

Original Research

Source Identification and Spatial Variation Analysis of Potentially Hazardous Elements in Topsoil Based on Multi-Source Data in Longkou City, China

Qing Yu^{1,2}, Quanyuan Wu^{1*}, Jianmin Qiao^{1**}, Xu Zhou³, Xibo Xu¹,
Jianfei Cao¹, Lei Yao¹, Yurong Cui¹

¹College of Geography and Environment, Shandong Normal University, Jinan 250358, China

²Key Laboratory of Water Cycle and Related Land Surface Processes, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China

³State Key Laboratory of Earth Surface Processes and Resource Ecology, Beijing Normal University, Beijing 100875, China

Received: 10 May 2021

Accepted: 25 June 2021

Abstract

This study identified the sources and analyzed the spatial variation of potentially hazardous elements (Cr, Cu, Ni, Pb, Zn, Cd, As, and Hg) using land use, soil parent material, slope and elevation, combined with geostatistics and multi-source statistics in Longkou city, China. The results showed that Cr, Ni, and As were mainly controlled by natural factors such as granite and elevation, and high-concentration areas were mainly in the northeast and south of the study area. Higher Hg levels were derived from mixed sources, including industrial and transportation activities in the central urban area. Cd was mainly controlled by mixed sources, and its spatial variation was affected by elevation, industrial activities, and granite. Pb came from mixed sources, and its spatial variation was influenced by smelting and gilding industries and alluvial sediments. Cu originated from artificial sources and was affected by industrial and agricultural activities; it was mainly distributed in mining areas in the northwest and orchards in the southeast of the study area. Zn came from mixed sources, and its spatial variation was affected by industry, agriculture, and alluvial sediments; high concentrations were mainly found in mining areas in the northwest and orchards and farmland in the south.

Keywords: potentially hazardous elements (PHEs), multi-source data, topsoil, source identification, spatial variation

*e-mail: wqy6420582@163.com

**e-mail: qjmwilson@163.com

Introduction

Soil potentially hazardous elements (PHEs) pollution is a hidden problem, which has become a global environmental problem due to its irreversibility and intractability. The increase of available fraction of PHEs in soil can not only directly change the internal properties of the soil, but can also spread into other compartments (e.g., air, water, and organisms), and may even endanger human health via exposure from water and food intake, dermal contact, and other ways [1-3]. Therefore, PHE pollution has become an important indicator of regional environmental quality assessments. It is particularly important to identify the sources of PHEs and study the spatial variations of various PHEs in soil, which can provide important decision-making information for the protection and treatment of soil.

Differences in soil concentrations of PHEs are driven by natural or anthropic factors. For instance, the natural sources are mainly influenced by the geological background, such as the own PHE concentration of rocks and minerals [4]. Meanwhile, anthropogenic sources include direct or indirect discharge of PHEs from human activities, such as mining, smelting, burning of fossil fuels, sewage irrigation, use of leaded gasoline and motor vehicle traffic, and widespread use of chemical fertilizers and pesticides [5, 6]. With the continuous development of social economy, human activities have a significant impact on PHE concentrations, exceeding the contributions of natural sources. Studies are increasingly using a combination of multivariate statistics and geostatistical methods to analyze the distribution and sources of pollutants in the environment [7-10]. However, most studies on the sources of PHEs have only used geostatistics and multivariate analysis to analyze the sample data, without considering the environmental factors and actual situation of the study area. Land use and soil parent materials generally represent human activities and geological background factors [11, 12]. So, statistics on the type of land use and soil parent material at the sampling site can reflect the local geographical environment more directly, offering in-depth identification of pollutant sources with increased accuracy and validation of whether the results of multivariate analysis are reasonable. The geostatistics method combined with multi-source data can better support the mapping of spatial distribution of PHEs and the modeling of spatial variability. It is a new statistical method combining geographic information science and remote sensing.

Shandong peninsula blue economic zone, as China's first regional development strategy with the theme of marine economy, is a major strategic move of China's regional development from land economy to marine economy. Longkou city is one of the nine core areas in this economic zone, and is also the main grain and economic crop production area in Shandong Province, China. However, the rapid development of both

agriculture and industry in recent years has driven soil PHE accumulation, and soil PHE pollution in Longkou is an urgent problem to be solved.

