Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Spatial Heterogeneity in Soil Contamination and Ecological Risk Near Copper Mining Sites in the Zambian Copperbelt

Elias Nyanza
Catholic University of Health and Allied Sciences
Paula Fernandez Musso
University of Oxford
Mulu Gebreyohannes
Barnard College
Obrian Ndhlovu
University of Zambia
Anja Benshaul-Tolonen
atolonen@barnard.edu

Barnard College

Article

Keywords:

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11393517/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Spatial Heterogeneity in Soil Contamination and Ecological Risk Near Copper Mining Sites in the Zambian Copperbelt

Elias C. Nyanza[1], Paula Fernandez Musso[2], Mulu Gebreyohannes[3], Obrian Ndhlovu[4], and Anja Benshaul-Tolonen[3,*]

[1]Department of Community Medicine, Environmental and Occupational Health, School of Public Health, Catholic University of Health and Allied Sciences, Mwanza, Tanzania
[2]Department of Economics, University of Oxford, Oxford, UK
[3]Department of Economics, Barnard College, Columbia University, New York City, USA
[4]Department of Economics, University of Zambia, Lusaka, Zambia
[*]Corresponding author: atolonen@barnard.edu

ABSTRACT

The Zambian Copperbelt Province, part of the world's largest sedimentary rock-hosted copper province, has experienced over a century of intensive mining activities, raising concerns about the accumulation of potentially toxic elements (PTEs) in soils. This study performed dry-season soil sampling in human-use areas near selected active or legacy copper mining sites. Fifty-nine samples were collected from the 10–30 cm depth interval and analysed by ICP-SFMS. Contamination and ecological-risk indices were calculated from raw concentrations against Upper Continental Crust background values. Spatial clustering was assessed using Moran's I and local indicators of spatial association (LISA). Copper was the dominant contaminant (mean 512 mg kg^{-1}; 18.3× background), with 93% of samples exceeding Canadian agricultural guidelines. The mean Pollution Load Index was 0.98 (31% of sites polluted), while the mean ecological Risk Index was 184 (moderate class), driven by Cu (≈50%) and, despite their lower concentrations, Cd (≈25%) and Hg (≈13%) owing to their high toxic-response factors. LISA analysis identified High-High hotspots for Cu, Co, Pb, and As in Wusakile, co-located with active mining and the Black Mountain slag heap. The findings reveal spatial heterogeneity in PTE contamination, suggesting that spatially targeted remediation may yield greater marginal benefits than uniformly distributed interventions.

Key Words: *Copper mining, potentially toxic elements (or PTEs), soil contamination, spatial hotspots, ecological risk assessment, Zambia Copperbelt*

Introduction

Copper is indispensable to the global green energy transition, underpinning electric vehicles (EVs), renewable power infrastructure, and energy storage technologies. As the world's seventh-largest copper producer[1], and with a geography indicating large untapped reserves and secondary levels of the critical mineral cobalt, Zambia is well positioned to benefit from the rapidly growing global demand. This opportunity is particularly significant given the mining sector's central role in the national economy, having contributed between 11 and 19% of national GDP over the past decade[2], while copper alone accounted for approximately 65% of export earnings, with the sector continuing to drive employment and infrastructure development[3]. Recognizing both the country's resource endowment and the sector's economic significance, the government has set an ambitious target of increasing annual copper production to 3 million metric tonnes by 2031, more than tripling recent production levels[4]. The country took a step toward this goal in 2025, achieving a record production of 890,346 metric tonnes, a 7.9% increase over 2024 output, though this still fell short of the interim milestone of 1 million tonnes[3].

Realizing this growth ultimately depends on geology. Zambia's resource wealth is concentrated in the Katangan basins of the Central African Copperbelt, Neoproterozoic sedimentary systems hosting the world's largest sedimentary rock-hosted copper province[5,6]. Mineralization in the Roan Group was controlled by basin architecture, evaporite distribution, and halokinetic processes[7]. Mining and processing of these copper- and cobalt-rich ores have generated large volumes of tailings and other mine wastes in which weathering and mineralogical transformations can influence the mobility of potentially toxic elements[8].

Over more than a century, the extraction and processing of these ores have released metal-laden dust, acid mine drainage (AMD), sulfur dioxide emissions, while generating large volumes of mine waste enriched in potentially toxic elements (PTEs),

including arsenic (As), cadmium (Cd), cobalt (Co), copper (Cu), iron (Fe), mercury (Hg), manganese (Mn), nickel (Ni), lead (Pb), selenium (Se), and zinc (Zn). These contaminants have been documented in soils, surface water, groundwater, and sediments across the Copperbelt[8–11]. Reported copper concentrations in tailings reach up to ≈12,223 mg kg^{-1} near the major mining centers of Kitwe and Mufulira[9,10], while arsenic concentrations in groundwater exceed WHO drinking-water guidelines in Kitwe[12]. From these source areas, the Kafue River—the dominant regional drainage traversing the mining corridor—carries contaminants further downstream, affecting aquatic ecosystems and agricultural lands[11,13]. The vulnerability of this system was starkly demonstrated by the February 2025 collapse of the Sino-Metals Leach tailings dam, which released approximately 50 million liters of contaminated effluent into the Kafue River, causing widespread aquatic mortality and severe contamination of riparian pastures and sediments[14]. These risks can be expected to intensify as mining activity continues to expand.

Environmental contamination ultimately translates into human exposure, particularly for communities living adjacent to mines, smelters, and tailings facilities. These populations face multiple exposure pathways, including inhalation of contaminated dust, ingestion of PTEs through locally grown crops and fish, and consumption of polluted drinking water[15]. Bio-monitoring studies have revealed substantial health burdens: children near a decommissioned lead mine in Kabwe, Central Province, frequently exceed 10 $\mu g\,dL^{-1}$, with approximately half at or above 45 $\mu g\,dL^{-1}$[16], far above the 5 $\mu g\,dL^{-1}$ level at which the World Health Organization recommends action to identify and reduce exposure[17], raising serious concerns about neuro-developmental impairment, and adults near copper mines in Zambia had higher concentrations of cobalt, copper, lead and selenium in toenail samples than control populations[18]. These findings are consistent with broader evidence linking mining-related pollution to increased respiratory disease, chronic metal toxicity, and long-term health complications among mining-adjacent populations across the Copperbelt and other Zambian mining districts[15,19,20].

Despite this well-documented health burden, effective mitigation remains challenging because PTE contamination can originate from multiple geogenic and anthropogenic sources and may be highly localized. Identifying the spatial distribution and potential sources of contamination is therefore important for developing effective mitigation strategies[21]. This is particularly relevant in the Copperbelt, where active and legacy mining infrastructure is interspersed with densely populated communities and agricultural land[22,23], and as copper production expands toward the 2031 target.

Addressing this gap requires spatially explicit, multi-element assessments that identify contamination hotspots, characterize patterns consistent with potential sources, quantify ecological risk, and delineate priority areas for human-exposure follow-up. Previous work in the Copperbelt has largely focused on site-specific tailings characterization, single-compartment studies of soil, water or sediment, or health surveys at individual locations[9,10,24,25]. These studies offer important but spatially limited insights, leaving landscape-scale patterns, multi-element co-occurrence, and the relative contribution of different sources across mining districts comparatively unexplored.

This study addresses that gap by characterizing the distribution, contamination status, and spatial clustering of eight PTEs (As, Cd, Co, Cu, Hg, Ni, Pb, and Zn) in near-surface (10–30 cm) soils of three adjoining districts, Chingola, Kalulushi and Kitwe, that together form the core of Zambia's Copperbelt copper-cobalt mining corridor. Using a dry-season sampling design in human-use areas, quality-controlled ICP-SFMS analysis, established contamination and ecological-risk indices, and geostatistical hotspot detection via local indicators of spatial association, we aim to (i) quantify the magnitude and spatial variability of PTE enrichment relative to geochemical background and agricultural screening guidelines, (ii) evaluate toxicity-weighted ecological risk and identify which elements drive it, and (iii) locate statistically significant multi-element hotspots that can serve as priority areas for targeted remediation, enhanced monitoring, and subsequent human-exposure assessment as Zambia pursues its copper production goals.

Methods

Study area

Fieldwork was conducted in Zambia's Copperbelt Province, the historic heart of its copper-cobalt mining industry[26]. The region's mineral deposits are primarily sediment-hosted stratiform deposits of the Roan Group[6,7], with ores processed through flotation and smelting for sulfide ores and acid leaching followed by solvent extraction–electrowinning (SX/EW) for oxide ores[27]. With a population of 2.7 million people according to the 2022 Zambia Census[28], the province contains Zambia's largest concentration of people living alongside mining activity, underscoring the scale of potential human exposure to mining-related contamination.

Within the Copperbelt Province, the study focused on three adjoining districts—Chingola, Kalulushi, and Kitwe—selected because they contain among the highest densities of major active and legacy copper-cobalt mining operations in the province[29] and support large human populations living in close proximity to mining infrastructure. Sampling covered all three districts (Chingola, n = 21; Kalulushi, n = 23; Kitwe, n = 15), focusing on human-use areas where mining activity and human settlement are most densely intermingled. Figure 1b situates the study area in its regional hydrological context, highlighting the Kafue River as the principal pathway for potential off-site contaminant transport.

Ethical approval

All methods were carried out in accordance with relevant guidelines and regulations, further described below.

This study forms part of the broader research project, *Measuring and Mitigating Exposure to Toxic Metals in Copper Mining Areas in Zambia*. Ethical approval for the soil samples and focus group research was obtained from Barnard College (USA; approval no. 2425-1110-092) and ERES Converge (Zambia; approval no. 2025-08-08). Research authorization was granted by the Zambia National Health Research Authority (research license no. NHRA-2491/18/07/2025).

