Geosystems and Geoenvironment 4 (2025) 100396

Contents lists available at ScienceDirect

Geosystems and Geoenvironment

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

Source apportionment of potentially toxic elements in soils from an urbanising region: Insights from multivariate analysis in Singida, Tanzania

Raymond Webrah Kazapoe[a,*], Benatus Norbert Mvile[b], John Desderius Kalimenze[c,d], Daniel Kwayisi[e,f], Samuel Dzidefo Sagoe[g], Kwabina Ibrahim[f], Obed Fiifi Fynn[h]

[a] Department of Geological Engineering, University for Development Studies, Nyankpala, Ghana
[b] Department of Physics, College of Natural and Mathematical Sciences, University of Dodoma, P. O. Box 259, Dodoma, Tanzania
[c] Department of Geography and Geology, University of Turku, FI-20014, Turku, Finland
[d] Geological Survey of Tanzania (GST), P.O. Box 903, Dodoma, Tanzania
[e] Department of Geology, University of Johannesburg, Auckland Park Kingsway Campus, South Africa
[f] Department of Earth Science, University of Ghana, Legon-Accra, Ghana
[g] Department of Environment and Sustainability Sciences, University for Development Studies, Tamale, Northern region, Ghana
[h] Department of Geological Sciences, University of Energy and Natural Resources, Sunyani, Ghana

ARTICLE INFO

Article history:
Received 1 January 2025
Revised 28 March 2025
Accepted 31 March 2025

Handling Editor: Dr. S Sanzhong Li

Keywords:
Soil pollution
Potentially toxic elements
Self-organising maps
Urbanisation and pollution hazards

ABSTRACT

This study evaluates the spatial distribution and geochemical characteristics of potentially toxic elements (PTEs) in soil samples across the Singida area, Central Tanzania, highlighting the environmental implications of rapid urbanisation and contributing to a deeper understanding of soil pollution in urbanising landscapes. A total of 1884 soil samples were analysed with an Inductively Coupled Plasma Mass Spectrometer (ICP-MS). The results of the study show that the background concentrations of the PTEs exceeded their corresponding Upper Continental Crustal (UCC) values in this order; Pb (86.25 %) > Ba (65.23 %) > As (45.65 %) > Cr (15.92 %) > Zn (15.18 %) > V (8.60 %) > Co (7.86 %) > Cu (5.68 %). However, only Cu (17 samples), Pb (2 samples), and Zn (1 sample) reached contaminant thresholds of 200 mg/kg, 200 mg/kg and 150 mg/kg, respectively in some samples. Agricultural practices and soil conditions are possible explanations for the high Cu values, which may be combined with other factors. This study also found that the Co, Cr, Ba and V concentrations vary greatly and even in some samples exceed the recommended levels. The principal component analysis, hierarchical cluster analysis, self-organising maps and positive matrix factorisation analysis revealed two main clusters: Ba, Zn and Pb (Factor 1) and Co, Cu, As, Cr and V (Factor 2). Cluster 1 is more prominent across most of the area, particularly the south. Cluster 2 is shown to be more prominent in the Northern part of the area such as Sekenke, Shelui, Lambi, Mtinko and New Kiomboi. Due to the growing rate of urbanisation, these areas have become relatively populous and have a high level of anthropogenic activities, such as gold mining, sunflower oil milling and agricultural activities which have been shown in the study to influence the spatial patterns of PTEs in the area. The level of anthropogenic influence on the PTEs calls for remediation and educative measures to be implemented.

1. Introduction

Soil is a very important element in the delivery of ecosystem services to human beings and their environment. Soil resource utilisation as an economic factor, such as urbanisation, industrialisation, and intensive agricultural production systems, has become an issue of concern. This is due to the myriad of effects the degradation has on the environment including the release of potentially toxic elements (PTEs) (Kumar et al., 2020; Ahmad et al., 2022). PTEs pollution in soils is of particular concern due to its persistence, concealment and potential irreversibility (Sethi and Gupta, 2020). In addition, these pollutants are spatially and temporally heterogeneous which depends on parameters such as industrial discharge, agricultural activities and urbanisation (Yang et al., 2022).

* Corresponding author.
E-mail address: rkazapoe@yahoo.com (R.W. Kazapoe).

https://doi.org/10.1016/j.geogeo.2025.100396

Globally, PTE contamination in soil is attributed to two primary sources: natural (geogenic) inputs and anthropogenic activities. For natural inputs, background concentrations vary between and within regions, although they are derived mainly from the parent material of the soil (Kabata-Pendias et al., 2017; Kazapoe and Arhin, 2021). In contrast, anthropogenic sources are varied and consist of industrial production waste, agricultural chemicals, and atmospheric deposition from transportation and energy production (Akhtar et al., 2021). However, they often overlap, leading to such complex contamination patterns that a robust analytical framework is needed to identify sources and implement mitigation strategies (Alloway, 2012; Kowalska et al., 2018; Kwayisi et al., 2024).

The rapid urbanisation in Tanzania during the last few decades has led to the emergence of pollution hazards often associated with ineffective environmental management. The country remains exposed to a significant level of industrialisation, urbanisation, energy production, transportation, and mechanised agricultural expansion zones, which have promoted PTE pollution (Mng'ong'o, 2022; Nyika and Dinka, 2023). Urbanization has become a substantial trend throughout Singida region during the recent decades as its urban districts continue to grow. Statistical data from the 2022 Population and Housing Census indicates that Singida region holds a total population of 2008,058 while its main urban district contains 232,459 residents (National Bureau of Statistics (NBS), 2022). The population of Singida urban district rose significantly between 2012 and 2022 according to the census data, increasing from 150,379 to 232,459 people, accounting for an annual growth rate of 4.4 % (NBS, 2022). The district currently experiences accelerating urbanisation because of population shifts from inside the country as well as infrastructure expansion and new economic opportunities located in the area.

These accompanying anthropogenic coupled with the extensive mineral deposits throughout Tanzania's geological landscape combine to produce background concentrations of PTEs in its soils. Mafic-ultramafic rocks in Central Tanzania regions make significant contributions to soil concentrations of manganese (Mn), zinc (Zn) and chromium (Cr), according to Mvile et al. (2023). Small-scale artisanal mining operations have grown throughout central Tanzania producing high levels of PTEs in adjacent soils. Research documents excessive levels of lead (Pb), arsenic (As), and cadmium (Cd), which exceed safety standards because of unsafe mining activities combined with flawed waste disposal methods (Mvile et al., 2023). Studies have also shown that soil contamination rises when fertilizers and pesticides containing trace metals are used. Zaller and Zaller (2020) found evidence that agrochemical exposure leads soil to accumulate hazardous PTEs which then threaten agricultural products as well as people who consume them. Dar es Salaam, along with other urban centres, experiences environmental degradation due to industrial operations releasing untreated waste into natural resources and non-compliant waste management practices. The identified industrial practices produce elevated concentrations of PTEs in both urban soils and aquatic systems which in turn pose risks to nearby agricultural areas during irrigation (Kibassa et al., 2013).

The sources of these heterogeneities and their interrelations are effectively analysed by employing both geostatistical and multivariate methodologies. These methods mainly analyse the spatial distribution but fail to provide comprehensive information on the types of heterogeneity across PTE pollutants (Kazapoe et al., 2021b). Additionally, the use of soil pollution indices is essential in addressing many of these challenges. Numerous studies have successfully applied various pollution indices to evaluate the degree of soil contamination (Kowalska et al., 2018; Baran, 2022). For instance, PTE contamination in soil and associated ecological risks have been assessed using pollution indices (Weissmannová and Pavlovský, 2017; Mahvi et al., 2022; Xiang et al., 2021; Hoque et al., 2023).

To identify sources and PTEs interactions in sediments, advanced statistical and modeling approaches are viable methods including principal component analysis (PCA), positive matrix factorisation (PMF), and the UNMIX model (Gulgundi and Shetty, 2019). For example, PMF quantifies source contributions from concentrations and uncertainty of chemical species; PCA is used to reduce data dimensionality to facilitate the interpretation of pollution patterns (Mas et al., 2010). However, these methodologies are increasingly complemented by ecological risk assessment methodologies to assess the potential human and environmental impacts of soil-bound PTEs (Kazapoe et al., 2022; Liao et al., 2021). Although significant progress has been made, the spatial distribution and ecosystem risk assessment of PTEs often remain regionally fragmented and rarely studied on a comprehensive scale (Ding et al., 2018; Zhang et al., 2019; Gan et al., 2023; Wang, 2023). The fragmented nature of current research and the importance of conducting large-scale studies to develop effective mitigation strategies necessitates a study of this nature. Additionally, temporal trends highlight the dynamic nature of PTE pollution over time, yet the underlying drivers, such as urbanisation and shifts in industrial practices, require further investigation to inform sustainable management strategies (Sharley et al., 2016; Li et al., 2020; Pan et al., 2021; Kazapoe et al., 2023, 2024). Understanding these dynamics is particularly critical in regions including Singida and its surroundings, where PTE contamination poses a potential threat to environmental quality and public health.

The hypothesis of this study is that soils in the Singida region are contaminated with potentially toxic elements (PTEs) from both natural and human-induced sources, with anthropogenic activities such as urbanisation, small-scale mining, and agriculture contributing significantly to elevated levels. Therefore, the objectives of this study are to PTE pollution to determine (i) concentrations of PTE in the soil of the study area, (ii) sources and interactions of these PTEs, (iii) spatial distribution, and (iv) ecological risk related to elements.

This study is framed within the context of rapid urbanisation in Tanzania, a perspective rarely emphasized in similar PTE assessments. The study uses a relatively high-resolution sampling approach and combines SOM with standard multivariate methods for enhanced source identification as well as detects new urban pollution sources. This research serves as a PTE baseline study for Singida region through which policymakers can use findings for environmental management and urban policy design. The research contributions unify geochemical data analysis with social-environmental findings to enhance existing knowledge about an underdeveloped yet growing vulnerable area.

2. Materials and methods

2.1. Area of study

The study area is situated in the core of the Tanzania Craton (Fig. 1) that is concealed by Neoarchaean granitoids, metasedimentary, and metavolcanic rocks. The Neoarchaean magmatism was discontinuous and took 160 million years (from 2775 to 2612 Ma; GST, 2015). Between 2712–2683 Ma, magmatic activity was terminated due to crustal uplift and late intrusive rocks were removed resulting in siliciclastic deposits and volcanic activity, possibly in an island arc setting. Granite volcan-sedimentary stages were the last to form, being intruded by large granite plutons and their associated vein system, syenite plutons were the last formed in the Neoarchaean magmatic event (Mvile et al., 2023). The invading granites were affected by deformation and low-grade metamorphism of the volcano-sedimentary rocks as evidenced