The specific objectives of this paper were to: (a) identify the sources of PHEs (Cr, Cu, Ni, Pb, Zn, Cd, As, and Hg) using multivariate statistical techniques combined with land-use and soil parent material types at sampling sites, and (b) analyze the spatial structure, variation, and causes of the spatial variation of PHEs through geostatistical analysis combined with multi-source data. The present study attempts to provide a statistical method combining geographic information science and remote sensing to explore the distribution and spatial variation of PHEs in soil.

Materials and Methods

Study Area

The study area (120°13'14"E~120°44'46"E, 37°27'30"N~37°47'24"N) is located in the north of Longkou city, Shandong Province, China, covering an area of about 500 km². According to the Köppen-Geiger climate classification [13], the study area belongs to warm temperate climate with dry winter. The annual average temperature is 12°C, and the annual average rainfall is about 600 mm. The dominant wind direction is northwest wind. The types of agricultural land in the study area include farmland, orchards, etc. Crops include winter wheat and summer corn, and economic crops are apples and grapes. This is the main food and economic crop-producing area in the Jiaodong region of China.

Longkou is the core area of Shandong peninsula blue economic zone. It is one of the top 100 industrial cities in China, and has experienced rapid industrial development. Industries in the area mainly include petrochemical, power energy, machinery manufacturing, smelting foundry, processing trade, and shipbuilding, and its transportation industry is well developed.

Soil Samples Collection and Chemical Analysis

According to a land use and field survey map, sampling sites were initially determined using a S-shaped method, and 125 topsoil samples were collected in Longkou City in September 2017. During the sampling process, a square cell of 30 m² was demarcated at each sampling site, and 5 soil samples were collected by diagonal sampling method and mixed evenly to form a soil sample. When collecting the samples, we used a wooden spatula to collect soil to a depth of 0-20 cm and removed coarse stones and plant debris and roots from the soil. Finally, the collected soil samples were put into polyethylene bags and labelled. The location of the study area and the distribution of sampling sites are shown in Fig. 1.

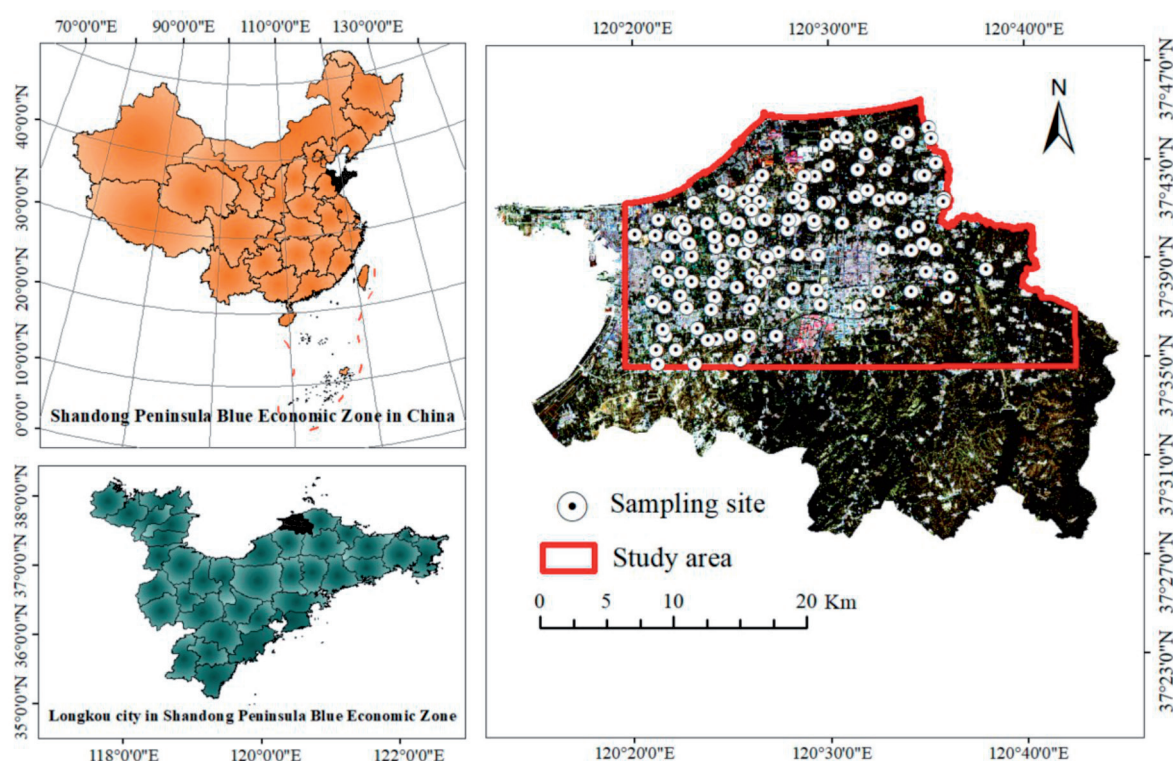


Fig. 1. Location of the sampling sites in Loukou City, China.

