法定；速效钾采用醋酸铵浸提-火焰光度法定；含水量采用烘干法定；容重通过测定烘干土壤的质量和体积来计算；颗粒组成中黏粒(<0.002 mm)、粉粒(0.002~0.05 mm)和砂粒(0.05~2.0 mm)采用 Horiba Partica 激光粒度分析仪测定。

### 1.4 土壤质量评价方法

**1.4.1 最小数据集构建** 本研究采用两种构建最小数据集(MDS)的方法，分别为主成分分析(PCA)法和 Shapely 加性解释(SHAP)法。其中，PCA 法是一种常用的方法，被用于简化土壤属性的数量，从而选择具有代表性的指标<sup>[14]</sup>。一般来说，当主成分的特征值  $\geq 1$  时，具有高因子载荷的土壤属性更能反映土壤质量的变化。对于特定的主成分，每个变量均有一个权

重或因子载荷，表明其对该主成分的贡献。在每个主成分中，只保留权重高的因子(即绝对权重前 10% 以内的因子)纳入 MDS。当一个主成分中保留了多个变量时，则需要通过相关性分析结果来剔除每个主成分中的冗余属性。具体来说，如果同一主成分中多个指标之间的相关系数大于 0.7，则只选择权重最高的指标<sup>[15]</sup>。基于以上原则，本研究选择了各主成分中因子载荷最高的土壤属性纳入 MDS。SHAP 法是通过解释特征变量对模型预测的贡献度来获取特征重要性进而筛选关键指标。本研究首先构建 11 项土壤属性与产量的随机森林预测模型，随后利用 SHAP 法计算每个土壤属性对于模型预测结果的影响程度，进而确定关键指标。通过选取累计贡献度超过 90% 的特征变量作为初步筛选的关键指标，确保这些指标能够捕捉到大部分模型预测能力<sup>[16]</sup>。最后对选出的指标进行相关性分析，剔除与其他指标高度相关的冗余变量( $r > 0.7$ )，构建最终的 MDS。

**1.4.2 指标评分** 指标评分用于提供明确的指标解释，并解决使用不同数字标度表达的不同指标所造成的偏差。本研究根据各指标对土壤生产力的影响对其进行解释，将所选指标评分分为 S 型、反 S 型和抛物线型 3 种类型。其中，有机质、全氮、全磷、全钾、速效氮、有效磷、速效钾、含水量对土壤生产力有积极影响，应用了 S 型函数；而颗粒组成、pH 和容重则采用了抛物线型函数。本研究中，指标评分阈值的确定参考《第三次全国土壤普查耕地质量等级评价技术规范》。指标评分的上下限值为指标的最大值和最小值。本研究利用非线性评分方程(公式(1))将不同土壤质量指标测定值标准化为 0~1 之间的无量纲值。

$$S_i = \frac{a}{1 + (X/X_0)^b} \quad (1)$$

式中： $S_i$  是变量在 0~1 之间的非线性指标得分值； $a$  是最大分值，在本文中等于 1； $X$  是变量值； $X_0$  是变量的平均值； $b$  是斜率，假定“越多越好”的斜率为 -2.5，“越少越好”的斜率为 2.5<sup>[17]</sup>。

**1.4.3 指标权重确定** 本研究使用公因子方差法来确定土壤质量评价指标的权重。各指标权重值等于该指标的公因子方差与所有指标公因子方差和的比值。计算公式如下：

$$W_i = \frac{C_i}{\sum_i^n C_i} \quad (2)$$

式中： $W_i$  为指标权重； $n$  为土壤指标个数； $C_i$  为公因子方差。

**1.4.4 土壤质量指数计算** 本研究中, SQI 均采用以下加权加性土壤质量指数公式计算:

$$SQI = \sum_{i=1}^n (W_i \cdot S_i) \quad (3)$$

式中: SQI 为土壤质量指数;  $n$  为土壤指标个数;  $W_i$  为指标权重;  $S_i$  为指标得分。

## 1.5 模型训练与评估

**1.5.1 随机森林模型** 随机森林模型是一种数据挖掘技术,它采用集成学习方法来进行数据分类和回归分析。本文使用 Python(版本 3.12)中的 sklearn.ensemble 模块(版本 1.5.2)所提供的 RandomForestRegressor 类来构建和训练随机森林模型,采用 Optuna 框架(版本 4.0.0)通过 100 次试验来优化并确定最佳的超参数组合。所调整的超参数包括:森林中树的数量(n\_estimators)、决策树的最大深度(max\_depth)、进行拆分一个节点所需的最小样本数(min\_samples\_split)、叶子节点所需的最小样本数(min\_samples\_leaf)以及在寻找最佳分割点时考虑的最大特征数(max\_features)。

