

DOI:10.7524/j.issn.0254-6108.2023041502

CSTR:32061.14.hjhx.2023041502

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 Shan, YANG Jini, SU Guijin, 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.

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*

李珊^{1,2} 杨济妮³ 苏贵金² 李倩倩² 孟晶² 史斌^{2**}

(1. 郑州大学河南先进技术研究院, 郑州, 450003; 2. 环境纳米技术与健康效应重点实验室, 中国科学院生态环境研究中心, 北京, 100085; 3. 生态环境部对外合作与交流中心, 北京, 100035)

摘要 土壤是保障人类生存和发展的物质基础, 农田土壤与人类生产生活密切相关, 农田土壤污染问题值得关注. 本文选取东北三省、京津冀和长江经济带作为研究区域, 通过文献调研筛选出 2000—2020 年发布的 400 篇相关文献, 收集到 2052 表层土壤样本, 通过统计学、空间分析和机器学习等方法对其污染特征进行分析. 结果表明, 与农用地最严格的筛选值标准相比, 3 个研究区域 Cd 污染问题较为突出, 东北三省、京津冀和长江经济带的超标率分别为 37.5%、34.0% 和 45.8%, 长江经济带也存在 Cu 污染问题, 其超标率为 30.6%; 进一步将采样点位分为矿区周边点位、市郊区点位和其他农田点位, 采用潜在生态风险指数法评价表明, 矿区周边样点严重和重风险占比 (54.8%) 显著高于市郊区样点 (37.2%) 和其他农田样点 (36.6%). 采用 Catboost 模型进行影响因子识别发现, 矿区周边农田主要受选矿、尾矿暴露等造成的 Cd 和 Hg 污染影响, 而市郊区农田主要受商业、工业和交通运输业等人类活动带来的 Hg 污染影响, 其他农田土壤中主要受农药化肥的施用以及大型器械化带来的 Cd 污染影响.

关键词 农田土壤, 重金属, 污染特征, Catboost 模型, 影响因素.

中图分类号 X-1; O6 文献标识码 A

Analysis of heavy metal pollution characteristics and influencing factors in China's typical agricultural areas based on Catboost algorithm

LI Shan^{1,2} YANG Jini³ SU Guijin² LI Qianqian² MENG Jing² SHI Bin^{2**}

(1. Henan Institutes of Advanced Technology, Zhengzhou University, Zhengzhou, 450003, China; 2. Key Laboratory of Environmental Nanotechnology and Health Effect, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing, 100085, China; 3. Foreign Cooperation and Exchange Center of the Ministry of Ecology and Environment, Beijing, 100035, China)

Abstract Soil is the material foundation for human survival and development. Agricultural soil is closely related to human production and life, and the issue of agricultural soil pollution deserves attention. This study focuses on the research areas of the three northeastern provinces, the Beijing-Tianjin-Hebei region, and the Yangtze River Economic Belt. By conducting literature research, 400

2023 年 4 月 15 日收稿(Received: April 15, 2023).

* 国家重点研发计划 (2018YFC1800300) 和国家自然科学基金 (42277387) 资助.

Supported by National Key Research and Development Program (2018YFC1800300) and Natural Science Foundation of China (42277387).

** 通信联系人 Corresponding author, E-mail: binshi@rcees.ac.cn

relevant articles published between 2000 and 2020 were selected, and 2,052 topsoil samples were collected. Pollution characteristics were analyzed using statistical, spatial analysis, and machine learning methods. The results indicate that, compared to the strictest screening values for agricultural land, Cd pollution is more prominent in the three research areas, with exceedance rates of 37.5%, 34.0%, and 45.8% for the three regions, respectively. The Yangtze River Economic Belt also exhibits Cu pollution with an exceedance rate of 30.6%. Further classifying the sample sites into mining area adjacent sites, suburban area sites, and other farmland sites, an assessment using the potential ecological risk index (PERI) method shows that mining area adjacent sites have significantly higher proportions of severe and high-risk pollution (54.8%) compared to suburban area points (37.2%) and other agricultural field points (36.6%). The Catboost model was employed to identify influencing factors, revealing that agricultural fields adjacent to mining areas are mainly affected by Cd and Hg pollution caused by ore selection and tailings exposure. Suburban area fields are primarily influenced by Hg pollution resulting from human activities such as commerce, industry, and transportation. Other agricultural fields' soil is mainly affected by Cd pollution due to the application of pesticides and fertilizers and the use of large-scale mechanization.

Keywords agricultural soils, heavy metals, contamination characteristics, catboost model, influencing factors.

土壤是陆地生态系统的重要组成部分,可提供维系植物生长的营养物质和微生物生存的家园,是保障人类生存和发展的物质基础^[1]。然而,随着城市化进程的加快,矿产资源开发、金属加工冶炼、化工生产、污水灌溉以及不合理的化肥农药施用等因素导致重金属在农田土壤中不断富集^[2],由此造成的粮食减产等问题引起关注。土壤中重金属主要来源于自然因素和人为因素。其中,自然因素主要为成土母质,包括土壤自身理化性质的不同所产生的土壤背景值的差异性,同时区域土壤侵蚀程度的不同也会影响成土母质中重金属的释放。人为因素主要由工业排放、农业活动和交通运输等人类活动等途径产生并在土壤中累积。重金属的过量累积对土壤孔隙结构、土壤氧化还原环境以及土壤微生物活性产生影响,从而危害整个农田生态安全。

农田土壤与人类生产生活密切相关,一旦受到污染不仅会造成农作物产量和品质的下降,更会对人类健康产生威胁。中国土壤污染总的超标率为 16.1%,其中无机型污染即 8 种重金属污染为主要类型,占全部超标位点的 82.8%^[3]。多项研究表明多种重金属对人体有强烈的致癌致畸毒性:铅中毒会引起人体贫血、便秘、腹痛、呕吐及食欲减退等的临床症状,儿童中毒时会出现智力障碍和行为异常等症状^[4]。长期暴露在镉(Cd)污染的环境下会导致肺癌、肾功能发育不良和骨折等人体疾病的高发^[5]。早在 20 世纪 30 年代,日本富山县就发生过严重的重金属镉污染稻米引起的数百人骨痛病的土壤污染公害事件^[6]。另外,食用铬(Cr)超标的农产品,会引起克汀病、肝癌等人体疾病。因此,如果农田土壤受到重金属污染,由此产生的食品安全问题将难以忽视,加强农田土壤重金属来源因素的研究是区域重金属风险防控及监测管控的必要条件和重要基础。