Air-dried soil samples were gently crushed and then the fraction <2 mm was pulverized. And then the concentrations of eight PHEs were analyzed. The samples used to analyze Cr, Cd, Ni, Cu, Pb, and Zn were digested with $\text{HSO}_4\text{-HNO}_3\text{-HF}$. The concentrations of Cr, Cu, Ni, Pb, and Zn were analyzed using inductively coupled plasma-mass spectrometry (ICP-MS; Thermo Fisher Scientific iCAP 6300, USA), while the concentration of Cd was determined by graphite furnace atomic absorption spectrometry (GF-AAS; PerkinElmer PinAAcle 900T, USA). The samples used to analyze Hg and As were digested with $\text{HSO}_4\text{-HNO}_3\text{-KMnO}_4$, and Hg and As were determined by hydride generation atomic fluorescence spectrometry (HG-AFS; Titan AFS 830, China). All chemicals used were of guaranteed reagent grade. Standard reference materials were obtained from the Center of National Standard Reference Material of China. The recoveries of all eight PHEs were within the range of $100\pm 10\%$.

Data Sources and Processing

Landsat 8 OLI data and ASTER GDEM data covering the study area were downloaded. The remote sensing images were pretreated for atmospheric correction, geometric correction, splicing, and cutting. Then, the images were classified into eight land-use types: orchards, construction land, farmland, traffic land, meadow, water area, woodland, and other (Fig. 2a). The map of the soil parent material types in the study area was extracted from the Shandong Geological

Map (1:150,000) and the mapping was performed in ArcGIS 10.2 (Esri, Redlands, CA, USA) (Fig. 2b). For the digital elevation model data, the attribute values of elevation and slope corresponding to the sampling sites were extracted through the spatial analysis module of ArcGIS, and the slope map (Fig. 2c) and elevation map (Fig. 2d) created. Table 1 presents additional information on this data.

Multivariate Statistical Analysis

Multivariate statistical analysis is the most commonly used method in the study of soil PHEs. In this study, it was used to analyze the source of PHEs in topsoil. Multivariate statistical methods such as correlation analysis, PCA, and one-way analysis of variance (ANOVA) were used to process the data. The software used for this purpose was the IBM SPSS Statistics 22.0 (SPSS Inc., Chicago, IL, USA).

Geostatistical Analysis

Variogram

Variogram is a unique analysis tool in geostatistics, which is the basis of geostatistical analysis. The coefficient of variation is a key quantitative value for studying the spatial variability of a variable. Because sampling sites are often discrete, the variation function is difficult to obtain directly. Therefore, after selecting the semivariogram model, the experimental variation