**1.5.2 环境协变量** 环境变量数据包括地形变量、气候变量、遥感影像数据、土壤母质、土地利用数据和土壤类型图。其中,利用 Google Earth Engine(GEE)从 NASADEM 数据集获取中国东北地区 30 m 分辨率的数字高程模型(DEM)数据,并基于此通过自动化地球科学分析系统(SAGA-GIS)(<http://www.saga-gis.org/>)计算地形变量。气候变量利用 GEE 从 ERA5 陆地月平均数据集中采集,分辨率为 11 132 m,涵盖了 2022 年全年的数据。遥感数据主要包括卫星波段、光谱指数和植被指数,利用 GEE 从 Landsat 卫星下载的 2022 年 6—11 月期间中国东北地区的 Landsat 7 影像计算获取,分辨率为 30 m。土壤母质数据来源于基于 USGS 全球生态系统地图的全球地貌和岩性分类数据集,分辨率为 250 m。土壤类型数据来源于基于中国土壤发生分类系统的土类图,分辨率为 250 m。土地利用数据来源于中国 Landsat 衍生年度土地覆盖产品,分辨率为 30 m。所有环境协变量均通过重采样被标准化成投影坐标系,分辨率为 250 m。在重采样过程中,对于分类变量采用最邻近插值法,而对于连续变量则采用双线性插值法。

**1.5.3 产量预测模型** 作物高产表明土壤质量良好,本研究使用作物产量来验证不同 SQI 的有效性。研究区 2022 年的作物产量数据通过反演全球历史产量数据集(GDHYv1.2+v1.3) (<https://doi.pangaea.de/10.1594/PANGAEA.909132>)获得<sup>[18]</sup>。GDHY 提供了 1981—

2016 年期间 0.5°网格单元中全球主要农作物(玉米、水稻、小麦和大豆)产量估算的年度时间序列数据。本研究使用 1981—2016 年的环境变量数据与作物产量数据来建立随机森林模型。除地形、母质及土壤类型外,其余环境变量皆为时间序列数据。本研究提出了两个关键假设:一是时间序列稳定性,即认为作物产量随时间变化表现出一定程度的稳定性和可预测性,从而可以基于历史数据推测未来或缺失年份的产量;二是空间一致性,即假设在具有相似环境条件(如气候和土壤状况)的区域,作物产量趋势相似。为此,本研究首先通过重采样调整了所有环境变量的分辨率以匹配原始产量栅格;接着,为玉米、水稻和大豆 3 种作物开发随机森林模型;之后,利用 2022 年的环境变量数据,在 250 m 分辨率上进行了产量预测。产量预测结果分别使用市级和县级的统计年鉴数据进行验证;最后,对 3 种作物的产量利用以下公式进行归一化处理,求平均值,获得相对产量。

$$RY = \frac{Y}{Y_{\max}} \quad (4)$$

式中: RY 为相对产量;  $Y$  为作物产量;  $Y_{\max}$  为作物最大产量。本研究采用 RY 作为生产能力的综合指标,并采用线性回归模型来验证 SQI。

**1.5.4 模型评估方法** 本研究采用 10 折交叉验证法来评估产量预测模型的性能。该算法重复 10 次,确保每个子集轮流作为验证集一次。最后,这 10 次的结果被平均化,以生成一个单一的准确度指标。通过计算决定系数( $R^2$ )、平均绝对误差(MAE)、均方根误差(RMSE)来评估所有模型的精度。其中,  $R^2$  表示在因变量的总变异中,模型所能解释的比例; MAE 表示预测值与真实值之间绝对差异的平均值; RMSE 表示预测值与真实值之间平方差的平均值的平方根。具体的计算公式如下:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (5)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

式中:  $\hat{y}_i$  和  $y_i$  分别表示单个预测值和相应的实际值;  $n$  表示数据点总数;  $\bar{y}$  表示所有实际值的平均值。其中, MAE 和 RMSE 越低,同时  $R^2$  越高,则模型性能越好。为了确保模型评估的稳健性,本研究执行了

100 次重复的 10 折交叉验证，然后计算上述 3 个精度评价指标的平均值以全面衡量模型的性能。

### 1.6 数据处理与统计分析

所有数据通过 Excel 2016 整理、汇总，统计分析使用 Python 和 Origin 2022 进行。

## 2 结果与分析

### 2.1 土壤属性描述性统计

由表 1 可知，研究区土壤 pH 范围为 4.04~8.87，均值为 6.13，表明土壤呈微酸性。土壤有机质的平均含量为 37.03 g/kg，范围为 2.69~94.81g/kg。速效氮、有效磷、速效钾的均值分别为 154.79、45.05、184.17 mg/kg。全氮、全磷、全钾的均值分别为 1.90、

0.73、26.13 g/kg。颗粒组成分析显示，粉粒占主导地位，平均含量为 670.26 g/kg，砂粒和黏粒的平均含量分别为 263.86 g/kg 和 68.62 g/kg。土壤容重的均值为 1.32 g/cm<sup>3</sup>，含水量的均值为 215.83 mL/kg。从各属性的变异程度来看，变异性较高的属性包括速效养分中的速效氮(44.03%)、有效磷(75.16%)和速效钾(47.42%)，全量养分中的全氮(45.56%)和全磷(44.49%)，颗粒组成中的砂粒(64.03%)和黏粒(40.98%)以及土壤有机质(51.08%)。相比之下，容重和全钾显示出较低的变异性，变异系数分别为 8.88% 和 11.96%。粉粒、含水量和 pH 的变异系数分别为 21.12%、29.84% 和 19.46%，显示出中等变异性。