The research team sought approval from the Ministry of Health at the national level, the provincial office in Ndola, and the district offices in Kitwe, Kalulushi, and Chingola to obtain stamped authorization. In addition, we received support for the study from the Ministry of Mines and Minerals Development (MMMD) and the Mine Safety Department in Ndola. A representative from MMMD accompanied the research team during sampling. The Ministry of Local Government and Rural Development, through the Office of the Town Clerk in each research district (Kitwe, Kalulushi, and Chingola), provided authorization to access municipal land. These offices also facilitated introductions to members of the Ward Development Committees (WDCs). Consistent with the institutional context of Copperbelt Province, customary authorities did not require formal authorization, although the team met with headmen in some communities to introduce the research as a matter of courtesy. Before fieldwork, community engagement was undertaken through governmental authorities in accordance with established participatory protocols[30].

Soil samples were collected from the study sites with the relevant permission or authorization from the landowners or authorities.

Participants in the focus group discussions provided oral informed rather than written consent, in accordance with the ethics approval, due to literacy barriers and the need to preserve anonymity.

Field sampling

Soil sampling and focus group discussions were conducted in eight wards selected in collaboration with the local Offices of the Town Clerk. The research team identified areas of interest based on their proximity to large-scale mining operations (including mining, smelting, refining, and waste disposal) and the ownership of these operations (e.g., Chinese-owned, state-owned, and Indian-owned). Government representatives then proposed specific wards and corresponding local contacts, typically chairpersons of the Ward Development Committees (WDCs).

A total of 59 soil samples were collected from residential, agricultural, and community lands in close proximity to 24 mine sites across the districts of Chingola (n = 21), Kalulushi (n = 23), and Kitwe (n = 15). The 59 soil samples were collected from eight spatial clusters near the eight locations where concurrent focus group discussions were conducted[31]. On a few occasions, the research team did not receive permission from the rightful landowners to collect samples; these locations were therefore neither sampled nor recorded to protect privacy.

We identify the sampling clusters using letter codes and report the corresponding administrative wards in Table 1. Because environmental contamination can extend across administrative boundaries, soil sampling was not restricted to the wards in which the focus group discussions took place. Thus, the sampling sites and focus group locations were geographically proximate but did not necessarily fall within the same administrative wards.

Sampling was conducted during the dry season (August 2025) to minimize the influence of seasonal runoff and the mobilization of mining waste. At each site, the surface litter layer was removed before soil was collected from the 10–30 cm depth interval (plough zone), consistent with guidelines for assessing contaminated sites with agricultural or surface-exposure relevance[32]. Given the targeted sampling of areas used by people near mining operations, the findings characterize contamination at the sampled locations and should not be interpreted as representative estimates of soil contamination across the broader Copperbelt region.

Laboratory analysis and quality control

Sample bags were labeled immediately after collection using permanent markers. Samples were subsequently dried, sieved to <2 mm, homogenized, and pulverized. To prevent cross-contamination, the soil auger was thoroughly cleansed with 18.2 $M\Omega \cdot cm$ ultrapure water between sampling locations. All field samples were analyzed at ALS Scandinavia AB (Luleå, Sweden), an ISO/IEC 17025-accredited laboratory, using Inductively Coupled Plasma Sector Field Mass Spectrometry (ICP-SFMS).

Analytical quality was assessed through two complementary exercises. First, six preparation blanks were carried through the full digestion and analysis sequence to establish detection limits. Two measures were calculated: the laboratory limit of detection (LOD), defined as three times the standard deviation of the blank measurements, and the method detection limit (MDL), defined as the blank mean plus three standard deviations[33,34] (Supplementary Table S1). Second, short-term analytical repeatability was assessed using six independently digested duplicate field samples (KL01, LW11, LG21, MS31, ND41, KB52), following established repeatability guidance[35,36]. For each element, the difference between paired measurements was used to assess whether the duplicates showed systematic bias. The mean difference and its 95% confidence interval were calculated using the t-distribution (df = 5), with an interval containing zero indicating no evidence of systematic bias[37]. Agreement

Table 1. Sampling localities in Copperbelt Province, Zambia

Locality	Reference Ward	District	n
A	Wusakile	Kitwe	6
B	Luangwa	Kitwe	9
C	Luapula	Kalulushi	5
D	Ichimpe	Kalulushi	10
E	Lulamba	Kalulushi	4
F	Musakashi	Kalulushi	4
G	Kabundi	Chingola	10
H	Maiteneke	Chingola	11

Note: Localities are lettered A–H by district (Kitwe, Kalulushi, Chingola), from Kitwe in the southeast to Chingola in the northwest. The reference ward column identifies the administrative ward in which the concurrent focus group discussions (FGDs) were conducted[31]. Soil sampling was conducted near these locations but was not restricted to the reference ward boundaries.

between duplicate measurements was further assessed using the intraclass correlation coefficient, ICC(C,1)[38], with values above 0.90 interpreted as excellent reliability[39]. Analytical precision was also summarized using the repeatability coefficient, $RC = 1.96\sqrt{2}s_w$, which represents the difference below which 95% of paired replicate measurements are expected to fall[37] (Supplementary Table S2). Finally, the mean concentration of the duplicated samples was compared with the MDL to determine whether measured concentrations were sufficiently above the blank-based detection limits (Supplementary Table S3).

Preparation-blank concentrations were low for As, Cd, Co and Hg (means 0.003–0.009 mg kg^{-1}) and higher for Pb, Ni, Cu and especially Zn (0.776 mg kg^{-1}). Zn was also the most variable blank (SD 0.890 mg kg^{-1}) and therefore had the highest detection limits (Supplementary Table S1). Duplicate analyses showed no systematic bias for any element: all 95% confidence intervals for the mean replicate differences included zero (paired t-tests, all $p > 0.20$). Intraclass correlation coefficients ranged from 0.930 (Hg) to 0.9997 (Cu), while repeatability coefficients were small relative to the observed concentration ranges (Supplementary Table S2). Measured concentrations exceeded the MDL by approximately 23- to 840-fold for As, Cd, Co, Cu, Ni, Pb and Zn, indicating negligible blank contribution (Supplementary Table S3). Hg was the sole exception: its mean concentration in the duplicated samples was only about 1.5 times the MDL and 2.4 times the laboratory LOD. Hg also had the widest reliability interval of the eight elements (ICC 0.930, 95% CI 0.588–0.990). Taken together, these quality-control analyses support the fitness-for-purpose of the soil digestion and ICP-SFMS procedure used, with the caveat that mercury concentrations lie near the method's working limit and should therefore be interpreted accordingly.

Descriptive statistics and inter-element associations

Descriptive statistics, including measures of central tendency, dispersion, and coefficients of variation, were calculated from the raw concentration data (mg kg^{-1}). Distributional properties were assessed using skewness coefficients and Shapiro–Wilk tests applied to both raw and $\log_{10}$-transformed concentration data. The raw data exhibited strong right-skewness, while $\log_{10}$ transformation substantially improved the approximation to normality and stabilized variance[40].

Inter-element associations were therefore examined using log10-transformed concentrations using three complementary techniques: (i) Pearson product-moment correlation analysis, with Spearman rank correlation used as a non-parametric robustness check[41]; (ii) principal component analysis (PCA) performed following column standardization (z-scoring), with components retained according to the Kaiser criterion (eigenvalues > 1)[42,43]; and (iii) hierarchical cluster analysis using average linkage and dissimilarity measure of $1 - |\rho|$, where ρ denotes the Pearson correlation coefficient[44]. Together, these approaches were used to identify inter-element relationships and potential common sources of the PTEs in the soils.

Contamination and ecological risk indices

Contamination severity and ecological risk were evaluated using five established indices (Table 2): the contamination factor (CF), geo-accumulation index (I_{geo}) and single-factor ecological risk (E_r) as single-element measures, and the pollution load index (PLI) and potential ecological risk index (RI) as integrated multi-element measures. All indices were calculated from raw, unstandardized PTE concentrations (mg kg^{-1}) against Upper Continental Crust (UCC) background values[45]. In the absence of national soil quality guidelines in Zambia, measured soil PTE concentrations were compared with the Canadian Council of Ministers of the Environment (CCME)[46] soil quality guidelines for agricultural land use, which were used as screening benchmarks to identify samples at potential risk to agriculture and human health.

The contamination and ecological risk assessments were further evaluated through two sensitivity analyses: (1) all indices were recalculated using Average Shale background values[47] in place of UCC values to assess the influence of background selection and (2) a leave-one-out analysis of the PLI was performed by sequentially excluding individual elements to assess their relative impact on the overall PLI.