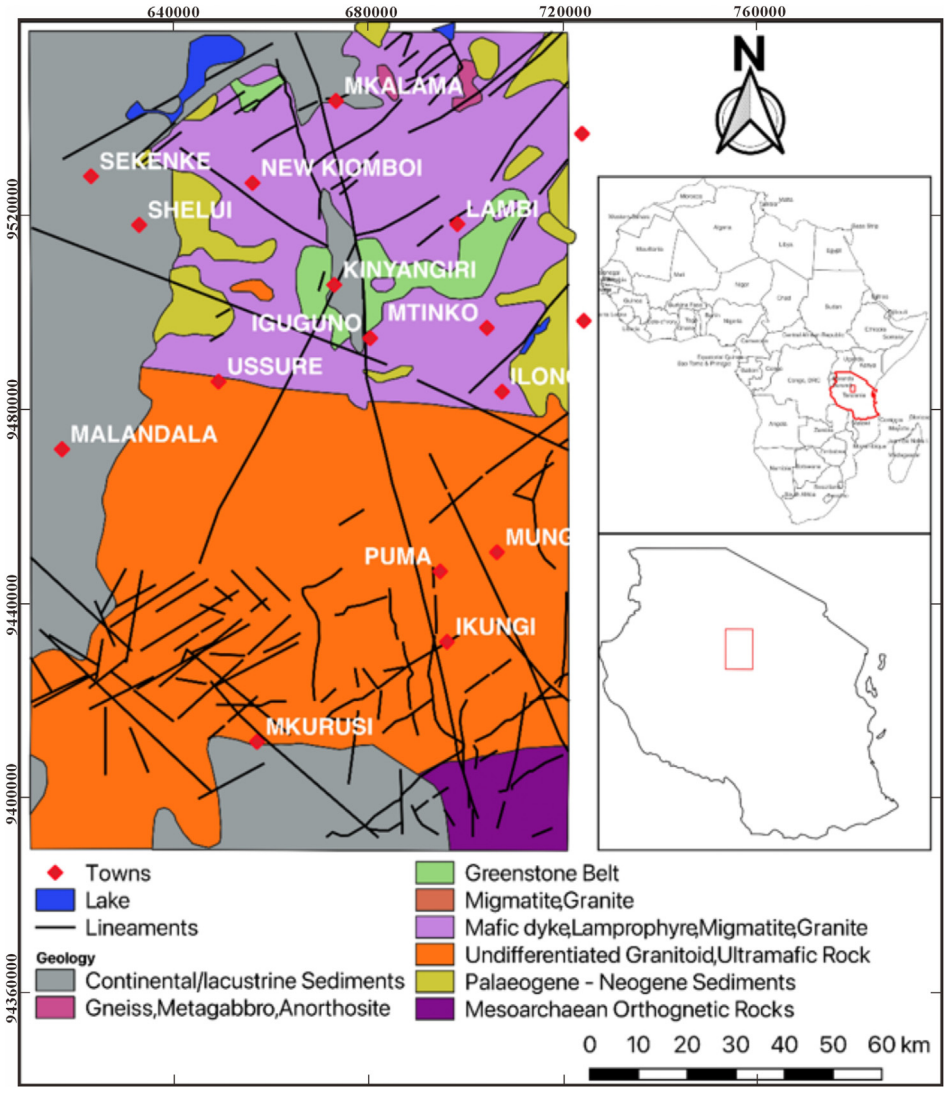

Fig. 1. Geology and Location map of the study area.

by folded granite veins. However, the duration of the tectono-metamorphic event is not clear. Syenite plutons may in places have some slight signs of deformation or metamorphism, whereby making the tectono-metamorphic event pre-dated at about 2612 Ma (Abu et al., 2024).

Ferruginous sandstones and conglomerates of Ndago are located above a peneplained granite surface that can be the regional Miocene Surface (Eades and Reeve, 1938). The clastic and chemical layer of Quaternary sediments indicating unending erosion due to the tectonic activities in East African Rift System. The highland areas are characterised by residual soils and the lowlands and the Wembere Mbuga occur in deposited, mainly alluvium. The chemical sediments are confined to small calcrete and silcrete layers in alluvium and ferricrete in soil. The study area and Singida region in particular have active artisanal mining for gold, copper, and detrital zircons-rare earth concentration. Additionally, the mining of building materials including granite and clay brick production from local clays is the customary activities (Kalimenze and Mvile, 2024).

2.2. Soil sampling

A systematic grid-based sampling design enabled the collection of 1884 soil samples to examine potentially toxic elements distribution and sources along with their concentrations across the Singida region of Central Tanzania. The sampling locations used a geographic coordinate system for placement and followed a set of intervals between points, which resulted in an average distance of 5.5 km. The method enabled space consistency while acquiring extensive area coverage that supported upcoming spatial analysis and source identification examinations. Surface soil samples were extracted at each sampling location through a composite collection process starting from the top 0–20 cm layer using hand augers and stainless-steel shovels to characterise agriculturally important and human-exposure areas. The collection of at least three sub-samples using a triangular layout with 50 to 100-meter spacing between points helped to overcome micro-level variability and boost sampling representativeness. The collected sub-samples were carefully combined to form a single 5-kilogram bulk sample. The High-Density Polyethylene (HDPE) bags were used to store individual samples, which received distinct pre-printed identification codes. After collection, the researchers used field sieving with a 2 mm mesh to eliminate stones, roots and coarse organic matter from the samples. The cleaned soil fractions were conveyed to the base station immediately after sealing them before storage under ambient dry conditions until laboratory analysis began.

A set of control samples stemmed from locations lacking human-caused impacts (outside mining areas and agricultural fields) utilizing land use data together with local expertise. The

Table 1
A description of the indices used in this study.

Equations	Definition of terms	References
$C_{f=\frac{C_i}{C_n}}$	where C_i is the mean concentration of pollutants and C_n is the pre-industrial reference value. C_f is the contamination factor.	(Hakanson, 1980)
$C_d = \sum_{i=1}^{n} C_f^i$	where C_i is the mean concentration of pollutants and C_n is the pre-industrial reference value. C_d is the degree of contamination.	(Hakanson, 1980)
$mC_d = \frac{1}{n}\sum_{i=1}^{n} C_f^i$	C_f is the contamination factor, n represents the number of pollutants considered, and mC_d is the modified contamination factor.	(Abrahim and Parker, 2008)
$PLI = \sqrt[n]{CF_1} \times CF_2 \times CF_3 \times \ldots\ldots CF_n$	C_f is the contamination factor, n represents the number of pollutants considered, and PLI is the pollution Load Index.	(Rodrigue et al., 2016)
$PI = \frac{C_n}{GB}$	C_n represents the concentration of heavy metals and GB represents the geochemical background values, and PI is the pollution index.	(Gong et al., 2008)
$PIsum = \sum_{i=1}^{n} PI$	PI is the pollution index, PIsum is the sum of pollution index, and n represents the number of heavy metals considered.	(Kowalska et al., 2018)
$PIaverage = \frac{1}{n}\sum_{i=1}^{n} PI$	PI is the pollution index, PIaverage is the average of pollution index, and n represents the number of heavy metals considered.	(Qingjie et al., 2008; Inengite et al., 2015)

researchers handled the control samples exactly like they handled the typical grid samples. During the sampling operation period, the team conducted rigorous quality assurance and quality control (QA/QC) procedures. The sampling process incorporated five percent duplicate sample collection from field stations as well as blank tests for contamination monitoring and official reference standards during lab-based examinations. The procedure followed for collecting geochemical data achieved reliability and reproducibility standards that match best practices for environmental geochemistry.

2.3. Analytical procedures

For chemical digestion and elemental analysis, representative aliquots (0.25 g) of the <75 µm fraction were sent to the internationally accredited ACME Laboratories in Canada. The four-acid mixture of hydrofluoric acid (HF) and perchloric acid ($HClO_4$), along with nitric acid (HNO_3) and deionised water operated at a ratio of 2:1:1:2 performed a near-total digestion on the samples. The chemical digestion follows the U.S. Geological Survey's (USGS) ICMPS81 procedure alongside total digestion methods, which is regularly used in high-precision environmental geochemistry research (e.g., USGS Open-File Report 03–024). A hot block applied heat to samples, which received 50 % hydrochloric acid (HCl) treatment before complete solid dissolution. Dilute HCl solutions were used to prepare test tubes which contained cooled solutions before the final dilution step to standard volume. The analysis was conducted on a 0.25 g split of the samples. The Inductively Coupled Plasma Mass Spectrometry (ICP-MS) used to determine the main elements also had detection limits of 0.0001 % to 0.01 % for oxides and 0.002 mg/kg to 2 mg/kg for trace elements. Duplicate samples were used to ensure the accuracy of laboratory analytical methods and the results. This research comprised 1884 soil samples and 209 replicates. The original and substantiation samples exhibited an outstanding correlation. The analysed original and replicate samples showed an acceptable variation of 1.2–9.3 %, thereby reflecting good quality control. The analysis procedure and sampling protocol were conducted as described by Kalimenze et al. (2023), Kazapoe et al. (2021a) and Abu et al. (2024).

2.4. Assessment of PTE pollution in the soil

Five indices were analysed to assess the extent of PTE pollution in the study area (Table 1). These included the pollutant accumulation index (PGI), PTE enrichment index (HMEI), ecological risk index (ERI), and PTE pollution load index (HMPLI). Table 3 shows the formula for each index.

2.5. Statistical analysis

The R software was used for the geostatistical analysis. Descriptive statistics helped determine the distribution of elemental concentrations through calculation of minimum, maximum, mean, median and standard deviation (SD) and coefficient of variation (CV) as well as kurtosis and skewness. The research concentrated on nine potential toxic elements (PTEs). A study has been conducted which looks at the correlations between these elements as well as their relationships with As, Ba, Co, Cr, Cu, Pb and V geochemical factors. Python version 3.10.12 was predominantly used for the statistical analysis in the study. This encompasses the correlation between parameters and the heatmap generated for its visualisation, the self-organising map (SOM) and its associated component planes, principal component analysis (PCA) and the hierarchical cluster analysis (HCA), as well as the chord diagram and dendrogram heatmap used to visualise them. Excel 2010 was employed for the summary statistics and EPA 5.0 was used to conduct the positive matrix factorisation analysis (PMF).

2.6. Self-organising map (SOM)

The self-organising map (SOM) algorithm, developed by Kohonen (1982), was employed in this study as a powerful tool for clustering and visualising relationships within the dataset. By employing a two-dimensional grid topology, the SOM algorithm facilitates the identification of patterns, groups, and similarities among the variables (Kohonen, 2001). This unsupervised machine learning approach enables the organisation of high-dimensional data into an interpretable low-dimensional representation, effectively highlighting inherent clusters and correlations between the input data. The SOM model is trained on the input data, the model updates its weight vectors iteratively based on the input data, using a neighbourhood function to ensure that adjacent neurons in the grid represent similar data points. Once the SOM training was complete, the non-hierarchical K-means classification algorithm was applied to refine and finalise the cluster assignments (Astel et al., 2007). This process facilitates the grouping of data into clusters that reflect their interdependencies, spatial distributions, and potential sources. To ensure the robustness of the clustering process, quantitative evaluation metrics were incorporated. The Davies-Bouldin index (DBI) was used to optimise the SOM parameters, such as the learning rate, neighbourhood size, and grid dimensions. This helps identify the parameter set that minimised within-cluster variance while maximising the separation between clusters (Davies and Bouldin, 1979). Additionally, Silhouette analysis was employed to determine the optimal number of clusters by evaluating the consistency and compactness of the resulting groupings. Higher Silhou-

Table 2
Statistical summary of the results from the Singida area.

	Cu(mg/kg)	Pb(mg/kg)	Zn(mg/kg)	Co(mg/kg)	As(mg/kg)	Cr(mg/kg)	Ba(mg/kg)	V(mg/kg)
Min	1.1	2.1	2	0.6	0.5	5	30	1
Max	246.8	302	424	68	84	679	4048	440
Median	11.8	22.7	38	8.7	1	48	539	47
Average	19.25	25.32	44.01	11.25	1.85	62.55	575.06	61.18
SD	20.84	17.38	29.79	8.99	3.03	50.45	330.83	49.94
Skewness	3.06	5.71	2.85	2.06	13.28	2.84	1.96	2.1
Kurtosis	15.06	62.27	20.86	5.75	309.89	16.18	10.71	6.06
CV (%)	108.24	68.63	67.7	79.89	183.03	80.66	57.53	81.62
Exceedance (%)	5.68	86.25	15.18	7.86	45.65	15.92	65.23	8.6
TBS	*200*	*200*				*100*		
CCME	*63*	*70*	*200*		*12*	*64*		
USEPA	*1600*	*400*	*3100*		*0.39*	*0.3*		
VROM	*36*	*85*	*140*		*29*	*100*		

*TBS: Tanzanian Bureau of Standards contaminant levels for heavy metals.
*CCME: Canadian Council for Ministers of the Environment standards for Agricultural soils.
*USEPA: The United States Environmental Protection Agency Regional Screening Level for residential soils.
*VROM: Circular on target values and intervention values for soil remediation of The Netherlands's Ministry of Housing, Spatial Planning and Environment.

ette scores indicated well-separated and cohesive clusters, which guided the final selection of clusters.

2.7. Positive matrix factorisation (PMF)

The EPA PMF 5.0 software (Song et al., 2006) was used to complete source apportionment analysis. The normalised data with its respective uncertainties were able to input to properly represent the model. Many possible factors were experimented with in order to find the most appropriate number to reflect the order of the data. The model performance was done by the Q-value and residual analysis (Karakas et al., 2017). The resulting PMF model was run with multiple initial conditions to ensure stability and the solution of lowest Q value and consistent factor profiles between runs was selected as the final solution.