目前对于农田土壤重金属污染的研究大多集中在小区域农田,对于大区域污染源问题仍在探索^[7]。近年来,机器学习能够分析土壤特征和重金属浓度大数据集,识别出与重金属浓度密切相关的因素,在处理非线性问题上表现出优势。目前,已有随机森林、支持向量机、神经网络等多种机器学习方法被广泛应用于土壤重金属预测和监测农田重金属污染方面,随着越来越多的数据积累,这些技术准确度和可信度都将得到提升^[8-10]。Catboost 模型是一种机器学习算法,能够处理高维度数据并自动选择最重要的特征指标,降低模型的复杂度,在噪声数据或异常值时也能够保持较好的预测能力具有较好的鲁棒性。相对于其他机器学习算法,Catboost 模型的训练速度较快,能够自动处理缺失数据,减少因数据缺失而导致的误差。因此,在土壤重金属源解析中,Catboost 模型是一种具有潜力和优势的机器学习算法。

本文选择中国具有代表性的三大区域,东北三省区域、京津冀和长江经济带作为研究区域。综述了三大区域的农田重金属研究文献,分析了三大区域重金属的污染特征和潜在生态风险特征,整合与

土壤重金属关系密切的辅助因素,通过 CatBoost 模型分析对农田土壤重金属风险因子进行影响评价,为农田土壤重金属影响因素风险评估提出一种分析方法,以期为我国农田土壤污染防治和环境管理提供一定的评估思路。

1 材料和方法 (Materials and methods)

1.1 研究区域

东北三省区域、京津冀和长江经济带作为研究区域。东北三省(包括黑龙江、吉林、辽宁)是中国的重要粮食生产区之一,2021 年粮食总产量为 1445.6 万吨,占全国总产量超过 21%^[12]。东北三省的主要产业有重工业、化工、冶金、机械制造、煤炭等,工业用地面积从 2007 年的约 20840 km² 增加到了 2019 年的约 33315 km²,增加了近 60.0%。农用地面积从 2007 年约 148780 km² 减少至 2019 年 64573 km²,减少了近 60.0%^[11-12]。长江经济带包括长江沿岸的 11 个省市,占全国总面积 21.4%,是我国重要的水稻、小麦、玉米等粮食作物和棉花、油菜等经济作物的生产基地。该区域的工业发展经历了从传统制造业向高技术制造业的转变,江苏、浙江、安徽、湖北等地区的工业总产值占全国工业总产值的比重均超过 6%^[13]。2019 年长江经济带地区的农用地总面积为 381740 km²,比 2007 年 1701010 km² 减少了 77.5%。工业用地由 2007 年的 120390 km² 增加至 136488 km² ^[11-12]。京津冀区域拥有以北京和天津为中心的城市群,是我国重要的钢铁和有色金属生产基地,石化和化工、机械和装备制造、电子信息等行业也在该地区占据重要地位。京津冀地区城市化的不断发展占据了大量的农用地资源,农业用地的面积从 2007 年的 148780 km² 减少到了 2019 年的 64573 km²,减少了约 56.6%。工业用地的面积从 2007 年的 24750 km² 增加到了 2019 年的 24787 km² ^[11-12]。

1.2 数据来源

在中国知网 (CNKI) 和 Web of Science (WOS) 上收集了 2000—2020 年发表的关于京津冀、长江经济带和东北三省区域的农田重金属研究论文。以“农田”、“农用地”、“农业”、“重金属”和单个重金属名称的中英文名称为关键词搜索,对收集到的文献数据按照以下条件进行筛选,将符合标准的数据纳入分析:(1) 采集的样品为农田耕层表层土壤样品 (0—15 cm 或者 0—20 cm);(2) 有明确的研究区域、有具体的采样点位置以及采样时间;(3) 土壤样品的采集、预处理和分析方法要使用行业认证的方法;共筛选收集到 2000—2020 年共 400 篇文献,包括东北三省区域 71 篇,京津冀 62 篇,长江经济带 267 篇,数据主要集中在 2006—2009 年和 2013—2019 年。

1.3 污染评价指标

潜在生态风险指数法考虑不同元素的毒性和环境敏感性差异,使用毒理学知识来解释重金属元素的生态效应和环境效应,为土壤环境质量评估提供更准确、科学、合理的参考依据,其计算公式如下:

$$PERI(i) = \frac{S_m^i}{S_b^i} \times T_r^i \tag{1}$$

$$PERI = \sum PERI(i) \tag{2}$$

$$R(i) = \frac{PERI(i)}{PERI} \tag{3}$$

式中,PERI(*i*) 为土壤中重金属元素 *i* 的单潜在生态风险指数;*S_mⁱ* 和 *S_bⁱ* 分别为重金属元素 *i* 在土壤中的实测值和参比值,参比值采用研究区土壤背景值;*T_rⁱ* 为重金属元素 *i* 的毒性响应系数,铬、铅、镉、汞、砷、铜、镍和锌元素的毒性响应系数分别是 2、5、30、40、10、5、5 和 1;PERI 是总生态风险指数,为潜在生态风险指数为各元素风险指数和。潜在生态风险指数划分标准如表 1 所示。

1.4 风险因子重要性评估方法

CatBoost 模型是一种基于梯度提升的决策树算法,能够高效合理地处理类别型特征,可自动将类别特征处理为数值型特征,并将特征之间的联系进行组合,丰富了特征的维度。Catboost 通过采用排序提升 (Ordered boosting) 代替梯度估计法减轻由于梯度估计偏差造成的预测偏移,以及改进标称属性训

练方法, 解决由于贪心算法(greedy)导致的条件偏移^[14-15]. CatBoost 模型在处理类别特征时, 将类别特征进行标签化, 将标签对应的平均值作为节点分裂的标准, 在分类的基础上添加先验分布项做平滑处理, 减少噪声和低频率数据对数据分布的影响. 其计算公式如下:

$$\hat{x}_k^j = \frac{\sum_{j=1}^{p-1} \left[x_{\sigma_j,k} = x_{\sigma_{p,k}} \right] Y_{\sigma_j} + a \times p}{\sum_{j=1}^{p-1} \left[x_{\sigma_j,k} = x_{\sigma_{p,k}} \right] + a}$$

(4)

其中, $\hat{x}_k^j$ 表示标签化的类属特征 i , 使用不包含样本 x_i 的训练集生成一个随机序列 $\sigma = (\sigma_1, \cdots, \sigma_{i-1})$, 从 σ_1 到 σ_{i-1} 一次遍历随机序列, 用遍历到的前 p 个记录 $x_{\sigma_{p,k}}$ 计算标称特征的数值. p 表示添加的先验项, a 是权重系数, 其取值均大于 0.

