



OPEN

Characterization and source apportionment of heavy metal contamination in agricultural soils in the complex genesis region of western Yunnan

Yingmei Li^{1,3,4,6}, Sheng Wang^{2,4,6}, Xiaoyan Shang^{1,3,4}, Hongyin Zhou^{4,5}, Jianyang He^{1,4}, Wen Chen^{1,4}, Lijuan Wang^{1,4}, Xiangmei Zhao^{1,4}, Li Bao^{1,4} & Naiming Zhang^{1,4}✉

The genesis of heavy metal contamination in arable soils is complex, and scientifically identifying risks and precisely analyzing contamination sources are essential for safely using contaminated arable land. In this study, we systematically evaluated the pollution characteristics of Cu, Zn, As, Hg, Cd, Pb, Ni, and Cr in soil, and then applied the APCS-MLR and PMF models to jointly analyze pollution sources and their contributions. The results showed that the concentrations of the eight heavy metals were significantly higher than the background values for soils in Yunnan Province, exhibiting clear spatial heterogeneity. The overall pollution level ranged from mild to severe, with Cd and Pb being the most critical contaminants. Four major pollution sources (industrial transportation, parent material, agriculture, and mining) were identified through the dual modeling approach. The results of both models corroborated each other, and the accuracy of the analysis was significantly improved compared to using a single method. This study not only provides a scientific basis for the safe utilization of contaminated arable land in western Yunnan, an area with a complex genesis of soil contamination, but also offers a generalized framework for source analysis in areas affected by geological-anthropogenic composite pollution.

Keywords High geological background, Heavy metal contamination, Cropland soil, Source analysis

Soil is an important resource for safeguarding human survival and development, and it is closely related to food security and ecological sustainability^{1,2}. With the rapid development of urbanization and industrialization, soil heavy metal pollution has become a serious environmental problem^{3,4}. In China, the limited per capita availability of arable land has led to the continued use of soils contaminated with heavy metals for agricultural production, in order to maintain food supply and meet land-use demands⁵. According to statistics, about 19.4% of China's arable land has varying degrees of heavy metal contamination^{6,7}. Because of its unique geologic background⁸ and the impact of human activities, the western Yunnan region has become a high-risk area for soil heavy metal pollution^{9,10}. The complex causes of the soil heavy metal pollution in this region have made pollution tracing more challenging and pollution management more complex^{11–13}. Compared with other high-background regions such as Greece and India, western Yunnan exhibits more intricate pollution characteristics due to its highly variable parent materials, intensive geological activity, and the overlapping impacts of mining, agriculture, and urbanization. However, current studies focusing on this region remain limited, leaving a gap in source identification research under such complex conditions. Therefore, it is crucial to investigate the risks and sources of heavy metal pollution in arable land in this region. Such research provides a scientific basis for effective source control and targeted pollution management.

¹College of Resources and Environment, Yunnan Agricultural University, Kunming 650201, China. ²School of Chemistry and Geographical Sciences, Chuxiong Normal University, Chuxiong 675000, China. ³College of Biology and Chemistry, Pu'er University, Pu'er 665000, China. ⁴Yunnan Soil Fertility and Pollution Remediation Engineering Research Center, Kunming 650201, China. ⁵School of Agriculture and Life Sciences, Zhaotong University, Zhaotong 657000, China. ⁶Yingmei Li and Sheng Wang contributed equally to this work and share first authorship. ✉email: zhangnaiming@sina.com

Spatial heterogeneity is a key feature of soil heavy metal pollution. Numerous studies have demonstrated that spatial variation is influenced by both natural factors—such as geological background, topography, and landforms—and anthropogenic activities¹⁴. Chen et al.¹⁵ evaluated the spatial distribution characteristics and contamination levels of soil heavy metals in various administrative regions in China, and found that their concentrations in cultivated land in the south were significantly higher than those in the north. Hu et al.¹⁶ demonstrated that concentrations of multiple heavy metals, with the exception of Hg, surpassed local background values, indicating that agricultural activities played a predominant role in the enrichment of heavy metals in surface soils. A variety of technical tools and modeling analysis methods have been widely used in source analysis of soil heavy metal pollution^{17–20}. Among them, the absolute principal component scores-multiple linear regression (APCS-MLR) model, which combines principal component analysis (PCA) and multiple linear regression (MLR), is able to quantify the contributions of different pollution sources under simple pollution scenarios⁴. The positive matrix factorization (PMF) (orthogonal minimum matrix factor) model has been widely used in soil heavy metal pollution research in recent years due to its advantage of evaluating the uncertainty of the source components²¹. For example, in studies focusing on coking industry regions, the PMF model effectively attributed the majority of soil heavy metal pollution to industrial emissions²². Lui et al.²³ combined the results of the APCS-MLR and PMF models and pointed out that industrial and agricultural activities were the main sources of soil heavy metal pollution in the Yunnan Pb–Zn mining area. Chen et al.^{24,25} reached a similar conclusion and pointed out that PMF could achieve finer identification of the pollution sources. Western Yunnan, China, is globally recognized for its widespread high geological background levels of heavy metals in soils, particularly Cd, Pb, and As. This region is characterized by a combination of natural and anthropogenic pollution sources, a unique distribution of geological sources, and a wide variety of human activities, which makes it difficult for a single model to comprehensively assess the contributions of different sources^{26,27}. Although both APCS-MLR and PMF have been successfully applied in other areas, their joint application has not yet been reported for complex farmland soils under high-background conditions in western Yunnan. This highlights the need for more refined analytical approaches to understand pollution sources in such regions, and provides the rationale for applying a combined model framework in this study^{28,29}.

In this study, we analyzed and evaluated the distribution characteristics of heavy metals and their comprehensive pollution in contaminated cropland under the combined influence of high geological background values for heavy metals and anthropogenic pollution in western Yunnan, China. The APCS-MLR and PMF models were used to investigate the sources and contributions of heavy metals, with the aim of providing a scientific basis for the safe utilization of contaminated arable land under the combined influence of high geological background values for heavy metals and anthropogenic pollution in western Yunnan, China.

Materials and methods

Overview of the study area

The study area is located in the mountainous western Yunnan region of Yunnan Province, China (100°01′–100°29′ E, 25°57′–26°42′ N) (Fig. 1). This area has diverse geomorphological characteristics and is dominated by mountains and hills. The predominant soil type in the area is red soil. The region has a subtropical monsoon climate, with an average annual temperature of 14.7 °C, an average annual precipitation of 939.7 mm, and a mean elevation of 2696 m. The study area is rich in mineral resources, with 18 types of minerals and 12 deposits (mountains) discovered, and is an important mineral resource-rich area in western Yunnan. High-quality manganese-rich ores, rock-gold ores, and bauxite are the primary resources. The area covered by mining, beneficiation, smelting, metallurgical processing, and the production and processing of mineral resources has gradually grown. With high-quality manganese-rich ore and bauxite as the core, it has gradually formed a complete mining industry chain covering mining, ore dressing, smelting, and comprehensive utilization of resources. Long-term mineral mining and smelting activities have been provided the backbone of local economic development, though it has simultaneously impacted the surrounding ecological environment and soil quality. Through data collection, investigation, and monitoring analysis, it was found that there were 78,021.32 acres of contaminated arable land in this area, most of which was in seven townships (Fig. 2).

Sample collection and processing

The distribution of the sampling points is shown in Fig. 2. Cropland soils surrounding the contaminated sites were selected as representative sampling areas. At each sampling site, five locations were selected for soil sampling within the 0–20 cm tillage layer using the double-diagonal five-point sampling method. The five sub-samples were mixed into a homogeneous sample immediately after sampling, and a total of 400 soil samples were collected. In the laboratory, the samples were air-dried at room temperature, and foreign materials such as gravel and film were removed. The samples were ground and passed through a 100-mesh nylon sieve, and sealed in polyethylene self-sealing bags prior to testing.