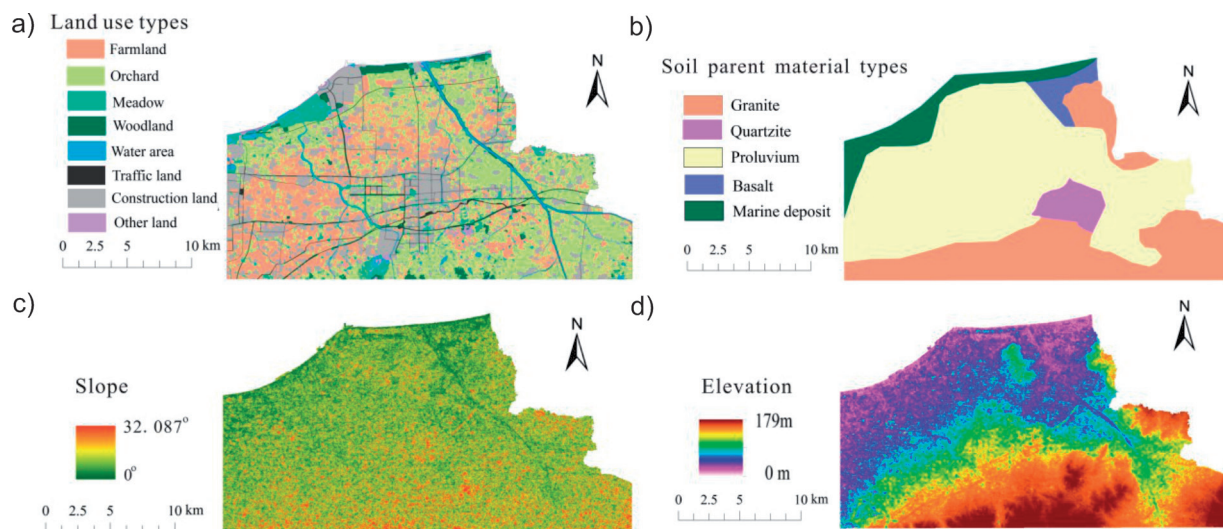


Fig. 2. Various topography and geological factors extracted from multi-source data.

Table 1. Relevant data information used to analyze PHEs.

	Source	Links
Land use	Landsat 8 OLI	http://glovis.usgs.gov
Elevation, Slope	ASTER GDEM	http://glovis.usgs.gov
Soil parent material	Geological Map	http://www.sdgtyzylg.com

values are fitted and the parameters of the theoretical model are solved. The semivariogram formula is:

$$r(h) = \frac{1}{2N(h)} \cdot \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (1)$$

...where $N(h)$ is the total number of data pairs separated by distance h , Z represents the measured value of a given soil property, and x is the sampling position of the soil samples. The software used for this purpose was the GS+9.0 (Gamma Design Software, Plainwell, MI, USA).

Kriging

Kriging method is the most widely used interpolation method and is based on a stochastic spatial variation model. The underlying assumption is that soil properties appear to be inherently stable regionalized random variables. Therefore, the kriging method can be used to estimate the confidence intervals at unsampled locations, and the equation is described as follows [14]:

$$z^*(x_0) = \sum_{i=1}^n \lambda_i z(x_i) \quad (2)$$

...where x_0 is the unknown point, $z^*(x_0)$ is the estimated value, $z(x_i)$ is the observed value at location i , n is the number of known points, and λ_i is the weight value

with sum of 1 to ensure that there is no deviation. The software used for this purpose was ArcGIS 10.5 (ESRI Co, Redlands, USA).

Results and Discussion

Sources Identification of PHEs in Soils

Descriptive Statistics of PHE Concentration in Soil

Table 2 shows the statistical results of the concentrations of various PHEs in soil. The means of these PHEs exceeded the background values (BVs) of PHEs in eastern Shandong Province (CNEMC (China National Environmental Monitoring Centre)). The average concentrations of Cu, Pb, Hg, As, Zn, Ni, and Cr reached 1.46, 1.35, 1.38, 1.21, 1.14, 1.11, and 1.02 times their respective BVs. In particular, the concentration of Cd was 2.13 times the BV of eastern Shandong Province, indicative of greater enrichment of Cd as affected by human activities. Although these PHEs showed a degree of enrichment, they did not exhibit serious pollution. Coefficient of variation (CV) can reflect data fluctuation, where larger values indicate greater influences of external factors on PHE concentrations. For instance, Wilding divided CVs into three levels: high variability ($CV \geq 0.36$), moderate variability ($0.16 < CV < 0.36$), and low variability

Table 2. Statistical results of PHE concentrations in soil.

Element	Concentration (mg·kg ⁻¹)			CV	Background values (mg·kg ⁻¹)
	Min	Max	Mean		
Cr	31.62	108.10	57.28	0.23	56.20
Cu	14.43	112.33	28.64	0.78	19.60
Ni	12.04	74.61	26.18	0.27	23.50
Pb	11.39	195.70	34.23	0.53	25.40
Zn	29.44	187.70	64.00	0.38	56.10
Cd	0.03	0.87	0.23	0.65	0.11
As	3.52	20.48	7.65	0.26	6.30
Hg	0.01	0.16	0.04	0.75	0.03

($CV \leq 0.16$) [15]. Cu, Hg, Cd, Pb, and Zn showed high variability, and Ni, As, and Cr moderate variability. This indicated that the spatial distributions of Cu, Hg, Cd, Pb, and Zn were discrete and uneven. These results indicated that human activities may have an effect on the enrichment of Cu, Hg, Cd, Pb and Zn, while Ni, As and Cr were less affected by human activities. We need to further analyze the sources of the PHEs.