表 1 潜在生态风险指数划分标准

Table 1 Criteria for classifying potential ecological risk indices

分级 (单) Classification (Single)	标准 Standard	分级 (总) Classification (Total)	标准 Standard
低	<40	低度	<150
中	40—80	中度	150—300
较重	80—160	重度	300—600
重	160—320	严重	≥600
严重	≥320		

本研究通过调用 python 软件中的 Catboost 包进行建模和贝叶斯参数优化, 主要使用的函数为 CatBoost 和 Bayes-opt.

2 结果与讨论 (Results and discussion)

2.1 区域农田重金属含量特征

共收集整理了 2052 组数据, 统计了八种常见的重金属的含量范围、中位值、标准差和变异系数, 如表 2 所示. 从整体上看, 长江经济带的八种重金属浓度均明显高于东北三省区域和京津冀区域, 东北三省区域和京津冀区域 Cd 和 Hg 的变异系数均超过 100%, 长江经济带 Cd、Hg、Zn 和 As 的变异系数均超过 100%, 说明样点重金属含量连续性变化较差, 样点重金属具有较强的空间变异性, 属于外界因素输入影响的可能性较大.

表 2 研究区域农田土壤重金属含量特征统计

Table 2 Statistics on the heavy metal content of agricultural soils in the study area

研究区域 Study area	重金属 Heavy metals	浓度范围/(mg·kg ⁻¹) Concentration range	平均浓度/(mg·kg ⁻¹) Average concentration	变异系数/% CV	筛选值/(mg·kg ⁻¹) Filter value	超标率/% Excess rate
东北三省 区域	Cr	1.5—155.0	59.9	48.3	150.0	6.3
	Ni	1.5—56.0	26.8	36.5	60.0	4.0
	Cu	1.5—92.5	26.9	71.6	50.0	11.4
	Cd	0.0—4.4	0.5	163.2	0.3	37.5
	Pd	1.2—89.5	26.1	65.1	80.0	8.1
	Hg	0.0—1.0	0.2	148.21	0.5	9.1
	Zn	0.6—367.1	97.3	78.4	200.0	13.3
	As	1.7—24.9	10.6	57.4	25.0	14.3
京津冀	Cr	0.0—99.5	56.1	43.6	150.0	2.1
	Ni	0.0—52.0	25.8	41.3	60.0	0.0
	Cu	0.0—141.1	31.5	77.5	50.0	14.0
	Cd	0.03—1.79	0.3	106.2	0.3	34.0
	Pd	0.0—85.9	26.9	59.7	80.0	7.7

续表 2

研究区域 Study area	重金属 Heavy metals	浓度范围/(mg·kg ⁻¹) Concentration range	平均浓度/(mg·kg ⁻¹) Average concentration	变异系数/% CV	筛选值/(mg·kg ⁻¹) Filter value	超标率/% Excess rate
京津冀	Hg	0.0—1.2	0.2	141.4	0.5	8.6
	Zn	0.0—255.9	87.1	54.0	200.0	5.4
	As	0.0—30.0	9.4	57.6	25.0	7.3
长江经济带	Cr	0.8—232.5	76.1	54.2	150.0	6.3
	Ni	0.9—89.5	36.6	43.0	60.0	13.3
	Cu	1.2—214.5	48.2	83.5	50.0	30.6
	Cd	0.0—9.1	0.7	148.4	0.3	45.8
	Pd	0.9—163.0	39.7	54.7	80.0	9.8
	Hg	0.0—4.3	0.3	169.1	0.5	9.5
	Zn	11.8—422.6	122.9	64.8	200.0	12.4
	As	0.1—21.3	13.4	63.5	25.0	14.9

将上述数据以 5 年为 1 个周期划分为 4 个时期 (图 1), 发现在 20 年的研究期间, 重金属 As、Cd、Pb 含量持续上升, 而 Ni 和 Cr 的含量则没有明显的变化规律. 然而, 自 2015 年以后, Hg、Cu 和 Zn 的含量呈下降趋势. 这一现象的原因一方面与我国污染防治工作的全面推进有关, 包括环保执法力度的提升和高污染工艺的改造, 从源头上减少了污染的排放. 另一方面, 不同重金属在土壤中的迁移赋存特征的差异导致呈现出各异的变化趋势^[16], 其中 As、Cd、Pb 重金属多吸附-沉淀在土壤颗粒表面, 迁移现象不明显, 因此其浓度持续上升. 而 Cr、Hg 在土壤中受吸附-沉淀作用影响较小, 进入土壤后能够依靠分子扩散-机械弥散进行迁移, 因此在污染防治措施实施切断污染源后其在土壤中的浓度有逐渐减少的趋势^[17-18].

