

DOI: 10.13205/j.hjgc.202508018

崔释文,陈坚,廖磊,等. 基于随机森林的区域土壤重金属污染影响因素识别研究[J]. 环境工程, 2025, 43(8): 195–203.

CUI S W, CHEN J, LIAO L, et al. Identification of influencing factors of regional soil heavy metal pollution based on random forest[J]. Environmental Engineering, 2025, 43(8): 195–203.

基于随机森林的区域土壤重金属污染影响因素识别研究

崔释文^{1,2} 陈 坚² 廖 磊³ 周 睿¹ 钱江波⁴ 黄国鑫^{2*}

(1. 吉林大学 新能源与环境学院, 长春 130021; 2. 生态环境部环境规划院, 北京 100041; 3. 核工业二九〇研究所, 广东 韶关 512029; 4. 浙江科环环境工程技术有限公司, 杭州 311200)

摘 要:准确识别重金属污染来源是土壤重金属污染防治的重要前提。由于缺乏土壤重金属污染影响因素有关信息, 污染溯源分析的效果常受到限制。以广东省某典型工业地区为研究对象, 基于577个采样点的土壤重金属浓度实测数据和18项环境协变量数据, 采用随机森林(RF)和双变量局部空间自相关方法, 识别了土壤Cd、Pb和Cr浓度的影响因素, 确定了18项影响因素对Cd、Pb和Cr浓度的定量贡献, 并提出了有针对性的土壤环境管理对策。结果表明: 3种重金属浓度的预测模型决定系数分别为0.93(Cd)、0.97(Pb)、0.93(Cr)时, RF预测模型性能最佳; 3种重金属相对浓度空间分布存在差异; Cd的主要影响因素为铁路(贡献率0.119), Pb的主要影响因素为土壤pH值(贡献率0.099), Cr的主要影响因素为危险废物集中处置场(贡献率0.100); 相较于研究区的其他地区, 中部的3种重金属浓度较高, 人类活动复杂, 且他们与其最主要影响因素的高-高区较多。在实施土壤重金属污染源头防控时应优先关注中部。

关键词:区域尺度; 土壤; 重金属; 随机森林; 影响因素; 双变量局部莫兰指数

Identification of influencing factors of regional soil heavy metal pollution based on random forest

CUI Shiwen^{1,2}, CHEN Jian², LIAO Lei³, ZHOU Rui¹, QIAN Jiangbo⁴, HUANG Guoxin^{2*}

(1. College of New Energy and Environment, Jilin University, Changchun 130021, China;

2. Chinese Academy of Environmental Planning, Beijing 100041, China;

3. Research Institute No. 290, CNNC, Shaoguan 512029, China;

4. Zhejiang Kehuan Environmental Engineering Technology Co., Ltd., Hangzhou 311200, China)

Abstract: Accurate identification of heavy metal pollution sources is an important prerequisite for soil pollution prevention and control. However, due to the lack of information on the influencing factors of heavy metal pollution, the effectiveness of pollution source tracing analysis is often constrained. In this paper, a typical industrial area was taken as the study area. On the basis of the actual measurements of soil heavy metal concentrations from 577 sampling points and data on the 18 environmental covariates, random forest (RF) and bivariate local spatial autocorrelation methods were applied to identify the influencing factors of the Cd, Pb, and Cr concentrations, determine the quantitative contributions of the 18 influencing factors to these heavy metal concentrations, and further propose targeted strategies for soil environmental management in an industrialized study area. The results showed that the optimal performance of the RF prediction models for the three heavy

收稿日期: 2024-12-03; 修改日期: 2025-01-14; 接收日期: 2025-01-20

基金项目: 国家重点研发计划(2023YFC3708901)

第一作者: 崔释文(2000—), 女, 硕士研究生, 主要研究方向为土壤和地下水污染治理。1176668387@qq.com

* 通信作者: 黄国鑫(1980—), 男, 研究员, 主要研究方向为土壤和地下水污染防治。huanggx@caep.org.cn

metals was achieved when the coefficient of determination (R^2) reached 0.93 with a root mean square error (RMSE) of 0.43 mg/kg for Cd, 0.97 with an RMSE of 48.57 mg/kg for Pb, and 0.93 with an RMSE of 18.57 mg/kg for Cr, namely. There were differences in the spatial distributions of the relative concentrations of the three heavy metals. The regions with relatively high Cd concentration were concentrated in the central and southern parts of the study area, whereas relatively high Pb relative concentration were predominantly found in the central and eastern parts. The regions with relatively high Cr relative concentration were concentrated in the southwestern and northeastern parts. Railway was identified as the most significant factor influencing Cd concentration with a contribution rate of 0.119. Soil pH was identified to be the most significant factor influencing Pb concentrations with a contribution rate of 0.099. The hazardous waste disposal site was identified to be the most significant factor influencing Cr concentrations with a contribution rate of 0.100. Compared to the rest of the study area, the central region exhibited higher concentrations of Cd, Pb, and Cr, more complex human activity, and a greater number of high-high cluster zones, where Cd, Pb, and Cr concentrations and their most significant influencing factors were concentrated. When source prevention and control measures for heavy metal pollution in soil were implemented, particular attention should be prioritized towards the central region of the study area.

