



Receptor model-based source-specific health risks of toxic metal(loid)s in coal basin-induced agricultural soil in northwest Bangladesh

Md. Ahasan Habib · Abu Reza Md. Towfiqul Islam · Memet Varol · Khamphe Phoungthong · Rahat Khan · Md. Saiful Islam · Md. Hasanuzzaman · Md. Yousuf Mia · Romulus Costache · Subodh Chandra Pal

Received: 14 May 2023 / Accepted: 17 August 2023 / Published online: 30 August 2023
© The Author(s), under exclusive licence to Springer Nature B.V. 2023

Abstract Toxic metal(loid)s (TMLs) in agricultural soils cause detrimental effects on ecosystem and human health. Therefore, source-specific health risk apportionment is very crucial for the prevention and control of TMLs in agricultural soils. In this study, 149 surface soil samples were taken from a coal mining region in northwest Bangladesh and analyzed for 12 TMLs (Pb, Cd, Ni, Cr, Mn, Fe, Co, Zn, Cu, As, Se, and Hg). Positive matrix factorization (PMF)

and absolute principal component score-multiple linear regression (APCS-MLR) receptor models were employed to quantify the pollution sources of soil TMLs. Both models identified five possible sources of pollution: agrochemical practice, industrial emissions, coal-power-plant, geogenic source, and atmospheric deposition, while the contribution rates of each source were calculated as 28.2%, 17.2%, 19.3%, 19% and 16.3% in APCS-MLR, 22.2%, 13.4%, 24.3%, 15.1% and 25.1% in PMF, respectively. Agrochemical practice was the major source of non-carcinogenic risk (NCR) (adults: 32.37%, children: 31.54%),

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10653-023-01740-3>.

M. Habib · K. Phoungthong
Industrial Ecology in Energy Research Center, Faculty of Environmental Management, 10 Prince of Songkla University, Songkhla 90112, Thailand

M. Habib
Geological Survey of Bangladesh, Government of the People's Republic of Bangladesh, 153 Pioneer Road, Seghunbaghicha, Dhaka 1000, Bangladesh

A. R. M. Islam · M. Hasanuzzaman · M. Mia
Department of Disaster Management, Begum Rokeya University, Rangpur 5400, Bangladesh

A. R. M. Islam
Department of Development Studies, Daffodil International University, Dhaka 1216, Bangladesh

M. Varol (✉)
Agriculture Faculty, Department of Aquaculture, Malatya Turgut Özal University, Malatya, Turkey
e-mail: mvarol23@gmail.com

R. Khan
Institute of Nuclear Science and Technology, Bangladesh Atomic Energy Commission, Savar, Dhaka 1349, Bangladesh

M. Islam
Department of Soil Science, Patuakhali Science and Technology University, Dumki, Patuakhali 8602, Bangladesh

R. Costache
Department of Civil Engineering, Transilvania University of Brasov, 5, Turnului Str, 500152 Brasov, Romania

R. Costache
Danube Delta National Institute for Research and Development, 165 Babadag Street, 820112 Tulcea, Romania

S. C. Pal
Department of Geography, The University of Burdwan, Bardhaman, West Bengal 713104, India

while atmospheric deposition was the highest source of carcinogenic risk (CR) (adults: 48.83%, children: 50.11%). NCR and CR values for adults were slightly higher than for children. However, the trends in NCR and CR between children and adults were similar. As a result, among the sources of pollution, agrochemical practices and atmospheric deposition have been identified as the primary sources of soil TMLs, so prevention and control strategies should be applied primarily for these pollution sources in order to protect human health.

Keywords Soil toxic metal(loid)s · Barapukuria coalfield · APCS-MLR and PMF models · Source apportionment · Source-specific health risk assessment

Introduction

Coal is one of the most important energy sources in many parts of the world, particularly in developing countries like Bangladesh (Habib & Khan, 2021; Tozsin, 2014). Although Bangladesh owns a minimal reserve of gas and coal, its effect has always been a significant concern for the ecosystem via water, air, and soil pollution (Bhuiyan et al., 2010; Bilgen, 2014). With the rapid development of the coal industry and urbanization, as well as coal production and combustion activities, soil toxic metal(loid)s (TMLs, hereinafter) pollution has attracted global attention due to its toxicity, distribution, and long-term persistence (Adimalla, 2020; Liu et al., 2021; Parzenty & Róg, 2021). Coal mine-induced soils may degrade soil quality (Wang et al., 2020). They enter the environment via geogenic sources, including weathering and pedogenic processes (Duan et al., 2020). Anthropogenic sources such as coal combustion activities (Singh et al., 2018), coal-power-plants (Zhang et al., 2021); agrochemical practice (Huang et al., 2018), industrial emissions (Wang et al., 2021), coal exhaust (Yang et al., 2019) have released a significant fraction of TMLs into the soils. Besides, TMLs can cause a possible health risk to humans via external exposure, breathing, and the food web (Guan et al., 2018; Liu et al., 2019). TMLs may harm human health because some (i.e., As, Cd, Hg, and Cr) are carcinogenic and mutagenic (Qi et al., 2020). To lessen the health hazard of TMLs, it is crucial to identify and understand

their abundance, distribution, origin, and health issues in the natural ecosystem (Yang et al., 2021). Therefore, noticeably assessing the source-specific health problems caused by TMLs is the preliminary prerequisite for developing targeted risk mitigation management and strategies (Sun et al., 2021).

To figure out where TMLs come from in soil, several different methods (like chemometric techniques and receptor models) have been used in recent years (Duan et al., 2020; Guan et al., 2018; Guo et al., 2021). USEPA-recommended positive matrix factorization (PMF) was used to measure metal(loid) pollution source apportionment (Khan et al., 2023; USEPA, 2018). PMF reliably identifies metal pollution sources (Islam et al., 2020; Kumar et al., 2021). The multivariate curve resolution weighted alternating least-squares model and chemical mass balance vary from PMF. Unlike previous models, it employs a weighted least-square fit using the well-known error measures of the metal(loid)s to generate composition and contribution profiles. It ensures that all profile values are non-negative (Zhang et al., 2020). This model can quickly determine TMLs and construct metal(loid) remediation solutions. The PMF model has classified metal(loid) sources in several environmental systems (Lv, 2019; Xiao et al., 2020; Yang et al., 2019, 2021).

Recently, a new chemometric technique, the absolute principal component score-multiple linear regression (APCS-MLR) model, has been introduced to quantify pollution sources (Chen et al., 2016). APCS-MLR, like the PMF model, may categorize samples and identify pollution sources by reducing a large dataset into indicators with fewer dimensions (Lv, 2019; Proshad et al., 2021). Several earlier studies found that the APCS-MLR model could produce similar results to the PMF method when identifying TML sources and suggested that it might be a better option for pollution source identification studies (Duan et al., 2020; Proshad et al., 2021; Zhang et al., 2020). A recent study indicated that PMF with positive values was more accurate for source identification than APCS-MLR (Guo et al., 2021), whereas APCS-MLR was better for source identification than PMF (Su et al., 2021). Thus, both techniques yield more consistent source distribution and identification with different ranking and contribution rates (Proshad et al., 2021). Early research focused on pollution source contributions to soil metal(loid) abundance.

However, they overlooked source-related risks. TMLs have a wide range of spatial patterns, toxicity levels, and environmental hazards, therefore various studies have shown that high TML concentrations in soils may not necessarily harm to the ecosystem and living things (Guan et al., 2018; Huang et al., 2018). Source-oriented risk assessment improves targeted risk reduction since risk levels vary by source (Yang et al., 2019). Few studies have examined the source-oriented health risk of TMLs using source apportionment and health risk evaluation (Duan et al., 2020; Guo et al., 2021). The literature lacks a scientific understanding of the combined nexus between source apportionment and the soil TML risk assessment. The literature on TML source-specific health concerns from soil surrounding the chosen coal basin, northwest Bangladesh, is still scarce.

The research region has undergone rapid urbanization and industrial and agricultural activity. The coal industrial region of Barapukuria may threaten ecology, the environment, and human health since coal production and electricity generation harm soil, water, and agriculture. With significant crop output, the area's agricultural economy dominates the coal industry's fast growth (BBS, 2018). Several studies have found moderate to severe TML contamination caused by coal processing, preparation, and utilization activities and combustion residuals on soil, water, sediments, crops, and plants (e.g., Bhuiyan et al., 2010; Habib et al., 2019a; Halim et al., 2015; Hossen et al., 2021; Rahman et al., 2017). This research hypothesized that agrochemical application, air deposition, and industrial pollutants may harm population health. Source-oriented risk apportionment of soil TMLs might confirm this. The prime objectives are to (1) determine the concentrations of soil physicochemical and TMLs and their spatial variations using principal coordination analysis (PCoA) and a self-organizing map (SOM) in the coal mine region and (2) investigate the co-occurrence relationships of metal(loid)s using network inference. (3) apportion the possible sources of TMLs via comparing the PMF and APCS-MLR models along with ranking and contribution rates; (4) measure the source-specific health risks of soil TMLs by coupling their source apportionment into a probabilistic human health risk model; (5) address the probability distribution of carcinogenic risk factors using Monte Carlo simulation to lessen the uncertainties of health risk appraisal.

Source apportionment and health risk assessment of TMLs in the soil surrounding the designated coal basin in northwest Bangladesh are yet to be scientifically understood. This study is the first to determine soil metal(loid)s contamination in the designated coal basin using both PMF and APCS-MLR models and the source-specific health risks of TMLs in a coal basin-derived agricultural soil from northwest Bangladesh. This research will help identify toxic element sources in coal industry-associated soils and how to regulate toxic element contamination. Finally, this study will help scientists, particularly geologists, planners, and politicians, manage coal mining soil in Bangladesh, reduce TML health concerns, and meet the UN SDGs.

Material and methods

Study area

Structurally the sampling area is situated within the southern slope of the Rangpur Saddle, which is a part of the stable platform of Bengal Basin (Habib et al., 2019a) (Fig. 1). The area is located in a humid subtropical region (mean annual precipitation: 1104–2985 mm) in a moderate to densely populated (823 people per km²) and an agriculture dominant farming area in the NW region of Bangladesh (Barapukuria, Dinajpur) (Fig. 1). Municipalities (Parbatipur and Phulbari) and Maddhapara Granite Mine are distributed within the sampling area. Physiographically, the area comprises dominantly Pleistocene Barind Clay Residuum (terrace deposits) and Holocene Alluvial flood plains of the Tista-Karotoya River system (major tributaries of the Brahmaputra-Jamuna River). It is level to gently sloping southward and comparatively well-drained, and elevation ranges from 27 to 31 m above mean sea level. The soils contain light-brown to orange, light-red, yellowish-grey to bluish-grey, various shades of yellow, red, orange, and white non-calcareous sticky silty clay; subordinated with a minor proportion of grey to yellowish-grey fine-grained sand having root tubes, burrows, vugs, and fractures and Fe-oxides nodules with a considerable fraction of grey soils primarily used as rice cultivation. Most samples belong to the silty loam or silt classes (Habib et al., 2019a). According to USDA taxonomy, it may correlate to Aerie Haplaquept. It is

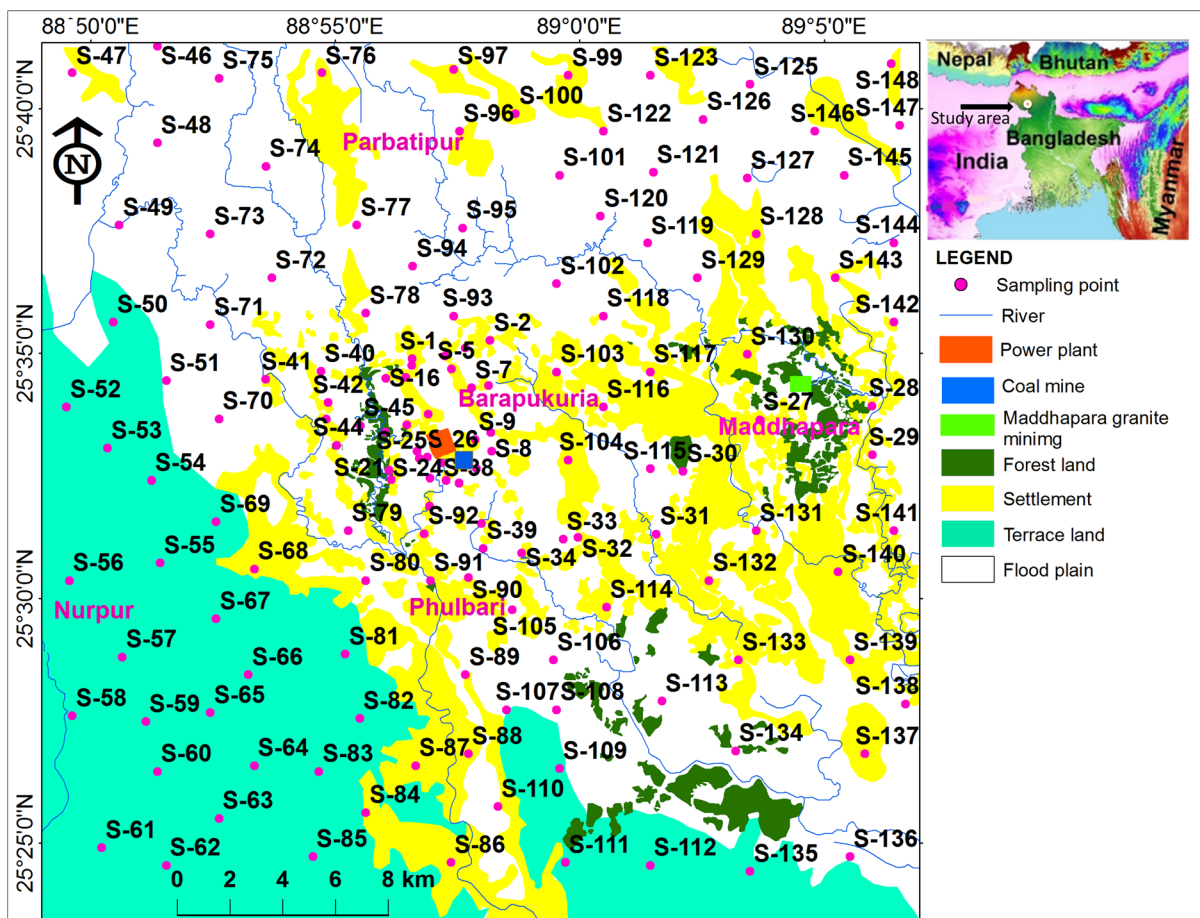


Fig. 1 Map showing the study area and soil sampling sites from coal mine and hard rock mine region, Bangladesh with the regional context

of Amnura, Belabo, Noadda, Jagdal soil series association. The predominant wind direction is generally from east to west (40%), followed by the west to east (25%) and north-east (18%) and the speed is relatively moderate. The nearby active rivers are namely Atrai, Banglai, Jabuneswari, Kala, Kharkhari, Tilai, Chirnai, and Little Jamuna-Tulsiganga (distributaries of the large river Tista), flowing over the area from north to south (Habib & Khan, 2021; Habib et al., 2019b).

