ORIGINAL PAPER

Combination of enrichment factor and positive matrix factorization in the estimation of potentially toxic element source distribution in agricultural soil

Prince Chapman Agyeman (ORCID) · **Kingsley John** · **Ndiye Michael Kebonye** ·
Luboš Borůvka · **Radim Vašát**

Abstract The study intended to assess the level of pollution of potential toxic elements (PTEs) at different soil depths and to evaluate the source contribution in agricultural soil. One hundred and two soil samples were collected for both topsoil (51), and the subsoil (51) and the content of PTEs (Cr, Cu, Cd, Mn, Ni, Pb, As and Zn) were determined using inductively coupled plasma–optical emission spectroscopy (ICP–OES). The concentrations of Zn and Cd in both soil horizons indicated that the current study levels were higher than the upper continental crust (UCC), world average value (WAV), and European average values (EAV). Nonetheless, the concentration values

of PTEs such as Mn and Cu for EAV, As, Cu, Mn, and Pb for UCC, and Pb for WAV were lower than the average values of the corresponding PTEs in this study. The single pollution index, enrichment factor, and ecological risk revealed that the pollution level ranged from low to high. The pollution load index, Nemerow pollution index, and risk index all revealed that pollution levels ranged from low to high. The spatial distribution confirmed that pollution levels varied between the horizons; that is, the subsoil was considered slightly more enriched than the topsoil. Principal component analysis identified the PTE source as geogenic (i.e. for Mn, Cu, Ni, Cr) and anthropogenic (i.e. for Pb, Zn, Cd, and As). PTEs were attributed to various sources using enrichment factor-positive matrix factorization (EF-PMF) and positive matrix factorization (PMF), including geogenic (e.g. rock weathering), fertilizer application, steel industry, industrial sewage irrigation, agrochemicals, and metal works. Both receptor models allotted consistent sources for the PTEs. Multiple linear regression analysis was applied to the receptor models (EF-PMF and PMF), and their efficiency was tested and assessed using root-mean-square error (RMSE), mean absolute error (MAE), and R^2 accuracy indicators. The validation and accuracy assessment of the receptor models revealed that the EF-PMF receptor model output significantly reduces errors compared with the parent model PMF. Based on the marginal error levels in RMSE and MAE, 7 of the 8 PTEs (As, Cd, Cr, Cu, Ni, Mn, Pb, and Zn) analysed performed

P. C. Agyeman (✉) · K. John · L. Borůvka · R. Vašát
Department of Soil Science and Soil Protection, Faculty of Agrobiology, Food and Natural Resources, Czech University of Life Sciences Prague, 16500 Prague, Czech Republic
e-mail: agyeman@af.czu.cz

N. M. Kebonye
Department of Geosciences, Chair of Soil Science and Geomorphology, University of Tübingen, Rümelinstr 19-23, Tübingen, Germany

N. M. Kebonye
DFG Cluster of Excellence "Machine Learning", University of Tübingen, AI Research Building, Maria-von-Linden-Str. 6, 72076 Tübingen, Germany

better under the EF-PMF receptor model. The EF-PMF receptor model optimizes the efficiency level in source apportionment, reducing errors in determining the proportion contribution of PTEs in each factor. The purpose of building a model is to maximize efficiency while minimizing inaccuracy. The marginal error limitation encountered in the parent model PMF was circumvented by EF-PMF. As a result, EF-PMF is feasible and useful for apparently polluted environments, whether farmland, urban land, or peri-urban land.

Keywords Enrichment factor-positive matrix factorization · Potentially toxic elements · Principal component analysis · Spatial distribution · Multiple linear regression · Frýdek-Místek District

Introduction

Potentially toxic elements (PTEs) are a generalized term referring to toxic metals (loids) that threaten human health, the environment, or both (Adimalla et al., 2020a, 2020b; Agyeman et al., 2020). Studies on agricultural soil pollution have gained particular attention due to PTEs triggered by natural and anthropogenic sources (Song et al., 2018; Zheng et al., 2018). However, anthropogenic variables contribute even more to the PTE pollution of soils than natural sources. PTEs are typically found in virtually all environmental matrices, including but not limited to soil, plants, and water (Agyeman et al., 2020). Over the past few years, the pollution of agricultural land by PTEs has occurred in diverse ways, such as industrial wastewater, and car exhaust emissions have worsened this situation (Chen et al., 2015a; Wang et al., 2012). According to Liu et al., (2013), soil PTE pollution is nonbiodegradable and tenacious. However, soil PTE pollution has therefore been regarded as a global threat (Mamat et al., 2014).

Many studies have been carried out in recent years in many parts of the world on PTE toxicity, source detection, distribution patterns, and pollution degree (Adimalla et al., 2020a, 2020b; Mazhari et al., 2018; Rastegari Mehr et al., 2017). Ruiz-Fernández et al. (2019) stressed that PTE concentrations are spatially and temporally distributed and enriched in soil. Atafar et al., (2010) reported on the effects of fertilizer use on soil PTE concentrations in agricultural soil,

emphasizing that excessive use of manure and phosphate fertilizers increases PTE content in agricultural soil. According to Li et al. (2014), anthropogenic sources of PTEs in soil pollution include those from agricultural, urbanization, domestic waste, industrialization, and mining activities. Li et al. (2014) further stated that PTE leaching causes food quality degradation, posing environmental risks to the ecosystem. The widespread use of fertilizers, manures, and agricultural waste in agricultural soils has the potential to degrade soil quality while also destroying the terrestrial ecosystem (Hu et al., 2019). The toxicity level of soil geochemistry at a location, as well as anthropogenic effects, is important sources of PTE pollution (Song et al., 2018).

There has been much research done on assessing the distribution of PTE sources in agricultural soil and looking into the processes that contribute to the decrease in PTE pollution in agricultural soil. The estimation of PTE concentrations using pollution indices (enrichment factor, pollution load index, ecological risk assessment) and the use of multivariate analyses, such as principal component analysis (PCA), positive matrix factorization (PMF), and Pearson correlation matrix (PCM), provide quantitative knowledge. However, the findings of using pollution indices are reliable, allowing soil scientists to provide realistic solutions to ecological problems. The identification of sources, the relationship that exists between PTEs, and the distribution of sources have primarily been a technical approach in assessing soil quality in the soil science community, and soil scientists use it most frequently to examine and determine the PTE fraction of contribution in polluted soil. Numerous authors, such as Rodríguez Martín et al. (2013), Huang et al., (2015); and Lü and He, (2018), indicated that multivariate statistical analysis, such as PCA, had been used in recognizing the pattern of PTE sources in soil and that it is extensively used globally. However, to further determine the percentage distribution of PTEs, most authors never hesitate to apply positive matrix factorization (PMF) or principal component analysis/absolute principal component score analysis-multilinear regression (PCA/APCS-MLR). Many soil scientists and researchers, such as Xue et al. (2014) and Chen et al., (2016), Agyeman et al., (2020), have used the PMF approach to determine the potential source distribution in agricultural soil.

Receptor models are constantly applied in source apportionment studies, and some of the famous approaches used comprise positive matrix factorization (PMF), UNMIX, principal component analysis/absolute principal component score analysis-multi-linear regression (PCA/APCS-MLR), and chemical mass balance (CMB). In recent literatures, Fei et al., (2020; Hossain Bhuiyan et al., (2021); Salim et al., (2019); Wu et al., (2020); Zhang et al., (2020) relied mainly on PMF, APCS/PCA-MLR or both to compute and detect the elemental source distribution of PTEs. PMF and APCS/PCA-MLR are preferred over other receptor models due to the following competitive advantages: (i) they use effective monitoring procedures, and they establish a substantial database that has become a universal practice; (ii) these receptor models do not need prequantified source profiles (i.e. backwards tracking) in disparity with chemical mass balance (CMB); and (iii) the receptor model's capacity allows it to cope with large amounts of monitoring data (Lee et al., 2016).

The ability of PMF to apportion sources to PTEs is not in doubt, but other authors have also raised concerns about its limitations regarding efficiency and its ability to minimize error when applied. Among some of the errors reported in some papers, according to Yuanan et al. (2020), if the PTEs detected in surface soils have experienced significant selective enrichment, PMF may produce incorrect estimates; for instance, when the elements come from the same sources, they may no longer be coassorted. In another vein, Wu et al. (2020) also reported that PMF was unable to adequately determine the nature of the differences in Cr, Cu, Ni, and Pb detected in surface soils across the entire area, where variances caused by soil parent materials may be large. In addition, Gholizadeh et al. (2016) also outlined that the findings and R^2 values of the predicted/observed plots for the majority of the water quality indicators revealed a greater goodness of fit with the APCS-MLR to the pollution source apportionment in the tested river waters.

PTE investigation is not a new thing since it poses a colossal menace in the environment and can cause a devastating effect on flora, fauna, and humans at large. The region under investigation is composed of several towns, including Havirov, Terlicko, Trinec, Bystrica, Jablunkov, Mainly Jabunkova, and Hrcava. Trinec and Vitkovice are important areas for assessing and determining the distribution of PTEs and the associated ecological implications. Steel industry operates in and around the Ostrava neighbourhood. This study seeks to combine a pollution index (that is, enrichment factor (EF)) and PMF to determine the source distribution of PTEs. This conduit allows source apportionment to be computed using the estimated EF values of each respective PTE rather than using the raw data. The study area is primarily an agrarian community that is also home to several industries as well as the steel industry and metal works. Based on the productive agricultural sector and active industrial activities, it is important to determine the soil health and quality of the study area. Against this backdrop, this current study seeks to explore the following objectives to (i) evaluate the concentration of PTEs in agricultural soil at different depths, (ii) compare the pollution levels at different depths, (iii) investigate the concentration of the PTEs that are spatially distributed at different soil depths, and (iv) quantify, identify, and compare the source contributions of PTEs in agricultural soil using EF-PMF and PMF receptor models.

Materials and methods

The research area is in the Frydek Mistek district in the Czech Republic, Europe, on the border with the Moravian–Silesian area (Fig. 1). The district is a rather rugged landscape, primarily with the Moravian-Silesian Beskydy, part of the outer carpathian mountain and the highest mountains. The carpathians have an undulating relief and a natural rock indicating highlands and valleys, which splits the main inland depressions. Larger portions of the district are skewed to the outer west carpathians, and the carpathians have only a small portion to the north and north-west. Geographically, the territory of the district is predominantly carbon-producing, providing a conducive haven for mining activities by the Paskov and Staříč mines, which are currently out of operation (Czso, 2019). The study area is positioned within the geographical coordinates of latitude 49° 41 ′0 'north and longitude 18° 20′ 0′ east at an altitude ranging between 225 and 327 m above sea level, which is characterized by a cold temperate climate and a high amount of rainfall even in dry months. Frýdek-Místek has humid, partly wet summers and cool, dry, windy
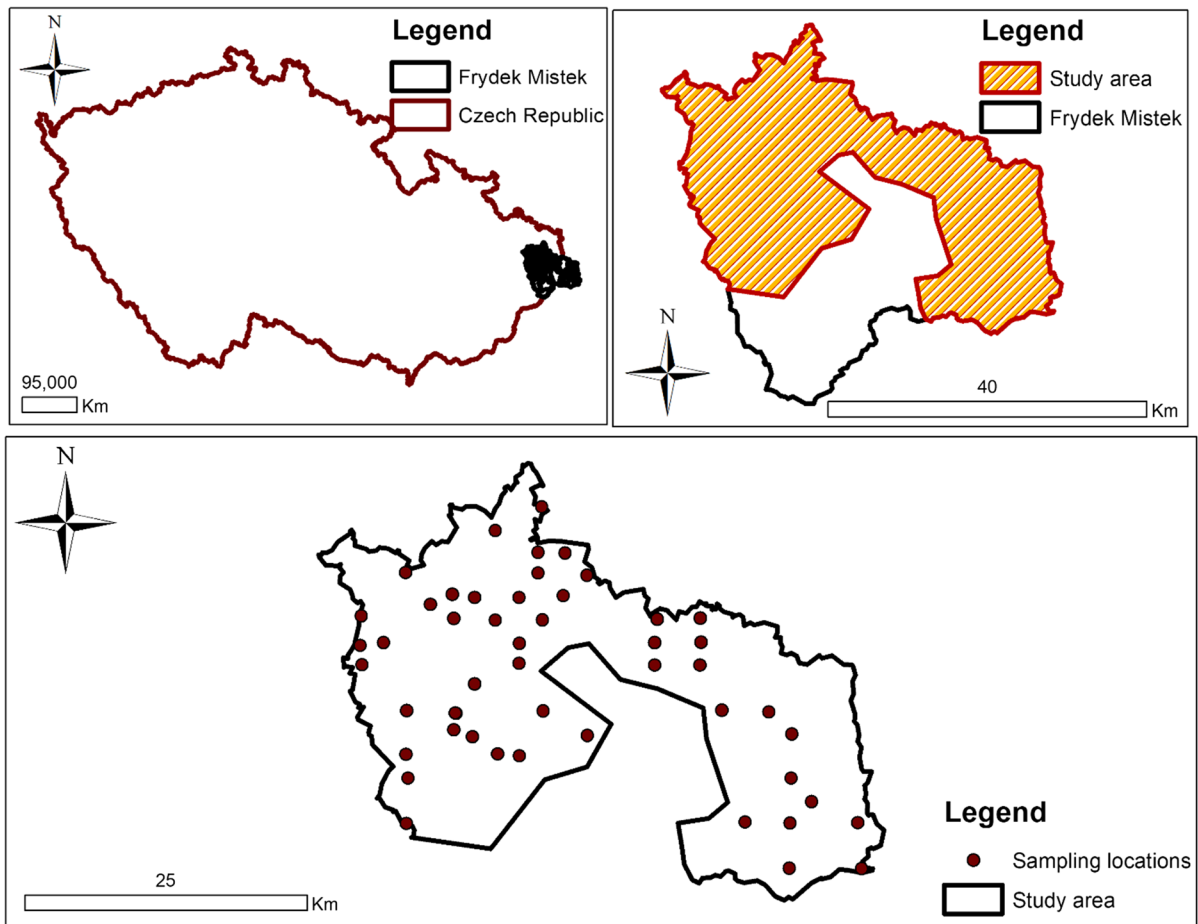
Fig. 1 Study area map showing sampling points

winters, and most of the winters are cloudy. Temperatures typically vary between − 5 °C and 24 °C over the year and are rarely lower than 14 °C or above 30 °C, while average annual precipitation is between 685 and 752 mm (Weather Spark, 2016). The area survey of the district is estimated at 1208 km², with 39.38% of the land area for cultivation and 49.36% being forests. The soil properties are clearly distinguished from the colour, structure, and carbonate content of the soil. The soil has a medium and fine texture content that is derived from parent materials. They are mainly colluvial, alluvial, or aeolian deposits. Some areas of the soil reveal mottles in the top and subsoil, which are mostly followed by concrete and bleaching. However, the predominant soil types in the region are Cambisols and stagnosols (Kozák et al., 2010). These soils are dominant in the Czech Republic, with

an elevation range of 160.6m to 455.1m for stagnosol and 59.6 to 493.5 m for cambisol (Vacek et al., 2020).