表 1 土壤属性描述性统计  
Table 1 Descriptive statistics of soil properties

土壤指标	最小值	最大值	均值	标准差	变异系数(%)
砂粒(g/kg)	20.30	746.40	263.86	168.95	64.03
粉粒(g/kg)	256.80	881.90	670.26	141.53	21.12
黏粒(g/kg)	5.60	145.70	68.62	28.12	40.98
容重(g/cm <sup>3</sup> )	1.03	1.61	1.32	0.12	8.88
含水量(mL/kg)	43.62	387.90	215.83	64.41	29.84
pH	4.04	8.87	6.13	1.19	19.46
有机质(g/kg)	2.69	94.81	37.03	18.92	51.08
速效氮(mg/kg)	22.60	347.10	154.79	68.16	44.03
有效磷(mg/kg)	0.50	144.20	45.05	33.86	75.16
速效钾(mg/kg)	30.70	433.30	184.17	87.34	47.42
全氮(g/kg)	0.17	4.52	1.90	0.87	45.56
全磷(g/kg)	0.11	1.62	0.73	0.32	44.49
全钾(g/kg)	18.90	35.00	26.13	3.12	11.96

斯皮尔曼相关分析表明，速效氮、有效磷、速效钾、有机质、全氮、全磷、含水量、粉粒和黏粒之间均呈显著正相关，而与砂粒之间均呈显著负相关(图 2)。pH、全钾、容重、砂粒之间呈显著正相关，但与速效氮、有机质、全氮、全磷、含水量、粉粒和黏粒呈显著负相关。这表明大多数指标之间存在很强的共线性，因此需要对指标进行筛选，以尽量减少冗余。

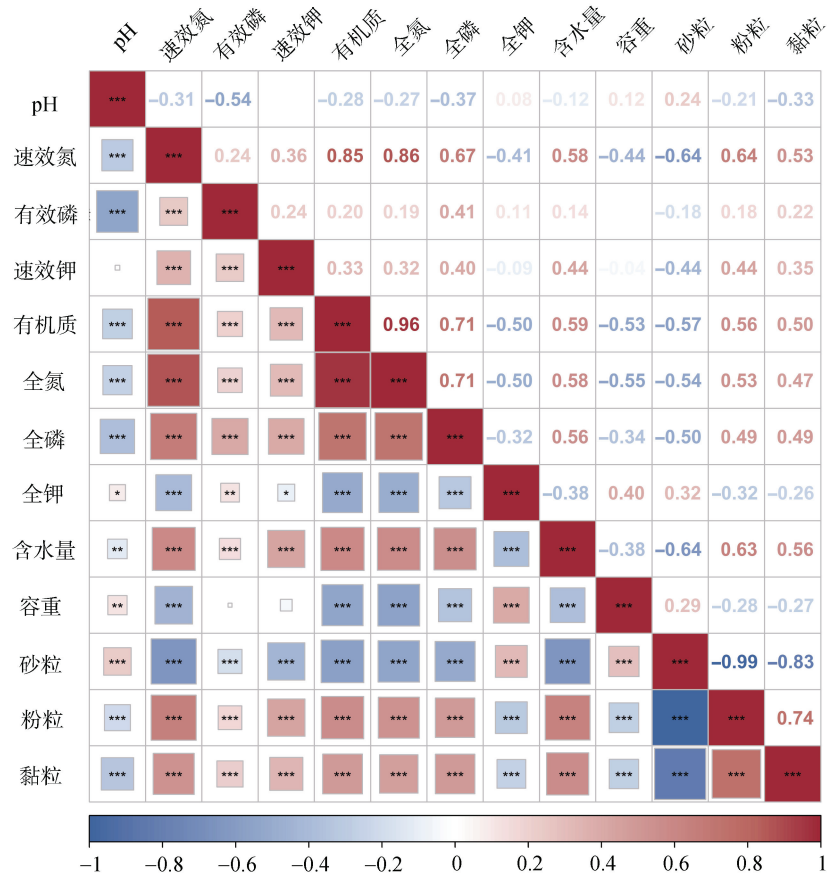
### 2.2 基于 PCA 法的最小数据集筛选

PCA 结果显示，有 3 个主成分(PC1、PC2、PC3)的特征值 $\geq 1$ ，且 PC4 的特征值较高(0.90)，这些主成分共解释了原始数据集中约 79% 的方差(表 2)。其中，PC1 解释了 49.426% 的方差，而 PC2、PC3 和 PC4 分别解释了 11.960%、10.327% 和 6.955% 的方差。在 PC1 中，有机质和全氮的载荷值均位于最高

