ORIGINAL PAPER

Source-oriented risk assessment of heavy metal(loid)s in agricultural soils around a multimetal smelting area near the Yellow River, China

Hui Chen · Shuo Qiao · Chang Li · Yong Wu · Donghao Li · Ling Li · Jianwei Liu

Abstract Heavy metal(loid) (HM) contamination in agricultural soils, particularly in areas severely impacted by smelting industries, has attracted worldwide attention. In this study, agricultural soils were collected in a flourishing multimetal smelting area near the Yellow River in central China. By an integrated approach encompassing the positive matrix factorization model, ordinary kriging interpolation and hierarchical clustering analysis (PMF-OK-HC), a total of four major sources and their mass contributions were identified, namely, soil parent material (56.6%), industrial waste and Mo smelting (24.0%), metal smelting and traffic emissions (12.8%), and coal combustion (6.7%). On this basis, the health risk of HMs was evaluated by Monte Carlo simulations and showed that a higher risk, with a higher proportion of exceeding-thresholds risk, was observed for children than for adults in terms of both noncarcinogenic and carcinogenic risks. Exposure pathways of oral ingestion in children could result in a higher attributed risk than other pathways. Furthermore, source-oriented risk assessment (SORA) revealed that the sources of coal combustion, industrial waste and Mo smelting had the highest contributions to noncarcinogenic and carcinogenic risks. Overall, for effective environmental management in agricultural soil, the framework of SORA was verified as an effective tool in the identification of the priority control of HMs and their sources.

Introduction

Heavy metal(loid) (HM) pollution in agricultural soil is of widespread concern due to the toxicity and bioaccumulation of HMs and their potential risk to crops and human health (Bhuiyan et al., 2021; Gui et al., 2023; Yang et al., 2018); this is especially true for the issue surrounding metal smelters because of the much higher risk faced by the soil than in other areas (Zhang et al., 2023a). High concentrations of soil HMs are frequently found near smelting sites, including As, Cd, Cu, Mn, Ni, Pb, and Zn (Haab et al., 2021; Li et al., 2022, 2023). The National Soil Pollution Survey Bulletin issued by the Chinese government indicated that 19.4% of farmland soil exceeded Chinese environmental quality standards, especially considering inorganic toxic metals in areas surrounding smelting enterprises and parks (Zhou et al., 2022a). HMs resulting from metal smelting can enter

H. Chen · S. Qiao · C. Li · Y. Wu · D. Li · L. Li · J. Liu (✉)
College of Resources and Environmental Sciences, Henan Agricultural University, Zhengzhou 450046, China
e-mail: liujianwei@henau.edu.cn

agricultural soil through wastewater discharge, atmospheric deposition and leaching from tailing residues (Li et al., 2023). Through inhalation, ingestion and skin contact or bioaccumulation into the food chain, these compounds can have various adverse effects on human health (Tan et al., 2023; Zhang et al., 2023a). It should be noted that, along with the metal smelting, soil HMs can also derive from soil-forming parent materials, atmospheric deposition, industrial waste, inputs of fertilizers and agrochemicals, and traffic emissions (Li et al., 2022; Ma et al., 2023). So, the key issue in metal smelting area is, which HMs are most polluted and which source ranks as the top risk contributor? Therefore, effectively identifying the sources and assessing the potential risk of HMs and different sources surrounding metal smelters is of vital importance to control the level of soil pollution (Li et al., 2023).

The existence of various polluting metallic elements due to the geological features and metallogenetic processes, complex production and processing processes of metallic elements (such as material preparation, smelting, fuming, and refining), and spatial transport behaviors of different HMs before soil entry make it challenging to accurately identify the sources of HMs in soil influenced by multimetal smelting activities. At present, multiple methods have been developed to trace the sources of HMs in soil, such as principal component analysis (Wang et al., 2020), positive matrix factorization (PMF) (Li et al., 2023), isotope fingerprinting (Wang et al., 2021a, 2021b), machine learning (Shi et al., 2022) and geographical detection (Sheng et al., 2023). Among these methods, isotope fingerprinting is high-cost and can only trace the source of a single or limited metal. Machine learning and geographical detectors require a relatively large quantity of data of different independent variables on the basis of a large sample size. Principal component analysis and the PMF model are both typical receptor source apportionment models, and the PMF model has the advantages of nonnegativity constraints on the factor matrixes, error estimations (bootstrap, displacement, and bootstrap enhanced by displacement), requirement of small sample size, no requirement for detailed emission profiles of specific sources prior to analysis, and the ability to obtain more accurate apportionment results (Jiang et al., 2017; Liu et al., 2018; Tan et al., 2023). Spatial interpolation analysis by a geographical information

system (GIS) and statistical analysis (e.g., cluster analysis and correlation analysis) are complements to source apportionment because some valuable information can be obtained, such as the correlation between different metals and the distribution of highly polluted areas and hotspots (He et al., 2023). Hence, an integrated method combining source apportionment models and statistical analysis is superior to a single method in identifying the sources of soil HMs (Tan et al., 2023).

Studies have shown that long-term exposure to HMs, particularly Cd, As, and Pb, increases the risk of various diseases and even carcinogenic effects, such as liver, kidney, nervous system, blood and bone diseases (Li et al., 2023; Tan et al., 2023). Most current studies associated with ecological and health risk assessment of soil HMs focus on the assessment of the elemental content but lack the consideration of the risk contributions of different pollution sources (He et al., 2023), which is detrimental to effective pollution prevention and risk reduction and safeguarding of the quality of regional agricultural products (Jiang et al., 2017). In contrast, the framework of source-oriented risk assessment (SORA), by combining health risk evaluation with source apportionment to resolve the risk contribution of different pollution sources, has been gradually recognized to facilitate the control of pollution sources (He et al., 2023; Li et al., 2023; Liu et al., 2018; Tan et al., 2023; Zhang et al., 2023a); however, the application of this framework in soil around metal-smelting industries is still limited.

Therefore, the main objectives of this study were (1) to analyze the characteristics of HM pollution in agricultural soil around a typical multimetal smelting area; (2) to identify the main sources of HMs by an integrated approach of the positive matrix factorization model, ordinary kriging interpolation and hierarchical clustering analysis (PMF-OK-HC); (3) and to further build the framework of source-oriented risk assessment (SORA) to identify the priority control sources through PMF-OK-HC and probabilistic health risk assessment by Monte Carlo simulations. Overall, this study provides an effective framework for the environmental management of HM-related pollution sources in multimetal smelting areas.

Materials and methods

Study area, sample collection and chemical analysis

The study area is located in western Henan Province, Central China, on the southern side of the Yellow River (Fig. 1). This area exhibits a continental monsoon semiarid climate, with an annual average temperature of 14.2 °C and an annual rainfall generally varying between 400 and 700 mm. The main crops grown include wheat, maize and peanuts, while apples and forsythia are well-known local farm products in terms of their planting area and production. Due to its rich mineral resources, this area has become a flourishing smelting base for lead (Pb), zinc (Zn), gold (Au), silver (Ag), copper (Cu), molybdenum (Mo), iron (Fe) and aluminum (Al) in China.