Soil sampling and soil analysis

A total sample size of 102 topsoils (51) and subsoils (51) was collected from agricultural land in the district of Frydek Mistek. The standard grid was the sample pattern adopted, and the soil sample intervals were 2 X 2 km using a handheld GPS unit (Leica Zeno 5 GPS) at a depth of 0–20 cm for topsoil and 20–40 cm for subsoil. The samples obtained were packaged in Ziploc bags, correctly labelled, and transported to the laboratory. The samples were air-dried, crushed by a mechanical device (Fritsch disk mill pulverize), and then sieved (< 2 mm) to obtain a pulverized sample. A gram of the dried, homogenized,

and sieved soil sample (sieve size < 2 mm) was inserted into a Teflon bottle and well labelled. Seven millilitres of 35% HCl and 3 ml of 65% HNO_3 (using automatic dispensers, a special dispenser for each acid) were dispensed in each bottle of Teflon, and the cap was gently closed to enable the sample to remain overnight for reactions (aqua regia procedure)(Cools & De Vos, 2016). The mixture was placed on a hot metal plate for 2 h to stimulate the process of digestion of the sample and left to cool. The mixture was then filtered to obtain a supernatant. The supernatant was transferred to a prepared 50-ml volumetric flask and then diluted with deionized water to 50 ml. The diluted supernatant was then filtered into 50-ml PVC tubes. Additionally, 1 ml of the diluted solution was further diluted with 9 ml of deionized water and filtered into a 12-ml test tube prepared for PTE pseudo-concentration of PTEs in this sample (Milićević et al., 2017). Metal concentration was measured by ICP–OES (inductively coupled plasma–optical emission spectrometry) (Thermo Fisher Scientific company, USA) in compliance with standard procedures and protocols. In addition to each study, the quality control and quality assurance processes were ensured (SRM NIST 2711a Montana II soil) by checking the reference criteria. Duplicate analyses were carried out to ensure that the error was minimized.

Pollution indices assessment

The consistency of agricultural soils must be assessed to determine the impact and toxicity of PTE pollution. Based on this, various pollution indices, such as the pollution index (PI), pollution load index (PLI), Nemerow pollution load (NPI), comprehensive ecological risk (ER), and risk assessment (RI), were used to assess the pollution status of the study area. Huang et al., (2018) and Sawut et al., (2018) argue that indices can reliably measure the status of soil pollution and the degree to which human activity impacts the soil environment. These indices are widely used in the assessment of PTE pollution in agricultural soil. The local background values of the PTEs in the study area are As (10 mg/kg), Cd (0.2 mg/kg), Cr (70 mg/kg), Cu (25 mg/kg), Ni (30 mg/kg), Mn (545 mg/kg), Pb (50 mg/kg), Zn (80 mg/kg).

Single pollution index (PI)

The single pollution index (PI) is the ratio of the soil PTE concentration to the geochemical background values. PI was introduced by Tomlinson et al., (1980), and the equation is given by

$$PI = \frac{C_n}{B_n} \tag{1}$$

where B_n is the geochemical background value of the PTE in the soil (mg/kg) and C_n is the concentration of the PTE in the soil (mg/kg). PI is categorized into a low level (PI ≤ 1), moderate level (1 < PI ≤ 3), considerable level (3 < PI ≤ 6), or high level (PI ≥ 6).

Pollution load index (PLI)

PLI is used for the overall assessment of the degree of soil pollution. This index proposes a simple way to display the soil deterioration resulting from the accumulation of PTEs. This equation was introduced by Tomlinson et al., (1980), and the equation is given by

$$PLI = \sqrt[n]{PI_1 \times PI_1 \times PI_1 \times \ldots \ldots \times PI_1} \tag{2}$$

where n represents the number of analysed PTEs and PLI is categorized into a low level (PLI ≤ 1), moderate level (1 < PLI ≤ 2), high level (2 < PLI ≤ 5), or extremely high level (PLI > 5) based on the degree of pollution.

Nemerow pollution index (PINemerow)

PINemerow computes the overall degree of pollution of the soil that consists of the concentration of all analysed PTEs (Qingjie et al. 2008). The index is used in the assessment of both the A horizons. The formula is given by

$$PI_{Nemerow} = \frac{\sqrt{(1/n \sum_{i-1}^{n} PI)^2 + PI_{max}^2}}{n} \tag{3}$$

where PI represents the computed values for the single pollution index, Pmax is the maximum values for the single pollution index of all the PTEs, the interpretation of PINemerow class values is given as ≤ 0.7 = clean, 0.7–1 = warning list, 1–2 = slight

pollution, 2–3 = moderate pollution, and ≥ 3 = heavy pollution.

Ecological risk assessment (ER and RI)

Ecological risk (RI) is an index used for the assessment of the degree of ecological risk caused by PTE concentrations in the soil. The index (RI) was introduced and applied by Hakanson, (1980), and the equation is given by

$$RI = \sum_{i=1}^{n} E_r^i \tag{4}$$

where n is the number of PTEs and E_r^I is the ecological risk index factor, which is given by

$$E_r^I = T_r^i \times PI \tag{5}$$

T_r^i is the toxicity response coefficient of a specific PTE (Hakanson, 1980), and PI represents the single pollution index. The toxicity response coefficients of the PTEs used are 30 (Cd), 10 (As), 5 (Cu), 5 (Pb), 2 (Cr), 2 (Zn), 2 (Ni), and 1 (Mn). The EI has five classifications: low risk (EI ≤ 40), moderate risk (40 < EI ≤ 80), considerable risk (80 < EI ≤ 160), high risk (160 < EI ≤ 320), and very high risk (EI ≥ 320). The RI has four categories, namely, low risk (RI ≤ 150), moderate risk (150 < RI ≤ 300), considerable risk (300 < RI ≤ 600), or very high risk (RI > 600).

PMF receptor model

A receptor model PMF was carried out using US EPA PMF 5.0 software with further description of the U.S. Environmental Protection Agency User Guide (Norris et al., 2014). The receptor model PMF is a multivariate method for factor analysis to solve the CMB, and the original data matrix X is represented in the following order $m \times n$, which can be given as

$$X = GF + E \tag{6}$$

G ($m \times p$) is a factor contribution matrix, F ($p \times n$) is a factor profile matrix, and E ($m \times n$) is a residual error matrix. E is given as

$$e_{ij} = \sum_{k=1}^{p} g_{ik} f_{ki} - x_{ij} \tag{7}$$

where i represents elements 1 to m, j represents elements 1 to n, and k represents the source from 1 to p.

The discharged factor contributions and profiles are acquired by the PMF model, which minimizes objective function Q under the constraint of nonnegative contributions, and the solution in the US-EPA PMF program is approximated by the Multilinear Engine-2 (ME-2)(Paatero, 1999)

$$Q = \sum_{i-1}^{n} \sum_{j-i}^{m} (e_{ij}/u_{ij})^2 \tag{8}$$

where u_{ij} is the uncertainty in the jth chemical element for sample i.

The uncertainty is computed based on the element-specific method detection limit (MDL), and the error percentage is measured by the standard reference materials. Since all the calculated contents are above the MDL, the uncertainty equations are taken as follows:

$$Unc = \sqrt{(ErrorFraction \times concerntration)^2 + (0.5 \times MDL)} \tag{9}$$

The US-EPA PMF 5.0 software provides a rotation method that sets the Fpeak value to boost oblique edges (Paatero & Hopke, 2009; Paatero et al., 2002). Positive Fpeak values are sharpened by F and G, while negative Fpeak values are transformed by comparison.

Enrichment factor (EF) and EF-PMF source apportionment

This is the index used to assess the concentration of PTEs in soil. For each PTE, the EF was determined to determine the concentration levels of the elements that are caused by anthropogenic activities. EF is mostly used to differentiate the source of PTEs that may be natural or anthropogenic (Kowalska et al., 2018). This includes the stabilization of the soil relative to the reference elements. Enrichment factor estimation is used to estimate the impact of anthropogenic activities related to metal abundance in sediments or soils. Ergin et al. (1991) defined EF by the following equation:

$$EF = \frac{C_n/C_{ref}}{B_n/B_{ref}} \qquad (10)$$

where Cn is the concentration of the examined element in the examined environment, Cref is the concentration of the examined element in the reference environment, Bn is the concentration of the reference element in the examined environment, and Bref is the concentration of the reference element in the reference environment. The threshold values used were the world average values (Kabata-Pendias, 2011). The interpretation of the estimated values is defined as follows: $EF < 2$ denoting deficiency to minimal enrichment, $2 < EF < 5$ representing moderate enrichment, $5 < EF < 20$ signifying significant enrichment, $20 < EF < 40$ suggesting very high enrichment, and $40 >$ indicating extremely high enrichment.

With the estimated EF values of the respective sampled points for each PTE, the source distribution was determined based on the values. In the normal sense of calculating PMF, raw data acquired after field analysis are used to determine the source apportionment. The novelty here is that instead of using the raw data for source analysis, the computed EF values will be used to ascertain the source distribution of the area under study. This approach is novel and rarely applied in PMF. The receptor model is based on the traditional PMF approach that combines enrichment factor with PMF to obtain a hybridized model EF-PMF. The receptor model EF-PMF is given as

$$(EF)_{ij} = \frac{(C_n/C_{ref})_{ij}}{(B_n/B_{ref})_i} \qquad (11)$$

where EF_{ij} is the calculated total enrichment factor of the PTEs from the jth source in the ith sampling site, $(C_n/C_{ref})_{ij}$ is both the concentration of the examined element in the examined environment, and Cref is the concentration of the examined element in the reference environment in the jth source from the ith sampling site, and $(B_n/B_{ref})_i$ is Bn, the concentration of the reference element in the examined environment, and Bref, the concentration of the reference element in the reference environment of the reference element of the respective PTEs.

Multiple linear regression model

A multiple linear regression model (MLR) is a regression model that describes the relationship between the response variable and multiple predictor variables by utilizing linearly inserted parameters determined using the least-squares approach. In MLR, the least square model is a prediction function that is directed towards a soil attribute following the selection of an explanatory variable. PTEs were employed as response variables to build a linear relationship using the explanatory variable (that is, the factor contribution from both models). The MLR equation is given as

$$y = a + \sum_{i-1}^{n} b_1 X x_i \pm \varepsilon_i \qquad (12)$$

where y signifies the response variable, an indicates the intercept, n connotes the number of predictors, b_1 denotes the partial regression of the coefficient, x_i implies the predictors or the explanatory variables, and ε_i signifies the error in the model, which is also called the residual. The model was utilized in R (K = tenfold cross-validation, which was repeated 5 times).

Data partitioning

A random data split approach was used to divide the data into a test dataset (with 25% for validation) and a training dataset (75% for calibration). The training dataset was used to calibrate the regression models, while the test dataset was utilized to assess generalization capabilities. (Kooistra et al., 2003). This was done to assess the suitability of the various models used to estimate PTE source apportionment. All of the models were subjected to a tenfold cross-validation process that was repeated five times. To predict the target variables, the factor contributions for each receptor model were employed as predictors or explanatory variables. *R* was used to carry out the modelling procedure.

Accuracy assessment and validation

A variety of validation criteria were utilized to determine the best and finest model suitable for

the estimation of source apportionment with pollution assessment-based positive matrix factorization receptor models while analysing the accuracy of the model and its validation. The receptor models were evaluated using the mean absolute error (MAE), root-mean-square error (RMSE), and R square, or coefficient of determination (R^2). R^2 describes the variation of the proportion in the response and is expressed by the regression model. The RMSE and the magnitude of the variability within the independent measurement define the model prediction capacity, while MAE determines the true quantitative value. The R^2 value must be high to establish the optimum receptor model using the validation criteria, and the closer the value is to 1, the higher the accuracy. According to Li et al., (2016), an0 R^2 criteria value of 0.75 is considered a satisfactory prediction. Methods for evaluating validation requirements using RMSE and MAE a lower obtained value is appropriate and deemed optimum for model selection. The following equation describes the validation procedures.

Mean absolute error

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} Y_i - \hat{Y}_i \tag{13}$$

R square

$$R^2(\%) = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \hat{Y}_i)^2} \tag{14}$$

Root-mean-square error

$$\text{RMSE}(\%) = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2 \tag{15}$$

where n represents the size of the observations, Y_i represents the measured response, and $\hat{Y}_i$ is also stated as the predicted response value, accordingly, for the ith observation term.

Data analysis

Statistical analyses were performed using KyPlot. PMF EPA 5.0 was used for source distribution estimation and is considered an excellent tool for estimating source apportionment. RStudio was used for both the principal component analysis and the assessment of the Pearson correlation matrix. Modelling and spatial distribution maps of the PTEs were analysed using ArcGIS, and inverse distance weighting (IDW) interpolation techniques were employed.

The IDW method for spatial interpolation is used to estimate cell values by weighing geometric data (points) in the vicinity of a processed cell. IDW interpolation is based on the idea that objects are more similar to each other than objects that are more subtly different. The effect of the variable entered on the map is assumed to decrease with an increasing distance from the sampling location. More weights are allocated to those points nearest to the target position, thus changing the assigned weights as the opposite 'pth distance' function, where the power function (*p*) is a positive actual number (Shukla et al., 2020). The forecast parameter for the target position consists of the sum of the 'allocated weights' and 'measured values' for each region.