Correlation Analysis Results

Pearson's correlation coefficients of PHEs in soils are summarized in Table 3. The highly positive correlations between heavy metals may indicate that these elements have similar sources [16-18]. The correlation coefficients ($p < 0.01$) between any two of Cr, Ni, and As were above 0.48. The correlation coefficient between Cr and Fe_2O_3 was 0.809, and that between Cr and MgO was 0.638. The correlation coefficient between Ni and Fe_2O_3 was 0.883, while

that between Ni and MgO was 0.766. As also had high correlations with Fe_2O_3 and MgO. Fe_2O_3 and MgO are important products in the formation of parent rock, and are correlated with natural sources [12, 16]. Therefore, Cr, Ni, and As likely mainly originated from the parent material. Moreover, the correlation between Cd and Pb was very strong; Pb and Cd likely came from the same source [19-21]. And the correlation coefficient between Zn and Cu reached 0.587, indicative of a strong relationship and suggesting a shared source. The correlation between Hg and other PHEs were relatively weak, and the degree of Hg variation was very high. Therefore, Hg appeared to be an isolated element and further analysis was required to determine its source more accurately.

PCA Analysis Results

The results of the PCA of PHE concentrations are shown in Table 4. The extracted three principal

Table 3. Correlation coefficient matrix of PHEs in soils.

	Cr	Cu	Ni	Pb	Zn	Cd	As	Hg	Fe_2O_3	MgO
Cr	1									
Cu	-0.050	1								
Ni	0.813**	-0.022	1							
Pb	0.265**	0.012	0.173	1						
Zn	0.153	0.587**	0.296**	0.102	1					
Cd	0.450**	0.17	0.438**	0.485**	0.123	1				
As	0.579**	0.237**	0.484**	0.329**	0.192*	0.349**	1			
Hg	-0.014	-0.030	-0.161	0.042	0.163	-0.207	0.102	1		
Fe_2O_3	0.809**	-0.043	0.883**	0.103	0.221	0.123	0.573**	0.273*	1	
MgO	0.638**	-0.10	0.766**	0.051	0.222	0.091	0.266*	0.232	0.823**	1

** Correlation is significant at the 0.01 level

* Correlation is significant at the 0.05 level

components together explained 71.83% of the data variance. This indicates that these three principal component factors essentially reflect the pollution levels of the PHEs. It can be intuitively seen from the dispersion degree of each element in Fig. 3 that the PHEs in the study area mainly come from three sources. As, Ni, and Cr were strongly associated with the first principal component (PC1), the second principal component (PC2) included Cu, Pb, Zn, Cd, and Hg, and Zn and Cu were in the third principal component (PC3).

PC1 accounted for 34.33% of the variance, and showed significant loadings of Cr (0.754), Ni (0.715), and As (0.656). The mean values of these three PHEs were similar to the BVs, and the correlations between among three elements were strong. Moreover, the correlation analysis showed that the correlation coefficients between these elements with Fe_2O_3 and MgO are relatively high. Therefore, these three elements were affected by the geological background. Many studies have shown that the use of chemical fertilizers, pesticides, and organic fertilizers in alluvial plains areas have lower effects on Cr and Ni concentrations than the soil background [12]. These two PHEs exhibit the lowest soil pollution levels in China among PHEs, and are not affected by human activities in most regions. Therefore, when other PHEs are grouped with Cr and Ni, they are considered to originate from natural sources. As also came from a natural source and, overall, it seemed reasonable to infer that PC1 represented natural origins.

PC2 explained 25.19% of the variance, and showed elevated loadings of Cu, Pb, Zn, Cd, and Hg. The concentrations of Cu, Pb, Zn, Cd, and Hg exceeded their BVs, and these PHEs are usually affected by human inputs [20]. For instance, Pb mainly originates from automobile exhaust, coal combustion, etc. Meanwhile, Cd, Zn, and Cu may be affected by surrounding coal mines and factories, such as smelting and electroplating. The load of Hg in PC2 was the largest (0.623), and its correlation with Fe_2O_3 was relatively low, indicating that Hg was affected by human activities. Many studies have proven that agricultural production has a relatively low impact on Hg; therefore, electroplating, metallurgical, and other industrial emissions in the study area may have been the main sources of Hg. Similarly, Cai et al. [22] found that Hg mainly came from industrial and transportation sources in the identification of eight harmful PHEs in agricultural soil in Huizhou city, Guangdong province, China. Therefore, PC2 mainly represented emissions from industrial and mining industries, as well as traffic emissions.