A total of 90 soil samples were collected at depths from 0 to 20 cm in the agricultural surface layer in October 2022. Soil samples were collected by mixing subsamples obtained at 5 points, placed in self-sealing polyethylene bags and returned to the lab. The locations of the sampling points were recorded by a global positioning system (GPS). All soil samples were naturally dried at room temperature and thoroughly mixed before sieving with 0.15-mm mesh. The soil samples were digested ($HCl-HNO_3-HF$), and the concentrations of metallic constituents were then determined via inductively coupled plasma−mass spectrometry (ICP−MS), including a total of 16 metallic constituents, namely, Cd, Pb, As, Cu, Cr, Zn, Ni, Al, Sc, Mn, Fe, Ge, Se, Rb, Sr and Mo. Quality control was performed by parallel samples and standard reference substances (GBW07401). The confidence level of the metals analyzed in the standards was 90%, the relative standard deviation of the repeated measurements was less than 10%, and the recoveries of the various HMs ranged from 82.5 to 118.3%. The instrumental detection limits of HMs are listed in Table S1.

Pollution assessment

Nemerow integrated pollution index

The Nemerow integrated pollution index (P_n) was used to reflect the pollution levels of HMs at each site (Nemerow, 1974; Pecina et al., 2021). This index highlights the impact of the major HMs and can be calculated as follows:

Fig. 1 Location of the sampling sites
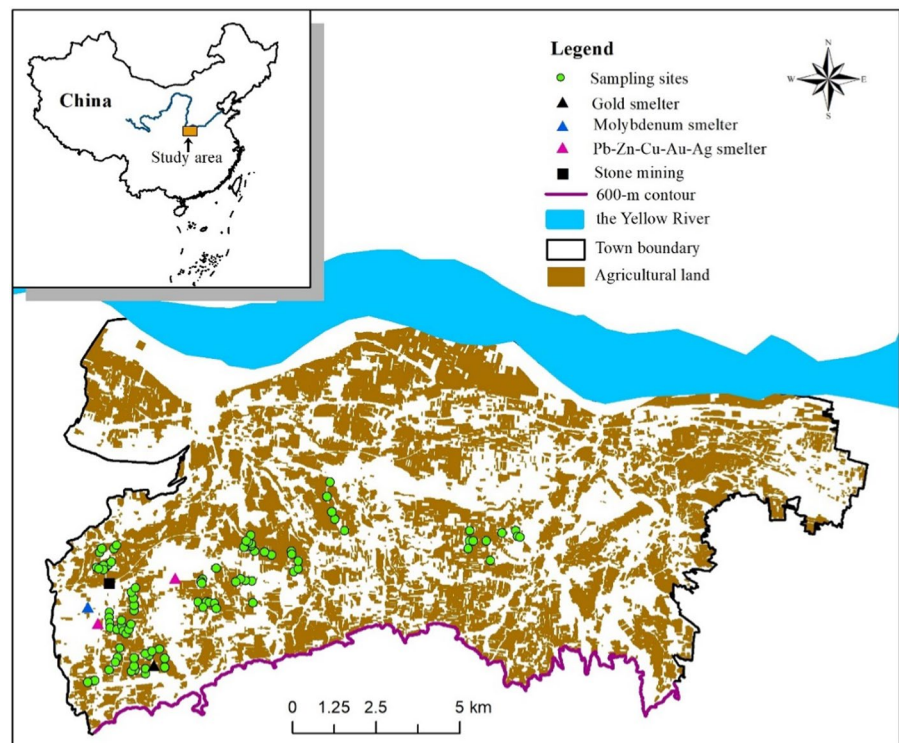
$$P_i = C_i / \text{BV}_i \tag{1}$$

$$P_n = \sqrt{\frac{(P_{\text{iave}})^2 + (P_{\text{imax}})^2}{2}} \tag{2}$$

where P_i is the single-factor pollution index of HM i; C_i is the measured concentration of HM i in the soil; BV is the background concentration of HM in the study area (Table 1); and P_{imax} and P_{iave} are the maximum and average values of P_i, respectively, for all the measured HMs in the soil. The pollution levels were divided into five classes according to the P_n value (Table S2).

Geoaccumulation index

The geoaccumulation index (I_{geo}) is a suitable indicator of the effects of natural variability and anthropogenic factors on the distribution of HMs (Muller et al., 1969). I_{geo} can be calculated as follows:

$$I_{\text{geo}} = \text{Log}_2 \left[\frac{C_i}{k \times \text{BV}_i} \right] \tag{3}$$

where C_i is the measured content of soil HM i (mg/kg); BV is the regional background value of a given HM (Table 1); and k is the correction factor, generally

chosen as 1.5. The pollution levels of the HMs were divided into seven classes according to the I_{geo} value (Table S3).

PMF source apportionment analysis

The PMF model is a typical receptor source apportionment model based on the least squares method (Jiang et al., 2017; Liu et al., 2018; Paatero & Tapper, 1994). In the PMF model, the original concentration matrix (E_{ij}) is decomposed into a factor profile matrix (A_{ik}), a contribution matrix (B_{kj}) and a residual error matrix (ε_{ij}). ε_{ij} can be obtained by minimizing the objective function Q, and in the case of a minimal Q value, the PMF model outputs the optimization result. The PMF model can be expressed as follows:

$$E_{ij} = \sum_{k=1}^{p} A_{ik} B_{kj} + \varepsilon_{ij} \tag{4}$$

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{m} \left(\frac{\varepsilon_{ij}}{\sigma_{ij}} \right)^2 \tag{5}$$

where E_{ij} is the concentration of the jth element in soil sample i; A_{ik} is the contribution of the kth source to soil sample i; B_{kj} is the concentration of the jth element in source k; σ_{ij} is the uncertainty in the jth

Table 1 Statistics of the HM concentration (mg/kg) and exceedance rate

BV denotes the background values of the HMs in the study area (HNDSS, 2004; CNEMC, 1990); exceedance rate 1 is based on the BV values as a reference standard; RSV denotes the risk screening values of HMs derived from the Soil Environmental Quality Risk Control Standard for Soil Contamination of Agricultural Land (GB 15618-2018); and exceedance rate 2 is based on the RSV as a reference standard. *CV* Coefficient of variation

HMs	Range	Mean	BV	CV (%)	Exceedance rate 1 (%)	RSV	Exceedance rate 2 (%)
Cr	62.2 ~ 109	74.2	64.4	11.5	92.2	250	0.0
Ni	25.3 ~ 55.0	30.0	28.4	12.3	66.7	190	0.0
Cu	25.2 ~ 578	101	22.2	112	100	100	23.3
Zn	65.6 ~ 449	110	87.6	51.5	60.0	300	2.2
As	6.9 ~ 64.6	25.5	12.4	76.7	67.8	25.0	32.2
Cd	0.2 ~ 7.1	1.3	0.078	82.1	100	0.6	76.7
Pb	29.0 ~ 3721	487	22.5	154	100	170	57.8
Al	16,787 ~ 57,139	32,796	6.34	33.5	100	–	–
Sc	12.8 ~ 26.5	18.4	11.63	12.3	100	–	–
Mn	569 ~ 1026	690	531	10.4	100	–	–
Fe	25,492 ~ 53,757	32,668	2.79	15.4	100	–	–
Ge	1.9 ~ 2.7	2.3	1.7	8.4	100	–	–
Se	5.6 ~ 42.6	21.3	0.201	41.9	100	–	–
Rb	69.0 ~ 128	95.3	94	10.3	54.4	–	–
Sr	101 ~ 306	163	160	23.5	43.3	–	–
Mo	0.2 ~ 6.6	1.1	0.78	84.1	60.0	–	–

element in sample i; n is the number of soil samples; and m is the number of chemical elements.