载荷值的前 10%。尽管全氮的载荷值相对较高，但考虑到有机质对土壤质量的重要性，本研究仍然选择有机质来代表 PC1 入选 MDS。在 PC2 中，砂粒、粉粒和黏粒均在最高因子载荷 10% 以内，这些指标主要与土壤颗粒组成有关。由于砂粒的载荷值最高，且与粉粒( $r=-0.99$ )和黏粒( $r=-0.83$ )的相关性较高，本研究只保留砂粒来代表 PC2 入选 MDS。在 PC3 中，pH 和有效磷均在最高因子载荷值 10% 以内。由于 pH 和有效磷的相关性较小( $r=0.54<0.7$ )，二者均入选 MDS。速效钾是 PC4 中权重最高的指标，因此入选 MDS。最终，基于 PCA 法的 MDS 包括有机质、pH、砂粒、有效磷和速效钾 5 个指标。

### 2.3 基于 SHAP 法的最小数据集筛选

对产量预测模型进行 100 次重复的 10 折交叉验证结果显示，玉米、水稻和大豆的产量预测模型  $R^2$



(图中\*、\*\*、\*\*\*分别表示相关性达  $P < 0.05$ 、 $P < 0.01$ 、 $P < 0.001$  显著水平，方框的大小和颜色的深浅反映了指标之间相关性的强弱)

图 2 11 种土壤指标间的斯皮尔曼相关性分析  
Fig. 2 Spearman correlation analysis of 11 soil indicators

表 2 土壤质量指标的主成分分析结果  
Table 2 Results of principal component analysis of soil quality indicators

土壤指标	PC1	PC2	PC3	PC4
pH	-0.178	-0.333	<b>-0.790</b>	0.215
速效氮	0.748	0.365	0.255	0.207
有效磷	-0.006	0.004	<b>0.868</b>	0.216
速效钾	0.074	0.290	0.050	<b>0.871</b>
有机质	<b>0.865</b>	0.280	0.193	0.185
全氮	<b>0.883</b>	0.254	0.189	0.191
全磷	0.632	0.218	0.404	0.340
全钾	-0.627	-0.272	0.232	0.143
含水量	0.463	0.595	0.011	0.292
容重	-0.724	-0.122	0.061	0.133
砂粒	-0.288	<b>-0.912</b>	-0.107	-0.173
粉粒	0.285	<b>0.889</b>	0.085	0.190
黏粒	0.222	<b>0.838</b>	0.202	0.044
特征值	6.425	1.555	1.342	0.904
方差百分比	49.426	11.960	10.327	6.955
累计方差百分比	49.426	61.385	71.712	78.667

注：表中字体加粗者代表该指标的载荷值在最高因子载荷值 10% 以内，具下划线者代表对应指标入选最小数据集。

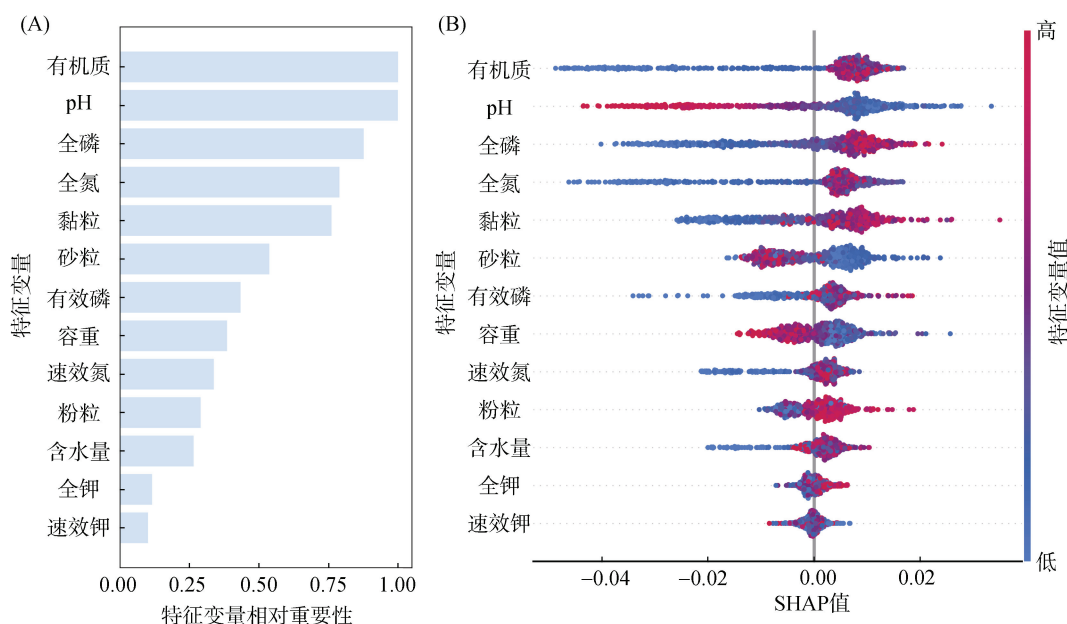
分别为 0.94、0.93 和 0.95(表 3)。为了实际验证产量模型的预测结果，将预测产量分别与中国东北各市、县级地区统计年鉴中的实际产量进行比较。这些年鉴中的产量数据已标准化为  $t/hm^2$ ，因此可与预测产量进行直接比较。在模型中，每个地区的市级和县级预测产量为这些地区的产量中值。精度验证结果表明，分市和分县验证中，预测产量与实际产量之间的  $R^2$  分别为 0.72 和 0.49，表明两者之间存在显著的正相关关系。因此，本研究用于土壤质量指数(SQI)验证的预测产量具有较高的可信度。