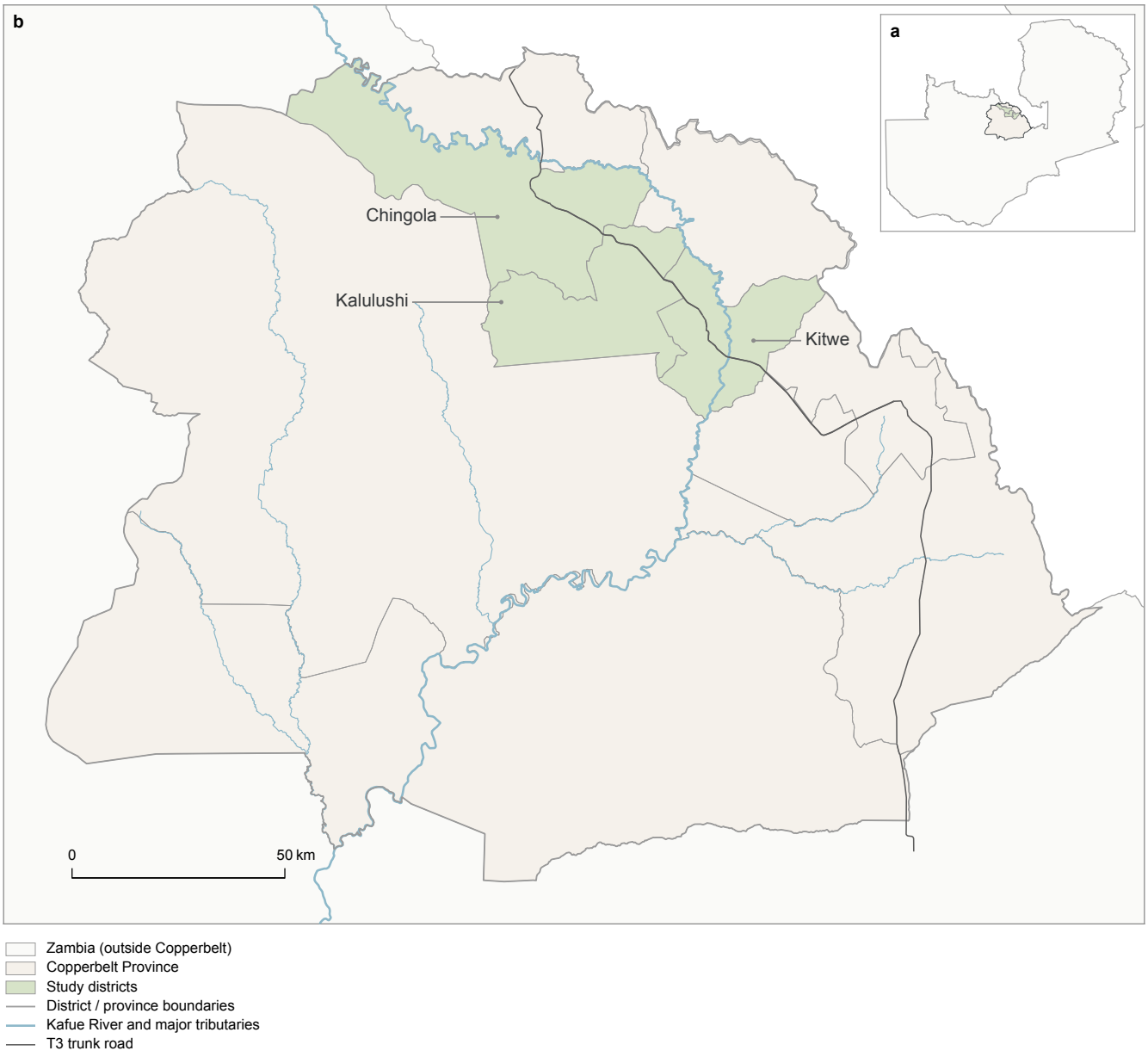


Figure 1. Location of the study area. (a) Location of Copperbelt Province and the three study districts within Zambia (inset). (b) Copperbelt Province, showing district and provincial boundaries, the three study districts (Chingola, Kalulushi and Kitwe), the Kafue River and its major tributaries, and the T3 trunk road.

Table 2. Formulas and classification criteria for the contamination and ecological risk indices

Index	Formula	Classes
Contamination factor (CF)[48]	C_i/B_i	<1 (low) 1–3 (moderate) 3–6 (considerable) ≥ 6 (very high)
Geo-accumulation index (I_{geo})[49]	$\log_2 [C_i/(1.5\,B_i)]$	≤ 0 (unpolluted) 0–1 (unpolluted to moderate) 1–2 (moderate) 2–3 (moderate to strong) 3–4 (strong) 4–5 (strong to extreme) >5 (extreme)
Pollution load index (PLI)[50]	$(\prod_{i=1}^{n} \mathrm{CF}_i)^{1/n}$	≤ 1 (unpolluted) >1 (polluted)
Single-factor ecological risk (E_r^i)[48]	$T_r^i \times \mathrm{CF}_i$	<40 (low) 40–80 (moderate) 80–160 (considerable) 160–320 (high) ≥ 320 (very high)
Potential ecological risk index (RI)[48]	$\sum_{i=1}^{n} E_r^i$	<150 (low) 150–300 (moderate) 300–600 (considerable) ≥ 600 (very high)

Note: C_i = measured concentration of element i in soil (mg kg^{-1}); B_i = geochemical background concentration of element i, taken from Upper Continental Crust (UCC) values[45]; n = number of elements ($n = 8$); T_r^i = toxic-response factor of element i (Hg = 40, Cd = 30, As = 10, Co = Cu = Ni = Pb = 5, Zn = 1)[48].

Spatial analysis

Spatial analyses were conducted using the 59 geo-referenced soil sampling locations collected at eight sampling localities (A–H; Table 1), which defined the study extent. Sampling coordinates, recorded using a handheld GPS device, were imported into QGIS 3.28 together with administrative boundary shapefiles and projected to the WGS 84 / UTM Zone 35S coordinate reference system (EPSG:32735).

Sample locations were mapped over ward boundaries to visualize the spatial distribution of contamination. Continuous concentration surfaces for priority PTEs were generated using inverse-distance weighting (IDW; power = 2, 10 nearest neighbors) on $\log_{10}$-transformed concentrations in projected coordinates, with interpolated surfaces clipped to the study-area boundary. Spatial autocorrelation was evaluated using Global Moran's I, while Local Moran's I (LISA) was used to identify statistically significant clusters and spatial outliers[51]. Both analyses employed row-standardized eight-nearest-neighbor spatial weights and 999 permutations (conditional permutations for LISA).

Results

PTE concentrations

Table 3 summarizes the descriptive statistics and mean contamination factors for the eight PTEs. Copper exhibited by far the strongest enrichment (18.3× UCC), whereas Pb, Cd, and Co were only modestly enriched (1.5–1.6× UCC). The remaining PTEs occurred at or below UCC background concentrations. Variability among samples was high, with coefficients of variation exceeding 100% for Pb (174%), Cu (161%), Hg (138%), Zn (122%), and Co (120%), consistent with spatially heterogeneous, localized inputs. The raw concentration data exhibited strong right-skewness and approximate log-normal distributions (Supplementary Figure S1), supporting the use of $\log_{10}$ transformation in subsequent analyses. Spatial variability was pronounced, with the highest concentrations of most elements observed at locality A, roughly corresponding to Wusakile (Figure 2).

Table 3. Descriptive statistics of potentially toxic element concentrations in soils (mg kg^{-1}; $n = 59$)

Element	UCC	Mean	Median	CV (%)	Mean CF
Cu	28	512	253	161	18.30
Pb	17	26.8	10.6	174	1.58
Cd	0.09	0.14	0.09	93	1.53
Co	17.3	25.9	13.6	120	1.50
Zn	67	62.5	27.2	122	0.93
Hg	0.05	0.03	0.02	138	0.62
As	4.8	1.74	1.17	89	0.36
Ni	47	14.7	9.77	86	0.31

Note: UCC = Upper Continental Crust background concentration; CV = coefficient of variation; Mean CF = mean contamination factor (mean concentration/UCC). Elements are ranked by decreasing mean CF.

Inter-element associations

Pearson product-moment correlation (Figure 3a) analysis revealed strong inter-element associations, including Co–Cu (r = 0.82), Pb–Zn (r = 0.81), As–Cu (r = 0.78), As–Co (r = 0.77), and Cd–Zn (r = 0.77). Spearman rank correlations, used as a non-parametric robustness check, showed the same associations (e.g., Co-Cu $\rho = 0.77$, Pb-Zn $\rho = 0.83$; Supplementary Figure S1). Principal component analysis (PCA) further characterized the multivariate relationship among PTEs (Figure 3b). Two components were retained according to the Kaiser criterion (Supplementary Figure S1), together explaining 80.6% of the total variance. PC1 accounted for 56.2% of variance and exhibited positive loadings for all eight PTEs. PC2 explained a further 24.4% of the variance and contrasted positive Ni-Co loadings against negative Pb-Zn-Cd-Hg loadings. Hierarchical cluster analysis (HCA) broadly reproduced this structure (Figure 3c), separating a closely associated As–Co–Cu group from a second cluster comprising Pb–Zn–Cd–Hg; Ni was the most distinct element and joined the other seven only at the highest linkage level (distance 0.72).

Collectively, the correlation, PCA, and HCA results (Figure 3) identified two broad elemental associations. The first comprised As-Co-Cu, with strong associations among all three elements and a weaker link to Ni (Ni–Co r = 0.60; positive PC2 loadings for both), a pattern consistent with the Katangan lithology and associated mining activities. The second comprised Pb-Zn-Cd-Hg, a pattern potentially consistent with atmospheric deposition and urban activities.