When the concentration of chemical elements is lower than or equal to their corresponding method detection limits (MDLs; refer to Table S1), the uncertainty can be calculated as:

$$u_{nc} = 5/6\,MDL \tag{6}$$

Conversely, it can be calculated as:

$$u_{nc} = \sqrt{(\sigma \times c)^2 + MDL^2} \tag{7}$$

where σ is the relative standard deviation and c is the elemental concentration.

The concentrations of 16 elements and their uncertainty data were substituted into the PMF model. The total mass of the metals was defined as the total variable (SUM). The fitting performance for each element was evaluated by the combination of categories, signal-to-noise ratio (S/N), and coefficient of determination (r^2) (Table S4), and many samples exhibited residuals beyond $+3$ and -3, while the species of Ni and Ge were categorized as weak species (Table S4). The number of factors ranging from 3 to 6 was examined by checking the Q value and the residual analysis, and then a total of 200 base model runs were employed under different numbers of factors. For each PMF model employed, the lowest Q value (robust) was chosen as the optimal solution. Afterward, fpeak analysis of this solution was performed to evaluate the rotational ambiguity (Table S5). Model uncertainties were estimated by the displacement and bootstrap techniques (Table S6, Table S7). Finally, a 4-factor analysis model yielded an invariably optimal solution.

Probabilistic health risk assessment

The USEPA risk assessment model (USEPA, 2011) was used to quantify the carcinogenic and noncarcinogenic risks of the soil HMs. Regarding noncarcinogenic risks, the HMs include As, Cd, Cu, Cr, Ni, Pb, Zn and Mn, while regarding carcinogenic risks, the HMs include As, Cd, Cr, Ni and Pb. Different exposure pathways of oral ingestion of soil particles, inhalation of soil particles and dermal contact with soil were considered for the adult and child populations in health risk assessment.

The noncarcinogenic risk can be obtained as:

$$ADD_{ing} = \frac{C \times R_{ing} \times EF \times ED}{BW \times AT} \times 10^{-6} \tag{8}$$

$$ADD_{inh} = \frac{C \times R_{inh} \times EF \times ED}{PER \times BW \times AT} \tag{9}$$

$$ADD_{derm} = \frac{C \times SA \times AF \times ABS \times EF \times ED}{BW \times AT} \times 10^{-6} \tag{10}$$

$$HQ_i = ADD_i / RfD_i \tag{11}$$

$$HI = \sum HQ_i \tag{12}$$

where ADD_{ing}, ADD_{inh}, and ADD_{derm} denote the average daily exposure of HMs via the ingestion, inhalation, and dermal contact pathways, respectively, mg/(kg-d); C denotes the content of HMs (mg/kg); EF is the exposure frequency, d/a; ED is the exposure duration (year); BW is the body weight (kg); AT is the average exposure time (day); HQ_i is the noncarcinogenic risk for HM i; and RfD is the reference dose mg/(kg-d) of HM. Moreover, HI is the sum of the HQ values for all HMs. $HQ \leq 1$ or $HI \leq 1$ indicates a low noncarcinogenic risk, while $HQ > 1$ or $HI > 1$ indicates a high noncarcinogenic risk and requires attention (NHCPRC, 2021).

The carcinogenic risk can be obtained as:

$$LADD_{ing} = \frac{C \times R_{ing} \times EF \times ED}{BW \times LT} \times 10^{-6} \tag{13}$$

$$LADD_{inh} = \frac{C \times R_{inh} \times EF \times ED}{PER \times BW \times LT} \tag{14}$$

$$LADD_{derm} = \frac{C \times SA \times AF \times ABS \times EF \times ED}{BW \times LT} \times 10^{-6} \tag{15}$$

$$CR_i = LADD_i \times SF_i \tag{16}$$

$$TCR = \sum CR_i \tag{17}$$

where LADD is the lifetime average daily exposure dose, mg/(kg-d); LT is the lifetime exposure time; SF_i is the slope factor (kg-d/mg); CR is the carcinogenic risk of a given HM; and TCR is the sum of the CR values for all HMs. The other parameters

were the same as those in the above noncarcinogenic risk assessment. Generally, a CR or TCR lower than 1×10^{-6} indicates that there is no obvious carcinogenic risk, a CR or TCR between 1×10^{-6} and 1×10^{-4} indicates that exposure to HMs may cause negative effects on human health, and a CR or TCR higher than 1×10^{-4} indicates a potentially severe cancer risk that requires caution (NHCPRC, 2021).

Probabilistic health risks were determined by Monte Carlo simulations (Oracle® CrystalBall (version 11.1.2.4)). The descriptions and values of the above parameters are listed in Table S8. The parameters of the chronic reference dose (RfD, mg/(kg-d)) and slope factor (SF, (kg-d)/mg) for each HM are given in Table S9. The parameters were obtained via 10,000 Monte Carlo simulations to ensure robust results.

Data analysis

Statistical analysis was performed in SPSS 22.0 (IBM Corp. Armonk, NY). Origin 9.1 software (OriginLab, Northampton, MA) was used to generate graphs. A hierarchical clustering heatmap was analysed by the "pheatmap" package in R studio. Maps of the study area and interpolation analysis (ordinary kriging) of the HM concentrations were performed in ArcGIS 10.2 software.

Results and discussion

Concentrations of HMs in agricultural soil

The mean concentrations of Cr, Ni, Cu, Zn, As, Cd and Pb were 74.2 mg/kg, 30.0 mg/kg, 100.6 mg/kg, 110.3 mg/kg, 25.5 mg/kg, 1.3 mg/kg and 486.5 mg/kg, respectively (Table 1). Compared to the BVs, the exceedance rates for most metals were high (even reaching 100% for some metals), including 16.7 and 21.6 times higher rates than the corresponding BV values for Cd and Pb, respectively, and even thousands of times higher rates than the corresponding BV values for Fe and Al. This occurs because the study area is a high-background metal area. Compared to the screening values of agricultural land, the exceedance rate of Cd was the highest, at 76.7%. The exceedance rate of Pb was also high, at 57.8%, while the exceedance rates of As, Cu and Zn were lower,

at 32.2%, 23.3%, and 2.2%, respectively. The contents of Cu and Pb greatly fluctuated, with coefficients of variation (CVs) of 112% and 154%, respectively, while Cr and Ni showed the least fluctuation, with CVs lower than 15%. The high CV values indicate high spatial heterogeneity and verify that some of the sampling points may have been notably influenced by pollution sources. In previous studies (Zhou et al., 2023), relatively high CV values of Cd, As, Pb, Cu and Zn were found and verified that these HMs were strongly influenced by external inputs such as mining activities, vehicle exhaust, smelting waste, pesticides and fertilizers.

Evaluation of HM pollution

The P_n index value at the different points was 15.7 ± 21.7 (mean $\pm$ standard deviation), reaching the heavy pollution level overall (Fig. 2a) because the study area is rich in metallic mineral resources and has high BVs for the metals. I_{geo} (Fig. 2b) showed that 97.8% of the soil samples did not exhibit Cr and Mn pollution ($I_{geo} \leq 0$), 15.6% and 2.1% of the soil samples exhibited no to moderate Zn and As pollution ($0 < I_{geo} \leq 1$), 40% and 20% of the soil samples were heavily polluted by Cd and Pb ($3 < I_{geo} \leq 4$), respectively, and 5.2% and 11.5% of the soil samples were extremely polluted by Cd and Pb ($I_{geo} > 5$), respectively. Li et al. (2023) found that long-term nonferrous industrial activities resulted in high I_{geo} values (2–8) of As, Cd and Pb in surrounding agricultural and woodland soils in Southwest China (Li et al., 2023).