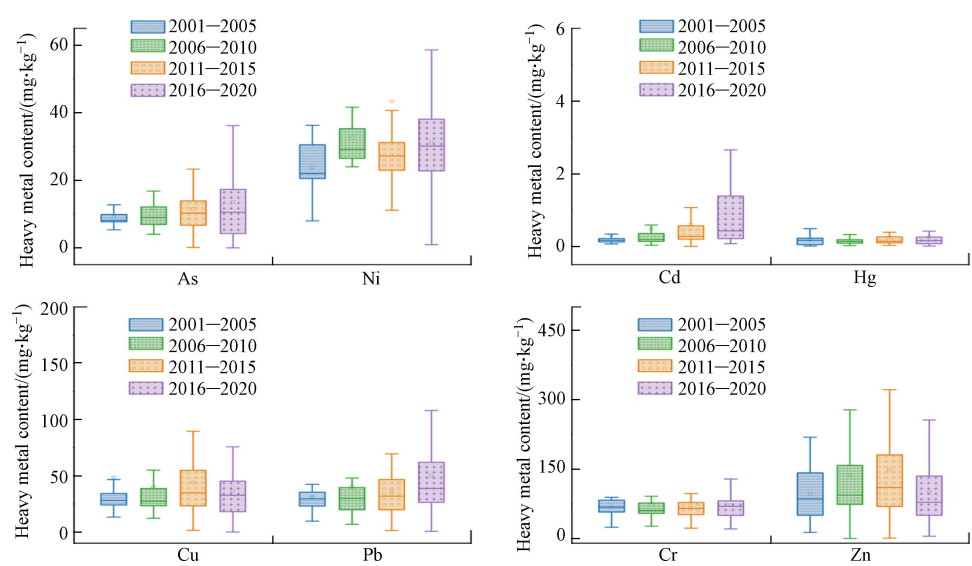


图 1 重金属时间维度浓度变化图

Fig.1 Concentration distribution of heavy metals over time

以《土壤环境质量农用地土壤污染风险管控标准》(GB 15618—2018)中较严格的筛选值为评价标准, 发现研究区域范围内重金属 Cd 存在超标现象, 其中 Cd 在长江经济带的超标率达到 45.8%, Cd 污染程度较为严重的样品主要采集自长江经济带上游区域, Cd 含量高不仅与区域矿区活动相关, 也与该区域有较高的地质背景值相关, 长江经济带上游区域中云南省 (0.22 mg·kg⁻¹) 和贵州省 (0.66 mg·kg⁻¹) Cd 土壤背景值相对于中国土壤背景值 (0.074 mg·kg⁻¹) 明显偏高^[19]. 长江经济带还存在 Cu 超标, Cu 含量超过农用地标准样点占比为 30.6%, Cu 污染程度较严重的样品采集自西南地区矿区周围农田土壤^[20], 西南区域矿资源丰富, 矿山开采历史悠久, 铜矿多为中小型矿床, 小规模开采、开采技术水平低等原

因,是造成区域浓度较高的原因之一.东北三省 Cd 超标率达 37.5%,Cd 含量的平均值为农用地标准筛选值的 2.94 倍,污染程度最大的样品采集自葫芦岛市,葫芦岛市有色金属工业发达,锌冶炼过程中产生的烟尘颗粒物中含有 Cd,是土壤中重金属的主要来源^[21].京津冀区域样点中 Cd 浓度超过农用地标准值筛选值的样点占比达 34.0%,化石燃料燃烧、工业生产和交通运输等人类活动引起的大气沉降是主要因素^[22].

正常情况下,土壤重金属来源于自然地质背景和人类活动两种途径,各个区域土壤在正常地质地理条件和地球化学条件下会产生一定量的金属元素,成土母质和岩石风化等自然来源对土壤重金属含量有很大影响,同时风力和水力搬运的土壤侵蚀过程、土壤母质风化过程以及土壤黏粒组成等因素也会对土壤重金属浓度值产生较大影响.为了剔除自然背景值对重金属含量的影响,参考 1990 年出版的我国土壤背景值.表 3 给出了各个研究区域重金属背景值及样点浓度超过区域背景值的占比,从表中可以看出所有的 8 种重金属均存在一定的超过背景值的现象(占比 0.7%—97.4%),Cd、Ni、Cu 和 Hg 四种重金属在 3 个区域超过背景值的比例最大,均在 60.0% 以上.除了上述 4 种重金属外,长江经济带的 Zn 和东北三省区域的 As 和 Zn 也值得关注.

表 3 研究区域农田土壤重金属背景浓度(mg·kg⁻¹)及超过背景值的比例(%)

Table 3 Background concentrations of heavy metals in the study area (mg·kg⁻¹) and the proportion exceeding background values (%)

重金属 Heavy metals	东北三省 Heilongjiang-Jilin-Liaoning		京津冀 Beijing-Tianjin-Hebei		长江经济带上游 ¹ Upstream region of the Yangtze River Economic Belt		长江经济带下游 Downstream region of the Yangtze River Economic Belt	
	背景值 Background value	比例/% Proportion	背景值 Background value	比例/% Proportion	背景值 Background value	比例/% Proportion	背景值 Background value	比例/% Proportion
Cr	51.1	0.7	71.2	29.2	79.8	47.1	67.2	56.0
Cd	0.1	90.9	0.1	94.1	0.3	64.7	0.1	90.9
Ni	21.9	83.3	29.9	96.0	36.7	67.9	28.4	67.5
Cu	17.9	71.4	23.9	65.9	35.1	72.4	23.7	83.8
Hg	0.0	95.5	0.0	94.4	0.1	92.1	0.1	71.5
Pb	23.0	59.5	21.9	58.5	14.8	89.7	11.3	97.4
As	7.0	60.7	10.5	23.8	34.4	69.7	27.2	11.6
Zn	66.9	64.5	81.9	43.2	90.6	44.4	74.8	75.5

¹ 长江经济带上游包括云南、四川、贵州、重庆,下游包括湖北、湖南、安徽、江西、江苏、浙江和上海.

¹ The upstream region of Yangtze River Economic Belt encompasses Yunnan, Sichuan, Guizhou and Chongqing. Its downsream includes Hubei, Hunan, Anhui, Jiangxi, Jiangsu, Zhejiang and Shanghai.

长江经济带中上游地区中重金属主要来源于工农业生产和矿山开采,而位于中下游的湖南、湖北、江西、江苏和上海等地的土壤重金属主要来自农业活动,因此将长江经济带分为上游和下游区域讨论.表 3 数据显示长江经济带上游区域超过背景值占比最高的重金属为 Hg 和 Pb,下游超过背景值占比最高的重金属为 Cd 和 Pb.京津冀区域超过背景值占比最高的重金属为 Cd、Hg 和 Ni,东北三省区域超过区域背景值占比最高的两种重金属为 Cd 和 Hg.

分析其原因,人类的工业生产和农业活动等过程对土壤重金属产生的影响很大,工业活动、污水灌溉以及化肥农药施用等方式产生的重金属直接或间接通过大气沉降、降雨、地表径流以及地下径流等方式进入土壤.

首先,采矿、金属冶炼加工会引起 Cd、Pb 和 Hg 等重金属污染,化工、机械制造、电镀及火力发电等工业活动会释放大量的 Cd、Cu、Zn、As 和 Hg 等重金属元素.过去的 40 年,工矿企业粗放式的增长,其污水、尾矿渣、粉尘污染直接排放进入到土壤和环境当中.采矿、石化、金属冶炼、加工高耗能的产业占工业总产值的 78.3%,由此产生的累计量污染量不可忽视.另外,汽车尾气排放、垃圾焚烧以及能源、冶金和建筑材料粉尘等因素产生的重金属微粒通过特定途径沉降到农田土壤,进一步加剧土壤重金属污染^[2].