Project measurement and methodology

The pH values of the soil samples were determined according to the Determination of Soil pH standard (NY/T 1377-2007), and the heavy metal contents were analyzed according to the following standard methods: Cd and Pb were analyzed according to the Determination of Lead and Cadmium in Soil Quality Graphite Furnace Atomic Absorption Spectrophotometric Method (GB/T 17141-1997); Cr, Cu, Zn, and Ni were analyzed according to the Determination of Copper, Zinc, Lead, Nickel, and Chromium in Soil and Sediment via Flame Atomic Absorption Spectrophotometry (HJ 491-2019); as was analyzed according to the Determination of Arsenic in Soil via Atomic Fluorescence Spectrometry (HJ 694-2014); and Hg was analyzed according to the Determination of Mercury in Soil via Cold Atomic Absorption Spectrophotometry (Cold AAS) (HJ 597-2011). The analyses were performed using a Thermo Scientific iCE 3000 AAS atomic absorption spectrometer. All of the reagents

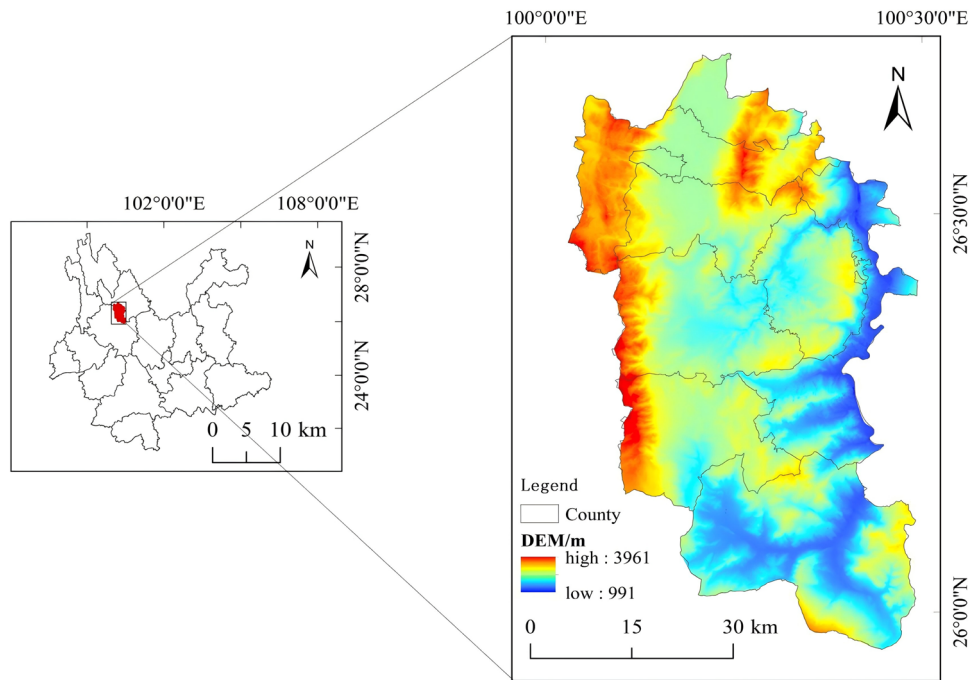


Fig. 1. Overview map of the study area. National Catalogue Service for Geographic Information of China, National county-level administrative boundaries shapefile data (<https://www.webmap.cn/main.do?method=index>); Geospatial Data Cloud, 30-m resolution digital elevation data (<https://www.gscloud.cn/>). The map was created using ArcGIS 10.8 (Environmental Systems Research Institute, Redlands, CA, USA; <https://www.esri.com/>).

were of superior purity, and the water used was ultrapure water. Quality control was strictly carried out using blank samples, parallel samples, and spiked recovery samples. The correlation coefficients of the standard curves of the elements were greater than 0.999, the relative standard deviations between the parallel samples were less than 10%, and the recoveries of the spiked samples were 80% ~ 110%, demonstrating the accuracy and reliability of the data.

Evaluation method

The ground accumulation index (Igeo)^{26,30} was used to comprehensively evaluate the soil heavy metal pollution. The Igeo is mainly used to reflect the enrichment level of individual heavy metals in the soil. All of the reference values in this study were selected from the background values of soil in Yunnan Province^{4,31}.

Source resolution models

APCS-MLR modeling

The APCS-MLR model is a pollutant source analysis method based on PCA³², and has been widely used in studies of soil heavy metal pollution³³. It first extracts potential components via PCA, then converts factor scores into absolute principal component scores (APCSs), and finally uses multiple linear regression (MLR) to quantify the contribution of each source, with pollutant concentrations as the dependent variable and APCSs as independent variables. The number of components is typically determined by eigenvalues (> 1) or scree plot, and model performance is evaluated using R^2 and significance tests.

PMF model

The PMF model, recommended by the U.S. EPA³⁴, is based on factor analysis and decomposes a matrix of pollutant concentrations to identify source profiles and their contributions³⁵. The number of factors is determined by Q-value comparison, residual analysis, and the interpretability of results. FPEAK values are adjusted to reduce rotational ambiguity. Uncertainties are estimated from detection limits and analytical errors and input into the model. Model performance is assessed by correlation between observed and predicted values, signal-to-noise ratios (S/N), and bootstrap or DISP tests.

Data processing and analysis

Excel was used to carry out the statistics and calculations of the heavy metal contents and ground accumulation index. ArcGIS 10.8 was used to produce the overview map of the study area and the spatial distribution maps of the heavy metals. Origin 2024 was used to plot the ground accumulation index. SPSS 26 was used to conduct the principal component and correlation analyses; and EPA PMF 5.0 was used to analyze the heavy metal pollution sources.

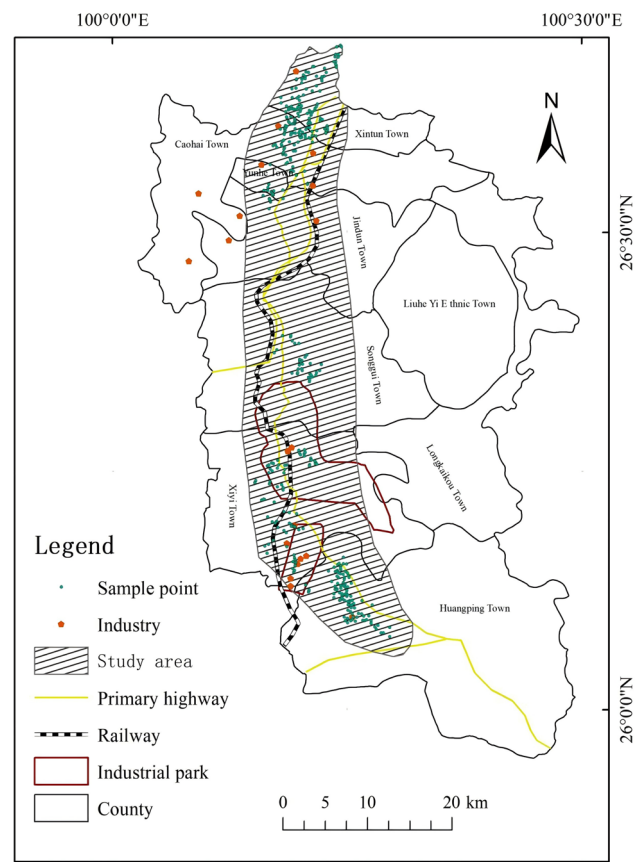


Fig. 2. Map showing the distribution of the sampling points. National Catalogue Service for Geographic Information of China, National County-level Administrative Boundaries Shapefile Data (<https://www.webmap.cn/main.do?method=index>). The map was created using ArcGIS 10.8 (Environmental Systems Research Institute, Redlands, CA, USA; <https://www.esri.com/>).