Source apportionment of HMs

As shown in Fig. 3 (red marks), some elements were significantly correlated ($p < 0.01$), such as the groups of Cu–Pb−Zn and Al–Sc–Cr–Mn–Fe–Ni–Rb−Sr; the higher the correlation, the stronger the homology. Furthermore, the HC analysis in Fig. 3 shows that these metals can be divided into 3–5 categories, indicating a potential common source of the metals in the same category. The study area contains abundant metallic mineral resources and densely distributed nonferrous metal smelters. Different types of metallic smelters have various processing procedures, pollutant emissions, and discharge paths. Before HMs migrate into soil, their transportation characteristics
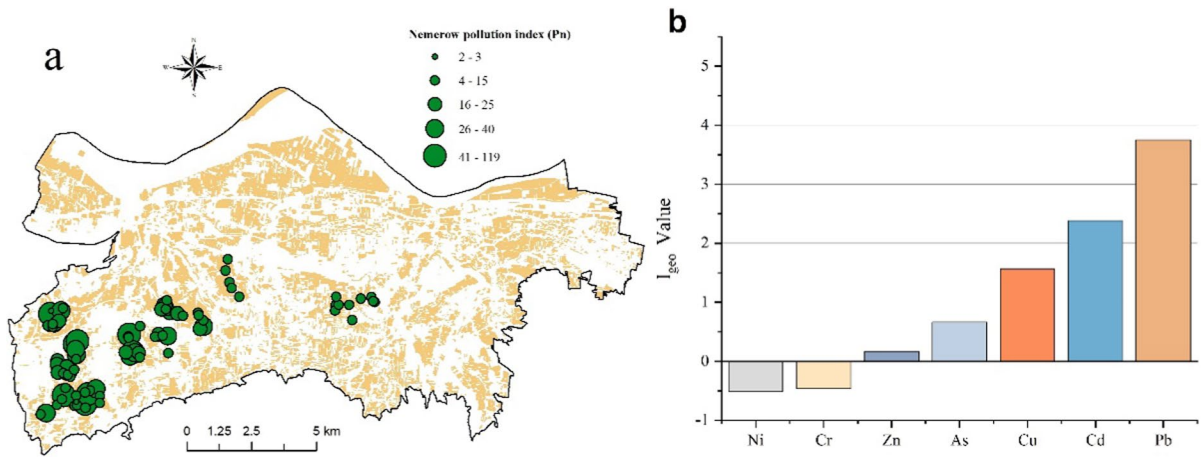
Fig. 2 Spatial distribution of the Nemerow pollution index (**a**) and geoaccumulation index of HMs (**b**)
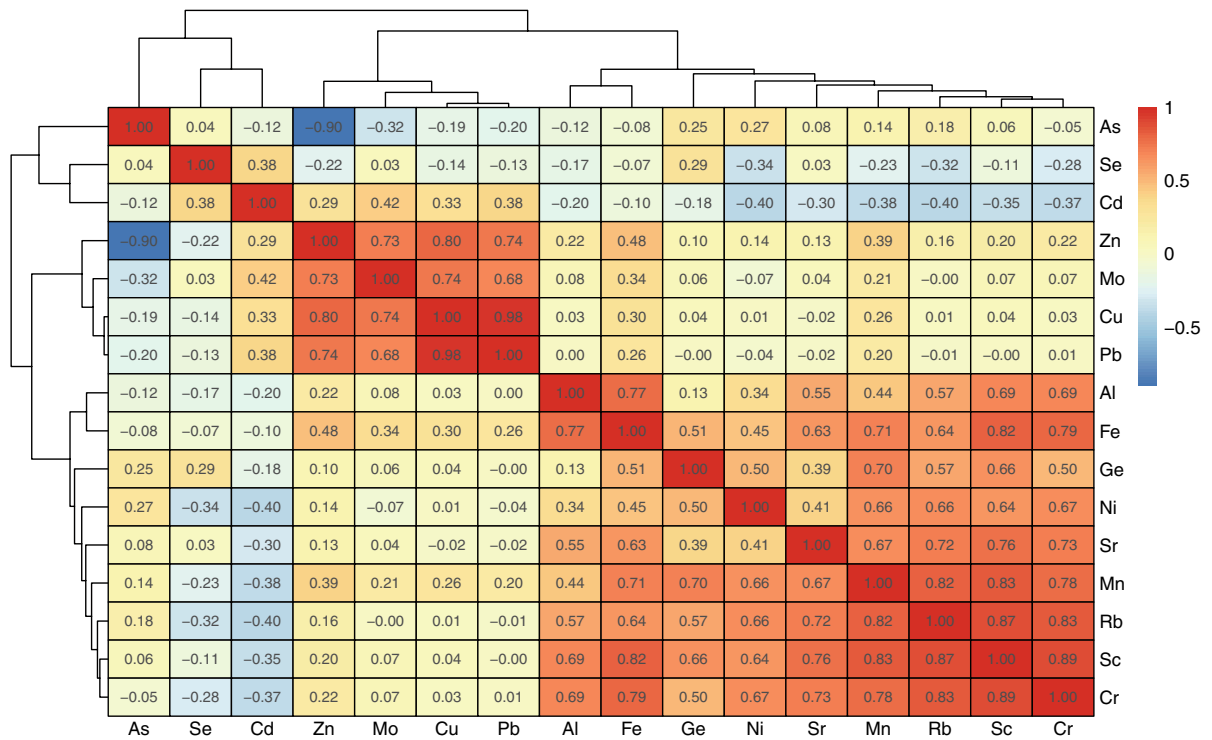

Fig. 3 Hierarchical clustering heatmap of Pearson correlation coefficients of the metals

may also be influenced by regional topographical conditions and meteorological conditions. In contrast, considering the associated nature of intermetals from polymetallic ores and the similar transport pathways before entering soil (Huang et al., 2023), auxiliary information on the type and distribution of industrial enterprises, topographical and geomorphological conditions, and meteorological conditions were also introduced into PMF analysis. Finally, four sources were identified and are shown as follows.

Factor 1 exhibited a notably high contribution of As (63.8%) and minor contents of other elements (lower than 20%). As is a typical tracer of coal combustion (Liu et al., 2018), and a large amount of coal is required as power during the smelting of minerals, of which the process may release substantial amounts of As (Zhang et al., 2023c). The study area is located in the valley of Yellow River (China) with an altitude of approximately 300 m, while the elevation increases to more than 2000 m on both sides of the valley. The prevailing winds in the area are predominantly in the directions of east and west (Fig. S2). Therefore, atmospheric deposition of As from coal combustion in smelting industries was speculated to be the major source of soil As (Zhang et al., 2023b).

Factor 2 exhibited a relatively high contribution to several metals (ranging from 45 to 65%), including Al, Fe, Mn, Sc, Cr, Ni, Rb, Sr, and Ge, except for HMs with a strong influence originating from anthropogenic and industrial activities, such as Cd, Cu, Pb, Mo and As. Fe and Al are widely present in the Earth's crust and are the main elements in soil dust (Xiao et al., 2023), which is a natural source of soil-forming parent material. Cr, Mn, and Ni are often used as indicators of natural sources of soil HMs (Li et al., 2023; Liang et al., 2017). Considering that the study area has high geological BVs for these metals and that the spatial distribution of these metals is relatively homogeneous, a smaller influence of regional industrial sources was identified. Therefore, this factor could be interpreted as a soil parent material source.