Sample collection and processing

In the selected sites, top soils (uncultivated and agricultural) ($n = 149$) were collected near an operational coal mine and power-plant surrounding (Fig. 1) by following the systematic random sampling technique (Ćujić et al., 2016) where regularly spaced intervals

were ensured carefully. Each of the sampling site was separated from each other by ~ 2 km and the total sampling area covers approximately 1000 km² considering the all-directional areas centering the coal-mine and coal-based power plant (Fig. 1). A 10 cm \times 10 cm sampling auger (length: 30 cm) was used to collect the soil samples. From each site, three sub-samples were taken from the same area up to a depth of 10–20 cm and thoroughly blended to form a representative composite replicate sample of that particular spot (Rodríguez-Irretagoiena et al., 2015). After sampling, the samples were immediately placed and sealed in polythene bags with appropriate labeling to avoid probable weathering and cross-contamination, and transported to the laboratory and stored in the refrigerator until further analysis (Özkul, 2016). All samples were then dried (except for samples for

Hg analysis) at room temperature and well mixed, then milled by pre-cleaned mortar and pestle, ensuring no cross-contamination followed by sieving and homogenization.

Analytical processes and quality control

Physicochemical parameters, e.g., pH of an aqueous suspension of soil samples (1:2.5 = soil: distilled water, w/v), were determined by pH meter (Inesa, PHS-2F, China; Habib et al., 2019b) and total OM (OM) of samples were measured by mass loss upon ignition of dried samples in a furnace at 550 °C for at least six-hour. Additionally, the traditional pipette method was applied for particle size analysis. However, elemental analysis of Cr, Mn, Fe, Co, Zn and As were conducted using instrumental neutron activation analysis (INAA). In contrast, Pb, Cd, Ni, Cu and Se were measured by inductively coupled plasma-optical emission spectrometry (ICP-OES: Optima 800, Perkin Elmer Instruments, USA). Briefly, in INAA, ~60 mg of sample was irradiated by neutrons (of TRIGA Mark II research reactor, Bangladesh Atomic Energy Commission) followed by γ -ray counting (High purity Ge-detector) with systematic nuclear decay schemes (Khan et al., 2021). On the other hand, ~HNO₃+HClO₄+HF+HCl based digested sample (Khan et al., 2015) solution was used for ICP-OES measurements (Rodriguez-Iruretagoien et al., 2015). However, a Mercury analyzer (FIMS 200, Perkin Elmer, USA) was used to determine the Hg-contents in the samples. Triplicate measurements of reference material (IAEA-Soil-7) were used to check the data quality of different analytical techniques. Analytical data of Soil-7 (of this study) were in good agreement (2–15%) with those of certificate values, whereas reproducibility (RSDs in %) was within 10%. Detection limits (3 σ) of Pb, Cd, Ni, Cr, Mn, Fe (%), Co, Zn, Cu, As, Se, and Hg were 1.2, 0.005, 0.08, 1.5, 10, 0.005%, 0.05, 1.0, 0.2, 0.01, 0.05, and 0.003 mg/kg, respectively.

Source contributions to soil TMLs

PMF model

In this study, the USEPA PMF5.0 (USEPA, 2014) was used to identify the source distribution of soil metal(loid)s. Under the non-negative limit, the

original matrix value was distributed/fragmented into a contribution matrix and a factor profile (Frie et al., 2017). The performance of this model is based on precise factorization algorithms that calculate the source profile and input, and it is often used to categorize metal(loid) sources in soils (Proshad et al., 2021). It uses uncertainty analysis as a receptor model to assist balance the dataset, because it doesn't need a source profile (Guo et al., 2021). The computed procedure and the mathematical function of PMF 5.0 are outlined in the Supplementary Material Text S1.

APCS-MLR model

The source distribution of soil TMLs was computed using an APCS-MLR model based on absolute principal component scores (APCS) and a multiple linear regression model (MLR). The APCS-MLR model was merged into the APCS and MLR models, both of which were based on the PCA approach. The PCA was used to get components from the same parameters using data dimensionality reduction. The rotational factor load of the metal(loid) was estimated as a starting point for pinpointing the metal's source (Guo et al., 2020). The PCA data was clearly not used to estimate the contribution rate of metal(loid) sources; hence the APCS technique had to be used as a non-standardized APCS method (Zhang et al., 2021). MLR includes the number of metals (loids) in soils as a dependent parameter, while APCS links the anticipated and actual values to ensure the model's accuracy (Proshad et al., 2021). The APCS-MLR model was described in detail in the Supplementary Material Text S2.

After the appropriate model was chosen, the mass contribution of soil metal(loid)s in each sampling location from definite sources was determined as follows:

$$C_{ij}^k = C_{ij}^{k*} \times C_i \tag{1}$$

where C_{ij}^k is the mass contents of the j th TMLs from k th source in the i th sample (mg/kg); C_{ij}^{k*} is the computed contribution of j th metal(loid)s from ij kth source in the i th sample (mg/kg), C_i is the concentrations of the soil TMLs in the i th sample (mg/kg).

The PMF and APCS-MLR models were both run 20 times with different beginning seeds, and the solution with the lowest objective function value was

chosen as the final solution (Zhang et al., 2020). The bootstrap (BS), displacement (DISP), and bootstrap-displacement tools pioneered by Paatero et al. (2014) were used to do the uncertainty analysis of the solution (BS-DISP).

Source-oriented health risk evaluation of soil TMLs

In this study, health risk evaluation and the source distribution were integrated to measure the source-oriented health risks to children (<6 years old) and adults (>18 years old) (USEPA, 2011). The average daily dose (ADD) of the j th TMLs from k th source in i th sample locations through ingestion (ADD_{ijing}^k), inhalation (ADD_{ijinh}^k), and dermal route (ADD_{ijder}^k) was computed by using Eqs. 2–4 (USEPA, 2011).

$$ADD_{ijing}^k = \frac{C_{ij}^k \times \text{IngR} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}} \times \text{CF} \quad (2)$$

$$ADD_{ijinh}^k = \frac{C_{ij}^k \times \text{InhR} \times \text{EF} \times \text{ED}}{\text{PEF} \times \text{BW} \times \text{AT}} \quad (3)$$

$$ADD_{ijder}^k = \frac{C_{ij}^k \times \text{SA} \times \text{ABS} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}} \times \text{CF} \quad (4)$$

The hazard quotient ($HQ_{ij,p}^k$) embodies the p th exposure route from k th source of j th TMLs in i th sample. The total hazard indexes (THI) is computed as the sum of hazard index (HI) (USEPA, 2011) in the following Eqs. 5–6:

$$\text{HI} = \sum \text{HQ}_{ij,p}^k \frac{ADD_{ijing}^k}{\text{RfD}_{ing}} + \frac{ADD_{ijinh}^k}{\text{RfD}_{inh}} + \frac{ADD_{ijder}^k}{\text{RfD}_{der}} \quad (5)$$

$$\text{THI} = \sum \text{HI} \quad (6)$$

where RfD_{ing} , RfD_{inh} , and RfD_{der} (mg/(kg day)) are the reference doses of resultant TMLs through ingestion, inhalation, and dermal route, respectively. If HQ, HI and THI are > 1, there is a possible detrimental impact on human health (USEPA, 2009).

The carcinogenic risk (CR) of the j th TMLs in the i th sample from k th source could be computed based on Eq. 7 (USEPA, 2011). Total lifetime carcinogenic risk (TCR) could be computed by the sum of CR according to Eq. 8 (USEPA, 2011).

$$\text{CR} = \sum \text{CR}_{ij,p}^k = \text{ADD}_{ijing}^k \times \text{SF}_{ing} + \text{ADD}_{ijinh}^k \times \text{SF}_{inh} + \text{ADD}_{ijder}^k \times \text{SF}_{der} \quad (7)$$

$$\text{TCR} = \sum \text{CR} \quad (8)$$

where SF_{ing} , SF_{inh} , and SF_{der} are the slope factors for CR, (mg/ (kg d)) for soil metal(loid)s. If CR and TCR surpassed 1×10^{-4} , there is a noteworthy carcinogenic risk, while the CR and TCR are $< 1 \times 10^{-6}$, no substantial carcinogenic impacts exist. When CR and TCR vary within 1×10^{-4} and 1×10^{-6} , indicating acceptable risk of cancer (USEPA, 2009). The detailed variables adopted in the human health risk evaluation are listed in Table S1. The standard reference dose and slope factors for each metal(loid) are given in Table S2.

Source-oriented health risk model

To report the source-specific risk assessment, the source distribution outcomes of both PMF and APCS-MLR models are equally integrated into the NCR and CR models. First of all, the contribution contents of soil TMLs for each source recognized by PMF and/or APCs-MLR can be computed in Eqs. 9–11.

$$\left(\begin{array}{l} \text{CF}_{ij}^{k1} = f_{ij}^{k1} \times C_{ij} \\ \sum_{k1=1}^p f_{ij}^{k1} = 1 \end{array} \right) \quad (9)$$

$$\left(\begin{array}{l} \text{CF}_{ij}^{k2} = f_{ij}^{k2} \times C_{ij} \\ \sum_{k2=1}^p f_{ij}^{k2} = 1 \end{array} \right) \quad (10)$$

$$\text{CF}_{ij}^k = \text{CF}_{ij}^{k1} \otimes \text{CF}_{ij}^{k2} \quad (11)$$

where CF^{k1} and CF^{k2} denote the content fraction of the j th TMLs in the i th sample attributed by the k th source originated from the PMF and APCS-MLR models, respectively; f^{k1} and f^{k2} embody the resultant contribution fraction of the k th pollution source to the j th TMLs in the i th sample; C_{ij} donates the measurement content of the j th TMLs in the i th sample; p is the number of recognized pollution source; CF_{ij}^k is the coupled estimation of content fraction from the both

models ($k=1,2,\dots,p$); \otimes is the composition operator (i.e., arithmetical men in our research).

Subsequently, the obtained content fractions of soil TMLs for each of the pollution sources in each sample are integrated into the NCR and CR models as represented in earlier Eqs. (2)–(8). Particularly, C_{ij} is substituted by CF_{ij} and the total carcinogenic risk (TCR) attributed by the k th identified source can be computed in the following Eq. 12:

$$TCR_j^k = \sum_{i=1}^n CR_{ij}^k \quad (12)$$

Statistical analyses

The TML concentrations were correlated using Spearman’s rank correlation. Using Gephi software (ver. 0.9.2), co-occurrence network analysis showed the TMLs relationship. For co-occurrence network analysis, only strong relationships ($r>0.2$, $p<0.01$) were evaluated (Islam et al., 2020). PCA was used to evaluate the best number of components for soil TML source distribution. All statistical tests were considered significant if the p value was 0.05 or below and the data were normalized using Origin Pro 2020 (Origin Lab, Northampton, MA). PCoA resulting from $\log(x+1)$ transformation was used to indicate overall TML change across sample sites and PRIMER (ver. 7.0.17) was used for this analysis (Clarke & Gorley, 2015). Source distribution was done by using PMF 5.0 (USEPA, 2014) and APCS-MLR was done by using SPSS (ver. 26.0). Both models were applied to get more consistent source distribution and identification with various ranking and contribution rates (Khan et al., 2023). Human body variability creates uncertainty in risk assessments (Chen et al., 2016; Qi et al., 2020). Monte Carlo simulation was used to lower the uncertainty of human health risk assessments (Islam et al., 2020). Table S1 lists the probabilistic variables used in this study. Crustal Ball software carried out the Monte Carlo analysis (ver. 11.1.2.3) with 10,000 iterations. To visualize a multi-dimensional dataset into a lower-dimensional representation, a self-organizing map (SOM), a non-linear approach, was applied to analyze the soil TMLs, which offers a graphical representation of the relationship between parameters. The SOM was drawn in MATLAB software (ver. R2019b) by a neural clustering toolbox. The detailed process of the SOM model can be found in the supplementary text S3.

Results and discussion

Descriptive statistics of physicochemical properties and TMLs in the soil

Table 1 provides an overview of the basic soil properties including pH, EC, OM, distribution/composition/pattern, and Pb, Cd, Ni, Cr, Mn, Fe, Co, Zn, Cu, As, Se, and Hg of soils. The pH, minerals (typically aluminosilicate, carbonates, clays, arsenopyrite, mica, pyrites, chlorite, biotite, feldspars, Fe–Mn Oxides), redox-states, organometallic complexes, and organic matter (OM) play a significant role in controlling the leaching and abundances/levels of some concern/toxic/hazardous/endangered chemical species in the soil. TMLs are accessible to the crops when the soil pH is acidic (Wang et al., 2018). The pH value ranged from 3.9 to 8.7 with a mean of 5.8 (Table 1). Most of the samples in the present investigation were acidic; 85% of them had pH values below 7.0, 14% were between 7.0 and 8.0, and just 1% were over 8.0. Acidic soils have developed in the region as a consequence of the large and widespread usage of nitrogen-containing fertilizers (Islam et al., 2021), and vegetation’s production of H⁺ ions in the root zone has also contributed to the acidity of the topsoil (Geetha et al., 2017). Moreover, the absorption of TMLs in the soil owing to the action of OM causes the cation exchange capacity of organics and subsequent generation of carbonic acid. (Islam et al., 2015).

Due to the cation exchange capacity of organics and the ensuing generation of organic acid, OM has also been shown to change the absorption of metals (loids) in soils (Islam et al., 2015). The samples had modest OM concentration, averaging 2.7% and ranging from 0.50 to 12.4%. According to the current research, low to medium levels of OM in soils suggest that TMLs are likely linked to OM to form complexes, which may also result in decreased TML availability to the surrounding ecosystem’s living organisms. The amount of organic carbon detected in the soils of the moderately farmed fields near the industrial area in Dhaka City, Bangladesh, Islam et al. (2014) reported to be found 0.62–3.6 mg/kg, somewhat higher than the level identified in the current study. According to Cai et al. (2015), OM in agricultural soil ranged from 0.48 to 7.33% with a mean value of 2.72%, which is comparable to our finding. When calculating the number of mobile/labile salts present in soils, EC is

Table 1 Descriptive summary of the TMLs in soils of coal mining region, Bangladesh ($n = 149$, mg/kg)

Parameter	Mean	S.E. Mean*	Std. Deviation*	RSDs (%)*	Skewness	Kurtosis	Minimum	Maximum	Percentiles
EC (dS/m)	0.2	0.02	0.23	115	3.42	14.38	0.01	1.4	0.42
pH	5.8	0.08	0.92	16	0.39	-0.2	3.9	8.7	7.09
OM (%)	2.66	0.2	2.41	91	2.02	4.13	0.5	12.4	6.3
Sand (%)	10.25	0.54	6.61	64	1.2	1.83	2	38	19
Silt (%)	71.14	0.96	11.68	16	-0.54	1.06	28	95	88.5
Clay (%)	21.95	1.21	14.83	68	0.54	-0.8	3.7	52.1	46.6
Pb (mg/kg)	39.73	2.49	30.42	77	1.7	2.06	14.9	160	97
Cd (mg/kg)	0.8	0.08	0.92	115	1.91	3.14	0.02	4.23	2.3
Ni (mg/kg)	47.15	2.67	32.64	69	2.5	12.11	3.9	262	84.3
Cr (mg/kg)	71.7	2.76	33.72	47	1.02	4.04	9.5	243.4	107.2
Mn (mg/kg)	533.6	23.35	284.97	53	0.46	-1	126.3	1103	972.3
Fe (%)	2.46	0.09	1.14	46	1.09	2.06	0.49	7.2	3.9
Co (mg/kg)	12.65	0.66	8.01	63	1.79	4.93	2	51.2	22.5
Zn (mg/kg)	92.46	6.29	76.78	83	1.86	4.42	5.77	443	178.6
Cu (mg/kg)	42.06	2.06	25.1	60	0.95	0.55	1.57	127.02	84.1
As (mg/kg)	5.73	0.41	5.02	88	3.13	14.84	0.12	40	13.13
Se (mg/kg)	1.87	0.08	0.97	52	0.42	-0.47	0.3	4.5	3.1
Hg (mg/kg)	0.54	0.09	1.15	213	9.03	95.36	0.02	13	1.1

*S.E. Mean standard error mean, Std. Deviation standard deviation, RSDs relative standard deviations

taken into account. Samples' EC values ranged from 0.01 to 1.40 dS/m, with an average SD of 0.20–0.23 dS/m (Table 1). The granulometric analysis reveals that the samples belong to the silty clay class, in accordance with the US soil classification. Sand, silt, and clay content averaged 10.25, 71.14, and 21.95%, respectively (Table 1). More so than coarser fractions (sand), the finer-grained soils (clay slit) seem to affect

TML solubility and mobility, weathering-oxidation, and leaching. The ranking of sand, silt, and clay in terms of adsorption illustrates the great affinity of TMLs and other metals for the finest fraction.