PC3 with the 12.31% variance was dominated by the loading Zn and Cu, their loads were 0.459 and 0.407, respectively. Moreover, the correlation between these two elements reached 0.587 ($p < 0.01$) (Table 3). Sterilization pesticides and water-soluble fertilizers contain high concentrations of Cu and Zn. There is a large area of agricultural land in the study area, and farmers tend to use excessive amounts of pesticides

and chemical fertilizers, which can increase PHE enrichment [23]. In particular, Cai et al. [22] found that Zn and its compounds were widely used in agricultural fertilizers. Zhao et al. [24] found that intensive use of fertilizers and agrochemicals had profound impacts on the concentration of Cu and Zn concentrations. Therefore, the PC3 mainly represented the use of pesticides and fertilizers in agricultural production processes and the effect of sewage irrigation on the soil PHE concentrations.

The total variance contribution rate of pollution source factors (PC2 and PC3) was 37.50%, accounting for 52.21% of the total accumulated variance, indicating that the topsoil of the study area was affected by human industrial and agricultural activities.

Cu had considerable loadings in PC2 and PC3, indicating that it mainly originated from human activities. Zn had similar loading in PC1, PC2, and PC3, indicating that Zn came from mixed sources and was considered to be jointly controlled by the natural geological background and industrial and agricultural pollution. Pb had the same loadings in PC1 and PC2, as did Cd, indicating that these two PHEs were influenced by both the natural geological background and industrial activities, such as electroplating and metallurgy. Lv et al. [8] found that Cd, Cu, Pb, and Zn were jointly affected by the natural background as well as human activities, such as industrial, agricultural traffic, and transportation emissions. Meanwhile, Ghrefat and Yusuf [7] found that Pb and Cd originated from mixed sources when studying the sources of PHEs in Wadi Al-Arab sediments, and Pb and Cd were influenced by anthropogenic and natural sources. Overall, the results of these studies were consistent with our findings.

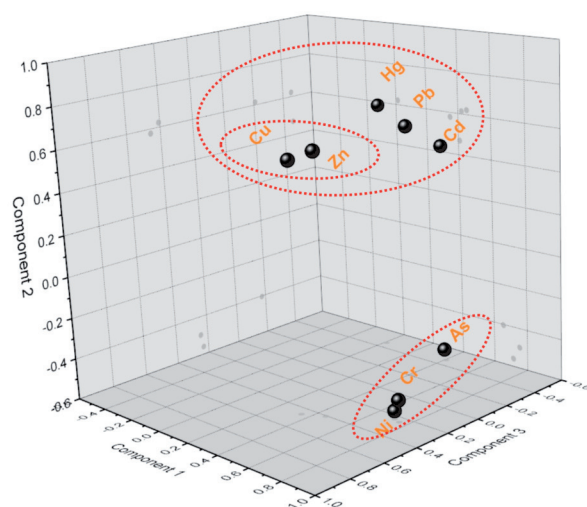


Fig. 3. Loading plot of principal component analysis based on PHE concentrations.

Table 4. Total variance explained and component matrix for PHEs in the soils.

Elements	PC1	PC2	PC3
Cr	0.754	-0.496	0.245
Cu	0.351	0.578	0.459
Ni	0.715	-0.561	0.232
Pb	0.423	0.613	-0.209
Zn	0.455	0.622	0.407
Cd	0.408	0.470	-0.461
As	0.656	-0.404	-0.175
Hg	0.013	0.623	-0.457
% of variance	34.33	25.19	12.31
Cumulative (%)	34.33	59.52	71.83

The Difference of PHE Concentrations in Different Land Uses and Soil Parent Material Types

Land use and parent materials usually represent human activities and geological background factors, respectively. We used ANOVA to compare PHE concentrations of different groups for a more in-depth analysis of the sources of PHEs, as shown in Figure 4. Different lowercase letters (a, b, c, and d) indicated that there were significant differences in the level of 0.05 among different land uses ($n = 4$). Comparing the various PHE concentrations of different land-use types (Fig. 4a), the land-use types with the highest mean concentrations of Cd, Pb, and Hg was construction land, with mean concentrations of 0.288, 52.22, and 0.0585 $\text{mg}\cdot\text{kg}^{-1}$, respectively. The concentrations of these PHEs in construction land were significantly higher than in the other land-use types, indicating that construction land had a significant effect on Cd, Pb, and Hg enrichment. The mean concentration of Cu in orchard was 54.1045 $\text{mg}\cdot\text{kg}^{-1}$, which was much higher than the mean concentration in other land-use types. Pesticides contain large amounts of PHEs, such as Cu, and are applied to orchard soil, which can cause soil PHE pollution. The concentration of Cu in farmland was higher than in construction land and grassland, but was highest in orchard. Pesticide use in orchards is often higher than in general field crops. Overall, the above analysis showed that Cd, Pb, Cu, and Hg were mainly controlled by varying degrees of human activity.