Factor 3 was characterized by high loadings of Pb (75.0%), Cu (60.5%), Mo (37.1%) and Zn (30%). Cu and Pb in soils are strongly influenced by anthropogenic sources such as smelter dust deposits, slag and tailings (He et al., 2023; Zhou et al., 2022a). According to the survey, the proven reserves of Cu and Pb metal ores in the area are high, distributed across several large smelters. From the pollution index of I_{geo} (Fig. 2b), these metals were deeply influenced by anthropogenic pollution. The spatial distribution (Fig. S1) also showed highly polluted areas surrounding the smelters. Due to the past use of leaded gasoline, tire wear and lead-containing antidetonants, lubricants and engines, Cu, Pb and Zn have become indicators of emissions from transportation (Jiang et al., 2017; Liu et al., 2023a; Wang et al., 2020). Considering that the main road (G310) is inside the

study area, traffic emissions may also be one of the main sources of these metals. Therefore, this factor could be identified as a source of metal smelting and traffic emissions.

Factor 4 exhibited a relatively high contribution of Cd (54.9%), Se (48.2%), and Mo (38.7%). Cd was spatially concentrated around the smelters. Cd can be released as tailing waste due to primitive associated crude smelting processes and inefficient recovery methods, followed by leaching, release and eventual deposition in nearby soil (He et al., 2023; Li et al., 2022). The HMs contained in dry smelting slag are unstable and can easily be leached, causing pollution. In addition, accounting for the main road (G310) for raw material transportation, it is reasonable that spills during transportation could lead to raw materials or waste residues entering nearby agricultural soil under the effect of dust. The study area also contains abundant molybdenum ore deposits. From the regional survey, a large molybdenum smelter (*Zhong-Shan molybdenum smelter*) is located on the western side of the sampling area, with an annual production of approximately 30,000 tons of molybdenum oxide and 20,000 tons of ferromolybdenum. The spatial distribution of the Mo concentrations (Fig. S1) showed that high values are located around this smelter. Therefore, this factor could be explained as industrial waste and Mo smelting (Fig. 4).

Health risk assessment of HMs

The noncarcinogenic and carcinogenic risks of HMs for different groups (adults and children) were estimated by health risk assessment and Monte Carlo simulations associated with three exposure pathways, namely, direct oral ingestion, inhalation of suspended soil particles and dermal exposure (Fig. 5 and Table 2). Considering the risk attributed to the different exposure pathways (Fig. S3), oral ingestion was the prominent pathway over inhalation and dermal contact for both the noncarcinogenic and carcinogenic risks, which is consistent with previous studies conducted near smelting sites (He et al., 2023; Zhou et al., 2022a).

In terms of the noncarcinogenic risk (Fig. 5 and Table 2), the HI values (mean) for adults and children were 1.51 and 1.85, respectively, and the probability of exceeding the threshold (HI > 1) reached 12.0% and 52.9%, respectively. The noncarcinogenic risks
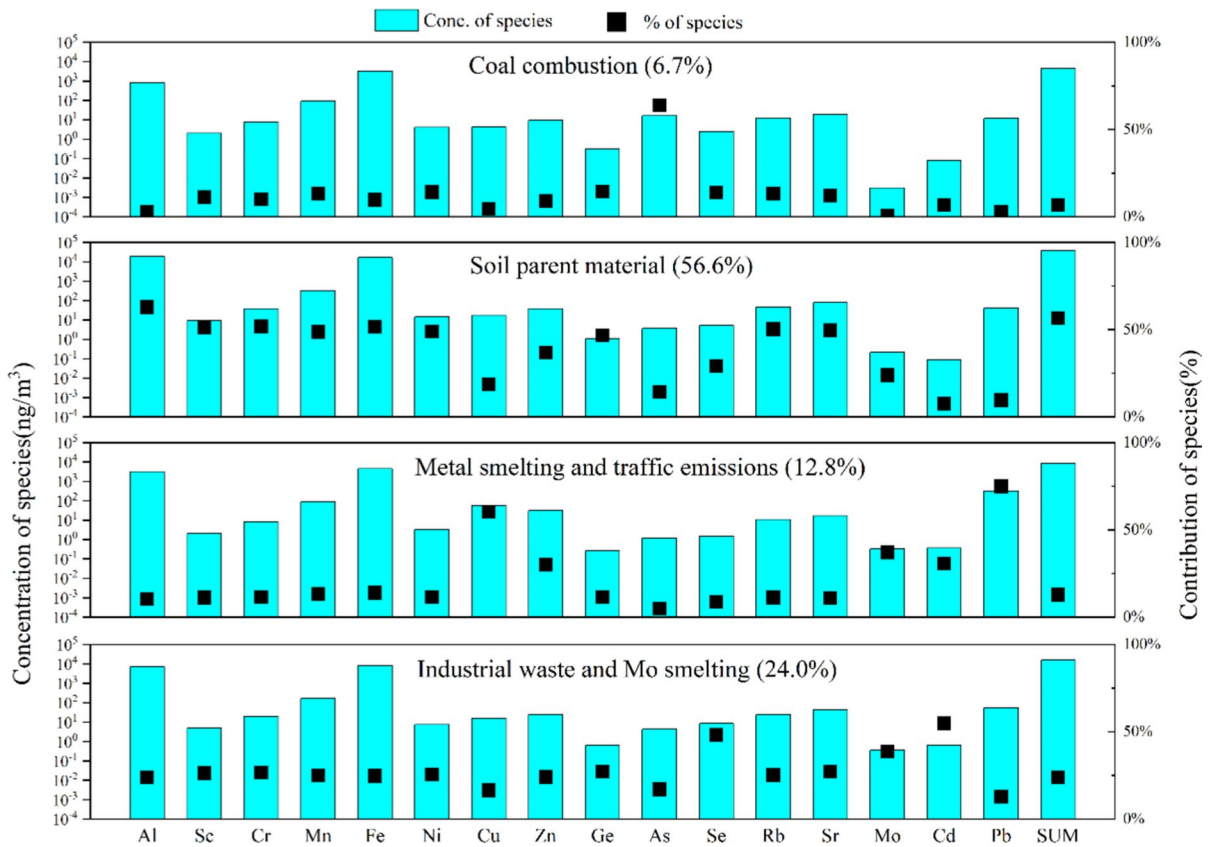
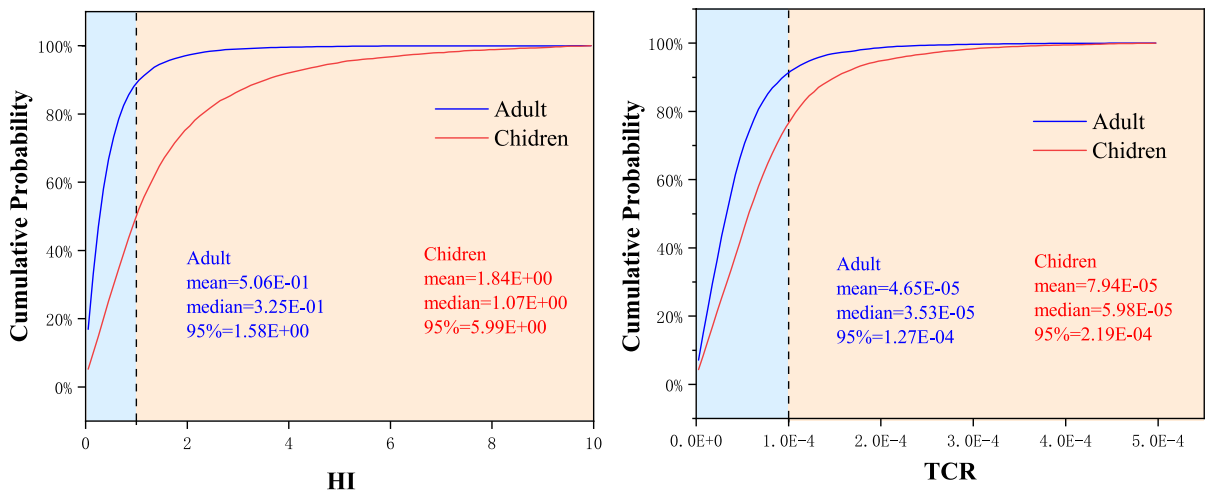**Fig. 4** Contribution spectrum of the different sources and corresponding contribution