3. Results and discussions

3.1. Pollutants and elemental concentrations and spatial patterns

A statistical summary of the soil quality is presented in Table 2. Exceedance as shown in Table 2, represents the percentage of samples which are above the Upper Continental Crustal (UCC) averages for the PTEs. The average concentration in mg/kg of the analysed PTEs in decreasing order are Ba (575.06) > Cr (62.55) > *V* (61.18) > Zn (44.01) > Pb (25.32) > Cu (19.25) > Co (11.25) > As (1.65) as seen in Fig. 2a and b. This marks a significant deviation from the findings of Rudnick and Gao (2003), who documented the elemental order for the UCC as Ba > *V* > Cr > Zn > Cu > Co > Pb, particularly differing from the results of this current study in the relative positions of Cr versus V, Pb versus Cu, and Cu versus Co. This relative change in the order of the natural abundance of the PTEs is suggestive of a degree of anthropogenic influence. This is corroborated by the coefficient of variation (CV %) recorded for the PTEs in this study (57.53–183.03 %) which are classified as high (i.e., 51–100 %) following Wang et al. (2024). The CV % in decreasing order are As > Cu > *V* > Cr > Co > Pb > Zn > Ba. Ba (57.53 %), Pb (68.63 %) and Zn (67.7 %) have high CV % which suggests a mixed influence of geogenic and anthropogenic factors on their spatial variance across the area. V (81.62 %), Cr (80.66 %) and Co (79.89 %) have very high CV % which implies a more pronounced anthropogenic control on their variance within the area. The significantly high CV % of Cu (108.24 %) and As (183.03 %) indicate that anthropogenic factors are largely responsible for the variability of these metals across the area. CV % > 100 % are typically classified as denoting extrinsic sources (Wu et al., 2014; Zhu et al., 2023; Adimalla et al., 2024; Cai et al., 2024; Gao et al., 2024). As concentrations across the area recorded a maximum value of 84 mg/kg. This relatively high average value of As (1.85 mg/kg) is identical to the UCC of As which is 1.8 mg/kg. Nearly half (45.65 %) of all samples had As values exceeding the UCC threshold. This value for As fell below the threshold for both the Canadian Council for Ministers of the Environment (CCME, 2007; NBS, 2013) standards for agricultural soils of 12 mg/kg and the circular on target values and intervention values for soil remediation of the Netherlands's Ministry of Housing, Spatial Planning and Environment (VROM, 2000) value of 29 mg/kg. However, it was above the United States Environmental Protection Agency's (USEPA, 2023) Regional Screening Level for residential soils of 0.39 mg/kg. The determined As values in this study were much lower than values determined in Geita district of northwestern Tanzania (63.8 ± 53.2 mg/kg) by Kaaya et al. (2025). Hotspots of As covers a significant portion of the Northern part of the area from Sekenke to Lambi (Fig. 3a). Ba concentration in the area ranged from 30 to 4840 mg/kg. The average concentration of Ba was 1.3 times more than the UCC of Ba (425 mg/kg). This high value is consistent for most of the samples, with a majority of them (65.23 %) exceeding this threshold. Hotspots of Ba can be found at Ikungi, Malandala, Kinyanguri, Iguguno and Lambi (Fig. 3b). The high spatial variability of Ba is consistent with geochemical heterogeneity that is largely driven by both geogenic and anthropogenic factors. Ba may be elevated as a consequence of weathering of Ba-rich Neoarchaean granitoid and syenites, tectonic redistribution of sediments or artisanal mining activities. The area of Ba hotspots corresponds to geological complex areas with mining activity, indicating local enrichment. Cu concentration ranges between 1.10 to 246.80 mg/kg (avg. 19.25). All the samples except one were below the contaminant level set by the Tanzanian Bureau of Standards (TBS, 2007a, 2007b) limit of 200 mg/kg as reported in Kibassa et al. (2013). These values are relatively higher than those reported by Kibassa et al. (2013) (8.98 mg/kg) for a study that was carried out in Dar-es-Salaam city. The results also mirror studies by Banzi et al. (2015) and Mng'ong'o et al. (2021) who reported values of 8.70 mg/kg and 3.34 mg/kg in Southern Tanzania. The relatively high concentration of Cu in the area may be because Cu concentration generally increases with Cu application, soil pH, soil salinity and exchangeable Na and Ca (Wightwick et al., 2006). The results presented in Table 2 show that Cr ranges from 5 to 679 mg/kg with an average concentration of 62.55 mg/kg which shows that about 60 % of the samples exceeded the threshold of 80.74 mg/kg that

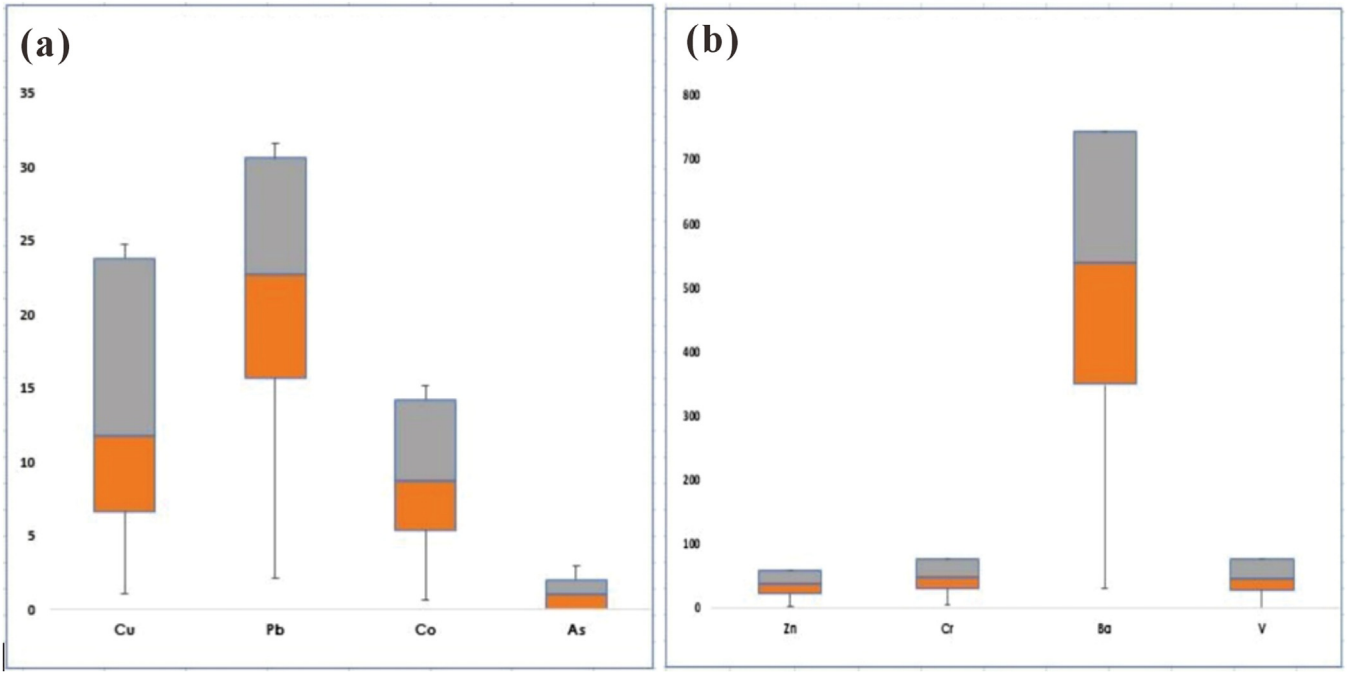

Fig. 2. Box and Whisker plots showing the concentrations of (a) Cu, Pb, Co and As in the study area and (b) the concentrations of Zn, Cr, Ba and V in the study area.

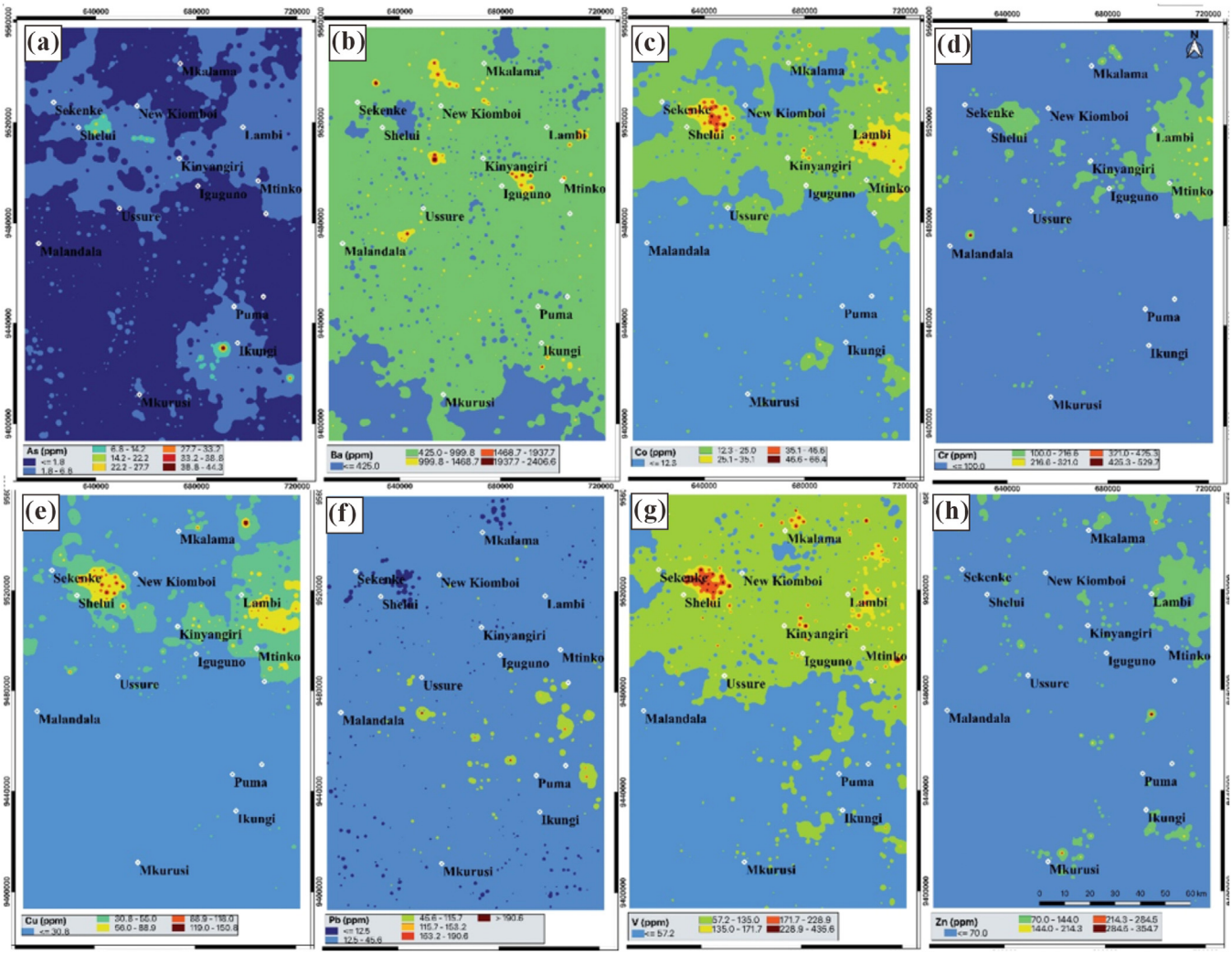

Fig. 3. Spatial distribution maps of (a) As (b) Ba (c) Co (d) Cr (e) Cu (f) Pb (g) V and (h) Zn across the Singida area.