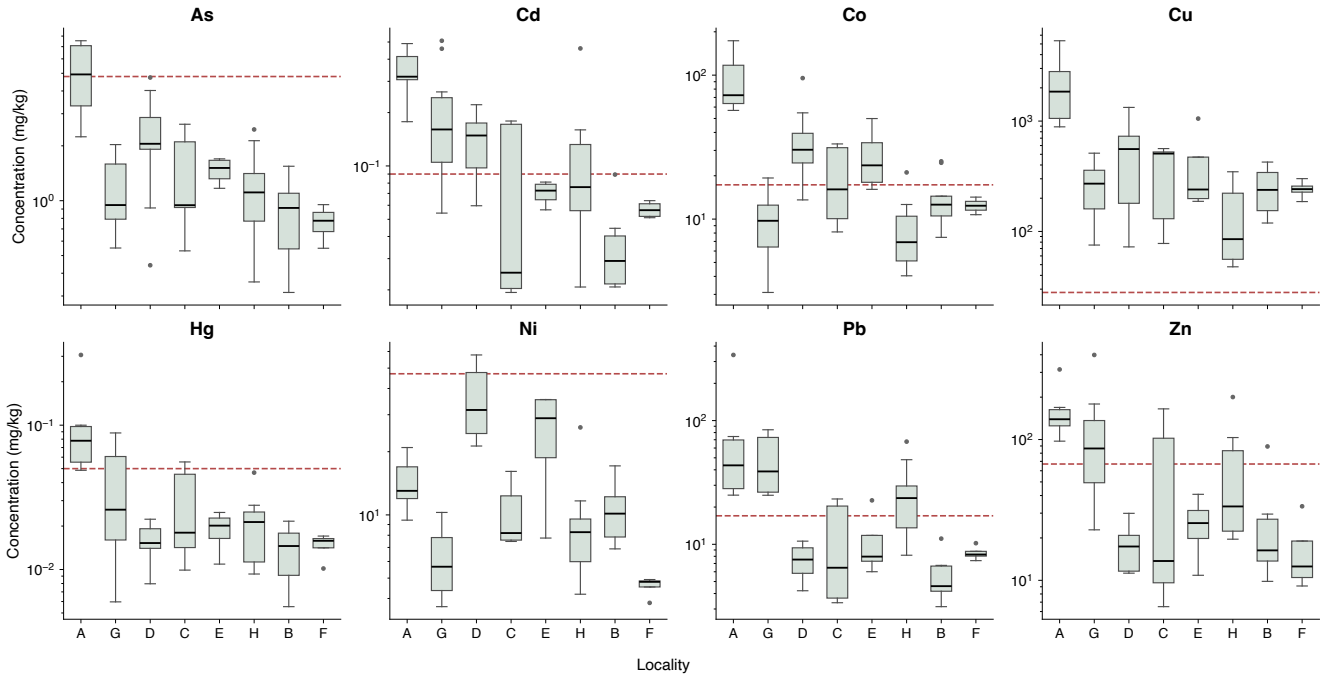


Figure 2. Concentrations of the eight PTEs (As, Cd, Co, Cu, Hg, Ni, Pb, Zn) by sampling locality ($n = 59$), on a logarithmic scale. Localities are lettered A–H by district (Kitwe: A, Wusakile; B, Luangwa; Kalulushi: C, Luapula; D, Ichimpe; E, Lulamba; F, Musakashi; Chingola: G, Kabundi; H, Maiteneke) and are ordered by decreasing mean PLI, as in Table 4. Boxes show the interquartile range (IQR) with the median line; whiskers extend to the most extreme sample within 1.5×IQR of the box, and dots are samples beyond the whiskers. The red dashed line is the Upper Continental Crust background[45]. Wusakile (A) carries the highest Cu, Co and Pb burdens; Ni is highest in Ichimpe (D) and Lulamba (E) rather than in the Wusakile hotspot.

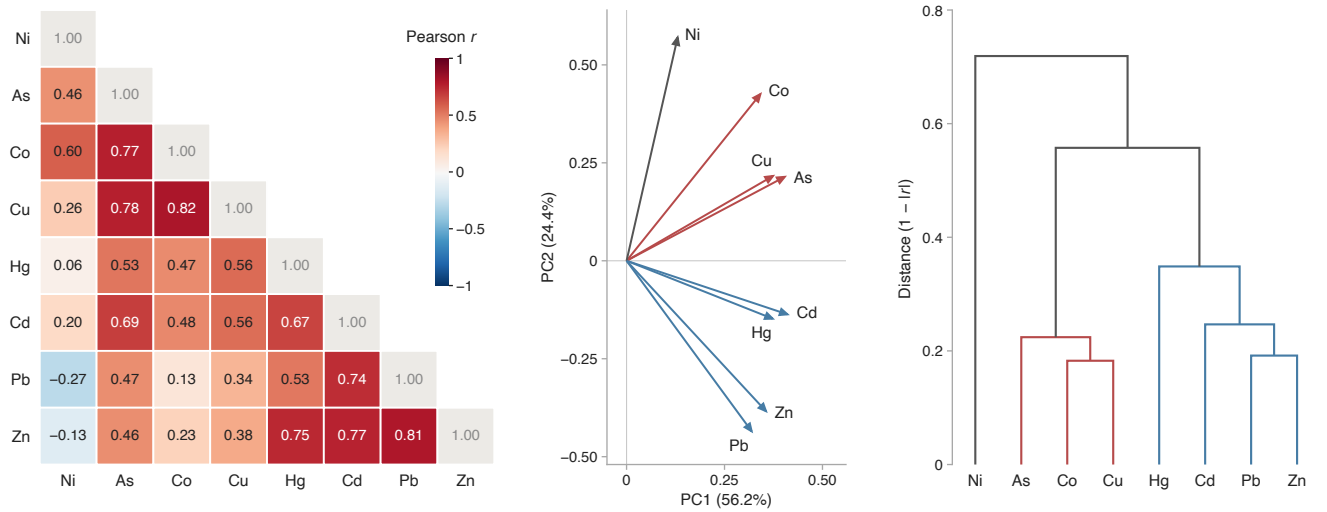


Figure 3. Multivariate analysis of potentially toxic elements ($n = 59$). (a) Pearson correlation coefficients (r) of $\log_{10}$-transformed concentrations (lower triangle); elements are ordered as in (c). (b) Loadings on the first two principal components (PC1, 56.2%; PC2, 24.4% of the variance) from principal component analysis of standardized $\log_{10}$-transformed concentrations. (c) Hierarchical clustering of elements (average linkage; distance $1 - |r|$). In (b) and (c), red marks the As–Co–Cu group and blue the Hg–Cd–Pb–Zn group; Ni (grey) joins the other elements only at the top level of the dendrogram.

Contamination status

Mean contamination factors placed copper in the very high contamination class (CF = 18.3), lead, cadmium and cobalt in the moderate class (CF 1.5–1.6), and zinc, mercury, arsenic and nickel in the low class (CF < 1) (Table 3). The geo-accumulation index (I_{geo}) identified copper as the only element within an elevated contamination class ("moderate-to-strong") (Figure 4a). This pattern was consistent with the CCME agricultural soil quality guideline comparisons, with 93% of samples exceeding the guideline for copper. In contrast, cobalt exceeded its respective guideline at 17% of sites, while exceedances for lead, nickel, and zinc occurred at 8% or fewer sites. No exceedances were observed for arsenic, cadmium, or mercury (Figure 4b).

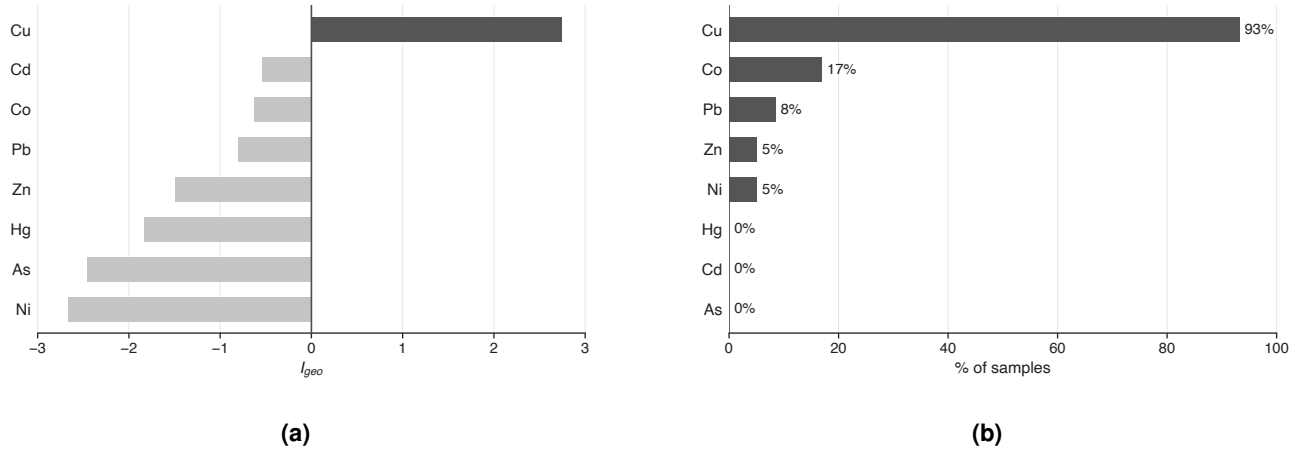


Figure 4. Contamination status of potentially toxic elements. (a) Mean geo-accumulation index (I_{geo}) for each element; Cu, the only element with a positive mean I_{geo}, is shown in dark grey. (b) Percentage of samples exceeding the CCME agricultural soil quality guideline for each element.

Pollution load and ecological risk

The integrated Pollution Load Index (PLI) had a mean value of 0.98 across all sampling sites, with 31% of sites exceeding the polluted threshold (PLI >1). Locality A, roughly corresponding to Wusakile, was the only area with a mean PLI above this threshold (mean = 2.97; Table 4; Figure 5a). Furthermore, every sample collected in locality A was classified as polluted, with the highest reaching 4.03. No other locality showed a mean PLI above 1 or had all of its samples classified as polluted. These results demonstrate that integrated contamination is strongly concentrated in locality A, Wusakile, while most of the study sites remain below the pollution threshold.

The contribution of individual elements to ecological risk reflected both their contamination levels and their toxic-response factors. Copper had the highest mean single-factor ecological risk ($E_r \approx 91.5$), driven by its high contamination factor, whereas cadmium ($E_r \approx 46$) and mercury ($E_r \approx 25$) contributed disproportionately because of their high toxic-response factors (T_r = 30 and 40, respectively) (Table 5; Figure 6a). The mean E_r of all other elements remained in the low-risk range (< 40), although Pb and Co each exceeded $E_r = 40$ at one site (maxima 99.8 and 50.1, respectively).

Copper, cadmium, and mercury together accounted for approximately 88% of the total integrated Risk Index (RI), contributing 49.9%, 24.9% and 13.5%, respectively (Table 5; Figure 6b). Copper reached the moderate single-factor ecological risk threshold ($E_r \geq 40$) at 59% of sampling sites. Despite their comparatively low concentrations, cadmium reached the moderate single-factor ecological risk threshold ($E_r \geq 40$) at 42% of sampling sites, while mercury exceeded this threshold at 17% of sites (Table 5).