There were differences in PHE concentrations of in different parent materials (Fig. 4b). The mean concentrations of Cr, Ni, Cd, and As in the granite parent material were significantly higher than in other parent materials. The concentrations of Pb and Zn in alluvial sediments were also slightly higher than those in other parent materials. The concentrations of Cu and Hg were very similar among the parent materials, without significant differences. Thus, Cr, Ni, Cd, As,

and Pb were affected by the parent material, which was consistent with the PCA results.

According to the ANOVA, it could be concluded that Cu and Hg were mainly affected by human activities, Cr, Ni, and As were closely related to the parent material, and Cd, Pb, and Zn were jointly affected by the parent material and industrial and agricultural activities. In general, the results of the ANOVA further supported the multivariate statistical analysis results.

Spatial Variations of PHEs

Spatial Structures of PHEs

Variogram can analyze the spatial structures of regional variables [12, 25]. In this study, the investigated PHEs were fitted to the variogram function, and kriging interpolation was selected for the variogram model to analyze their distributions.

Various types of variograms exist, including exponential, spherical, linear, and gaussian models, with parameters including nugget (C_0), sill (C_0+C), coefficient of determination (R^2), residual sum of squares (RSS), and range value (R) [23-27]. In this study, C_0 indicates the degree of spatial distribution of PHEs due to unnatural factors, whereas C_0+C indicated

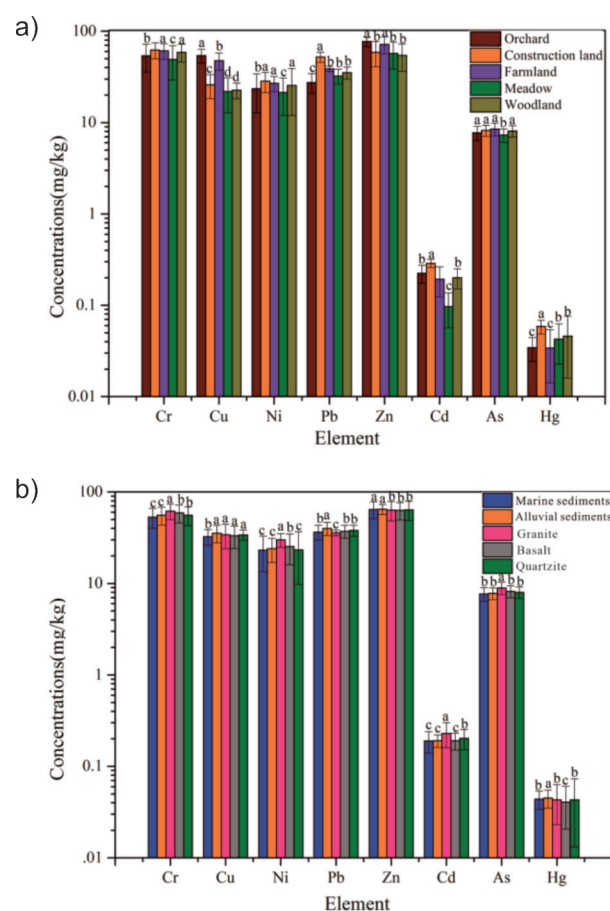


Fig. 4. Statistics for PHE concentrations in various land uses and parent materials.

Table 5. The parameters and the cross-validation results of the fitted variogram models.

	Model	Co	Co+C	Co/Co+C	R/m	Rss	R ²
Cr	Spherical	0.0278	0.1367	0.2034	3874	0.0510	0.702
Cu	Spherical	0.1340	0.1570	0.8535	6306	0.0039	0.901
Ni	Exponential	0.1000	0.6735	0.1485	4447	0.0942	0.678
Pb	Spherical	41.7520	67.3005	0.6204	4792	0.0173	0.929
Zn	Spherical	10.2460	35.1213	0.2904	6281	0.0439	0.776
Cd	Exponential	0.0560	0.1198	0.4706	6385	0.0245	0.813
As	Spherical	0.0780	0.3322	0.2348	7213	0.0721	0.685
Hg	Exponential	0.0980	0.1240	0.7903	3416	0.0087	0.894

the degree of total variation within the system. Therefore, Co/Co+C indicated the spatial variation of PHE concentrations caused by artificial and other non-natural random factors accounting for the proportion

of the total variation. If the ratio is less than 25%, the spatial variation is dominated by structural variations, and each sampling site has a strong spatial correlation, between 25% and 75%, the variable has moderate