其次,农业生产过程中,农药和化肥的过量使用也是导致农田重金属浓度升高的主要因素. 磷肥、氮肥和复合肥料中含有一定量的 As、Cd、Pb、Zn、Cr 和 Hg 等重金属, 早期施用的杀虫剂等农药中也含有 Hg、Cd、Cu 和 As 等重金属^[23]. 如图 2(a)、(b) 为三大研究区域近 20 年农药化肥施用统计图, 表 4 为本文文献调研中 2000—2020 年间重金属年均含量与农药化肥年总施用量的相关系数表, 如表 4 所示, 化肥的施用量与 As、Hg 和 Ni 三种重金属的年均浓度变化相关性较高, 相关系数分别为 0.70、0.47 和 0.42. 农药的施用量与 As 和 Pb 两种重金属的年均浓度变化相关性较高, 分别为 0.68 和 0.42.

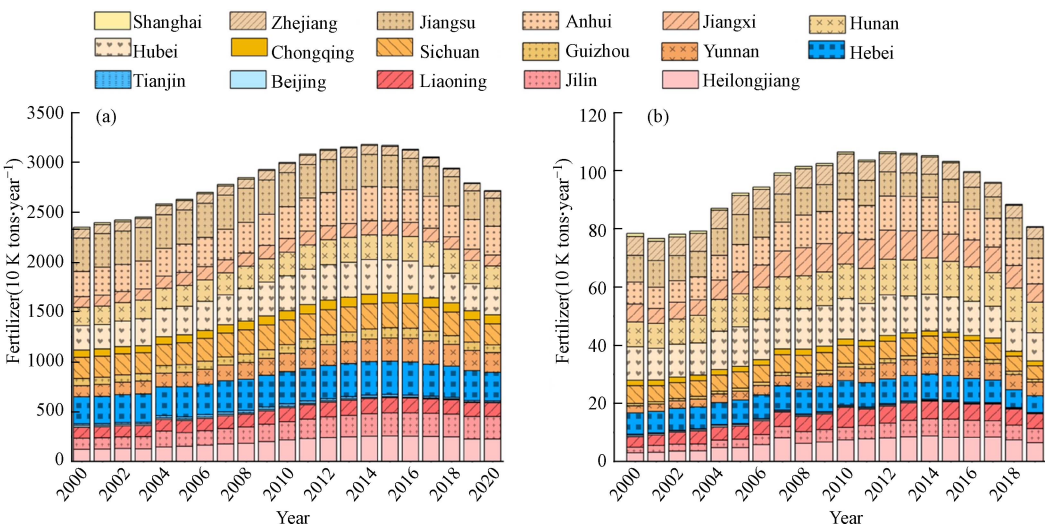


图 2 (a)研究区域 2000-2020 年化肥施用统计图; (b) 研究区域 2000—2020 年农药施用统计图

Fig.2 (a) Statistical map of fertilizer application in the study area, 2000—2020; (b) Pesticides in the study area 2000-2020 Statistical map of pesticide application

表 4 重金属浓度与农药化肥施用量 Person 相关系数 ($P<0.05$)

种类 Species	重金属 Heavy metals							
	Cr	Ni	Cu	Cd	Pb	Zn	As	Hg
化肥施用量	0.05	0.42	0.14	0.25	0.32	0.19	0.70	0.47
农药施用量	0.13	0.31	0.26	0.14	0.42	0.36	0.68	0.16

2.2 潜在生态风险指数

根据采样点所处位置不同将采样点位分为矿区周边农田 (24.4%)、市郊区农田 (22.7%) 和其他农田 (52.9%) 3 类 (图 3), 矿区周围农田土壤中重金属主要来源于采矿冶炼过程中产生的工业废弃物、裸露的废弃矿坑以及矿车运输过程等, 市郊区农田土壤中重金属主要来源于工业生产、商业活动, 人类生活所产生的生活垃圾以及交通运输所产生的尾气等, 其他农田土壤中重金属主要来源于农业活动, 农药化肥的施用、大型农业器械的使用等.

选取潜在生态风险指数法对 3 类区域的单个重金属生态风险指数和综合生态风险指数进行评估 (图 4). 从综合风险指数来看, 矿区周边农田土壤样点里重、严重风险土壤样点占比 (54.8%) 显著高于市郊区样点 (37.2%) 和其他农田土壤样点 (36.6%). 从表 5 中单生态风险指数来看 (样点单生态风险指数均值), 3 种样点区域中 Ni 和 Zn 的单生态风险均较低, Cd 单生态风险均较高. 矿区周边农田样点中 $PERI(Cr)=48.47$ 分别为市郊区农田样点 $PERI(Cr)=1.88$ 的约 25 倍和其他农田样点 $PERI(Cr)=1.78$ 的 27 倍, Cd 单生态风险相较于市郊区和农田土壤样点明显偏高, 同时 Hg 也存在一定的生态风险. 市郊区农田样点中 $PERI(Hg)=155.36$ 的单生态风险相较于其他重金属明显偏高. 而其他农田中主要为 Cd 和 Hg 带来的潜在生态风险.

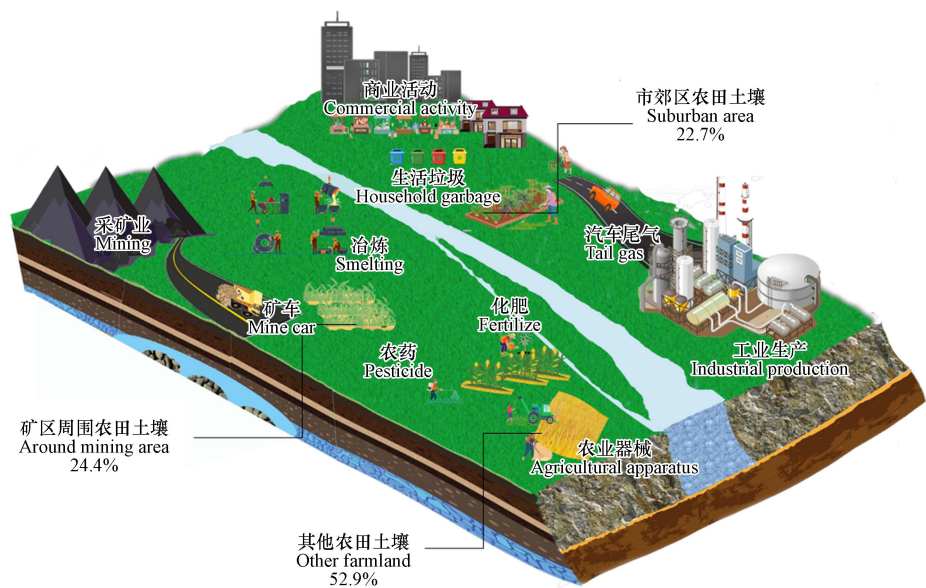


图 3 不同区域样点类型及潜在污染来源

Fig.3 The distribution of styles of sampling points and potential sources of pollution in different regions