Element	Maximum (mg/kg)	Minimum (mg/kg)	Mean (mg/kg)	Median (mg/kg)	Standard deviation	Coefficient of variation	Soil background values in Yunnan Province
Cu	468.00	24.10	123.54	89.75	78.89	0.64	46.3
Zn	1310.00	45.30	232.27	148.00	196.56	0.85	89.7
As	177.00	2.70	26.38	20.75	19.58	0.74	18.4
Hg	1.98	0.02	0.14	0.10	0.14	1.00	0.06
Cd	9.95	0.16	1.26	0.70	1.26	0.99	0.22
Pb	3780.00	14.50	244.18	51.80	442.39	1.81	40.6
Ni	337.00	20.30	115.89	115.00	54.56	0.47	42.5
Cr	682.00	54.10	205.36	200.00	99.70	0.49	65.2
pH	8.60	4.08	7.17	7.49	0.91	0.13	5.7

Table 1. Statistics of heavy metal contents of soil in the study area.

Results and analysis
Characteristics of soil heavy metal contents

The environmental quality of the soil in the study area is presented in Table 1, which shows that the soil in the study area is generally neutral. Eight heavy metal elements were within the medium–high intensity variation range. Hg and Pb were within the high intensity variation range, and the order of the values of the coefficients of variation are as follows: Pb>Hg>Cd>Zn>As>Cu>Cr>Ni. Compared with the background values of the soil heavy metals in Yunnan Province, the average contents of the eight heavy metals were higher than the background values of the soil, i.e., 1.26 to 6.01 times higher. The pollution levels of the eight heavy metals in the soils in the study area were evaluated using the soil pollution screening values in the Soil Environmental Quality Risk Control Standards for Soil Pollution on Agricultural Land (Trial) (GB 15618-2018) as reference values (Table 2). The heavy metal contents of all of the soil samples collected from the study area exceeded the

Element	Number of exceedances of risk screening values				
	pH≤5.5	5.5<pH≤6.5	6.5<pH≤7.5	pH>7.5	Total
Cu	20 (76.9)	52 (7.5)	42 (38.5)	94 (47.7)	208 (52.0)
Zn	7 (26.9)	10 (14.7)	19 (17.4)	59 (30.0)	95 (23.8)
As	1 (3.8)	4 (5.9)	44 (40.4)	93 (47.2)	142 (35.5)
Hg	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Cd	25 (96.2)	68 (100.0)	107 (98.2)	120 (61.0)	320 (80.0)
Pb	13 (36.1)	27 (39.7)	28 (25.7)	52 (26.4)	120 (30.0)
Ni	15 (57.7)	30 (44.1)	59 (54.1)	14 (7.1)	118 (29.5)
Cr	6 (23.1)	22 (3)	56 (51.4)	60 (30.5)	144 (36.0)
pH	26	68	109	197	400

Table 2. Frequency and proportion of soils exceeding risk screening values.

standard values, except for Hg, and their exceedance rates were Cd (80.0%), Cu (52.0%), Cr (36.0%), As (35.5%), Pb (30.0%), Ni (29.5%), and Zn (23.8%).

Characterization of spatial distributions of heavy metals in soil

Inverse distance weighted (IDW) interpolation was used to map the spatial distributions of the heavy metal contents. By analyzing the natural turning points and characteristic points, the heavy metal contents were classified into clusters with similar properties to maximize the differences between the classes (Fig. 3). Cd, Pb, Zn, and Cu exhibited similar piecewise enrichment characteristics in the study area, and the high-value areas were mainly concentrated in the southern part of the study area, demonstrating that human activities had a significant influence on the distributions of the soil heavy metals in the areas close to the industrial concentration areas and near the main transportation routes. Cr and Ni exhibited similar piecewise enrichment, and there was a strong spatial correlation between the distributions of As, Hg, and pH. In the high-value heavy metal content areas, the pH value was correspondingly higher. Based on analysis of the coefficients of variation, the spatial distribution patterns of these elements in the study area are similar, indicating that the distributions of these elements are influenced by similar factors.

Evaluation of soil land accumulation index in the study area

The results of the evaluation of the soil geoaccumulation index of the cultivated land in the study area are shown in Fig. 4. The results show that the ranges of the land accumulation indices for Cu, Zn, As, Hg, Cd, Pb, Ni, and Cr were −1.5–2.8, −1.6–3.3, −3.4–2.7, −2.1–4.5, −1.0–4.9, −2.1–6.0, −1.7–2.4, and −0.9–2.8, with mean values of 0.6, 0.5, −0.4, 0.3, 1.5, 0.6, 0.7, and 0.9, respectively. The order of the pollution degrees of the elements was $Cd \geq Cr \geq Ni \geq Pb \geq Cu \geq Zn \geq Hg \geq As$. The spatial distributions of land accumulation indices for the eight heavy metals exhibited different degrees of variations (Fig. 5). Among them, Cd and Pb had the highest land accumulation pollution indices, reaching 4.9 and 6.0, respectively, indicating that there was a significant anthropogenic pollution source in localized areas³⁶. These areas were mainly concentrated in the southern part of the study area and near the industrial and transportation areas, suggesting that industrial activities and transportation emissions were the key factors influencing the accumulation of these heavy metals. Comparatively speaking, the geoaccumulation index of As was lower, indicating that the degree of accumulated pollution was slight and was closer to the natural background value. The overall level of the soil heavy metal accumulation in the study area was mainly slight to mildly polluted, and it did not reach heavy or serious pollution. The pollution level was greater in the southern part of the study area than in the northern part of the study area.

Pollution source analysis modeling

Correlation analysis

Pearson correlation analysis was performed on the contents of the eight heavy metal elements. The results are shown in Fig. 6. The results showed that many of the heavy metal elements and the pH exhibited significant correlations, indicating that they had common sources and similar migration and transformation processes. Cu, Zn, Cd, and Pb exhibited significant positive correlations, indicating that these four elements had similar input pathways, such as industrial emissions, mining activities, and/or agricultural inputs. Additionally, the correlations of As with Hg, Cd, and Pb further revealed the contributions of deposition and/or anthropogenic disturbances to the accumulation of these heavy metals. In addition, the independence of Ni and Cr indicated that they originated from natural geochemical processes rather than from significant exogenous pollution inputs. In addition, the pH, as an important environmental factor, profoundly affected the morphology, solubility, and transport behaviors of the heavy metals in the soil.

APCS-MLR modeling

Multiple linear regression analysis of the eight heavy metal elements revealed that the Bartlett’s sphericity test companion probability was 0, which was lower than the significance level of 0.05, and the Kaiser–Meyer–Olkin (KMO) test statistic value was 0.753, indicating that the sample data for the study area was suitable for PCA³⁷. Through PCA, three principal components were extracted (Table 3). The first principal component (F1) had a variance contribution rate of 44.78%, and it explained the largest proportion of the variance in soil heavy metal