Keywords: regional scale; soil; heavy metals; random forest; influencing factors; bivariate local Moran index

0 引言

土壤重金属污染已成为世界性的环境和公共卫生问题^[1]。许多国家如中国^[2]、美国^[3]、印度^[4]均发现部分区域土壤重金属浓度超过规定限值。重金属会随着食物链累积迁移进入人体,引起人体免疫系统、神经系统、心血管系统和肾脏系统的病变,对人体健康构成威胁^[5]。生态环境部、国家发展和改革委员会、工业和信息化部等七部委联合印发了《土壤污染防治行动计划》,旨在解决长期积累的土壤重金属严重污染问题,全面管控土壤重金属污染风险。因此,开展土壤重金属污染溯源分析对精准实施土壤污染风险管控和修复,保障人民群众吃得放心、住得安心具有重要意义。

土壤重金属来源复杂,包含人为因素和自然因素,其中采矿活动、工业排放、污水灌溉等人为因素加剧了土壤重金属富集^[6],岩石风化、沙尘暴、生物过程等自然因素促进了土壤重金属累积^[7]。近年来,土壤重金属污染溯源分析越来越受到重视^[8]。多元统计分析^[9]、正定矩阵分解模型(PMF)^[10]、UNMIX模型^[11]、地理探测器^[12]、空间分析^[13]等方法在该领域已被广泛应用。上述传统方法通常依靠已有研究成果和专家经验来确定影响因素^[14]。但若没有先验条件,将无法准确识别影响因素。而且,由于土壤重金属污染具有高度随机性,很难定量表征影响因素的贡献程度^[15]。

针对上述问题,以土壤样品的实测数据和环境协变量为基础,随机森林(RF)^[16]、CatBoost^[17]和人工神经网络^[18]等大数据挖掘方法被逐渐应用于污染溯

源分析研究。例如,冯锋等^[19]基于170个土壤样品和20个环境协变量(如土壤侵蚀度、人口密度、植被),运用RF和XGBoost预测了土壤Cd浓度的空间分布特征,确定了土壤重金属浓度来源,并建立了重金属浓度与其影响因素之间的复杂相关性。与其他大数据挖掘方法相比,RF对样本量的要求更宽松,处理数据间非线性和层次关系的能力更强,对影响因素贡献的定量分析也更好^[20,21]。即便如此,仍无法直观地建立重金属浓度与其影响因素之间的空间相关性,也无法进一步定性提供影响因素的空间交互信息。

基于此,本研究以广东某典型工业区域为研究对象,基于577个采样点的土壤重金属浓度实测数据和18个环境协变量数据,采用RF定量确定土壤Cd、Pb和Cr浓度的影响因素,利用双变量局部莫兰指数(BLMI)定性确定土壤Cd、Pb和Cr浓度与其影响因素之间的空间相关关系,从定量和定性的角度确定土壤Cd、Pb和Cr浓度的影响因素,以期精准实施土壤重金属污染风险管控和修复提供理论依据和技术支持。

1 材料与方法

1.1 研究区概况

研究区(113° 05'E—113° 56'E, 24° 28'N—24° 53'N)位于广东省某工业化地区,总面积1670 km²(图1)。研究区属于典型的亚热带湿润性季风气候,冬短夏长,雨量充沛,雨热基本同季。年平均温度21℃,年极端最高气温为40℃(7月),年极端最低气温为-5℃(1月),年无霜期约310 d,年平均降水量1400~2400 mm。研究区属山地丘陵、盆地、冲积平原

地貌;河网密布,所有河流均发源于山区,向中部汇合后注入北江。研究区主要从事采矿、冶炼等工业活动,大巴山矿、凡口铅锌矿、乐昌铅锌矿附近土壤重金属浓度比背景值高8倍左右^[22]。本研究重点关注土壤Cd、Pb和Cr的污染。

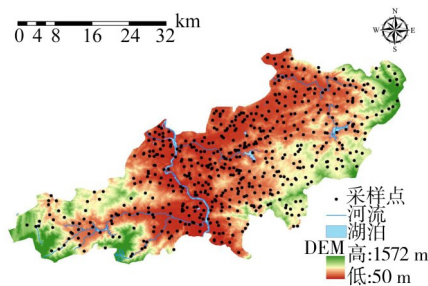


图1 研究区与采样点分布

Figure 1 Study area and sampling sites

1.2 样品采集

在研究区采集577个表层土壤样品(0~20 cm深),利用“对角线五点采样法”在每个采样区域(25 cm×25 cm)采集5个子样品,子样品混合得到均匀、有代表性的样品(平均重量1.5 kg),并放置在密封袋中。冷藏运输送至实验室后,去除石块和植物碎屑,将样品在室温下风干研磨,过尼龙筛(100目),装入聚乙烯袋,低温保存(4℃),以备用于重金属浓度测定和理化性质分析。

1.3 分析测试

采用数显pH计(中国,雷磁PXSJ-216F)测定土壤pH值(固液比1:2.5);采用滴定法,参照NY/T 1121.6—2006《土壤检测 第6部分 土壤有机质的测定》在重铬酸钾-硫酸体系中测定土壤有机质含量;采用石墨炉原子吸收光谱法测定土壤Cd浓度(日本,岛津AA-6880G);采用火焰原子吸收分光光度法测定土壤Pb和Cr浓度(美国,安捷伦240AA)。同时,为保证检测方法与分析结果的准确度和可靠性,采取重复空白样品和标准对照样品作为参考。采用超纯水制备试剂溶液。如无特殊说明,试剂均为分析纯或优级纯。

1.4 数据收集与预处理

信访举报数据反映了公众对环境污染问题的关注度,可以表征污染源信息;夜间灯光指数数据反映了人类活动的强度和分布,可以表征人类活动对土壤重金属的影响^[23]。土壤pH值、土壤有机质数据来自实验室检测;地下水埋深数据(文本,152条)来源