Fig. 5 Probability distribution of the noncarcinogenic risk (HI) and carcinogenic risk (TCR) of HMs

Table 2 Noncarcinogenic and carcinogenic risk of HMs based on Monte Carlo simulations

Risk	HMs	Mean		SD		95%CI	
		Adult	Children	Adult	Children	Adult	Children
HQ	As	6.35E−02	5.52E−01	1.17E−03	9.38E−03	2.13E−01	1.80E+00
	Cd	7.74E−04	3.77E−02	9.22E−06	8.95E−04	2.48E−03	1.35E−01
	Cr(VI)	1.06E−E−02	5.94E−02	1.23E−04	8.72E−04	3.19E−02	1.84E−01
	Cu	1.92E−03	1.63E−02	5.40E−05	3.98E−04	6.33E−03	5.27E−E−02
	Mn	2.99E−01	1.65E−01	3.99E−03	1.92E−03	9.73E−01	4.50E−01
	Ni	1.10E−03	9.27E−03	7.98E−06	6.35E−05	2.55E−03	2.01E−02
	Pb	1.29E−01	1.00E+00	3.39E−03	2.64E−02	5.02E−01	3.85E+00
	Zn	2.89E−04	2.34E−03	2.67E−06	2.10E−05	7.51E−04	5.89E−03
HI	Total	5.06E−01	1.84E+00	6.30E−03	3.01E−02	1.58E+00	5.99E+00
CR	As	9.80E−06	2.13E−05	1.80E−07	3.62E−07	3.29E−05	6.96E−05
	Cd	1.56E−06	3.73E−06	1.87E−08	4.27E−08	5.02E−06	1.14E−05
	Cr(VI)	3.53E−06	4.19E−06	4.82E−08	8.55E−08	1.20E−05	1.60E−05
	Ni	3.08E−05	4.82E−05	3.19E−07	5.71E−07	8.68E−05	1.31E−04
	Pb	8.41E−07	2.01E−06	2.23E−08	5.53E−08	3.26E−06	7.62E−06
TCR	Total	4.65E−05	7.94E−05	4.57E−07	8.44E−07	1.27E−04	2.19E−04

SD standard deviation, *CI* confidential interval, *HQ* hazard quotient, *HI* hazard index, *CR* carcinogenic risk, *TCR* total carcinogenic risk

of HMs for adults and children could be ranked as Pb > As > Mn > Cr(VI) > Cd > Cu > Ni > Zn, with Pb and As dominating the noncarcinogenic risk. The 5th percentile value is usually considered the lowest hazard, while the 95th percentile value is considered the worst scenario. The 95th HI values notably exceeded the threshold (HI > 1), with values of 1.58 and 5.99 for adults and children, respectively, indicating that the HMs in the study area posed a notable noncarcinogenic risk that cannot be ignored.

In terms of the carcinogenic risk (Fig. 5 and Table 2), the TCR values for adults and children were 4.65E−05 and 7.94E−05, respectively, with probabilities of exceeding the thresholds of 9.2% and 24.9%, respectively, while the carcinogenic risks of HMs for adults and children could be ranked as Ni > As > Cr(VI) > Cd > Pb. The 95th percentile TCR values for adults and children were 1.27E−04 and 2.19E−04, respectively, exceeding the thresholds (TCR > 1.0E−04), with Ni and As posing the highest carcinogenic risk, which is a major concern.

Both the noncarcinogenic and carcinogenic risks were significantly higher in children than in adults, which is consistent with previous studies (Gorka et al., 2022; Liu et al., 2023b; Yang et al., 2019; Zhou et al., 2022a); in addition, children are more vulnerable to toxic substances than adults because their hand-to-mouth behavior, physiological characteristics and exposure times are different from those of adults, while children possess higher intake rates, exposure frequencies and smaller body sizes and are more likely to be exposed to HMs in contaminated soil. Therefore, it is particularly important to give attention to the oral ingestion behavior of children to protect their health (Gorka et al., 2022).

Source-oriented health risks

Source-oriented risks were assessed by combining health risk assessment and source apportionment to identify the sources that dominate the health risk of HMs for priority control (He et al., 2023). Because the trend of the risk for children was the same as that for adults, only the results for the adult population are described. As shown in Fig. 6a, the noncarcinogenic risks of the four sources and the contributions followed the decreasing order of coal combustion (36.4%), industrial waste and Mo smelting (30.5%), metal smelting and traffic emissions (17.7%), and soil parent material (15.4%). Considering the contribution of different metals, As in coal combustion was the top contributor, with contributions reaching 33.0%, and Cd in industrial waste and Mo smelting also contributed a risk of 18.5%. Regarding the carcinogenic risk (Fig. 6b), the decreasing order of the sources was industrial waste and Mo smelting (42.0%), metal smelting and traffic emissions (22.1%), soil parent material (18.7%), and coal combustion (17.2%).
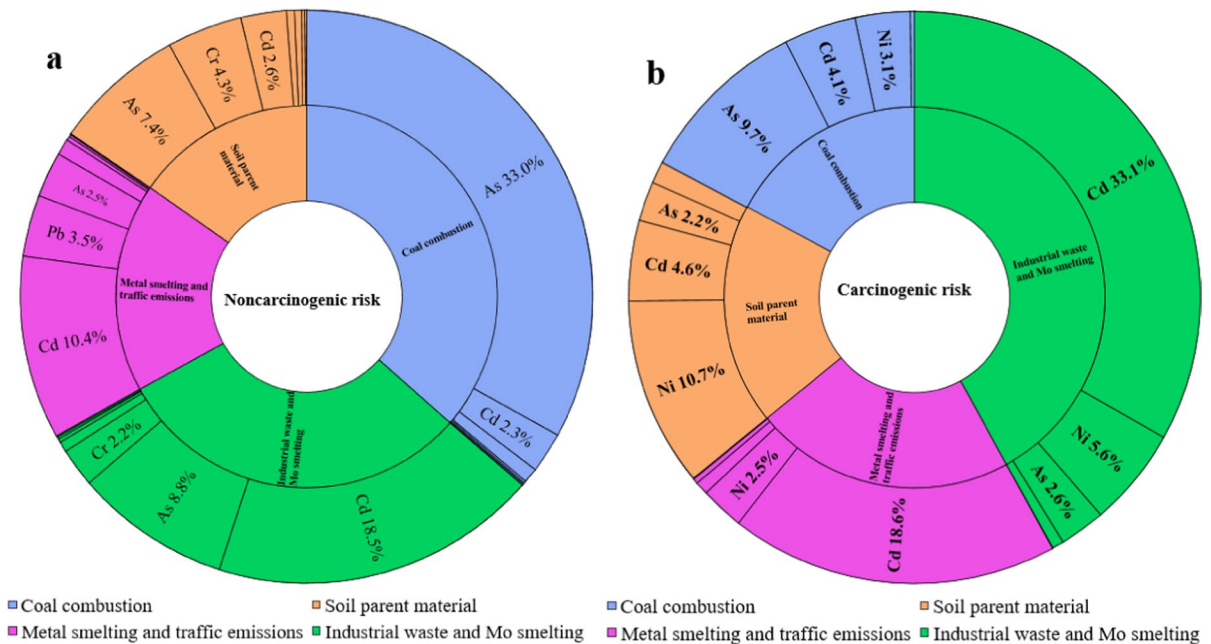
Fig. 6 Source-oriented health risks of the different sources