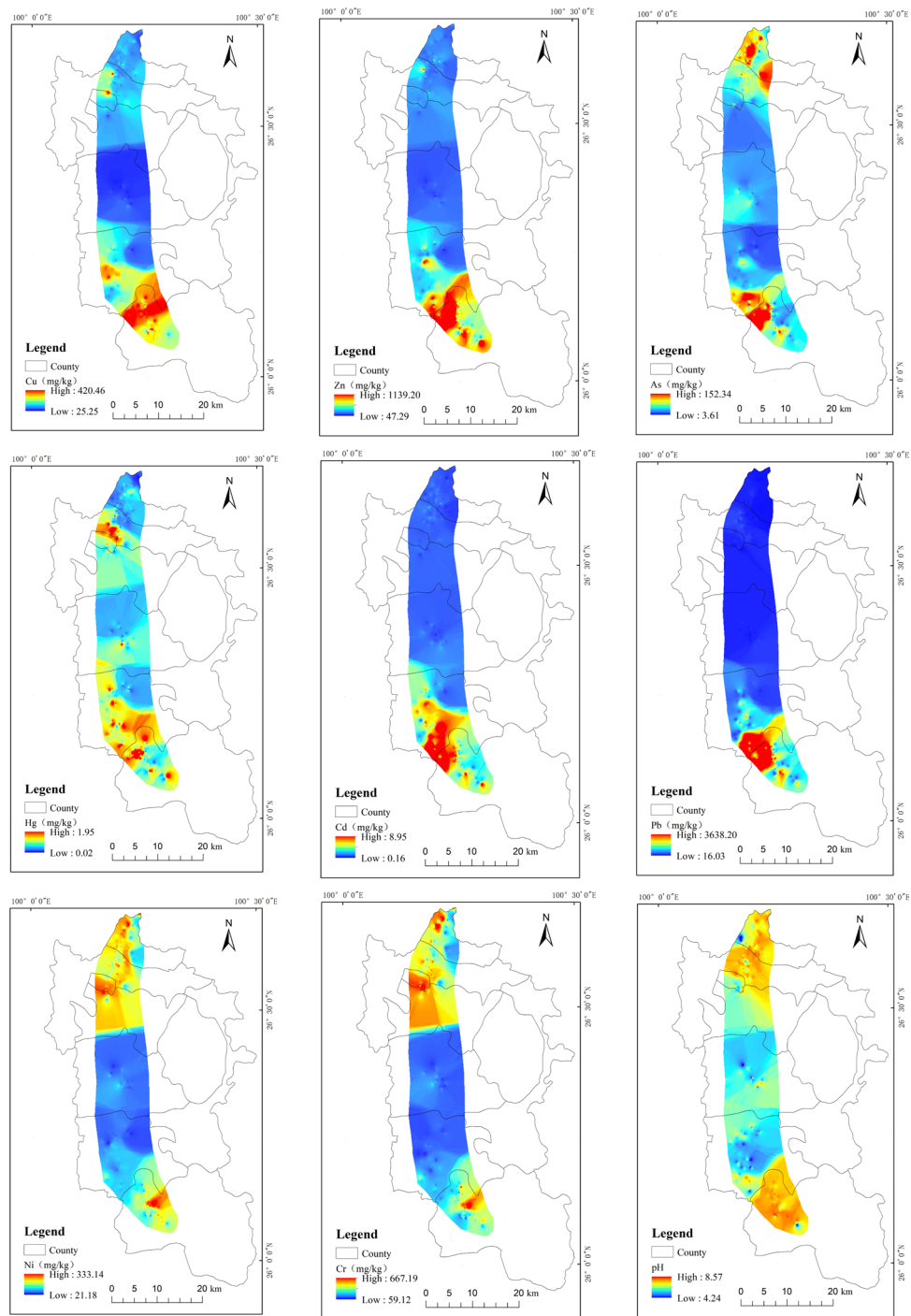


Fig. 3. Spatial distributions of heavy metal contents and pH in arable soil. National Catalogue Service for Geographic Information of China, National County-level Administrative Boundaries Shapefile Data (<https://www.webmap.cn/main.do?method=index>). The map was created using ArcGIS 10.8 (Environmental Systems Research Institute, Redlands, CA, USA; <https://www.esri.com/>).

distributions, corresponding to high concentrations or strong symbiosis of the heavy metal fractions. The second principal component (F2) and the third principal component (F3) contributed 24.53% and 11.22%, respectively, with a cumulative variance contribution of 80.53%, which explained most of the information about the heavy metal pollution.

The explanatory degree of F1 was significantly higher than those of the other principal components, and the heavy metals with higher loadings were Hg, Cu, Zn, Cd, and Pb. Lv et al. (2013) and Wang et al. (2021) showed that when Cd, Pb, and Zn were classified into the same principal component, they were mainly affected by anthropogenic sources, and the contribution of Hg was the highest (66.99%)^{38,39}. Studies have also shown

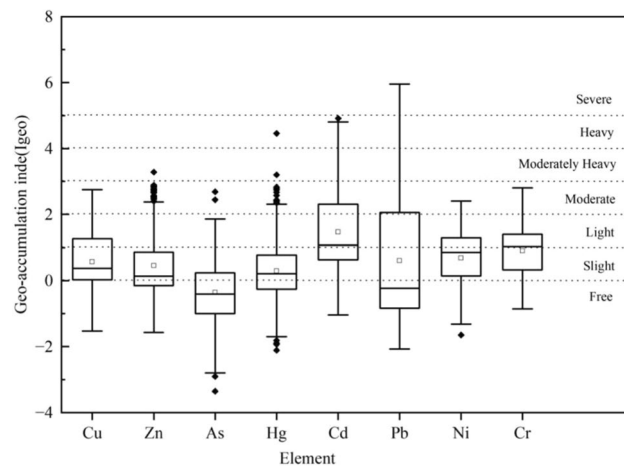


Fig. 4. Box plot of land accumulation index.