As listed in Table 1 and Fig. 2, there was a distinct variability in the concentration of TMLs among the studied samples along with their descriptive statistics and the relevant literature data. The average contents

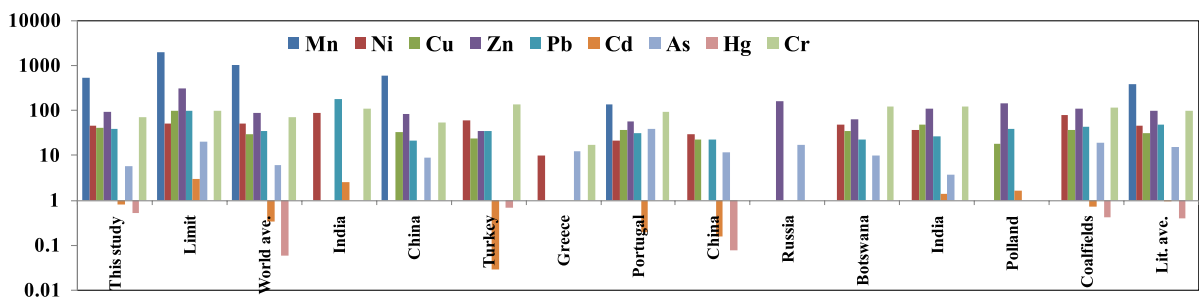


Fig. 2 Comparison of selected potentially toxic elements concentrations ($\mu\text{g/g}$) found in coal industry-associated soils in this study and other contaminated soils worldwide. Safe-ecotoxicological limit, Kabata-Pendias (2000); World median, Bowen (1979); Ledo-Tinsukia (India), Reza et al. (2015); Huainan (China), Tang et al. (2018); Oltu (Turkey), Tozsin

(2014); Ptolemais (Greece), Modis et al. (2013); Douro (Portugal), Ribeiro et al. (2010); Yongcheng (China), Shi et al. (2022); Tula-Moscow (Russia), Komnitsas and Modis (2006); Palapye (Botswana), Zhai et al. (2009); Jharia-Dhanbad (India), Masto et al. (2017); Smolnica (Poland), Pietrzykowski et al. (2014); Coalfields ($n = 13$, global), Xiao et al. (2020)

of these potentially toxic metals were 39.7 ± 30.4 , 0.80 ± 0.92 , 47.1 ± 32.6 , 71.7 ± 33.7 , 533.6 ± 285.0 , 2.5 ± 1.1 , 12.7 ± 8.0 , 92.5 ± 76.8 , 42.1 ± 25.1 , 5.7 ± 5.0 , 1.9 ± 0.97 , and 0.54 ± 1.2 mg/kg for Pb, Cd, Ni, Cr, Mn, Fe, Co, Zn, Cu, As, Se and Hg, respectively. The highest measured TMLs (115% and 213%, respectively) for Cd and Hg indicate that those metals exhibited the most variability in the samples, which anthropogenic influences may have contributed to (Zhang et al., 2018). TMLs with values over one that are severely skewed are Mn, Cu, and Se exceptions. The majority of the samples were clustered at low values, which resulted in extremely acute kurtosis (Islam et al., 2019).

The concentrations of potentially toxic elements in coal industry-associated soils in the current study are compared with the trace elements in soil in Bangladesh and other studies worldwide (Fig. 2 and Table S3). Concentrations of Cd (8.9 times), Hg (10.8 times), and Pb (2.3 times) in soils of the current study are considerably higher than the typical UCC values proposed by Rudnick and Gao (2014), and the levels of other elements in soils of the present study are nearly close to the UCC values. The two most toxic elements, Cd (2.3 times) and Hg (9.0 times), are considerably higher than the typical world median value of trace elements in soil (Bowen, 1979), indicating coal industry-associated soils are polluted by trace elements, especially Cd and Hg (Fig. 2 and Table S3). Our obtained values are higher for Mn (3.8 times), Ni (2.2 times), Cd (4.0 times) than (Douro, Portugal); for Zn (2.6 times) and Cd (26.7 times) (Oltu, Turkey); for Ni (4.7 times) and Cr (4.1 times) (Ptolemais, Greece); for Cd (5.0 times) and Hg (6.8 times) (Yongcheng, China); Cu (2.3 times), respectively in several-order of magnitudes (Fig. 2 and Table S3). Ribeiro et al. (2010) observed elevated levels of As, Cd, Cu, Pb, Ni, and Zn in the Serrinha coal waste pile site soils, Douro Coalfield, Portugal. They also confirmed that the acid mine drainage and associated leaching of heavy metals are responsible for elevated levels of toxic elements in the soil. Tozsin (2014) observed Cr, Ni, and Hg concentrations in soil samples collected from Turkey's Oltu coal mine district were above the limit values. Tozsin (2014) also concluded that only Cr contamination of the Oltu coal mine soil was related to the leachate derived from coal samples. Modis et al. (2013) analyzed trace elements in soils from the lignite opencast mining and industrial area

of Greece, and they concluded that the generation of acidic leachates from the mining waste is responsible for the solubilization and mobilization of inorganic contaminants. Tang et al., 2018 determined contamination levels of several important toxic heavy metals in coal gangue-reclaimed soils in China. They observed high levels of Cu, Pb and Zn in soil from coal gangue. The elements concentrations of the current study were higher than the above mention studies in the world, indicating the severe contamination of trace elements in coal mine soil in Bangladesh. Compared to non-contaminated soils, the observed amounts of Mn are two times higher (Kabata-Pendias, 2000). Most metal(loid) concentrations did not surpass the European Communities soil quality regulations (ECD, 1986).

The concentrations of the studied elements in soil samples were higher than the background values of trace elements in the soil of Bangladesh (Kashem & Singh, 1999). The study by Kashem and Singh (1999) concluded that soils from abandoned sites with industrial waste are highly contaminated by toxic elements, where As, Cd, and Pb are the most serious. The current study's findings also indicate the contamination of coal industry-associated soils by toxic trace elements in soils of the study area. A recent study shows that the soils around the Barapukuria mining zone are extensively polluted with Fe, Mn, Zn, Pb, As, Cr, Ni, Cd, and Cu (Bhuiyan et al., 2010; Hossain et al. 2015; Halim et al., 2015; Siddique et al., 2020; Hossen et al., 2021). Pollution was worse around mines and power plants. Mining, combustion, leaching, weathering (alteration-dissolution), coal water mixing/interactions emission, and deposition may release a large amount of toxic mobile heavy metal(loid)s from the coal matrix (Cao et al., 2021). Compared to soil-enriched elements, Cd, Fe, Co, As, Se and Hg averaged around their local background values. The samples had TML concentrations similar to Bangladeshi and global soil levels (Table S3).

Co-occurrence relationships of TMLs in soils

Co-occurrence correlations and PCoA were performed through network inference utilizing non-parametric Spearman's rank coefficient (Barberán et al., 2012) to disseminate the total variability in physicochemical variables TMLs in samples (Fig. 3a, b). Thus, co-network correlations were used to assess

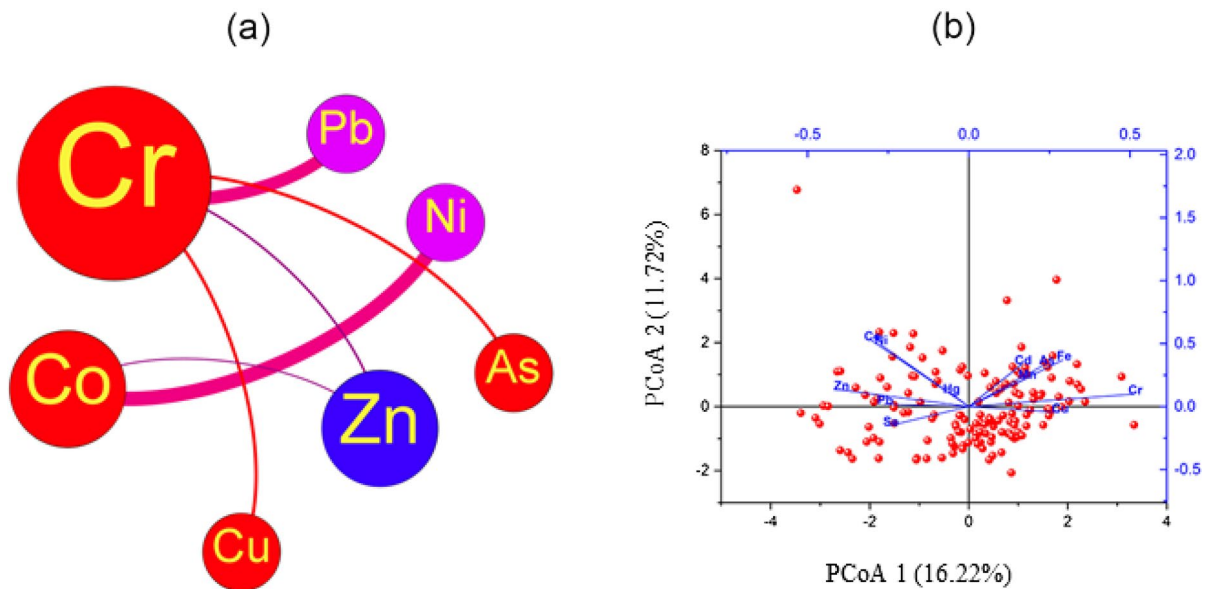


Fig. 3 **a** Co-occurrence network of toxic metal(loid)s ($p < 0.01$; $r > 0.2$), the node size based on concentration and line strength based on correlation value; **b** Principal coordinate

analysis (PCoA). Each dot represents each sampling point and vectors overlay of variables shown if $r > 0.3$ to PCoA axis;

measured/determined metal(loid)s in soils. Network analysis uses topological properties like the cluster coefficient and modularity index to discuss and quickly interpret complex relationships, associations, occurrences, and common occurrences between complicated objects, variables, and attributes. When the Spearman's coefficient for TML co-occurrence correlations was more than 0.2, it showed significant linkages. The average clustering coefficients were 0.62–0.95, indicating strong TML relationships (Islam et al., 2020). Raw data indicated relationships in several network characteristics. They also enabled TML co-occurrence and clustering in Bangladeshi research soils.

Figure 3a displays the samples' strong co-occurrence correlations of found TMLs ($r > 0.2$) Spearman's coefficients. Islam et al.'s (2020) study suggests using Cu, Zn, and As to identify lithogenic, atmospheric (coal flying dust, fly-ash), and external agrochemical factors. Since no positive correlation between reference metals and TMLs in soils has been found, Liu et al. (2015) conclude that this region was mostly influenced by external contamination. Cr and Pb are strongly linked by lithogenic/mineralogic and exogenous factors, according to co-occurrence network research. Nickel and Co were

very positive (Fig. 3a). TML co-occurrence correlations were abnormally high, especially for industrial contaminants as Cr, Pb, Co, and Ni. Tamim et al. (2016) attribute the elevated element concentrations in the study location to human activities with geo-environmental and pedogenic processes. The negative linkages between other components (Fig. 4a) show a strong relationship between most TMLs. Liu et al. (2015) found a similar result.

The PCoA explained 27.9% of soil sample TML variation (Fig. 3b). Principal component 1 (PCoA1) consisted of Cr, Cu, Mn, Cd, As, and Fe at similar quantities. According to the data's rotational component matrix, Pb, Zn, Co, Hg, Ni, and Se were strongly associated to PCoA2. Singh and Kumar (2017) reported that the PCoA1 is primarily Cr, Cu, Mn, Cd, As, and Fe, industrial and agricultural components. Human activities like mining, smelting, burning, and others may alter this 16.2% of the variance input rate. Ca-, Na-, and Pb-arsenate, which are commonly employed in agrochemicals to boost pesticide effectiveness for local crop production, are given special attention. According to Chen et al. (2016), frequent and excessive pesticide usage is the major cause of As and Cu soil pollution.