Results and discussion

PTEs concentrations of soil samples

The descriptive statistics of the analysed PTE concentrations from the district of Frydek Mistek are presented in Table 1. The arithmetic mean concentrations of the PTEs (Cr, Cu, Cd, Mn, Ni, Pb, As, and Zn) in the topsoil and the subsoil were 27.33 mg/kg, 23.49 mg/kg, 1.65 mg/kg, 672.14 mg/kg, 16.88 mg/kg, 29.93 mg/kg, 4.02 mg/kg, and 81.67 mg/kg (topsoil), respectively, and those in the subsoil were 26.88 mg/kg, 23.22 mg/kg, 1.72 mg/kg, 694.29 mg/kg, 16.65 mg/kg, 31.41 mg/kg, 4.91 mg/kg, and 84.24 mg/kg, respectively. The maximum and minimum concentrations of the PTEs in the topsoil ranged from 3.21 mg/kg to 1691.76 mg/kg (maximum) and 0.95 mg/kg to 281.93 mg/kg (minimum), respectively, whereas those in the subsoil oscillated from 2.62 mg/kg to 1581.55 mg/kg (maximum) and 0.93 mg/kg to 213.23 mg/kg (minimum), respectively. Comparing the mean concentration of the topsoil and subsoil to the threshold levels (i.e. world average value, European average value, and upper continental crust) extracted from, Kabata-Pendias, (2011), the mean concentration levels of some of the PTEs in the current study (for both topsoil and subsoil) are higher than some of the PTEs (Zn and Cd) from all 3 geochemical background values. For instance, it is

Table 1 The concentration of PTEs in the study area, showing basic statistical and geochemical background values

Statistic	Mn	Ni	Pb	Zn	As	Cd	Cr	Cu
	< —————————————————-mg/kg————————————————— >							
Topsoil								
Mean	672.14	16.88	29.93	81.67	4.05	1.65	27.33	23.49
[e]S.D	272.37	7.5	8.84	27.96	2.2	0.5	9.53	11.53
[d]Coef.Var	40.52	44.4	29.53	34.23	54.31	30.19	34.87	49.08
Minimum	281.93	4.86	16.31	48.22	1.85	0.95	10.9	7.88
Maximum	1691.76	42.39	59.64	182.68	8.69	3.21	61.51	62.62
Kurtosis	2.5	3.18	0.82	1.42	− 1.24	− 0.55	11.05	8.13
Skewness	1.01	1.63	0.88	0.89	0.12	0.47	2.59	2.56
Subsoil								
Mean	694.29	16.65	31.41	84.28	4.91	1.72	26.88	23.22
S.D	247.39	6.81	9.28	24.95	2.44	0.42	9.66	13.04
[d]Coef.Var	36	41	30	30	50	25	36	56
Minimum	213.23	4.65	16.82	41.05	1.85	0.93	10.61	7.2
Maximum	1581.55	41.27	59.65	169.4	9.84	2.62	74.48	80.31
Kurtosis	2.5	3.18	0.82	1.42	− 1.24	− 0.55	11.05	8.13
Skewness	1.01	1.63	0.88	0.89	0.12	0.47	2.59	2.56
Thresholds								
[a] UCC	900	20	15	70	1.8	0.1	100	17.3
[b] WAV	488	29	27	70	6.83	0.41	59.5	38.9
[c] EAV	524	37	32	68.1	11.6	0.28	94.8	17.3

*(Kabata-Pendias, 2011) (pages 41 and 42) [a] Upper continental crust (UCC), [b]World average value (WAV) [c]European average value (WAV), [d]Coefficient of variability (CV), [e]standard deviation

evident from Table 1 that the Mn and Cu concentrations in the current study area are above the European average value (EAV). Similarly, the mean concentration value of Pb in this study is likewise higher than the mean value of the corresponding PTE in the world average value (WAV). On the other hand, the mean concentration values of As, Cu, Mn, and Pb in the upper continental crust (UCC) are equally lower than the obtained mean concentration values of PTEs in the present study.

Standard deviation values computed for both soil depths were found to be high due to the concentration of the PTEs having high variable heterogeneity in the study area. The normality and the abnormality of the PTE data distribution were determined using the skewness values. According to Chandrasekaran et al., (2015), if the PTE skew value ranges from 1 to − 1, it can be interpreted as a normal distribution; however, if the PTE value is slightly skewed positively (> 1), it is said to be an abnormal distribution. Generally, the skewness values of the following PTEs Pb, Zn, As, and Cd data were below 1; therefore, it can be interpreted that the distribution of PTE data is normal, whereas the other PTE data (Mn, Cr, Cu and Ni) distribution is abnormal and skewed in the right direction as well as leptokurtic.

Nezhad et al. (2015) reported that the coefficient of variance (CV) suggests the degree of variability within the concentrations of PTEs. Thus, if the CV varies between 0 and 20%, it is presumed that the PTEs are of a natural origin, and if it is above 20%, it implies that it is being influenced by anthropogenic activities. Thus, if $CV \leq 20\%$, this indicates low variability; $21\% \leq CV \leq 50\%$, it is considered moderate variability; $50\% \leq CV \leq 100\%$, it suggests high variability; and if CV is above 100%, it is regarded as exceptionally high variability. The coefficient of variation (CV %) of the PTEs in the current agricultural soils decreased in the following order: $As > Cu > Pb > Ni > Mn > Cr > Zn > Cd > Pb$. The results clarified that all the PTEs have moderate variability and are more homogeneous except As in both soils and Cu in the subsoil. Arsenic showed high variability in the topsoil, and Cu showed high variability in the subsoil. High variability in As and Cu indicates a nonhomogenous distribution, which explains the existence of a potential human-related impact.

Pollution characteristics

Single pollution index (PI) and pollution load index (PLI)

The single pollution index estimated revealed a heterogeneous distribution of the PTEs within the topsoil and the subsoil (Table S1). Out of the 51 samples collected for each soil (topsoil and the subsoil), most of the soils sampled showed low PTE pollution levels for all the metals except Mn and Cd. Thirteen soil samples showed low pollution levels for Mn in the topsoil and nine for the subsoil. Some of the PTE concentration levels in the topsoil exhibited moderate pollution levels, such as 31 samples for Mn, 29 for Pb, 7 for As, 1 for Cr, 4 for Ni, 31 for Zn, 12 for Cd, and 5 for Cr. In contrast, the moderate pollution level in the subsoil was as follows: Mn 41, Pb 35, As 13, Cr 1, Ni 3, Zn 33, Cd 4, and Cu 4. Cd and Mn pollution levels in both soils in some observed areas showed a considerable level of pollution, accounting for 37 (Cd),1 (Mn) in the topsoil and Cd (42), Mn (1) in the subsoil sample areas. Five areas sampled in the subsoil showed a high Cd pollution degree, whereas in the topsoil, only 2 samples indicated a high pollution rate. Regardless of the fact that the pollution levels in both soils range from moderate to high, the difference is significant. The subsoil level of enrichment was higher than the topsoil level in all the PTE pollution analysed except for Ni and Cu. The computed PI demonstrates that there is a downwards movement

of PTEs from the topsoil to the subsoil coupled with a geogenic base. Madrid et al., (2002) and Parveen et al., (2012) reported that vast amounts of PTEs accumulate mostly in surface layers of the soil (topsoil), but the results ascertained contradict that assertion due to leaching of PTEs from the topsoil to the subsoil eroding some level of accumulated concentration from the topsoil to the subsoil. Antoniadis and Alloway, (2002) findings corroborate our results that leaching plays a role in the mobility of PTEs from the topsoil to the subsoil.

The pollution load index computed displayed varied pollution levels, as indicated in Table S 1. Forty–eight out of the 51 sampled areas in the topsoil showed a low pollution load, while forty-nine were in the subsoil. Each of the soils displayed a moderate pollution load; however, the topsoil exhibited a high pollution load for 2 sampled areas, whereas the subsoil showed a high pollution load. From the PLI _IDW (pollution load index _inverse distance weighting Fig. 2) spatial distribution map, it was evident that the hotspot of the distribution of the pollution load in the subsoil is mainly centred in the middle of the north–western part of the map. The north–west part of the map is a primarily agrarian community, and therefore, the hotspot there may be attributed to the intensive agriculture in that vicinity. The topsoil also showed a hotspot in the middle of the north-western part of the map, as well as a hotspot in the eastern parts of the map. Comparatively, the pollution load in the topsoil appears to be denser than that
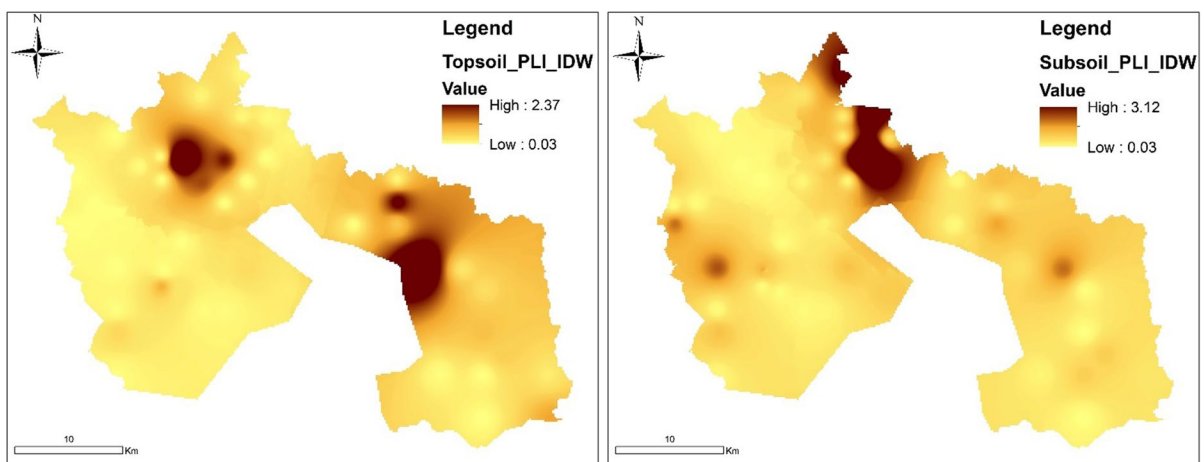
Fig. 2 Spatial distribution of the pollution load index (PLI _IDW) estimated values

of the subsoil. The steel industries and some metal works are in the north-eastern and eastern parts of the area of the study area. The spatial distribution map (PLI _IDW) displayed the pollution load on the topsoil rather than the subsoil. It clearly shows that anthropogenic activities such as industry, agriculture, vehicular traffic, sewage sludge, and atmospheric deposition are significant contributors to soil pollution in the upper layer (topsoil). The results from the PLI _IDW spatial map are coherent with the attained results from various studies, such as Zhu et al., (2006); Boyter et al., (2009); Gąsiorek et al., (2017); Zhu et al., (2017); Chen and Lu, (2018).

*Potential ecological risk index*

The potential ecological risk index for both soils exhibited varied pollution concentration levels, but all the PTEs showed a low ecological risk level except for cadmium (see table S2). The variability of cadmium in the topsoil demonstrates six moderate risk levels in the topsoil to one moderate risk level in the subsoil. The topsoil exhibited a considerable ecological risk level for 39 sampled locations compared to 42 locations showing a considerable risk level in the subsoil. Few locations within the study areas displayed a high level of ecological risk level for the topsoil and subsoil, accounting for 6 and 8 high ecological risk levels, respectively.

The risk index computed for both soils ranged from low risk to moderate risk, accounting for 31

and 20 for topsoil and 27 and 24 for subsoil, respectively, out of the 51 sampled locations for each soil (see Table S2). The RI_IDW spatial distribution map shows that the risk assessment level in the subsoil is higher than that in the topsoil. The hotspot pattern in the subsoil is denser than that of the topsoil from the north-eastern and south–eastern parts of the map. Regardless of the source of anthropogenic pollution, it may be responsible for the accumulation of PTEs on topsoil, which migrates from topsoil to subsoil, (Liu et al., 2016), it was complemented by the geogenic source.(Fig. 3).