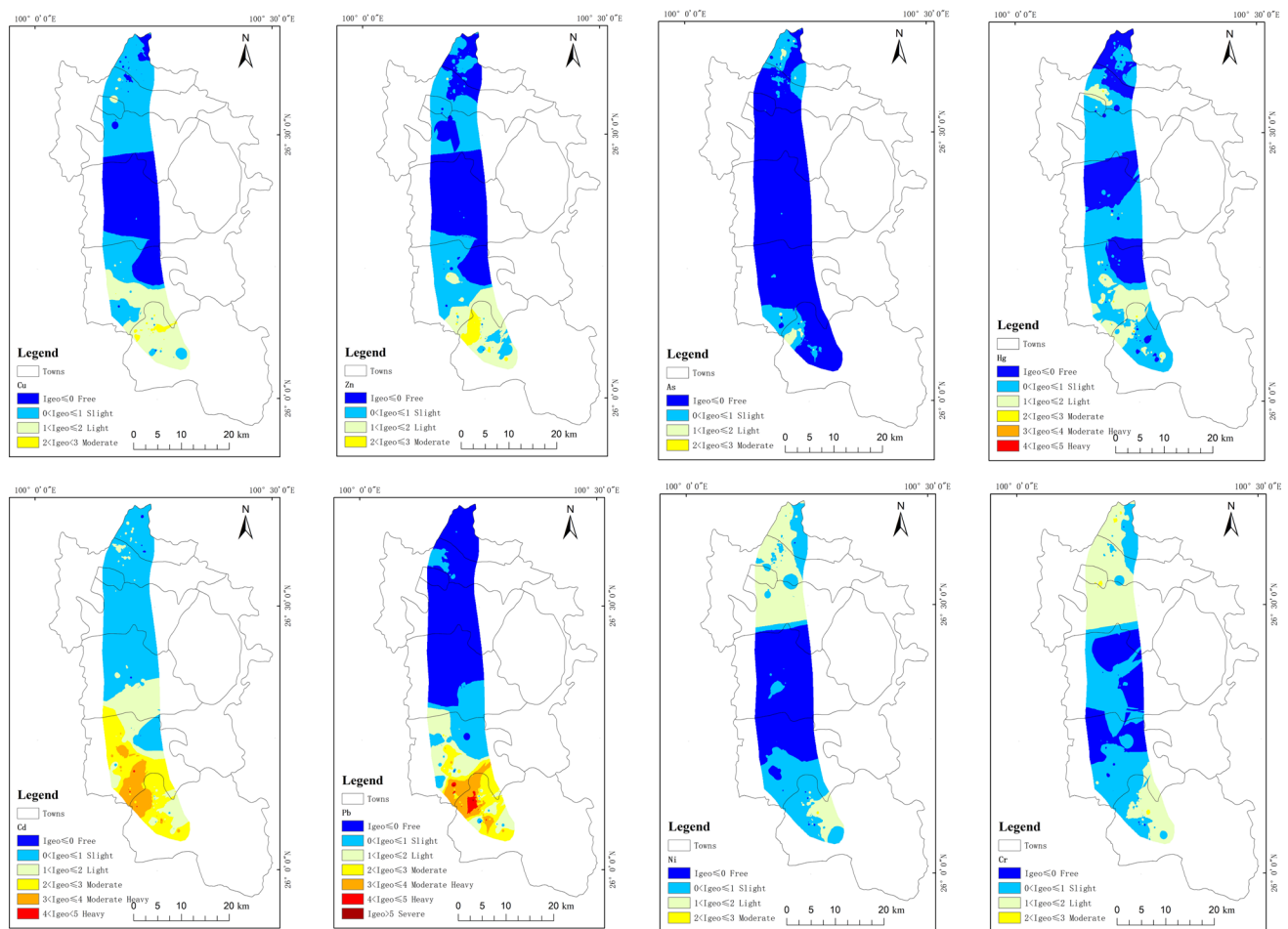


Fig. 5. Spatial distributions of the land accumulation index values of the heavy metals in arable soils. National Catalogue Service for Geographic Information of China, National County-level Administrative Boundaries Shapefile Data (<https://www.webmap.cn/main.do?method=index>). The map was created using ArcGIS 10.8 (Environmental Systems Research Institute, Redlands, CA, USA; <https://www.esri.com/>).

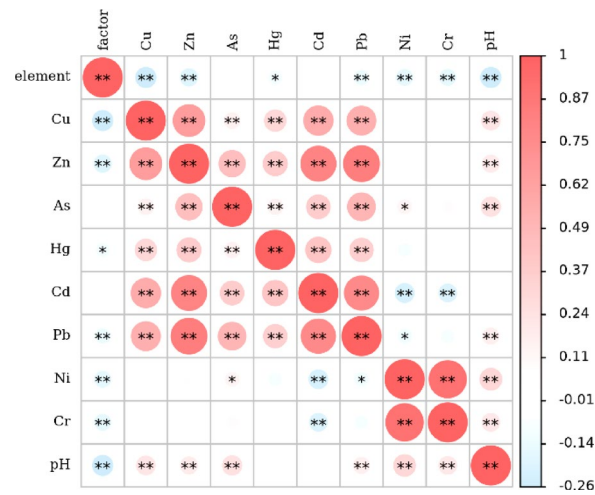


Fig. 6. Results of correlation analysis of heavy metal content of soil.

Components	Initial eigenvalues			Extracted load sum of squares			Rotated sum of squared loadings		
	Total	%Var	%Cum	Total	%Var	%Cum	Total	%Var	%Cum
F1	3.59	44.78	44.78	3.58	44.78	44.78	2.68	33.55	33.55
F2	1.97	24.53	69.31	1.96	24.53	69.31	1.99	24.87	58.43
F3	0.90	11.22	80.53	0.90	11.22	80.53	1.77	22.10	80.53
F4	0.75	9.39	89.91						
F5	0.35	4.43	94.34						
F6	0.22	2.76	97.09						
F7	0.14	1.78	98.87						
F8	0.09	1.13	100.00						

Table 3. Principal component analysis results: rotated component loadings and variance explained. Note: “Initial eigenvalues” indicate the variance explained by each principal component prior to rotation. “Extracted load sum of squares” refers to the variance explained by the extracted components before rotation. “Rotated sum of squared loadings” effects the redistributed variance contributions after Varimax rotation. %Var denotes percentage of variance explained by each component; %Cum represents cumulative explained variance. Only components with eigenvalues > 1 were retained, based on the Kaiser criterion.

that Hg is affected by multiple factors, such as soil-forming matrices, industrial emissions, transportation, and agriculture, and that bituminous coal has a significant effect on its accumulation in soil^{40,41}. In addition, transportation exhaust, dust, and particles can all contribute to the increase in soil Hg after atmospheric deposition. According to the spatial distribution characteristics and correlation analysis of the heavy metals, Cu, Hg, Zn, Cd, and Pb were highly enriched and correlated in the southern part of the study area, and the pollution was particularly serious in the industrial parks and near the transportation routes, indicating that in the study area, these metals were greatly influenced by industrial and transportation sources. Based on the high value areas in the spatial distributions of these five elements and the distributions of the industrial zones and transportation routes, it was concluded that F1 was an industrial and transportation source.

The F2 source contributed the most to Cr and Ni, reaching 96.99% and 93.51%, followed by Cd (25.26%). Some studies have shown that Cr and Ni are mainly affected by the soil parent material and soil-forming processes^{42,43}. The ground accumulation indexes and comprehensive pollution indexes of Ni and Cr in the study area did not indicate pollution, and their coefficients of variation were relatively small, further indicating that the Ni and Cr in the soil in the study area were less influenced by exogenous sources and mainly came from the parent material of soil formation processes. Therefore, we concluded that F2 was the parent material source.

Source F3 contributed the most to As, Zn, and Pb, with values of 51.75%, 26.18%, and 22.95%, respectively. According to the preliminary investigation, this area is rich in mineral resources, and there are non-ferrous metal resources such as gold, lead–zinc, and copper ores in the study area, which is an important mineral development zone in Yunnan Province. As is an accompanying element in gold ores. Based on the fact that the areas with high As, Zn, and Pb values overlapped with smelting plant locations, we concluded that F4 was a mining source.

Unknown source F4 mainly contributed to As, Pb, and Hg, with contribution values of 26.49%, 25.52%, and 19.86%, respectively, although these contribution values were not high. Hg and Pb are important components of pesticide fertilizers, can enter the soil through their application and become enriched on the soil surface⁴⁴.

The main land use type in the study area is arable land, suggesting that this part of the impact originates from agricultural activities; therefore, we conclude that F4 is an agricultural source.

The APCS-MLR model was used to analyze the eight heavy metals via multiple linear regression (Fig. 7). The predicted values fit well with the measured values, and the results show that the R^2 values of Cu, Zn, As, Hg, Cd, Pb, Ni, and Cr are 0.687, 0.867, 0.841, 0.497, 0.810, 0.838, 0.949, and 0.942, with an average R^2 value of 0.804. The estimated/measured values (E/O) of all of the elements were close to 1, indicating that the APCS-MLR model analysis results have a high credibility. Notably, the R^2 value of Hg was relatively moderate (0.497), which may be attributed to the discontinuous and nonlinear nature of agricultural Hg inputs, such as sporadic application of Hg-containing fertilizers or pesticides. The industrial and transportation activities contributed substantially to the Cu, Zn, Hg, Cd, and Pb pollution. The soil matrices had a significant influence on the Cr, Ni, and Cd pollution and were the main sources. The mining industry also had a large influence on the As, Zn, and Pb pollution.

PMF model

In order to verify the reliability of the source analysis results of the APCS-MLR model, the PMF model was used to conduct quantitative source analysis of the eight heavy metals. Three to five factors were selected for random iterative operations, and four factors were finally determined to obtain the contributions of the four heavy metal pollution sources to each heavy metal (Fig. 8).