表 3 基于 100 次重复 10 折交叉验证的不同预测模型性能平均值

模型	$R^2$	RMSE	MAE
玉米	0.94	0.72	0.54
水稻	0.93	1.05	0.78
大豆	0.95	0.21	0.16
相对产量	0.59	0.06	0.04

基于 11 项土壤属性与相对产量建立的随机森林模型精度结果为  $R^2=0.59$ (表 2), 这表明该模型中 11 项土壤属性解释了 59% 的产量变化。SHAP 法分析结果(图 3)表明, 对模型输出结果影响最大的前 3 个特征分别为有机质、pH 和全磷。其中, 有机质、全氮、速效氮和有效磷等营养指标的高值对产量具有显著的促进作用, 而高容重和高砂粒含量对产量表现出负面影响。相对特征重要性结果表明, 排名前 5 位的特征变量累计贡献度超过 60%, 排名前 8 位的特征变量累计贡献度超过 80%, 排名前 10 位的特征变量累计贡献度超过 90%。因此, 可以认定重要性排名前 10 位的特征变量已经在模型预测中发挥了主要作用。在此基础上, 结合相关性分析对重要性排名前

10 位的指标进行进一步筛选。相关性分析结果(图 2)表明, 有机质与全氮( $r=0.96$ )、全磷( $r=0.71$ )及速效氮( $r=0.85$ )之间存在显著的正相关关系, 反映出养分指标之间存在较强的冗余性。同时, 考虑到磷在土壤中易与钙结合形成难溶性磷酸盐, 其有效性受限<sup>[19]</sup>。因此, 在这几个指标中选取解释能力更强且综合性更高的有机质作为代表性指标。而在颗粒组成中, 黏粒与砂粒( $r=-0.83$ )和粉粒( $r=0.74$ )均具有较高相关性, 且黏粒在 SHAP 分析中的相对重要性最高, 因此选取黏粒作为土壤质地的代表变量。综合 SHAP 特征相对重要性排序与相关性分析结果, 在确保模型解释能力的同时降低变量冗余, 最终构建的 MDS 由有机质、pH、黏粒、有效磷和容重 5 个关键指标组成。



(图 A 为基于 SHAP 值的相对特征重要性排序。图 B 为 SHAP 摘要图, 其中每个点对应一个样本, 颜色(红色表示高值, 蓝色表示低值)则标识了特征值的大小; 点相对于中心线(底部横轴零点对应的垂线)的位置反映了该特征对模型输出的影响程度: 离中心线越远, 表示该特征变量对预测模型的影响越大。SHAP 值为正表示正面影响, 反之表示负面影响)

图 3 SHAP 法分析结果

Fig. 3 Results of SHAP analysis

## 2.4 土壤质量评价

结合指标权重(表 4)和基于非线性评分函数公式(1)计算的指标分数, 采用公式(3)分别计算基于全量数据集的 SQI(SQI-TDS)、基于 PCA 法获取的 MDS 的 SQI(SQI-PMDS)和基于 SHAP 法获取的 MDS 的 SQI(SQI-SMDS)。结果表明, 中国东北地区耕地 SQI-TDS 介于 0.205~0.679, 均值为 0.466; SQI-PMDS 介于 0.179~0.723, 均值为 0.452; SQI-SMDS 介于 0.194~0.718, 均值为 0.467。SQI-TDS 与 SQI-PMDS、SQI-SMDS 的回归分析结果显示, SQI-PMDS 和 SQI-SMDS 均与 SQI-TDS 呈显著正相关,  $R^2$  分别为 0.63

和 0.82 (图 4), 表明两种方法筛选的 MDS 均能够较好地实现评价指标的精简, 但基于 SHAP 法筛选的 MDS 评价的精度更高。此外, 与相对产量进行回归验证, 结果显示, 3 种 SQI 均与相对产量呈显著相关,  $R^2$  大小顺序为: SQI-SMDS( $R^2=0.40$ )>SQI-TDS( $R^2=0.37$ )>SQI-PMDS( $R^2=0.15$ )(图 5)。这表明基于 SHAP 法筛选的 MDS 的评价结果相较 PCA 法筛选的 MDS 的评价结果能更好地反映作物产量状况。

## 3 讨论

本研究以中国东北黑土区耕地为研究对象, 从

表 4 全量数据集和最小数据集的指标权重  
Table 4 Weight values of soil quality indicators of total and minimum data sets