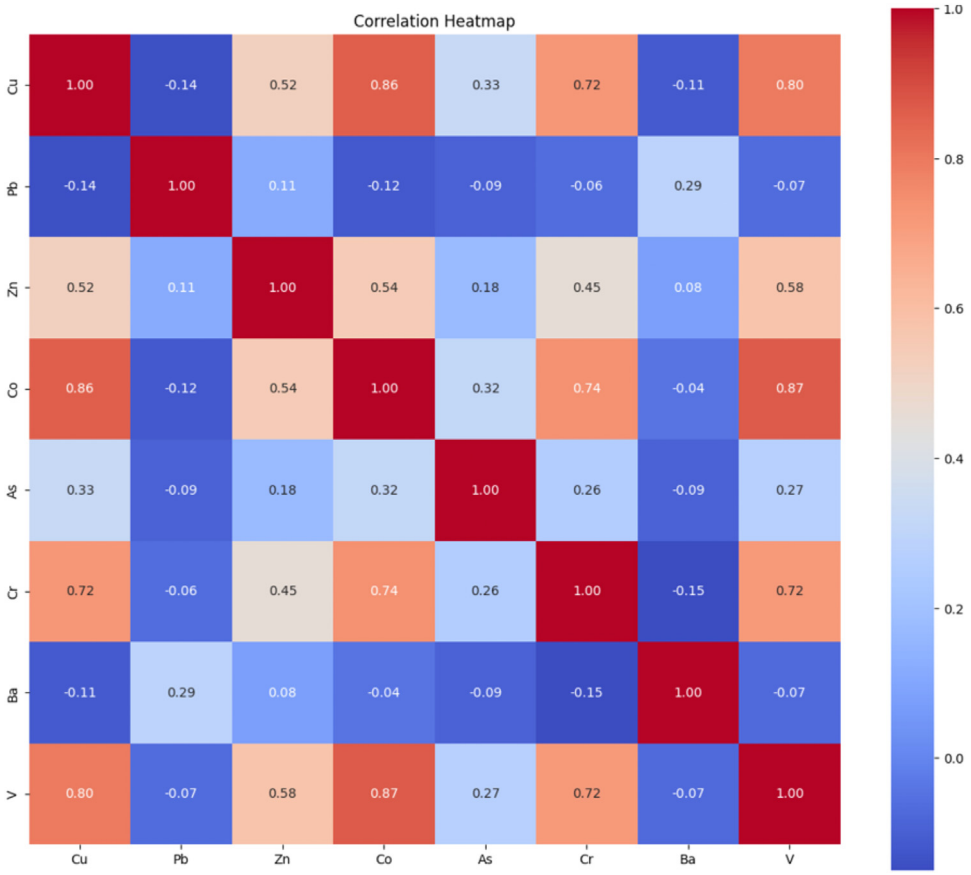

Fig. 4. Correlation matrix of the elements in the study area.

Kazapoe and Arhin (2021) set in West Africa. This also shows that 301 samples representing 15.98 % of the samples exceed the contaminant threshold for Cr set by the TDS and VROM (100 mg/kg). The average Cr value for this study (62.55 mg/kg) was nearly identical to the CCME value of 64 mg/kg but above the USEPA value foe residential soils of 0.39 mg/kg. This value for Cr is lower than what was determined for Nyarugusu in northwestern Tanzania (204.53 mg/kg) by Kaaya et al. (2025). V has a mean concentration of 61.18 mg/kg. The study showed that the mean concentration of V is below crustal values as set by Taylor (1964) which could be explained by the geology of the underlying rocks in the study area. Co, Cr, Cu and V show similar spatial dispersion patterns with hotspots around Shelui, Sekenke, Kinyangui and Lambi Mtinko in the northern part of the area (Fig. 3c-g). Pb concentration ranged from 2.10 to 302.00 mg/kg (Avg. 25.32) with only 2 samples exceeding the guideline value for contaminants set by the TBS for Pb. Machiwa (2010) also recorded similar mean values for a study that was done in the Lake Victoria Basin of Tanzania. The average Pb concentration fell below values for CCME (70 mg/kg), USEPA (400 mg/kg) and VROM (85 mg/kg). However, a majority of the samples (86.25 %) exceed the UCC of Pb (12.5 mg/kg). These values can be found across the entire area with the most intense spots found in the eastern part of the area around Malandala towards Puma and Mtinko (Fig. 3f). The results of Zn varied between 2.00 and 424.00 mg/kg (44.01 mg/kg). The results show that 17 of the samples were above the set contaminant level of Zn (150 mg/kg) for Tanzania. Similarly, high values of Zn have been reported for Tazara Mchichani and Temeke Wailes (57.10 mg/kg and 46.82 mg/kg) in southern Tanzania (Kibassa et al., 2013). Zn hotspots are sparse, comprising 15.18 % of the samples, and can be found around Lambi in the northeast and Ikungi and Mkurusi in the southern part of the area (Fig. 3h).

3.2. Relationships among the pollutant and trace elements

The covariance-matrix provides a means to assess the association between the various elements considered in the study. It is a measure of the closeness of dissimilar variables (Mugheri et al., 2019). In the output, only positive (direct) and negative (inverse) relationships were highlighted. Out of all the PTE considered in this study, Ba and Pb showed no strong correlation with the other elements. Ba was weakly correlated with Pb (0.29), Zn (0.08) and weakly negatively correlated with As (−0.09) and Cr (−0.15). Zn is strongly associated with V (0.58) and Cr (0.75) and less so with Cr (0.45). Cr on the other hand is strongly correlated with V (0.72). From Fig. 4, it is clear that the correlation of elemental concentrations could be grouped in three categories: high (r = over 0.5, $p < 0.05$), medium (r = positive but <0.5, $p < 0.05$) and low (r = negative values). Those in the first category includes: Zn-Cu, Cr-Cu, Cr-Co, V-Cu, V-Zn, Co-Zn, Cu-Co, V-Co, and V-Cr. The first group signals PTEs with likely common origins. This in combination with the summary and spatial characteristics suggests they are likely linked to anthropogenic sources, which are agricultural and mining activities. These elements are commonly associated and occur naturally in the environment more closely in association with the mafic gneisses, metagabbros and anorthosites characteristic of the neighbouring areas such as New Kiomboi and Mkalama. This suggests they may have originated from elemental mobility processes. Those in the second category include Zn-Pb, As-Cu, As-Zn, Cr-Zn, Ba-Pb, Ba-Zn, and V-As. The last group consist of the correlation between Ba and Pb on one hand and the other PTEs, such as, Pb-Cu, As-Pb, Cr-Pb, Ba-Cu, Ba-Cr, V-Pb and V-Ba. The third group suggests an association with the local geology, influenced to a greater degree by anthropogenic processes.

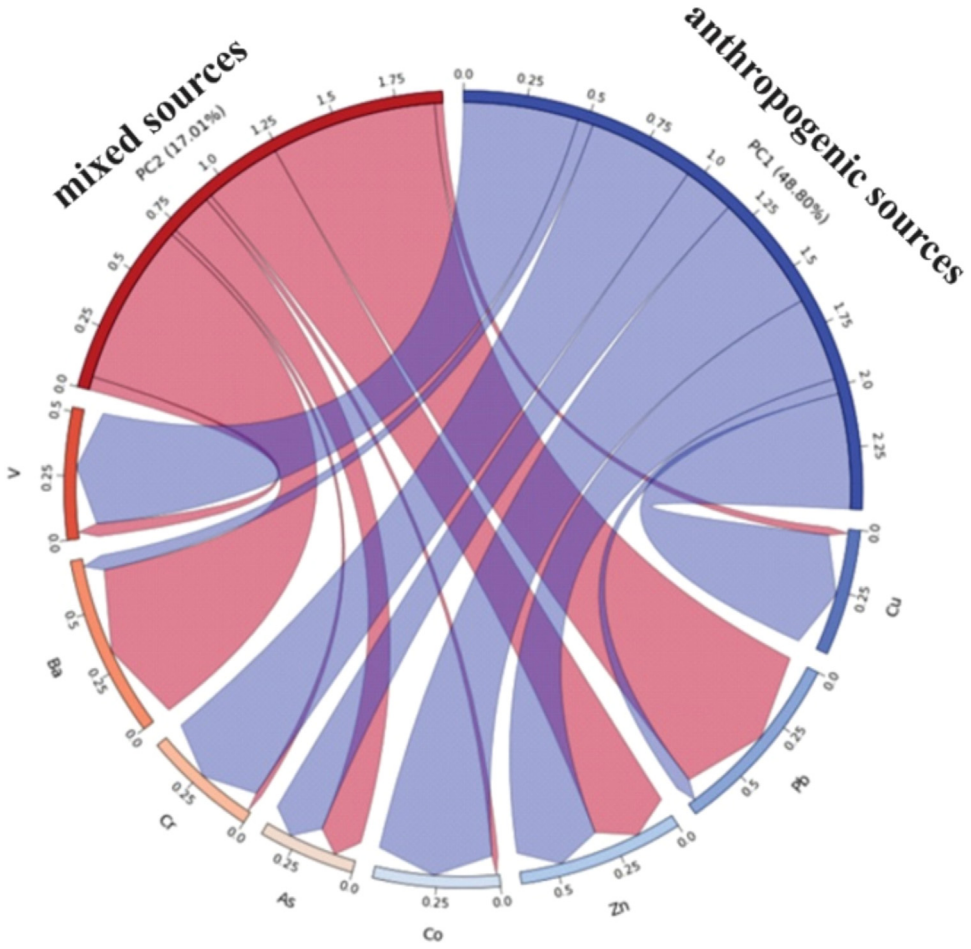

Fig. 5. The loadings of the analysed potentially toxic elements (PTEs) in the study onto the first two principal components derived from the dataset.

Table 3
Results for the KMO and Bartlett's test.

KMO and Bartlett's test		0.857
KMO measure of sampling adequacy	Approx. Chi-square	8351.361
Bartlett's test of sphericity	Sig	0.000

In terms of the principal component analysis, results from Table 3 present the outcomes of the KMO measure of sampling adequacy and Bartlett's test of sphericity. The sampling adequacy measure was determined to be 0.857, exceeding the required threshold of >0.5. Additionally, the significance value for Bartlett's test was 0.000, meeting the condition of ($p < 0.001$).

Fig. 5 and Table 3 show the loadings of potentially toxic elements (PTEs) analysed in the study (Cu, Pb, Zn, Co, As, Cr, Ba and V) onto the first two principal components (PC1, PC2) derived from the dataset. Contributions to PC1 are shown in blue, PC2 in red. The thickness and direction of the connecting chords represent the strength and nature of the relationships between the metals and the components. PC1 accounts for the largest variance (48.80 %), with Cu, Co, Cr and V having the most significant connections. PC2 explains 17.01 % of the variance, with notable connections to Pb and Ba. Factor analysis results support those of the element association in cluster 1 (Fig. 4) and hence suggest lithological control of the elements present in the samples, particularly those associated with mafic rocks.

The HCA analysis (Fig. 6) outlines 2 main clusters. The first is composed only of Ba while the second has As, Co, Cu. Cr, Pb, V and Zn. The high concentration of Ba may be linked to agricultural practices as well as the surrounding country rocks which include ferruginous sandstones and conglomerates, which dominate the area, and often contain Ba-bearing minerals such as barite ($BaSO_4$) or Ba-rich feldspars. These minerals release Ba into the soil during weathering processes. The ferruginous components of these rocks, characterised by their high iron content and large reactive surface area, further enhance Ba retention through adsorption mechanisms (Tarawneh et al., 2011; Jones et al., 2023; Jiménez-Vázquez et al., 2025). The second cluster could be linked to anthropogenic activities which induce localised concentration of these PTEs in the soil such as mining and agricultural practices Fig. 7.

Leveraging the Davies-Bouldin index (Best DBI: 0.9607) and Silhouette score for the identification of the best set of parameters (grid dimensions 10, 10, learning rate 0.5, sigma 2, iterations, 50,000) and the number of clusters (2). The SOM is composed of 100 hexagons each representing a neuron, the number within each neuron helps identify the samples within each hexagon thus identifying the clusters the samples belong to as well. The image shows different number of clusters ranging from 2 to 10 and its associated Silhouette score (Fig. 8). The number of clusters with the best Silhouette score is the best for the SOM model. The highest score occurs at 2 clusters suggesting that 2 is the best number of clusters for the SOM model.