Furthermore, Cd emission from the source of industrial waste and Mo smelting was the dominant risk contributor of 33.1%. In addition, Cd from the source of metal smelting and traffic contributed 18.6% of the risk. Therefore, with the aim of environmental management, As in coal combustion (Wang et al., 2021a, 2021b; Yan et al., 2022) and Cd from metal smelting-related emissions (Micó et al., 2006; Yang et al., 2019) should be the most critical objectives to control in agricultural soils in the study area.

However, there are still some limitations of this study, such as the behavior and fate of metals during the smelting processes and the lack of information related to other influencing variables (physical and chemical properties of soil, emission feathers of smelting industries, agricultural and inhabits activities). Apart from the integrated approach (i.e., PMF-OK-CH) used in this study, there are some robust auxiliary solutions. Stable isotope and multi-isotope tracing techniques, such as Pb isotopes ([204]Pb, [206]Pb, [207]Pb and [208]Pb), are widely used isotope tracing elements in soil analysis (Cheema et al., 2020). Some rare element (e.g., [203]Tl/[205]Tl) isotopes can be used to determine the behavior of metals at different stages of smelting processes (Vaněk et al., 2018; Zhou et al., 2022b); Pb–Sr isotope compositions have been used

to analyze the sources of metal pollution in urban soils (Sun et al., 2018); and Sr–Nd–Pb isotopes have been employed to trace sediment sources (Deng et al., 2021; Wu et al., 2022). In addition, machine learning models provide the advantages of managing the complex nonlinear relationships between soil HMs and environmental variables and identifying the key influencing factors, which supplies valuable information in source apportioning (Ru et al., 2016; Shi et al., 2022; Zhang et al., 2021). For example, Shi et al. (2022) used a conditional inference tree algorithm and revealed that the influence of altitude on soil Cr, Cu, Hg, Ni and Zn, as well as of soil pH on Cd, indicated their primary origin of natural processes. By applying a random forest algorithm to identify the key role of environmental variables, Yang et al. (2021) revealed that the distance to the nearest industrial site was the most important factor (accounting for the largest proportion of the total variation) in determining Cd and Cu concentrations, which indicated that atmospheric deposition of industrial waste smoke and dust is an obvious potential source of Cd and Cu in surrounding agricultural soils, while the distance to the nearest road accounted for most of the variations in the concentrations of both Pb and Zn, suggesting that road traffic is an important source of both metals.

In future research, the PMF model will be combined with isotope tracer technology and machine learning to further trace the pollution sources of soil more accurately.

Conclusion

HMs in agricultural soils around a nearby multimetal smelting area were systematically analyzed and it was determined that Cd and Pb were heavily polluted due to their high exceedance rates of screening values of agricultural land. Based on an integrated approach of PMF-OK-CH, the sources of HMs and their contributions were identified as soil parent material (56.6%), industrial waste and Mo smelting (24.0%), metal smelting and traffic emissions (12.8%), and coal combustion (6.7%). The probabilistic health risk assessment showed that both noncarcinogenic and carcinogenic risks were higher for children than for adults, with the proportion of values exceeding the thresholds (HI > 1 and TCR > 1.0E−04) in children vs. adults reaching 52.9% versus 12.0% for the noncarcinogenic risk and 24.9% versus 9.2% for the carcinogenic risk. Then, the framework of source-oriented risk assessment (SORA) was built and showed that coal combustion dominated the noncarcinogenic risk (36.4%), while industrial waste and Mo smelting contributed the most to the carcinogenic risk (42.0%). Overall, Cd and As and their related sources should be treated as a priority in this multimetal smelting area.

Declarations

References

Bhuiyan, H. M. A., Shamal, C. K., Md, B. D., Md, A. R., & Bidyu, B. S. (2021). Enrichment, sources and ecological risk mapping of heavy metals in agricultural soils of dhaka district employing SOM, PMF and GIS methods. *Chemosphere, 263*, 128339.

Cheema, A. I., Liu, G., Yousaf, B., et al. (2020). A comprehensive review of biogeochemical distribution and fractionation of lead isotopes for source tracing in distinct interactive environmental compartments. *Science of the Total Environment, 719*, 135658.

CNEMC, China National Environmental Monitoring Centre (1990) Background concentrations of elements in soils of China. Chinese Environment Science Press, Beijing

Deng, F. F., Hellmann, S., Zimmermann, T., et al. (2021). Using Sr-Nd-Pb isotope systems to trace sources of sediment and trace metals to the Weser River system (Germany) and assessment of input to the North Sea. *Science of the Total Environment, 791*, 148127.

Gorka, R., Kumar, R., Yadav, S., et al. (2022). Health implications, distribution and source apportionment of heavy metals in road deposited dust of Jammu City in northern India. *Chemosphere, 308*, 136475.

Gui, H., Yang, Q. C., Lu, X. Y., et al. (2023). Spatial distribution, contamination characteristics and ecological-health risk assessment of toxic heavy metals in soils near a smelting area. *Environmental Research, 222*, 115328.

Haab, C., Pm, D., Cs, A., et al. (2021). Enzyme activities and microbial functional diversity in metal(loid) contaminated soils near to a copper smelter. *Science of the Total Environment, 779*, 146423.

He, J., Peng, Z. H., Zeng, J. Q., et al. (2023). Source apportionment and quantitative risk assessment of heavy metals at an abandoned zinc smelting site based on GIS and PMF models. *Journal of Environmental Management., 336*, 117565.

HNDSS, Department of Soil Survey of Henan Province. (2004). *Henan Soil*. China Agricultural Press.

Huang, C. Y., Guo, Z. H., Li, T. S., et al. (2023). Source identification and migration fate of metal(loid)s in soil and groundwater from an abandoned Pb/Zn mine. *Science of the Total Environment, 895*, 165037.

Jiang, Y. X., Chao, S. H., Liu, J. W., et al. (2017). Source apportionment and health risk assessment of heavy metals in soil for a township in Jiangsu Province. *China. Chemosphere., 168*, 1658–1668.

Li, X. X., Liu, H. Y., Meng, W., et al. (2022). Accumulation and source apportionment of heavy metal(loid)s in agricultural soils based on GIS, SOM and PMF: A case study in superposition areas of geochemical anomalies and zinc smelting, Southwest China. *Process Safety and Environmental Protection, 159*, 964–977.