The mean RI across all 59 sampling sites was 184 (median 113; range 36–1,273), corresponding to moderate ecological risk. However, risk varied substantially among sites, with a pronounced upper tail: 38 sites were classified as low risk, 12 as moderate, 7 as considerable, and 2 as very high risk (Figure 6c). The maximum RI was 1,273, and both very-high-risk sites were located in locality A, Wusakile (Figure 5b), indicating that the most severe ecological risk was spatially concentrated in this area.

Robustness of contamination and risk indices

The robustness of the contamination and ecological-risk results was assessed by varying the geochemical background and the set of elements included in the PLI calculation.

1. Sensitivity to geochemical background. Recalculating the indices using Average Shale[47] rather than UCC background values reduced the magnitude of the contamination and ecological-risk estimates, but did not alter the overall pattern (Figure

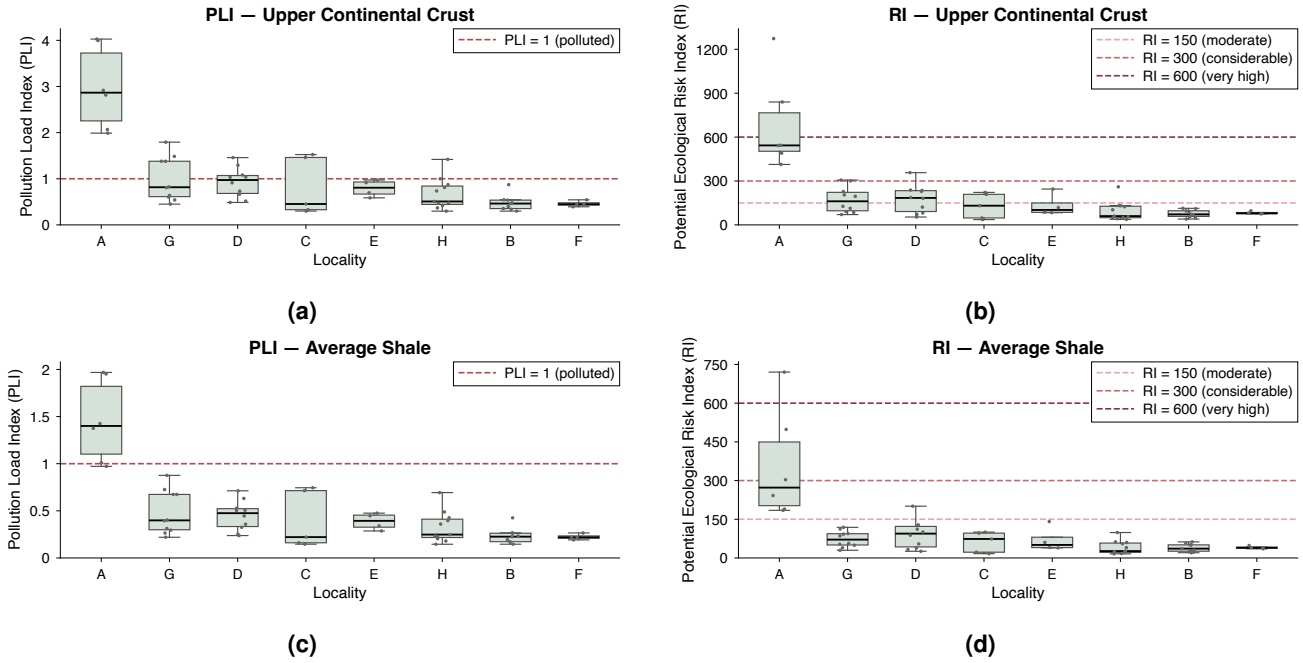


Figure 5. Locality-level PLI and RI under two geochemical backgrounds. (a, b) Upper Continental Crust[45]; (c, d) Average Shale[47]. Boxes show the inter-quartile range with the median; points are individual samples. Localities are lettered as in Figure 2. Red dashed line, PLI = 1 (polluted); light to dark red dashed lines, RI class boundaries at 150, 300 and 600.

Table 4. Pollution load index (PLI) and potential ecological risk index (RI) by locality

Locality	n	Mean PLI	Max PLI	PLI status	Mean RI	Max RI	Max RI class
A	6	2.97	4.03	polluted	684	1273	very high
G	10	0.99	1.79	locally polluted	172	306	considerable
D	10	0.92	1.46	locally polluted	176	357	considerable
C	5	0.81	1.52	locally polluted	129	222	moderate
E	4	0.79	0.97	unpolluted	133	245	moderate
H	11	0.67	1.42	locally polluted	95	260	moderate
B	9	0.48	0.87	unpolluted	77	113	low
F	4	0.46	0.54	unpolluted	82	96	low

Note: Localities (lettered as in Table 1) are ranked by decreasing mean PLI. PLI status: polluted, mean PLI > 1; locally polluted, mean PLI ≤ 1 but at least one sample with PLI > 1; unpolluted, all samples with PLI ≤ 1. Max RI class is the risk class of the highest-RI sample in each locality, following the RI classes in Table 2.

Table 5. Single-factor ecological risk (E_r) and contribution of each element to the potential ecological risk index (RI)

Element	T_r	Mean CF	Mean E_r	Max E_r	$E_r \geq 40$ (% sites)	Share of RI (%)
Cu	5	18.30	91.5	954	59	49.9
Cd	30	1.53	45.8	169	42	24.9
Hg	40	0.62	24.8	245	17	13.5
Pb	5	1.58	7.9	99.8	2	4.3
Co	5	1.50	7.5	50.1	2	4.1
As	10	0.36	3.6	15.7	0	2.0
Ni	5	0.31	1.6	6.1	0	0.9
Zn	1	0.93	0.9	6.0	0	0.5

Note: T_r = toxic-response factor; CF = contamination factor; $E_r = T_r \times \text{CF}$. $E_r \geq 40$ corresponds to moderate or higher single-factor risk (Table 2). Share of RI = mean E_r of each element divided by the mean RI across all sites (184). Elements are ranked by decreasing mean E_r.

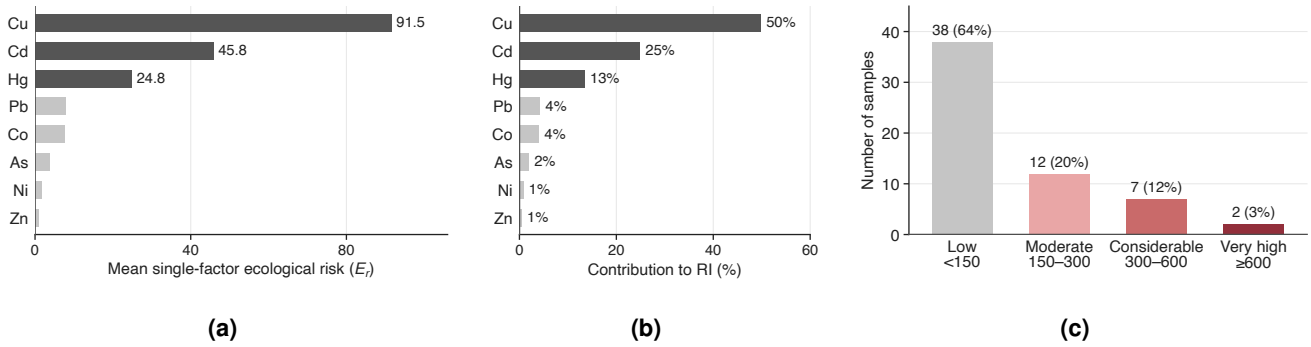


Figure 6. Ecological risk characteristics. (a) Mean single-factor ecological risk (E_r) by element. (b) Percentage contribution of individual elements to the integrated Risk Index (RI). In (a) and (b), Cu, Cd and Hg, which together account for 88% of RI, are shown in dark grey. (c) Distribution of the 59 soil samples across the four ecological-risk classes, defined by RI thresholds of 150, 300 and 600.

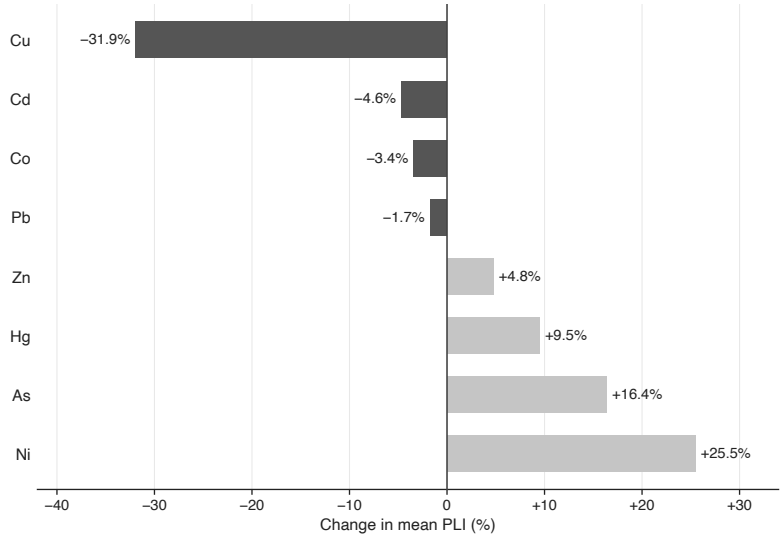


Figure 7. Sensitivity analysis: leave-one-out sensitivity of mean PLI to the exclusion of each metal. Bars show the percentage change in mean PLI when each element is excluded from the index in turn; dark bars, decrease; light bars, increase.

5). Copper remained the dominant contaminant under both reference backgrounds, although its mean CF decreased from 18.3 using UCC to 11.4 using Average Shale. Mean PLI decreased from 0.98 to 0.48 and mean RI from 184 to 90, moving the study-wide mean RI from the moderate to the low ecological-risk class, and the number of polluted sites (PLI $>$ 1) fell from 18 to 5 (Figure 5c, 5d compared with 5a, 5b). All five remaining polluted sites were at Wusakile (A), which remained the only locality with a mean PLI above 1 (1.45). These results indicate that absolute index values were sensitive to the choice of geochemical background, while the identification of copper as the dominant contaminant remained robust, as did the identification of Wusakile as the principal hotspot.

