Science of the Total Environment 954 (2024) 176359

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Risk hotspots and influencing factors identification of heavy metal(loid)s in agricultural soils using spatial bivariate analysis and random forest

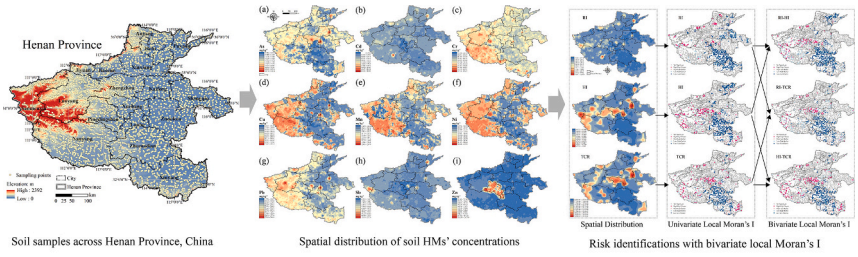
Xiaohang Xu [a,b], Zhidong Xu [b], Longchao Liang [c], Jialiang Han [b], Gaoen Wu [b], Qinhui Lu [d], Lin Liu [b], Pan Li [b], Qiao Han [b], Le Wang [a], Sensen Zhang [e,*], Yanhai Hu [f], Yuping Jiang [f], Jialin Yang [f], Guangle Qiu [b,*], Pan Wu [a]

[a] Key Laboratory of Karst Georesources and Environment, Ministry of Education, College of Resources and Environmental Engineering, Guizhou University, Guiyang 550025, China
[b] State Key Laboratory of Environmental Geochemistry, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang 550081, China
[c] School of Chemistry and Materials Science, Guizhou Normal University, Guiyang 550001, China
[d] The Key Laboratory of Environmental Pollution Monitoring and Disease Control, Ministry of Education, Guizhou Provincial Engineering Research Center of Ecological Food Innovation, School of Public Health, Guizhou Medical University, Guiyang 550025, China
[e] Henan Academy of Geology, Zhengzhou 450016, China
[f] No.6 Geological Unit Team, Henan Provincial Non-ferrous Metals Geological and Mineral Resources Bureau, Luoyang 471002, China

HIGHLIGHTS

- An integrated investigation on HMs in agricultural soils of Henan province was reported firstly.
- Multivariate statistics, geo-statistical methods, receptor model, and random forest were employed.
- Cd was identified as the priority pollutant in agricultural soils of Henan province.
- Sensitivity analyses of exposure factors to TCR and NCR were performed.
- Traffic source is the dominant source of HMs in agricultural soils in Henan province.

GRAPHICAL ABSTRACT



Heavy metal(loids) in agricultural soils from Henan Province

ARTICLE INFO

Editor: Paola Verlicchi

Keywords:
Heavy metal(loid)s
Soil pollution
Machine learning
Risk identifications
Spatial autocorrelation
Environmental factors

ABSTRACT

Heavy metal(loid)s (HMs) in agricultural soils not only affect soil function and crop security, but also pose health risks to residents. However, previous concerns have typically focused on only one aspect, neglecting the other. This lack of a comprehensive approach challenges the identification of hotspots and the prioritization of factors for effective management. To address this gap, a novel method incorporating spatial bivariate analysis with random forest was proposed to identify high-risk hotspots and the key influencing factors. A large-scale dataset containing 2995 soil samples and soil HMs (As, Cd, Cr, Cu, Mn, Ni, Pb, Sb, and Zn) was obtained from across Henan province, central China. Spatial bivariate analysis of both health risk and ecological risks revealed risk hotspots. Positive matrix factorization model was initially used to investigate potential sources. Twenty-two environmental variables were selected and input into random forest to further identify the key influencing factors impacting soil accumulation. Results of local Moran's I index indicated high-high HM clusters at the

* Corresponding authors.
E-mail addresses: xuxh@gzu.edu.cn (X. Xu), xuzhidong@vip.skleg.cn (Z. Xu), lianglc139@gznu.edu.cn (L. Liang), liulin@xju.edu.cn (L. Liu), lipan@gig.ac.cn (P. Li), hanqiao@gyig.ac.cn (Q. Han), zsgucas@sina.com (S. Zhang), qiuguangle@vip.skleg.cn (G. Qiu), pwu@gzu.edu.cn (P. Wu).

https://doi.org/10.1016/j.scitotenv.2024.176359
Received 29 June 2024; Received in revised form 12 September 2024; Accepted 16 September 2024
Available online 19 September 2024

western and northern margins of the province. Hotspots of high ecological and health risk were primarily observed in Xuchang and Nanyang due to the widespread township enterprises with outdated pollution control measures. As concentration and exposure frequency dominated the non-carcinogenic and carcinogenic risks. Anthropogenic activities, particularly vehicular traffic (contributing ~37.8 % of the total heavy metals accumulation), were the dominant sources of HMs in agricultural soils. Random forest modeling indicated that soil type and PM2.5 concentrations were the most influencing natural and anthropogenic variables, respectively. Based on the above findings, control measures on traffic source should be formulated and implemented provincially; in Xuchang and Nanyang, scattered township enterprises with outdated pollution control measures should be integrated and upgraded to avoid further pollution from these sources.

1. Introduction

Soil acts as a key sink of pollutants from many sources and may be adversely affected by them, particularly heavy metal(loid)s (HMs). Soil HM contamination may directly affect soil function, with potential threats to human health arising through direct exposure and food consumption (Rinklebe et al., 2020; Tokatlı et al., 2023). Heavy metal contamination in soil can deplete nutrients and degrade soil structure and functionality (Jiang et al., 2020a). Long-term exposure to low concentrations of heavy metals poses significant health risks, potentially causing cancer, reproductive disorders, skin lesions, lung diseases, and gastrointestinal issues (Wang et al., 2022), and in severe cases, fatalities through skin contact, inhalation, or ingestion (Huang et al., 2022; Liu et al., 2021a). Generally, HMs in soils primarily originate from weathering of parent materials but can also accumulate from industrial and agricultural activities (Hu and Cheng, 2013). Soil HM contamination levels have increased globally due to accelerated industrialization and urbanization in recent decades (Akopyan et al., 2018; Kulikova et al., 2019; Liu et al., 2021b).

The management of soil HM contamination requires a thorough understanding of the spatial characteristics, sources, and influencing factors. Multivariate statistical analysis (Zeng et al., 2024b), geostatistical analysis (Baltas et al., 2020), quantitative receptor models (Hu et al., 2020), and stable-isotope fingerprints are currently applied in identifying HM sources (Wang et al., 2021), while multivariate statistical methods, such as correlation analysis (CA) and principal component analysis can qualitatively identify sources based on dimensionality reduction (Han et al., 2022; Liang et al., 2023). Receptor models such as positive matrix factor (PMF) analysis, absolute principal component score-multiple linear regression (APCS–MLR), and UNMIX could quantitatively apportion the sources of HM in soil (Guan et al., 2019; Han et al., 2022). Stable-isotope techniques may accurately identify sources on the basis of isotopic ratio variations among different emission sources (Liu et al., 2022; Wang et al., 2021), but are not suitable for large scale application due to the complexity and expense of analyses. Compared with other receptor models, the PMF model assumes that the observed data are a linear combination of contributions from multiple sources, with both source contributions and source profiles being non-negative. PMF analysis provides an "uncertainty weighting that is useful in" assessing the final results (Sun et al., 2020), with factor loadings being statistically scaled in terms of their relative contributions, thereby improving the reliability of the analysis. Though the primary potential types of contamination sources can be distinguished by receptor models, discerning the direct and indirect factors governing HM accumulation in soils is essential for crafting targeted control measures. Increasingly, previous researches have emphasized the synergistic influences of soil properties and climate on soil HM contamination (Wu et al., 2021; Zhong et al., 2020). The involvement of various contamination sources, along with numerous environmental variables, complicates the source-sink relationships (Yang et al., 2021). Most receptor models identify sources through linear relationships, however, the source–sink relationship of HMs in soil not always linear, which may bias the results. The random forest (RF) method is adept at handling the non-linear relationships between driving factors behind environmental pollutants. It utilizes multiple independent decision trees, full–size trees without pruning, bootstrap sampling, random split selection, an unbiased measure of error rate, and covariate importance analysis (Zhao et al., 2023). Thus, combining the RF and PMF models can provide a comprehensive approach to identifying potential driving sources of pollutants.

Accurate assessment of ecological and human health risks posed by soil HMs is crucial for developing refined control measures and reducing costs in soil HM management. While previous studies have commonly utilized global Moran's I and local Moran's I (LMI) to identify soil hotspots, these methods primarily focus on single-variable spatial relationships. In contrast, bivariate LMI spatial association analysis offers a novel approach by coupling HM risks based on spatial interactions between different risks. This innovative method allows for a more comprehensive understanding of the interplay between multiple contaminants, thereby providing deeper insights into the complex spatial patterns of soil contamination (Abokifa et al., 2020), describing the spatial clustering characteristics of 'hotspots', and aiding in understanding of combined large-scale contamination risks.

Henan Province in central China is an important grain-producing areas with an annual grain yield of ~6.8 million tonnes, accounting for ~10 % of national production (HSB, 2023). The province is also rich in mineral resources such as coal and polymetallic minerals (HSB, 2023). Soil HMs contamination events have occurred frequently in recent years (Zhang et al., 2015), challenging food security (Xing et al., 2016; Xu et al., 2022b; Yang et al., 2022). Concern has focused on HM contamination around typical sources such as battery factory in Xinxiang (Jiang et al., 2020c), molybdenum–polymetallic and tungsten mining in Luoyang (Chen et al., 2023; Hui et al., 2021), coal mining areas (Li et al., 2018), a lead smelting plant in Jiyuan (Wu et al., 2020; Xing et al., 2019b), and soil HM levels at the county (Tanghe and Yongcheng county) and city (Zhengzhou) scales (Liu et al., 2022c; Meng et al., 2021b; Zhang and Zhang, 2021; Zhang and Li, 2021). In studies conducted in Henan, Cd, Pb, Cu, Zn, Cr, and As were highlighted as the primary HM exceeding the limits set by Chinese Government. However, no study have examined soil HM levels on a provincial scale, limiting the understanding of pollution patterns and hindering policy-making for contamination prevention and control in Henan Province. Ensuring soil quality in Henan Province is important to food security in China.

This study aimed to fill these knowledge gaps by collecting agricultural soil samples from across the province in 2018, with the objectives: (1) to assess the concentration and distribution of HMs in agricultural soils; (2) to investigate the spatial variability of soil HMs; (3) to describe the spatial distribution of associated ecological and human-health risks; and (4) to identify the main contributing sources and clarify the primary factors influencing soil HM levels and their corresponding risks. The results of this study should enhance the understanding of HM contamination of agricultural soils across the province, thereby guiding soil contamination management.

2. Methods

2.1. Study area and soil sampling

Henan Province is located in central China (31°23′–36°22′N,

110°21′–116°39′E) and has an area of 165,700 km^2. Henan province has a transitional climate from subtropical to warm temperate zones, with an annual temperature of 15.8 °C and annual precipitation of 513–1129 mm. Landforms include a plain basin, mountain, hill that respectively account for 55.7 %, 26.6 %, and 17.7 % of the total area. Some 144 mineral types have been found in the province and 93 mined, predominantly for molybdenum, gold, aluminum, silver, trona, salt, refractory clay, fluorite, perlite, cement limestone, and graphite. The population in Henan Province is 98.72 million, and the gross domestic product is RMB 6135 billion (USD 912 billion) (HSB, 2023).

Surface soil samples (2995 samples of 0–20 cm depth) were collected from agricultural land across the province (Fig. 1). The soil samplings, including sites selection and sampling procedures, were conducted according to the Technical Rules for Monitoring Environmental Quality of Farmland Soil (MARAPRC, 2012). Samples were collected from areas distant from obvious sources of pollution. Five subsamples were collected within a 50 m radius at each site, mixed, and stored as a single sample in a zip-lock polyethylene bag. During sampling, a plastic shovel was used to remove gravel, grass roots and other decries. A handheld GPS device was used to capture the latitude and longitude coordinates for each sample. After air-drying at room temperature, stones, roots, and litter were removed and samples ground in an agate mortar to 200 mesh (Xu et al., 2017).

2.2. Chemical analysis

For HM analysis, 0.05 g of soil was weighed into a Teflon tube and digested with HNO_3 + HF at 185 °C for 48 h. After cooling to room temperature, H_2O_2 was added and the mixture evaporated to near dryness at 110 °C. Ultrapure water and HNO_3 were added and kept at 185 °C for 16 h. HM concentrations in the residual solution were determined by inductively coupled plasma–mass spectrometry (ICP–MS, NexIONTM 300×, Perkin Elmer, USA) (Han et al., 2022).

An accuracy was based on analysis of internal standards, a National Soil Reference Material (GBW07405), duplicates, and reagent blanks. Recoveries of reference material was between 94 %–109 %.

2.3. Geo-accumulation index

The geo-accumulation index (I_{geo}) indicates the variations in the natural distribution of HMs in soil, reflecting the historical accumulation of contaminants, calculated as follows:

$$I_{geo} = log_2 \frac{C_i}{KB_i} \quad (1)$$

where C_i (mg kg^{-1}) represents the measured concentration of element i; B_i (mg kg^{-1}) is the background concentration of element I in Henan Province, and K is the correction coefficient of B_i, which is a constant value of 1.5.