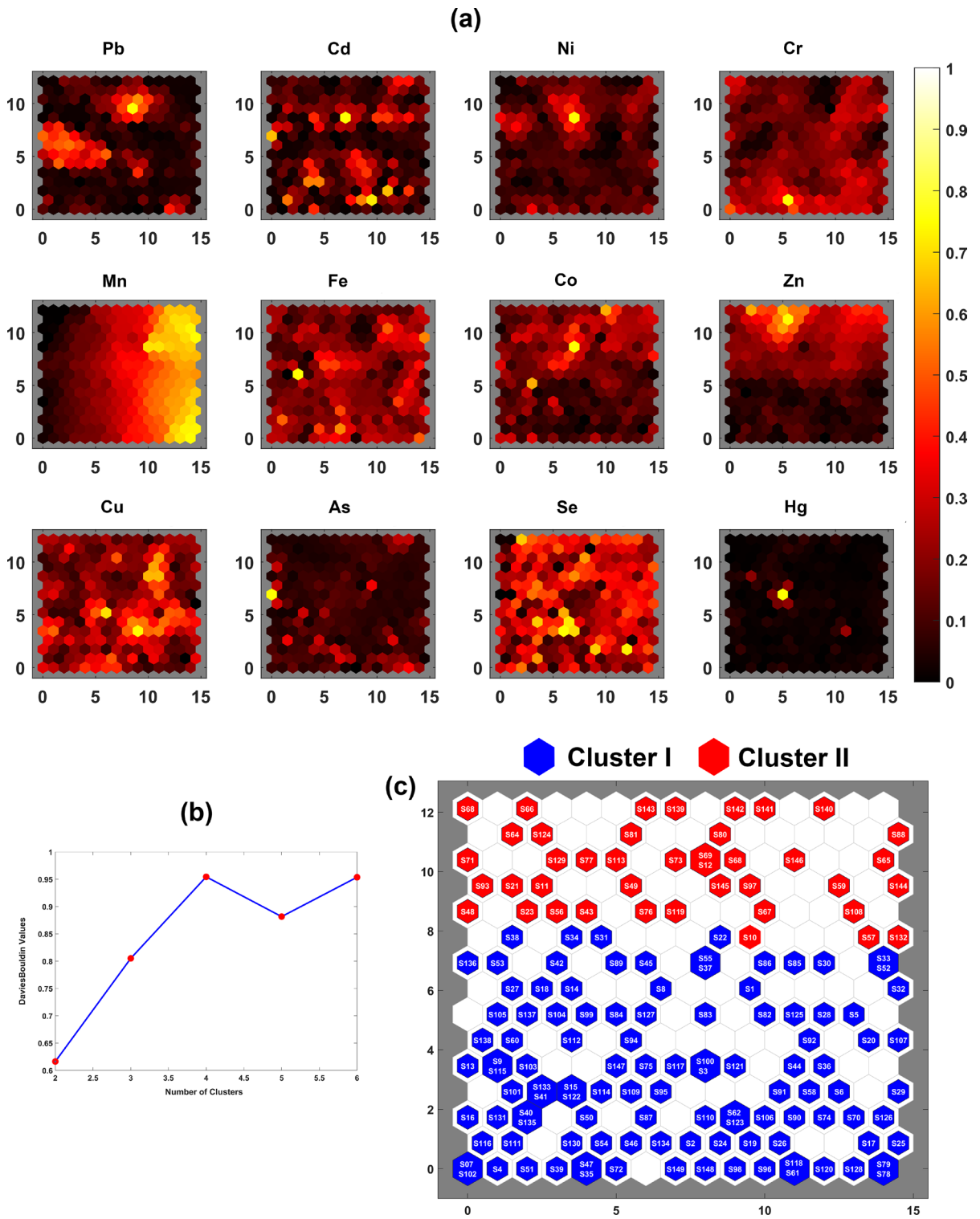


Fig. 4 **a** SOM analysis of toxic metal(loid)s in the soils; **b** Cluster estimated by DBI index **c** Cluster of soil sampling sites from SOM analysis

The soil matrix absorbed Cd, As, and Fe (Lv et al., 2015). Pb dominated the PCoA2, whereas Zn, Co, Hg, Ni, and Se accounted for 11.72% of the variance addition rate and may represent extrinsic factors. For this investigation, soil chemical species averages were increased. Hg, Zn, and Pb had a high coefficient of variation, showing that extrinsic inputs are the major sources of those TMLs in soils, even if the average levels of all examined TMLs did not vary. Our research also uses PMF and APCS-MLR models to investigate soil TML interactions and determine source apportionments.

Spatial pattern recognition of TMLs in soils using SOM model

SOM component planes are shown in Fig. 4. SOM planes were created using colored hexagons mimicking benzene to identify, demonstrate, show, and convey the significance of particular variables for each SOM class. The hexagon's inside must be smaller the more similar the samples are. The plane's colors signify positive and negative TML connections. Figure 4a shows twelve-TML component planes. Five-color patterns resembled metal(loid) groups. In the first pattern, neurons represented Se, Cu, and Pb in ascending sequence from top to bottom left. Metalloids found a source. Zn and Co analogized the second pattern of neuron growth from bottom left to top right. Combining fundamental element regulations governs both metals. Geogenic attributions ordered neurons from left to right in the third pattern, and Mn and Fe matched this spatial arrangement. The fourth pattern showed that the largest and lowest neuron percentages are in the top right corner and bottom left corner, respectively. Fitting Cd, Hg, and Similar to these geographical trends, pesticides used in agro-farming, air deposition, electroplating, painting, and other operations regulate coal basin soil quality. SOM neurons with the greatest and lowest percentages are in the fifth pattern's bottom left and top right corners. Cr and Ni were matched to this geographical pattern, which may be due to a lot of human-induced, untreated waste and sewage in the studied areas. In the following part, SOM analysis findings show that PMF and APCS-MLR TMLs are comparable.

After the training phase, the DBI values for SOM analysis using 1 to 6 clusters were obtained and are shown in Fig. 4b to help choose the best number of

clusters. After choosing the ideal cluster number, the SOM map offered two distinct clusters of sample locations. The sample collection locations were grouped using the SOM analysis (Fig. 4c). Only 109 samples were added to cluster I, while 40 samples were added to cluster II. Significantly colored zones served as a marker for the cluster borders. A semi-qualitative link between the measured TMLs may be seen in color variations within the same cluster. The level of contamination separates clusters from sample sites with high to low levels of pollution (Kumar et al., 2021). Consequently, the SOM map is a useful tool for assessing the TMLs properties of soil samples.

Comparisons of APCS-MLR with PMF models for apportioning soil TMLs

APCS-MLR and PMF models were used to quantify the sources of soil TMLs, and the results are shown in Table 2 and Fig. 5. PCA was used to confirm the model's complexity before the PMF model was used. The usual TMLs from a certain source were regarded to be soil TMLs with a strong positive loading in the factors. The APCS-MLR model yielded five components, as shown in Table 2 and Fig. 5a. Cu, Se, Fe, Hg, and Pb were predominantly blamed for the first factor (F1); their average contributions were 67.77%, 29.10%, 23.74%, 17.19%, and 16.20%, respectively. The second factor (F2) had greater quantities of Zn, Co, and As (51.77%, 51.62%, and 49.42%). The 70.14% and 37.31% Mn and Fe concentrations were explained by the third factor (F3). The fourth factor (F4) made preliminary contributions of 47.33%, 33.70%, 33.15%, and 35.29%, respectively, to Cd, Pb, Ni, and Hg. The fifth factor (F5) explained the concentrations of Cr and Pb by 65.44% and 30.92%, respectively, but only partially explained the other TMLs.

The PMF approach also identified five factors based on the least and most stable Q values (Table 2; Fig. 5b). TMLs were confined to a strong category since the signal-to-noise ratios (S/N) were greater than 4, proving the model's predictions were logical. In addition, the fitting determination of coefficients (R^2) of the PMF technique for each metal(loid) were greater than 0.61, showing an acceptable fit for the PMF method (Fig. S1). It is important to note that the residual value of the majority of the samples

Table 2 Contribution of each factor originated from APCS-MLR and PMF models

Metal(loid)s	APCS-MLR						PMF					
	R ²	IE*	AP*	GS*	CPP*	AD*	R ²	IE*	AP*	GS*	CPP*	AD*
Pb	0.61	16.2	4.83	14.35	33.7	30.92	0.62	10.07	19.08	19.07	16.33	35.45
Cd	0.95	3.2	44.11	2.14	47.33	3.22	0.99	86.1	5.83	7.63	0	0.43
Ni	0.62	1.63	46.14	10.83	33.15	8.26	0.64	8.46	27.56	13.88	25.49	24.6
Cr	0.68	13.59	6.34	6.32	8.3	65.44	0.76	16.66	0	4.7	28.96	49.68
Mn	0.72	11.55	2.82	70.14	9.88	5.6	0.78	10.25	0	13.66	74.46	1.62
Fe	0.64	23.74	9.69	37.31	17.5	11.75	0.66	13.67	9.11	0	28.45	48.77
Co	0.81	2.99	51.62	23.11	16.24	6.04	0.77	0.15	74.09	1.7	2.15	21.91
Zn	0.66	6.5	51.77	4.69	15.53	21.5	0.72	1.37	91.56	3.91	3.16	0
Cu	0.65	67.77	13.3	2.76	8.92	7.25	0.67	4.08	8.15	8.54	0	79.23
As	0.82	12.58	49.42	25.32	2.17	10.52	0.81	8.5	14.39	21.48	26.44	29.19
Se	0.98	29.1	25.49	18.14	4.01	23.26	0.99	1.31	4.63	73.53	10.42	10.11
Hg	0.7	17.46	33.25	12.58	35.29	1.41	0.71	0	11.44	12.5	75.55	0.5

*IE industrial emission, AP agrochemical practice, GS geogenic source, CPP coal power plant, AD atmospheric deposition

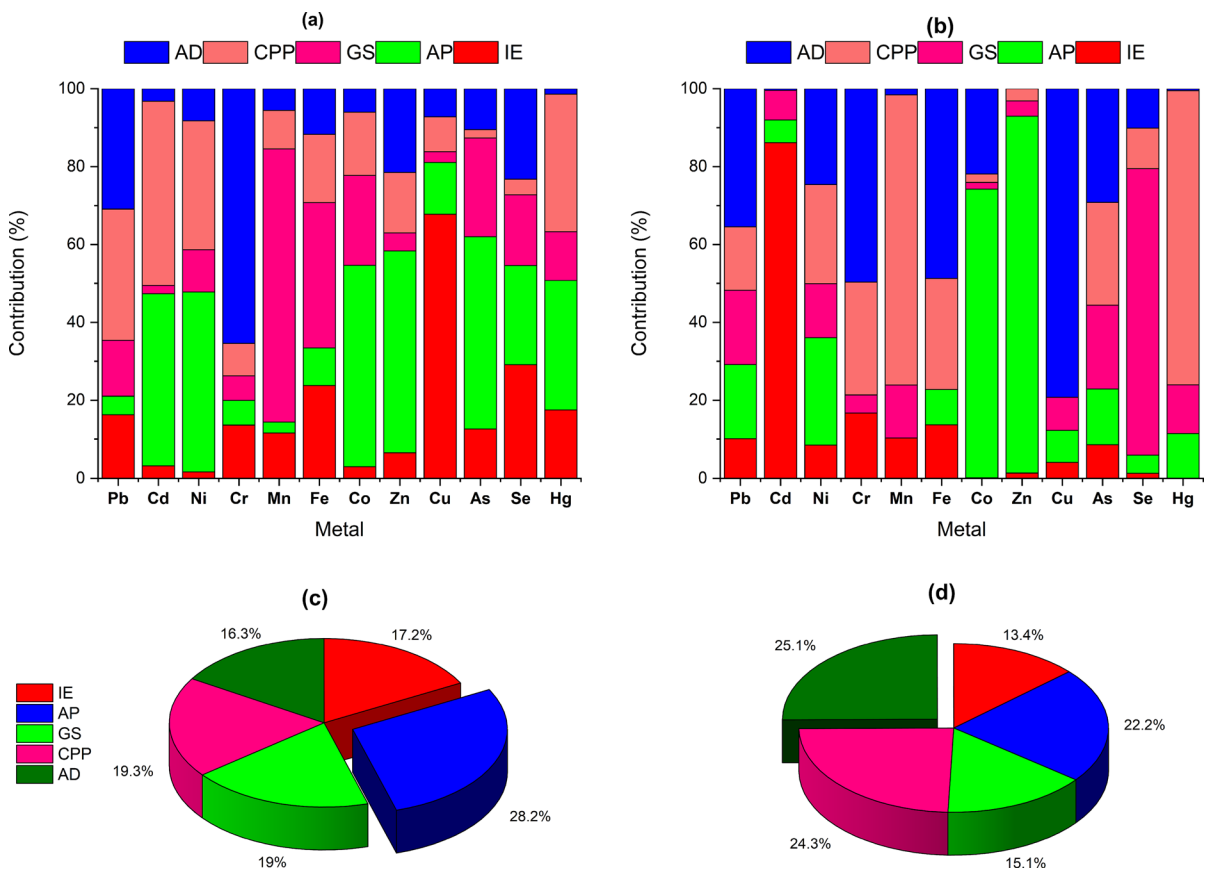


Fig. 5 Factor profiles of toxic metal(loid)s in the soils originated from the APCS-MLR model **a** and PMF model **b**, and average contributions for the five pollution sources detected **c**

and **(d)**. (Industrial emission: IE, Agrochemical practice: AP, Geogenic source: GS, Coal power plant: CPP, Atmospheric deposition: AD)

that were analyzed ranged from 3 to 3 (Duan et al., 2020). Cu, Cr, and Fe had the highest attribution in the first factor (F1), as shown in Table 2, with average values ranging from 48.77% to 79.23%. The contents of Co and Zn made up 74.09% and 91.56% of the second factor (F2). With an average increase of 86.10%, Cd had the highest input, which was mostly due to the third factor (F3). The largest average contributions in the fourth factor (F4) were from Mn and Hg, with 74.46% and 75.55%, respectively. The fifth factor (F5) clarified less of the previous TMLs while explaining 73.53% of the Se content.

Five variables were found for the APCS-MLR and PMF models, which both had success identifying the source categories. The factors 2 and 4 retrieved by the PMF model, which have been identified as the agrochemical practice and coal-related sources, are comparable to the factors 2 and 4, as can be shown in Fig. 5. Overall, the APCS-MLR model and the PMF model have produced results that are quite comparable. The APCS-MLR model also recognizes the potential contributors of atmospheric deposition (AD), agrochemical practice (AP), geogenic source (GS), coal power plant (CPP), and industrial emission (IE), with average contributions of 16.3%, 28.2%, 19%, 19.3%, and 17.2%, respectively (Fig. 5c). The findings suggest that the two chosen models for apportioning the sources of TMLs in the studied samples are consistent. Agrochemical practices have been identified as a key factor influencing the distribution of metals (loids) in soils, with contributions from PMF and APCS-MLR totaling 22.2% and 28.2%, respectively. According to many studies (Bhuiyan et al., 2010; Halim et al., 2015; Hossen et al., 2021; Rahman et al., 2017), Bangladesh has been and continues to be incredibly dependent on the agricultural sector.

Table 2 provides information on the R^2 of detected soil TMLs. All TMLs had high R^2 values ranging from 0.61 to 0.99 in the receptor models, indicating that both the PMF and APCS-MLR models showed satisfactory results. However, the R^2 values of most TMLs in the PMF model were higher than those in the APCS-MLR model, indicating that the PMF model is more ideal for source distribution of TMLs in this research. Also, negative contributions that affect the accuracy of the results were observed in the APCS-MLR model. According to several prior research (Guo et al., 2021; Zhang et al., 2021) the

performance of R^2 of the PMF model in recognizing a variety of sources was superior than APCS-MLR. The APCS-MLR and PMF investigations both used a number of mathematical restrictions to identify the likely sources of pollution. Compared to the APCS-MLR model, the PMF model identified more sources of pollution with low percentage error. According to Deng et al. (2018), the PMF model's source's ideal numbers had the lowest Q value. Therefore, the ideal number of sources was five possible components. The apportionment and contribution rate of the prospective source varied according to different theoretical tenets. The results were thus inconsistent. In addition, the APCS-MLR failed to provide accurate previous estimates of the error-scale for the uncertain measurements. The lower data constraint was in this model.

In contrast, the PMF model may provide reliable results since it gives weight to uncertainty estimations for the dataset (USEPA, 2014; Luo et al., 2019). Notably, uncertainty analysis was given to evaluate the strength and coherence of the results of the PMF model. Thus, compared to the APCS-MLR model, the PMF model may examine source analysis results of soil TMLs in complex agricultural districts affected by contaminated coal more precisely.