Fig. 9 is a component plane for each PTE considered in the study. Each square within each component plane represents a neuron, neurons with high values are represented by bright colours (yellow) and neurons with low values, dark colours (blue). Ar-

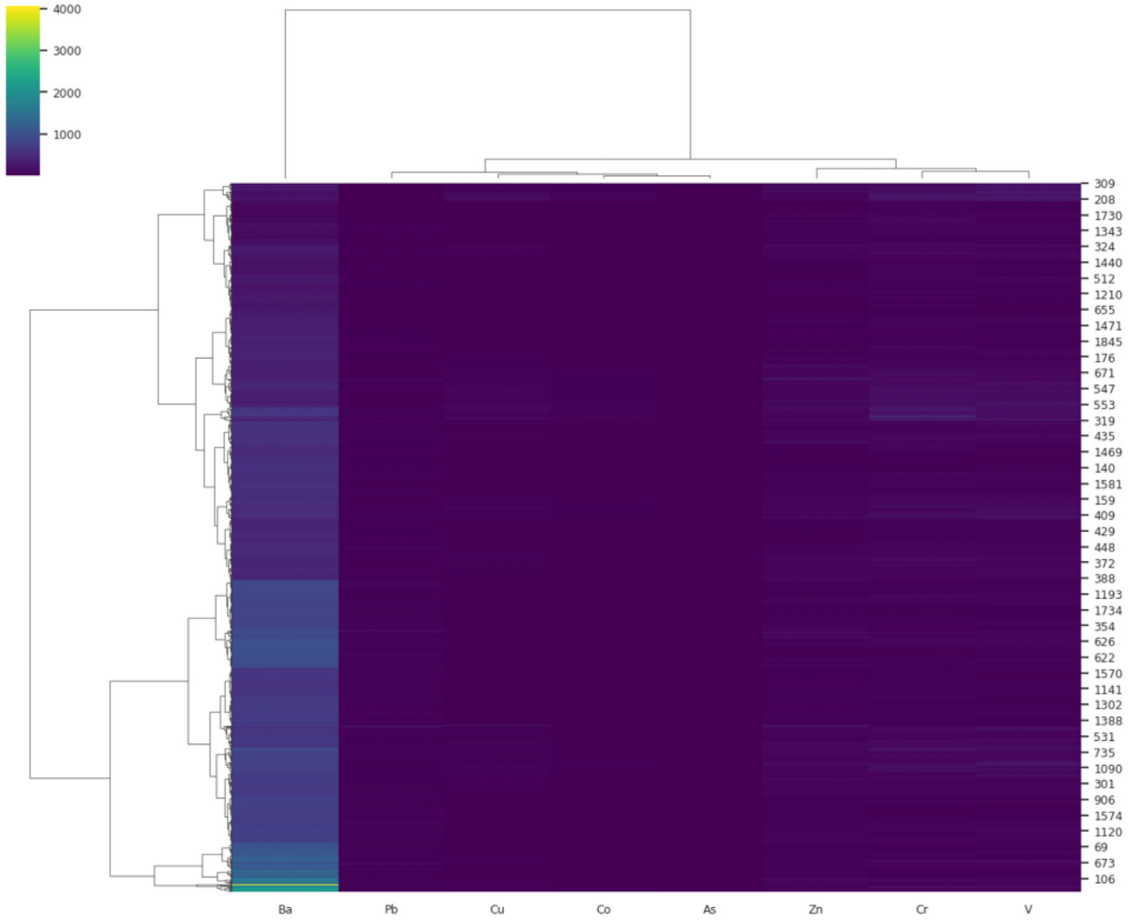

Fig. 6. Dendogram showing the main clusters of PTEs from the study.

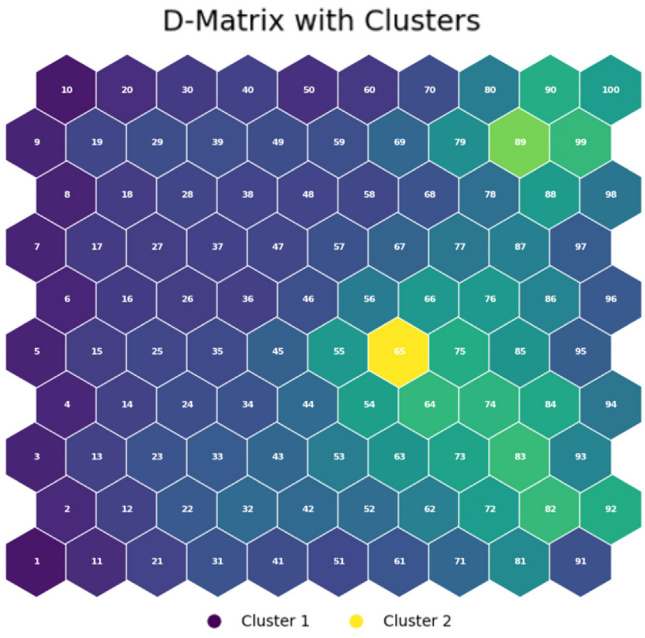

Fig. 7. Self-organizing map (SOM) D-matrix visualizing clusters based on input data. Sample IDs are displayed within each node, illustrating the distribution of data points across the SOM.

eas with high values suggest high activation indicating the associated neuron is strongly represented by the PTE, areas with low values suggest less activation thus the PTE is less prominent. The results indicate that Co, Cu, As, Cr and V have high activations in relatively similar regions. Ba, Zn and Pb are the outliers. These findings mirror the results from Fig. 9, where Co, Cu, As, Cr, and V all find themselves in one cluster (1) and Ba, Zn, Pb in another (2). The SOM analysis corroborates the results from the PCA, HCA and spatial analysis in outlining these two clusters. Both clusters show the impact of anthropogenic activities to varying degrees. The influence of anthropogenic activities, chiefly mining and agricultural practices, on the concentrations of these PTEs in that part of Tanzania has been attested by Mihale (2019), Abu et al. (2024) and Karungamye et al. (2023). Fig. 10 shows cluster 1 covering most of the area, particularly the south. This underscores the mixed nature of the sources for these elements, where they naturally occur in these areas but have been significantly influenced by human-induced activities (Abu et al., 2021; Kalimenze et al., 2023). Cluster 2 is shown to be more prominent in the Northern part of the area such as Sekenke, Shelui, Lambi, Mtinko and New Kiomboi. These areas are known to be relatively populous and have high level of anthropogenic activities such as gold mining, sunflower oil milling and agricultural activities (Kalimenze and Mvile, 2024). Studies by Manya (2012), Lawley et al. (2014) and Henckel et al. (2016) demonstrate gold mineralisation patterns in the area that establish mafic rocks as primary hosts for gold deposits. The designation of these areas as hubs for Artisanal and Small-Scale Mining (ASSM) has further amplified their importance as gold production centre. Kalimenze and

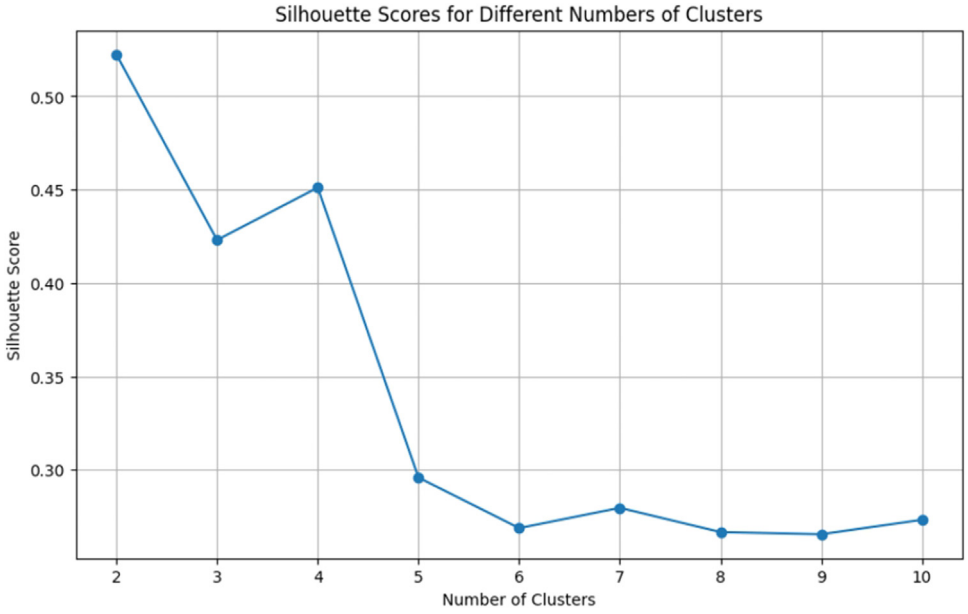

Fig. 8. Figure showing the results from the Silhouette analysis.

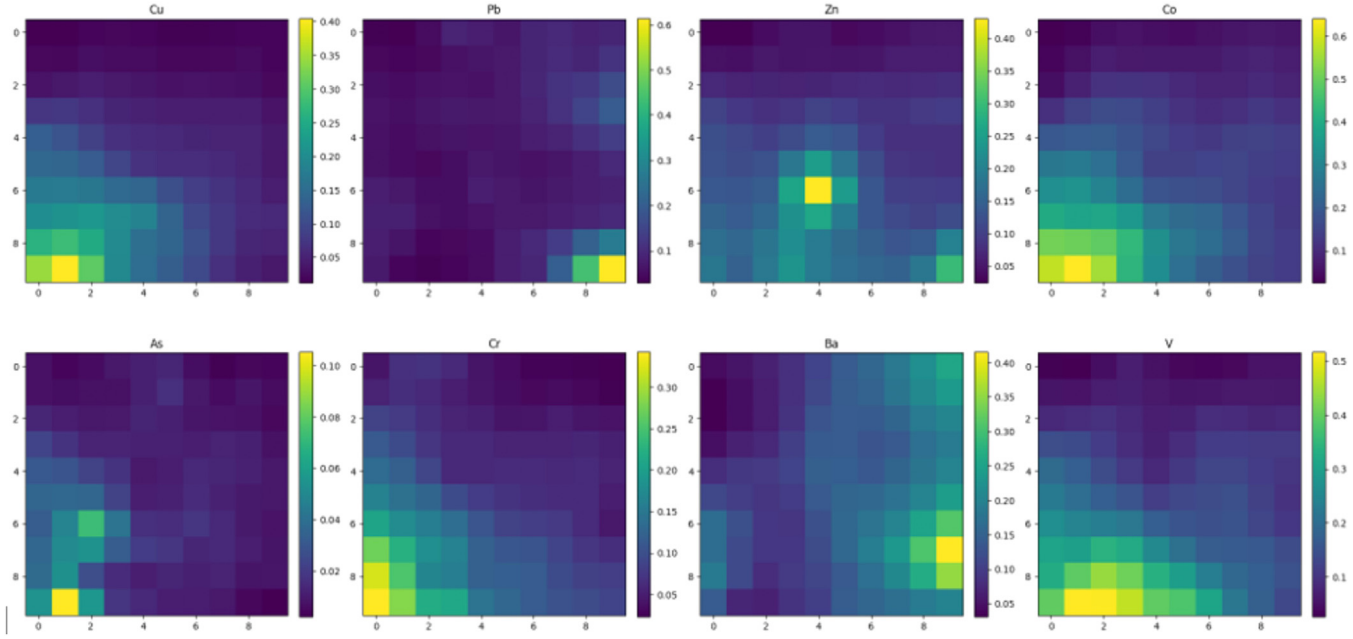

Fig. 9. Component planes for each of the PTEs.

Mvile (2024) report that the presence of these deposits has attracted numerous miners to Singida causing ASSM activities to expand quickly. The absence of proper regulations in these activities creates major environmental problems. Rising mining activities have increased concerns as this may introduce PTEs into the environment and thereby threaten local ecological systems. Research in similar gold-mining territories including Lake Victoria Goldfields shows that mining operations release toxic metals such as arsenic along with Pb and Hg into surrounding natural systems (Van Straaten, 2000). This aligns with the presence of these clusters centred around the known mining areas.

The EPA PMF model (V. 5) was applied with the elements analysed in the soil samples to identify and quantify the probable origins of the PTEs in the study area as well as the effect of every element. In the present research, the PMF model was used 20 times for which the chosen factors were 2 or 3. In this study, two factors were chosen according to the level of pollutant for each of the PTEs. The value of R^2 (fitness) of predicted and observed concentration for values above 0.94 mark out the fitness of the model. The concentration and contribution rate of every factor are shown in Figs. 11 and 12 for the PTEs.