The elements with high loadings for Factor 1 were Pb, Cd, Zn, and Cu, which were mainly found in mining areas, transportation routes, and other areas with intensive human activities. This aligns well with the results of the APCS-MLR model analysis. The results of both models indicated that Factor 1 was an industrial and transportation source. Factor 2 contributed the most to Cr and Ni, which had small coefficients of variation. The APCS-MLR model results were highly consistent with the results of the APCS-MLR model, and both models indicated that Factor 2 was a parent material source. Factor 3 had the highest loading for As, which aligned with the fact that the geologic background of the study area was rich in mineral resources and is an important development zone for non-ferrous metals, such as gold, in Yunnan Province. This is essentially the same as the results from the APCS-MLR model, and both models indicate that Factor 3 is a mining industry source. Factor 4 had high loadings for Hg and Cd. Several studies have shown that fertilizers and pesticides often contain elements, such as Hg, Cd, and Pb, which can be indicative of agricultural activities such as the application of pesticides and fertilizers⁴⁵. Thus, we concluded that Factor 4 was an agricultural source. The PMF model identifies the four sources of contamination of the eight heavy metals as industrial traffic, parent material, mining, and agricultural sources.

The results of the pollution source contribution analysis using the PMF model (Fig. 8) correspond well with the results of the correlation analysis (Fig. 6) and the APCS-MLR model analysis (Fig. 7). The industrial and transportation activities contributed more to the Cu, Zn, Hg, Cd, and Pb pollution. The soil matrix had an important influence on the Cr and Ni, which aligned with the independence of several elements in the correlation analysis and the natural sources corresponding to F2 in the APCS-MLR model. The mining sources had an influence on As, Hg, and Cd, which aligned with the analysis of F3 in the APCS-MLR model, and the fact that agricultural sources had a greater influence on Hg and Cd. This conclusion is reflected by the results presented in Fig. 6, Table 3, and Fig. 7, which show the main composition of the soil heavy metal pollution sources and their relative contributions in the study area.

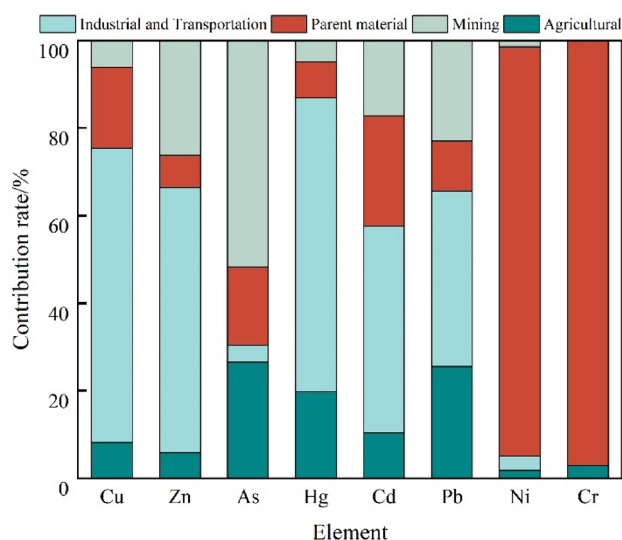


Fig. 7. Contributions of pollution sources in the APCS-MLR model.

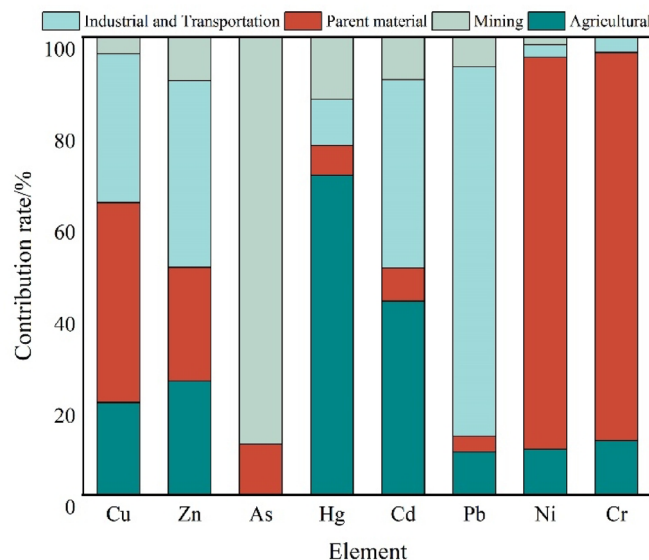


Fig. 8. PMF model pollution source contributions.

Discussion

Universal analysis of heavy metal pollution characteristics and causes

This study reveals the combined factors behind heavy metal contamination in cropland soils of the western Yunnan high geological background region: both natural parent material contributions and anthropogenic activities have jointly shaped the contamination patterns. Similar to other geological regions with high background levels worldwide (e.g., the ophiolite region of Greece, the Deccan Plateau of India), elements such as Cr and Ni in the study area are primarily sourced from the parent material (with contribution rates >90%), and their spatial distribution closely aligns with the weathering characteristics of the regional bedrock^{46,47}. In contrast, elements such as Cd, Pb, and Cu exhibit significant enrichment (with exceedance rates ranging from 30 to 80%), with Cd and Pb concentrations reaching 6.01 and 4.2 times their background values, respectively, indicating that human activities are the primary driving force behind the intensification of pollution¹⁰. Spatial analysis further reveals that the accumulation coefficients of Cd and Pb around industrial areas and transportation corridors are as high as 4.9 to 6.0, with pollution patterns similar to those observed in typical mining areas worldwide, such as the Mount Isa Pb–Zn mining region in Australia^{47,48}, and the contribution of Pb and Zn from transportation emissions in the Mississippi River Basin in the United States^{25,49}, collectively confirming the significant role of industrial and transportation sources. Notably, the pollution characteristics of As in western Yunnan are distinctive: despite a low geocumulative index (−0.4), areas of high As concentration overlap significantly with the distribution of gold mining belts, suggesting a synergistic effect between geological background and mining activities. This type of “implicit high background superimposed with anthropogenic disturbance” pollution pattern is widely observed in lateritic gold mining areas across Southeast Asia (e.g., Laos, Myanmar)⁵⁰, indicating that the “geological-anthropogenic dual-source coupling” analytical framework proposed in this study holds substantial potential for regional application.

Advantages of multi-modeling for analyzing complex pollution areas

The combined application of the APCS-MLR and PMF models effectively addresses the limitations associated with using a single model to analyze complex pollution sources. ① Complementarity in source apportionment: Both models yielded consistent results, indicating that the combined contribution of Cd and Pb ranged from 58 to 63%. This is in close agreement with the emission inventory data from lead–zinc smelting areas in western Yunnan, where Cu, Zn, and Cd collectively account for more than 50% of total emissions¹⁸. APCS-MLR identifies the direct contribution of industrial and traffic sources to Cd and Pb through the principal component loadings ($R^2 = 0.84$). In contrast, the PMF model captures the more complex and indirect inputs of agricultural sources to Hg, with an estimated contribution of 19%–26% derived from factor decomposition. The combined use of both models provides a comprehensive view of the continuous and discrete nature of pollution sources⁵¹. ② Reduction of Geological Background Interference: To address the issue of “false-positive contamination” from Cr and Ni in high-background areas, the dual-model approach has been validated through both coefficient of variation analysis ($CV < 15\%$) and elemental correlation networks ($r = 0.92$), successfully removing the interference from parent material sources on the contamination index. This approach can be applied to similar geological regions, such as the Iron Quadrangle in Brazil and the Bushveld Complex in South Africa⁵². ③ Optimization of Agricultural Source Quantification: The PMF model captures the nonlinear relationship between Hg and fertilizer application intensity ($R^2 = 0.47$), addressing the limitations of APCS-MLR in resolving discrete sources. This is particularly important in cases like Hg, where the APCS-MLR model showed only moderate fit ($R^2 = 0.497$), likely due to irregular agricultural inputs. The improved resolution offered by PMF provides

a valuable methodological reference for similar agriculture–mining complex regions. ④ Accuracy of Policy Response: Model analysis indicates that industrial and transportation sources contribute 58% (APCS-MLR) to 63% (PMF) of the Cd pollution, significantly higher than the 35% contribution observed in typical agricultural areas of southern China¹⁸. This suggests that western Yunnan should prioritize targeted strategies such as “cleaner production reform in key industries + interception of particulate matter from transportation emissions” rather than simply applying agronomic control measures used in other regions. In addition, Cd also showed a moderate contribution (25.26%) from the parent material source (F2), indicating that geogenic background still plays a non-negligible role. This reflects the complex dual-source characteristics of Cd in the region, suggesting the need for differentiated control measures that consider both anthropogenic and natural sources.