于现场钻孔采样;信访举报数据(文本,20条)通过网络爬虫爬取;工业企业数据(文本,310条)、加油站数据(文本,34条)及危险废物集中处置场数据(文本,2条)来自百度地图服务(<http://map.baidu.com>);尾矿库数据(文本,12条)来源于绿网(<http://www.lvwang.org.cn>);固体废物堆场数据(文本,105条)来源于生态环境部卫星环境应用中心(<http://www.secmep.cn>);高程(栅格,分辨率90 m,1:10⁵)、地形地貌(栅格,1:10⁵)、年平均降水量(栅格,1:10⁵)、年平均气温(栅格,1:10⁵)、夜间灯光指数(栅格,1:10⁵)、归一化植被指数(栅格,1:10⁵)、陆地生态系统服务总价值(栅格,1:10⁵)、居民点(矢量)、铁路(矢量)等数据均来源于中国资源环境科学与数据中心(<https://www.resdc.cn>)。使用ArcGIS10.6软件的Euclidean Distance工具,计算每个采样点与信访举报、工业企业、尾矿库、铁路、固体废物堆场、危险废物集中处理场、加油站和居民点之间的距离。在统计采样点的地形地貌类型(平原、台地、丘陵、小起伏山地、中起伏山地、高起伏山地)基础上,采用描述性统计方法将地形地貌数据从字符串型转换为数值型。

1.5 随机森林模型

RF是一种非参数、非线性、监督和多元集合学习方法,具有良好的鲁棒性,可防止发生过拟合情况^[24]。在构建RF模型时,先集成每个决策树的预测结果,再将每个预测结果进行平均计算得到最终的预测结果[式(1)]。利用训练集之外的剩余样本,即袋外样本(OOB),通过内部交叉验证[式(2)],评估环境协变量(即影响因素)对预测结果的重要性^[25]。

$$f(x_i) = \frac{1}{n_{tree}} \sum_{i=1}^{n_{tree}} h_i(x_i) \quad (1)$$

式中: $f(x_i)$ 为决策树 t 的预测结果; x_i 为决策树 t ,无量纲; $h_i(x_i)$ 为决策树 t 的预测结果 i ; n_{tree} 为决策树的数量。

$$VI_i = \frac{\sum (\text{errOOB}_2 - \text{errOOB}_1)}{n_{tree}} \quad (2)$$

式中: VI_i 为环境协变量 i 的重要性; errOOB_1 为决策树的袋外样本准确率; errOOB_2 为对袋外样本中任意特征 X_j 添加随机扰动后的袋外样本准确率。

1.6 双变量局部空间自相关

BLMI由Anselin首次提出^[26],其可揭示局域空间内2个变量间的空间关联特征,呈现5种空间聚类关系(高-高聚类、高-低聚类、低-低聚类、低-高聚类和

不显著)[式(3)]^[27]。

$$I_{kl}^i = \left(\frac{X_k^i - \bar{X}_k}{\sigma_k} \right) \sum_{j=1}^n W_{ij} \left(\frac{X_l^j - \bar{X}_l}{\sigma_l} \right) \quad (3)$$

式中: I_{kl}^i 是双变量局部莫兰指数; X_k^i 和 X_l^j 分别是位置*i*和位置*j*上变量*k*和变量*l*的值; $\bar{X}_k$ 和 $\bar{X}_l$ 分别是变量*k*和变量*l*的平均值; σ_k 和 σ_l 分别是变量*k*和变量*l*的方差; W_{ij} 是空间权重矩阵,可根据位置*i*和位置*j*之间的距离权重来表示; n 是网格单元数量。

1.7 实验设计

土壤重金属实测浓度的描述性统计分析。运用SPSSPRO软件提取研究区土壤Cd、Pb和Cr实测浓度的最小值、最大值、中位数、平均值和变异系数等,进行描述性统计分析。

基于RF的土壤重金属浓度空间分布,基于577个采样点的土壤Cd、Pb和Cr浓度实测数据,以18个环境协变量作为输入变量,以实测土壤Cd、Pb和Cr浓度作为输出,分别建立基于RF的土壤Cd、Pb和Cr浓度预测模型。随机将519个土壤样本作为训练集,58个土壤样本作为验证集。选择决策树个数(n_{tree})(100、150、200、250、300、350、400、450)和节点变量数(m_{try})(1、3、5、7、9)作为RF预测模型的训练参数,进行参数调整,提升模型的预测精度。利用决定系数 R^2 [式(4)]和均方根误差RMSE[式(5)]作为判断依据,评估模型的预测性能。模型输出的 R^2 均值越高, RMSE均值越低,则模型越精确。以RF预测结果为基础,利用ArcGIS 10.6进行克里金插值,绘制Cd、Pb和Cr相对浓度[式(6)]的空间分布图。

$$R^2 = 1 - \frac{\sum_{i=1}^n (\hat{z}_i - \bar{z})^2}{\sum_{i=1}^n (z_i - \bar{z})^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{z}_i - \bar{z})^2}{n}} \quad (5)$$

式中: n 为数据数量; z_i 和 $\hat{z}_i$ 分别为样本*i*的实测值和预测值,mg/kg; $\bar{z}$ 为测量值的平均值,mg/kg。

$$P_i = \frac{C_i}{S_i} \quad (6)$$

式中: P_i 为土壤重金属相对浓度,无量纲; C_i 为重金属*i*预测浓度,mg/kg; S_i 为GB 15618—2018《土壤环境质量农用地土壤污染风险管控标准(试行)》相应风险筛查值,mg/kg。

基于RF的土壤重金属浓度影响因素定量识别。

利用RF预测模型,拟合环境协变量与土壤Cd、Pb和Cr浓度数据;对环境协变量进行计算、归一化,采用袋外样本的准确率计算环境协变量重要性,揭示不同环境协变量在RF预测模型中的相对重要性,即识别影响因素对土壤Cd、Pb和Cr浓度的贡献率。

基于BLMI的土壤重金属浓度影响因素定性识别。利用GeoDa 1.22(<http://geodacenter.github.io>),采用BLMI,建立土壤Cd、Pb和Cr浓度与其对应的最主要影响因素的空间自相关性。考虑数据的可用性和计算效率,将研究区域划分为12262个网格单元(0.4 km×0.4 km)。