2.4. Spatial autocorrelation

Spatial autocorrelation is commonly denoted by Moran's I, including global and local Moran's indices (Zhang et al., 2019). The global Moran's I may cover the whole range of Moran's I from −1 to +1, denoting a perfect negative or perfect positive spatial autocorrelation, respectively (Eq. (2)). Values of Moran's I near 0 indicate no spatial autocorrelation near the mean. The LMI indicates the degree of spatial clustering of regional units, identifies hotspots (Eq. (3)) (Yuan et al., 2018).

$$\text{Global Moran's I} = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n}(x_i-\overline{x})(x_j-\overline{x})}{S^2\sum_{i=1}^{n}\sum_{j=1}^{n}w_{ij}} \quad (2)$$

$$\text{Local Moran's I} = \frac{(x_i-\overline{x})}{S^2}\sum_{j=1}^{n}w_{ij}(x_j-\overline{x}) \quad (3)$$

where n is the sample size; x_i and x_j are observed concentration (mg kg^{-1}) at sampling sites i and j, respectively; $\overline{x}$ is the mean value of x; S^2 is the variance of samples; and w_{ij} indicates the distances between sites i and j.

Based on the Moran's I index, the bivariate LMI is a method of characterizing spatial correlation between different variables (Wu et al.,

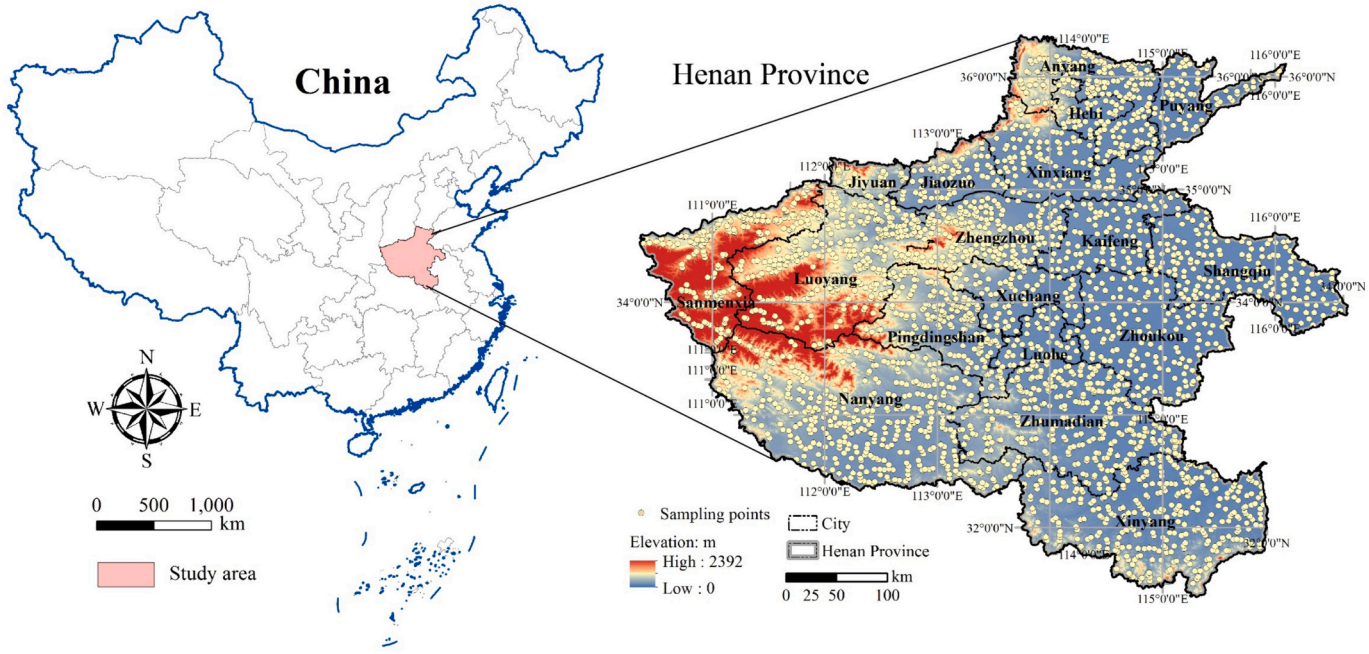


Fig. 1. Locations of sampling sites in agricultural lands across Henan Province.

2019). Bivariate LMI was applied in characterizing the spatial association of ecological, non-carcinogenic, and carcinogenic risks to probe the points with high risks:

$$I_i^{ab} = X_i^a \sum_{j=1, j\neq i}^{n} w_{ij} X_j^b \quad (4)$$

where X_i^a and X_j^b denote values variables a and b on the grid i and j, respectively; I_i^{ab} represents the bivariate LMI at location i, and w_{ij} is the spatial weight matrix based on the distance weighting between locations i and j. If I_i^{ab} is significantly positive or negative, variable a at grid point I is strongly correlated with variable b in the adjacent area (otherwise there is no evident correlation) (Shi et al., 2023).

2.5. Human health risk

Human exposure to soil HMs may occur through three main pathways: ingestion, inhalation, and dermal contact. The average daily exposure dose (ADD) of each pathway and its corresponding health effects were evaluated as follows, with both non-carcinogenic and carcinogenic risks considered.

$$ADD_{ing}^i = \frac{C_i \times IR_S \times EF \times ED}{BW \times AT} \times 10^{-6} \quad (5)$$

$$ADD_{derm}^i = \frac{C_i \times SA \times SL \times ABF \times EF \times ED}{BW \times AT} \times 10^{-6} \quad (6)$$

$$ADD_{inh}^i = \frac{C_i \times IR_a \times EF \times ED}{BW \times AT \times PEF} \times 10^{-6} \quad (7)$$

$$HI = \sum HQ_j^i = \sum \frac{ADD_{ij}}{RfD_{ij}} \quad (8)$$

$$TCR = \sum CR = \sum ADD_{ij} \times SF_{ij} \quad (9)$$

where HQ_j^i is the average hazard quotient (HQ) of exposure pathway j for element i, and C_i is the soil concentration (mg kg^{-1}) of element i; other parameters are described in Table S1–2. The hazard index (HI) is the sum of all HQ. HI and HQ values of >1 indicate a potential adverse health risk (MohseniBandpi et al., 2018). CR represents carcinogenic risk, and TCR is the total carcinogenic risk. Values of CR or TCR of >1E–4 indicate a potential carcinogenic risk; values of <1E-4 indicate no significant carcinogenic risk; and values of <1E-6, indicate a negligible risk (USEPA, 2009).

2.6. Ecological risk

The potential ecological risk of HMs to ecosystems was evaluated as follows (Arfaeinia et al., 2019).

$$E_r^i = T_r^i C_i / C_b^i \quad (10)$$

where E_r^i indicates the monomial ecological risk index of an element, and T_r^i the toxic response coefficient of an HM. Values of T_r^i for As, Cd, Cr, Cu, Mn, Ni, Pb, Sb, and Zn were 10, 30, 2, 5, 1, 5, 5, 40, and 1, respectively (Hakanson, 1980; Jiang et al., 2020b). The sum of E_r^i values of all studied HMs represent the ecological risk index (RI).

2.7. PMF model

The PMF model provides a mathematical approach to quantifying the contribution of sources based on the composition ('fingerprint') of sources. PMF 5.0 could resolve the concentration matrix into two matrices: factor contributions and factor profiles. Each factor was then analyzed further (Guan et al., 2019). The specific number of factors is not usually given directly, and much debugging may be necessary.

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + e_{ij} \quad (11)$$

where x_{ij} denotes the HM concentration; i represents the sample number; j is the chemical species; p is the source number; e is the error of each sample; and u is the uncertainty calculated as follows:

$$\text{For } x_{ij} \leq MDL,\ u_{ij} = \frac{5}{6} \times MDL \quad (12)$$

$$\text{For } x_{ij} > MDL,\ u_{ij} = \sqrt{(\sigma \times x_{ij})^2 + (0.5 \times MDL)^2} \quad (13)$$

where MDL is the method detection limit of each element, and σ is relative standard deviation. The contribution and profile of each sources were obtained by minimizing the sum of an 'object function' Q as follows (Zhang et al., 2020).

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{e_{ij}}{u_{ij}}\right)^2 \quad (14)$$

2.8. Random Forest method

The RF model is an integration algorithm or decision tree classifier (Bremian, 2001), which first employs self-help sampling for training data to form samples before, a regression tree model for each sample is developed. The predictive value of the dependent variable is determined by the average of results of n regression tree models. A random data split technique was employed to partition the dataset into a training dataset (comprising 75 % for calibration) and a test dataset (with 25 % allocated for validation). The training dataset was utilized to calibrate the regression models, whereas the test dataset was employed to evaluate their final outcomes (Agyeman et al., 2022).

In 2019, 22 selected variables were used as contributing factors, selected for quantitative contribution evaluation. The factors were classified into two main categories: 'natural' factors including temperature, elevation, fine sand, soil type, precipitation, clay particles, total nitrogen, silty sand, organic matter, coarse sand, pH, total phosphorus, and total potassium; and anthropogenic factors including population, railway, chemical fertilizers, $PM_{2.5}$, residential zones, factories and mines, pesticides, river related to anthropogenic activities. Relevant data were all downloaded from the Chinese Resource and Environmental Science and Data Center (https: www.resdc.cn).

String environmental variables were digitized into numeric variables based on sampling sites. The importance score of each variable was determined using RF model (Huang et al., 2022).

2.9. Statistical analysis

Raw data was processed using Microsoft Excel 2016, and all the figures produced with Origin 2023 (©OriginLab Corporation). Outliers of raw data were screened using a threshold of more or less than three times the standard deviation away from the mean, and the outliers were replaced with the left highest or lowest values for further analysis (Yang et al., 2019). Assessing spatial autocorrelation for each variable is essential to determine the optimal interpolation algorithm for map generation. HMs exhibiting spatial autocorrelation and conforming to a normal distribution are suitable for interpolation using kriging interpolation method in ArcGIS 10.8 (ESRI, Redlands, USA). This involves computing the corresponding semivariogram, leading to maps of the variables accompanied by information regarding estimated value uncertainty. Conversely, HMs lacking spatial autocorrelation are interpolated with inverse distance weight (IDW) (Swidwa-Urbanska and Batlle-Sales, 2021). Spatial autocorrelations were obtained as global Moran's I, LMI, and bivariate LMI with GeoDa. To assess the impact of data

transformation on the results of spatial cluster and outlier identification, while also accounting for the compositional nature of environmental data, three data transformation methods were chosen: log (10) transformation, centered log-ratio transformation, and normal score transformation. The Shapiro-Wilk test was used to examine the normality of both raw and transformed datasets. RF modeling were conducted in Rstudio (https://www.rstudio.com). Exposure health risks were estimated by Monte Carlo simulation (MCS) with Crystall Ball softwere (©Oracle, Redwood City, CA, USA). All the raster-based variables were available at 1 km resolution, and these raster data were resampled to all the individual soil sampling points in ArcGIS 10.8. The shortest distances of soil sampling points to industrial enterprises, river, and railway were obtained with 'Near' in analysis tools in ArcGIS software.

3. Results and discussion

3.1. HM characteristics

Concentrations of Cd, Ni, Sb, and Zn in soils generally followed a lognormal distribution, and other HMs a skewed distribution (Table 1). Mean concentrations of Cd, Ni, Sb, and Zn are therefore expressed as geometric mean values, and those of all other HMs as median values. Geometric mean values of Cd, Ni, Sb, and Zn were 0.26, 27.6, 1.40, and 70.0 mg kg^{-1}, respectively, and the median concentrations of As, Cr, Cu, Mn, and Pb were 13.6, 66.0, 21.6, 537, and 24.3 mg kg^{-1}, respectively.

The highest mean concentrations of As were observed in Xuchang (Fig. 2); those of Cd, Pb, and Sb were observed in Jiyuan; and those of Cr, Cu, Mn, and Ni in Nanyang. For Zn, both the arithmetic mean (5040 mg kg^{-1}) and geometric mean (525 mg kg^{-1}) concentrations in Pingdingshan were much higher than those in other cities, possibly due to the presence of coal, non-ferrous metal mining, and related processing industries (Jiang et al., 2019). Relative to previous studies in Henan Province, concentrations of As, Cd, Cr, Cu, Ni, Pb, and Zn were consistent or slightly higher (Chen et al., 2015).

While geometric mean Cd concentrations were higher than those in Zhejiang (0.23 mg kg^{-1}) and Jiangsu Province (0.18 mg kg^{-1}) areas (Yang et al., 2020a; Yang et al., 2020b), they were lower than those reported in Guizhou Province (0.577 mg kg^{-1}) (Zhang et al., 2022). Concentrations of Cd, Cu, Mn, Ni, Pb, Sb, and Zn were lower than those adopted in recently developed soil-quality criteria, but slightly higher than recommended safe values (Zhang et al., 2023).