Nemerow pollution index

Using the Nemerow pollution index to determine the pollution level in both soils, 9 (topsoil) and 2 (subsoil) sampled locations each were clean, 19 and 25 samples each fell within the warning perimeter category, and 22 and 24 were slightly polluted, respectively. Only a sample from the topsoil was found to be moderately polluted (see Table S1). Similarly, the spatial distribution map (Fig. 4) showed hotspots and higher pollution levels in the subsoil than in the topsoil. Pockets of hotspots in the subsoil were found in some parts of the map, except in the north–western area. The southern part of the subsoil spatial distribution map exhibited relatively sporadic pollution. The pollution patterns from the RI _IDW and $P_{nemerow}$ _IDW spatial distribution maps exhibit some coherency.
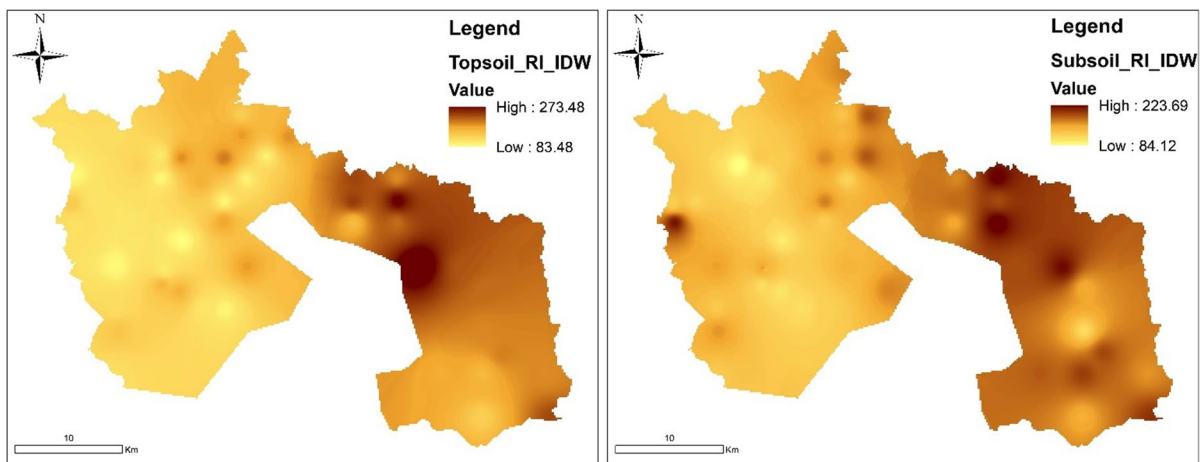
Fig. 3 Spatial distribution of the risk index (RI _IDW) estimated values
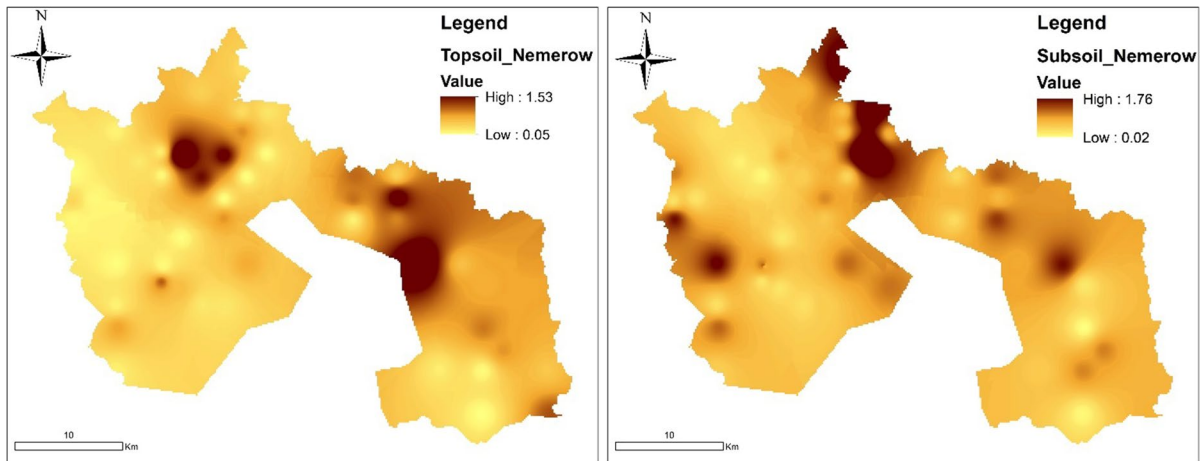
Fig. 4 Spatial distribution of the Nemerow pollution load ($P_{nemerow}$ _IDW) estimated values

Enrichment factor

The PTEs showed diverse enrichment levels ranging from a deficiency or minimal enrichment level to a significant enrichment level (see Table S3). All the PTEs displayed low or minimal enrichment levels in both soil horizons except for Cd. Low enrichment levels were computed for Mn, Ni, Pb, Zn, As, Cr, and Cu, which accounted for 15, 47, 24, 25, 45, 51, and 47 in the topsoil and 11, 48, 22, 20, 39, 50, and 47 in the subsoil. Thirty–four of the sampled locations fell within a moderate enrichment level for Mn, 4 for Ni, 27 for Pb, 26 for Zn, 6 for As, 6 for Cd, and 4 for Cu(topsoil), whereas in the subsoil, PTEs exhibited the following 39(Mn), 48(Ni), 29(Pb), 31(Zn), 12 (As), 1 (Cr), 1 (Cd), and 4 (Cu). PTEs showed a moderate enrichment level accordingly. Only Cd exhibited a significant enrichment level for both soils, with 45 and 50 sampled areas being significantly enriched for the topsoil and the subsoil, respectively. According to Zhang and Liu (2002), EF values of $0.5 \leq EF \leq 1.5$ indicate that PTE concentrations can occur totally from natural weathering processes. However, if the EF values are above 1.5, a large portion of PTEs has been delivered from noncrustal materials as well as a divergent source, such as point and nonpoint emissions and biota (Sautherland et al., 2000; Zhang & Liu, 2002). It was evident that most of the EFs calculated from the sampled areas were above 1.5, signifying that anthropogenic activities played a crucial role aside from the natural source. However, the movement of PTEs that accumulated in the topsoil to subsoil explains the differences in the enrichment of the subsoil compared to topsoil. Despite this, the disparity in enrichment level may be attributed to the transfer of PTEs from the topsoil to the subsoil; however, it may also be due to the high level of PTEs related to parental materials with potential uplift due to leaching.

Multivariate analysis of PTEs

Pearson correlation matrix (PCM)

The correlation matrix (Table 2) demonstrates that there is a relationship between the PTEs under analysis in both soils. The PTEs PbCd, ZnCd, and AsCd showed a high degree of connection in the topsoil, with r values ranging between 0.7 and 0.81. It appears that the subsoil showed no high correlation but rather a moderate correlation between ZnPb, AsPb, NiCr, ZnCd, and AsCd, with r values varying between 0.52 and 0.61. Similarly, the relationship between ZnAs, AsPb, NiZn, NiCd, and NiCu in the topsoil also showed moderate connections. Generally, the correlation in the topsoil is stronger than that in the subsoil. This indicates that the PTEs from both soil levels may share a closely related source.

Table 2 Correlation matrix between the PTEs

Topsoil	Mn	Ni	Pb	Zn	As	Cd	Cr	Cu
Mn	1.00							
Ni	0.36	1.00						
Pb	0.45	0.31	1.00					
Zn	0.32	**0.59**	**0.81**	1.00				
As	−0.08	0.06	**0.60**	**0.52**	1.00			
Cd	0.20	**0.53**	**0.73**	**0.79**	**0.70**	1.00		
Cr	0.46	0.40	0.00	0.10	−0.33	0.03	1.00	
Cu	0.47	**0.65**	0.27	0.38	−0.02	0.21	0.25	1.00
Subsoil								
Mn	1.00							
Ni	0.36	1.00						
Pb	0.28	0.00	1.00					
Zn	0.11	0.39	**0.60**	1.00				
As	−0.08	−0.11	**0.52**	0.44	1.00			
Cd	0.08	0.33	0.48	**0.58**	**0.61**	1.00		
Cr	0.35	**0.58**	−0.14	0.18	−0.16	0.09	1.00	
Cu	0.30	0.49	0.25	0.48	−0.09	0.02	0.24	1.00

Bold indicates the optimal performance of some PTE interactions in the top and subsoil, with moderate and strong correlations

Table 3 Correlations between PTEs and principal components

Elements	PC1/Topsoil	PC2/Topsoil	PC 1/Subsoil	PC 2/Subsoil
Mn	0.52	**0.55**	0.41	**0.45**
Ni	**0.71**	0.43	0.59	**0.63**
Pb	**0.85**	−0.27	**0.69**	−0.44
Zn	**0.91**	−0.15	**0.86**	−0.13
As	0.55	−0.72	**0.51**	−0.70
Cd	**0.85**	−0.34	**0.73**	−0.37
Cr	0.22	**0.75**	0.35	**0.68**
Cu	**0.56**	0.52	**0.54**	0.45
Eigenvalue	**3.71**	**2.04**	**2.93**	**2.09**
% Variance explained	**46.40**	**25.49**	**36.65**	**26.12**
Cumulative % Total		**71.89**		**62.76**

Bold indicates the optimal performance of some PTEs' relationships with each in the projected principal component, revealing moderate and high r values per principal component of each soil level

Identification of sources based on PCA

The PCA results are displayed in Table 3 and projected in Fig. 5a. Hou et al. (2013) reported that PCA is a useful tool that can provide informative suggestions on pathways for PTEs and primary sources. The characteristics of the extracted principal components (PCs) selected have eigenvalues all equal to or greater than 1. Based on the criteria, PC 1 and PC 2 were found to be statistically significant, accounting for 71.89% and 62.76% of the data variations for topsoil and subsoil, respectively. The groupings of the PTEs in the projection of components 1 and 2 from

both soils suggest that Pb, Zn, Cd, and As are polluting PTEs and that other PTEs Mn, Cu, Ni, and Cr are more geogenic elements. This is consistent with Borůvka et al., (2005) report stating that the positive correlation between this group of elements and their place within the primary component projection denotes their origin. Even though some of the PTEs are more geogenic, anthropogenic factors augment their enrichment in both soils. The topsoil and the subsoil accounted for the principal component (PC 1) 46.40 and 36.65%, respectively.

The principal component values for the PTEs are in the order of $Zn < Cd < Pb < Ni < Cu < As$ in the
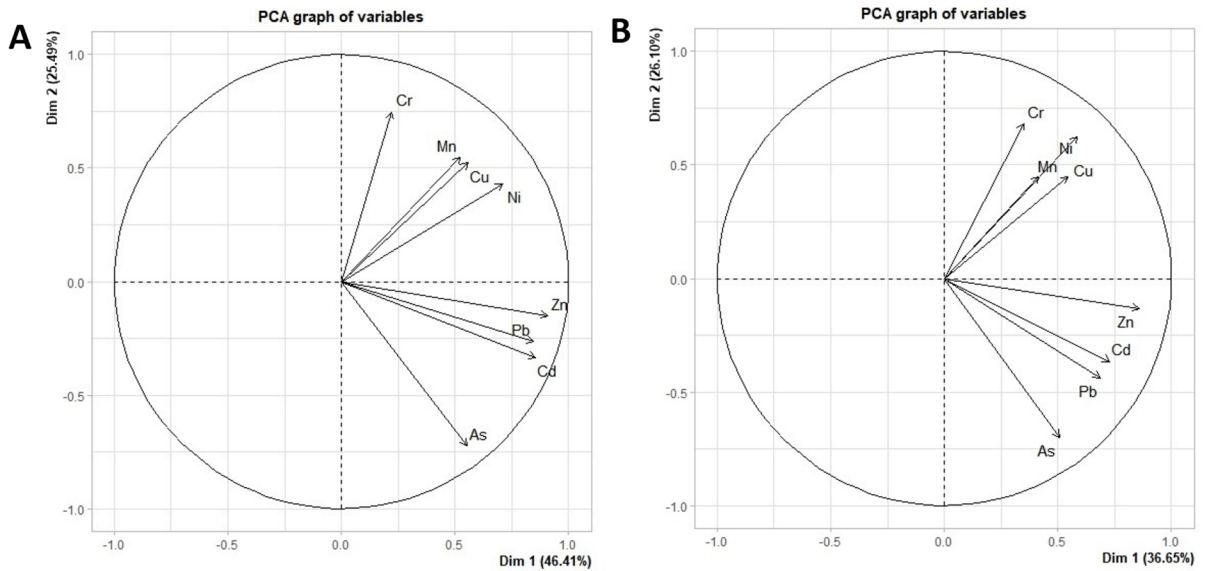
(a) projection of principal components 1 and 2 for the topsoil (A) and the subsoil (B)
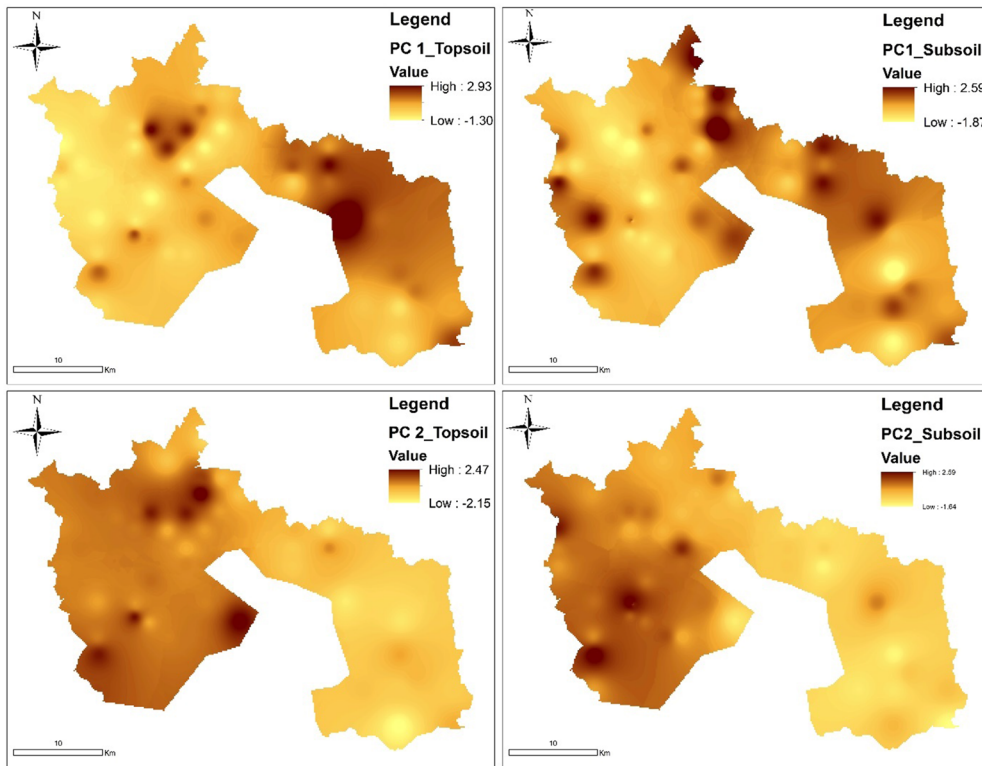
(b) Spatial distribution of the principal components (PCA _IDW) estimated values.