土壤指标	全量数据集		PCA 法最小数据集		SHAP 法最小数据集	
	公因子方差	权重	公因子方差	权重	公因子方差	权重
pH	0.813	0.08	0.775	0.227	0.763	0.216
速效氮	0.801	0.078				
有效磷	0.800	0.078	0.749	0.220	0.753	0.213
速效钾	0.851	0.083	0.512	0.150		
有机质	0.899	0.088	0.631	0.185	0.752	0.213
全氮	0.917	0.09				
全磷	0.725	0.071				
全钾	0.542	0.053				
含水量	0.653	0.064				
容重	0.56	0.055			0.712	0.202
砂粒	0.957	0.094	0.741	0.218		
粉粒	0.914	0.089				
黏粒	0.794	0.078			0.547	0.155

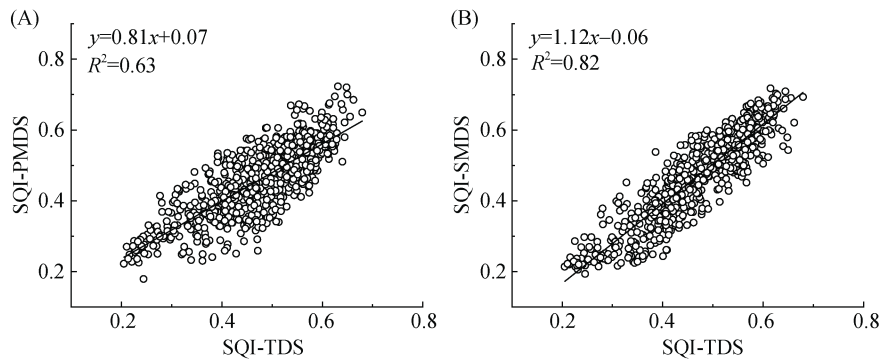


图 4 基于最小数据集的土壤质量指数与基于全量数据集的土壤质量指数的相关性  
Fig. 4 Correlations of soil quality indexes based on minimum data set and total data set

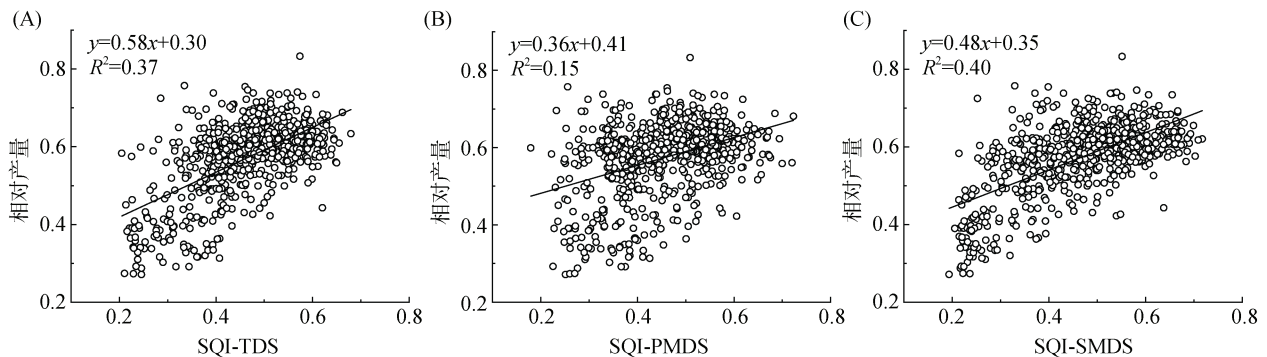


图 5 基于不同方法的土壤质量指数与相对产量的关系  
Fig. 5 Relationships between soil quality indexes based on different methods and relative yields

11 项土壤属性中分别采用 PCA 法和 SHAP 法筛选包含 5 个指标的 MDS, 旨在优化区域耕地土壤质量评估体系。两种方法均明确了有机质、pH、土壤养分以及颗粒组成是影响黑土区耕地质量的关键指标<sup>[20]</sup>。其中, 有机质作为土壤的重要组成部分, 不仅影响多种土壤性质, 还参与调控生物地球化学循环, 是评价

土壤肥力和退化程度的核心指标<sup>[21]</sup>。pH 作为土壤酸碱度的表征指标, 其变化直接影响土壤生物地球化学过程, 对陆地生态系统的结构和功能具有重要调控作用<sup>[22]</sup>。土壤氮、磷、钾等养分元素是植物生长发育的必需营养元素, 对作物产量和品质具有决定性影响<sup>[23]</sup>。土壤颗粒组成作为决定土壤过程和功能

的关键基础性指标,影响着土壤电导率、有机碳、阳离子交换容量等许多理化性质的动态<sup>[24]</sup>。两种方法构建的 MDS 与耕地质量等级国家标准(GB/T 33469—2016)<sup>[25]</sup>中的评价指标相比,均强调了有机质、土壤养分状况、土壤酸碱度的重要性。因此,总体而言,两种方法选取的关键指标体系均具有一定的合理性,且具备一定的相似性。