Coefficients of variation (C.V.) followed the order of Cd (68.2 %) > Sb (53.1 %) > Zn (52.5 %) > As (38.3 %) > Pb (36.2 %) > Mn (34.7 %) > Cu (29.2 %) > Ni (24.7 %) > Cr (16.9 %). Hence, the HM concentrations in agricultural soils exhibit significant degrees of variation and extensive ranges. Moreover, the HM concentrations of the agricultural soil is determined by not only parental materials but also external factors. The higher C.V. values indicated the higher impact of anthropogenic activities, while the lower C.V. values demonstrated the higher impact of natural factors (Baltas et al., 2020). Generally, the C.V. values of Cd, Sb, and Zn were > 50 %, which was more likely to be affected by human activities. The application of chemical fertilizers, pesticides, polluted irrigation and dry deposition from traffic sources and industrial factories (Han et al., 2022).

All HM geometric mean and median soil concentrations, except for Mn, were higher than their corresponding background values, but lower than risk screening values for contamination of agricultural land (GB15618–2018; Table 1). However, several points had concentrations above the risk intervention values due to the impacts of point source (Zhang et al., 2022).

Single-pollutant indices indicated that most of the HMs could be categorized into uncontaminated or low-contamination zones, with values of <2. However, high proportions of the Cd (67.3 %), Pb (33.2 %), Sb (27.4 %), and Zn (18.1 %) data points were classed as being at heavily contaminated levels (Fig. S1; Table S3). Geo-accumulation indices indicated that >67 % of the Cd data points could be classed as moderately–heavily to extremely polluted levels (Table S4). Several Pb, Sb, and Zn samples had concentrations in the heavy to extreme polluted levels, requiring further attention.

Geo-accumulation indices also indicated that most HM concentrations correspond to unpolluted levels, although many samples had Cd concentrations in the unpolluted–moderately polluted or were moderately polluted (Fig. 3). Therefore, Cd should be regarded as a priority-control HM in Henan Province. The fact that 39.5 % of the samples from Pingdingshan city have Zn concentrations at extremely polluted levels indicates that extreme emission sources should not be ignored.

3.2. Spatial distribution of HM

The spatial distribution of soil HMs is illustrated in Fig. 4. Barring Cd, Sb and Zn, the other studied HMs have variable spatial distributions, with most of the higher concentrations being observed in the western, southern, and northern margins of Henan Province. HM hotspots are mainly in mountainous areas, including Taihang Mountain (northern margin of the province), Funiu Mountain (western margin), and Dabie Mountain (southern margin). Specifically, these hotspots were

Table 1
Concentrations of soil HMs' in agricultural lands in Henan Province.

Categories		Concentrations (mg kg^{-1})								
		As	Cd	Cr	Cu	Mn	Ni	Pb	Sb	Zn
Arithmetic mean		14.2	0.292	67.1	22.7	551	28.5	26.5	1.49	74.6
Geometric mean		13.1	0.26	66.2	21.8	460	27.6	25.2	1.40	70.0
Median		13.6	0.257	66.0	21.6	537	27.6	24.3	1.42	67.6
Standard deviation		5.43	0.199	11.4	6.64	191	7.02	9.59	0.793	39.2
Variation coefficient (%)		38.3 %	68.2 %	16.9 %	29.2 %	34.7 %	24.7 %	36.2 %	53.1 %	52.5 %
Skewness		1.26	6.28	0.707	1.156	−0.069	0.749	3.58	10.5	8.11
Kurtosis		5.87	59.3	2.06	2.02	2.03	1.18	20.4	160	104
Range		61.4	3.58	87.9	51.9	1219	52.0	101	18.5	814
Minimum		0.624	0.001	2.44	0.898	2.82	1.04	0.883	0.083	2.48
Maximum		62.1	3.6	112	52.8	1222	56.2	102	18.6	814
Henan geochemical baselines		11.4	0.074	63.8	19.7	579	26.7	19.6	1.37	60.1
The mean China geochemical baselines		9	0.137	53	20	569	24	22	0.73	66
Risk screening values for soil contaminated of agricultural land (GB15618–2018)	pH ≤5.5	40	0.3	150	50	/	60	70	/	200
	5.5<pH ≤6.5	40	0.3	150	50	/	70	90	/	200
	6.5<pH ≤7.5	30	0.3	200	100	/	100	120	/	250
	pH > 7.5	25	0.6	200	100	/	100	170	/	300
Distribution		Skew	Lognormal	Skew	Skew	Skew	Lognormal	Skew	Lognormal	Lognormal

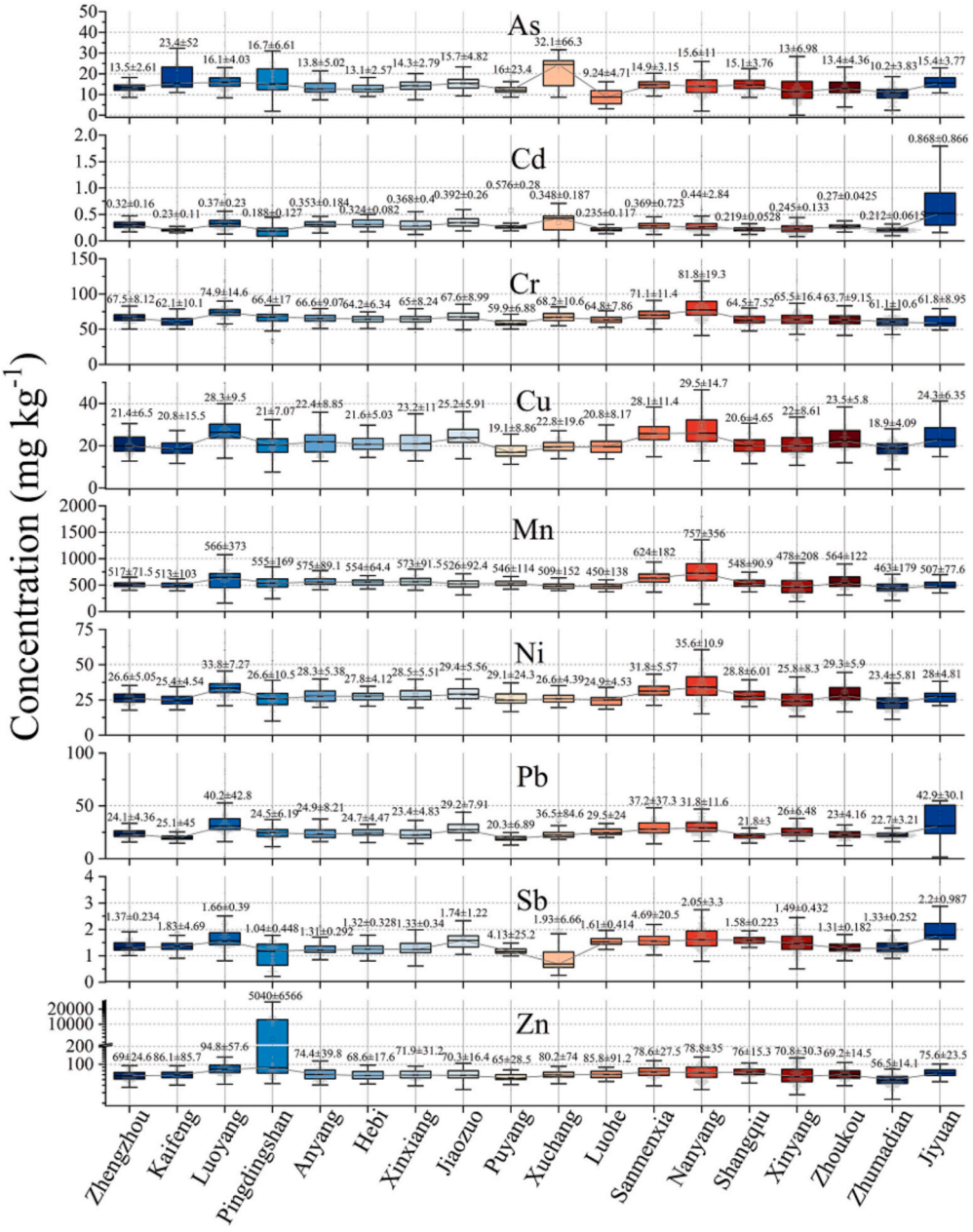


Fig. 2. Box plots for the HMs' concentrations in different cities in the agricultural soils of Henan Province.

distributed in the Luoyang, Sanmenxia, Jiyuan, Jiaozuo, Xinxiang, Anyang, and Nanyang cities.

The HMs As, Cr, Cu, Mn, Ni, Pb, and Sb were primarily distributed in mountain areas, which may be due to the high background values in mountainous areas (Zhu et al., 2023). Moreover, the corresponding non-ferrous-metal mining (Cheng et al., 2023) and processing industries are located in western Henan (Lai et al., 2021; Xing et al., 2018). In spite of the implementation of strict environmental management policies, long term mining and processing produced large amounts of HM-bearing waste and slag dumps, releasing HMs to surrounding soil continuously (Xu et al., 2019). The high soil concentrations of Zn in Pingdingshan mean that hotspots are likely to occur in this area, which is known for its coal mining and Pb—Zn smelting industries (Zhu et al., 2023). Nanyang basin, in southwest Henan Province, also has elevated soil HM levels, likely due to the accumulation of air pollution in the topographically low area, also to runoff from surrounding higher-elevation areas (Li et al., 2022).

As, Cd, Pb, and Sb concentrations are variable over the eastern plain of Henan Province especially in the Xuchang area, possibly due to local industries that produce rubber and tires, fibers and textiles, and building materials with associated solid wastes, dust deposition, and drainage from nearby higher-elevation agricultural land (Liu et al., 2023; Wang, 2014; XCSY, 2017). Peak concentrations of As, Cd, Sb and Zn occur in eastern cities, possibly due to local point sources (Xu et al., 2017). Thus, point sources in plain areas should be considered, although the plain areas where most crops are grown have lower HM concentrations in soil (Zhou et al., 2014).

Global Moran's I analysis indicated that the studied HMs are categorized as random zones (Fig. S2), reflecting the diverse industrial characteristics of the different cities. The LMI provide a clearer representation of spatial patterns (Fig. 5) that is similar to those of soil HMs (Fig. 4). It I important to acknowledge that LMI is sensitive to the outliers (Yuan et al., 2018). LMI results are classed into five categories: high-high (HH), low-low (LL), high-low (HL), low-high (LH), and 'not

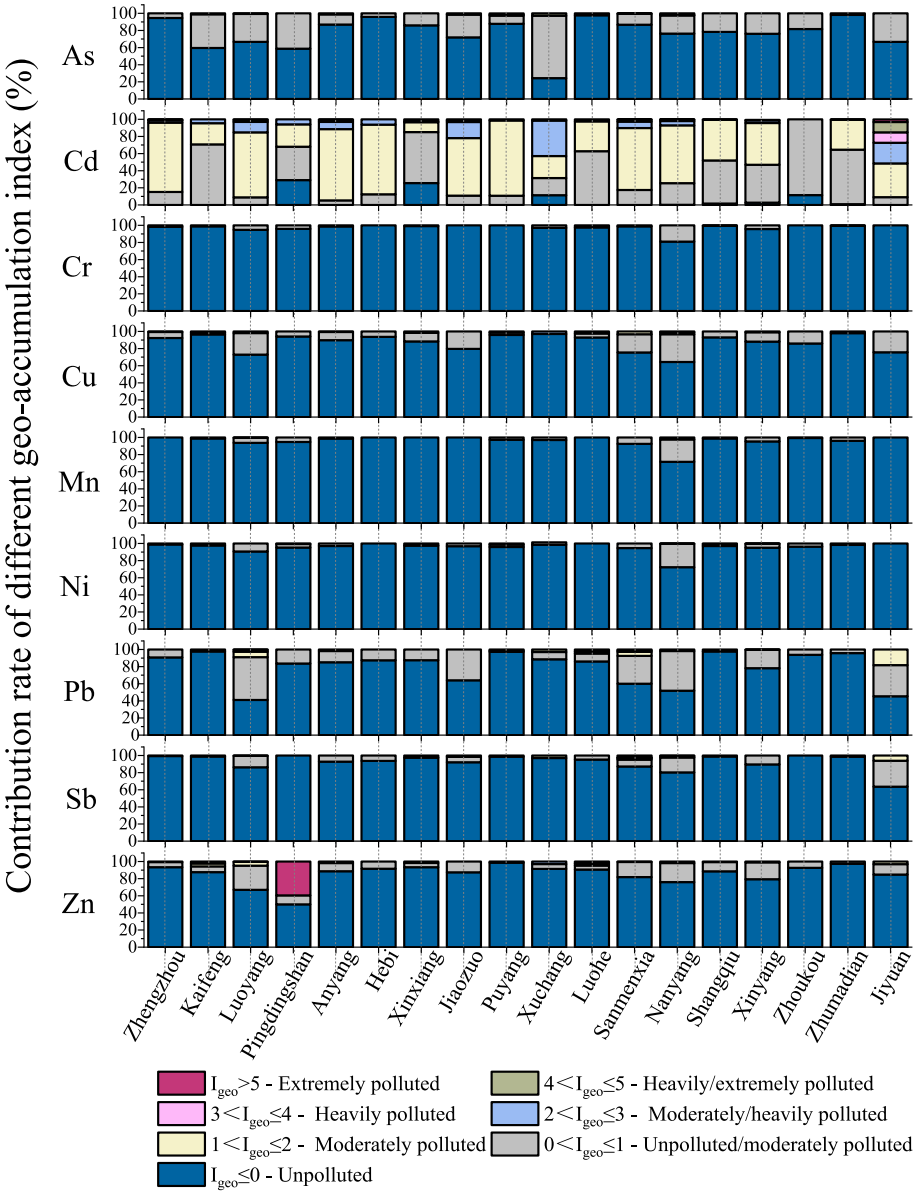


Fig. 3. Distribution of HMs' geo-accumulation indexes in different cities in Henan Province.

significant' (Han et al., 2022). HH clusters are regarded as contamination hotspots, and HL outliers individual hotspots. Similar spatial patterns were observed for As and Cd, with HH clusters mainly in northern and western Henan Province, including Sanmenxia, Luoyang, Jiaozuo, Jiyuan, and Anyang cities. For Cr, Cu, Mn, Ni, Pb, and Sb, HH clusters occur mainly in western Henan Province, including Samenxia, Luoyang, and Nanyang cities. For Zn, HH clusters occur mainly in Pingdingshan city. Similar HH distribution patterns of Cr, Cu, Mn, and Ni, together with the comparable or slightly higher levels than Henan geochemical baselines, it can be inferred that these HH clusters were governed by natural sources (Zhou et al., 2014). HL outliers occur near or in LL clusters, while LH outliers occur mainly near HH clusters, indicating the impact from anthropogenic activities (Jin et al., 2021). LL clusters of most HMs, particularly Cd, Pb, and Zn, occur mainly in the eastern plains, which is the main cropping area and is remote from intensive industrial activities.