Li, H., Yao, J., Sunahara, G., et al. (2023). Quantifying ecological and human health risks of metal(loid)s pollution from non-ferrous metal mining and smelting activities in Southwest China. *Science of the Total Environment, 873*, 162364.

Liang, J., Feng, C., Zeng, G., et al. (2017). Spatial distribution and source identification of heavy metals in surface soils in a typical coal mine city, Lianyuan, China. *Environmental Pollution, 225*, 681–690.

Liu, J. W., Chen, Y. J., Chao, S. H., et al. (2018). Emission control priority of $PM_{2.5}$-bound heavy metals in different seasons: A comprehensive analysis from health risk perspective. *Science of the Total Environment, 644*, 20–30.

Liu, F. T., Wang, X. Q., Dai, S., et al. (2023a). Spatial variations, health risk assessment, and source apportionment of soil heavy metals in the middle Yellow River Basin of northern China. *Journal of Geochemical Exploration, 252*, 107275.

Liu, Z., Du, Q. Q., Guan, Q. Y., et al. (2023b). A Monte Carlo simulation-based health risk assessment of heavy metals in soils of an oasis agricultural region in northwest China. *Science of the Total Environment, 857*, 159543.

Ma, J. W., Chen, L., Chen, H. S., et al. (2023). Spatial distribution, sources, and risk assessment of potentially toxic elements in cultivated soils using isotopic tracing techniques and Monte Carlo simulation. *Ecotoxicology and Environmental Safety, 259*, 115044.

Micó, C., Recatalá, L., Peris, M., et al. (2006). Assessing heavy metal sources in agricultural soils of an European Mediterranean area by multivariate analysis. *Chemosphere, 65*, 863–872.

Muller, G. (1969). Index of geoaccumulation in sediments of the rhine river. *GeoJournal, 2*, 109–118.

Nemerow, N.L. (1974). *Scientific stream pollution analysis*. McGraw-Hill Book Company.

NHCPRC. Health Industry Standard of the People's Republic of China: Technical Guide for Environmental Health Risk Assessment of Chemical Substances WS/T777–2021.

Paatero, P., & Tapper, U. (1994). Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics, 5*, 111–126.

Pecina, V., Brtnický, M., Baltazár, T., et al. (2021). Human health and ecological risk assessment of trace elements in urban soils of 101 cities in China: A meta-analysis. *Chemosphere, 267*, 129215.

Ru, F., Yin, A. J., Jin, J. X., et al. (2016). Prediction of cadmium enrichment in reclaimed coastal soils by classification and regression tree. *Estuarine, Coastal and Shelf Science, 177*, 1–7.

Sheng, Y., Wang, Z. Q., & Feng, X. (2023). Potential ecological risk and zoning control strategies for heavy metals in soils surrounding core water sources: A case study from Danjiangkou Reservoir, China. *Ecotoxicology and Environmental Safety, 252*, 114610.

Shi, T. R., Zhang, J. R., Shen, W. J., et al. (2022). Machine learning can identify the sources of heavy metals in agricultural soil: A case study in northern Guangdong Province, China. *Ecotoxicology and Environmental Safety, 245*, 114107.

Sun, J. W., Yu, R. L., Hu, G. R., et al. (2018). Tracing of heavy metal sources and mobility in a soil depth profile via isotopic variation of Pb and Sr. *CATENA, 171*, 440–449.

Tan, C., Wang, H., Yang, Q. C., et al. (2023). An integrated approach for quantifying source apportionment and source-oriented health risk of heavy metals in soils near an old industrial area. *Environmental Pollution, 323*, 121271.

USEPA (2011) Exposure Factors Handbook, Final. U.S. Environment Protection Agency, Washington, DC

Vanek, A., Grosslova, Z., Mihaljevic, M., et al. (2018). Thallium isotopes in metallurgical wastes/contaminated soils: a novel tool to trace metal source and behavior. *Journal of Hazardous Materials, 343*, 78–85.

Wang, Y., Zhang, L., Wang, J., et al. (2020). Identifying quantitative sources and spatial distributions of potentially toxic elements in soils by using three receptor models and sequential indicator simulation. *Chemosphere, 242*, 125266.

Wang, L. W., Jin, Y. L., Weiss, D. J., et al. (2021a). Possible application of stable isotope compositions for the identification of metal sources in soil. *Journal of Hazardous Materials, 407*, 124812.

Wang, Y., Yu, J. L., Wang, Z. H., et al. (2021b). A review on arsenic removal from coal combustion: Advances, challenges and opportunities. *Chemical Engineering Journal, 414*, 128785.

Wu, W. H., Qu, S. Y., Nel, W., et al. (2022). Tracing and quantifying the sources of heavy metals in the upper and middle reaches of the Pearl River Basin: New insights from Sr–Nd–Pb multi-isotopic systems. *Chemosphere, 288*(3), 132630.

Xiao, K., Yao, X. Q., Ren, X. C., et al. (2023). Transmission characteristics, source apportionment and health risk assessment of PM2.5 in a typical industrial city in Northwest China. *Urban Climate, 47*, 101367.

Yan, Q., Kong, S. F., Yan, Y. Y., et al. (2022). Emission and spatialized health risks for trace elements from domestic coal burning in China. *Environment International, 158*, 107001.

Yang, Q., Li, Z. Y., Lu, X. N., et al. (2018). A review of soil heavy metal pollution from industrial and agricultural regions in China: Pollution and risk assessment. *Science of the Total Environment, 642*, 690–700.

Yang, S. Y., Zhao, J., et al. (2019). Status assessment and probabilistic health risk modeling of metals accumulation in agriculture soils across China: A synthesis. *Environment International, 128*, 165–174.

Yang, S. Y., Taylor, D., Yang, D., et al. (2021). A synthesis framework using machine learning and spatial bivariate analysis to identify drivers and hotspots of heavy metal pollution of agricultural soils. *Environmental Pollution, 287*, 117611.

Zhang, H., Yin, A. J., Yang, X. H., et al. (2021). Use of machine-learning and receptor models for prediction of

and source apportionment of heavy metals in coastal reclaimed soils. *Ecological Indicators, 122*, 107233.

Zhang, X. H., Tian, K., Wang, Y. M., et al. (2023a). Identification of sources and their potential health risk of potential toxic elements in soils from a mercury-thallium polymetallic mining area in Southwest China: Insight from mercury isotopes and PMF model. *Science of the Total Environment, 869*, 161774.

Zhang, Y. X., Li, T. S., Guo, Z. H., et al. (2023b). Spatial heterogeneity and source apportionment of soil metal(loid)s in an abandoned lead/zinc smelter. *Journal of Environmental Sciences., 127*, 519–529.

Zhang, Y. X., Song, B., & Zhou, Z. Y. (2023c). Pollution assessment and source apportionment of heavy metals in soil from lead – Zinc mining areas of south China. *Journal of Environmental Chemical Engineering, 11*(2), 109320.

Zhou, Y., Jiang, D. D., Ding, D., et al. (2022a). Ecological-health risks assessment and source apportionment of heavy metals in agricultural soils around a super-sized lead-zinc smelter with a long production history, in China. *Environmental Pollution., 307*, 119487.

Zhou, Y. T., He, H. P., Wang, J., et al. (2022b). Stable isotope fractionation of thallium as novel evidence for its geochemical transfer during lead-zinc smelting activities. *Science of the Total Environment, 803*, 150036.

Zhou, H., Chen, Y., Yue, X. M., Ren, D. J., et al. (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. *Journal of Hazardous Materials, 445*, 130528.