2 结果与讨论

2.1 土壤重金属实测浓度特征分析

研究区土壤Cd、Pb和Cr实测浓度如表1所示。可知:3种重金属实测浓度的平均值均大于各自中位数,表明研究区内土壤重金属浓度存在极大值。Cd、Pb浓度值呈强变异特点(变异系数>1.0),Cr浓度值呈中等变异特点(变异系数为0.86),说明3种重金属在研究区内分布极不均匀。此外,经数据统计,Cd-Pb浓度的Spearman秩相关系数为0.571($P<0.01$),表明这两种重金属可能具有共同或相似来源;Pb-Cr浓度的Spearman秩相关系数为-0.339($P<0.01$),表明这两种重金属可能具有不同来源。

表1 土壤重金属浓度描述性统计

Table 1 Descriptive statistics of soil heavy metal concentrations

重金属	最小值/ (mg/kg)	最大值/ (mg/kg)	中位数/ (mg/kg)	平均值/ (mg/kg)	变异 系数
Cd	0.02	13.96	0.26	0.66	1.99
Pb	9.39	2588.11	56.06	81.72	1.73
Cr	4.00	885.60	56.16	57.46	0.86

2.2 土壤重金属浓度预测模型优化

在不同的最优 n_{tree} 值和 m_{try} 值下,基于RF的Cd、Pb和Cr预测浓度与实测浓度之间的对比情况如图2所示。可知:当 n_{tree} 为300、 m_{try} 为1时,Cd的RF预测模型拟合效果最佳(R^2 为0.93, RMSE为0.43 mg/kg)(见图2a);当 n_{tree} 为350、 m_{try} 为1时,Pb的RF预测模型拟合效果最佳(R^2 为0.97, RMSE为48.57 mg/kg)(见图2b);当 n_{tree} 为450、 m_{try} 为1时,Cr的RF预测模型拟合效果最佳(R^2 为0.93, RMSE为18.57 mg/kg)(见图2c)。本研究中3种重金属的RF预测模型的 R^2 均大于已有研究中的 R^2 (Cd为0.85、Pb为0.79、Cr为

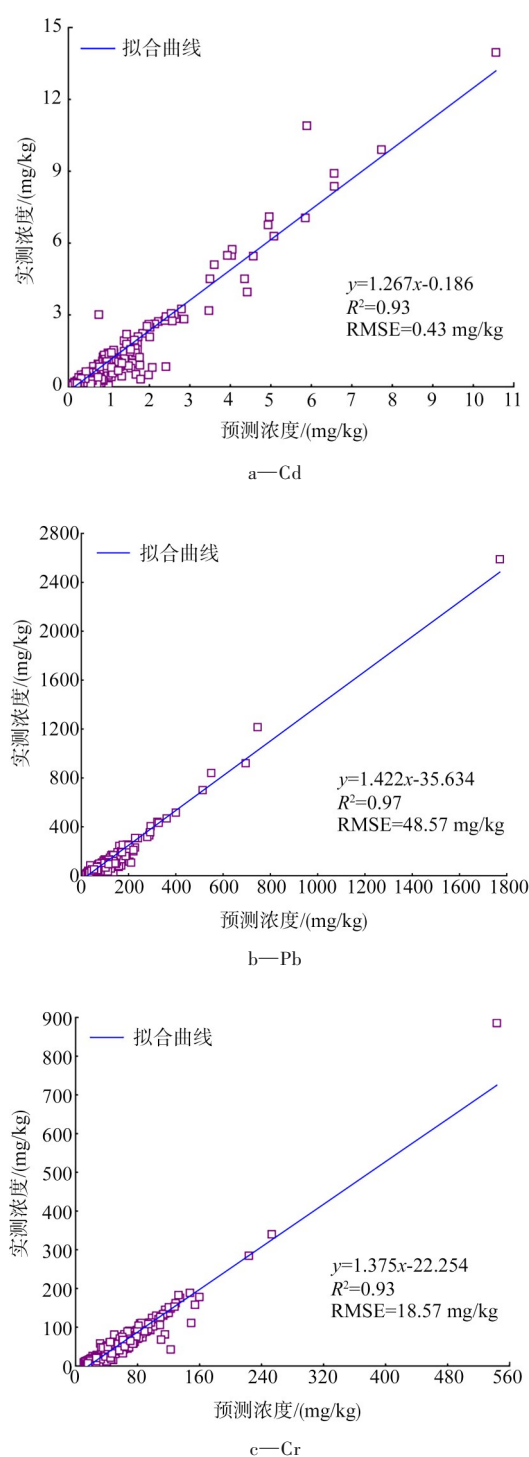


图2 基于随机森林的土壤重金属浓度预测值与实测值散点图

Figure 2 Scatter plot of predicted vs. measured heavy metal concentrations in soil based on random forest

0.78)^[28], RMSE均小于已有研究获得的 RMSE(Cd为 0.98 mg/kg、Pb为 59.88 mg/kg、Cr为 21.57 mg/kg)^[28], 可见拟合效果更佳。这可能缘于本研究利用了更多

的环境协变量来克服土壤空间异质性,从而降低了 RF 预测模型的不确定性^[29]。值得注意的是,选择更合适的重金属浓度数据集和获得更高的环境协变量空间分辨率,有望进一步提升 RF 预测模拟的准确性。