3.3. Ecological and human health risks

3.3.1. Ecological risk

The potential ecological risk of single elements (E_i) followed the order Cd (131) > Sb (53.2) > As (13.1) > Pb (7.17) > Zn (6.33) > Cu (6.00) > Ni (5.43) > Cr (2.14) > Mn (0.975) (Table S5). Overall, the ecological risk values of individual elements were categorized into low risk (As, Pb, Zn, Cu, Ni, Cr, and Mn) to moderate risk level (Sb), based on the classification criteria of pollution assessment methods. However, the mean potential ecological risks of Cd and Sb were notably higher than those of other HMs, due to the high percentage of Cd E_i in considerable risks (67.2 %) and the notable Sb E_i proportion within the moderate risk level (53.6 %). This may be due to their high toxicity response factors of 30 and 50, respectively. Some values were even >320 for As, Cd, Sb, and Zn, indicating extremely high risk. Overall, the RI values ranged from 6.95 to 23,324, with a mean of 226 (Table S5). And 74.5 % and 5.58 % of RI values categorized as moderate and considerable risk.

The IDW interpolation method was applied in characterizing the spatial distribution of ecological risk (Fig. 6). From provincial scale, most of peak RI values were observed in the western Henan and Taihang

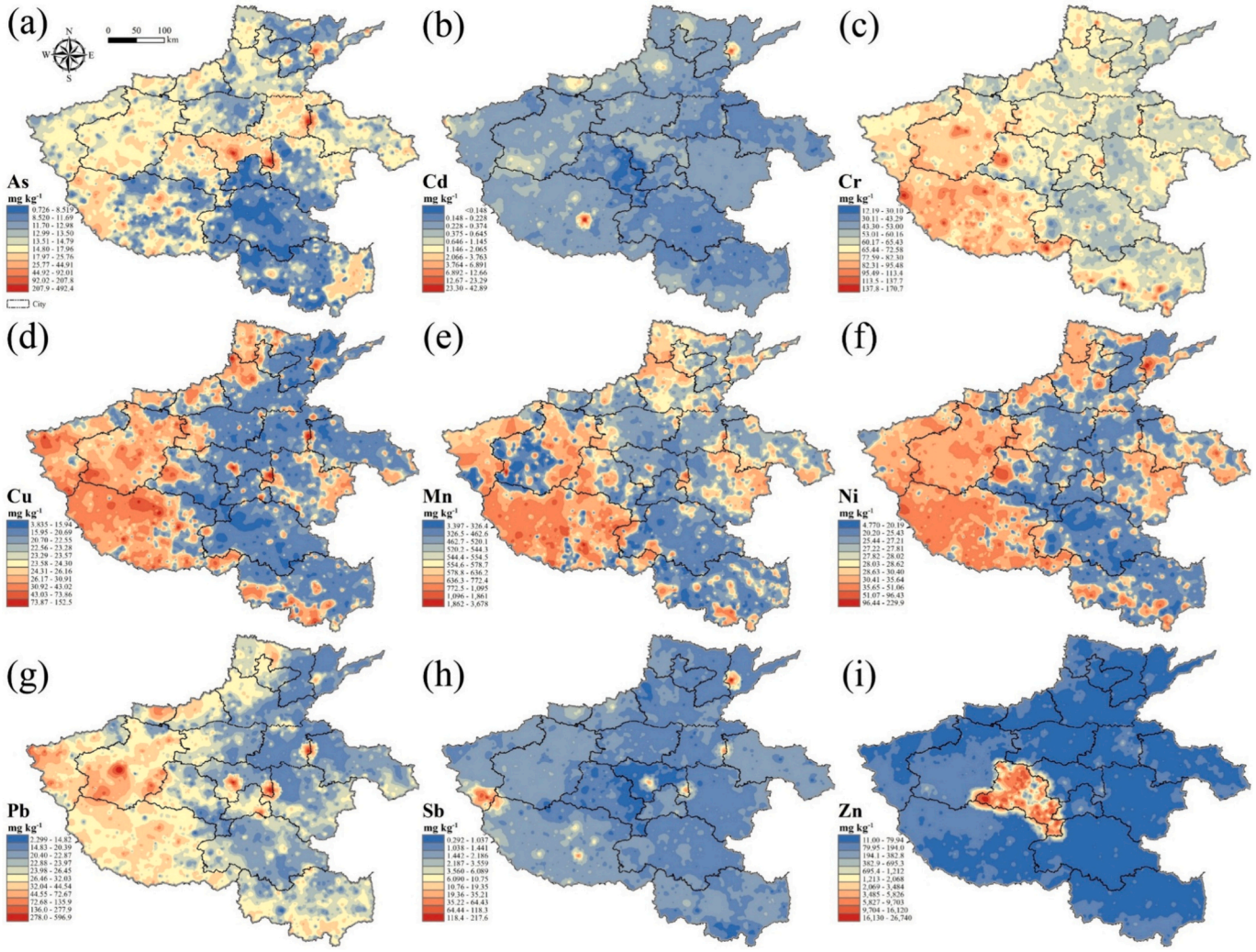


Fig. 4. Spatial distribution of HMs' concentrations in agricultural soils of the Henan Province. (a) As; (b) Cd; (c) Cr; (d) Cu; (e) Mn; (f) Ni; (g) Pb; (h) Sb; (i) Zn.

mountain areas, and from city scale peak RI values were primarily in Luoyang, Samenxia, Jiyuan, Jiaozuo, and Anyang cities. The scatter peak levels in southwest (Nanyang), central (Xuchang), and northeast (Puyang) Henan Province are consistent with the spatial distribution of As and Cd in soils, likely due to the higher toxicity response factors of As (10) and Cd (30) (Liang et al., 2023). Zhu et al. (2023) observed ecological risks in cities in northern Taihang Mountain areas, such as Anyang, Xinxiang, Jiaozuo, Jiyuan, and Luoyang, which may be due to the superposition of geological background and industrial activities (Han et al., 2023). Moreover, a previous study also highlighted the dominant role of Cd in ecological risks in grain producing areas in Nanyang city (Zhang and Zhang, 2021). Ecological risks posed by soil HMs were thus notably high in Henan Province, highlighting the need for effective pollution control, at least at peak value sites.

3.3.2. *Health risk*

The mean health-risk values for children, and adult female and male were 2.64E–1, 4.04E–1, and 3.6E–1, respectively. These values are less than the USEPA threshold value of 1, and indicate that although soil HM exposure is higher for children than adults, there is no potential NCR risk of soil HM in Henan Province. Regarding the contribution of different exposure pathways, ingestion is dominant for both NCR (>90 %) and TCR (>80 %) (Fig. S3), followed by inhalation and dermal contact (Liu et al., 2022b). The contribution of the different HMs to NCR followed the order As > Cr > Mn > Pb > Sb > Ni > Cu > Zn > Cd (Fig. S4). Most TCR values were at acceptable levels, with a contribution order of As > Cd > Cr > Ni > Pb (Fig. S5).

Sensitivity analysis identified key factors influencing health risks, revealing that As concentration, ingestion rate, and exposure frequency significantly affect both non-carcinogenic and carcinogenic risks across all population groups (Fig. 7). For children, soil As concentration had the highest impact on HI values (0.670), followed by ingestion (0.523) and exposure frequency (0.416). In adult females and males, ingestion was the most significant factor for non-carcinogenic risk (0.727 and 0.720), followed by As concentration (0.488 and 0.467) and exposure frequency (0.336 and 0.358). As concentration had the greatest effect on TCR values (>0.6) for all groups, followed by ingestion and exposure frequency. Although both TCR and NCR were at acceptable levels, key factors like As emissions, exposure frequency, and ingestion should be controlled to prevent further risk (Zhu et al., 2023). Additionally, considering the bioavailability of HMs can improve the accuracy of health-risk evaluations (Liang et al., 2023). Individuals in occupations such as agriculture or construction, as well as children who frequently play outdoors, may have elevated soil ingestion rates due to their activities (Hubbard et al., 2022). Hence, more attentions and measures should be made to prevent these population from HM exposure.

3.3.3. *Risk identifications*

An understanding of the spatial distribution of risk is helpful in contamination prevention and control. For ecological risk and human-

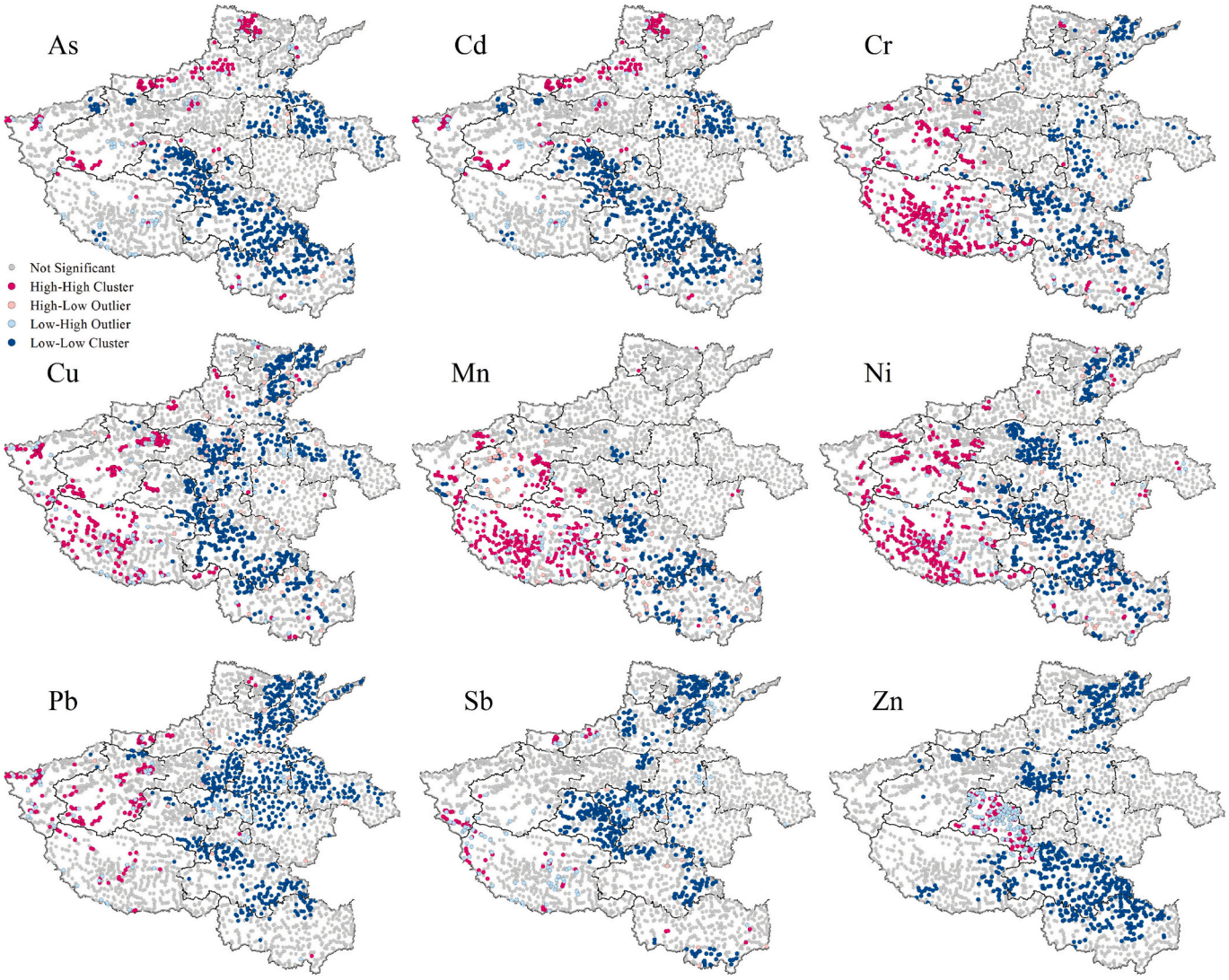


Fig. 5. Local Moran's I mapping of studied HMs in Henan Province.

exposure risk (both TCR and NCR) evaluation, it is necessary to elucidate the spatial clustering of combinations of these risks, with high-risk areas then being prioritizes. IDW spatial interpolation was applied in investigating risk hotspots in Henan Province. IDW is suitable for simpler applications, especially when the data is sparse or the spatial correlation is unclear (Xu et al., 2022a). Bivariate LMI analysis has proved effective in apportioning spatial clustering patterns of soil HMs (Shi et al., 2023; Yang et al., 2021).