Decision-support for differentiated management of safe use of agricultural soils

This “pollution characterization–source analysis–risk zoning” framework supports the development of refined soil management strategies tailored to varying contamination levels and specific agricultural land use patterns, including crop sensitivity and planting objectives. ① High Geological Background Control Zone (Cr and Ni Enrichment Zones): Implement low-accumulation crop cultivation and pH regulation to prevent the activation of parent material and its release⁵³. ② Compound Pollution Remediation Zone (Cd, Pb Exceedance Zones): Focus on structural adjustments, while adopting plant passivator and blocking technologies to mitigate risks to human health. Additionally, source interception should be applied to industrial and transportation sources⁵⁴. ③ Compound Risk Early-Warning Zone (Cd and Pb Compound Pollution Zones): Establish a dynamic monitoring network to track real-time changes in soil and environmental quality, and prioritize the closure or upgrading of non-compliant smelting enterprises to cut off major anthropogenic pollution sources. This hierarchical control strategy has been validated in typical karst regions of Southwest China and has been shown to reduce pollution migration risks in areas with complex geological backgrounds and compound pollution⁵⁵.

Conclusions

- (1) The soil in the study area was overall neutral, and the average contents of the eight heavy metals, namely, Cu, Zn, As, Hg, Cd, Pb, Ni, and Cr, ranging from 1.26 to 6.01 times the background values in Yunnan Province. Among them, Cd and Pb pollution was relatively severe, with concentrations in 80% and 30% of the samples, respectively, exceeding the risk screening values for agricultural land (GB 15618-2018).
- (2) The spatial distributions of heavy metals showed considerable heterogeneity. Based on the cumulative geo-accumulation index, which integrates metal enrichment and potential ecological risk, Cd and Pb exhibited the highest levels, indicating stronger anthropogenic influence and higher management priority. In contrast, Hg had the lowest index, suggesting limited accumulation and relatively low environmental concern.
- (3) This study demonstrated that heavy metal contamination in the study area was primarily attributed to industrial and traffic-related emissions (accounting for 58%–63% of Cd and Pb) and to geogenic sources, with parent materials contributing over 90% of Cr and Ni. A nonlinear contribution of agricultural sources to Hg was identified by the PMF model, likely resulting from sporadic applications of Hg-containing agrochemicals. These findings highlight the necessity of strengthened regulation of agricultural inputs, the promotion of low-Hg alternatives, and targeted environmental monitoring in agricultural zones. The integrated use of dual-source apportionment models significantly improved source differentiation and clarified the respective roles of anthropogenic and natural influences, thereby providing a scientific basis for source-specific remediation and stratified management of contaminated croplands.

Data availability

The datasets used and/or analysis during the current study available from the corresponding author on reasonable request.

Received: 12 March 2025; Accepted: 18 August 2025

Published online: 01 September 2025

References

1. Foley, J. A. et al. Solutions for a cultivated planet. *Nature* **478**, 337–342 (2011).
2. Yang, Q. et al. A review of soil heavy metal pollution from industrial and agricultural regions in China: Pollution and risk assessment. *Sci. Total Environ.* **642**, 690–700 (2018).
3. Jansson, J. K. & Hofmockel, K. S. Soil microbiomes and climate change. *Nat. Rev. Microbiol.* **18**, 35–46 (2020).
4. Mu, D. et al. Source apportionment, source-specific health risks, and control factors of heavy metals in water bodies of a typical karst basin in southwestern China. *PLoS ONE* **19**, e0309142 (2024).
5. Su, K. et al. Heavy metals concentrations in commercial organic fertilizers and the potential risk of fertilization into soils. *Sci. Rep.* **15**, 1230 (2025).
6. Khan, S. et al. Global soil pollution by toxic elements: current status and future perspectives on the risk assessment and remediation strategies—A review. *J. Hazard. Mater.* **417**, 126039 (2021).
7. Wu, Y. et al. Review of soil heavy metal pollution in China: spatial distribution, primary sources, and remediation alternatives. *Resour. Conserv. Recycl.* **181**, 106261 (2022).
8. Ministry of Environmental Protection of the People's Republic of China. *National Soil Contamination Survey Report*. (2014).
9. Wang, M. et al. A modified model for quantitative heavy metal source apportionment and pollution pathway identification. *Toxics* **12**, 382 (2024).
10. Huang, J. et al. Source apportionment and spatial and quantitative ecological risk assessment of heavy metals in soils from a typical Chinese agricultural county. *Process Saf. Environ. Prot.* **126**, 339–347 (2019).
11. Wuepper, D., Borrelli, P. & Finger, R. Countries and the global rate of soil erosion. *Nat. Sustain.* **3**, 51–55 (2020).