Fig. 5 a A projection of principal components 1 and 2 for the topsoil (A) and the subsoil (B). **b** Spatial distribution of the principal components (PCA _IDW) estimated values

topsoil, and those for the subsoil are in the order of Zn < Cd < Pb < Cu < As. The following PTEs, Ni, Pb, Zn, and Cd, exhibited high correlations ranging from 0.7 to 0.91 in the topsoil, while Cd and Zn also displayed strong positive loads (Table 3). In PC1, the control sources are more anthropogenic than geogenic. Consequently, this suggests anthropogenic pollution arising from farming practices, industrial activities, atmospheric deposition, and soil manure/fertilizer application. This is consistent with the claim of Chen et al. (2015). The topsoil revealed a hotspot around the north-eastern part of the map (PC1) that may be linked to industrial activities (steel industry). The spatial distribution of the PC-IDW (principal component of inverse distance weighting) map for the PC1 subsoil indicates the hotspot on the north-western part of the map, which may be attributed to both geogenic and anthropogenic sources (Fig. 5b). Virtually all the north-eastern to south-eastern parts of the map (subsoil PC 1) showed a moderate spatial distribution of PTEs. Subsoil enrichment is not limited to the geogenic source but rather to the mobility and leaching of PTEs from the topsoil to the subsoil. This is consistent with similar results captured by Borůvka et al. (2005). In PC 2, 25.49% and 26.12% of the total variance were explained, and their positive loads were satisfactory, with $r = 0.55$ (Mn) and 0.75 (Cr) for topsoil and 0.63 (Ni), 0.68 (Cr), and 0.45 (Mn) for subsoil. The PTE positive loads in PC2 are more geogenically inclined. The PC2 spatial distribution map shows more hotspot and moderate PTE distributions in the topsoil than in the subsoil. This is the consequence of anthropogenic activities that occur on topsoil. The north–west and south–western portions of the PC2 map (topsoil) are mainly agrarian vicinities, which use many agro-related products coupled with the geogenic source, which likely accounted for the high spatial distribution level of PTEs in that region.

Spatial distribution of PTEs

Adimalla et al. (2019) reported that spatial distribution maps play a crucial role in defining safe and hazardous areas as well as providing the basic details required to avoid and monitor further soil pollution. The spatial maps of both soil levels of As, Zn, Pb, Cr, Ni, Cu, Mn, and Cd are shown in Figs. 6 and 7. The concentration of As in both soils seen in Fig. 6

indicates a high concentration of As in both soils. Hotspots of As seen in the north-east and south-east parts of the map, but in the topsoil, more hotspots can be realized in the south-east part, which may be attributed to agricultural activities. Hu and Cheng, (2013) and Liang et al., (2015) indicated that some phosphate fertilizers contain As and are used as essential pesticide ingredients that tend to increase As in agricultural soils. The spatial distribution map of Cd shown in Fig. 6 exhibits more Cd pollution content in the subsoil than in the topsoil. The hotspots of Cd can be seen in the north-eastern and central parts of the south-eastern part of the map. The high content of Cd in the soil can be attributed to the multiplicity of sources, such as the parent materials, metal works, and steel industry waste discharges that have leached into the subsoil. Chai et al., (2015) and Sun et al., (2011) confirmed similar cases in agricultural soil in which the reasonably high concentration of Cd is due to the steel and smelting industries, and it has been shown that the waste discharged from the steel and smelting industries ultimately contributes to the accumulation of Cd in the surrounding soil. The topsoil shared similar hotspots pattern in the north–western part of the map with the subsoil (Fig. 6). The north-west part of the study area is known predominantly for intensive agricultural activities. The hotspot of Cr seems to be more geogenic with support from agro-related activities (agrochemicals) in both soils. The variable copper distribution in the topsoil is more pronounced than that in the subsoil. The hotspots displayed in the south–eastern part, the central part of the north–west, and the south-western part of the topsoil Cu spatial distribution map are more of geogenic origin with a boost from vehicular traffic (Fig. 6). Zhao et al., (2015) findings corroborate the present research results that vehicle exhaust deposits large contents of Cu on the topsoil. Both soils displayed varied spatial variability in Mn concentrations; nevertheless, the distribution and the hotspots in the subsoil were denser than those in the topsoil in the north–west parts of the map (Fig. 7). The distribution of Mn may be ascribed to the parent material. Nickel showed a hotspot in the north-western part of the map in the topsoil, whereas in the subsoil, it showed a hotspot in the south-western part of the map. In addition, both soils showed a sparse distribution of Mn on the map (Fig. 7). The distribution of Mn in both soils is more of a geogenic source. The distribution of
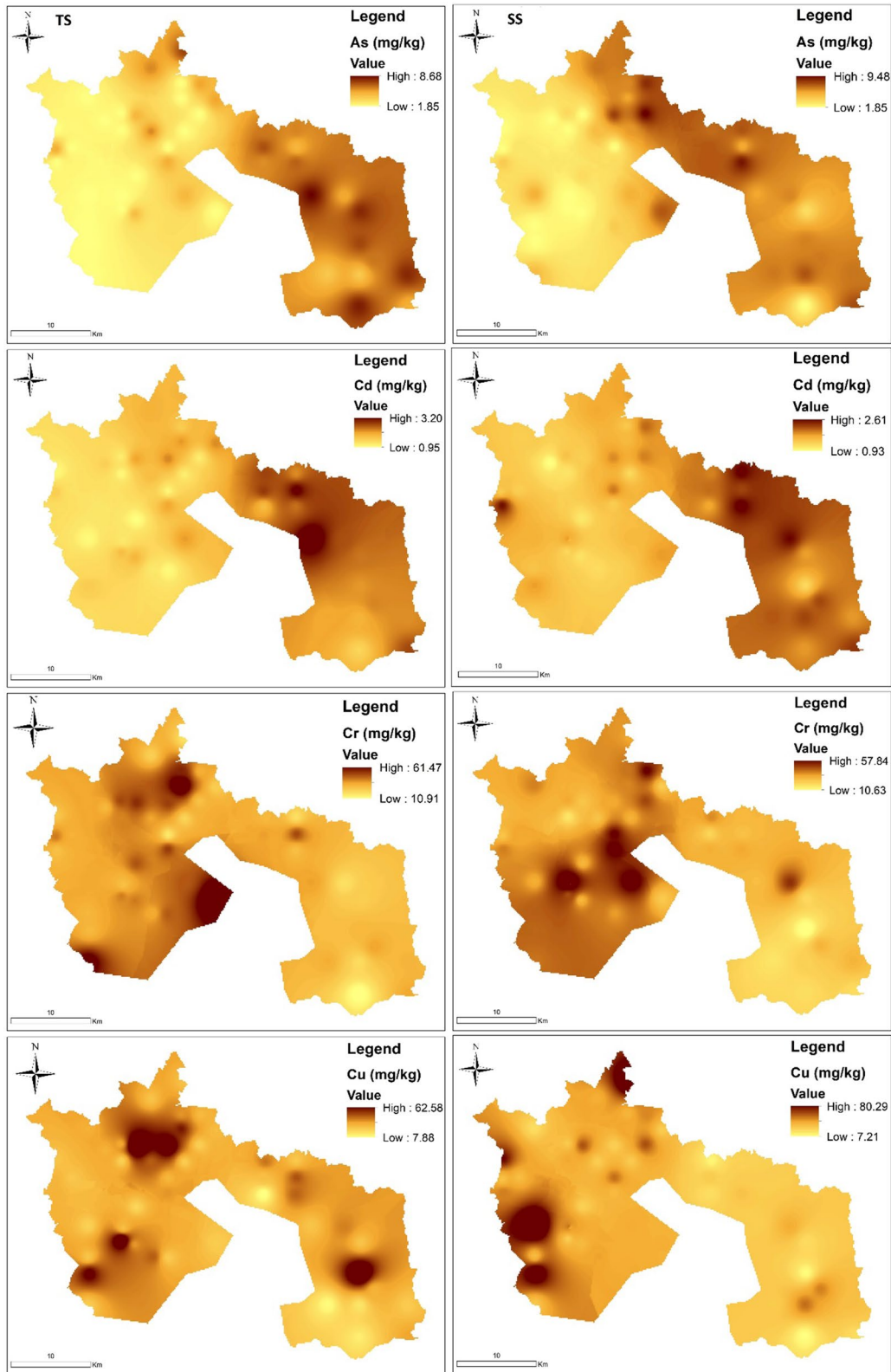
◄**Fig. 6** Spatial distribution of potentially toxic elements using IDW (TS—topsoil & SS—subsoil)

Pb concentration in the subsoil in the north–western part is more pronounced than that in the topsoil, and Pb is denser in the south-eastern part of the map in the topsoil than in the subsoil. The pollution in the north–western part of the subsoil is more of agro-related pollution as a result of mobility and leaching. Nevertheless, the pollution in the north-west part of the topsoil may be attributed to the steel industry and the metalwork within that vicinity. Pb enrichment in the subsoil may be attributed to the leaching of Pb-based pesticides and fertilizers (e.g. lead arsenate) used in agricultural land. Atafar et al., (2010) argues that fertilizer and pesticide application increases the concentration of Pb in agricultural soil. Zinc distribution varied in both soils (Fig. 7), showing hotspots in the north-western part (subsoil) and central part of the eastern part (topsoil) of the map. Zinc enrichment might be attributed to the nonferric metal industry and agricultural practices. Kabata-Pendias, (2011) affirm that Zn's anthropogenic sources relate to the metal industry, steel industry, and agriculture.

Source analysis using EF-PMF and PMF receptor models

EPA PMF software (version. 5.0) was used for the PMF and ER-PMF (enrichment factor positive matrix factorization model) receptor modelling analysis. The computed EFs and raw data of the respective PTEs, as well as the data for uncertainty, were used as input data for both receptor models. We considered the optimal number of variables in both receptor model studies and followed the set of guidelines. A gradual decline in the Q/Qexp index and a tendency to stabilize the Q value were chosen. The value of freak does not seem to boost sources. The results of both receptor model analyses are illustrated in Table 4. For a PTE to dominate a factor, it must attain 40% or more percentage apportionment in a factor.

The first factor was dominated predominantly by As (77.8%), Cd (50.6%), and Zn (43.7%) in the subsoil for the EF-PMF receptor model, while in the PMF receptor model, it was controlled by As (75.7%), Cd (49.3%), Pb (48.8%), and Zn (42.8%). This pattern of dominance by the PTEs in the subsoil for both receptor models suggests the influence

of anthropogenic activities. This result is consistent with the PCA projection loadings in PC 1 for the subsoil (see Fig. 5). The first four PTEs that dominated the subsoil of Factor 1 (Pb, Zn, Cd, As) suggest more anthropogenic sources of pollution (such as agro-chemicals and industrial activities) through leaching. Previous studies found elevated levels of Pb, Zn, Cd, and As above their tolerable soil limits from a range of activities, such as metalworking, the steel industry, industrial waste, and agriculture (Atafar et al., 2010; Belon et al., 2012; Sun et al., 2013; Tang et al., 2017). The spatial distribution maps (Figs. 6 and 7) of Pb, Zn, Cd, and As show the enrichment of these PTEs in the subsoil compared to the topsoil. Even though anthropogenic activities are directly incident on topsoil, the migration of PTEs from the topsoil to the subsoil due to soil-related properties (chemical and physical) plays a major role in the enrichment of the subsoil. Moreover, activities such as leaching may be responsible for these findings. Arsenic dominance and enrichment in the subsoil is a result of the successive application of fertilizer to the soil every crop season to ensure a bumper harvest that leaches from the topsoil to the subsoil as a result of rainfall, excessive irrigation, organic matter content, and soil texture. Soil physical and chemical factors may considerably alter the migration from the topsoil to the subsoil, such as Cu and Cd (Slavich et al., 2013). Studies carried out by Duan et al., (2015) have shown that Pb, Ni, and Cu increased in the subsoil compared to the topsoil. This is because the PTEs are leached to the subsoil from the topsoil (Shiva Kumar & Srikantaswamy, 2014). The topsoil was controlled by Cu (61.6%) and Ni (59.5%) for the EF-PMF receptor model, and the PMF model was dominated by Cu (71.5%), Ni (40.3%), Pb (40.0%), and Zn (42.7). The EF-PMF dominance PTEs (Ni, Cu) in the topsoil based on the source identification by PCA suggest that Ni and Cu are more of geogenic origin but with an anthropogenic source augmenting it. In contrast, the prevalence of PMF model PTEs (Cu, Ni, Pb, Zn) indicates that topsoil factor 1 is a composite of anthropogenic and geogenic sources. Topsoil enrichment based on the controlling PTEs, especially Ni and Cu, in both models may be inclined to a geogenic source with assistance from agro-related activities (agrochemicals). Previous studies by Liang et al., (2017) and Yang et al., (2017) have discovered that the use of fertilizers and pesticides in agricultural fields elevates
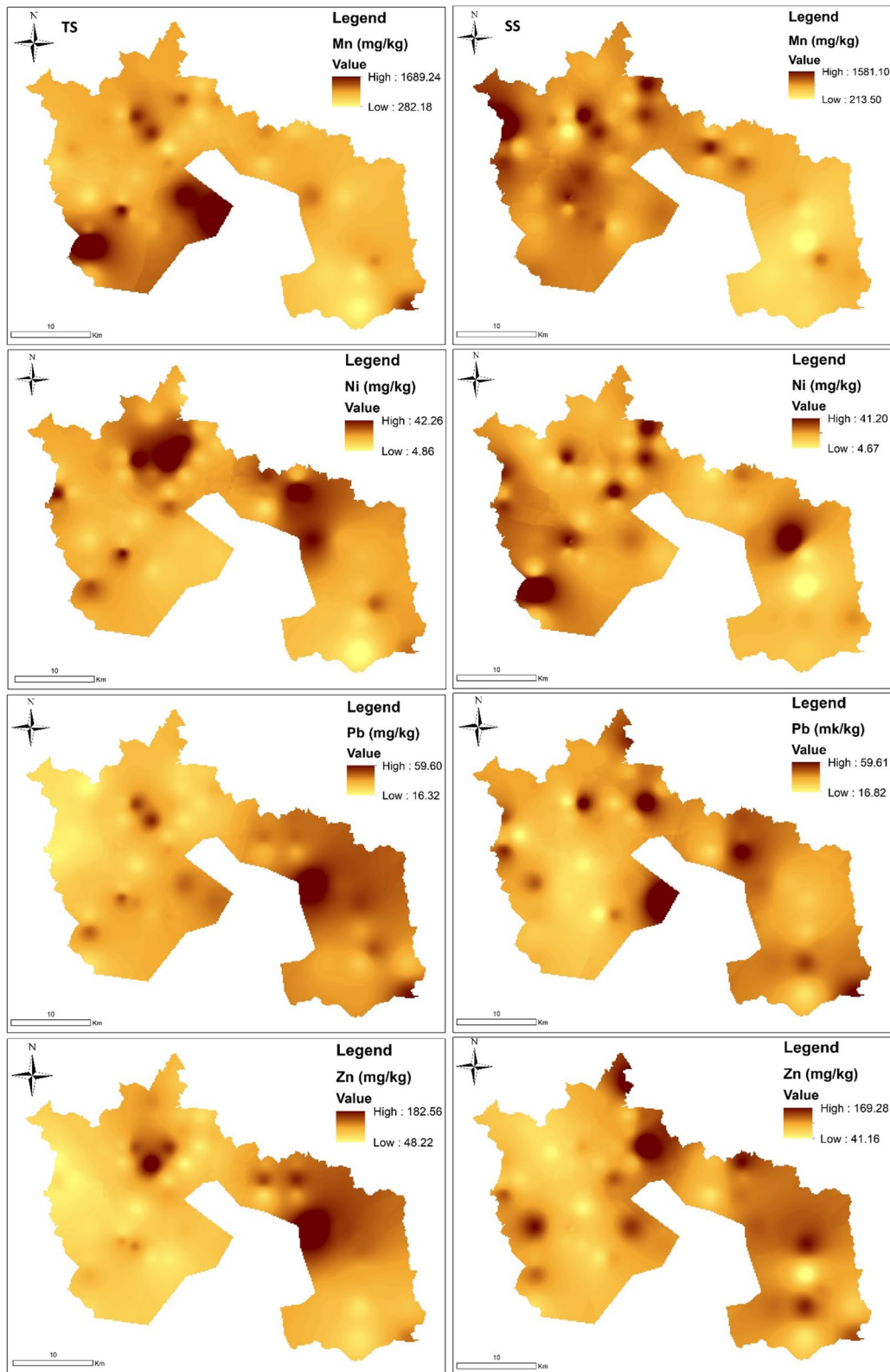
the normal content of copper in the soil. Huang et al., (2019) reported that nickel enrichment mostly in the surface of the soil is more of a natural source.