2.3 土壤重金属浓度空间分布

研究区土壤 Cd、Pb 和 Cr 的相对浓度空间分布图如图 3 所示。

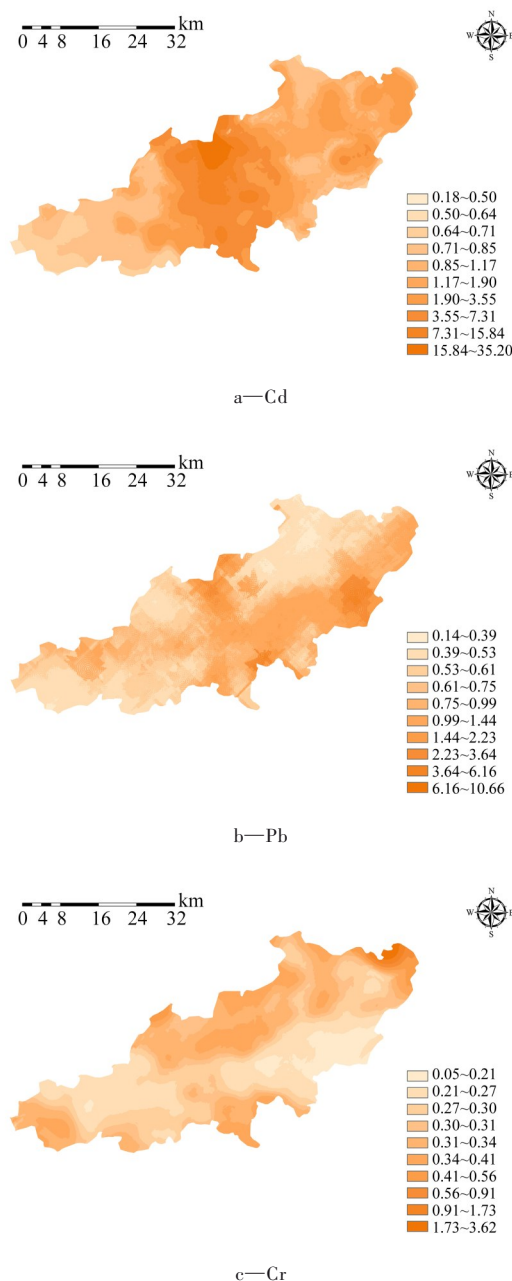


图3 土壤重金属相对浓度空间分布

Figure 3 Spatial distributions of relative heavy metal concentrations in soil

由图3可知:3种重金属浓度空间分布各异,但中部的Cd、Pb和Cr浓度均出现较高值。相对浓度>1.00的高Cd浓度区域主要集中在中部和南部(图3a),该区域有较多加油站,还有铁路覆盖,且人口密集^[30]。相对浓度>1.00的高Pb浓度区域主要分布在中部和东部(图3b),该区域的大部分土壤有机质含量较高^[31]。相对浓度较高的高Cr浓度区域主要分布在西南部 and 东北部(图3c),该地区拥有北江、马坝河和枫湾河。显然,高Cr浓度区域呈现沿着河流分布的趋势(图1)。3种重金属浓度的空间分布说明其污染来源存在较为明显的差异,但也受部分相同因素的影响。

2.4 土壤重金属浓度影响因素定量识别

影响因素对研究区土壤Cd、Pb和Cr浓度的定量贡献如图4所示。可知: Cd浓度的主要影响因素为铁路(贡献率0.119)、加油站(0.089)、年降水量(0.080)、危险废物集中处置场(0.074)、尾矿库

(0.069)和固废堆场(0.069),而工业企业(0.001)和地形地貌(0.000)对Cd浓度的影响较小(图4a); Pb浓度的主要影响因素为pH值(贡献率0.099)、铁路(0.091)、气温(0.081)、加油站(0.064)和危险废物集中处置场(0.061),而地形地貌(0.017)和地下水埋深(0.028)对Pb浓度的影响较小(图4b); Cr浓度的主要影响因素为危险废物集中处置场(贡献率0.100)、年降水量(0.099)、尾矿库(0.082)、夜间灯光(0.073)、高程(0.067)和加油站(0.067),而有机质(0.005)和信访举报(0.026)对Cr浓度的影响较小(图4c)。显然,各影响因素的贡献率在Cd、Pb和Cr浓度之间存在显著差异。另一方面,对Cd、Pb和Cr浓度影响均较大的影响因素有危险废物集中处置场和加油站。针对前者,应加强对危险废物转移和运输的监管以及对危险废物的安全处置,减少重金属的浸出率;针对后者,应设置排水沟和油水分离池,减少含油废水直接排入土壤。

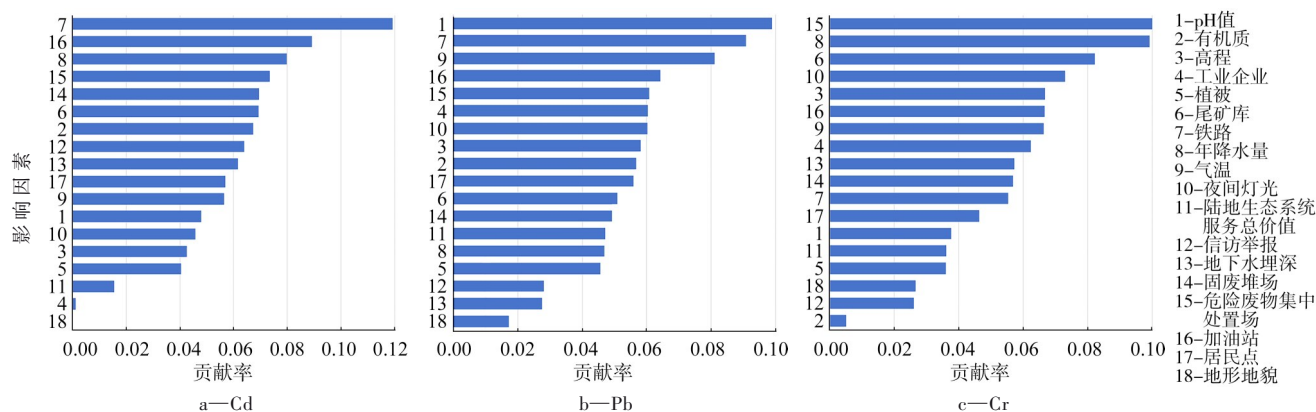


图4 影响因素对土壤重金属浓度的贡献率

Figure 4 Contribution rates of influencing factors to soil heavy metal concentrations