The results from the PMF analysis aligns with the results of the PCA (Table 4). Factor 1 shows high loadings by Pb (94.4 %) and Ba (94.2 %) as the sole members. The CV % of Pb (68.63 %) and Ba (57.53 %) as well as their spatial characteristics shown in Fig. 3 distinguishes them from the other PTEs. These characteristics outlines that these elements are sourced from the local geology but their dispersal across the area have been influenced by anthropogenic activities. Consistent with findings by Mvile et al. (2023), we infer that the mining and agricultural practices in central Tanzania play

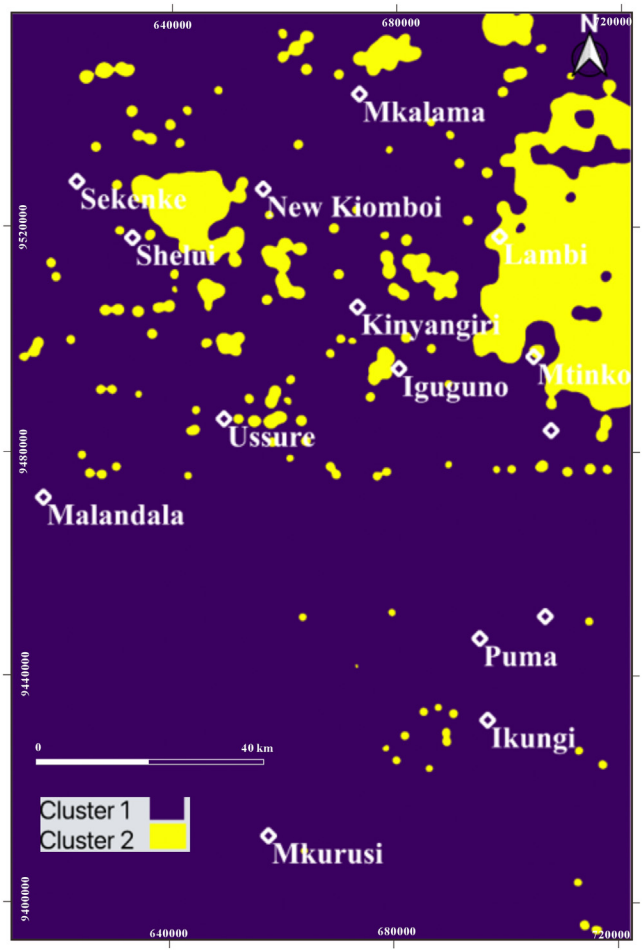

Fig. 10. Spatial representation of SOM U-matrix across the study area.

Table 4
Factor loadings for Components 1 and 2.

	PC1	PC2
Eigen values	3.906	1.361
Percentage of variance (%)	48.80	17.01
Cumulative percentage (%)	48.80	65.81
Cu	0.462	0.033
Pb	0.061	0.666
Zn	0.338	0.311
Co	0.473	0.025
As	0.209	0.172
Cr	0.426	0.023
Ba	0.062	0.651
V	0.463	0.054

a role in soil pollution with PTEs. Factor 2 has strong loadings for Cu (96.1 %), V (90.9 %), Cr (90.5 %), Co (84.8 %, As (72.7 %) and Zn (70.4 %). These PTEs show similar variance and spatial characteristics in the area. Additionally, their very high CV % values (67.7–183.03 %) suggests anthropogenic activities linked to the gold mining and farming practices. Similar associations were documented in the Singida region by Herman and Kihampa (2015) with small-scale gold mining suspected to be the source of elevated levels of PTEs in soils and water.

Table 5
Summary of the results of the indices.

Index	Min	Max	Mean	Range	Class	Number of Samples (%)
Degree of Contamination (Cd) (Hakanson, 1980)	0.851	23.997	6.070	< 8	Low	1506 (79.94 %)
				8≤Cd<16	Moderate	368 (19.53 %)
				16≤Cd<32	Considerable	10 (0.53 %)
Modified Degree of Contamination (mCd) (Abrahim and Parker, 2008)	0.085	2.340	0.607	< 1.5	nil to very low	1868 (99.15 %)
				1.5 ≤ mCd< 2	low	15 (0.80 %)
				2 ≤ mCd < 4	moderate	1 (0.05 %)
				4 ≤ mCd < 8	high	
				8 ≤ mCd< 16	very high	
				16 ≤ mCd<32	extremely high	
				mCd≥ 32	ultra-high	
Pollution Load Index (PLI) (Rashed, 2010; Rai et al., 2019)	0	79.49065	0.150969	0	None	1850 (98.20 %)
				1	None to medium	15 (0.80 %)
				2	Moderate	7 (0.40 %)
				3	Moderate to strong	3 (0.20 %)
				4	Strongly polluted	2 (0.11 %)
				5	Strong to very strong	1 (0.05 %)
				6	Very strong	6 (0.32 %)
Sum of Pollution Index (PIsum) (Haque et al., 2022)	0.925756	53.88584	10.19577	<1	Low	2 (0.11 %)
				1<Pisum<3	Moderate	105 (5.57 %)
				3<Pisum<6	High	488 (25.90 %)
				>6	Very high	1289 (68.42 %)
Average of Pollution Index (Plaverage) (Gong et al., 2008)	0.077146	4.490487	0.849647	<1	High quality soil	1271 (67.46 %)
				>1	Low quality soil	631 (33.49 %)

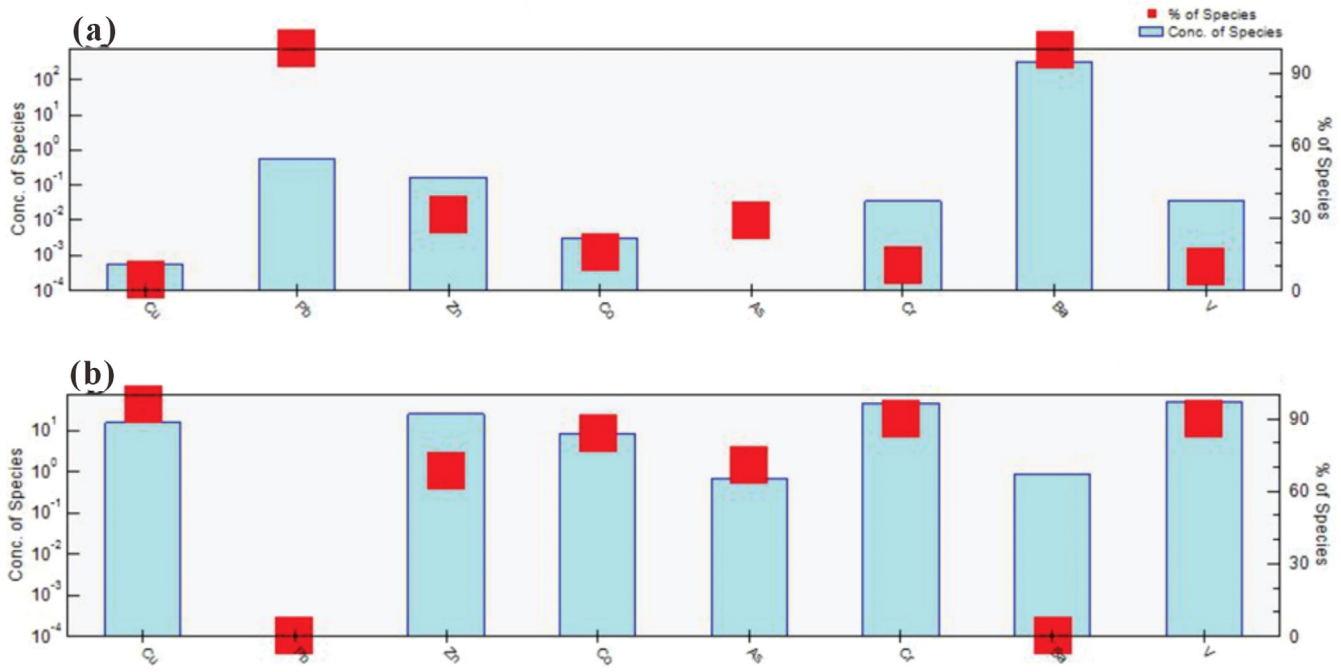

Fig. 11. PTEs profile source and contribution from PMF (a) Factor 1 and (b) Factor 2.

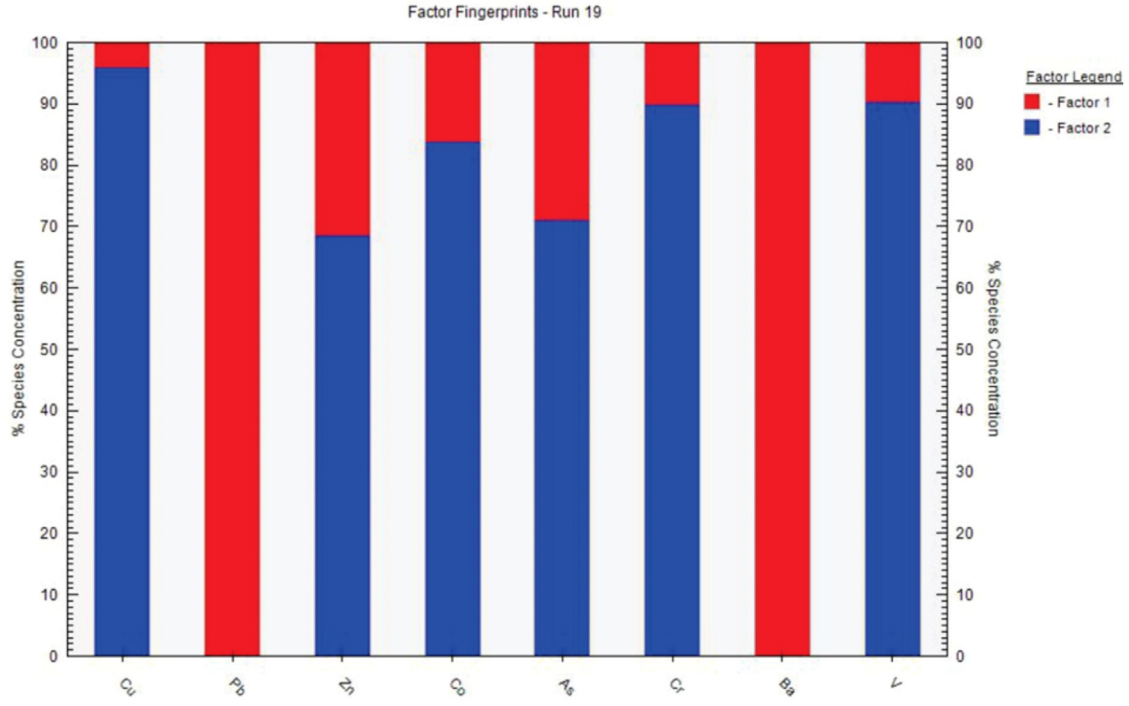

Fig. 12. PTEs source and factor fingerprint from PMF.

3.3. Potentially toxic element contamination in singida urban and non-urban areas

A summary of the results of the indices used to assess the level of contamination or pollution in the soil samples of the study area are shown in Table 5. The calculation for degree of contamination shows that 19.53 % and 0.53 % of the samples are classed as "moderate" to "considerable" contaminated respectively, while the majority of the samples (79.94 %) of the samples are classed as "low" in terms of contamination. These results fall more in line with Sivakumar et al. (2016) in coastal sediment from South East Coast of Tamilnadu. The values determined in this study are however lower than what was determined by Mihale (2019), who found that all samples from Mtoni estuary in Dar es Sallam were severely contaminated (DC > 48).

Similarly, Table 5 shows that most of the samples (99.15 %) fall under the "nil to very low" depicting insignificant contamination among the soil samples in the area. The PIsum analysis indicates that most samples (68.42 %) are polluted. However, only two samples (0.11 %) record low pollution. This is in contrast to the PIaver-

age which shows that 67.46 % of the samples are of high quality. The Nemerow Pollution Index reflects a mixed picture where 33 % of the samples are classed under the Heavy Polluted class while 22.56 % and 18.95 % are classed in the Slight and Moderate Polluted categories respectively. The HMPLI records the majority of the samples (69.43 %) under the "control" classification. 28.77 % are reported under the baseline level with only 1.8 % reported under continuous degradation class. The HMEI analysis identified no enrichment in Cu (91.14 %), Zn (95.28 %), Co (85.56 %), As (99.10 %), Cr (81.48 %) and V (90.34 %). The analysis also shows moderate enrichment in Pb (58.12 %) and Ba (43.84 %). The PGI analysis shows that the significant majority of the samples record no pollution in the order of Ni > Zn > Co > Cu > Sr > Cr > Mn > Pb. Similar findings have been reported by studies from central Tanzania by Mvile et al. (2023) which showed that some parts of the coastal zone had elevated concentrations of PTEs such as Pb and Cr (Mvile et al., 2023). Similarly, Abu et al. (2021) identified As, Cd and Pb as polluted in Singida. The authors assert that these PTEs are predominantly linked to geogenic source related to the mafic ore bearing with some input from the mining processes within the area.

The moderate enrichment shown by the PTE enrichment index (HMEI) could be explained by anthropogenic activities such as mining and industrial processes that have been reported in other parts of Tanzania. Studies have shown, for example, the role of small-scale gold mining in soil PTEs contamination such as Pb and Ba (Mnali, 2001).