Source apportionment of soil TMLs

The factor I (F1) was responsible for 25.1% of the contribution rate; Factor 1 was predominantly contributed by Cu (79.23%), Cr (49.68%), Fe (48.77%), and Pb (35.45%) (Fig. S1). Correlation coefficient analysis indicated that there was a significant association between Cu vs Cr ($r=0.523$, $p<0.05$), Cr vs Pb ($r=0.491$, $p<0.05$). Besides, Yang et al. (2019) and Duan et al. (2020) also noted that Cu, Cr, and Pb might be related to atmospheric deposition. Earlier literature has revealed that Pb and Cr are encompassed in brakes, roadside tires, and mine waste (Huang et al., 2015). These toxic metals can be liberated into the ambient atmosphere and then landed into the topsoils. In addition, the power-plant emission is a potential contributor of Pb in atmospheric deposition. Lin et al. (2018) also stated that Pb content might get into soils via atmospheric deposition. Thus, Cu, Cr, Fe, and Pb in F1 are considered as atmospheric deposition (coal flying dust, fly-ash).

Factor 2 (F2) accounted for 22.2% of the input rate, was mostly positively loaded on Co (74.09%),

Zn (91.56%), and As (27.56%) (Fig. S1). A significant positive association was identified between Zn and As ($r=0.490$, $p<0.01$) and Co and Zn ($r=0.61$, $p<0.05$). Previous research showed that these elements' concentration is relevant to applying agrochemical products (Cai et al., 2019; Islam et al., 2020). A recent study by Lin et al. (2018) stated that Zn and As are intimately linked to applying agrochemical products and improving pesticide efficiency in crop fields for high yield. Similarly, Xiao et al. (2020), Jiang et al. (2020), and Yang et al. (2019) have noticed an increased concentration of As, Co, and Zn in soils, which may be attributed to agrochemical inputs. Some studies pointed out that agrochemical fertilizer also triggered the high accumulation of Zn and As in the soil (Duan et al., 2020; Xiao et al., 2020). Therefore, Factor 2 was considered as originated from an agrochemical practice source.

Factor 3 (F3) made up 13.4% of the total contribution and was dominated mainly by Cd (86.10%) and Cr (16.66%). Earlier researches noted that Cd and Cr would source from industrial activities (e.g., battery production, metal manufacturing, mineral exploration and utilization, effluents, electroplating, etc.) (Islam et al., 2015). The mean concentrations of Cd exceeded the background value of soil in Bangladesh (Table S3). The hot spot of Cd was linked with industrial contributions. In addition, Cd was well-known as fingerprints connected to industrial and municipal contamination where unplanned and undesired activities might be a considerable contribution of Cd in soils (Ogundele et al., 2020). Previous investigations have stated that Cd in the examined samples is likely derived from emissions, including smelting and metal-and alloy processing (Jiang et al., 2020). According to the field survey, metal manufacturing plants, combustion-based and other residuals/byproducts disposal sites, and mineral, coal processing and preparation plants (coal flying dust, fly-ash) at different scales are available in the area. Therefore, F3 was identified as industrial emission.

Factor 4 (F4) accounted for 75.55% and 74.46% of the total content of Hg and Mn (Table 2), which is liable for 24.3% of the total input. The results of descriptive statistics (Tables 1 and S3) showed that the level of Mn and Hg in studied samples was higher than in soils around coal industries worldwide. In addition, the elevated level of Mn and Hg areas were situated/occurred in the Barapukuria

coal mining region, Bangladesh, which is further supported by a significant relationship between Mn and Hg ($r=0.531$, $p<0.01$). Previous researches have reported that coal-power-plant contains enhanced contents of Hg (Cai et al., 2019), are volatile and may have the possibility to migrate to the surrounding soils (Men et al., 2020). hence/thus, the use of coal-power-plant coal-power plants can/might potentially affect the surrounding soil health/environment/systems for the input of Mn and Hg. Mercury is generally associated with industrial activities (e.g., coal-based brickfield and power production, municipal wastes, and medical-waste (Xiao et al., 2021). Hence, F4 was regarded as the coal-power-plant input (typically carbon nano-tube, nano-minerals, ultrafine particulates in coal flying dust and coal combustion fly-ash.

Factor 5 (F5) was described by Se and As in the PMF method, with increased loadings of 73.53% and 21.48%, respectively, responsible for 15.1% of the total contribution. There was a considerable relationship between Se and As ($r=0.598$, $p<0.01$). Selenium is mainly discharged to water bodies via mine effluents, coal-based power plant discharges, surface-runoff, weathering-erosion, and wastewater. Se and As are discharged into soil media primarily by leaching, mobilization, weathering, and dissolution of the parent bed-rock/minerals matrix, though Se and As levels also increased in soils receiving from the ambient atmosphere. Some studies have demonstrated that Se and As in soils might be affected due to geologic- and pedogenic processes such as erosion, oxidation, and weathering (Bhuiyan et al., 2010; Halim et al., 2015). Generally, TMLs associated with the composition of the soil parental rock materials often had a low degree of pollution (Wu et al., 2020). Furthermore, Se and As contents in F5 are considered the representative TMLs associated with soil parent granite rock. Thus, F5 was recognized as a geogenic source (rock-hosted minerals).

Overall, atmospheric deposition, agrochemical practice, geogenic source, coal-power-plant-, and industrial emissions are distributed as the probable contamination contributors of TMLs in the analyzed samples in the selected region (Fig. 5d). Comparatively, atmospheric deposition (25.1%) of soil particles (i.e., coal waste, mine tailings, and coal ash residuals) has contributed most, followed by agrochemical practice (22.2%) for the soil TMLs. The contributions

of atmospheric deposition and coal-power-plants were almost similar, with 24.3% and 25.1%.

Source-oriented health risk evaluation of soil TMLs

The source-specific health hazards linked to distinct exposure pathways to TMLs in the soils were analyzed using a probabilistic technique based on the apportionment pollution outputs of PMF and APCS-MLR models. As outlined in Eqs. (2)–(4), the concentration percentage of soil TMLs contaminations obtained from the two designated models for the five sources (agrochemical practice, atmospheric deposition, industrial emission, coal-power-plant, and geogenic source) are integrated into the non-carcinogenic (NCR) and carcinogenic risks (CR) models (8). The uncertainty of health risk assessment was measured using Monte Carlo simulation. In addition, the risk values for the estimated concentration, i.e., total factors, were calculated in this study. Using probability density functions, Fig. 6 depicted the computation/calculation of the danger presented by TMLs in soils based on the total variables and the quantified five contamination sources. The mean NCR and CR values offered by the soil TMLs are slower than their

respective guideline values for the total components (estimated concentrations), as illustrated in Fig. 6a, b, suggesting a negligible risk. The findings are in good agreement with those presented by earlier cited works that the health risk of soil TMLs was insignificant (Guo et al., 2021; Jiang et al., 2021).

For NCR, the total THI outcomes/results of TMLs were $1.66\text{E}-01$ for adults and $1.28\text{E}-01$ for children, suggesting a negligible NCR for both age groups in the study region (Table 3). For Pb, Cr, Fe, and Cu, atmospheric deposition was primarily pollutant sources of HI, which was dominated with $5.30\text{E}-03$, $2.28\text{E}-02$, $1.46\text{E}-02$ and $6.31\text{E}-04$ for adults, and $4.20\text{E}-03$, $1.72\text{E}-02$, $1.17\text{E}-06$, and $5.03\text{E}-04$ for children, respectively. For As, Zn, Co, and Ni, agrochemical practice was the major contributor for HI, and it was responsible for $8.44\text{E}-03$, $3.09\text{E}-04$, $3.66\text{E}-02$ and $1.24\text{E}-03$ for adults and $6.72\text{E}-03$, $2.45\text{E}-04$, $2.92\text{E}-02$ and $9.67\text{E}-04$ for children, respectively. For Cd, the elevated HI has come from industrial emission sources ($6.87\text{E}-04$ for adults and $4.96\text{E}-04$ for children). For Mn and Hg, the HI of coal-power-plant source was the leading, accounting for $4.42\text{E}-03$ and $4.09\text{E}-03$ for adults and $2.73\text{E}-03$ and $3.27\text{E}-03$ for children, respectively. For Se,

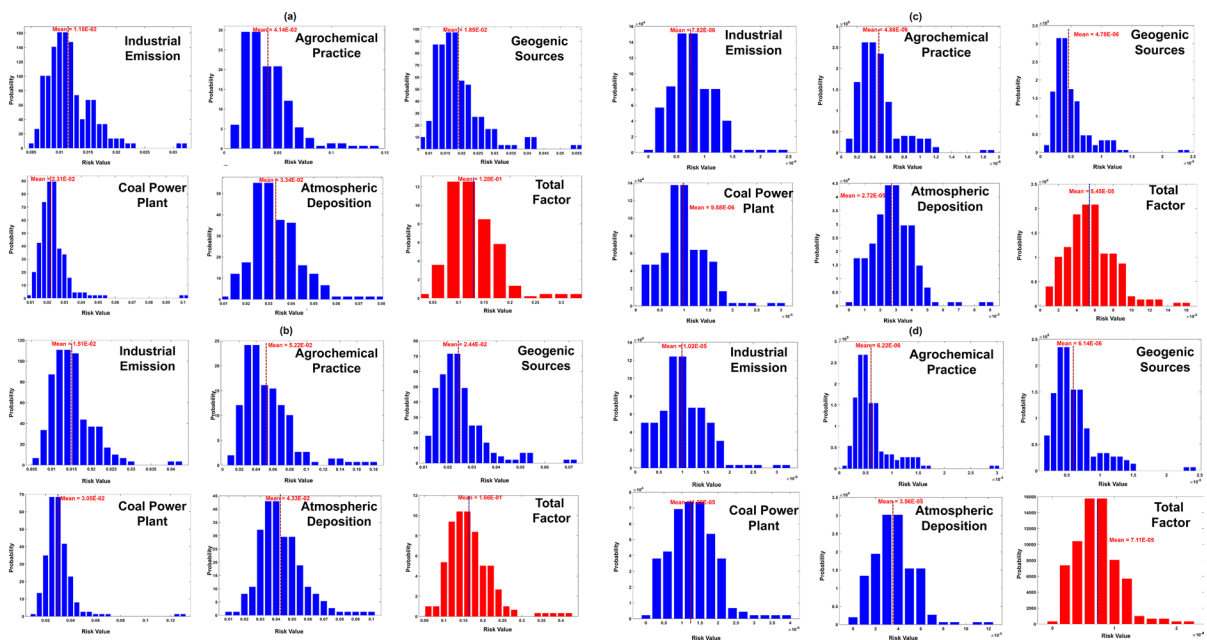


Fig. 6 Probability density function (PDF) for non-carcinogenic risk (N_{CR}) of **a** children and **b** adults and carcinogenic risk (C_R) of **c** children and **d** adults

Table 3 Human health risk estimation posed by the soil TMLs from total factors and the five potential pollution sources detected

Parameter	Children					Adult					Total	
	IE*	AP*	GS*	CPP*	AD*	Total	IE*	AP*	GS*	CPP*		AD*
Non-Carcinogenic												
Pb	1.66E-03	1.51E-03	2.12E-03	3.17E-03	4.20E-03	1.27E-02	2.10E-03	1.91E-03	2.67E-03	3.99E-03	5.30E-03	1.60E-02
Cd	4.96E-04	2.77E-04	5.43E-05	2.63E-04	2.03E-05	1.11E-03	6.87E-04	3.84E-04	7.52E-05	3.64E-04	2.81E-05	1.54E-03
Ni	1.32E-04	9.67E-04	3.24E-04	7.70E-04	4.31E-04	2.63E-03	1.69E-04	1.24E-03	4.15E-04	9.84E-04	5.51E-04	3.36E-03
Cr	4.51E-03	9.45E-04	1.64E-03	5.55E-03	1.72E-02	2.98E-02	6.00E-03	1.26E-03	2.19E-03	7.39E-03	2.28E-02	3.96E-02
Mn	7.06E-04	9.14E-05	2.71E-03	2.73E-03	2.34E-04	6.47E-03	1.14E-03	1.48E-04	4.39E-03	4.42E-03	3.78E-04	1.05E-02
Fe	7.23E-07	3.63E-07	7.21E-07	8.88E-07	1.17E-06	3.86E-06	9.04E-07	4.54E-07	9.02E-07	1.11E-06	1.46E-06	4.83E-06
Co	7.27E-04	2.92E-02	5.75E-03	4.26E-03	6.48E-03	4.64E-02	9.14E-04	3.66E-02	7.23E-03	5.36E-03	8.15E-03	5.83E-02
Zn	1.35E-05	2.45E-04	1.47E-05	3.20E-05	3.68E-05	3.42E-04	1.69E-05	3.09E-04	1.85E-05	4.03E-05	4.63E-05	4.31E-04
Cu	4.18E-04	1.25E-04	6.57E-05	5.19E-05	5.03E-04	1.16E-03	5.24E-04	1.57E-04	8.25E-05	6.51E-05	6.31E-04	1.46E-03
As	2.22E-03	6.72E-03	4.93E-03	3.01E-03	4.18E-03	2.11E-02	2.79E-03	8.44E-03	6.19E-03	3.79E-03	5.25E-03	2.65E-02
Se	6.28E-05	6.22E-05	1.89E-04	2.98E-05	6.89E-05	4.13E-04	7.87E-05	7.79E-05	2.37E-04	3.73E-05	8.63E-05	5.17E-04
Hg	5.16E-04	1.32E-03	7.41E-04	3.27E-03	5.67E-05	5.91E-03	6.45E-04	1.65E-03	9.26E-04	4.09E-03	7.09E-05	7.39E-03
THI	1.15E-02	4.14E-02	1.85E-02	2.32E-02	3.34E-02	1.28E-01	1.51E-02	5.22E-02	2.44E-02	3.05E-02	4.33E-02	1.66E-01
Carcinogenic												
Pb	5.13E-09	4.66E-09	6.52E-09	9.76E-09	1.29E-08	3.90E-08	6.98E-09	6.35E-09	8.88E-09	1.33E-08	1.76E-08	5.32E-08
Cd	1.64E-07	9.17E-08	1.79E-08	8.69E-08	6.71E-09	3.67E-07	2.15E-07	1.20E-07	2.35E-08	1.14E-07	8.78E-09	4.81E-07
Ni	5.23E-08	3.82E-07	1.28E-07	3.04E-07	1.70E-07	1.04E-06	6.57E-08	4.80E-07	1.61E-07	3.82E-07	2.14E-07	1.30E-06
Cr	6.60E-06	1.38E-06	2.40E-06	8.13E-06	2.51E-05	4.36E-05	8.68E-06	1.82E-06	3.16E-06	1.07E-05	3.30E-05	5.74E-05
As	9.99E-07	3.02E-06	2.22E-06	1.36E-06	1.88E-06	9.48E-06	1.25E-06	3.79E-06	2.78E-06	1.70E-06	2.36E-06	1.19E-05
THI	7.82E-06	4.89E-06	4.78E-06	9.88E-06	2.72E-05	5.45E-05	1.02E-05	6.22E-06	6.14E-06	1.29E-05	3.56E-05	7.11E-05

*IE industrial emission; AP agrochemical practice, GS Geogenic source, CPP coal power plant, AD Atmospheric deposition

the geogenic source ($2.37\text{E}-04$ for adults and $1.89\text{E}-04$ for children) was the key source with the elevated HI (Table S5). The highest source of contribution to the THI was agrochemical practice (adults: $5.22\text{E}-02$ and children: $4.14\text{E}-02$), followed by atmospheric deposition (adults: $4.33\text{E}-02$ and children: $3.34\text{E}-02$), coal-power-plant (adults: $3.05\text{E}-02$ and children: $2.32\text{E}-02$), geogenic source (adults: $2.44\text{E}-02$ and children: $1.85\text{E}-02$) and industrial emission (adults: $1.51\text{E}-02$ and children: $1.15\text{E}-02$), respectively. The THI of the agrochemical practice source was the largest, which was linked with the agricultural fertilizer and pesticides. These can bring a significant amount of toxic metals like As, Zn, and Cu that pose a potential threat to the local inhabitants. The total THI outcomes for adults were notably lower than those from other studies in China (2.81) (Guo et al., 2021) and India (13.8) (Singh et al., 2018). For CR, the total TCR value for children ($5.45\text{E}-05$) and adults ($5.37\text{E}-05$) was within the standard limit of $1\text{E}-06$ and $1\text{E}-04$, indicating the permissible risk level.