2.5 土壤重金属浓度与其第一影响因素之间的空间相关性分析

研究区内土壤Cd、Pb和Cr浓度与其第一影响因素的空间自相关聚类如图5所示。可知:高-高区中高重金属浓度与高影响因素值显著并存,应给予重点关注。对于Cd浓度与铁路空间关系来说,高-高区集中在研究区的中部和南部(图5a)。铁路运输货物泄漏、燃料燃烧、润滑油和枕木浸渍油的使用^[32],可能造成Cd在铁路周边土壤中累积。对于Pb与pH值的空间关系来说,高-高区集中在中部(图5b)。该区域的土壤pH值为7.0~9.1,呈弱碱性,而碱性pH会影响Pb的溶解度、生物可利用性、流动性^[33],进而影

响Pb的累积迁移和分布。对于Cr与危险废物集中处置场的空间关系来说,高-高区集中在东北部和南部(图5c)。Cr在危险废物贮存、焚烧、填埋的过程中浸出率较高^[34],会影响Cr的累积迁移和分布。与其他地区相比,中部的Cd、Pb和Cr浓度高(图3),人类活动复杂^[30],高-高区分布更多(图5)。因此,在采取土壤重金属污染风险防控措施时,应更多地关注中部。

3 结论

本研究构建的RF预测模型拟合效果优于已有研究。本研究中RF预测模型的最佳 R^2 (Cd)、 R^2 (Pb)、 R^2 (Cr)为0.93、0.97、0.93, RMSE(Cd)、RMSE(Pb)、

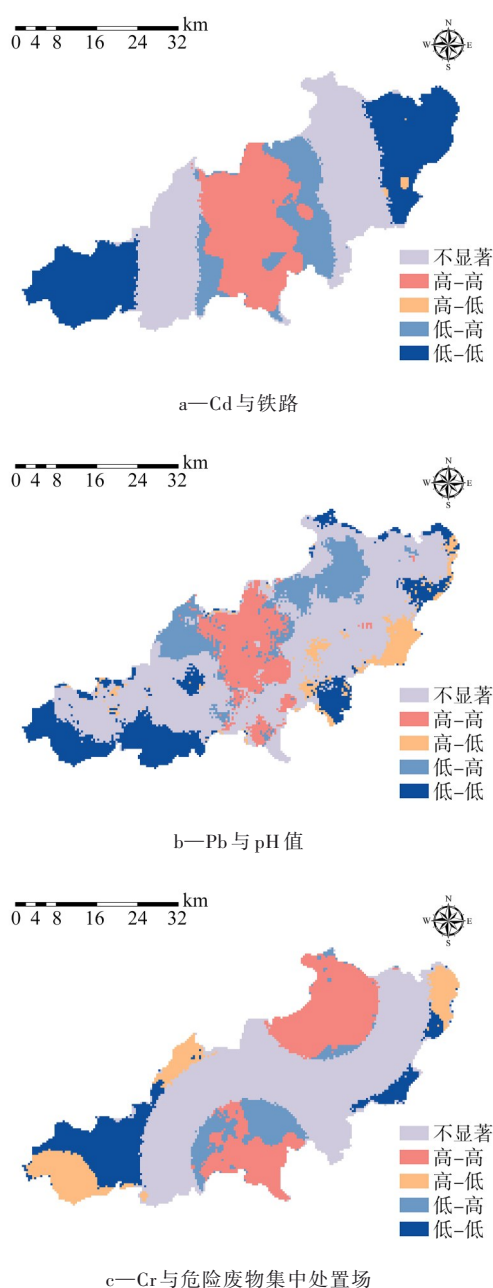


图5 土壤重金属浓度与其第一影响因素之间的空间聚类图

Figure 5 Spatial clustering maps of the relationship between soil heavy metal concentrations and their dominant influencing factors

RMSE(Cr)为0.43,48.57,18.57 mg/kg。

三种重金属浓度的空间分布存在显著差异: Cd的相对浓度高值区主要集中在研究区的中部和南部, Pb主要分布在中部和东部, Cr主要分布在西南部 and 东北部。

三种重金属浓度的主要影响因素存在明显差异: Cd的主要影响因素为铁路、加油站、年降水量、危

险废物集中处置场、尾矿库和固废堆场, Pb浓度的主要影响因素为土壤pH值、铁路、气温、加油站和危险废物集中处置场, Cr浓度的主要影响因素为危险废物集中处置场、年降水量、尾矿库、夜间灯光、高程和加油站。

三种重金属浓度与其影响因素的高-高区分布存在较大差异: Cd与铁路的高-高区集中在研究区的中部和南部, Pb与pH值的高-高区集中在中部, Cr与危险废物集中处置场的高-高区集中在东北部和南部。