12. Xiang, M. et al. Heavy metal contamination risk assessment and correlation analysis of heavy metal contents in soil and crops. *Environ. Pollut.* **278**, 116911 (2021).
13. Jia, Z., Cai, A., Li, R., Wang, X. & Liu, Y. Knowledge map and hotspot analysis in source appointment of heavy metals from 1994 to 2022: A scientometric review. *Front. Environ. Sci.* **12**, 1443633 (2024).
14. Xia, F., Zhao, Z., Niu, X. & Wang, Z. Integrated pollution analysis, pollution area identification and source apportionment of heavy metal contamination in agricultural soil. *J. Hazard. Mater.* **465**, 133215 (2024).
15. Chen, W., Li, Q., Wang, Z. & Sun, Z. Spatial distribution characteristics and pollution evaluation of heavy metals in arable land soil of China. *Environ. Sci. (China)* **41**, 2822–2833 (2020).
16. Hu, W. et al. Source identification of heavy metals in peri-urban agricultural soils of southeast China: An integrated approach. *Environ. Pollut.* **237**, 650–661 (2018).
17. Borůvka, L., Vacek, O. & Jehlička, J. Principal component analysis as a tool to indicate the origin of potentially toxic elements in soils. *Geoderma* **128**, 289–300 (2005).
18. Chen, W. et al. Evaluation and source analysis of heavy metal pollution in cultivated soil of a lead–zinc mining area in Huize County Yunnan Province China. *J. Agro-Environ. Sci.* **43**, 1036–1044 (2024).
19. Shen, Z. et al. Pollution source apportionment of heavy metals in cultivated soil around a Red Mud Yard based on APCS-MLR and PMF models. *Environ. Sci. (China)* **45**, 1058–1068 (2024).
20. Shi, B. et al. The main strategies for soil pollution apportionment: a review of the numerical methods. *J. Environ. Sci.* **136**, 95–109 (2024).
21. Huang, Y. et al. A modified receptor model for source apportionment of heavy metal pollution in soil. *J. Hazard. Mater.* **354**, 161–169 (2018).
22. Wang, C. et al. Impact of coking plant to heavy metal characteristics in groundwater of surrounding areas: Spatial distribution, source apportionment and risk assessments. *J. Environ. Sci.* **149**, 688–698 (2025).
23. Liu, D., Meng, F., Duan, H., Li, Y. & Ma, J. Analysis of heavy metal sources in farmland soil of sewage irrigation and industrial complex area based on APCS-MLR and PMF. *Environ. Sci. (China)* **45**, 4812–4824 (2024).
24. Liu, Y. et al. Enrichment and environmental availability of cadmium in agricultural soils developed on cd-rich black shale in southwestern China. *Environ. Sci. Pollut. Res.* **29**, 36243–36254 (2022).
25. Hu, P. et al. Research progress on the causes, risks, and control of high geological background of heavy metals in soils. *Acta Pedologica Sin.* **60**, 1363–1377 (2023).
26. Shifaw, E. Review of heavy metals pollution in China in agricultural and urban soils. *J. Health Pollut.* **8**, 180607 (2018).
27. Mu, L. et al. The effects of the combined application of biochar and phosphogypsum on the physicochemical properties of cd-contaminated soil and the yield quality of Chinese cabbage. *Agriculture* **14**, 1865 (2024).
28. Daferera, O., Soulis, K. X., Kairis, O. & Gasparatos, D. Environmental and health risk assessment of potentially toxic elements and spatial distribution in soils of Lavrio Mining area, Greece: A meta-synthesis. *Environ. Earth Sci.* **84**, 265 (2025).
29. Mir, I. A., SreePrabash, M., Sridhar, V. & Maruthi, K. V. Distribution, sources and toxicity of heavy metals in surface sediments of north western Karnataka, south India. *Sci. Rep.* **12**, 15782 (2022).
30. Hoque, M. M. et al. Assessment of soil heavy metal pollution and associated ecological risk of agriculture dominated mid-channel bars in a subtropical River Basin. *Sci. Rep.* **13**, 11104 (2023).
31. China National Environmental Monitoring Centre. Background values of soil elements in China (1990).
32. Henry, R. C. & Hidy, G. M. Multivariate analysis of particulate sulfate and other air quality variables by principal components-part I: Annual data from Los Angeles and New York. *Atmos. Environ.* **1967**(13), 1581–1596 (1979).
33. Fei, X., Lou, Z., Xiao, R., Ren, Z. & Lv, X. Contamination assessment and source apportionment of heavy metals in agricultural soil through the synthesis of PMF and GeogDetector models. *Sci. Total Environ.* **747**, 141293 (2020).
34. Huang, H., Hu, G., Yu, R., Hao, C. & Chen, F. Source appointment of heavy metals in agricultural soils of the Jiulong River Basin based on positive matrix factorization. *Environ. Sci. (China)* **41**, 430–437 (2020).
35. Zeng, T. et al. Source apportionment of soil heavy metals with PMF model and pb isotopes in an intermountain basin of Tianshan mountains, China. *Sci. Rep.* **12**, 19429 (2022).
36. Zhang, J. et al. Evaluation of heavy metal contamination and associated human health risk in soils around a battery industrial zone in Henan province, central China. *Agriculture* **14**, 804 (2024).
37. Yu, L. et al. Spatial distribution, source apportionment, and ecological risk assessment of soil heavy metals in Jianghugongmi Producing Area, Shandong Province. *Environ. Sci. (China)* **43**, 4199–4211 (2022).
38. Lv, J., Liu, Y., Zhang, Z. & Dai, J. Factorial kriging and stepwise regression approach to identify environmental factors influencing spatial multi-scale variability of heavy metals in soils. *J. Hazard. Mater.* **261**, 387–397 (2013).
39. Wang, Q. et al. Source identification and spatial distribution of soil heavy metals in Western Yunnan. *China Environ. Sci.* **41**, 3693–3703 (2021).
40. Ariya, P. A., Khalizov, A. & Gidas, A. Reactions of gaseous mercury with atomic and molecular halogens: Kinetics, product studies, and atmospheric implications. *J. Phys. Chem. A* **106**, 7310–7320 (2002).
41. Wei, M., Pan, A., Ma, R. & Wang, H. Distribution characteristics, source analysis and health risk assessment of heavy metals in farmland soil in Shiquan county, Shaanxi province. *Process Saf. Environ. Prot.* **171**, 225–237 (2023).
42. Šajin, R., Halamić, J., Peh, Z., Galović, L. & Alijagić, J. Assessment of the natural and anthropogenic sources of chemical elements in alluvial soils from the drava river using multivariate statistical methods. *J. Geochem. Explor.* **110**, 278–289 (2011).
43. Zeng, W., Wan, X. & Lei, M. Heavy metal contents of soil and surface dust and Its ecological risk analysis in a multifunctional industrial park. *Environ. Sci. (China)* **42**, 1105–1113 (2021).
44. Liu, F. et al. Spatial variations, health risk assessment, and source apportionment of soil heavy metals in the middle Yellow River Basin of northern China. *J. Geochem. Explor.* **252**, 107275 (2023).
45. Lv, J. et al. Identifying the origins and spatial distributions of heavy metals in soils of Ju country (eastern China) using multivariate and geostatistical approach. *J. Soils Sediments* **15**, 163–178 (2015).
46. Melfos, V. & Voudouris, PCh. Geological, mineralogical and geochemical aspects for critical and rare metals in Greece. *Minerals* **2**, 300–317 (2012).
47. Alloway, B. Heavy Metals in Soils: Trace Metals and Metalloids in Soils and Their Bioavailability (2013).
48. Mackay, A. K., Taylor, M. P., Munksgaard, N. C., Hudson-Edwards, K. A. & Burn-Nunes, L. Identification of environmental lead sources and pathways in a mining and smelting town: Mount Isa, Australia. *Environ. Pollut.* **180**, 304–311 (2013).
49. Callender, E. & Rice, K. C. The urban environmental gradient: anthropogenic influences on the spatial and temporal distributions of lead and zinc in sediments. *Environ. Sci. Technol.* **34**, 232–238 (2000).
50. Krishnan, A., Nadia, A. & Chong, M. Ecological risk assessment of heavy metal pollution in mangrove sediments of the Sepang Besar River, West Coast Peninsular Malaysia. *Environ. Ecol. Res.* **10**, 497–507 (2022).
51. Duan, H. et al. Pollution characteristics and source analysis of heavy metals in soils in Yellow River Cultural Park based on APCS-MLR and PMF receptor model. *Environ. Sci. (China)* **44**, 4406–4415 (2023).
52. Li, J. et al. Multiproxy comprehensive analysis for source apportionment and pollution of heavy metals in urban drinking-water source soils from the Lanzhou Reach of the Yellow River. *Environ. Sci. (China)* **45**, 6723–6733 (2024).
53. Wu, S. et al. Cr and Ni geochemistry and some suggestions on soil pollution risk prevention control in the Zhelong-Gasa area, Yunnan Province. *Geophys. Geochem. Explor.* **45**, 517–527 (2021).

54. Wang, Y., Jiang, Z., Nie, W. & Zhang, J. Phytoremediation of the soil contaminated by Pb, Cd and secondary salinization with the enhancement of EDTA. *J. Agro-Environ. Sci.* **37**, 1866–1874 (2018).
55. Jiang, Z., Yang, S. & Luo, S. Source analysis and health risk assessment of heavy metals in agricultural land of multi-mineral mining and smelting area in the karst region—A case study of Jichangpo town, southwest China. *Heliyon* **9**, e17246 (2023).

Acknowledgements

All samples were tested by the Yunnan Soil Fertility and Pollution Remediation Research Center, and Naiming Zhang provided valuable advice on the source apportionment of heavy metals. We would like to express our gratitude to our research group colleagues for their support, suggestions, and invaluable input.

Author contributions

A.B. wrote the main manuscript text and prepared the figures; C. provided funding support and project management; D. contributed to the conceptualization of the manuscript; E. designed the research plan and performed data analysis; F. conducted data analysis; G.H. carried out experimental operations and data collection; I. reviewed the manuscript; J. contributed to the conceptualization, provided funding support, and reviewed the manuscript. All authors reviewed and revised the manuscript.

Funding

This work was supported by the Yunnan Provincial Science and Technology Talent and Platform Program (Project Number: 202405AM340004) and the Scientific Research Foundation of the Yunnan Provincial Department of Education (Project Number: 2024J1098).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to N.Z.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025