The contribution of five carcinogenic TMLs (As, Cr, Cd, Ni, and Pb) to CR was computed as outlined in Table 3. For Cr, the atmospheric deposition was the key source apportionment in the elevated CR, $2.51\text{E}-05$ for children and $3.30\text{E}-05$ for adults. Similar to Cr, for Pb, atmospheric deposition was an elevated CR ($1.29\text{E}-08$ for children and $1.76\text{E}-08$ for adults). For As and Ni, the agrochemical practice source was also had an elevated CR ($3.02\text{E}-06$ for children and $3.79\text{E}-06$ for adults; $3.82\text{E}-07$ for children, and $4.80\text{E}-07$ for adults). Cd was from industrial emission, of which CR was accounted for $1.64\text{E}-07$ for children and $2.15\text{E}-07$ for adults. Hence, the atmospheric deposition should be insulated to lessen the CR. Based on Monte Carlo simulation results of CR, Fig. S4 shows that Cr content was the most contributed metal due to the highest variance and high-rank correlation for industrial emission, coal-power-plant, and atmospheric deposition sources, while As concentration was the most attributed metalloid due to the largest variance and high-rank correlation for agrochemical practice and geogenic sources (Fig. S2). The TCR of apportionment pollution sources decreased in the following order: atmospheric deposition (adults: $3.56\text{E}-05$ and children: $2.72\text{E}-05$), coal-power-plant (adults: $1.29\text{E}-05$ and children: $9.88\text{E}-06$), industrial emission (adults:

$1.02\text{E}-05$ and children: $7.82\text{E}-06$), agrochemical practice (adults: $6.22\text{E}-06$ and children: $4.89\text{E}-06$) and geogenic source (adults: $6.14\text{E}-06$ and children: $4.78\text{E}-06$), respectively (Table S6).

In general, our research's NCR and CR values for adults and children were considerably smaller than those from other researchers obtained values (Sun et al., 2022; Yang et al., 2019). Spatially, concomitant with the distribution pattern of concentration levels, the NCR of soil TMLs are commonly higher in the western and southwestern parts and much lower in the eastern and north-central parts of the study region (Fig. 7). Similarly, CR of soil TMLs is usually higher in the central and western parts and much lower in the eastern part of the investigated region (Fig. 8). These TMLs, As, Cr, and Cd pose elevated risks owing to their enhanced concentrations (Table 3). Analogous findings were also reported previously in soils in different parts of many countries (Duan et al., 2020; Islam et al., 2020; Wang et al., 2018). Reasonably, ingestion is the prime exposure route, contributing to most NCR and CR risks compared to the other two (inhalation and dermal) routes (Table S7). This observation is consistent with earlier cited works that the health risk posed by ingestion was many folds of magnitude higher than those via inhalation and external pathways (Chen et al., 2016; Guo et al., 2020; Zhang et al., 2018). Overall, adults demonstrate marginally higher NCR and CR triggered by these five sources than children, implying adults are more susceptible to the health risk of soil TMLs (Zhang et al., 2021) (Fig. 7; Table 3). The main reason is that adults could be associated with their higher body weight, hand-to-mouth behavior pattern, less skin surface area, shorter duration of exposure pathway, and higher respiration rate (Huang et al., 2018; Jiang et al., 2021).

Furthermore, similar trends of NCR and CR from five sources were observed for both age groups (Fig. S3). Source-specific health risk reveals that agrochemical practices (AP) were considered the key contributor to NCR for 32.37% of adults and 31.54% for children, which is slightly higher than the additions to contents of TMLs in soils (22.2% for PMF and 28.2% for APCS-MLR). On the other hand, industrial emission contributed the least (13.4% for PMF and 17.2% for APCS-MLR) to the deposition of TMLs in the soil surfaces; for that reason, it only attributed a comparatively small percentage to the NCR (adults: 9.1%; children: 8.96%). This might be

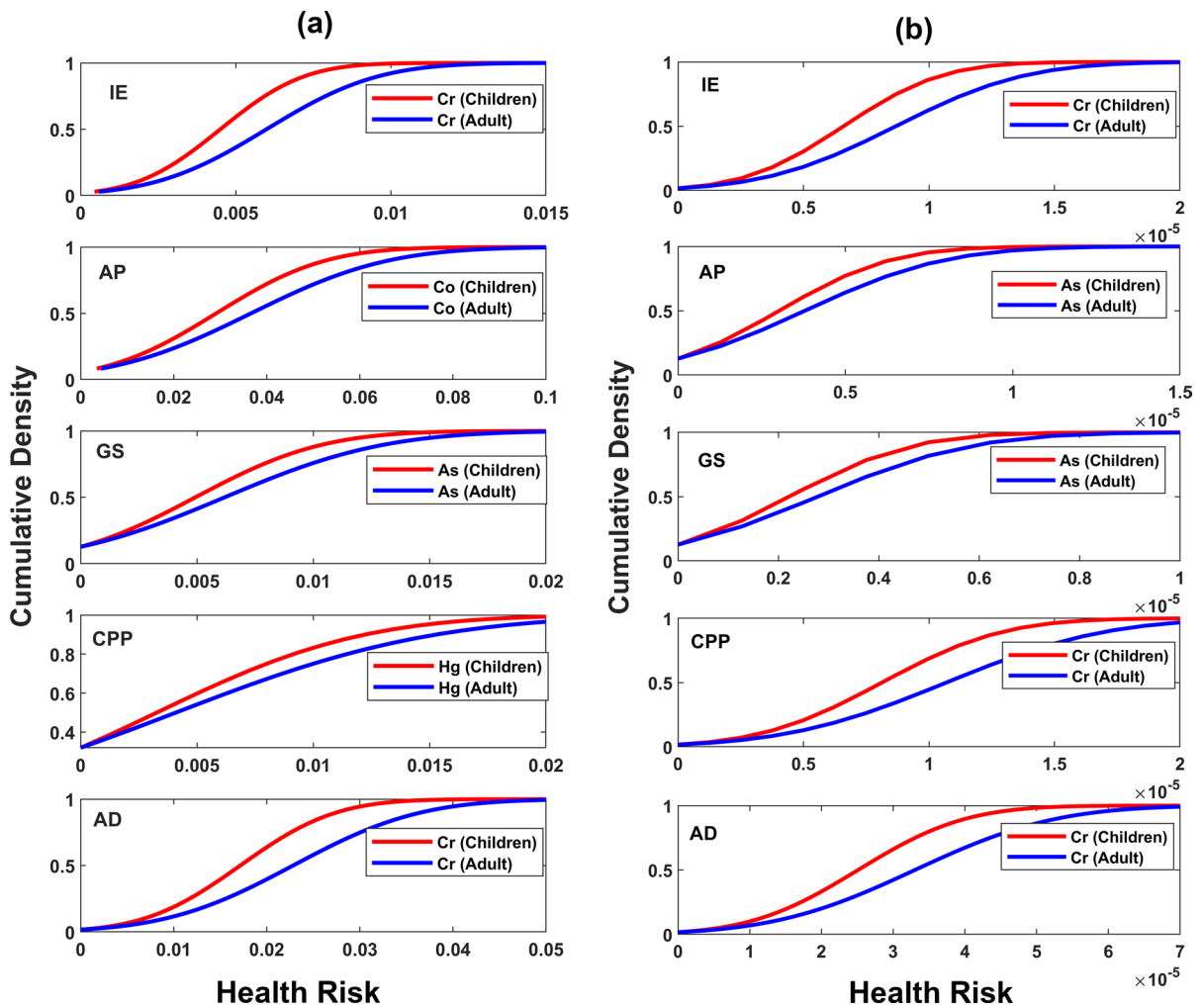


Fig. 7 Cumulative density function (CDF) with normal distribution fitting for **a** non-carcinogenic risk and **b** carcinogenic risk of children and adults for most contributed metal(loid)s according to Monte Carlo Simulation

elucidated by agricultural inputs containing Zn, As, and Cu with the elevated toxicity, which led to cause increased/enhanced NCR compared to other TMLs (Bhuiyan et al., 2010; Islam et al., 2021; Yang et al., 2019). Besides, the elevated SF and low RfD of As might interpret its high NCR compared with other TMLs (USEPA, 2011). Previous studies confirmed that bio-organic fertilizers could be used instead of agrochemical fertilizers which can efficiently lessen the deposition of TMLs in soils (Jiang et al., 2020). It is assumed that decreasing the using agrochemical fertilizers and pesticides can efficiently reduce the health risks of inhabitants. However, to restrict the human health risks of soil TMLs contamination in the

study region for safeguarding public health perspective, specific attention need be paid to the agrochemical practice, particularly for fertilizers and pesticides used in an agricultural crop field in the rural areas. Achieving sustainable development goals, this study recommends the application of low-toxic bio-organic fertilizers in place of agrochemical fertilizers.

Conversely, atmospheric deposition was regarded as the primary contributor to CR for 50.11% for adults and 49.83% for children, which is two times higher than the contributions to abundances of TMLs in the soils (25.1% for PMF and 16.3% for APCS-MLR). Although geogenic sources contributed moderately (15.1% for PMF and 19% for APCS-MLR) to

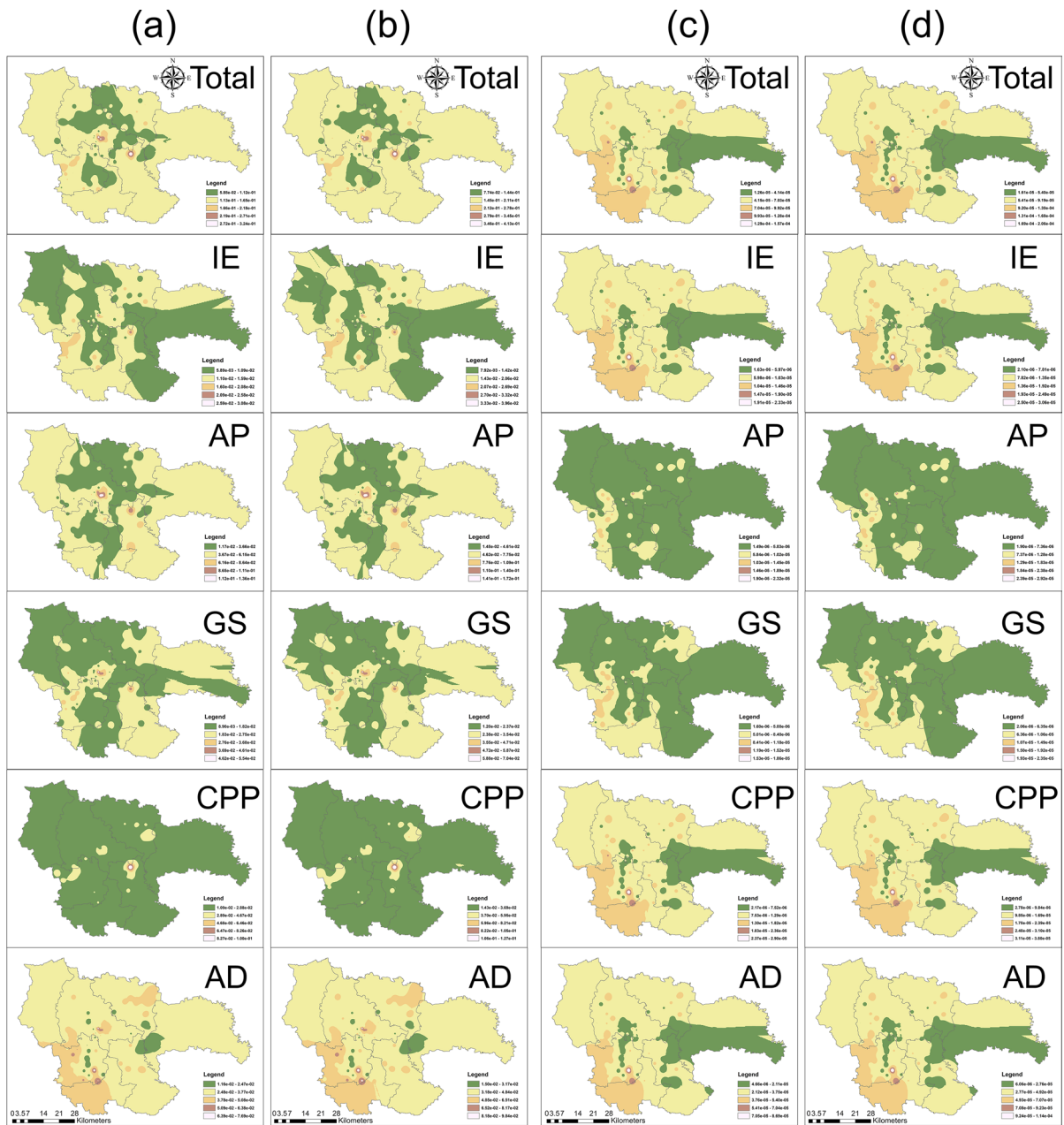


Fig. 8 Spatial distributions of the health risks of soil toxic metal(loid)s from the total factor (measured concentrations) and the five pollution sources detected for non-carcinogenic risk of **a** children and **b** adults and carcinogenic risk of **c** children and **d** adults

the deposition of toxic metal(loid)s in the soil layers, it only ascribed a reasonably small amount to the CR (adults: 8.63%; children: 8.76%). The high SF might explain this and low RfD of Cr and its high CR compared with other TMLs (USEPA, 2011). Many earlier studies have stated the high human health risks

caused by As, Pb, and Cr in soils (Guo et al., 2021; Habib et al., 2019a). It is worth mentioning that there is a big difference between the source distribution and source-oriented carcinogenic risk assessment, which is possibly caused by the diverse soil toxicities of elements. Hence to acquire inclusive information to

counteract and control the probable carcinogenic risk of TMLs, it is crucial to focus the source-oriented human health risk appraisal as/since the risk linkage with diverse sources will differ as to the contents and geochemical species may vary from source to source and distance to distance (Jiang et al., 2020; Sun et al., 2021; Zhang et al., 2021). For instance, recent research carried out by Guo et al. (2021) revealed that the inputs from the distribution of contamination sources to the total health risk could be diverse from their attributions to the entire mass contents. Thus, this study highlighted that community/civic health might not be well-protected if the remedial measures only released the TMLs in the soils. To better safeguard human health risk perspective, agrochemical practices and atmospheric deposition need to be prioritized for counteracting and risk control as the best alternative option.