Points with considerable risk RI values ($300 \leq RI < 600$) are scattered over different areas of Henan Province, with two main peak areas in Nanyang and Puyang cities (Fig. 6), reflecting the high soil As and Cd concentrations and their high toxic-response coefficients. Regarding HI values, there was a belt of peak values across the central and western parts of Henan Province, while for TCR values there were peak areas in central, mid-east, and southeast Henan Province and a high-value belt from Jiyuan to western Nanyang. For LMI, RI had several HH clusters in central and western Henan Province (Xuchang, Pingdingshan, and Luoyang), similar to those with high RI values. LL clusters of RI, HI, and TCR values were located mainly in the southeast the Henan Province.

Previous studies usually use the bivariate LMI to characterize the spatial correlation between two concentration categories (Wu et al., 2019). Here, two risk categories, calculated by Monte Carlo simulations, were considered, and high risks hotspots were identified. HH clusters of RI-HI, RI-TCR, and HI-TCR are distributed mainly in central and western Henan Province, indicating a need for further management and control measures. HH clusters of RI-HI occur mainly in Xuchang, Pingdingshan, Luoyang, and western Nanyang, which are areas of ferrous and non-ferrous metal smelting, coal mining, and cement production. It follows that reduction in soil HM risks are needed in Xuchang, Pingdingshan, Luoyang, Nanyang, and Sanmenxia. Some cities that are characterized by intensive industries would require greater monitoring and management (He et al., 2022; Zhu et al., 2023). Regarding other scattered RI-HI HH clusters in different cities, most occur in or adjacent to industrial or mineral mining areas. RI-HI HH clusters of have a spatial distribution similar to that of RI-TCR, although fewer RI-TCR HH clusters were observed due to fewer slope factors being available for HM carcinogenic risks (Han et al., 2022). Overall, however, bivariate LMI results for HI-TCR indicate more HH clusters than for other patterns, likely due to having similar key influencing factors-soil As, ingestion rate, exposure frequency in the assessment of exposure risks (Han et al., 2022).

LMI results indicate that HH clusters also occur in central and western Henan Province, with most in Xuchang and Nanyang, where high–high values are observed. Low-Low TCR values are clustered in Zhumadian and Xinyang cities.

3.4. HM source apportionment

Spearman correlation analysis was initially used to probe their

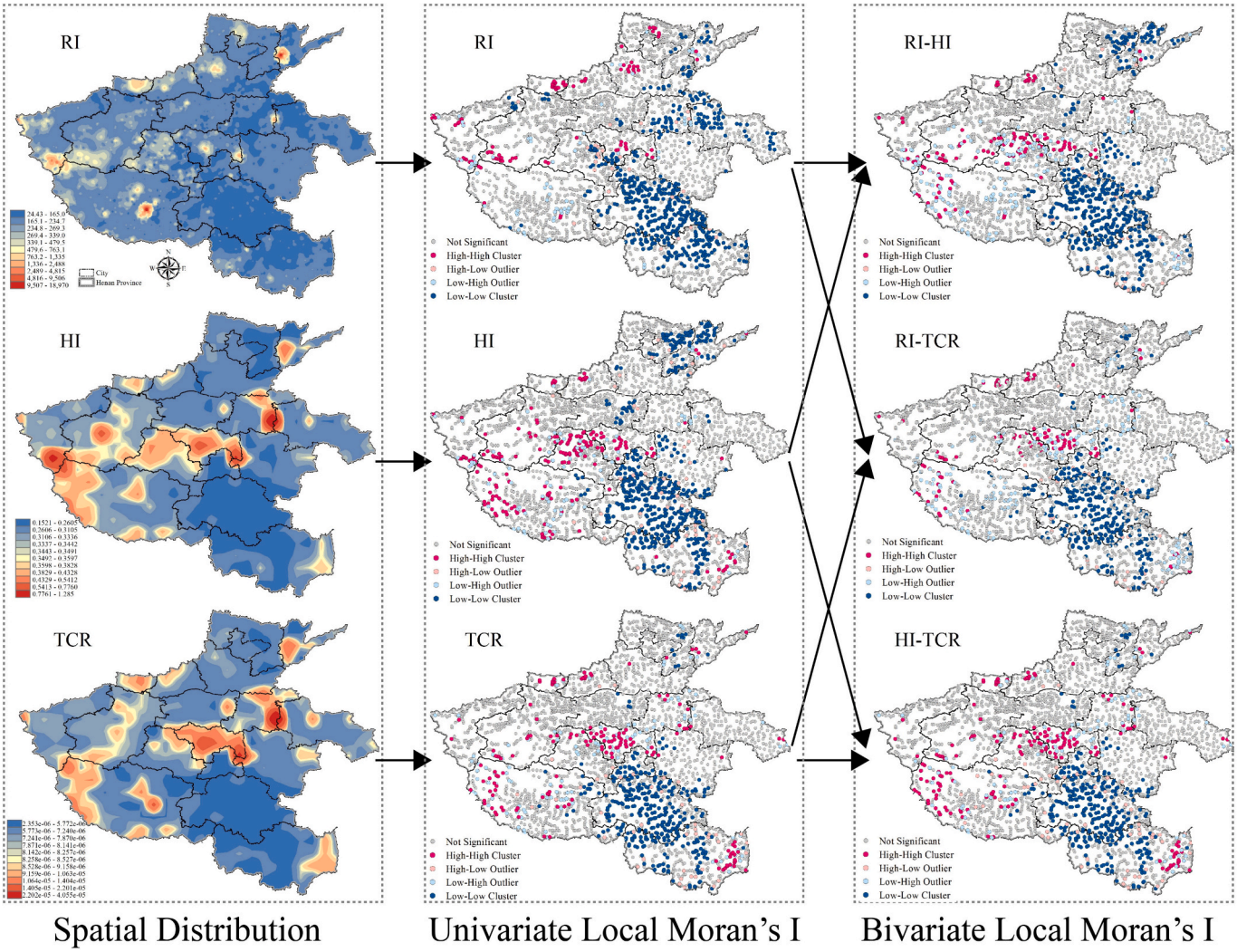


Fig. 6. Spatial distribution, univariate Local Moran's I and bivariate Local Moran's I of HMs in Henan Province.

relationships of HMs in agricultural soils to lend additional support to source attributions. A general correlationship can be observed due to the similarity in spatial distribution characteristics mentioned above. Specifically, Ni had the highest correlation with Cr ($r = 0.838$, $p < 0.01$) and Cu ($r = 0.808$, $p < 0.01$). Since Cr and Ni were considered as indicators of natural sources, thus, they may share the same natural sources (Han et al., 2022). Moreover, Cu is significantly correlated with Pb ($r = 0.702$, $p < 0.01$), followed by Zn ($r = 0.688$, $p < 0.01$), implying they may originate from same sources. Cd had the highest correlation with Pb ($r = 0.506$, $p < 0.01$), which is similar to the results obtained in PMF model (Fig. 8a). To verify source identifications, PMF factor contributions were obtained and correlated with the Pearson correlation coefficients (Zeng et al., 2024a). Four factors were apportioned, as follows.

Factor 1 contributed 23.3 % of HMs in agricultural soils in Henan Province, and is dominated by Mn (92.3 %), Ni (33.0 %), and Cr (31.0 %) (Fig. 8). In Henan Province, Mn concentrations averaged 539 mg kg^{-1}, lower than the Henan geochemical baselines (579 mg kg^{-1}) and mean Chinese geochemical baseline (569 mg kg^{-1}), indicating a lower impact of anthropogenic activities. Mn is a lithogenic element that occurs at relatively high concentrations in Earth's crust, and the soil Mn is not substantially affected by anthropogenic input (Hu and Cheng, 2016). Ni and Cr in soil originate from the weathering of parental materials (Zhu et al., 2023). Hence, Factor 1 could be interpreted as natural sources.

Factor 2 is weighted mainly by Sb (100 %), Cr (54.3 %), Ni (49.1 %), Pb (48.6 %), and Zn (38.9 %), explaining 37.8 % of total contribution. With the increasing use of motor vehicles that have Sb in brake linings, the environmental contribution of Sb from vehicular traffic activities is substantial (Harrison et al., 2012). Vehicle brakes and clutch systems are also Pb sources (Jeong et al., 2022). Although China has promoted unleaded gasoline since 2000, some soil Pb may still be sourced from gasoline use, given its environmental persistence, thus, Pb is the signature element of a traffic source (Zhou et al., 2023). Zn serves as an antioxidant and detergent in lubricants, or as a vital constituent of vehicle tires (Wang et al., 2019). Previous studies have confirmed Pb, Cr, and Zn as being derived from traffic sources (Zhou and Wang, 2019). Simultaneously, in Henan Province, there are many vehicles (26.23 million vehicles, ranking 3rd in China) with 0.27 million km of highways (5th in China) and 3.52 billion passenger movements per year (1st in China) (NBS, 2023). Therefore, Factor 2 represents traffic sources.

Factor 3 is characterized by As (90.2 %), explaining 18.4 % of the total variance. Elevated soil As levels may be related to deposition from coal combustion and industrial activity (Bhuiyan et al., 2015). Henan Province is rich in coal deposits (total reserves 1.01 billion tonnes), with 622 operational coal mines (HSB, 2023). Sites of elevated As concentrations are located mainly in central and western Henan Province (Fig. 4), where intensive industries activity occurs. In particular, smallpapermaking, tanning, printing, dyeing, smelting, refining, and electroplating related to As emissions were reported in these areas (Gong et al., 2020). Moreover, in these areas, coal's coking process (Yang and

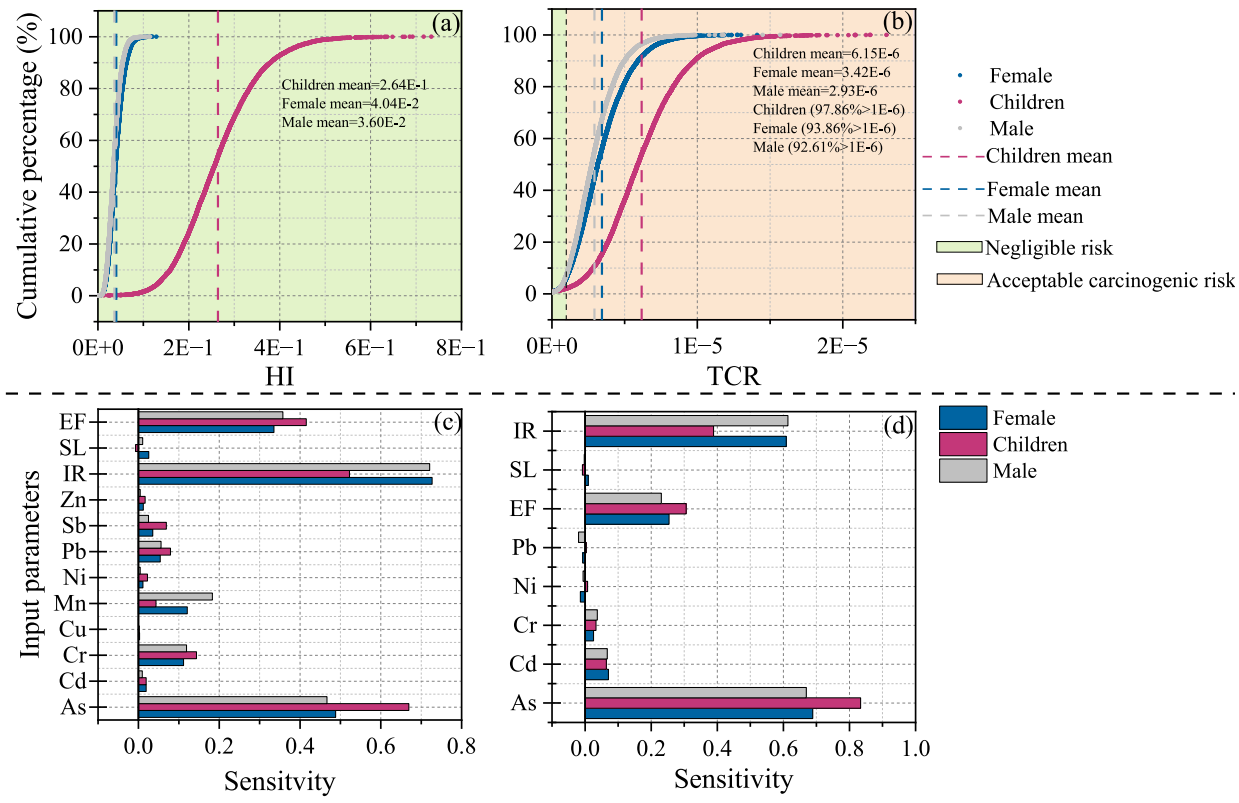


Fig. 7. The probability distribution for non-carcinogenic risk (a) and carcinogenic risk (b), sensitivity analyses of exposure parameters to non-carcinogenic risk (c) and carcinogenic risk (d).