4. Conclusions

The present study is a comprehensive risk assessment of PTE pollution in Singida and its surrounding areas. The study was designed to investigate the environmental quality in relation to PTE pollution as follows: pollutants, concentrations and sources, intervening patterns, distributions and ecological risks. The results of the study show that although the background concentrations the PTEs exceeded their corresponding UCC values in this order: Pb (86.25 %) > Ba (65.23 %) > As (45.65 %) > Cr (15.92 %) > Zn (15.18 %) > *V* (8.60 %) > Co (7.86 %) > Cu (5.68 %); only Cu (17 samples), Pb (2 samples), and Zn (1 sample) had reached contaminant thresholds of 200 mg/kg, 200 mg/kg and 150 mg/kg respectively in some samples. Nearly half (45.65 %) of all samples had As values exceeding the UCC threshold. The average concentration of Ba was 1.3 times more than the UCC of Ba (425 mg/kg). While results for Cr shows that about 60 % of the samples exceeded the threshold of 80.74 mg/kg. Agricultural practices and soil conditions are possible explanations for the high-Cu values, which may be combined with other factors. This research has found that the Co, Cr, Ba andVconcentrations vary greatly and even in some samples exceed the recommended levels. The covariance-matrix suggests strong possibilities for some of these factors having the same geological origins or anthropogenic impacts based on their strong correlation. The PCA, HCA, SoM and PMF analysis revealed two main cluster; Ba, Zn and Pb (Factor 1) and Co, Cu, As, Cr, and V (Factor 2). Both clusters show the impact of anthropogenic activities to varying degrees. Cluster 1 is more prominent across most of the area particularly the south. cluster 2 is shown to be more prominent in the Northern part of the area such as Sekenke, Shelui, Lambi, Mtinko and New Kiomboi. These areas are known to be relatively populous and have high level of anthropogenic activities such as gold mining, sunflower oil milling and agricultural activities. Results of contamination assessment indices indicate that the majority of soil samples possess low nor negligible levels of contamination, 79.94 % fall under the 'low' degree of contamination and 99.15 % in the 'nil to very low' contamination categories. Various indices show certain differences, indicating that 68.42 % Polluted samples, according to Pollution Index (PIsum), mixed pollution from Nemerow pollution index and PTEs enrichment in general absent except slight enrichment in Pb and Ba.

Based on the above, the following is recommended:

- Implement phytoremediation using metal-absorbing plants. Use soil washing and stabilisation to reduce PTE concentrations.
- Establish health surveillance programs in high-risk areas. Ensure access to safe drinking water with filtration systems or alternative sources.
- Strengthen environmental laws to align with international PTE standards. Mandate periodic environmental assessments with transparent reporting.
- Conduct awareness campaigns on PTE pollution risks and prevention. Engage communities in remediation efforts, such as phytoremediation projects.
- Deploy advanced monitoring systems such as GIS and real-time sensors. Analyse data using machine learning for pollution prediction, and optimised remediation.

Future research should prioritise longitudinal monitoring of PTE pollution to capture seasonal variations and long-term trends, providing a more comprehensive means of environmental monitoring. Researchers may also need to study the bioavailable fraction and mobility of PTEs within soil systems to better evaluate potential risk impacts on human health and ecosystems. It is also vital to utilise advanced source apportionment techniques to determine exactly how industrial activities and agricultural practices lead to soil contamination when designing targeted interventions. Furthermore, to combat pollution and guarantee sustainable soil management, stakeholders in the area need to develop and evaluate ecofriendly remediation methods which fit specific local environmental conditions and socio-economic situations for the most affected parts of the area.

Declaration of competing interest

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

CRediT authorship contribution statement

Raymond Webrah Kazapoe: Writing – review & editing, Writing – original draft, Visualization, Supervision, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Benatus Norbert Mvile:** Writing – review & editing, Writing – original draft, Validation, Supervision, Methodology, Investigation, Data curation, Conceptualization. **John Desderius Kalimenze:** Writing – review & editing, Writing – original draft, Formal analysis, Data curation. **Daniel Kwayisi:** Writing – review & editing, Writing – original draft, Methodology. **Samuel Dzidefo Sagoe:** Writing – review & editing, Writing – original draft, Methodology. **Kwabina Ibrahim:** Writing – review & editing, Writing – original draft. **Obed Fiifi Fynn:** Writing – original draft, Visualization, Software.

References

Abrahim, G.M.S., Parker, R.J., 2008. Assessment of heavy metal enrichment factors and the degree of contamination in marine sediments from Tamaki Estuary, Auckland, New Zealand. Environ. Monit. Assess. 136 (1), 227–238.

Abu, M., Kalimenze, J., Mvile, B.N., Kazapoe, R.W., 2021. Sources and pollution assessment of trace elements in soils of the central, Dodoma region, East Africa: implication for public health monitoring. Environ. Technol. Innov. 23, 101705.

Abu, M., Mvile, B.N., Kalimenze, J.D., 2024. Provenance studies of Au-bearing stream sediments and performance assessment of machine learning-based models: insight from whole-rock geochemistry central Tanzania, East Africa. Environ. Earth. Sci. 83 (3), 105.

Adimalla, N., Gao, Y., Wang, Z., Qian, H., 2024. Spatial distribution, contamination characteristics, and potential ecological risk assessment of trace metals in surface soils of south-central stretch of India. Water, Air, & Soil Pollution 235 (8), 522.
Ahmad, F., Saeed, Q., Shah, S.M.U., Gondal, M.A., Mumtaz, S., 2022. Environmental sustainability: challenges and approaches. Nat. Resources Conservation and Adv. Sustain. 243–270.
Akhtar, N., Syakir Ishak, M.I., Bhawani, S.A., Umar, K., 2021. Various natural and anthropogenic factors responsible for water quality degradation: a review. Water. (Basel) 13 (19), 2660.
Alloway, B.J., 2012. Heavy Metals in soils: Trace metals and Metalloids in Soils and Their Bioavailability. Springer Science & Business Media.
Astel, A., Tsakovski, S., Barbieri, P., Simeonov, V., 2007. Comparison of self-organizing maps classification approach with cluster and principal components analysis for large environmental data sets. Water Res. 41 (19), 4566–4578.
Banzi, F. P., Msaki, P. K., Mohammed, N. K., 2015. Distribution of heavy metals in soils in the vicinity of the proposed Mkuju uranium mine in Tanzania. Environ Pollut. 4 (3), 42.
Wieczorek, J., Baran, A., 2022. Pollution indices and biotests as useful tools for the evaluation of the degree of soil contamination by trace elements. J. Soils Sediments. 1–18.
Cai, S., Shen, Z., Zhou, S., Wang, Q., Cheng, J., Yan, X., Tan, M., Tu, G., Cen, Y., 2024. Health risk assessment and potential sources of metals in riparian soils of the Wujiang River, China. Environ. Geochem. Health 46 (3), 106.
CCME, 2007. Canadian Soil Quality Guidelines for the Protection of Environmental and Human Health. Canadian Council of Ministers of the Environment, Winnipeg.
Davies, D.L., Bouldin, D.W., 1979. A cluster separation measure. IEEE Trans. Pattern Anal. Mach. Intell. 1, 224–227.
Ding, X., Ye, S., Yuan, H., Krauss, K.W., 2018. Spatial distribution and ecological risk assessment of heavy metals in coastal surface sediments in the Hebei Province offshore area, Bohai Sea, China. Mar. Pollut. Bull. 131, 655–661.
Eades, N.W., Reeve, W.H., 1938. Explanation of the geology of degree sheet No. 29 (Singida). Tanganyika Department of Lands and Mines. Geological Division.
Gan, T., Zhao, H., Ai, Y., Zhang, S., Wen, Y., Tian, L., Mipam, T.D., 2023. Spatial distribution and ecological risk assessment of heavy metals in alpine grasslands of the Zoige Basin, China. Front. Ecol. Evol. 11, 1093823.
Gao, Z., Tan, M., Liu, J., Zhang, Y., Niu, Y., Jiang, B., 2024. Characterization of soil trace metal pollution, source identification, and health risk assessment in the middle reaches of the Guihe River Basin. Environ. Monit. Assess. 196 (2), 122.
Gong, Q., Deng, J., Xiang, Y., Wang, Q., Yang, L., 2008. Calculating pollution indices by heavy metals in ecological geochemistry assessment and a case study in parks of Beijing. J. China Univ. Geosci. 19 (3), 230–241.
Gulgundi, M.S., Shetty, A., 2019. Source apportionment of groundwater pollution using unmix and positive matrix factorization. Environ. Processes 6, 457–473.
Hakanson, L., 1980. An ecological risk index for aquatic pollution control. A sedimentological approach. Water Res. 14 (8), 975–1001.
Haque, S. E., Shahriar, M. M., Nahar, N., Haque, M. S., 2022. Impact of brick kiln emissions on soil quality: A case study of Ashulia brick kiln cluster, Bangladesh. Environ. Chall. 9, 100640.
Henckel, J., Poulsen, K. H., Sharp, T., Spora, P., 2016. Lake victoria goldfields. Episodes 39 (2), 135–154.
Herman, A., Kihampa, C., 2015. Heavy metals contamination in soils and water in the vicinity of small scale gold mines at londoni and sambaru, Singida region, Tanzania. Int. J. Environ. Monitoring and Analysis 3 (6), 397.
Hoque, M. M., Islam, A., Islam, A. R. M. T., Pal, S. C., Mahammad, S., Alam, E., 2023. Assessment of soil heavy metal pollution and associated ecological risk of agriculture dominated mid-channel bars in a subtropical river basin. Sci. Rep. 13 (1), 11104.
Inengite, A. K., Abasi, C. Y., Walter, C., 2015. Application of pollution indices for the assessment of heavy metal pollution in flood impacted soil. Int. Res. J. Pure Appl. Chem 8 (3), 175–189.
Jiménez-Vázquez, A., Jaimes-López, R., Morales-Bautista, C.M., Pérez-Rodríguez, S., Gochi-Ponce, Y., Estudillo-Wong, L.A., 2025. Catalytic applications of natural iron oxides and hydroxides: a review. Catalysts. 15 (3), 236.
Jones, K.L., Ziegler, B.A., Davis, A.M., Cozzarelli, I.M., 2023. Attenuation of barium, strontium, cobalt, and nickel plumes formed during microbial iron reduction in a crude-oil-contaminated aquifer. ACS Earth and Space Chemistry 7 (7), 1322–1336.
Kaaya, N.I., Vegi, M.R., Macheyeki, A.S., 2025. Health risks of geogenic contaminants in gold mining areas in Geita, Tanzania. J. Trace Elem. Med. Biol., 100222.
Kabata-Pendias, A., Dudka, S., Chlopecka, A., 2017. Background levels and environmental influences on trace metals in soils of the temperate humid zone of Europe. In: Biogeochemistry of Trace Metals. CRC Press, pp. 73–96.
Kalimenze, J.D., Mvile, B., 2024. The investigations of ore and tailing samples for improving extracting methods and enhancing gold recovery of artisanal and small-scale miners: a case study of Singida Region in Tanzania, East Africa. J. Mines Geosci. 1, 44–72.
Kalimenze, J.D., Abu, M., Mvile, B.N., 2023. Soil geochemistry and multivariate statistical assessment of copper–Gold-PGEs mineralization in parts of Singida Region of the Tanzania Craton, Tanzania, East Africa. Arab. J. Geosci. 16 (1), 59.
Karakas, F., Imamoglu, I., Gedik, K., 2017. Positive matrix factorization dynamics in fingerprinting: a comparative study of PMF2 and EPA-PMF3 for source apportionment of sediment polychlorinated biphenyls. Environ. Pollution 220, 20–28.
Karungamye, J., Rwiza, M., Selemani, J., Marwa, J., 2023. Geochemistry of Potentially Toxic Elements in Soil and Sediments of a Tanzanian Small-Scale Gold Mining Area.
Kazapoe, R., Arhin, E., 2021. Determination of local background and baseline values of elements within the soils of the Birimian terrain of the Wassa Area of Southwest Ghana. Geology, Ecology, and Landscapes 5 (3), 199–208.
Kazapoe, R.W., Amuah, E.E.Y., Dankwa, P., Ibrahim, K., Mville, B.N., Abubakari, S., Bawa, N., 2021a. Compositional and source patterns of potentially toxic elements (PTEs) in soils in southwestern Ghana using robust compositional contamination index (RCCI) and k-means cluster analysis. Environ. Challenges 5, 100248.
Kazapoe, R.W., Arhin, E., Amuah, E.E.Y., 2021b. Known and anticipated medical geology issues in Ghana. Ecofeminism and Climate Change 2 (4), 169–184.
Kazapoe, R.W., Amuah, E.E.Y., Dankwa, P., 2022. Sources and pollution assessment of trace elements in soils of some selected mining areas of southwestern Ghana. Environ. Technol. Innov. 26, 102329.
Kazapoe, R. W., Addai, M. O., Amuah, E. E. Y., Dankwa, P., 2023. Hydrogeochemical characterization of groundwater in the Wassa Amenfi East and Prestea-Huni Valley areas of southern Ghana using GIS-based and multivariate statistical techniques. Sustain. Water Resour. Manag. 9 (5), 141.
Kazapoe, R.W., Amuah, E.E.Y., Dankwa, P., Fynn, O.F., Addai, M.O., Berdie, B.S., Douti, N.B., 2024. Fluoride in groundwater sources in Ghana: a multifaceted and country-wide review. Heliyon. 10 (13).
Kibassa, D., Kimaro, A. A., Shemdoe, R. S., 2013. Heavy metals concentrations in selected areas used for urban agriculture in Dar es Salaam, Tanzania. Sci. Res. Essays. 8 (27), 1296–1303.
Kohonen, T., 1982. Self-organized formation of topologically correct feature maps. Biol. Cybern. 43 (1), 59–69.
Kohonen, T., 2001. Self-Organizing Maps, 3rd ed Springer, Berlin/Heidelberg, Germany.
Kowalska, J.B., Mazurek, R., Gąsiorek, M., Zaleski, T., 2018. Pollution indices as useful tools for the comprehensive evaluation of the degree of soil contamination–A review. Environ. Geochem. Health 40, 2395–2420.
Kumar, V., Singh, J., Kumar, P. (Eds.), 2020. Environmental Degradation: Causes and Remediation Strategies (Vol. 1). Agro Environ Media, Publication Cell of AESA, Agriculture and Environmental Science Academy.
Kwayisi, D., Kazapoe, R.W., Alidu, S., Sagoe, S.D., Umaru, A.O., Amuah, E.E.Y., Kpiebaya, P., 2024. Exploring soil pollution patterns in Ghana's northeastern mining zone using machine learning models. J. Hazard. Mater. Adv. 16, 100480.
Lawley, C. J. M., Selby, D., Condon, D., Imber, J., 2014. Palaeoproterozoic orogenic gold style mineralization at the Southwestern Archaean Tanzanian cratonic margin, Lupa Goldfield, SW Tanzania: Implications from U–Pb titanite geochronology. Gondwana Res. 26 (3–4), 1141–1158.
Li, C., Sanchez, G.M., Wu, Z., Cheng, J., Zhang, S., Wang, Q., Li, F., Sun, G., Meentemeyer, R.K., 2020. Spatiotemporal patterns and drivers of soil contamination with heavy metals during an intensive urbanization period (1989–2018) in southern China. Environ. Poll. 260, 114075.
Liao, J., Cui, X., Feng, H., Yan, S., 2021. Environmental background values and ecological risk assessment of heavy metals in watershed sediments: a comparison of assessment methods. Water, 14 (1), 51.
Machiwa, J.F., 2010. Heavy metal levels in paddy soils and rice (Oryza sativa (L)) from wetlands of Lake Victoria Basin, Tanzania. Tanz. J. Sci. 36.
Mahvi, A. H., Eslami, F., Baghani, A. N., Khanjani, N., Yaghmaeian, K., Mansoorian, H. J., 2022. Heavy metal pollution status in soil for different land activities by contamination indices and ecological risk assessment. Int. J. Environ. Sci. Technol. 19 (8), 7599–7616.
Manya, S., 2012. SHRIMP zircon U-Pb dating of the mafic and felsic intrusive rocks of the Saza area in the Lupa goldfields, southwestern Tanzania: Implication for gold mineralization. Nat. Sci. 4 (9), 724–730.
Mas, S., de Juan, A., Tauler, R., Olivieri, A.C., Escandar, G.M., 2010. Application of chemometric methods to environmental analysis of organic pollutants: a review. Talanta 80 (3), 1052–1067.
Mihale, M. J., 2019. Multivariate assessment of metal contamination in mtoni sedimentary environment, Tanzania. Tanz. J. Sci. 45 (2), 158–172.
Mnali, S.R., 2001. Assessment of heavy metal pollution in the lupa gold field, SW Tanzania. Tanzania J. Sci. 27 (2), 15–22.
Mng'ong'o, M., Comber, S., Munishi, L. K., Ndakidemi, P. A., Blake, W., Hutchinson, T. H., 2021. Potentially toxic elements status and distribution in Usangu agroecosystem–Tanzania. Environ. Chall. 4, 100200. doi:10.1016/j.envc.2021.100200.
Mng'ong'o, M., 2022. Comparative assessment of soil phosphate status, water eutrophication, and potentially toxic metal accumulation in Usagu agro-ecosystem (Doctoral dissertation, NM-AIST).
Mugheri, M.H., Pathan, A.M., Sayed, M.A., Maira, M., Soomro, D.B., Ali amur, S., Soomro, N.A., 2019. Assessment of drinking water quality district Jamshoro Sindh Pakistan: A case study. Int. J. Curr. Res. 11 (3), 1812–1816.
Mvile, B.N., Abu, M., Kalimenze, J.D., 2023. Assessment of heavy metals concentration in soils in the central parts of Tanzania using pollution indices and multivariate statistical approach: implication on the source and health. J. Sediment. Environ. 8 (3), 457–469.
National Bureau of Statistics, 2013. 2012 Population and Housing Census: Population distribution By Administrative Areas. Government of Tanzania, Dar es Salaam, Tanzania Retrieved from. https://www.nbs.go.tz.
National Bureau of Statistics, 2022. 2022 Population and housing census: Population distribution by administrative units. Government of Tanzania, Dodoma, Tanzania Retrieved from. https://www.nbs.go.tz.