在采取重金属污染风险防控措施时应多关注研究区中部。中部Cd、Pb和Cr浓度高, 人类活动复杂, 高-高区分布数量多。

参考文献

- [1] 王勇漠. 土壤重金属污染特征与状况调查分析[J]. 西部探矿工程, 2024, 36(10): 12-15.
WANG Y H. Investigation and analysis of soil heavy metal pollution characteristics and conditions [J]. West-China Exploration Engineering, 2024, 36(10): 12-15.
- [2] 董志芬, 梁策, 谢容生, 等. 有色冶炼企业地块非饱和带重金属分布迁移特征及影响因素[J]. 有色金属(冶炼部分), 2023(11): 135-143.
DONG Z F, LIANG C, XIE R S, et al. Distribution and migration characteristics and influencing factors of heavy metals in unsaturated zone of metallurgical enterprise [J]. Site Nonferrous Metals (Extractive Metallurgy), 2023(11): 135-143.
- [3] ADHIKARI K, MANCINI M, LIBOHOVA Z, et al. Heavy metals concentration in soils across the conterminous USA: Spatial prediction, model uncertainty, and influencing factors [J]. Science of the Total Environment, 2024, 919: 170972.
- [4] VASUDHEVAN P, MANIKANDAN E, JONATHAN M P, et al. Pollution assessment and source apportionment of metals in paddy field of Salem, South India [J]. Environmental Earth Sciences, 2022, 81(6): 184.
- [5] 张雨涵, 李瑜, 官开江, 等. 重庆某农业集中区土壤重金属污染评价与来源解析[J]. 中国土壤与肥料, 2024(7): 10-18.
ZHANG Y H, LI Y, GUAN K J, et al. Pollution assessment and source analysis of soil heavy metals in a concentrated distribution area of agricultural in Chongqing [J]. Soil and Fertilizer Sciences in China, 2024(7): 10-18.
- [6] 杨昱莹, 刘亮, 陈明, 等. 长三角地区南京市表土重金属污染特征及源解析[J]. 中国环境科学, 2024, 44(7): 3910-3918.
YANG Y Y, LIU L, CHEN M, et al. Characterization and source analysis of topsoil heavy metal pollution in Nanjing, Yangtze River Delta Region [J]. China Environmental Science, 2024, 44(7): 3910-3918.
- [7] 岳聪, 黄顺寅, 屠春宝, 等. 植物修复重金属镉污染土壤研究

- 进展[J]. 现代农业科技, 2024(17): 129-135.
- YUE C, HUANG S Y, TU C B, et al. Research progress on phytoremediation of soil contaminated by heavy metal cadmium [J]. Modern Agricultural Science and Technology, 2024(17): 129-135.
- [8] 李军, 高占栋, 马利邦, 等. 黄河兰州段城市饮用水源地土壤重金属污染及其溯源的多指标综合分析[J]. 环境科学, 2024, 45(11): 6723-6733.
- LI J, GAO Z D, MA L B, et al. Multiproxy comprehensive analysis for source apportionment and pollution of heavy metals in urban drinking-water source soils from the Lanzhou Reach of the Yellow River [J]. Environmental Science, 2024, 45(11): 6723-6733.
- [9] 赵恒谦, 常仁强, 金倩, 等. 河北西石门铁矿区土壤重金属污染空间分析及风险评价[J]. 岩矿测试, 2023, 42(2): 371-382.
- ZHAO H Q, CHANG R Q, JIN Q, et al. Spatial analysis and risk assessment of soil heavy metal pollution in the Xishimen iron mining area of Hebei Province [J]. Rock and Mineral Analysis, 2023, 42(2): 371-382.
- [10] 宋炉生, 孙振洲, 胡晶, 等. 废弃铁矿及下游农田土壤重金属污染特征及来源解析[J]. 环境工程, 2024, 42(10): 155-164.
- SONG L S, SUN Z Z, HU J, et al. Pollution characteristics and source apportionment of heavy metals in an abandoned iron ore and downstream farmland soil [J]. Environmental Engineering, 2024, 42(10): 155-164.
- [11] 董曼慧, 夏卫生, 周浩, 等. 湖南省地块尺度土壤重金属污染及成因研究[J]. 环境科学与技术, 2023, 46(增刊1): 60-66.
- DONG M H, XIA W S, ZHOU H, et al. Study on the pollution status and causes of heavy metals in soil at plot scale in Hunan Province [J]. Environmental Science & Technology, 2023, 46(S1): 60-66.
- [12] 孙思静, 董春雨, 张好, 等. 基于PMF模型和地理探测器的土壤重金属源解析及影响因素分析[J]. 环境科学, 2024, 45(9): 5474-5484.
- SUN S J, DONG C Y, ZHANG H. Source and influencing factor analysis of soil heavy metals based on PMF model and GeoDetector [J]. Environmental Science, 2024, 45(9): 5474-5484.
- [13] 耿治鹏, 宋颀, 王春林, 等. 污染场地土壤重金属污染空间特征分析—以某搬迁电镀厂为例[J]. 环境工程技术学报, 2023, 13(1): 295-302.
- GENG Z P, SONG J, WANG C L, et al. Spatial characteristics of soil heavy metal pollution in polluted sites: taking a relocated electroplating factory as an example [J]. Journal of Environmental Engineering Technology, 2023, 13(1): 295-302.
- [14] ZHENG J, WANG P, SHI H, et al. Quantitative source apportionment and driver identification of soil heavy metals using advanced machine learning techniques [J]. Science of the Total Environment, 2023, 873: 162371.
- [15] SHI T, MA J, WU X, et al. Inventories of heavy metal inputs and outputs to and from agricultural soils: a review [J]. Ecotoxicology and Environmental Safety, 2018, 164: 118-124.
- [16] 许洋, 陈健松, 王志栋, 等. 基于多源异构数据的典型场地土壤重金属污染模拟预测研究[J]. 环境科学学报, 2023, 43(9): 357-68.
- XU Y, CHEN J S, WANG Z D, et al. Simulation and prediction research of heavy metal pollution in soil of typical sites based on multi-source heterogeneous data [J]. Acta Scientiae Circumstantiae, 2023, 43(9): 357-68.
- [17] 李珊, 杨济妮, 苏贵金, 等. 基于Catboost算法的中国典型农业区重金属污染特征及影响因素分析[J]. 环境化学, 2024, 43(10): 3377-3387.
- LI S, YANG J N, SU G J, et al. Analysis of heavy metal pollution characteristics and influencing factors in China's typical agricultural areas based on Catboost algorithm [J]. Environmental Chemistry, 2024, 43(10): 3377-3387.
- [18] GUI H, YANG Q, LU X, et al. Spatial distribution, contamination characteristics and ecological-health risk assessment of toxic heavy metals in soils near a smelting area [J]. Environmental Research, 2023, 222: 115328.
- [19] 冯锋, 王育红, 左雨芳. 基于RF-XGBoost的土壤镉污染影响因素及空间分布研究[J]. 农业环境科学学报, 2023, 42(4): 811-819.
- FENG F, WANG Y H, ZUO Y F. A study on factors that influence the spatial distribution of soil cadmium pollution based on RF-XGBoost [J]. Journal of Agro-Environment Science, 2023, 42(4): 811-819.
- [20] 王爱琼, 萨日娜, 张巧丽, 等. 基于随机森林的川崎病诊断关键实验室指标预测研究[J]. 检验医学与临床, 2024, 21(19): 2887-2891.
- WANG A Q, SA R N, ZHANG Q L, et al. Prediction of key laboratory indicators for diagnosis of Kawasaki disease based on random forest [J]. Laboratory Medicine and Clinic, 2024, 21(19): 2887-2891.
- [21] BI Z, SUN J, XIE Y, et al. Machine learning-driven source identification and ecological risk prediction of heavy metal pollution in cultivated soils [J]. Journal of Hazardous Materials, 2024, 476: 135109.
- [22] ZHOU M, LIAO B, SHU W, et al. Pollution Assessment and Potential Sources of Heavy Metals in Agricultural Soils around Four Pb/Zn Mines of Shaoguan City, China [J]. Soil & Sediment Contamination, 2015, 24(1): 76-89.
- [23] Zhen J, Pei T, Xie S. Kriging methods with auxiliary nighttime lights data to detect potentially toxic metals concentrations in soil [J]. Science of the Total Environment, 2019, 659: 363-371.
- [24] 李辉, 翟星, 李琛曦, 等. 河北省泥石流灾害易发性云模型评价方法: 以邢台赵沟村泥石流为例[J]. 科学技术与工程, 2024, 24(25): 10884-10891.
- LI H, ZHAI X, LI C X, et al. Development law of debris flow disaster and the evaluation method of its susceptibility: take