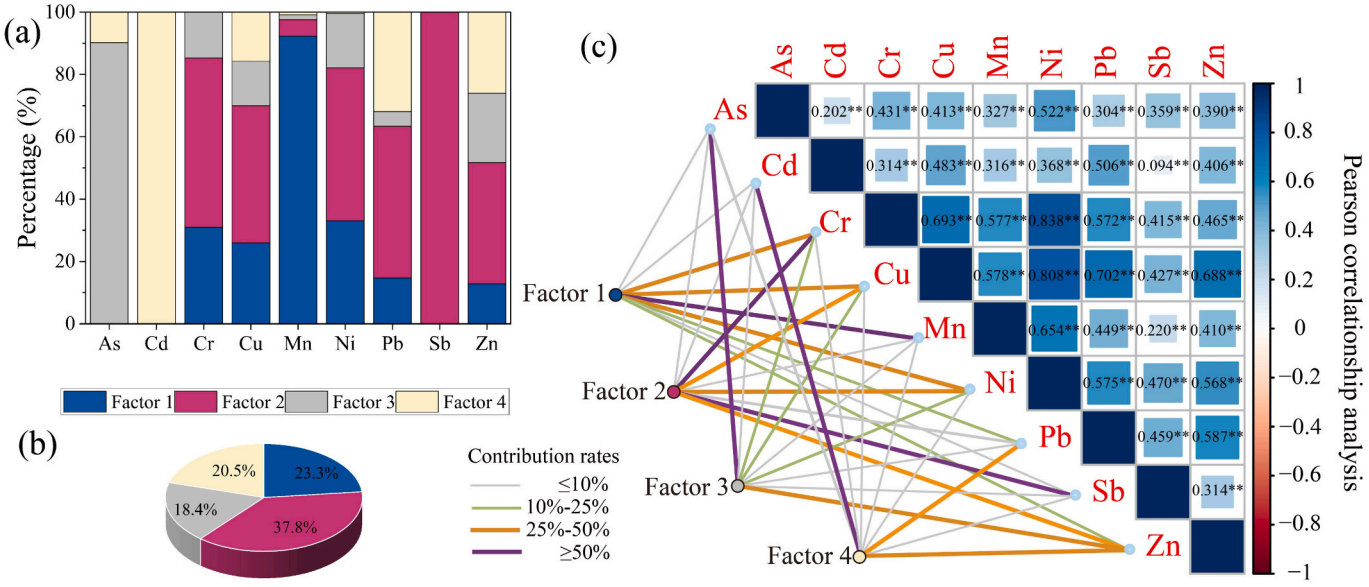


Fig. 8. Source apportionment of soil HMs (a) PMF factor profiles of HMs; (b) percentage contributions of each PMF factor; (c) Spearman correlationship among HMs and between HMs and PMF factors.

Liu, 2022) and leather-making can also contribute to elevated As in surrounding agricultural soils (Hua et al., 2019). Factor 3 is therefore assigned to industrial sources.

Factor 4 is dominated by Cd (100 %) and Pb (31.9 %), which together explain 20.5 % of the total variance. Soil Cd concentrations are higher than Henan background levels, indicating the impact of anthropogenic sources. The spatial distribution of Cd differs from that of other HMs, implying a different set of sources. Chemical fertilizers and pesticides are major sources of Cd and Pb (Meng et al., 2021a), and the annual application amounts were 6.67 million tonnes and 0.107 million tonnes in Henan Province, respectively (HSB, 2023). Moreover, additives containing HMs are often fed to livestock, so HMs are expected in livestock manure (Wang et al., 2013). The application of livestock manure results in enhanced levels of HM in soil (Liu et al., 2020). Cd concentrations show little variation over Henan Province (Fig. 6). Hence, Factor 4 was assigned to agricultural activities.

Vehicular traffic is generally the dominant source of HMs in agricultural soils in Henan Province (contributing 37.8 %), followed by natural (23.3 %) and industrial sources (20.5 %). Generally, the findings are consistent with some prior researches highlighting traffic as a primary contributor to soil contamination in Zhengzhou (Liu et al., 2022c) and Shangqiu city (Shi et al., 2022) in Henan Province. It's worth noting that the relative contributions of different pollution sources can vary depending on local factors such as land use patterns, industrial activities, and transportation infrastructure.

3.5. Factors influencing HM

Although the main potential sources of HM have been identified, it is necessary to identify priority environmental variables for pollution management, which will affect soil HM accumulation directly or indirectly. Factors affecting soil HM concentrations at a province scale were identified using RF models, and ranked by their quantitative contributions (Fig. 9). The 22 selected environmental variables were categorized into two main groups: (1) natural factors including temperature, elevation, fine sand, soil type, precipitation, clay particles, total nitrogen, silty sand, organic matter, coarse sand, pH, total phosphorus, and total potassium; and (2) anthropogenic factors including population, railways, chemical fertilizers, $PM_{2.5}$, residential zones, factories, pesticides, river, and GDP. Since HMs from industrial effluents and sewage can contaminate farmland soils via river irrigation, rivers are deemed an anthropogenic factor.

Overall, RF modeling confirmed that most of the studied HMs originated from both natural and anthropogenic sources, albeit with contributions of different variables varying considerably (Fig. 9). Soil type ranked in the top ten driving factors for all HMs; $PM_{2.5}$, GDP,

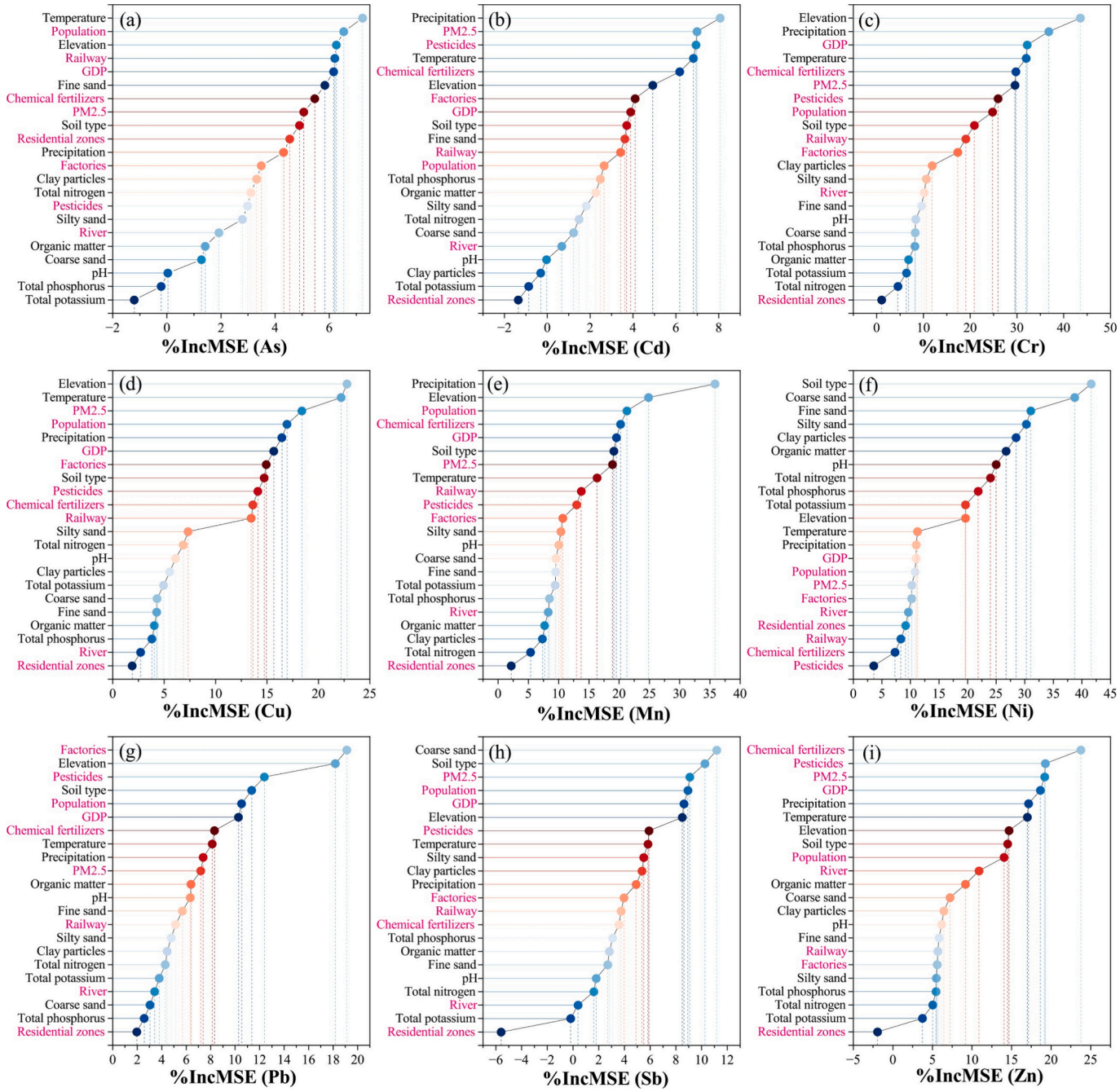


Fig. 9. HM specific relative importance of each environmental variable in Henan Province.

temperature, and elevation contributed to most HMs except Ni; chemical fertilizers and pesticides contributed to seven HMs; and precipitation to six HMs, with these being the dominant factors. Elevation was the predominant factor influencing Cr and Cu concentrations, and precipitation for Cd and Mn. For most of the studied HMs (excluding Pb and Zn), natural factors ranked above anthropogenic variables. All of the various anthropogenic variables were related to different HMs, indicating their role in HM accumulation. In this study, 48 soil types were observed (Fig. S6). For the dominant risk contributing elements, the highest mean As (25.0 ± 26.8 mg kg^{-1}) and Pb (42.5 ± 28.5 mg kg^{-1}) were both observed in gray tide soil; and Cd (0.897 ± 6.05 mg kg^{-1}) in sticky plate yellow brown soil. The standard deviations of As and Cd were even higher than the mean values, suggesting the high variation in soil types.

Anthropogenic sources of As, Cd, Cu, Pb, Sb, and Pb require special attention in controlling HM contamination. Population, railway, GDP, chemical fertilizers, and $PM_{2.5}$ were the main anthropogenic variables affecting As levels; $PM_{2.5}$, pesticides, chemical fertilizers, factories, and GDP affected Cd levels; GDP, chemical fertilizers, $PM_{2.5}$, pesticides, and population had the highest impact on Cu concentrations; factories, pesticides, population, GDP, and chemical fertilizers on Pb; $PM_{2.5}$, population, GDP, pesticides, and factories on Sb; and chemical fertilizers, pesticides, $PM_{2.5}$, GDP, and population on Zn concentrations, with contributions from both agricultural and industrial sources (Fig. S7-S27).

GDP and population were not direct pollution sources, but could indicate the influence of anthropogenic activities (Hu and Cheng, 2013). Of all anthropogenic variables, $PM_{2.5}$ had the highest contribution to all studied HMs. In Henan Province, As in $PM_{2.5}$ is derived mainly from domestic and industrial fossil-fuel combustion. Considering the high proportion of industrial energy consumption (55.8 % of the total energy) industries likely contribute most of the $PM_{2.5}$ (Wang et al., 2016; Xing et al., 2019a). The high contributions of railway and chemical fertilizers indicate that traffic and agricultural sources are major contributors to As contamination (Baltas et al., 2020; Liu et al., 2022c). The predominant anthropogenic sources are agricultural and industrial.

RF analysis indicated non-linear relationships between the different above factors, highlighting the importance of environmental variables. The lack of emission data for the selected variables means that the results in this study may only serve as a reference to guide pollution control. Future studies should develop detailed emission inventories to allow accurate evaluation of soil HM source. It should be noted that historical pollution was not considered here. Such shortcomings would have introduced bias to the results in this study. However, since Henan is a traditional farming province, and few legacies polluted sites (Peng et al., 2022), thus, the bias, arisen by historical pollution sites, on the key findings is limited. Due to the diversity of industries structure in different cities, high resolution covariates are required to predict soil HM concentrations and evaluate contributing factors.

4. Conclusion

Agricultural soils in Henan Province generally present low risks from HMs, with Cd being the primary concern, followed by As and Pb. The highest HM concentrations are located in the western mountainous regions, with notable peaks in Xuchang and Nanyang, and elevated Zn levels in Pingdingshan. The main ecological risks stem from Cd and Sb. Both carcinogenic and non-carcinogenic risks are within safe limits. Areas with significant combined risks to both ecology and human health have been identified in Xuchang and Nanyang. Vehicular traffic is the main source of HMs in agricultural soils (37.8 %), followed by natural and industrial sources. Soil type and $PM_{2.5}$ contribute as natural and anthropogenic factors, respectively. Since Cd mainly originates from agricultural activities, it is essential to promote the proper disposal of agricultural wastes like pesticides, herbicides, and fertilizers.