Factor 2 was dominated by Mn (56%) and Pb (41.5%) for the EF-PMF receptor model (subsoil) and Cr (56.2%) and Mn (57.5%) in the PMF receptor model (subsoil). However, in the topsoil, As (71.7%), Cd (42.3%), and Pb (40.6%) were included in the EF-PMF receptor model, and As (88.8%), Cd (52.1%), Pb (54.3%), and Zn (43.1%) were included in the PMF receptor model. The dominant PTEs (particularly Cr and Mn) in both receptor models in the subsoil share a similar pathway, as revealed by PCA, which is more of a geogenic source supplemented by anthropogenic activities such as the steel industry. Chromium is an ideal element for the formation of alloys in the steel industry, and it is abundant in topsoil in most study areas, indicating the existence of industrial activities (Yeung et al., 2003) and agricultural practices. The leachability of Cr from the topsoil to the subsoil due to rainfall augments the concentration level in the subsoil. Previous studies have shown that Cr is a significant source of elements discharged by a variety of industrial activities that pollute soils through waste disposal (Choppala et al., 2013; Guan et al., 2018; Pan et al., 2016). Zhang et al., (2016) performed a comprehensive study collecting numerous articles in China (464), and the findings in the papers collected indicated that the average amount of Cr in the agricultural soil was 78.94 mg/kg, which surpasses the geochemical background level (57.30 mg/kg). Further studies by Li et al., (2009) and Liu et al., (2015) have indicated that the use of sewage as a source of irrigation is likely to increase the Cr content of agricultural soil. The elevation of Cr in the topsoil can come from a multiplicity of sources, but in the current study area, Cr elevation will predominantly be attributed to steel plant and agricultural activities.

According to Foo et al., (2008) and Yang et al., (2011), manganese (Mn) is widely regarded as the most abundant lithospheric element and as a major soil element. The excessive manganese in topsoil might be attributed to the anthropogenic source supplementing the geogenic source. Various soils in Mn are enriched by anthropogenic inputs that may influence the function of the environment (Herndon et al., 2011).

The final factor (factor 3) was primarily dominated by Cr (59.9%), Cu (62.4%), Mn (44%), and Ni (65.2%) in the EF-PMF receptor model and Cu (59.4%) and Ni (57.5%) in the subsoil. The topsoil was eclipsed by Cr (74.7%), Mn (61.8%), and Pb (41.1%) for EF-PMF and for the PMF receptor model; it was largely influenced by Cr (48.4%) and Mn (48%). The consistency in both models in source apportioning the dominant PTEs in each factor is inline. Both models apportioned Cr and Mn as dominant PTEs in factor 3 for the subsoil, but EF-PMF goes a step further to include Mn and Cu. The dominant PTEs in the subsoils aligning with the PCA projection loading suggest that the dominant PTEs are more of geogenic origin. The source distribution in the topsoil by both models also exhibits a blend of both anthropogenic and geogenic sources. The spatial distribution of Pb in both topsoil and subsoil indicates anthropogenic-induced Pb concentrations in the soil. Research conducted by Zhao et al., (2018) indicated that PTE (Cd, Cu, Pb) intake from chemical fertilizers and manure increased by 3–4 percent yearly. Wang et al.,(2006) stated that anthropogenic activities, such as metal smelting, pollute the environment with PTEs such as Pb, but rain leaching carries the PTEs from the soil surface to the subsurface of the soil.

Performance of EF-PMF and PMF

The source analysis suggested that both models apportioned sources efficiently and showed a generally consistent trend of the same PTEs selected in the same factors. Nevertheless, based on the fixed criteria chosen by this paper in selecting the appropriate PTE to have a controlled factor (obtaining 40% or more), the traditional PMF apportioned more PTEs than the novel receptor model EF-PMF. Both models were subjected to multiple linear regression analysis based on the source apportionment accuracy, marginal error, and percentage efficiency. An intracomparison is performed by calculating the regression analysis between the models to see whether the model fits well (Song et al., 2008; Stanimirova et al., 2011) or which model optimizes efficiency and minimizes error. The estimated validation and efficiency assessment using R^2, RMSE, and MAE suggested that the novel receptor performed better than the parent model

Table 4 Percentage contribution of each factor to PTEs derived from the ER-PMF receptor model

PTEs	F1%	F2%	F3%	F1%	F2%	F3%	F1%	F2%	F3%	F1%	F2%	F3%
	EF-PMF Subsoil			PMF Subsoil			EF-PMF Topsoil			PMF Topsoil		
As	77.8	22.2	–	75.7	13.8	10.4	10.1	71.7	18.2	8.6	88.8	2.6
Cd	50.6	16.5	32.9	49.3	23.7	27	27.1	42.3	30.5	24.7	52.1	23.2
Cr	33.5	6.6	59.9	24.9	56.2	18.9	17.1	8.2	74.7	28.5	23.4	48.4
Cu	24.4	13.1	62.4	20.9	19.6	59.4	61.6	14.8	23.7	71.5	19.2	9.3
Mn	–	56	44	28.5	46.8	24.7	23.7	14.5	61.8	24.5	27.5	48
Ni	34.7	0.1	65.2	23.4	19.1	57.5	59.5	17.4	23.1	40.3	22.4	37.3
Pb	34.5	41.5	24	48.8	31.2	19	18.3	40.6	41.1	40	54.3	5.7
Zn	43.7	16.4	39.8	42.8	23.6	33.7	34.1	35.7	30.1	42.7	43.1	14.2

PMF. According to Singh et al., (2017), to determine the performance of diverse modelling techniques, RMSE, R^2, and MAE are most often employed. The coefficient of determination (R^2) suggested that both model source apportionment efficiencies were consistent with each other. However, the root-mean-square error (RMSE) computed for both models and soil levels suggested that out of the 8 PTEs (As, Cd, Cr, Cu, Ni, Mn, Pb, and Zn) assessed, 7 (As, Cr, Cu, Ni, Mn, Pb, and Zn) from the EF-PMF receptor model showed a significant reduction in errors compared with the PMF. Similarly, the mean absolute error (MAE) computation also established the same results as RMSE when both receptor models were compared. Thus, this result implies that the novel EF-PMF model can apportion sources to PTEs with minimal errors compared to PMF. The hybridization of the enrichment factor to the parent receptor model PMF has exhibited its applicability and efficiency in source analysis of PTEs in agricultural soil. The pollution index enrichment factor has strengthened the base model PMF, thereby minimizing errors in source apportionment in this study. A similar comparative analysis was performed by Gholizadeh et al. (2016), who concluded that the R^2 values of the observed/predicted plots for the majority of the water quality variables demonstrated superior goodness-of-fit with the APCS-MLR receptor modelling approach to pollution source apportionment than PMF. Contrary to the concluding remarks of Gholizadeh et al. (2016), Zhang et al., (2019) compared three receptor models (PMF, PCA-MLR, and Unmix), and based on the R^2 values obtained, the authors established that PMF demonstrated superior goodness-of-fit in their studies than the other receptor models. Similarly, Larsen and

Baker, (2003) and Yang et al., (2013) likewise upheld the results by using model accuracy and validation criteria to suggest that PMF was the finest model to expound PAHs in Huanghuai Plain soils.(Table 5).

Conclusion

This study investigates the source contribution of PTEs using a novel EF-PMF receptor model approach against a parent receptor model (PMF) and further evaluates the pollution characteristics and determines the spatial variability of PTE in agricultural soil at different depths. PMF has been used in various papers quantifying the source apportionment in sediments, determining air quality assessment (an air pollutant) and equally applied in the soil to determine the source contribution at different levels.

The study ascertained that both soils (topsoil and subsoil) Zn and Cd mean concentrations values of this present study are higher than all 3 geochemical background levels presented in Table 1. Similarly, the study suggested that Mn and Cu were also higher than the respective PTEs in EAV; likewise, Pd, mean values were higher than the same PTE in the WAV level. As, Cu, Mn, and Pb average values from the UCC were found to be lower than the mean values of the same PTEs in this study. The pollution indices, such as PI, EF, ER, hinted that some of the computed PTE concentration values fell within the low to moderate pollution/risk levels. Only the concentrations of Mn and Cd exhibited high pollution or considerable risk of significant pollution levels. The calculated Nemerow pollution index, pollution load index, and risk index revealed varied pollution

Table 5 Comparison of the EF-PMF and PMF receptor models via multiple linear regression

	As			Cd			Cr			Cu		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
EF-PMF TOPSOIL	0.877	0.214	0.162	0.810	0.986	0.746	0.705	0.147	0.104	0.681	0.354	0.210
PMF TOP-SOIL	0.877	0.773	0.585	0.810	0.213	0.161	0.705	4.617	3.258	0.681	7.274	4.330
EF-PMF SUB-SOIL	0.918	0.199	0.157	0.663	1.132	0.941	0.691	0.126	0.091	0.871	0.190	0.137
PMF SUB-SOIL	0.917	0.720	0.566	0.665	0.244	0.203	0.690	3.956	2.858	0.871	3.914	2.821

	Mn			Ni			Pb			Zn		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
EF-PMF TOPSOIL	0.526	0.701	0.412	0.855	0.200	0.145	0.762	0.312	0.245	0.765	0.381	0.283
PMF TOP-SOIL	0.526	180.588	106.125	0.855	3.062	2.222	0.762	4.454	3.504	0.764	14.099	10.481
EF-PMF SUB-SOIL	0.494	0.567	0.414	0.608	0.286	0.208	0.659	0.342	0.264	0.716	0.378	0.269
PMF SUB-SOIL	0.494	146.113	106.846	0.608	4.369	3.178	0.659	4.880	3.773	0.715	14.010	9.951

levels ranging from low to moderate, with a few samples, indicating a high level of ecological risk in both soils. The PCM stated that there was a relationship between some of the metals, and the PCA identified the sources into two categories, namely geogenic (Mn, Cu, Ni, Cr) and anthropogenic (Pb, Zn, Cd, and As). The spatial distribution revealed hotspots of all the PTEs studied, showing that the pollution level in the subsoil is considerably higher than that in the topsoil. It further revealed a multiplicity of sources, such as agricultural activities, industrial activities, industrial waste, and geogenic sources, accounting for the pollution level in both soils. However, leaching and mobility of PTEs from the topsoil to the subsoil accounted for the higher level of PTE enrichment in the subsoil than in the topsoil. The receptor models EF-PMF and PMF showed a consistent source apportionment for PTEs in each factor. The receptor models suggested that geogenic, steel industry, agrochemicals, fertilizer application, metal works, and sewage irrigation were the prominent sources from which each PTE was inclined to be polluted. PTE pollution sources included a combination of anthropogenic sources such as chromium (agricultural activities, the steel industry, and industrial sewage irrigation). For instance, Cu and Ni were both dominant in both soils, which is factor 3 in the subsoil, and factor 1 in topsoil, which is also probably attributed to a geogenic source with a boost from agricultural activities such as fertilizers and pesticides. Mn, on the other hand, based on its richness in the lithosphere, is of more natural origin in both soils. The assessment of the receptor models using the validation and accuracy assessment criteria (RMSE, MAE, and R^2) via multiple linear regression suggested that the novel receptor model EF-PMF performed better than the parent model PMF. The EF-PMF receptor model was able to minimize the error ascertained in source apportionment substantially. Out of the 8 (As, Cr, Cd, Cu, Ni, Mn, Pb, and Zn) PTEs assessed from each depth of soil sampled; 7 (As, Cr, Cu, Ni, Mn, Pb and Zn) performed better in the EF-PMF receptor model than PMF. Hybridization of pollution indices such as the enrichment factor with the PMF receptor model in source apportionment minimizes error and optimizes efficiency. The outcomes of this study suggest that the EF-PMF receptor model be used in all polluted soils and environments. Furthermore, this study demonstrated the dire situation of soil health and quality in the study area and provided the necessary information for stakeholders to take proactive efforts to reduce increasing PTE levels in agricultural soils.

Author contributions Prince Chapman Agyeman prepared the manuscript and collected the data. Kingsley JOHN helped in visualization and proofreading. Ndiye Michael Kebonye contributed to data collection and proofreading. Luboš Borůvka: Supervision, Editing. Radim Vašát: Data Curation and Visualization.

Data availability The data are available upon request.

Declarations

Conflict of interests The authors declare that they have no known competing personal interests or relationships that could have appeared to influence the scientific work in this manuscript.

Animal research Not applicable.

Consent to participate Not applicable.

Consent to publish Not applicable.