- mudslide in Zhaogou Village, Xingtai as an example [J]. Science Technology and Engineering, 2024, 24(25): 10884–10891.
- [25] GUO Q, LI B, CHEN Y, et al. Intelligent Model for the Compressibility Prediction of Cement-Stabilized Dredged Soil Based on Random Forest Regression Algorithm [J]. KSCE Journal of Civil Engineering, 2021, 25(10): 3727–3736.
- [26] 张慧, 栾思雨, 丛蓉. 东北黑土区典型县域耕地质量对耕地水田化的空间响应 [J]. 水土保持研究, 2024, 31(01): 327–334.
- ZHANG H, LUAN S Y, CONG R. Spatial response of cultivated land quality to paddy farming in typical counties of northeast black soil region [J]. Research of Soil and Water Conservation, 2024, 31(1): 327–334.
- [27] 刘宝涛, 杨兴龙, 郭佳. 东北地区城镇化发展与土地健康利用空间关联特征 [J]. 水土保持研究, 2018, 25(4): 386–392.
- LIU B T, YANG X L, GUO J. Spatial correlation characteristics between urbanization development and land health utilization [J]. Research of Soil and Water Conservation, 2018, 25(4): 386–392.
- [28] CHEN D, WANG X, LUO X, et al. Delineating and identifying risk zones of soil heavy metal pollution in an industrialized region using machine learning [J]. Environmental Pollution, 2023, 318: 120932.
- [29] LI J, ZHONG M, JIANG L, et al. Assessing and mitigating health risks of workers exposed to volatile organic compounds in contaminated soils during active pit excavation: Accounting for exposure variability and uncertainty [J]. Journal of Cleaner Production, 2024, 469: 143227.
- [30] 钱江波, 陈涤, 王夏晖, 等. 基于机器学习的区域土壤重金属污染风险诊断 [J]. 环境工程, 2023, 41(12): 296–303.
- QIAN J B, CHEN D, WANG X H, et al. Risk diagnosis of heavy metal pollution in regional soil based on machine learning [J]. Environmental Engineering, 2023, 41(12): 296–303.
- [31] HUANG G, WANG X, CHEN D, et al. A hybrid data-driven framework for diagnosing contributing factors for soil heavy metal contaminations using machine learning and spatial clustering analysis [J]. Journal of Hazardous Materials, 2022, 437: 129324.
- [32] DING Y, YAO X, SU Z, et al. Research and engineering application of efficient pretreatment technology of initial rainwater in artificial wetland of the coastal industrial park [J]. Fresenius Environmental Bulletin, 2022, 31(6A): 6290–6304.
- [33] HOU S, ZHENG N, TANG L, et al. Effect of soil pH and organic matter content on heavy metals availability in maize (*Zea mays* L.) rhizospheric soil of non-ferrous metals smelting area [J]. Environmental Monitoring and Assessment, 2019, 191(10): 634.
- [34] 龙於洋, 邱钧健, 宋禹翮, 等. 危险废物焚烧灰-渣共热处理对 Cr 和 Pb 的原位固定 [J]. 环境科学学报, 2023, 43(6): 390–397.
- LONG Y Y, QIU J J, SONG Y Y, et al. In-situ immobilization of Cr and Pb by co-thermal treatment of hazardous waste incineration ash-slag [J]. Acta Scientiae Circumstantiae, 2023, 43(6): 390–397.