CRediT authorship contribution statement

Xiaohang Xu: Writing – review & editing, Writing – original draft, Visualization, Validation, Project administration, Investigation, Formal analysis, Data curation, Conceptualization. **Zhidong Xu:** Investigation, Data curation. **Longchao Liang:** Investigation, Data curation. **Gaoen Wu:** Investigation, Data curation. **Qinhui Lu:** Formal analysis, Data curation. **Lin Liu:** Formal analysis, Data curation. **Qiao Han:** Data curation. **Le Wang:** Investigation. **Sensen Zhang:** Project administration, Investigation, Conceptualization. **Yanhai Hu:** Conceptualization. **Yuping Jiang:** Conceptualization. **Jialin Yang:** Investigation. **Guangle Qiu:** Writing – review & editing, Supervision, Conceptualization. **Pan Wu:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The author (Xiaohang Xu) was supported by the Special Research Fund of Natural Science (Special Post) of Guizhou University (202303). This study was funded by Guizhou Provincial Key Technology R&D Program ([2020]1Y140), the High-Level Talent Training Program in Guizhou Province (GCC[2023]045), Henan Academy of Geology (Project no. 2023-904-XM009-KT01), Henan Non-ferrous Metals Geological and Mineral Resources Bureau (Project no. KCXM202018); National Natural Science Foundation of China (42003065, 42467029).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.scitotenv.2024.176359.

References

Abokifa, A.A., Katz, L., Sela, L., 2020. Spatiotemporal trends of recovery from lead contamination in Flint, MI as revealed by crowdsourced water sampling. Water Res. 171, 115442.

Agyeman, P.C., John, K., Kebonye, N.M., Ofori, S., Borůvka, L., Vašát, R., Kočárek, M., 2022. Ecological risk source distribution, uncertainty analysis, and application of geographically weighted regression cokriging for prediction of potentially toxic elements in agricultural soils. Process Saf. Environ. Prot. 164, 729–746.

Akopyan, K., Petrosyan, V., Grigoryan, R., Melkom Melkomian, D., 2018. Assessment of residential soil contamination with arsenic and lead in mining and smelting towns of northern Armenia. J. Geochem. Explor. 184, 97–109.

Arfaeinia, H., Dobaradaran, S., Moradi, M., Pasalari, H., Mehrizi, E.A., Taghizadeh, F., Esmaili, A., Ansarizadeh, M., 2019. The effect of land use configurations on concentration, spatial distribution, and ecological risk of heavy metals in coastal sediments of northern part along the Persian Gulf. Sci. Total Environ. 653, 783–791.

Baltas, H., Sirin, M., Gokbayrak, E., Ozcelik, A.E., 2020. A case study on pollution and a human health risk assessment of heavy metals in agricultural soils around Sinop province. Turkey. Chemosphere 241, 125015.

Bhuiyan, M.A., Dampare, S.B., Islam, M.A., Suzuki, S., 2015. Source apportionment and pollution evaluation of heavy metals in water and sediments of Buriganga River, Bangladesh, using multivariate analysis and pollution evaluation indices. Environ. Monit. Assess. 187 (1), 4075.

Bremian, L., 2001. Random Forests.

Chen, H., Teng, Y., Lu, S., Wang, Y., Wang, J., 2015. Contamination features and health risk of soil heavy metals in China. Sci. Total Environ. 512-513, 143–153.

Chen, Z., Zhao, Y., Chen, D., Huang, H., Zhao, Y., Wu, Y., 2023. Ecological risk assessment and cumulative early warning of heavy metals in the soils near the Luanchuan molybdenum polymetallic ore concentration area. Henan. China Geol. 6 (1), 15–26.

Cheng, X., Sun, J., Jia, X., Liu, X., Zhao, Y., 2023. Pollution characteristics and health risk assessment of heavy metals in farmland soil around the molybdenum mining area in Luanchuan. Henan Province. Geol. China 1–20.

Gong, Y., Qu, Y., Yang, S., Tao, S., Shi, T., Liu, Q., Chen, Y., Wu, Y., Ma, J., 2020. Status of arsenic accumulation in agricultural soils across China (1985-2016). Environ. Res. 186, 109525.
Guan, Q., Zhao, R., Pan, N., Wang, F., Yang, Y., Luo, H., 2019. Source apportionment of heavy metals in farmland soil of Wuwei, China: comparison of three receptor models. J. Clean. Prod. 237, 117792.
Hakanson, L., 1980. An ecological risk index for aquatic pollution control.a sedimentological approach. Water Res. 14 (8), 975–1001.
Han, J., Liang, L., Zhu, Y., Xu, X., Wang, L., Shang, L., Wu, P., Wu, Q., Qian, X., Qiu, G., Feng, X., 2022. Heavy metal(loid)s in farmland soils in the karst plateau, Southwest China - an integrated analysis of geochemical baselines, source apportionment and associated health risk. Land Degrad. Dev. 33 (10), 1689–1703.
Han, Q., Wang, M., Xu, X., Li, M., Liu, Y., Zhang, C., Li, S., Wang, M., 2023. Health risk assessment of heavy metals in road dust from the fourth-tier industrial city in Central China based on Monte Carlo simulation and bioaccessibility. Ecotoxicol. Environ. Saf. 252, 114627.
Harrison, R.M., Jones, A.M., Gietl, J., Yin, J., Green, D.C., 2012. Estimation of the contributions of brake dust, Tire Wear, and resuspension to nonexhaust traffic particles derived from atmospheric measurements. Environ. Sci. Technol. 46 (12), 6523–6529.
He, J., Xue, W., Yan, L., Shi, X., Wang, Y., Lei, Y., Zheng, Y., Zhang, Y., 2022. Multi-dimensional sources apportionment of PM2.5 in Zhongyuan urban agglomeration based on the CAMx mode. Aerosol Air Qual. Res. 22 (5), 210254.
HSB, 2023. Henan Statistical Yearbook. China Statistics Press, Beijing, Henan Statistical Bureau.
Hu, Y., Cheng, H., 2016. A method for apportionment of natural and anthropogenic contributions to heavy metal loadings in the surface soils across large-scale regions. Environ. Pollut. 214, 400–409.
Hu, Y., He, K., Sun, Z., Chen, G., Cheng, H., 2020. Quantitative source apportionment of heavy metal(loid)s in the agricultural soils of an industrializing region and associated model uncertainty. J. Hazard. Mater. 391, 122244.
Hu, Y.N., Cheng, H.F., 2013. Application of stochastic models in identification and apportionment of heavy metal pollution sources in the surface soils of a large-scale region. Environ. Sci. Technol. 47 (8), 3752–3760.
Hua, Y., Ma, Y., Yi, J., 2019. Current situation and countermeasures of soil pollution prevention and control in heavy metal industry in Henan Province. Sci. Technol. Inform. 17 (06), 96–98.
Huang, G., Wang, X., Chen, D., Wang, Y., Zhu, S., Zhang, T., Liao, L., Tian, Z., Wei, N., 2022. A hybrid data-driven framework for diagnosing contributing factors for soil heavy metal contaminations using machine learning and spatial clustering analysis. J. Hazard. Mater. 437, 129324.
Hubbard, H., Ozkaynak, H., Glen, G., Cohen, J., Thomas, K., Phillips, L., Tulve, N., 2022. Model-based predictions of soil and dust ingestion rates for U.S. adults using the stochastic human exposure and dose simulation soil and dust model. Sci. Total Environ. 846, 157501.
Hui, W., Hao, Z., Hongyan, T., Jiawei, W., Anna, L., 2021. Heavy metal pollution characteristics and health risk evaluation of soil around a tungsten-molybdenum mine in Luoyang. China. Environ. Earth Sci. 80 (7).
Jeong, H., Ryu, J.S., Ra, K., 2022. Characteristics of potentially toxic elements and multi-isotope signatures (cu, Zn, Pb) in non-exhaust traffic emission sources. Environ. Pollut. 292 (Pt A), 118339.
Jiang, H.H., Cai, L.M., Wen, H.H., Luo, J., 2020b. Characterizing pollution and source identification of heavy metals in soils using geochemical baseline and PMF approach. Sci. Rep. 10 (1), 6460.
Jiang, H.-H., Cai, L.-M., Wen, H.-H., Hu, G.-C., Chen, L.-G., Luo, J., 2020a. An integrated approach to quantifying ecological and human health risks from different sources of soil heavy metals. Sci. Total Environ. 701, 134466.
Jiang, N., Liu, X., Wang, S., Yu, X., Yin, S., Duan, S., Wang, S., Zhang, R., Li, S., 2019. Pollution characterization, source identification, and health risks of atmospheric-particle-bound heavy metals in PM10 and PM2.5 at multiple sites in an emerging megacity in the central region of China. Aerosol Air Qual. Res. 19 (2), 247–271.
Jiang, Y., Ma, J., Ruan, X., Chen, X., 2020c. Compound health risk assessment of cumulative heavy metal exposure: a case study of a village near a battery factory in Henan Province. China. Environ. Sci. Process Impacts 22 (6), 1408–1422.
Jin, Z., Zhang, L., Lv, J., Sun, X., 2021. The application of geostatistical analysis and receptor model for the spatial distribution and sources of potentially toxic elements in soils. Environ. Geochem. Health 43 (1), 407–421.
Kulikova, T., Hiller, E., Jurkovič, Ľ., Filová, L., Šottník, P., Lacina, P., 2019. Total mercury, chromium, nickel and other trace chemical element contents in soils at an old cinnabar mine site (Merník, Slovakia): anthropogenic versus natural sources of soil contamination. Environ. Monit. Assess. 191 (5).
Lai, S., Dong, Q., Song, C., Yang, Z., 2021. Distribution characteristics and ecological risk assessment of soil heavy metals in the eastern mountainous area of the Nanyang Basin. Environ. Sci. 42 (11), 5500–5509.
Li, K., Gu, Y., Li, M., Zhao, L., Ding, J., Lun, Z., Tian, W., 2018. Spatial analysis, source identification and risk assessment of heavy metals in a coal mining area in Henan, Central China. Int. Biodeterior. Biodegradation 128, 148–154.
Li, K., Wang, J., Zhang, Y., 2022. Heavy metal pollution risk of cultivated land from industrial production in China: spatial pattern and its enlightenment. Sci. Total Environ. 828, 154382.
Liang, L., Zhu, Y., Xu, X., Hao, W., Han, J., Chen, Z., Dong, X., Qiu, G., 2023. Integrated insights into source apportionment and source-specific health risks of potential pollutants in Urban Park soils on the karst plateau. SW China. Expo. Health 1–18.
Liu, J., Ouyang, Q.E., Wang, L., Wang, J., Zhang, Q., Wei, X., Lin, Y., Zhou, Y., Yuan, W., Xiao, T., 2022. Quantification of smelter-derived contributions to thallium contamination in river sediments: novel insights from thallium isotope evidence. J. Hazard. Mater. 424, 127594.
Liu, K., Gao, L., Zu, Z., Jia, G., 2023. Monitoring heavy metal content in rural soil of Xuchang city, Henan province from 2017-2021. Mod. Dis. Control Prevent. 34 (10), 767–786.
Liu, L., Xu, X., Han, J., Zhu, J.-M., Li, S., Liang, L., Wu, P., Wu, Q., Qiu, G., 2022b. Heavy metal(loid)s in agricultural soils in the world's largest barium-mining area: pollution characteristics, source apportionment, and health risks using PMF model and cd isotopes. Process. Saf. Environ. Prot. 166, 669–681.
Liu, W.R., Zeng, D., She, L., Su, W.X., He, D.C., Wu, G.Y., Ma, X.R., Jiang, S., Jiang, C.H., Ying, G.G., 2020. Comparisons of pollution characteristics, emission situations, and mass loads for heavy metals in the manures of different livestock and poultry in China. Sci. Total Environ. 734, 139023.
Liu, X., Gu, S., Yang, S., Deng, J., Xu, J., 2021a. Heavy metals in soil-vegetable system around E-waste site and the health risk assessment. Sci. Total Environ. 779, 146438.
Liu, X., Zhu, H., Zhang, B., Xu, C., Li, L., Xing, W., 2022c. Heavy metals (HMs) in soils of different land-use types in Zhengzhou, China: occurrence, source and ecological risk. Soil and sediment contamination: an. Int. J. 1–21.
Liu, Y.B., Ma, Z.H., Liu, G.N., Jiang, L., Dong, L.M., He, Y., Shang, Z.F., Shi, H.D., 2021b. Accumulation risk and source apportionment of heavy metals in different types of farmland in a typical farming area of northern China. Environ. Geochem. Health 43 (12), 5177–5194.
MARAPRC, 2012. Technical Rules for Monitoring of Environmental Quality of Farmland Soil (NY/T 395–2012).
Meng, M., Yang, L., Yu, J., Wei, B., Li, H., Cao, Z., Chen, Q., Zhang, G., 2021a. Identification of spatial patterns and sources of heavy metals in greenhouse soils using geostatistical and positive matrix factorization (PMF) methods. Land Degrad. Dev. 32 (18), 5412–5426.
Meng, X., Guo, J., Yang, J., Yang, J., Zheng, G., Qiao, P., Bian, J., Chen, T., 2021b. Spatial distribution and risk assessment of heavy metal pollution in farmland soils surrounding a typical industrial area of Henan Province. Environ. Sci. 42 (02), 900–908.
MohseniBandpi, A., Eslami, A., Ghaderpoori, M., Shahsavani, A., Jeihooni, A.K., Ghaderpoury, A., Alinejad, A., 2018. Health risk assessment of heavy metals on PM2.5 in Tehran air. Iran. Data in Brief 17, 347–355.
NBS, 2023. National Data Statistics of China (Last accessible December 1st, 2023). https://data.stats.gov.cn/english/.
Peng, J.Y., Zhang, S., Han, Y., Bate, B., Ke, H., Chen, Y., 2022. Soil heavy metal pollution of industrial legacies in China and health risk assessment. Sci. Total Environ. 816, 151632.
Rinklebe, J., Shaheen, S.M., El-Naggar, A., Wang, H., Du Laing, G., Alessi, D.S., Sik Ok, Y., 2020. Redox-induced mobilization of ag, Sb, Sn, and Tl in the dissolved, colloidal and solid phase of a biochar-treated and un-treated mining soil. Environ. Int. 140, 105754.
Shi, H., Wang, P., Zheng, J., Deng, Y., Zhuang, C., Huang, F., Xiao, R., 2023. A comprehensive framework for identifying contributing factors of soil trace metal pollution using Geodetector and spatial bivariate analysis. Sci. Total Environ. 857 (Pt 3), 159636.
Shi, X.M., Liu, S., Song, L., Wu, C.S., Yang, B., Lu, H.Z., Wang, X., Zakari, S., 2022. Contamination and source-specific risk analysis of soil heavy metals in a typical coal industrial city, Central China. Sci. Total Environ. 836, 155694.
Sun, X., Wang, H., Guo, Z., Lu, P., Song, F., Liu, L., Liu, J., Rose, N.L., Wang, F., 2020. Positive matrix factorization on source apportionment for typical pollutants in different environmental media: a review. Environ. Sci.: Processes Impacts 22 (2), 239–255.
Swidwa-Urbanska, J., Batlle-Sales, J., 2021. Data quality oriented procedure, for detailed mapping of heavy metals in urban topsoil as an approach to human health risk assessment. J. Environ. Manag. 295, 113019.
Tokatlı, C., Varol, M., Ustaoğlu, F., Muhammad, S., 2023. Pollution characteristics, sources and health risks assessment of potentially hazardous elements in sediments of ten ponds in the Saros Bay region (Türkiye). Chemosphere 340, 139977.
USEPA, 2009. (United States Environmental Protection Agency) Risk Assessment Guidance for Superfund (RAGS), Volume I, Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment).
Wang, H., Dong, Y., Yang, Y., Toor, G.S., Zhang, X., 2013. Changes in heavy metal contents in animal feeds and manures in an intensive animal production region of China. J. Environ. Sci. 25 (12), 2435–2442.
Wang, L., Jin, Y., Weiss, D.J., Schleicher, N.J., Wilcke, W., Wu, L., Guo, Q., Chen, J., O'Connor, D., Hou, D., 2021. Possible application of stable isotope compositions for the identification of metal sources in soil. J. Hazard. Mater. 407, 124812.
Wang, S., Cai, L.M., Wen, H.H., Luo, J., Wang, Q.S., Liu, X., 2019. Spatial distribution and source apportionment of heavy metals in soil from a typical county-level city of Guangdong Province, China. Sci. Total Environ. 655, 92–101.
Wang, X., 2014. The Study on Development Problems of Industry Agglomeration Area in Xuchang City.
Wang, Y., Li, J., Li, A., Xie, P., Zheng, H., Zhang, Y., Wang, Z., 2016. Modeling study of surface $PM_{2.5}$ and its source apportionment over. Henan in 2013—2014 36 (10), 3543–3553.
Wang, Y., Wang, S., Jiang, L., Ma, L., Li, X., Zhong, M., Zhang, W., 2022. Does the geographic difference of soil properties matter for setting up the soil screening levels in large countries like China? Environ. Sci. Technol. 56, 5684–5693.
Wu, H., Yang, F., Li, H., Li, Q., Zhang, F., Ba, Y., Cui, L., Sun, L., Lv, T., Wang, N., Zhu, J., 2020. Heavy metal pollution and health risk assessment of agricultural soil near a smelter in an industrial city in China. Int. J. Environ. Health Res. 30 (2), 174–186.