Uncertainty assessment

Uncertainty in source distributions and risk assessment may exist in this investigation. To highlight the uncertainty of PMF method outputs, both base model bootstrap (BS) and base model displacement (DISP) models were used in the PMF method. The factors generated using this model were submitted to the base run for BS evaluation based on the relationships among the factor contributions (USEPA, 2014). The fact that BS results for five factors exceeded 90% in our study suggests that the base run's factor profiles are consistent. In addition, Fig S5 shows the diversity of the source inputs to the soil TMLs derived/obtained using the BS model. The BS model was tested in our study with a total of 100 trials and 80 Bootstrap random seeds, suggesting that the PMF model's results are quite reliable. The inter-quartile limit was exceeded in the base run values of Fe and As in factor 1, Cr, Fe, and Co in factor 4, and Cr, Fe, Se, and Zn in factor 5, showing that multiple hazardous soil metal(loid) findings/outcomes could impact PMF outcomes. The DISP method was shown to be an effective screening method for the PMF model's vigor and dependability. There were no factor swaps for the lowest Q (percent Q) for five factors in the DISP evaluation, indicating that PMF results were credible. In addition, Fig S6 shows the heterogeneity in the source apportionment contribution to soil TMLs derived from the DISP model. Except for Mn

in factor 2, the base run values of soil metal(loid)s were all within the inter-quartile range, indicating no significant rotational uncertainty. Overall, the results of the BS and DISP models indicated that the PMF findings are reliable and acceptable (Fig. S7).

Additionally, uncertainty existed in the human health risk appraisal. For instance, only twelve TMLs were regarded in our research, and their total factors and potential health risks were computed according to the sum of the possible risk of each metal(loid)s (Augustsson et al., 2018). Besides, the bioavailability of TMLs was not accounted, which might have reasoned an overestimation/underestimation in the probabilistic health risk model (Islam et al., 2020; Mehta et al., 2020). These drawbacks might have influenced the precision of the outcomes of our research to a specific extent. Thus, the bioavailability and the relationships among the metal(loid)s should be accounted for PMF method in future studies. Generally, this research highlighted that the most significant source distribution of pollution should not be chaotic with the leading source of human health risks in probabilistic health risk appraisal. However, this is the first inclusive and integrated/comprehensive effort in the country to help control the pollution source of soil, target TMLs and aid policy-makers articulate policies to lessen coal field management costs and mitigate potential further soil pollution and reduce the associated health risk. The control factor linked with precise human-induced inputs, including pH, traffic emission, gross domestic product, population density, industrial growth, and land use/land cover, deserves further investigation into the source apportionment of TMLs aid local authority make precise decisions.

Conclusions

This research aims to report for the first time the source characterization and source-oriented health risk evaluation of TMLs in topsoil by combining advanced techniques and models and scientifically analyzing the TML content in the agricultural soils of northwest Bangladesh caused by the coal basin. The mean concentrations of the analyzed metal(loid)s were in the descending order of $Mn > Zn > Cr > Ni > Cu > Pb > Co > As > Fe > Se > Cd > Hg$. The source apportionment findings revealed that five sources, including agrochemical practices, industrial

emissions, coal-power-plant, geogenic sources, and atmospheric deposition, were recognized and quantified using PMF and APCS-MLR models, aided by co-occurrence networks, PCoA, and SOM analysis. Agrochemical practices were the critical pollution source for NCR, whereas atmospheric deposition was CR's largest pollution source for children and adults. Uncertainty assessment demonstrated that the PMF outcomes were more valid and suitable than APCS-MLR model. In future studies, the bioavailability and exchange of toxic soil metal(loid)s should be coupled with the PMF method to get the precise outcomes of source-oriented risk. The study outcomes give a novel insight into preventing and controlling soil TMLs in the coal mine basin. Also, in its first application in Bangladesh, the proposed integrated approach here will provide a helpful tool to revise the information in the risk assessment and enable consequent mitigation schemes. Hence, importance should be given to controlling and managing agrochemical practices and atmospheric deposition to safeguard the study basin's aquatic ecosystem and human health.

Acknowledgements This research was financially supported by Prince of Songkla University and Ministry of Higher Education, Science, Research and Innovation under the Reinventing University Project (Grant Number REV65007). We are grateful to the technical personnel who are dedicatedly assigned to operate the TRIGA Mark II research reactor at the center for research reactor of Bangladesh, Atomic Energy Commission. Authors would also like to greatly acknowledge the authority of the Geological Survey of Bangladesh for all other forms of support for the study.

Author contributions ARMTI, MAH, and MH, designed, planned, conceptualized, drafted the original manuscript, and RK, and MSI, was involved in statistical analysis, interpretation; ARMTI, and KP, contributed instrumental setup, data analysis, validation; MAH, SCP, CR, and MH, contributed to editing the manuscript, literature review, proofreading; MV, MSI, MYM, and ARMTI, were involved in software, mapping, and proofreading during the manuscript drafting stage.

Funding Not applicable.

Data availability Data are available upon request on the corresponding author.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Adimalla, N. (2020). Heavy metals pollution assessment and its associated human health risk evaluation of urban soils from Indian cities: A review. *Environmental Geochemistry and Health*, 42(1), 173–190.
- Augustsson, A., Uddh-Söderberg, T., Filipsson, M., Helmfriid, I., Berglund, M., Karlsson, H., Hogmalm, J., Karlsson, A., & Alriksson, S. (2018). Challenges in assessing the health risk of consuming vegetables in metal-contaminated environments. *Environment International*, 113, 269–280.
- Barberán, A., Bates, S. T., Casamayor, E. O., & Fierer, N. (2012). Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME Journal*, 6, 343–351.
- BBS. (2018). *Yearbook of agricultural statistics of Bangladesh, 2017*. Dhaka: Bangladesh Bureau of Statistics.
- Bhuiyan, M. A., Parvez, L., Islam, M. A., Dampare, S. B., & Suzuki, S. (2010). Heavy metal pollution of coal mine-affected agricultural soils in the northern part of Bangladesh. *Journal of Hazardous Materials*, 173(1–3), 384–392.
- Bilgen, S. (2014). Structure and environmental impact of global energy consumption. *Renewable and Sustainable Energy Reviews*, 38, 890–902.
- Bowen, H. J. M. (1979). *Environmental chemistry of the elements*. New York: Academic Press.
- Cai, L. M., Wang, Q. S., Wen, H. H., Luo, J., & Wang, S. (2019). Heavy metals in agricultural soils from a typical township in Guangdong Province, China: Occurrences and spatial distribution. *Ecotoxicology and Environmental Safety*, 168, 184–191.
- Cai, L., Xu, Z., Bao, P., He, M., Dou, L., Chen, L., Zhou, Y., & Zhu, Y. G. (2015). Multivariate and geostatistical analyses of the spatial distribution and source of arsenic and heavy metals in the agricultural soils in Shunde, Southeast China. *Journal of Geochemical Exploration*, 148, 189–195.
- Cao, Q., Yang, L., Ren, W., Song, Y., Huang, S., Wang, Y., & Wang, Z. (2021). Spatial distribution of harmful trace elements in Chinese coalfields: An application of Web-GIS technology. *Science of the Total Environment*, 755, 142527.
- Chen, H., Teng, Y., Lu, S., Wang, Y., Wu, J., & Wang, J. (2016). Source apportionment and health risk assessment of trace metals in surface soils of Beijing metropolitan, China. *Chemosphere*, 144, 1002–1011.
- Clarke, K. R., & Gorley, R. N. (2015). Getting started with PRIMER v7. *PRIMER-E: Plymouth, Plymouth Marine Laboratory*, 20.
- Ćujić, M., Dragović, S., Đorđević, M., Dragović, R., & Gajić, B. (2016). Environmental assessment of heavy metals around the largest coal fired power plant in Serbia". *CATENA*, 139(2016), 44–52.
- Deng, J., Zhang, Y., Qiu, Y., Zhang, H., Du, W., Xu, L., Hong, Y., Chen, Y., & Chen, J. (2018). Source apportionment of PM_{2.5} at the Lin'an regional background site in China with three receptor models. *Atmospheric Research*, 202, 23–32.

- Duan, Y., Zhang, Y., Li, S., Fang, Q., Miao, F., & Lin, Q. (2020). An integrated method of health risk assessment based on spatial interpolation and source apportionment. *Journal of Cleaner Production*, 276, 123218.
- ECD. (1986). Council of the European Communities, 1986. Directive (86/278/EEC) on the protection of the environment, and in particular of the soil, when sewage sludge is used in agriculture. *Official Journal of the European Communities*, L181, 6–12.
- Frie, A. L., Dingle, J. H., Ying, S. C., & Bahreini, R. (2017). The effect of a receding saline lake (The Salton Sea) on airborne particulate matter composition. *Environmental Science and Technology*, 51, 8283–8292.
- Geetha, S., Reddy, B., & Hemalatha, K. P. J. (2017). Physico-chemical analysis of selected agricultural soil samples in Kommangipanchyathi, Chintapallimandal, Visakhapatnam. *International Journal of Information Retrieval Research*, 4(1), 3530–3532.
- Guan, Q., Wang, F., Xu, C., Pan, N., Lin, J., Zhao, R., Yang, Y., & Luo, H. (2018). Source apportionment of heavy metals in agricultural soil based on PMF: A case study in Hexi Corridor, northwest China. *Chemosphere*, 193, 189–197.
- Guo, B., Hong, C., Tong, W., Xu, M., Huang, C., Yin, H., Lin, Y., & Fu, Q. (2020). Health risk assessment of heavy metal pollution in a soil-rice system: A case study in the Jin-Qu Basin of China. *Scientific Reports*, 10(1), 1–11.
- Guo, G., Wang, Y., Zhang, D., & Lei, M. (2021). Source-specific ecological and health risks of potentially toxic elements in agricultural soils in Southern Yunnan Province and associated uncertainty analysis. *Journal of Hazardous Materials*, 417, 126144.
- Habib, M. A., Basuki, T., Miyashita, S., Bekelesi, W., Nakashima, S., Phoungthong, K., Khan, R., Rashid, M. B., Islam, A. R. M. T., & Techato, K. (2019b). Distribution of naturally occurring radionuclides in soil around a coal-based power plant and their potential radiological risk assessment. *Radiochimica Acta*. <https://doi.org/10.1515/ract-2018-3044>
- Habib, M. A., Basuki, T., Miyashita, S., Bekelesi, W., Nakashima, S., Techato, K., Khan, R., Majlis, A. B. K., & Phoungthong, K. (2019a). Assessment of natural radioactivity in coals and coal combustion residues from a coal-based thermoelectric plant in Bangladesh: Implications for radiological health hazards. *Environmental Monitoring and Assessment*. <https://doi.org/10.1007/s10661-018-7160-y>
- Habib, M. A., & Khan, R. (2021). Environmental impacts of coal-mining and coal-fired power-plant activities in a developing country with global context (chapter 24). In R. J. Cabin (Ed.), *Environmental challenges and solutions, Springer book series*. Switzerland: Springer.
- Halim, M. A., Majumder, R. K., & Zaman, M. N. (2015). Paddy soil heavy metal contamination and uptake in rice plants from the adjacent area of Barapukuria coal mine, northwest Bangladesh. *Arabian Journal of Geosciences*, 8(6), 3391–3401.
- Hossain, M. N., Paul, S. K., & Hasan, M. M. (2015). Environmental impacts of coal mine and thermal power plant to the surroundings of Barapukuria, Dinajpur, Bangladesh. *Environmental Monitoring and Assessment*, 187, 202.
- Hossen, M. A., Chowdhury, A. I. H., Mullick, M. R. A., & Hoque, A. (2021). Heavy metal pollution status and health risk assessment vicinity to Barapukuria coal mine area of Bangladesh. *Environmental Nanotechnology, Monitoring and Management*, 16, 100469.
- Huang, Y., Chen, Q., Deng, M., Japenga, J., Li, T., Yang, X., & He, Z. (2018). Heavy metal pollution and health risk assessment of agricultural soils in a typical peri-urban area in southeast China. *Journal of Environmental Management*, 207, 159–168.
- Huang, Y., Li, T., Wu, C., He, Z., Japenga, J., Deng, M., & Yang, X. (2015). An integrated approach to assess heavy metal source apportionment in peri-urban agricultural soils. *Journal of Hazardous Materials*, 299, 540–549.
- Islam, A. R. M. T., Bodrud-Doza, Md., Rahman, M. S., Amin, S. B., Chu, R., & Al Mamun, H. (2019). Sources of trace elements identification in drinking water of Rangpur district, Bangladesh and their potential health risk following multivariate techniques and Monte-Carlo simulation. *Groundwater for Sustainable Development*, 9, 100275.
- Islam, A. R. M. T., Hasanuzaman, M., Islam, H. M. T., Mia, M. U., Khan, R., Habib, M. A., Rahman, M. M., Siddique, M. A. B., Moniruzzaman, M., & Rashid, M. B. (2020). Quantifying source apportionment, co-occurrence and ecotoxicological risk of metals from up-mid downstream river segments, Bangladesh. *Environmental Toxicology and Chemistry*, 39, 2041–2054.
- Islam, M. S., Ahmed, M. K., Al-Mamun, M. H., & Hoque, M. F. (2014). Preliminary assessment of heavy metal contamination in surface sediments from a river in Bangladesh. *Environment and Earth Science*, 73, 1837–1848.
- Islam, M. S., Idris, A. M., Islam, A. M. R. T., Phoungthong, K., Ali, M. M., & Kabir, M. H. (2021). Geochemical variation and contamination level of potentially toxic elements in land-uses urban soils. *International Journal of Environmental Analytical Chemistry*. <https://doi.org/10.1080/03067319.2021.1977286>
- Islam, S., Ahmed, K., & Masunaga, S. (2015). Potential ecological risk of hazardous elements in different land-use urban soils of Bangladesh. *Science of the Total Environment*, 512, 94–102.
- Jiang, H. H., Cai, L. M., Wen, H. H., Hu, G. C., Chen, L. G., & Luo, J. (2020). An integrated approach to quantifying ecological and human health risks from different sources of soil heavy metals. *Science of the Total Environment*, 701, 134466.
- Jiang, Y., Wen, H., Zhang, Q., Yuan, L., & Liu, L. (2021). Source apportionment and health risk assessment of potentially toxic elements in soil from mining areas in northwestern China. *Environmental Geochemistry and Health*, 14, 1551–1566.
- Kabata-Pendias, A. (2000). *Trace elements in soils and plants* (4th ed.). Boca Raton: Taylor and Francis, CRC Press.
- Kashem, M. A., & Singh, B. R. (1999). Heavy metal contamination of soil and vegetation in the vicinity of industries in Bangladesh. *Water, Air, and Soil Pollution*, 115, 347–361.
- Khan, R., Hossain, S., Anik, A. H., Phoungthong, K., Islam, A. R. M. T., Saha, N., Idris, A. M., & Alam, M. (2023). Indexical and statistical approaches to investigate the integrated origins of elements in the sediment of Teesta River, Bangladesh: Sediment quality and ecological risks