在 MDS 构建过程中,两种方法在指标选择上存在的显著差异在于:SHAP 法筛选的 MDS 包含了容重这一关键物理属性,而 PCA 法则未将其纳入。容重作为评价土壤压实敏感性和强度的核心参数,与土壤通气性、水分渗透性等关键功能密切相关<sup>[26]</sup>。相比之下,通过 PCA 法构建的 MDS 未能涵盖这一重要指标。这种差异可能源于两种方法的本质特性不同。PCA 法的核心目标是最大限度保留数据的变异性,因此倾向于选择变异性较高的指标,例如有效磷和速效钾。速效养分指标虽然在统计学上具有较强的解释力,但其对施肥等人为管理措施高度敏感,主要反映短期土壤肥力动态和管理干预效果,难以表征土壤固有的长期质量特征<sup>[27]</sup>。这种高变异性导致速效养分含量在短期内波动显著,甚至可能因过量施肥而丧失指示价值,从而降低了其作为土壤长期质量评价指标的可靠性。相比之下,容重作为土壤物理属性的关键指标,尽管空间变异性较低,但因其稳定性和与土壤有机质、质地等内在属性的密切关联,成为土壤质量评价中的重要参考<sup>[28]</sup>。SHAP 法通过量化土壤属性对作物产量预测的贡献度来筛选关键指标,相较于传统 PCA 法展现出显著优势。具体而言,SHAP 法以管理目标和土壤功能为导向,筛选的指标体系更贴近实际生产需求,更能准确地反映土壤的实际产能状况。产量验证结果也强有力地支持了这一观点。此外,虽然通过两种方法构建的 MDS 均能保留足够信息以评估管理措施对土壤健康的影响,但在相同指标数量下,基于 SHAP 法构建的 MDS 的 SQI 与 SQI-TDS 的相关性更强,这进一步证实了 SHAP 法在指标筛选中的优越性。

与传统 PCA 法相比,SHAP 法在 MDS 筛选中具有以下优势:首先,克服了 PCA 法过度依赖统计显著性的局限。PCA 法倾向于选择数据变异性高的指标,但这些指标与实际生产功能的关联性可能较弱,难以充分反映土壤生产功能的关键属性<sup>[9]</sup>。这可能是导致基于 PCA 法构建的 MDS 的 SQI 与作物产量相关性较低的原因。其次,SHAP 法采用以生态系统服务为导向的筛选策略,在指标选择过程中更注重其与

土壤功能的相关性,而非单纯依赖统计贡献率<sup>[29]</sup>。这种功能导向的筛选机制有效避免了关键信息的丢失,确保了所选指标能够准确反映土壤功能的实际状况及其对农业生产的影响<sup>[30]</sup>。因此,基于 SHAP 法构建的 MDS 不仅保留了土壤信息的完整性,还显著提升了评估结果的实践指导价值,为土壤管理和农业决策提供了更科学可靠的依据。

## 4 结论

本研究创新性地引入 SHAP 法筛选关键土壤质量指标并构建 MDS,再与传统的 PCA 法进行了系统性的对比研究。结果表明,基于 SHAP 法构建的 MDS 获取的 SQI 在模型拟合效果和实际应用表现上均显著优于传统 PCA 方法。SHAP 法优势主要表现在:① 相较于 PCA 法,基于 SHAP 法构建的 MDS 的 SQI 与基于全量数据集的 SQI 线性回归  $R^2$  提升了 30%;② 相应地,其与作物产量的线性回归  $R^2$  提高了近 1.7 倍。这证实了基于生产功能导向的 SHAP 法能够显著提升土壤质量评估的精准性和实用性,为区域尺度土壤质量评价提供了更为科学和可靠的技术支撑。然而,由于测试指标的局限性,本研究主要聚焦于土壤理化指标,未来需进一步整合生物指标以构建更全面的评价体系,从而更综合地反映土壤质量状况及其生态功能。

**致谢:** 本研究在东北黑土区的调查和采样工作得到了来自中国科学院南京土壤研究所黄标、赵玉国、李德成、于东升、刘峰、王昌昆、高磊、马利霞、谢恩泽、高张、陈剑、严国菁、徐兵等,来自中国科学院东北地理与农业生态研究所隋跃宇、陈一民等,来自黑龙江省黑土保护利用研究院迟凤琴、张久明、匡恩俊等,来自沈阳农业大学汪景宽、裴久渤、徐英德等,来自中国农业大学张超等的大力协助,在此一并表示衷心感谢。

## 参考文献:

- [1] Yu P J, Han D L, Liu S W, et al. Soil quality assessment under different land uses in an Alpine grassland[J]. *Catena*, 2018, 171: 280–287.
- [2] 丁明净, 脱云飞, 郑阳, 等. 基于最小数据集栗子坪自然保护区不同林分类型土壤质量评价[J]. *土壤通报*, 2024, 55(5): 1215–1228.
- [3] Bünemann E K, Bongiorno G, Bai Z, et al. Soil quality – A critical review[J]. *Soil Biology and Biochemistry*, 2018, 120: 105–125.
- [4] Larson W E, Pierce F J. Conservation and enhancement of soil quality // *Evaluation for sustainable land management*