2. Leave-one-out influence on the Pollution Load Index. Recalculating PLI after sequentially excluding individual metals showed that copper had the greatest influence on the index (Figure 7). Excluding Cu reduced mean PLI by 32%, substantially more than the exclusion of any other element. In contrast, removing elements with relatively low contamination factors increased mean PLI, because their inclusion lowered the geometric mean. These results indicate that the overall pollution signal was driven primarily by a limited number of enriched metals rather than by uniform multi-element enrichment.

Spatial patterns and hotspots

Global Moran's I showed significant positive spatial autocorrelation for all eight PTEs (all $p < 0.01$; Table 6), indicating non-random spatial distributions. The strength of spatial autocorrelation varied considerably among elements, with Moran's I ranging from 0.221 for Hg to 0.632 for Ni.

Local Moran's I identified High–High clusters for Cu, Co, As, and Pb in Wusakile, Kitwe, with Low–Low clusters occurring toward the periphery of the study area. The spatial overlap of High–High clusters across these four PTEs identifies Wusakile as the principal multi-element contamination hotspot.

IDW surfaces further showed that Cu, Co, As, and Pb concentrations were highest in and around Wusakile, in the southern part of the study area, and generally declined with distance from this area (Supplementary Figure S1). Although the spatial extent of elevated concentrations varied among the four PTEs, their highest concentrations occurred within the same general area.

Table 6. Global Moran's I (8-nearest-neighbour weights, 999 permutations).

Metal	Moran's I	p	Pattern
As	0.289	0.001	clustered
Cd	0.323	0.001	clustered
Co	0.553	0.001	clustered
Cu	0.333	0.001	clustered
Hg	0.221	0.003	clustered
Ni	0.632	0.001	clustered
Pb	0.537	0.001	clustered
Zn	0.346	0.001	clustered

Discussion

The present study provides a spatially explicit characterization of PTE contamination in soils across mining communities in Zambia's Copperbelt Province, one of the world's largest and most historically significant sedimentary rock-hosted copper provinces. Beyond documenting substantial copper enrichment, perhaps unsurprising in the Copperbelt, the contribution of this study lies in showing where the copper burden occurs in human-use soils, which elements accompany it, and how unevenly it is distributed across mining-adjacent communities. Among the residential, agricultural and community lands sampled in three districts, 93% of samples exceeded the CCME agricultural guideline for copper (Figure 4b). Yet contamination and ecological risk were far from uniform. This heterogeneity is important for understanding how contamination is distributed across human-use soils in the Copperbelt. Previous studies have documented substantial contamination associated with mine wastes, smelting and mineral processing in the region[9,10,25,52].

Both contamination and ecological risk were greatest in Wusakile (reference cluster A), where Cu, Co, As and Pb formed overlapping High–High clusters (Table 4). Wusakile represents the clearest legacy hotspot situated within the long-established Nkana mining and smelting landscape, where previous work has documented substantial metal enrichment in surrounding soils[52]. As a densely populated urban neighborhood[53], Wusakile is particularly important from a human-exposure perspective. Its status as one of Kitwe's oldest and most prominent mining neighborhoods also makes it an important setting for understanding the legacy of long-term mining activity[53]. Locality G, corresponding to Kabundi, and locality D, corresponding to Ichimpe also showed elevated contamination and ecological risk, although less pronounced than Wusakile, illustrating that the environmental

burden is not confined to a single hotspot. Differences across localities may reflect geology, current and historic mine and processing activities, historical accumulation, environmental management, or other site-specific factors; the present data cannot distinguish among these explanations.

The elemental and index-based results also show why contamination cannot be summarized by a single metric. Cu was widely enriched, while the Cu–Co–As association is consistent with the mineralization and mining and metallurgical activity characteristic of the Copperbelt[5,6,25], although these associations cannot independently distinguish geogenic from anthropogenic sources. Ni provides an instructive counterexample: it exhibited the strongest global spatial autocorrelation despite being depleted relative to UCC, demonstrating that spatial clustering alone should not be equated with contamination. Likewise, mean PLI was approximately one despite 93% of samples exceeding the agricultural guideline for Cu. Because PLI is a geometric mean, elements below background can offset a strongly enriched element, as confirmed by the leave-one-out analysis. Toxicity weighting produced a different pattern, with Cd and Hg contributing disproportionately to RI because of their higher toxic-response factors[48]; the Hg contribution warrants particular caution because measured concentrations approached the analytical working limit. The sensitivity of PLI and RI to background selection further supports interpreting composite indices alongside element-specific concentrations and spatial patterns, and underscores the value of establishing a locally characterized Copperbelt geochemical baseline.

The occurrence of PTEs in human-use soils also raises important questions about potential human exposure. Total soil concentrations cannot establish individual exposure or health effects, which depend on bioaccessibility and contact through pathways including soil and dust ingestion, inhalation, water and food. However, qualitative research conducted concurrently and independently across the same study communities provides complementary evidence on how such potential exposure pathways are experienced by residents[31]. Participants described persistent airborne dust, visibly deteriorated water quality, changes in crops and concerns about associated health effects, with particular concern for children and limited ability to avoid perceived environmental hazards[31]. Previous biomonitoring research has also reported elevated Co, Cu, Pb and Se concentrations among residents of Zambian copper-mining communities[18], while contamination of water, food crops and ambient air provides additional potential exposure pathways[11,15,20]. These findings do not establish a causal pathway from the soil contamination measured here to human disease, but their convergence provides a strong rationale for exposure-focused follow-up in the identified higher-burden communities.

Several limitations should guide interpretation. The 59 sampling locations provide coverage across three adjoining mining districts but are insufficient to delineate fine-scale contamination boundaries within individual communities; LISA and IDW therefore identify broad areas for follow-up rather than precise hotspot boundaries. Sampling at 10–30 cm characterized the cultivated soil layer but may not represent concentrations in surface soil and dust most relevant to direct human contact; previous work around the Nkana smelter has documented pronounced enrichment in surface soils[52]. Total concentrations do not measure bioavailability or bioaccessibility, and a single dry-season campaign cannot capture seasonal redistribution. Finally, standardized UCC and Average Shale backgrounds[45,47] cannot substitute for a locally characterized baseline, and the observational design cannot attribute contamination to individual mines or emission pathways.

Conclusions and recommendations

This study demonstrates widespread Cu enrichment in soils across mining communities in Zambia's Copperbelt, alongside pronounced spatial heterogeneity and localized multi-element contamination hotspots. These findings underscore the need for locally characterized geochemical reference values that account for the region's naturally mineralized geology, alongside repeated, spatially resolved environmental monitoring in mining-adjacent communities.

Future assessments should extend beyond soil concentrations to investigate potential human exposure pathways, including drinking water, locally produced food, and metal bioaccessibility. Where warranted, human biomonitoring could further establish whether observed spatial differences in environmental contamination translate into differences in human exposure.

Finally, environmental monitoring programs should incorporate systematic communication of findings to affected communities and relevant authorities. Residents participating in the concurrent qualitative study explicitly requested environmental testing and access to results[31]. Integrating community engagement and accessible reporting into environmental surveillance would help ensure that monitoring informs both public understanding and locally appropriate mitigation strategies.

Data availability

The datasets generated and analyzed during this study are not publicly available. Spatially masked data may be made available by the corresponding author upon reasonable request.

References

1. U.S. Geological Survey. Copper: Mineral commodity summaries 2026. Tech. Rep., U.S. Geological Survey (2026).

2. PwC Zambia. Zambia mining report 2025. Tech. Rep., PricewaterhouseCoopers, Lusaka, Zambia (2025).

3. Ministry of Finance and National Planning. 2025 annual economic report. Tech. Rep., Republic of Zambia, Lusaka, Zambia (2025).

4. Republic of Zambia, Ministry of Mines and Minerals Development. National three (3) million tonnes copper production strategy by 2031. Tech. Rep., Ministry of Mines and Minerals Development, Lusaka, Zambia (2024).

5. Selley, D. *et al.* A new look at the geology of the Zambian Copperbelt. In Hedenquist, J. W., Thompson, J. F. H., Goldfarb, R. J. & Richards, J. P. (eds.) *Economic Geology 100th Anniversary Volume*, 965–1000, DOI: 10.5382/AV100.29 (Society of Economic Geologists, 2005).

6. Hitzman, M. W. *et al.* The Central African Copperbelt: Diverse stratigraphic, structural, and temporal settings in the world's largest sedimentary copper district. In Hedenquist, J. W., Harris, M. & Camus, F. (eds.) *Geology and Genesis of Major Copper Deposits and Districts of the World: A Tribute to Richard H. Sillitoe*, vol. 16 of *Society of Economic Geologists Special Publication*, 487–514 (Society of Economic Geologists, 2012).

7. Hitzman, M. W., Kaemba, R. & MacIntyre, T. An introduction to the geology of the Central African Copperbelt (with a focus on the Zambian Copperbelt and the domes region, Zambia). In *Central African Copperbelt, Zambia*, vol. 70 of *SEG Guidebook Series*, 1–39, DOI: 10.5382/GB.70 (Society of Economic Geologists, Littleton, CO, 2025).

8. Sracek, O., Mihaljevič, M., Kříbek, B., Majer, V. & Veselovskỳ, F. Geochemistry and mineralogy of cu and co in mine tailings at the copperbelt, zambia. *J. Afr. Earth Sci.* **57**, 14–30 (2010).

9. Dusengemungu, L., Mubemba, B. & Gwanama, C. Evaluation of heavy metal contamination in copper mine tailing soils of Kitwe and Mufulira, Zambia, for reclamation prospects. *Sci. Reports* **12**, 11283 (2022).