- assessment. *Environmental Science Processes Impacts*. <https://doi.org/10.1039/D2EM00475E>
- Khan, R., Islam, H. M. T., & Islam, A. R. M. T. (2021). Mechanism of elevated radioactivity in a freshwater basin: Radiochemical characterization, provenance and associated hazards. *Chemosphere*. <https://doi.org/10.1016/j.chemosphere.2020.128459>
- Khan, R., Shirai, N., & Ebihara, M. (2015). Chemical Characteristics of R chondrites in the light of REE, Th, U and P abundances. *Earth and Planetary Science Letters*, 422, 18–27.
- Komnitsas, K., & Modis, K. (2006). Soil risk assessment of As and Zn contamination in a coal mining region using geostatistics. *Science of the Total Environment*, 371(1–3), 190–196.
- Kumar, S., Zhao, M., Zhang, H., Rahman, M. A., Luo, C., & Rahman, M. M. (2021). Distribution, contamination status and source of trace elements in the soil around brick kilns. *Chemosphere*, 263, 127882.
- Lin, Y., Ma, J., Zhang, Z., Zhu, Y., Hou, H., Zhao, L., Sun, Z., Xue, W., & Shi, H. (2018). Linkage between human population and trace elements in soils of the Pearl River Delta: Implications for source identification and risk assessment. *Science of the Total Environment*, 610, 944–950.
- Liu, H., Zhang, Y., Yang, J., Wang, H., Li, Y., Shi, Y., Li, D., Holm, P. E., Ou, Q., & Hu, W. (2021). Quantitative source apportionment, risk assessment and distribution of heavy metals in agricultural soils from southern Shandong Peninsula of China. *Science of the Total Environment*, 767, 144879.
- Liu, L., Wang, Z. P., Ju, F., & Zhang, T. (2015). Co-occurrence correlations of heavy metals in sediments revealed using network analysis. *Chemosphere*, 119, 1305–1313.
- Liu, Y. G., Gao, P., Su, J., Silva, E. B., Oliveira, L. M., Townsend, T., Xiang, P., & Ma, L. Q. (2019). PAHs in urban soils of two Florida cities: Background concentrations, distribution, and sources. *Chemosphere*, 214, 220–227.
- Luo, L., Mei, K., Qu, L., Zhang, C., Chen, H., Wang, S., Di, D., Huang, H., Wang, Z., Xia, F., Dahlgren, R. A., & Zhang, M. (2019). Assessment of the Geographical Detector Method for investigating heavy metal source apportionment in an urban watershed of Eastern China. *Science of the Total Environment*, 653, 714–722.
- Lv, J. (2019). Multivariate receptor models and robust geostatistics to estimate source apportionment of heavy metals in soils. *Environmental Pollution*, 244, 72–83.
- Lv, J., Liu, Y., Zhang, Z., Dai, J., Dai, B., & Zhu, Y. (2015). Identifying the origins and spatial distributions of heavy metals in soils of Ju country (Eastern China) using multivariate and geostatistical approach. *Journal of Soils and Sediments*, 15(1), 163–178.
- Masto, R. E., George, J., Rout, T. K., & Ram, L. C. (2017). Multi element exposure risk from soil and dust in a coal industrial area. *Journal of Geochemical Exploration*, 176, 100–107.
- Mehta, N., Cipullo, S., Cocerva, T., Coulon, F., Dino, G. A., Ajmone-Marsan, F., Padoan, E., Cox, S. F., Cave, M. R., & De Luca, D. A. (2020). Incorporating oral bioaccessibility into human health risk assessment due to potentially toxic elements in extractive waste and contaminated soils from an abandoned mine site. *Chemosphere*, 255, 126927.
- Men, C., Liu, R., Xu, L., Wang, Q., Guo, L., Miao, Y., & Shen, Z. (2020). Source-specific ecological risk analysis and critical source identification of heavy metals in road dust in Beijing, China. *Journal of Hazardous Materials*, 388, 121763.
- Modis, K., Vatalis, K. I., & Sachanidis, C. (2013). Spatiotemporal risk assessment of soil pollution in a lignite mining region using a Bayesian maximum entropy (BME) approach. *International Journal of Coal Geology*, 112, 173–179.
- Ogundele, L. T., Ayeku, P. O., Adebayo, A. S., Olufemi, A. P., Adejoro, I. A. (2020). Pollution indices and potential ecological risks of heavy metals in the soil: A case study of municipal wastes site in Ondo State, Southwestern, Nigeria. *Polytechnica* (pp. 1–9).
- Özkul, C. (2016). Heavy metal contamination in soils around the Tunçbilek thermal power plant (Kütahya, Turkey). *Environmental Monitoring and Assessment*, 188(5), 284.
- Paatero, P., Eberly, S., Brown, S. G., & Norris, G. A. (2014). Methods for estimating uncertainty in factor analytic solutions. *Atmospheric Measurement Techniques*, 7, 781–790.
- Parzenty, H. R., & Róg, L. (2021). Distribution and mode of occurrence of Co, Ni, Cu, Zn, As, Ag, Cd, Sb, Pb in the feed coal, fly ash, slag, in the topsoil and in the roots of trees and undergrowth downwind of three power stations in Poland. *Minerals*, 11(2), 133.
- Pietrzykowski, M., Socha, J., & van Doorn, N. S. (2014). Linking heavy metal bioavailability (Cd, Cu, Zn and Pb) in Scots pine needles to soil properties in reclaimed mine areas. *Science of the Total Environment*, 470, 501–510.
- Proshad, R., Kormoker, T., Sayed, A., Khadka, S., & Idris, A. M. (2021). Potential toxic metals (PTMs) contamination in agricultural soils and foodstuffs with associated source identification and model uncertainty. *Science of the Total Environment*, 789, 147962.
- Qi, P. Z., Qu, C. K., Albanese, S., Lima, A., Cicchella, D., Hope, D., Cerino, P., Pizzolante, A., Zheng, H., Li, J. J., & De Vivo, B. (2020). Investigation of polycyclic aromatic hydrocarbons in soils from Caserta provincial territory, southern Italy: Spatial distribution, source apportionment, and risk assessment. *Journal of Hazardous Materials*, 383, 121158.
- Rahman, M. M., Howladar, M. F., & Faruque, M. O. (2017). Assessment of soil quality for agricultural purposes around the Barapukuria coal mining industrial area, Bangladesh: Insights from chemical and multivariate statistical analysis. *Environmental Systems Research*, 6(1), 24.
- Reza, S. K., Baruah, U., Singh, S. K., & Das, T. H. (2015). Geostatistical and multivariate analysis of soil heavy metal contamination near coal mining area. *Northeastern India. Environmental Earth Sciences*, 73(9), 5425–5433.
- Ribeiro, J., Ferreira da Silva, E., Li, Z., Ward, C., & Flores, D. (2010). Petrographic, mineralogical and geochemical characterization of the Serrinha coal waste pile (Douro Coalfield, Portugal) and the potential environmental impacts on soil, sediments and surface waters. *International Journal of Coal Geology*, 83(4), 456–466.
- Rodríguez-Iruetagoiena, A., de Vallejuelo, S. F. O., Gredilla, A., Ramos, C. G., Oliveira, M. L., Arana, G., de Diego,

- A., Madariaga, J. M., & Silva, L. F. (2015). Fate of hazardous elements in agricultural soils surrounding a coal power plant complex from Santa Catarina (Brazil). *Science of the Total Environment*, 508, 374–382.
- Rudnick, R. L., & Gao, S. (2014). Composition of the continental crust (chapter 4). In H. D. Holland & K. K. Turekian (Eds.), *Treatise on geochemistry* (2nd ed., pp. 1–64). Amsterdam: Elsevier.
- Shi, X. M., Liu, S., Song, L., Wu, C. S., Yang, B., Lu, H. Z., Wang, X., & Zakari, S. (2022). Contamination and source-specific risk analysis of soil heavy metals in a typical coal industrial city, central China. *Science of the Total Environment*, 836, 155694.
- Siddique, M. A. B., Alam, M. K., Islam, S., Diganta, M. T. M., Akbor, M. A., Bithi, U. H., Chowdhury, A. I., & Ullah, A. A. (2020). Apportionment of some chemical elements in soils around the coal mining area in northern Bangladesh and associated health risk assessment. *Environmental Nanotechnology, Monitoring and Management*, 14, 100366.
- Singh, M., Thind, P. S., & John, S. (2018). Health risk assessment of the workers exposed to the heavy metals in e-waste recycling sites of Chandigarh and Ludhiana, Punjab, India. *Chemosphere*, 203, 426–433.
- Singh, U. K., & Kumar, B. (2017). Pathways of heavy metals contamination and associated human health risk in Ajay River basin, India. *Chemosphere*, 174, 183–199.
- Su, J., Qiu, Y., Lu, Y., Yang, X., & Li, S. (2021). Use of multivariate statistical techniques to study spatial variability and sources apportionment of pollution in rivers flowing into the Laizhou Bay in Dongying District. *Water*, 13, 772.
- Sun, J., Zhao, M., Huang, J., Liu, Y., Wu, Y., Cai, B., Han, Z., Huang, H., & Fan, Z. (2021). Determination of priority control factors for the management of soil trace metal(loid)s based on source-oriented health risk assessment. *Journal of Hazardous Materials*. <https://doi.org/10.1016/j.jhazmat.2021.127116>
- Sun, J., Zhao, M., Huang, J., Liu, Y., Wu, Y., Cai, B., Han, Z., Huang, H., & Fan, Z. (2022). Determination of priority control factors for the management of soil trace metal (loid) s based on source-oriented health risk assessment. *Journal of Hazardous Materials*, 423, 127116.
- Tamim, U., Khan, R., Jolly, Y. N., Fatema, K., Das, S., Naher, K., Islam, M. A., Islam, S. M. A., & Hossain, S. M. (2016). Elemental distribution of metals in urban river sediments near an industrial effluent source. *Chemosphere*, 155, 509–518.
- Tang, Q., Li, L., Zhang, S., Zheng, L., & Miao, C. (2018). Characterization of heavy metals in coal gangue-reclaimed soils from a coal mining area. *Journal of Geochemical Exploration*, 186, 1–11.
- Tozsın, G. (2014). Hazardous elements in soil and coal from the Oltu coal mine district, Turkey. *International Journal of Coal Geology*, 131, 1–6.
- USEPA (United States Environmental Protection Agency). 2009. Risk Assessment Guidance for Superfund (RAGS), Volume I, Human Health Evaluation Manual (Part E, Supplemental Guidance for Dermal Risk Assessment)
- USEPA (The United States Environmental Protection Agency). 2011. Exposure Factors Handbook, 2011 ed. (Final). U.S. Environmental Protection Agency, Washington, DC.EPA/600/R-09/052F.
- USEPA (The United States Environmental Protection Agency). 2014. EPA Positive Matrix Factorization (PMF) 5.0 Fundamentals and User Guide., United States Environmental Protection Agency, Washington, DC, USA. Available from, https://www.epa.gov/sites/production/files/2015-02/documents/pmf_5.0_user_guide.pdf. Accessed November 2021.
- USEPA. 2018. Positive Matrix Factorization Model for environmental data analyses. (Online) Accessed on 24th February <https://www.epa.gov/air-research/positive-matrix-factorization-model-environmental-data-analyses>
- Wang, C. H., Zhou, S. L., Song, J., & Wu, S. H. (2018). Human health risks of polycyclic aromatic hydrocarbons in the urban soils of Nanjing, China. *Science of the Total Environment*, 612, 750–757.
- Wang, F., Guan, Q., Tian, J., Lin, J., Yang, Y., Yang, L., & Pan, N. (2020). Contamination characteristics, source apportionment, and health risk assessment of heavy metals in agricultural soil in the Hexi Corridor. *CATENA*, 191, 104573.
- Wang, J., Wang, L. L., Wang, Y. X., Tsang, D., Yang, X., Beiyuan, J. Z., Yin, M. L., Xiao, T. F., Jiang, Y. J., Lin, W. L., Zhou, Y. C., Liu, J., Wang, L., & Zhao, M. (2021). Emerging risks of toxic metal(loid)s in soil-vegetables influenced by steel-making activities and isotopic source apportionment. *Environment International*, 146, 106207.
- Wu, J., Li, J., Teng, Y., Chen, H., & Wang, Y. (2020). A partition computing-based positive matrix factorization (PC-PMF) approach for the source apportionment of agricultural soil heavy metal contents and associated health risks. *Journal of Hazardous Materials*, 388, 121766.
- Xiao, W., Chen, W., & Deng, X. (2021). Coupling and coordination of coal mining intensity and social-ecological resilience in China. *Ecological Indicators*, 131, 108167.
- Xiao, X., Zhang, J., Wang, H., Han, X., Ma, J., Ma, Y., & Luan, H. (2020). Distribution and health risk assessment of potentially toxic elements in soils around coal industrial areas: A global meta-analysis. *Science of the Total Environment*, 713, 135292.
- Yang, J., Sun, P., Zhang, X., Wei, X. Y., Huang, Y. P., Du, W. N., Qadeer, A., Liu, M., & Huang, Y. (2021). Source apportionment of PAHs in roadside agricultural soils of a megacity using positive matrix factorization receptor model and compound-specific carbon isotope analysis. *Journal of Hazardous Materials*, 403, 123592.
- Yang, S., Zhao, J., Chang, S. X., Collins, C., Xu, J., & Liu, X. (2019). Status assessment and probabilistic health risk modeling of metals accumulation in agriculture soils across China: A synthesis. *Environment International*, 128, 165–174.
- Zhai, M., Totolo, O., Modisi, M. P., Finkelman, R. B., Kelesitse, S. M., & Menyatso, M. (2009). Heavy metal distribution in soils near Palapye, Botswana: An evaluation of the environmental impact of coal mining and combustion on soils in a semi-arid region. *Environmental Geochemistry and Health*, 31(6), 759–777.

- Zhang, H., Zhang, F., Song, J., Tan, M. L., & Johnson, V. C. (2021). Pollutant source, ecological and human health risks assessment of heavy metals in soils from coal mining areas in Xinjiang, China. *Environmental Research*, *202*, 111702.
- Zhang, M., Wang, X., Liu, C., Lu, J., Qin, Y., Mo, Y., Xiao, P., & Liu, Y. (2020). Quantitative source identification and apportionment of heavy metals under two different land use types: Comparison of two receptor models APCS-MLR and PMF. *Res Environmental Science and Pollution Research*, *27*, 42996–43010.
- Zhang, P., Qin, C., Hong, X., Kang, G., Qin, M., Yang, D., Pang, B., Li, Y., He, J., & Dick, R. P. (2018). Risk assessment and source analysis of soil heavy metal pollution

from lower reaches of Yellow River irrigation in China. *Science of the Total Environment*, *633*, 1136–1147.

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.