- in the developing world. Vol.2[M]. Bangkok, Thailand: International Board for Soil Research and Management, 1991: 175–203.
- [5] Shi J F, Qian W T, Zhou Z B, et al. Influence of acid mine drainage leakage from tailings ponds on the soil quality of desert steppe in the northwest arid region of China[J]. *Land*, 2023, 12(2): 467.
- [6] Guo L L, Sun Z G, Ouyang Z, et al. A comparison of soil quality evaluation methods for Fluvisol along the lower Yellow River[J]. *Catena*, 2017, 152: 135–143.
- [7] 段碧辉, 孙奥, 王芳, 等. 荆门市耕地不同利用方式土壤质量综合评价[J]. *土壤*, 2023, 55(6): 1371–1379.
- [8] 代先强, 周嵘, 周鑫斌, 等. 重庆市植烟土壤质量评价指标体系的构建与应用[J]. *土壤*, 2024, 56(2): 388–397.
- [9] Li P, Zhang H J, Deng J J, et al. Cover crop by irrigation and fertilization improves soil health and maize yield: Establishing a soil health index[J]. *Applied Soil Ecology*, 2023, 182: 104727.
- [10] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems. December 4 - 9, 2017, Long Beach, California, USA. ACM, 2017: 4768–4777.
- [11] 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.
- [12] 杨青松, 杨伟, 彭珏, 等. 典型黑土区坡耕地土壤微生物多样性及群落结构对侵蚀–沉积的响应[J]. *土壤学报*, 2024, 61(6): 1741–1754.
- [13] 鲁如坤. 土壤农业化学分析方法[M]. 北京: 中国农业科学技术出版社, 2000.
- [14] Biswas S, Hazra G C, Purakayastha T J, et al. Establishment of critical limits of indicators and indices of soil quality in rice-rice cropping systems under different soil orders[J]. *Geoderma*, 2017, 292: 34–48.
- [15] Choudhury B U, Mandal S. Indexing soil properties through constructing minimum datasets for soil quality assessment of surface and profile soils of intermontane valley (Barak, North East India)[J]. *Ecological Indicators*, 2021, 123: 107369.
- [16] 朱志畅, 葛焱, 臧晶荣, 等. 基于无人机图像和 SHAP 特征筛选的小麦田间产量预测方法研究[J]. *麦类作物学报*, 2025, 45(2): 264–274.
- [17] Zeraatpisheh M, Bakhshandeh E, Hosseini M, et al. Assessing the effects of deforestation and intensive agriculture on the soil quality through digital soil mapping[J]. *Geoderma*, 2020, 363: 114139.
- [18] Iizumi T, Sakai T. The global dataset of historical yields for major crops 1981–2016[J]. *Scientific Data*, 2020, 7: 97.
- [19] 冶赅康, 俄胜哲, 陈政宇, 等. 土壤中磷的存在形态及分级方法研究进展[J]. *中国农学通报*, 2023, 39(1): 96–102.
- [20] 郭伟, 李丹丹, 徐基胜, 等. 秸秆与有机无机肥配施对不同质地潮土土壤质量和小麦产量的影响[J]. *土壤学报*, 2024, 61(5): 1360–1373.
- [21] Zhang Z Y, Ai N, Liu G Q, et al. Soil quality evaluation of various microtopography types at different restoration modes in the loess area of Northern Shaanxi[J]. *Catena*, 2021, 207: 105633.
- [22] Hong S B, Piao S L, Chen A P, et al. Afforestation neutralizes soil pH[J]. *Nature Communications*, 2018, 9: 520.
- [23] 陈方正, 任健, 刘思涵, 等. 基于最小数据集的洞庭湖流域南部耕地土壤肥力综合评价[J]. *土壤通报*, 2021, 52(6): 1348–1359.
- [24] Meng C, Yang W, Ren X J, et al. *In-situ* soil texture classification and physical clay content measurement based on multi-source information fusion[J]. *International Journal of Agricultural and Biological Engineering*, 2023, 16(1): 203–211.
- [25] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. 耕地质量等级: GB/T 33469—2016[S]. 北京: 中国标准出版社, 2016.
- [26] Makovníková J, Širáň M, Houšková B, et al. Comparison of different models for predicting soil bulk density. Case study—Slovakian agricultural soils[J]. *International Agrophysics*, 2017, 31(4): 491–498.
- [27] Zhang W C, Zhang Y Y, Zhang X, et al. The spatiotemporal variability of soil available phosphorus and potassium in karst region: The crucial role of socio-geographical factors[J]. *Land*, 2024, 13(6): 882.
- [28] 杨丽扬, 张永清, 田静, 等. 基于最小数据集的生草覆盖对吉县苹果园土壤肥力质量的影响评价[J]. *生态与农村环境学报*, 2020, 36(3): 374–381.
- [29] Calzolari C, Ungaro F, Filippi N, et al. A methodological framework to assess the multiple contributions of soils to ecosystem services delivery at regional scale[J]. *Geoderma*, 2016, 261: 190–203.
- [30] Askari M S, Holden N M. Indices for quantitative evaluation of soil quality under grassland management[J]. *Geoderma*, 2014, 230/231: 131–142.

(责任编辑: 于 飞)