Wu, Q., Hu, W., Wang, H., Liu, P., Wang, X., Huang, B., 2021. Spatial distribution, ecological risk and sources of heavy metals in soils from a typical economic development area. Southeastern China. Sci. Total Environ. 780, 146557.
Wu, S., Zhou, S., Bao, H., Chen, D., Wang, C., Li, B., Tong, G., Yuan, Y., Xu, B., 2019. Improving risk management by using the spatial interaction relationship of heavy metals and PAHs in urban soil. J. Hazard. Mater. 364, 108–116.
XCSY, 2017. Xuchang Statistical Yearbook. China Statistics Press, Beijing, Xuchang Statistical Bureau.
Xing, W., Zhang, H., Scheckel, K.G., Li, L., 2016. Heavy metal and metalloid concentrations in components of 25 wheat (Triticum aestivum) varieties in the vicinity of lead smelters in Henan province. China. Environ. Monit. Assess. 188 (1), 23.
Xing, W., Cao, E., Scheckel, K.G., Bai, X., Li, L., 2018. Influence of phosphate amendment and zinc foliar application on heavy metal accumulation in wheat and on soil extractability impacted by a lead smelter near Jiyuan. China. Environ. Sci. Pollut. Res. Int. 25 (31), 31396–31406.
Xing, W., Zhao, Q., Scheckel, K.G., Zheng, L., Li, L., 2019a. Inhalation bioaccessibility of cd, cu, Pb and Zn and speciation of Pb in particulate matter fractions from areas with different pollution characteristics in Henan Province, China. Ecotoxicol. Environ. Saf. 175, 192–200.
Xing, W., Zheng, Y., Scheckel, K.G., Luo, Y., Li, L., 2019b. Spatial distribution of smelter emission heavy metals on farmland soil. Environ. Monit. Assess. 191 (2), 115.
Xu, H., Croot, P., Zhang, C., 2022a. Exploration of the spatially varying relationships between lead and aluminium concentrations in the topsoil of northern half of Ireland using geographically weighted Pearson correlation coefficient. Geoderma 409, 115640.
Xu, J., Hu, C., Wang, M., Zhao, Z., Zhao, X., Cao, L., Lu, Y., Cai, X., 2022b. Changeable effects of coexisting heavy metals on transfer of cadmium from soils to wheat grains. J. Hazard. Mater. 423 (Pt B), 127182.
Xu, X., Gu, C., Feng, X., Qiu, G., Shang, L., Xu, Z., Lu, Q., Xiao, D., Wang, H., Lin, Y., 2019. Weir building: a potential cost-effective method for reducing mercury leaching from abandoned mining tailings. Sci. Total Environ. 651, 171–178.
Xu, X.H., Meng, B., Zhang, C., Feng, X.B., Gu, C.H., Guo, J.Y., Bishop, K., Xu, Z.D., Zhang, S.S., Qiu, G.L., 2017. The local impact of a coal-fired power plant on inorganic mercury and methyl-mercury distribution in rice (Oryza sativa L.). Environ. Pollut. 223, 11–18.
Yang, L., Ren, Q., Zheng, K., Jiao, Z., Ruan, X., Wang, Y., 2022. Migration of heavy metals in the soil-grape system and potential health risk assessment. Sci. Total Environ. 806 (Pt 2), 150646.
Yang, S., He, M., Zhi, Y., Chang, S.X., Gu, B., Liu, X., Xu, J., 2019. An integrated analysis on source-exposure risk of heavy metals in agricultural soils near intense electronic waste recycling activities. Environ. Int. 133, 105239.
Yang, S., Qu, Y., Ma, J., Liu, L., Wu, H., Liu, Q., Gong, Y., Chen, Y., Wu, Y., 2020a. Comparison of the concentrations, sources, and distributions of heavy metal(loid)s in agricultural soils of two provinces in the Yangtze River Delta. China. Environ. Pollut. 264, 114688.
Yang, S., Taylor, D., Yang, D., He, M., Liu, X., Xu, J., 2021. A synthesis framework using machine learning and spatial bivariate analysis to identify drivers and hotspots of heavy metal pollution of agricultural soils. Environ. Pollut. 287, 117611.
Yang, W., Liu, Y., 2022. Heavy metal pollution assessment of key soil supervision units in Pingdingshan City. Leather Making and Environmental Protection Technology 3 (20).
Yang, Y., Yang, X., He, M., Christakos, G., 2020b. Beyond mere pollution source identification: determination of land covers emitting soil heavy metals by combining PCA/APCS, GeoDetector and GIS analysis. Catena 185.
Yuan, Y.M., Cave, M., Zhang, C.S., 2018. Using local Moran's I to identify contamination hotspots of rare earth elements in urban soils of London. Appl. Geochem. 88, 167–178.
Zeng, J., Han, G., Wu, Q., Qu, R., Ma, Q., Chen, J., Mao, S., Ge, X., Wang, Z.J., Ma, Z., 2024a. Significant influence of urban human activities and marine input on rainwater chemistry in a coastal large city. China. Water Res. 257, 121657.
Zeng, J., Han, G., Zhang, S., Zhang, Q., Qu, R., 2024b. Potentially toxic elements in rainwater during extreme rainfall period in the megacity Beijing: variations, sources, and reuse potential. Atmos. Environ. 318.
Zhang, H., Cheng, S., Li, H., Fu, K., Xu, Y., 2020. Groundwater pollution source identification and apportionment using PMF and PCA-APCA-MLR receptor models in a typical mixed land-use area in southwestern China. Sci. Total Environ. 741, 140383.
Zhang, J., Liu, Z., Tian, B., Li, J., Luo, J., Wang, X., Ai, S., Wang, X., 2023. Assessment of soil heavy metal pollution in provinces of China based on different soil types: from normalization to soil quality criteria and ecological risk assessment. J. Hazard. Mater. 441, 129891.
Zhang, M., Wang, J.M., Li, S.J., 2019. Tempo-spatial changes and main anthropogenic influence factors of vegetation fractional coverage in a large-scale opencast coal mine area from 1992 to 2015. J. Clean. Prod. 232, 940–952.
Zhang, Q., Zhang, X., 2021. Quantitative source apportionment and ecological risk assessment of heavy metals in soil of a grain base in Henan Province, China, using PCA, PMF modeling, and geostatistical techniques. Environ. Monit. Assess. 193 (10), 655.
Zhang, X., Zhong, T., Liu, L., Ouyang, X., 2015. Impact of soil heavy metal pollution on food safety in China. PLoS One 10 (8), e0135182.
Zhang, Y., Li, Y., 2021. Determination of heavy metal geochemical baseline value of surface soil in Yongcheng, Henan, and its accumulation characteristics. Glob. Geol. 40 (03), 728–737.
Zhang, Y., Wu, Y., Song, B., Zhou, L., Wang, F., Pang, R., 2022. Spatial distribution and main controlling factor of cadmium accumulation in agricultural soils in Guizhou. China. J. Hazard. Mater. 424, 127308.
Zhao, W., Ma, J., Liu, Q., Dou, L., Qu, Y., Shi, H., Sun, Y., Chen, H., Tian, Y., Wu, F., 2023. Accurate prediction of soil heavy metal pollution using an improved machine learning method: a case study in the Pearl River Delta, China. Environ. Sci. Technol. 57 (46), 17751–17761.
Zhong, X., Chen, Z., Li, Y., Ding, K., Liu, W., Liu, Y., Yuan, Y., Zhang, M., Baker, A.J.M., Yang, W., Fei, Y., Wang, Y., Chao, Y., Qiu, R., 2020. Factors influencing heavy metal availability and risk assessment of soils at typical metal mines in eastern China. J. Hazard. Mater. 400, 123289.
Zhou, H., Chen, Y., Yue, X., Ren, D., Liu, Y., Yang, K., 2023. Identification and hazard analysis of heavy metal sources in agricultural soils in ancient mining areas: a quantitative method based on the receptor model and risk assessment. J. Hazard. Mater. 445, 130528.
Zhou, L., Yang, B., Xue, N., Li, F., Seip, H.M., Cong, X., Yan, Y., Liu, B., Han, B., Li, H., 2014. Ecological risks and potential sources of heavy metals in agricultural soils from Huanghuai plain. China. Environ. Sci. Pollut. Res. Int. 21 (2), 1360–1369.
Zhou, X.-Y., Wang, X.-R., 2019. Impact of industrial activities on heavy metal contamination in soils in three major urban agglomerations of China. J. Clean. Prod. 230, 1–10.
Zhu, H., Liu, X., Wang, Q., Zhang, B., Xu, C., Wang, Z., Chen, H., 2023. Heavy metals pollution of soil in central plains urban agglomeration (CPUA), China: human health risk assessment based on Monte Carlo simulation. Environ. Geochem. 45, 8063–8079.