Nyika, J., Dinka, M.O. (Eds.), 2023. Global Industrial Impacts of Heavy Metal Pollution in Sub-Saharan Africa. IGI Global.
Pan, Y., Ding, L., Xie, S., Zeng, M., Zhang, J., Peng, H., 2021. Spatiotemporal simulation, early warning, and policy recommendations of the soil heavy metal environmental capacity of the agricultural land in a typical industrial city in China: case of Zhongshan City. J. Clean. Prod. 285, 124849.
Qingjie, G., Jun, D., Yunchuan, X., Qingfei, W., Liqiang, Y., 2008. Calculating pollution indices by heavy metals in ecological geochemistry assessment and a case study in parks of Beijing. J. China Univ. Geosci. 19 (3), 230–241.
Rai, R., Sharma, S., Gurung, D. B., Sitaula, B. K., 2019. Heavy metal contamination in sediments from vehicle washing: a case study of Olarong Chhu Stream and Paa Chhu River, Bhutan. Int. J. Environ. Stud. 76 (1), 66–83.
Rashed, M. N., 2010. Monitoring of contaminated toxic and heavy metals, from mine tailings through age accumulation, in soil and some wild plants at Southeast Egypt. J. Hazard. Mater. 178 (1–3), 739–746.
Rodrigue, K. A., Yao, B., Trokourey, A., Kopoin, A., 2016. Assessment of heavy metals contamination in sediments of the Vridi Canal (Côte d'Ivoire). Journal of Geoscience and Environment Protection 4 (10), 65–73.
Rudnick, R.L., Gao, S., 2003. Composition of the continental crust. Treatise on Geochemistry vol. 3, 1–64.
Sethi, S., Gupta, P., 2020. Soil contamination: a menace to life. Soil Contamination. IntechOpen.
Sharley, D.J., Sharp, S.M., Bourgues, S., Pettigrove, V.J., 2016. Detecting long-term temporal trends in sediment-bound trace metals from urbanised catchments. Environ. Poll. 219, 705–713.
Sivakumar, S., Chandrasekaran, A., Balaji, G., Ravisankar, R., 2016. Assessment of heavy metal enrichment and the degree of contamination in coastal sediment from South East Coast of Tamilnadu, India. J. Heavy Metal Toxicity Diseases 1 (2), 1–8.
Song, Y., Xie, S., Zhang, Y., Zeng, L., Salmon, L.G., Zheng, M., 2006. Source apportionment of PM2. 5 in Beijing using principal component analysis/absolute principal component scores and UNMIX. Sci. Total Environ. 372 (1), 278–286.
Tanzania Bureau of Standards (2007). Soil quality limits for Soil contaminants in habitat and agriculture TZS 972:2007 (E).
Tanzania Bureau of Standards. (2007). TZS 972:2007: soil quality – Limits for soil contaminants in habitat and agriculture. Dar es Salaam, Tanzania.
Tarawneh, K., Alnawafleh, H., Harahsheh, M., Radaideh, N., Shreideh, S., Moumani, K., Tarawneh, B., Abdelghafoor, M., 2011. Mineralogy, geochemistry and genesis of the ferruginous sandstone in Batn Al Ghul Area/Southern Jordan. Ann Univ Mining Geol 54 (I), 89–94.
Taylor, S.R., 1964. Abundance of chemical elements in the continental crust: a new table. Geochim. Cosmochim. Acta. 28 (8), 1273–1285.
USEPA, 2023. Regional Screening Levels (RSLs) – Generic Tables (May 2023). U.S. Environmental Protection Agency, Office of Superfund Remediation and Technology Innovation.
Van Straaten, P., 2000. Mercury contamination associated with small-scale gold mining in Tanzania and Zimbabwe. Sci. Total Environ. 259 (1–3), 105–113.
VROM, 2000. Circular On Target Values and Intervention Values For Soil Remediation. Ministry of Housing, Spatial Planning and Environment, Netherlands.
Wang, N., Liu, Z., Sun, Y., Lu, N., Luo, Y., 2024. Analysis of soil fertility and toxic metal characteristics in open-pit mining areas in northern Shaanxi. Sci. Rep. 14 (1), 2273.
Wang, Y., 2023. Ecological risk identification and assessment of land remediation project based on GIS technology. Environ. Sci. Poll. Res. 30 (27), 70493–70505.
Weissmannová, H.D., Pavlovský, J., 2017. Indices of soil contamination by heavy metals–methodology of calculation for pollution assessment (minireview). Environ. Monit. Assess. 189 (12), 616.
Wightwick, A., Mollah, M., Smith, J., MacGregor, A., 2006. Sampling considerations for surveying copper concentrations in Australian vineyard soils. Soil Res. 44 (7), 711–717.
Wu Jin, W.J., Teng YanGuo, T.Y., Lu SiJin, L.S., Wang YeYao, W.Y., & Jiao XuDong, J.X. (2014). Evaluation of soil contamination indices in a mining area of Jiangxi, China.
Xiang, M., Li, Y., Yang, J., Lei, K., Li, Y., Li, F., Zheng, D., Fang, X., Cao, Y., 2021. Heavy metal contamination risk assessment and correlation analysis of heavy metal contents in soil and crops. Environ. Poll. 278, 116911.
Yang, Z., Zou, L., Xia, J., Qiao, Y., Bai, F., Wang, Q., Cai, D., 2022. Spatiotemporal variation characteristics and source identification of water pollution: insights from urban water system. Ecol. Indic. 139, 108892.
Zaller, J. G., Zaller, J. G., 2020. Pesticide impacts on the environment and humans. *Daily poison: pesticides-an underestimated*. danger, pp. 127–221.
Zhang, Y., Hou, D., O'Connor, D., Shen, Z., Shi, P., Ok, Y.S., Tsang, D.C., Wen, Y., Luo, M., 2019. Lead contamination in Chinese surface soils: source identification, spatial-temporal distribution and associated health risks. Crit. Rev. Environ. Sci. Technol. 49 (15), 1386–1423.
Zhu, H., Liu, X., Wang, Q., Zhang, B., Xu, C., Wang, Z., Chen, H., 2023. Heavy metals pollution of soil in central plains urban agglomeration (CPUA), China: human health risk assessment based on Monte Carlo simulation. Environ. Geochem. Health 45 (11), 8063–8079.