References

Adimalla, N., Qian, H., & Wang, H. (2019). Assessment of heavy metal (HM) contamination in agricultural soil lands in northern Telangana, India: An approach of spatial distribution and multivariate statistical analysis. *Environmental Monitoring and Assessment*. https://doi.org/10.1007/s10661-019-7408-1

Adimalla, N., Chen, J., & Qian, H. (2020a). Spatial characteristics of heavy metal contamination and potential human health risk assessment of urban soils: A case study from an urban region of South India. *Ecotoxicology and Environmental Safety, 194*, 110406. https://doi.org/10.1016/j.ecoenv.2020.110406

Adimalla, N., Qian, H., Nandan, M. J., & Hursthouse, A. S. (2020b). Potentially toxic elements (PTEs) pollution in surface soils in a typical urban region of south India: An application of health risk assessment and distribution pattern. *Ecotoxicology and Environmental Safety, 203*, 111055. https://doi.org/10.1016/j.ecoenv.2020.111055

Agyeman, P. C., Ahado, S. K., Borůvka, L., Biney, J. K. M., Sarkodie, V. Y. O., Kebonye, N. M., & Kingsley, J. (2020). Trend analysis of global usage of digital soil mapping models in the prediction of potentially toxic elements in soil/sediments: A bibliometric review. *Environmental Geochemistry and Health*. https://doi.org/10.1007/s10653-020-00742-9

Antoniadis, V., & Alloway, B. J. (2002). Leaching of cadmium, nickel, and zinc down the profile of sewage sludge-treated soil. *Communications in Soil Science and Plant Analysis, 33*, 273–286. https://doi.org/10.1081/CSS-120002393

Atafar, Z., Mesdaghinia, A., Nouri, J., Homaee, M., Yunesian, M., Ahmadimoghaddam, M., & Mahvi, A. H. (2010). Effect of fertilizer application on soil heavy metal concentration. *Environmental Monitoring and Assessment, 160*, 83–89. https://doi.org/10.1007/s10661-008-0659-x

Belon, E., Boisson, M., Deportes, I. Z., Eglin, T. K., Feix, I., Bispo, A. O., Galsomies, L., Leblond, S., & Guellier, C. R. (2012). An inventory of trace elements inputs to French agricultural soils. *Science of the Total Environment, 439*, 87–95. https://doi.org/10.1016/j.scitotenv.2012.09.011

Borůvka, L., Vacek, O., & Jehlička, J. (2005). Principal component analysis as a tool to indicate the origin of potentially toxic elements in soils. *Geoderma, 128*, 289–300. https://doi.org/10.1016/j.geoderma.2005.04.010

Boyter, M. J., Brummer, J. E., & Leininger, W. C. (2009). Growth and metal accumulation of geyer and mountain willow grown in topsoil versus amended mine tailings. *Water, Air, and Soil Pollution, 198*, 17–29. https://doi.org/10.1007/s11270-008-9822-9

Chai, Y., Guo, J., Chai, S., Cai, J., Xue, L., & Zhang, Q. (2015). Source identification of eight heavy metals in grassland soils by multivariate analysis from the Baicheng-Songyuan area, Jilin Province, Northeast China. *Chemosphere, 134*, 67–75. https://doi.org/10.1016/j.chemosphere.2015.04.008

Chandrasekaran, A., Ravisankar, R., Harikrishnan, N., Satapathy, K. K., Prasad, M. V. R., & Kanagasabapathy, K. V. (2015). Multivariate statistical analysis of heavy metal concentration in soils of Yelagiri Hills, Tamilnadu, India - Spectroscopical approach. *Spectrochimica Acta - Part a: Molecular and Biomolecular Spectroscopy, 137*, 589–600. https://doi.org/10.1016/j.saa.2014.08.093

Chen, H., Teng, Y., Chen, R., Li, J., & Wang, J. (2016). Contamination characteristics and source apportionment of trace metals in soils around Miyun Reservoir. *Environmental Science and Pollution Research, 23*, 15331–15342. https://doi.org/10.1007/s11356-016-6694-1

Chen, H., Teng, Y., Lu, S., Wang, Y., & Wang, J. (2015). Contamination features and health risk of soil heavy metals in China. *Science of the Total Environment, 512–513*, 143–153. https://doi.org/10.1016/j.scitotenv.2015.01.025

Chen, X., & Lu, X. (2018). Contamination characteristics and source apportionment of heavy metals in topsoil from an area in Xi'an city, China. *Ecotoxicology and Environmental Safety, 151*, 153–160. https://doi.org/10.1016/j.ecoenv.2018.01.010

Choppala, G., Bolan, N., & Park, J. H. (2013). Chromium contamination and Its risk management in complex environmental settings. *Advances in agronomy* (pp. 129–172). Elsevier. https://doi.org/10.1016/B978-0-12-407686-0.00002-6

Cools, N., & De Vos, B. (2016). Part X: Sampling and analysis of soil. Manual on methods and criteria for harmonized sampling, assessment, monitoring and analysis of the effects of air pollution on forests. Thünen Institute of Forest Ecosystems, Eberswalde, Germany, 29.

czso, (2019). Characteristics of the Frýdek-Místek district CZSO in Ostrava [WWW Document]. URL https://www.czso.cz/csu/xt/charakteristika_okresu_frydek_mistek (Accessed 10.6.20).

Duan, X., Zhang, G., Rong, L., Fang, H., He, D., & Feng, D. (2015). Spatial distribution and environmental factors of catchment-scale soil heavy metal contamination in the dry-hot valley of Upper Red River in Southwestern China. *CATENA, 135*, 59–69. https://doi.org/10.1016/j.catena.2015.07.006

Ergin, M., Saydam, C., Baştürk, Ö., Erdem, E., & Yörük, R. (1991). Heavy metal concentrations in surface sediments from the two coastal inlets (Golden Horn Estuary and İzmit Bay) of the northeastern Sea of Marmara. *Chemical Geology, 91*, 269–285. https://doi.org/10.1016/0009-2541(91)90004-B

Fei, X., Lou, Z., Xiao, R., Ren, Z., & Lv, X. (2020). Contamination assessment and source apportionment of heavy metals in agricultural soil through the synthesis of PMF and GeogDetector models. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2020.141293

Foo, T., Fong, S., Chee, A., Azani, M., & Tahir, N. M. (2008). Possible source and pattern distribution of heavy metals content in urban soil at kuala terengganu town center. *The Malaysian Journal of Analytical Sciences., 12*(2), 458–467.

Gąsiorek, M., Kowalska, J., Mazurek, R., & Pająk, M. (2017). Comprehensive assessment of heavy metal pollution in topsoil of historical urban park on an example of the planty park in Krakow (Poland). *Chemosphere, 179*, 148–158. https://doi.org/10.1016/j.chemosphere.2017.03.106

Gholizadeh, M. H., Melesse, A. M., & Reddi, L. (2016). Water quality assessment and apportionment of pollution sources using APCS-MLR and PMF receptor modeling techniques in three major rivers of South Florida. *Science of the Total Environment, 566*, 1552–1567.

Guan, Q., Wang, F., Xu, C., Pan, N., Lin, J., Zhao, R., & 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.

Hakanson, L. (1980). An ecological risk index for aquatic pollution control.a sedimentological approach. *Water Research, 14*, 975–1001. https://doi.org/10.1016/0043-1354(80)90143-8

Herndon, E. M., Jin, L., & Brantley, S. L. (2011). Soils reveal widespread manganese enrichment from industrial inputs. *Environmental Science and Technology, 45*, 241–247. https://doi.org/10.1021/es102001w

Hossain Bhuiyan, M. A., Chandra Karmaker, S., Bodrud-Doza, M., Rakib, M. A., & Saha, B. B. (2021). Enrichment, sources and ecological risk mapping of heavy metals in agricultural soils of dhaka district employing SOM. *PMF and GIS Methods. Chemosphere, 263*, 128339. https://doi.org/10.1016/j.chemosphere.2020.128339

Hou, D., He, J., Lü, C., Ren, L., Fan, Q., Wang, J., & Xie, Z. (2013). Distribution characteristics and potential ecological risk assessment of heavy metals (Cu, Pb, Zn, Cd) in water and sediments from Lake Dalinouer, China. *Ecotoxicology and Environmental Safety, 93*, 135–144. https://doi.org/10.1016/j.ecoenv.2013.03.012

Hu, J., Lin, B., Yuan, M., Lao, Z., Wu, K., Zeng, Y., Liang, Z., Li, H., Li, Y., Zhu, D., Liu, J., & Fan, H. (2019). Trace metal pollution and ecological risk assessment in agricultural soil in dexing Pb/Zn mining area, China. *Environmental Geochemistry and Health, 41*, 967–980. https://doi.org/10.1007/s10653-018-0193-x

Hu, Y., & Cheng, H. (2013). Application of stochastic models in identification and apportionment of heavy metal pollution sources in the surface soils of a large-scale region. *Environmental Science and Technology, 47*, 3752–3760. https://doi.org/10.1021/es304310k

Huang, H., Lin, C., Yu, R., Yan, Y., Hu, G., & Li, H. (2019). Contamination assessment, source apportionment and health risk assessment of heavy metals in paddy soils of Jiulong River Basin, Southeast China. *RSC Advances, 9*, 14736–14744. https://doi.org/10.1039/c9ra02333j

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. https://doi.org/10.1016/j.jenvman.2017.10.072

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. https://doi.org/10.1016/j.jhazmat.2015.07.041

Kabata-Pendias, A. (2010). *Trace elements in soils and plants*. CRC Press. https://doi.org/10.1201/b10158

Kooistra, L., Wanders, J., Epema, G. F., Leuven, R. S. E. W., Wehrens, R., & Buydens, L. M. C. (2003). The potential of field spectroscopy for the assessment of sediment properties in river floodplains. *Analytica Chimica Acta, 484*, 189–200. https://doi.org/10.1016/S0003-2670(03)00331-3

Kowalska, J. B., Mazurek, R., Gąsiorek, M., & Zaleski, T. (2018). Pollution indices as useful tools for the comprehensive evaluation of the degree of soil contamination–A review. *Environmental Geochemistry and Health, 40*(6), 2395–2420.

Kozák J., Němeček J., Borůvka L., Kodešová R., Janků J., Jacko J., & Hladík J. (2010). *Soil Atlas of the Czech Republic*. Prague: CULS.

Larsen, R. K., & Baker, J. E. (2003). Source apportionment of polycyclic aromatic hydrocarbons in the urban atmosphere: A comparison of three methods. *Environmental Science and Technology, 37*, 1873–1881. https://doi.org/10.1021/es0206184

Lee, D. H., Kim, J. H., Mendoza, J. A., Lee, C. H., & Kang, J. H. (2016). Characterization and source identification of pollutants in runoff from a mixed land use watershed using ordination analyses. *Environmental Science and Pollution Research, 23*, 9774–9790. https://doi.org/10.1007/s11356-016-6155-x

Li, J., He, M., Han, W., & Gu, Y. (2009). Analysis and assessment on heavy metal sources in the coastal soils developed from alluvial deposits using multivariate statistical methods. *Journal of Hazardous Materials, 164*, 976–981. https://doi.org/10.1016/j.jhazmat.2008.08.112

Li, L., Lu, J., Wang, S., Ma, Y., Wei, Q., Li, X., Cong, R., & Ren, T. (2016). Methods for estimating leaf nitrogen concentration of winter oilseed rape (Brassica napus L.) using in situ leaf spectroscopy. *Industrial Crops and Products, 91*, 194–204.

Li, Z., Ma, Z., van der Kuijp, T. J., Yuan, Z., & Huang, L. (2014). A review of soil heavy metal pollution from mines in China: Pollution and health risk assessment. *Science of the Total Environment*. https://doi.org/10.1016/j.scitotenv.2013.08.090

Liang, J., Feng, C., Zeng, G., Gao, X., Zhong, M., Li, X., Li, X., He, X., & Fang, Y. (2017). Spatial distribution and source identification of heavy metals in surface soils in a typical coal mine city, Lianyuan, China. *Environmental Pollution, 225*, 681–690. https://doi.org/10.1016/j.envpol.2017.03.057

Liang, J., Hua, S., Zeng, G., Yuan, Y., Lai, X., Li, X., Li, F., Wu, H., Huang, L., & Yu, X. (2015). Application of weight method based on canonical correspondence analysis for assessment of anatidae habitat suitability: A case study in east dongting lake, middle China. *Ecological Engineering, 77*, 119–126. https://doi.org/10.1016/j.ecoleng.2015.01.016

Liu, J., Liu, Y. J., Liu, Y., Liu, Z., & Zhang, A. N. (2018). Quantitative contributions of the major sources of heavy metals in soils to ecosystem and human health risks: A case study of Yulin, China. *Ecotoxicology and Environmental Safety, 164*, 261–269. https://doi.org/10.1016/j.ecoenv.2018.08.030

Liu, X., Qiujin Song, Y., Tang, W. L., Jianming, X., Jianjun, W., Wang, F., & Brookes, P. C. (2013). Human health risk assessment of heavy metals in soil–vegetable system: A multi-medium analysis. *Science of the Total Environment, 463–464*, 530–540. https://doi.org/10.1016/j.scitotenv.2013.06.064

Liu, Y., Wang, H., Li, X., & Li, J. (2015). Heavy metal contamination of agricultural soils in Taiyuan, China. *Pedosphere, 25*, 901–909. https://doi.org/10.1016/S1002-0160(15)30070-9

Liu, Y. R., He, Z. Y., Yang, Z. M., Sun, G. X., & He, J. Z. (2016). Variability of heavy metal content in soils of typical Tibetan grasslands. *RSC Advances, 6*, 105398–105405. https://doi.org/10.1039/c6ra23868h

Lü, J. S., & He, H. C. (2018). Identifying the origins and spatial distributions of heavy metals in the soils of the jiangsu coast. *Huanjing Kexue/environmental Science, 39*, 2853–2864. https://doi.org/10.13227/j.hjkx.201707219

Madrid, L., Díaz-Barrientos, E., & Madrid, F. (2002). Distribution of heavy metal contents of urban soils in parks of

Seville. *Chemosphere, 49*, 1301–1308. https://doi.org/10.1016/S0045-6535(02)00530-1

Mamat, Z., Yimit, H., Eziz, A., & Ablimit, A. (2014). Oasis land-use change and its effects on the eco-environment in Yanqi Basin, Xinjiang, China. *Environmental Monitoring and Assessment, 186*, 335–348. https://doi.org/10.1007/s10661-013-3377-y

Mazhari, S. A., Bajestani, A. R. M., Hatefi, F., Aliabadi, K., & Haghighi, F. (2018). Soil geochemistry as a tool for the origin investigation and environmental evaluation of urban parks in Mashhad city. *NE of Iran. Environmental Earth Sciences*. https://doi.org/10.1007/s12665-018-7684-z

Milićević, T., Relić, D., Škrivanj, S., Tešić, Ž., & Popović, A. (2017). Assessment of major and trace element bioavailability in vineyard soil applying different single extraction procedures and pseudototal digestion. *Chemosphere, 171*, 284–293. https://doi.org/10.1016/j.chemosphere.2016.12.090

Nezhad, M. T. K., Tabatabaii, S. M., & Gholami, A. (2015). Geochemical assessment of steel smelter-impacted urban soils, Ahvaz, Iran. *Journal of Geochemical Exploration, 152*, 91–109.