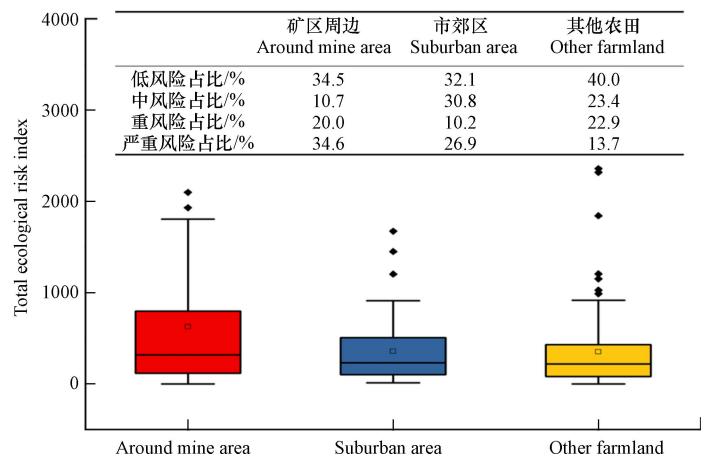


图 4 三类区域样点土壤重金属综合生态风险指数

Fig.4 Integrated ecological risk index for heavy metals in soils from three types of regional sample sites

表 5 三种样点来源单生态风险指数及单生态风险指数/总生态风险指数 (R)

Table 5 Single ecological risk index [PERI(i)] and R[PERI(i)/PERI]

参数 Parameter	区域 Area	重金属 Heavy metals							
		Cr	Ni	Cu	Cd	Pb	Hg	Zn	As
PERI(i)	矿区周边	48.47	3.23	13.73	197.97	14.64	81.23	1.58	9.93
	市郊区	1.65	2.09	6.40	95.10	7.37	155.36	0.92	8.72
	其他农田	1.77	1.92	4.43	155.48	6.82	106.27	1.10	8.91
R/%	矿区周围	14.77	0.95	3.89	50.50	4.10	22.45	0.45	2.89
	市郊区	0.63	0.78	2.49	34.54	2.90	54.87	0.36	3.43
	其他农田	0.64	0.69	1.57	54.00	2.40	39.13	0.39	1.19

对 3 个样区总生态风险指数与 8 种重金属单生态风险指数对总生态风险贡献比率分析,如表 5 所示. 从结果来看, 3 个样区中潜在生态风险均来源于 Hg 和 Cd,但占比略有不同, 其中矿区周边农田土壤样点中主要为 Cd 带来的风险, 而市郊区农田土壤样点中主要为 Hg 带来的风险, 其他农田土壤样点中, Cd 和 Hg 的贡献率相比前二者较为接近.

2.3 CatBoost 分析

分别选取了影响农田土壤重金属浓度的人均 GDP 和人口密度等人为因素和土壤质地、土壤类型和土壤侵蚀度等自然因素进行分析,采用 CatBoost 模型对潜在风险因子进行重要性评估,如图 5 所示,矿区周边采样的农田土壤中重金属的潜在生态风险的影响因素排序为: 人口密度(POP) > 砂土占比(Sand) > 人均 GDP(GDP) > 粉黏土占比(Silt) > 土壤类型(Type) > 土壤侵蚀程度(Erode). 人口密度与人均 GDP 对矿区周边土壤中重金属潜在生态风险影响较大,采矿业和有色金属冶炼作为支柱产业发展,为当地带来财政收入和就业机会的同时,也往往伴随着大量的工业排放和废弃物排放,矿区周围农田土壤重金属的积累受人为选矿的影响,导致除去所需矿种的其他重金属元素暴露于地表,经过渗透或地表水作用进入矿区土壤环境中. 自然因素主要表现在土壤砂土、粉砂土占比以及土壤类型的影响. 土壤中砂土、粉砂土和黏土占比对土壤吸附重金属的能力影响较大,砂土、粉砂土和黏土占比对重金属的物理吸附产生影响,黏土颗粒表面通常带负电,较高的黏土占比有利于吸附带正电的重金属阳离子,但不利于重金属的迁移;较高的砂土占比有利于重金属阳离子的迁移,但重金属的环境容量相对较低^[24-27]. 另外,土壤的 pH 和有机质含量等也会影响重金属在土壤中的迁移和转化,在酸性条件下(pH<7) 重金属多以离子态存在,迁移能力较强,而在碱性条件下(pH>7) 孔隙水中自由状态的重金属离子转化为难溶沉淀,不利于其迁移;土壤中高含量的有机质会与重金属发生螯合/络合等化学反应,限制其迁移转化^[28-29].

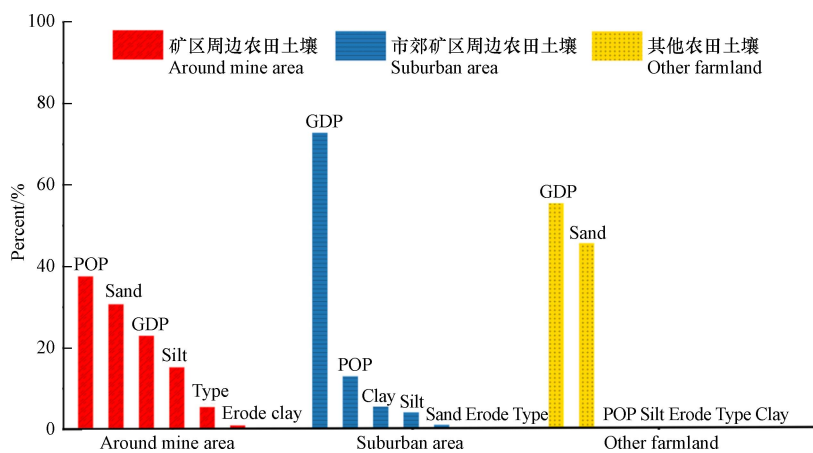


图 5 3 个样区影响因子重要程度分布图

Fig.5 Distribution of the importance of impact factors in the three sample areas

市郊区周边农田土壤潜在生态风险影响因素排序为: 人均 GDP(GDP) > 人口密度(POP) > 黏土占比(Clay) > 粉黏土占比(Silt) > 砂土占比(Sand). 主要受人为因素影响,人口密度和人均 GDP 主要与商业活动、人类生活影响和工业生产以及农业生产相关,市郊区人口密度较高,路网密集,汽车尾气排放颗粒物中含有大量重金属及其化合物^[30]. 汽车轮胎磨损和润滑油燃烧过程导致重金属添加剂的释放,也会加重重金属污染^[31].