10. Chileshe, M. N. *et al.* Physico-chemical characteristics and heavy metal concentrations of copper mine wastes in Zambia: Implications for pollution risk and restoration. *J. For. Res.* **31**, 1283–1293 (2020).

11. Sracek, O. *et al.* Mining-related contamination of surface water and sediments of the Kafue river drainage system in the Copperbelt district, Zambia: An example of a high neutralization capacity system. *J. Geochem. Explor.* **112**, 174–188 (2012).

12. Chabu, C. Mitigating the concentration of arsenic in drinking water in mining areas using ferric hydroxide soil samples in response to the incidence of cancer diseases: A case study in Kitwe, Copperbelt province. *JSM Chem.* **10**, 1062, DOI: 10.47739/2334-1831/1062 (2023).

13. Bábek, O. *et al.* Trends of Cu and Co contaminated sediment dispersal along the Kafue river, the Zambian Copperbelt. *J. Geochem. Explor.* 107900 (2025).

14. Chama, L. *et al.* Ecological disaster in the making: Assessing the impact of Sino-Metals Leach tailings dam collapse on the Kafue River ecosystem. *J. Hazard. Mater. Adv.* 101104 (2026).

15. Muimba-Kankolongo, A., Banza Lubaba Nkulu, C., Mwitwa, J., Kampemba, F. & Mulele Nabuyanda, M. Impacts of trace metals pollution of water, food crops, and ambient air on population health in Zambia and the DR Congo. *J. environmental public health* **2022**, 4515115 (2022).

16. Bose-O'Reilly, S. *et al.* Lead intoxicated children in Kabwe, Zambia. *Environ. research* **165**, 420–424 (2018).

17. World Health Organization. *Guideline for Clinical Management of Exposure to Lead* (World Health Organization, Geneva, 2021).

18. Ndilila, W., Callan, A. C., McGregor, L. A., Kalin, R. M. & Hinwood, A. L. Environmental and toenail metals concentrations in copper mining and non mining communities in Zambia. *Int. journal hygiene environmental health* **217**, 62–69 (2014).

19. Ngosa, K. & Naidoo, R. N. The risk of pulmonary tuberculosis in underground copper miners in Zambia exposed to respirable silica: A cross-sectional study. *BMC Public Heal.* **16**, 855 (2016).

20. Mwaanga, P., Silondwa, M., Kasali, G. & Banda, P. M. Preliminary review of mine air pollution in Zambia. *Heliyon* **5** (2019).

21. Ripanda, A. *et al.* Combatting toxic chemical elements pollution for Sub-Saharan Africa's ecological health. *Environ. Pollut. Manag.* **2**, 42–62 (2025).

22. Mususa, P. Mining, welfare and urbanisation: the wavering urban character of zambia's copperbelt. *J. Contemp. Afr. Stud.* **30**, 571–587 (2012).

23. Smart, J., Nel, E. & Binns, T. Economic crisis and food security in africa: Exploring the significance of urban agriculture in zambia's copperbelt province. *Geoforum* **65**, 37–45 (2015).

24. Nyanza, E. C. *et al.* Toxic metal contamination of soil, water and food crops and associated human health risks in Zambia's copperbelt: A systematic review. Preprint (2026).

25. Křibek, B. *et al.* Impacts of mining and processing of copper and cobalt ores on the environment and human health in the central-northern part of the Copperbelt province of Zambia: An overview. *Min. Environ. Afr.* 24 (2011).

26. Sikamo, J., Mwanza, A. & Mweemba, C. Copper mining in Zambia-history and future. *J. South. Afr. Inst. Min. Metall.* **116**, 491–496 (2016).

27. Tinkler, O. S. & Sole, K. C. Copper solvent extraction on the African Copperbelt: from historic origins to world-leading status. *J. South. Afr. Inst. Min. Metall.* **123**, 349–356 (2023).

28. Zambia Statistics Agency. Population and Social Statistics. https://www.zamstats.gov.zm/population-and-social-statistics/. Accessed: 2026-10-09.

29. Canadian Mining Report. Map of major copper mines in Zambia (2026): Locations, operators & production overview. https://www.canadianminingreport.com/blog/map-of-major-copper-mines-in-{Z}ambia-2026-locations-operators-production-overview (2026). Published May 2, 2026. Accessed July 2026.

30. Nyanza, E. C., Dewey, D., Thomas, D. S., Davey, M. & Ngallaba, S. E. Spatial distribution of mercury and arsenic levels in water, soil and cassava plants in a community with long history of gold mining in Tanzania. *Bull. environmental contamination toxicology* **93**, 716–721 (2014).

31. Benshaul-Tolonen, A., Nyanza, E. C., Fernandez Musso, P., Gebreyohannes, M. & Ndhlovu, O. Dust, discoloration, and distrust: Lived experiences in zambia's copper mining communities. Tech. Rep. (2026).

32. Columbia, B. Recommended guidance and checklist for ecological risk assessment of contaminated sites in British Columbia—chapter 6. urban park (2010).

33. U.S. Environmental Protection Agency. Definition and procedure for the determination of the method detection limit, revision 2. Tech. Rep. EPA 821-R-16-006, U.S. Environmental Protection Agency, Office of Water, Washington, DC (2016).

34. Currie, L. A. Nomenclature in evaluation of analytical methods including detection and quantification capabilities (iupac recommendations 1995). *Pure applied chemistry* **67**, 1699–1723 (1995).

35. International Organization for Standardization. Accuracy (trueness and precision) of measurement methods and results — part 1: General principles and definitions. Tech. Rep. ISO 5725-1:1994, Geneva, Switzerland (1994).

36. International Organization for Standardization. Accuracy (trueness and precision) of measurement methods and results — part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method. Tech. Rep. ISO 5725-2:2019, Geneva, Switzerland (2019).

37. Bland, J. M. & Altman, D. G. Measuring agreement in method comparison studies. *Stat. methods medical research* **8**, 135–160 (1999).

38. McGraw, K. O. & Wong, S. P. Forming inferences about some intraclass correlation coefficients. *Psychol. methods* **1**, 30 (1996).

39. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. chiropractic medicine* **15**, 155–163 (2016).

40. Reimann, C., Filzmoser, P., Garrett, R. & Dutter, R. *Statistical data analysis explained: Applied environmental statistics with R* (John Wiley & Sons, 2011).

41. Zar, J. H. *Biostatistical Analysis* (Pearson, 2010), 5 edn.

42. Kaiser, H. F. The application of electronic computers to factor analysis. *Educ. psychological measurement* **20**, 141–151 (1960).

43. Jolliffe, I. T. *Principal Component Analysis*. Springer Series in Statistics (Springer, 2002), 2 edn.

44. Everitt, B. S., Landau, S., Leese, M. & Stahl, D. *Cluster Analysis* (John Wiley & Sons, Chichester, UK, 2011), 5 edn.

45. Rudnick, R. L. & Gao, S. Composition of the continental crust. In Holland, H. D., Turekian, K. K. & Rudnick, R. L. (eds.) *Treatise on Geochemistry*, vol. 3, 1–64, DOI: 10.1016/B0-08-043751-6/03016-4 (Elsevier, Oxford, UK, 2003).

46. Canadian Council of Ministers of the Environment. Canadian environmental quality guidelines: Summary table. https://ccme.ca/en/summary-table (2026). Accessed October 7, 2026.

47. Turekian, K. K. & Wedepohl, K. H. Distribution of the elements in some major units of the earth's crust. *Geol. society Am. bulletin* **72**, 175–192 (1961).

48. Hakanson, L. An ecological risk index for aquatic pollution control. a sedimentological approach. *Water research* **14**, 975–1001 (1980).

49. Müller, G. Index of geoaccumulation in sediments of the rhine river. *GeoJournal* **2**, 108–118 (1969).

50. Tomlinson, D. L., Wilson, J. G., Harris, C. & Jeffrey, D. Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. *Helgoländer meeresuntersuchungen* **33**, 566–575 (1980).

51. Anselin, L. Local indicators of spatial association—lisa. *Geogr. analysis* **27**, 93–115 (1995).

52. Ettler, V., Mihaljevič, M., Kříbek, B., Majer, V. & Šebek, O. Tracing the spatial distribution and mobility of metal/metalloid contaminants in oxisols in the vicinity of the nkana copper smelter, copperbelt province, zambia. *Geoderma* **164**, 73–84 (2011).

53. Peša, I. Water, housing and (in)formality in Kitwe, Zambia: Infrastructure, citizenship and urban belonging. In Lemanski, C. (ed.) *Citizenship and Infrastructure: Practices and Identities of Citizens and the State*, 104–122, DOI: 10.4324/9781351176156-7 (Routledge, 2019).

Acknowledgments

We are grateful for the support from the International Growth Centre (IGC), including Benjamin Shawa, Lwandamo Muwema, and Shahrukh Wani; Zambia's Ministry of Mines and Mineral Development; Zambia's Ministry of Local Government and Rural Development; and Zambia's Ministry of Health.

Funding

This work was sponsored by the International Growth Centre (IGC), United Kingdom.

Ethical approval

This study forms part of the broader research project, *Measuring and Mitigating Exposure to Toxic Metals in Copper Mining Areas in Zambia*. Ethical approval for the soil samples and focus group research was obtained from Barnard College (USA; approval no. 2425-1110-092) and ERES Converge (Zambia; approval no. 2025-08-08). Research authorization was granted by the Zambia National Health Research Authority (research license no. NHRA-2491/18/07/2025). All participants in the focus group research provided oral consent, as per the ethics approval. Soil samples were collected from the study sites with the relevant permission or authorization from the landowners or authorities.

Author contributions statement

A.B.T. conceived the study, secured funding, and supervised the project. E.C.N. developed the methodology, collected the samples, conducted the primary analysis, and wrote the original draft of the manuscript. A.B.T. and P.F.M. contributed to the analysis and interpretation of the results. P.F.M. administered the fieldwork. A.B.T., E.C.N., P.F.M., M.G., and O.N. reviewed and edited the manuscript and approved the final version.