Norris, G., Duvall, R., Brown, S., Bai, S. (2014). EPA Positive Matrix Factorization (PMF) 5.0-Fundamentals and User Guide Prepared for the US Environmental Protection Agency Office of Research and Development, Washington.

Paatero, P. (1999). The multilinear engine—a table-driven, least squares program for solving multilinear problems, including the n-way parallel factor analysis model. *Journal of Computational and Graphical Statistics, 8*, 854–888. https://doi.org/10.1080/10618600.1999.10474853

Paatero, P., & Hopke, P. K. (2009). Rotational tools for factor analytic models. *Journal of Chemometrics, 23*, 91–100. https://doi.org/10.1002/cem.1197

Paatero, P., Hopke, P. K., Song, X. H., & Ramadan, Z. (2002). Understanding and controlling rotations in factor analytic models. *Chemometrics and Intelligent Laboratory Systems, 60*, 253–264. https://doi.org/10.1016/S0169-7439(01)00200-3

Pan, L. B., Ma, J., Wang, X. L., & Hou, H. (2016). Heavy metals in soils from a typical county in Shanxi Province, China: Levels, sources and spatial distribution. *Chemosphere, 148*, 248–254. https://doi.org/10.1016/j.chemosphere.2015.12.049

Parveen, N., Ghaffar, A., Shirazi, S. A., & Bhalli, M. N. (2012). A GIS based assessment of heavy metals contamination in surface soil of urban parks: a case study of Faisalabad City-Pakistan. *Journal of Geography & Natural Disasters, 02*, 1–5. https://doi.org/10.4172/2167-0587.1000105

Qingjie, G., Jun, D., Yunchuan, X., Qingfei, W., & Liqiang, Y. (2008). Calculating pollution indices by heavy metals in ecological geochemistry assessment and a case study in parks of Beijing. *Journal of China University of Geosciences, 19*(3), 230–241.

Rastegari Mehr, M., Keshavarzi, B., Moore, F., Sharifi, R., Lahijanzadeh, A., & Kermani, M. (2017). Distribution, source identification and health risk assessment of soil heavy metals in urban areas of Isfahan province. *Iran. Journal of African Earth Sciences, 132*, 16–26. https://doi.org/10.1016/j.jafrearsci.2017.04.026

Rodríguez Martín, J. A., Ramos-Miras, J. J., Boluda, R., & Gil, C. (2013). Spatial relations of heavy metals in arable and greenhouse soils of a Mediterranean environment region (Spain). *Geoderma, 200–201*, 180–188. https://doi.org/10.1016/j.geoderma.2013.02.014

Ruiz-Fernández, A. C., Sanchez-Cabeza, J. A., Pérez-Bernal, L. H., & Gracia, A. (2019). Spatial and temporal distribution of heavy metal concentrations and enrichment in the southern Gulf of Mexico. *Science of the Total Environment, 651*, 3174–3186. https://doi.org/10.1016/j.scitotenv.2018.10.109

Salim, I., Sajjad, R. U., Paule-Mercado, M. C., Memon, S. A., Lee, B. Y., Sukhbaatar, C., & Lee, C. H. (2019). Comparison of two receptor models PCA-MLR and PMF for source identification and apportionment of pollution carried by runoff from catchment and subwatershed areas with mixed land cover in South Korea. *Science of the Total Environment, 663*, 764–775. https://doi.org/10.1016/j.scitotenv.2019.01.377

Shiva Kumar, D., & Srikantaswamy, S. (2014). Factors affecting on mobility of heavy metals in soil environment. *International Journal for Scientific Research & Development., 2*, 201–203.

Singh, B., Sihag, P., & Singh, K. (2017). Modelling of impact of water quality on infiltration rate of soil by random forest regression. *Modeling Earth Systems and Environment, 3*, 999–1004. https://doi.org/10.1007/s40808-017-0347-3

Slavich, P. G., Sinclair, K., Morris, S. G., Kimber, S. W. L., Downie, A., & Van Zwieten, L. (2013). Contrasting effects of manure and green waste biochars on the properties of an acidic ferralsol and productivity of a subtropical pasture. *Plant and Soil, 366*, 213–227. https://doi.org/10.1007/s11104-012-1412-3

Song, Y., Dai, W., Shao, M., Liu, Y., Lu, S., Kuster, W., & Goldan, P. (2008). Comparison of receptor models for source apportionment of volatile organic compounds in Beijing China. *Environmental Pollution, 156*, 174–183. https://doi.org/10.1016/j.envpol.2007.12.014

Stanimirova, I., Tauler, R., & Walczak, B. (2011). A comparison of positive matrix factorization and the weighted multivariate curve resolution method. application to environmental data. *Environmental Science and Technology, 45*, 10102–10110. https://doi.org/10.1021/es201024m

Sawut, R., Kasim, N., Maihemuti, B., Hu, L., Abliz, A., Abdujappar, A., & Kurban, M. (2018). Pollution characteristics and health risk assessment of heavy metals in the vegetable bases of northwest China. *Science of the Total Environment, 642*, 864–878. https://doi.org/10.1016/j.scitotenv.2018.06.034

Shukla, K., Kumar, P., Mann, G. S., & Khare, M. (2020). Mapping spatial distribution of particulate matter using kriging and inverse distance weighting at supersites of megacity Delhi. *Sustainable Cities and Society, 54*, 101997. https://doi.org/10.1016/j.scs.2019.101997

Song, H., Hu, K., An, Y., Chen, C., & Li, G. (2018). Spatial distribution and source apportionment of the heavy metals in the agricultural soil in a regional scale. *Journal of Soils and Sediments, 18*, 852–862. https://doi.org/10.1007/s11368-017-1795-0

Sun, C., Liu, J., Wang, Y., Sun, L., & Yu, H. (2013). Multivariate and geostatistical analyses of the spatial distribution

and sources of heavy metals in agricultural soil in Dehui, Northeast China. *Chemosphere, 92*, 517–523. https://doi.org/10.1016/j.chemosphere.2013.02.063

Sun, G. X., Wang, X. J., & Hu, Q. H. (2011). Using stable lead isotopes to trace heavy metal contamination sources in sediments of Xiangjiang and Lishui Rivers in China. *Environmental Pollution, 159*, 3406–3410. https://doi.org/10.1016/j.envpol.2011.08.037

Sutherland, R. A., Tolosa, C. A., Tack, F. M. G., & Verloo, M. G. (2000). Characterization of selected element concentrations and enrichment ratios in background and anthropogenically impacted roadside areas. *Archives of Environmental Contamination and Toxicology, 38*, 428–438. https://doi.org/10.1007/s002440010057

Tang, Z., Chai, M., Cheng, J., Jin, J., Yang, Y., Nie, Z., Huang, Q., & Li, Y. (2017). Contamination and health risks of heavy metals in street dust from a coal-mining city in eastern China. *Ecotoxicology and Environmental Safety, 138*, 83–91. https://doi.org/10.1016/j.ecoenv.2016.11.003

Tomlinson, D. L., Wilson, J. G., Harris, C. R., & Jeffrey, D. W. (1980). Problems in the assessment of heavy-metal levels in estuaries and the formation of a pollution index. *Helgoländer Meeresuntersuchungen, 33*, 566–575. https://doi.org/10.1007/BF02414780

Vacek, O., Vašát, R., & Borůvka, L. (2020). Quantifying the pedodiversity-elevation relations. *Geoderma, 373*, 114441.

Wang, M., Bai, Y., Chen, W., Markert, B., Peng, C., & Ouyang, Z. (2012). A GIS technology based potential eco-risk assessment of metals in urban soils in Beijing, China. *Environmental Pollution, 161*, 235–242. https://doi.org/10.1016/j.envpol.2011.09.030

Wang, Y. B., Gou, X., Su, Y. B., & G.W. (2006). Wang: Risk assessment of heavy metals in soils and - Google Scholar [WWW Document]. *Journal of Environmental Science., 18*(6), 1124–1134.

Weather Spark, (2016). Average Weather in Frýdek-Místek, Czechia, Year Round - Weather Spark [WWW Document]. URL https://weatherspark.com/y/83671/Average-Weather-in-Frýdek-Místek-Czechia-Year-Round (Accessed 9.14.20).

Wu, J., Margenot, A. J., Wei, X., Fan, M., Zhang, H., Best, J. L., Wu, P., Chen, F., & Gao, C. (2020). Source apportionment of soil heavy metals in fluvial islands, Anhui section of the lower Yangtze River: Comparison of APCS–MLR and PMF. *Journal of Soils and Sediments, 20*, 3380–3393. https://doi.org/10.1007/s11368-020-02639-7

Xue, J. L., Zhi, Y. Y., Yang, L. P., Shi, J. C., Zeng, L. Z., & Wu, L. S. (2014). Positive matrix factorization as source apportionment of soil lead and cadmium around a battery plant (Changxing County, China). *Environmental Science and Pollution Research, 21*, 7698–7707. https://doi.org/10.1007/s11356-014-2726-x

Yang, B., Zhou, L., Xue, N., Li, F., Li, Y., Vogt, R. D., Cong, X., Yan, Y., & Liu, B. (2013). Source apportionment of polycyclic aromatic hydrocarbons in soils of Huanghuai Plain, China: Comparison of three receptor models. *Science of the Total Environment, 443*, 31–39. https://doi.org/10.1016/j.scitotenv.2012.10.094

Yang, Y., Christakos, G., Guo, M., Xiao, L., & Huang, W. (2017). Space-time quantitative source apportionment of soil heavy metal concentration increments. *Environmental Pollution, 223*, 560–566. https://doi.org/10.1016/j.envpol.2017.01.058

Yang, Z., Lu, W., Long, Y., Bao, X., & Yang, Q. (2011). Assessment of heavy metals contamination in urban topsoil from Changchun City, China. *Journal of Geochemical Exploration, 108*, 27–38. https://doi.org/10.1016/j.gexplo.2010.09.006

Yeung, Z. L. L., Kwok, R. C. W., & Yu, K. N. (2003). Determination of multielement profiles of street dust using energy dispersive X-ray fluorescence (EDXRF). *Applied Radiation and Isotopes, 58*, 339–346. https://doi.org/10.1016/S0969-8043(02)00351-2

Yuanan, H., He, K., Sun, Z., Chen, G., & Cheng, H. (2020). Quantitative source apportionment of heavy metal (loid)s in the agricultural soils of an industrializing region and associated model uncertainty. *Journal of Hazardous Materials, 391*, 122244.

Zhang, X., Zhong, T., Liu, L., Zhang, X., Cheng, M., Li, X., & Jin, J. (2016). Chromium occurrences in arable soil and its influence on food production in China. *Environmental Earth Sciences, 75*(3), 1–8.

Zhang, H., Li, H., Yu, H., & Cheng, S. (2020). Water quality assessment and pollution source apportionment using multistatistic and APCS-MLR modelling techniques in Min River Basin China. *Environmental Science and Pollution Research, 27*, 41987–42000. https://doi.org/10.1007/s11356-020-10219-y

Zhang, J., Li, R., Zhang, X., Bai, Y., Cao, P., & Hua, P. (2019). Vehicular contribution of PAHs in size dependent road dust: A source apportionment by PCA-MLR, PMF, and Unmix receptor models. *Science of the Total Environment, 649*, 1314–1322. https://doi.org/10.1016/j.scitotenv.2018.08.410

Zhang, J., & Liu, C. L. (2002). Riverine composition and estuarine geochemistry of particulate metals in China - Weathering features, anthropogenic impact and chemical fluxes. *Estuarine, Coastal and Shelf Science, 54*, 1051–1070. https://doi.org/10.1006/ecss.2001.0879

Zhao, F. J., Ma, Y., Zhu, Y. G., Tang, Z., & McGrath, S. P. (2015). Soil contamination in China: Current status and mitigation strategies. *Environmental Science and Technology*. https://doi.org/10.1021/es5047099

Zhao, S., Qiu, S., & He, P. (2018). Changes of heavy metals in soil and wheat grain under long-term environmental impact and fertilization practices in North China. *Journal of Plant Nutrition, 41*, 1970–1979. https://doi.org/10.1080/01904167.2018.1485158

Zheng, Y., Shen, D., Wu, S., Han, Y., Li, S., Tang, F., Ni, Z., Mo, R., & Liu, Y. (2018). Uptake effects of toxic heavy metals from growth soils into jujube and persimmon of China. *Environmental Science and Pollution Research, 25*, 31593–31602. https://doi.org/10.1007/s11356-018-2959-1

Zhu, G., Guo, Q., Xiao, H., Chen, T., & Yang, J. (2017). Multivariate statistical and lead isotopic analyses approach to identify heavy metal sources in topsoil from the industrial zone of Beijing Capital Iron and Steel Factory. *Environmental Science and Pollution*

Research, 24, 14877–14888. https://doi.org/10.1007/s11356-017-9055-9

Zhu, J. J., Cui, B. S., Yang, Z. F., Dong, S. K., & Yao, H. R. (2006). Spatial distribution and variability of heavy metals contents in the topsoil along roadside in the Longitudinal Range-Gorge Region in Yunnan Province. *Acta Ecologica Sinica, 26*, 146–153.

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