其他农田土壤重金属的潜在生态风险的风险因子重要程度排序为: 人均 GDP(GDP) > 砂土占比(Sand). 其他农田土壤样点影响较大的风险因子为人均 GDP 和砂土占比,高人均 GDP 区域,车辆、工农交通活动密集,导致 Cd、Pb 和 Hg 生态风险,As、Zn 一定程度上受人为活动影响,农药化肥的施用也会造成 As 积累^[32].

3 结论 (Conclusion)

本研究针对区域农田土壤重金属污染污染特征及影响因素分析,选择了代表我国经济发展过程中 3 个典型区域作为研究对象,采用统计学、空间分析和机器学习等分析手段,分析了研究区域表层土壤中重金属的分布规律与污染特征:

(1) 3 个研究区域 (东北三省区域、长江经济带和京津冀) 中 Cd 污染问题较为突出,东北三省、京

津冀和长江经济带的超标率分别为 37.5%、34.0% 和 45.8%, 长江经济带也存在 Cu 污染问题, 其超标率为 30.6%。

(2) 相对于全国土壤背景值, 3 个研究区域均有不同程度的污染, 长江经济带上区域样点超过背景值最多的重金属种类为 Hg 和 Pb, 下游为 Pb 和 Cd; 京津冀区域样点超过背景值最多的重金属为 Cd、Hg 和 Ni; 东北三省区域样点中超过背景值最多的重金属为 Cd 和 Hg。

(3) 进一步将样点来源分为矿区周边点位、市郊区点位和其他农田点位, 采用潜在生态风险指数法评价表明, 矿区周边样点中严重和重风险占比(54.8%)显著高于市郊区样点(37.2%)和其他农田样点(36.6%)。

(4) 采用 Catboost 模型进行影响因子识别发现, 矿区周边农田主要受选矿、尾矿暴露等造成的 Cd 和 Hg 污染影响, 而市郊区农田主要受商业、工业和交通运输业等人类活动带来的 Hg 污染影响, 其他农田土壤中主要受农药化肥的施用以及大型器械化带来的 Cd 污染影响。

参考文献 (References)

- [1] 宋伟, 陈百明, 刘琳. 中国耕地土壤重金属污染概况 [J]. 水土保持研究, 2013, 20(2): 293-298.
SONG W, CHEN B M, LIU L. Soil heavy metal pollution of cultivated land in China [J]. Research of Soil and Water Conservation, 2013, 20(2): 293-298(in Chinese).
- [2] 陈文轩, 李茜, 王珍, 等. 中国农田土壤重金属空间分布特征及污染评价 [J]. 环境科学, 2020, 41(6): 2822-2833.
CHEN W X, LI Q, WANG Z, et al. Spatial distribution characteristics and pollution evaluation of heavy metals in arable land soil of China [J]. Environmental Science, 2020, 41(6): 2822-2833 (in Chinese).
- [3] 环境保护部, 国土资源部. 全国土壤污染状况调查公报 [EB]. [2023-04-07].
- [4] 唐翔宇, 朱永官. 土壤中重金属对人体生物有效性的体外试验评估 [J]. 环境与健康杂志, 2004, 21(3): 183-185.
TANG X Y, ZHU Y G. Advances in *in vitro* tests in evaluating bioavailability of heavy metals in contaminated soil via oral intake [J]. Journal of Environment and Health, 2004, 21(3): 183-185 (in Chinese).
- [5] ZUKOWSKA J, BIZIUK M. Methodological evaluation of method for dietary heavy metal intake [J]. Journal of Food Science, 2008, 73(2): 21-29.
- [6] 贾晓琳. 区域土壤重金属污染的源汇空间分析和时空模拟研究 [D]. 杭州: 浙江大学, 2020.
JIA X L. Source-sink spatial analysis and spatio-temporal simulation of soil heavy metal pollution in regional scale [D]. Hangzhou: Zhejiang University, 2020 (in Chinese).
- [7] 黄颖. 不同尺度农田土壤重金属污染源解析研究 [D]. 杭州: 浙江大学, 2018.
HUANG Y. The exploring of heavy metal pollution source apportionment in various scale of agricultural soils [D]. Hangzhou: Zhejiang University, 2018 (in Chinese).
- [8] 金昭, 吕建树. 基于机器学习模型的区域土壤重金属空间预测精度比较研究 [J]. 地理研究, 2022, 41(6): 1731-1747.
JIN Z, LV J S. Comparison of the accuracy of spatial prediction for heavy metals in regional soils based on machine learning models [J]. Geographical Research, 2022, 41(6): 1731-1747 (in Chinese).
- [9] 黄煜韬, 施维林, 纪娟, 等. 基于 BP 神经网络对某电镀厂土壤重金属预测及人体健康风险评价 [J]. 生态毒理学报, 2022, 17(2): 278-289.
HUANG Y T, SHI W L, JI J, et al. Prediction of soil heavy metals based on BP neural network and assessment of human health risk of an electroplating plant [J]. Asian Journal of Ecotoxicology, 2022, 17(2): 278-289 (in Chinese).
- [10] HU Y A, CHENG H F. Application of stochastic models in identification and apportionment of heavy metal pollution sources in the surface soils of a large-scale region [J]. Environmental Science of Technology, 2013, 47(8): 3752-3760.
- [11] 国家统计局. 2008 年统计年鉴 [DB]. [2023-04-07].
- [12] 国家统计局. 2021 年统计年鉴 [DB]. [2023-04-07].
- [13] 中华人民共和国 2021 年国民经济和社会发展统计公报 [EB]. 国家统计局信息公开 (stats.gov.cn). 2021.
Statistical Bulletin of National Economic and Social Development of the People's Republic of China. [EB]. National Bureau of Statistics information disclosure(in Chinese).
- [14] 苗丰顺, 李岩, 高岑, 等. 基于 CatBoost 算法的糖尿病预测方法 [J]. 计算机系统应用, 2019, 28(9): 215-218.
MIAO F S, LI Y, GAO C, et al. Diabetes prediction method based on CatBoost algorithm [J]. Computer Systems & Applications, 2019, 28(9): 215-218 (in Chinese).
- [15] PROKHORENKOVA L, GUSEV G, VOROBIEV, A. et al. CatBoost: Unbiased Boosting with Categorical Features [C]. NIPS'18: Proceedings of the 32nd International Conference on Neural Information Processing Systems, 2018: 6639-6649.