Competing interests

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

Additional information

Correspondence and requests for materials should be addressed to A.B.T.

Supplementary Material

Table S1. Preparation-blank statistics and detection limits (n = 6 blanks; mg kg^{-1})

Element	Mean	SD	Min	Max	LOD (3 × SD)	MDL (Mean + 3 × SD)
As	0.00800	0.00487	0.00000	0.01386	0.01462	0.02262
Cd	0.00302	0.00232	0.000063	0.00585	0.00696	0.00998
Co	0.00820	0.00408	0.00348	0.01460	0.01223	0.02043
Cu	0.251	0.0821	0.121	0.342	0.246	0.498
Hg	0.00921	0.00513	0.00013	0.01487	0.01538	0.02459
Ni	0.0554	0.0421	0.00048	0.110	0.126	0.182
Pb	0.0372	0.0275	0.0145	0.0889	0.0826	0.120
Zn	0.776	0.890	−0.00282	2.028	2.671	3.447

Note: LOD = limit of detection, reported by the laboratory as three times the blank standard deviation; MDL = method detection limit, calculated as the blank mean plus three standard deviations[33]. The MDL exceeds the LOD for every element because the blank mean is non-zero. The negative minimum for Zn arises from blank subtraction. Elements are listed alphabetically.

Table S2. Measures of analytical precision for soil metals (N = 6 duplicate pairs; mg kg^{-1})

Element	Measurements (mg kg^{-1})			Reliability	Repeatability
	1st Mean (SD)	2nd Mean (SD)	Difference Mean[a] (95% CI)	ICC[b] (95% CI)	RC[c] (95% CI)
As	1.70 (0.72)	1.73 (0.79)	−0.030 (−0.250 to 0.191)	0.961 (0.752–0.994)	0.38 (0.25–0.84)
Cd	0.245 (0.198)	0.211 (0.151)	0.034 (−0.026 to 0.095)	0.947 (0.672–0.992)	0.12 (0.08–0.27)
Co	17.66 (8.24)	16.65 (7.09)	1.01 (−0.84 to 2.86)	0.974 (0.826–0.996)	3.73 (2.40–8.21)
Cu	366.3 (192.8)	367.2 (194.5)	−0.92 (−5.74 to 3.90)	0.9997 (0.998–1.000)	8.41 (5.42–18.52)
Hg	0.038 (0.024)	0.036 (0.025)	0.002 (−0.008 to 0.012)	0.930 (0.588–0.990)	0.017 (0.011–0.038)
Ni	11.42 (7.14)	11.18 (6.84)	0.25 (−0.47 to 0.97)	0.995 (0.966–0.999)	1.32 (0.85–2.90)
Pb	32.14 (32.52)	30.60 (30.17)	1.54 (−2.71 to 5.79)	0.992 (0.942–0.999)	7.85 (5.06–17.30)
Zn	91.43 (85.81)	92.31 (84.87)	−0.88 (−13.02 to 11.25)	0.991 (0.936–0.999)	20.8 (13.4–45.7)

Note: Duplicate pairs were six independently digested field samples (KL01, LW11, LG21, MS31, ND41, KB52). SD = standard deviation; CI = confidence interval; ICC = intraclass correlation coefficient; RC = repeatability coefficient.

[a]Mean of the signed differences (1st − 2nd measurement), $\bar{d} = \sum(x_{1i} - x_{2i})/n$, with a 95% CI from the t-distribution (df = 5). All CIs include zero, indicating no statistically significant systematic bias between replicates (paired t-tests, all $p > 0.20$).

[b]Two-way random-effects, single-measure, consistency-form ICC(C,1); the absolute-agreement form ICC(A,1) is near-identical because mean differences are negligible. Approximate 95% CIs derived from the F-distribution. Values > 0.90 indicate excellent reliability (Koo and Li, 2016). ICC is unitless.

[c]Value below which 95% of absolute differences between two replicate measurements are expected to lie under the same conditions, RC $= 1.96 \times \sqrt{2} \times s_w = 1.96 \times \mathrm{RMS}(d)$, where $s_w = \sqrt{\sum(x_{1i} - x_{2i})^2/(2n)}$. The 95% CI was obtained by scaling the chi-squared interval for the variance of the differences by 1.96. Smaller values indicate better short-term precision.

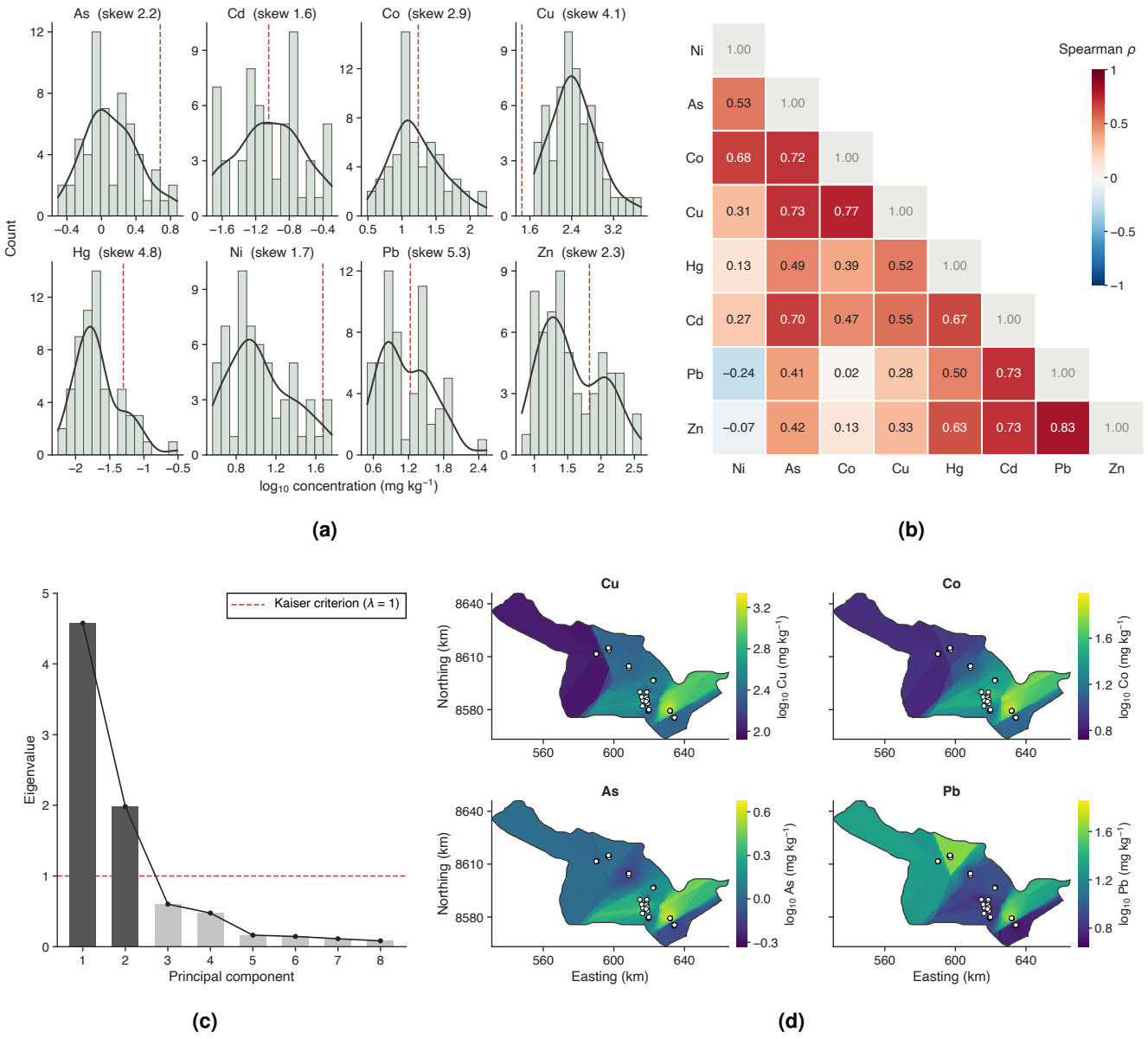


Figure S1. Supplementary analyses of potentially toxic element (PTE) concentrations. (a) Distributions of $\log_{10}$-transformed concentrations for each PTE, with kernel density curves; panel titles give the skewness of the untransformed concentrations, and the red dashed line marks the Upper Continental Crust (UCC) background. (b) Spearman rank correlation matrix of the eight PTEs. (c) Scree plot of principal component eigenvalues for the $\log_{10}$-transformed concentrations; the dashed horizontal line marks the Kaiser criterion (eigenvalue = 1), supporting the retention of two principal components. (d) Inverse-distance-weighted interpolation surfaces of $\log_{10}$-transformed Cu, Co, As and Pb concentrations, calculated in UTM Zone 35S; white dots mark sampling locations.

Table S3. Sample concentrations relative to blank-based detection limits (mg kg^{-1})

Element	Sample mean	LOD	MDL	Sample / MDL
As	1.71	0.0146	0.0226	∼76-fold
Cd	0.228	0.0070	0.0100	∼23-fold
Co	17.16	0.0122	0.0204	∼840-fold
Cu	366.8	0.246	0.498	∼740-fold
Hg	0.037	0.0154	0.0246	∼1.5-fold
Ni	11.30	0.126	0.182	∼62-fold
Pb	31.37	0.0826	0.120	∼260-fold
Zn	91.87	2.671	3.447	∼27-fold

Note: Sample mean is the average of the two replicate-group means in Table S2 and refers to the six duplicated field samples only, not to the full survey of 59 samples reported in Table 3 of the main text. LOD and MDL are reproduced from Table S1. Mercury (Hg) is the only element whose sample signal approaches the detection limits, at approximately 1.5 times the MDL and 2.4 times the laboratory LOD.