[16] 郑顺安. 我国典型农田土壤中重金属的转化与迁移特征研究 [D]. 杭州: 浙江大学, 2010.
ZHENG S A. Studies on the transformation and transport of heavy metals in typical Chinese agricultural soils[D]. Hangzhou: Zhejiang University, 2010 (in Chinese).

[17] 陈曼丽. 重金属在表层土壤中迁移特性研究 [D]. 郑州: 郑州大学, 2019.
CHEN M L. Study on migration characteristics of heavy metals in surface soil[D]. Zhengzhou: Zhengzhou University, 2019 (in Chinese).

[18] 梁成华, 刘学, 杜立宇, 等. 砷在棕壤中的吸附解吸行为及赋存形态研究 [J]. 河南农业科学, 2009, 38(4): 64-68.
LIANG C H, LIU X, DU L Y, et al. The studies on the adsorptive-desorptive behaviour of arsenic in brown soil and its form distribution[J]. *Journal of Henan Agricultural Sciences*, 2009, 38(4): 64-68 (in Chinese).

[19] 魏复盛, 杨国治, 蒋德珍, 等. 中国土壤元素背景值基本统计量及其特征 [J]. 中国环境监测, 1991, 7(1): 1-6.
WEI F S, YANG G Z, JIANG D Z, et al. Basic statistics and characteristics of background values of soil elements in China[J]. *Environmental Monitoring in China*, 1991, 7(1): 1-6 (in Chinese).

[20] 刘小红, 周东美, 郝秀珍, 等. 九华铜矿重金属环境污染状况研究 [J]. 土壤, 2007, 39(4): 551-555.
LIU X H, ZHOU D M, HAO X Z, et al. Heavy metal pollution of the environment in Jiuhua copper orefield[J]. *Soils*, 2007, 39(4): 551-555 (in Chinese).

[21] 中华人民共和国生态环境部. 铅锌冶炼行业危险废物环境管理指南 [S]. W020211227380770404835. pdf (mee. gov. cn Ministry of Ecology and Environment, PRC, Guide for environmental management of hazardous wastes in lead and zinc smelting industry. [S]. W020211227380770404835. pdf (mee. gov. cn) (in Chinese) .

[22] DONG S W, PAN Y C, GUO H, et al. Identifying influencing factors of agricultural soil heavy metals using a geographical detector: A case study in Shunyi district, China[J]. *Land*, 2021, 10: 1010.

[23] 刘孝严, 樊亚男, 刘鹏, 等. 基于文献计量分析的长江经济带农田土壤重金属污染特征 [J]. 环境科学, 2022, 43(11): 5169-5179.
LIU X Y, FAN Y N, LIU P, et al. Characteristics of heavy metal pollution in farmland soil of the Yangtze River economic belt based on bibliometric analysis[J]. *Environmental Science*, 2022, 43(11): 5169-5179 (in Chinese).

[24] 张丽娜. 山东省基本农田土壤重金属含量分布特征及其环境容量研究 [D]. 济南: 山东师范大学, 2010.
ZHANG L N. The research of the distribution characteristics of heavy metals and the environmental capacity of fundamental farmland in Shandong Province[D]. Jinan: Shandong Normal University, 2010 (in Chinese).

[25] 周璐. 重金属 Pb^{2+} 在砂土中的迁移及吸附特性 [D]. 吉林: 东北电力大学, 2019.
ZHOU L. Migration characteristic of heavy metal Pb^{2+} in sand[D]. Jilin: Northeast Dianli University, 2019 (in Chinese).

[26] WANG H Y, GAO B, WANG S S, et al. Removal of Pb(II), Cu(II), and Cd(II) from aqueous solutions by biochar derived from $KMnO_4$ treated hickory wood[J]. *Bioresource Technology*, 2015, 197: 356-362.

[27] 曾兴. 黏土屏障重金属迁移离心模拟相似性及击穿时间评估方法 [D]. 杭州: 浙江大学, 2015.
ZENG X. Similitude for centrifuge modelling of heavy metal migration in clay barrier and method for evaluating breakthrough time[D]. Hangzhou: Zhejiang University, 2015 (in Chinese).

[28] 刘闯. 重金属 Pb^{2+} 污染下红黏土变形特性及微观结构变化规律 [D]. 桂林: 桂林理工大学, 2022.
LIU C. Microstructure change and deformation characteristics of red clay pollution by heavy metal Pb^{2+} [D]. Guilin: Guilin University of Technology, 2022 (in Chinese).

[29] 李懿. 区域土壤重金属污染风险评价、驱动因子与管控策略研究 [D]. 杭州: 浙江大学, 2022.
LI Y. Risk evaluation, driving factors and management of soil contaminated by heavy metal in regional area[D]. Hangzhou: Zhejiang University, 2022 (in Chinese).

[30] 苏全龙. 城市土壤重金属污染特征及影响因素研究: 以南京市为例 [D]. 南京: 南京大学, 2017.
SU Q L. Study on characteristics and influencing factors of heavy metal pollution in urban soils-a case study of Nanjing city[D]. Nanjing: Nanjing University, 2017 (in Chinese).

[31] 李婧, 周艳文, 陈森, 等. 我国土壤镉污染现状、危害及其治理方法综述 [J]. 安徽农学通报, 2015, 21(24): 104-107.
LI J, ZHOU Y W, CHEN S, et al. Actualities, damage and management of soil cadmium pollution in China[J]. *Anhui Agricultural Science Bulletin*, 2015, 21(24): 104-107 (in Chinese).

[32] 罗耀东. 嘉陵江上游典型县域耕地土壤重金属污染特征分析及风险评价 [D]. 兰州: 西北师范大学, 2021.
LUO Y D. Characteristic analysis and risk assessment of heavy metal pollution in cultivated soil of a typical county in the upstream of the Jialing River[D]. Lanzhou: Northwest Normal University, 2021 (in Chinese).