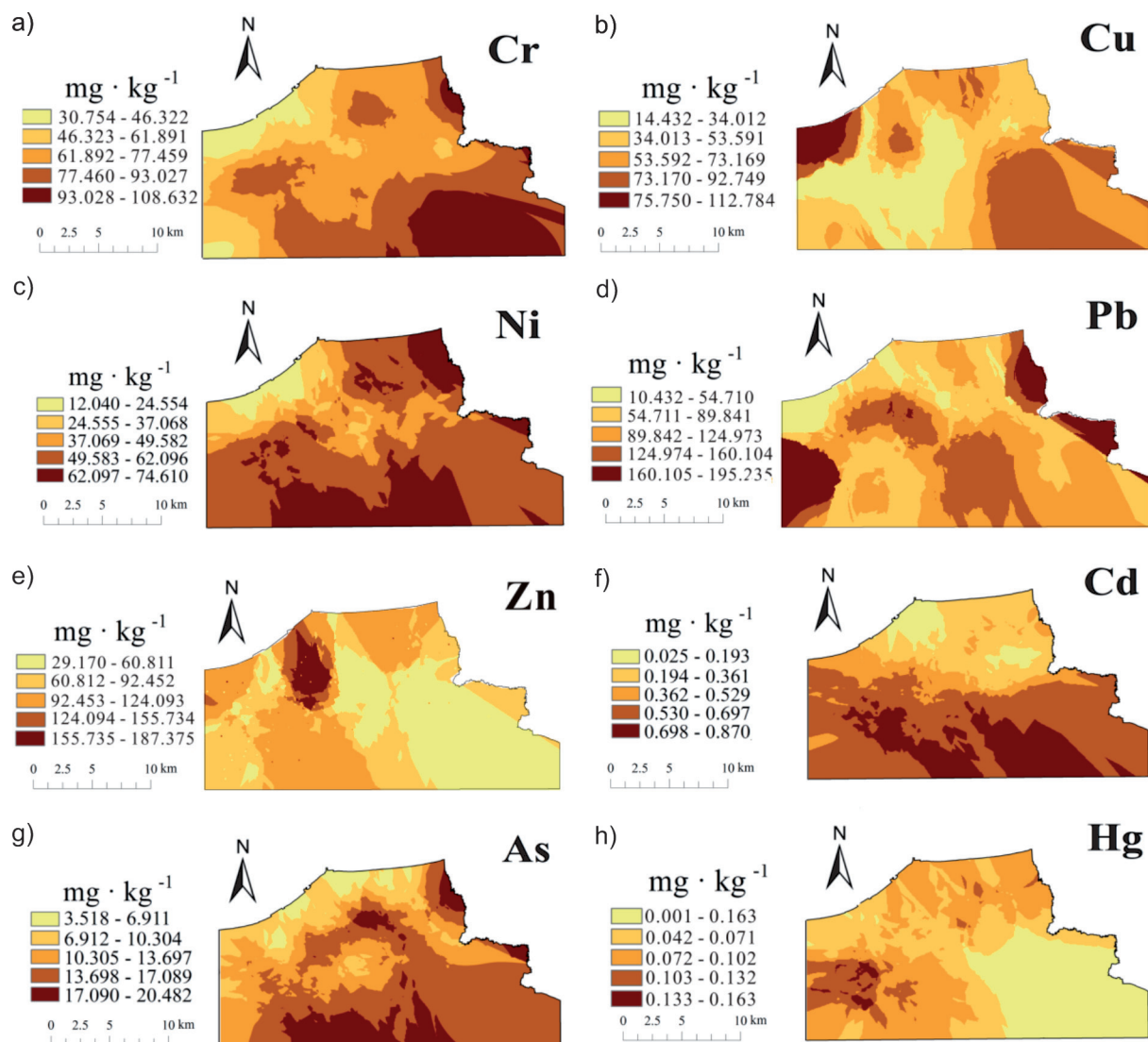


Fig. 5. Spatial distribution of PHE concentrations in the study area.

Table 6. Correlation coefficient matrix between PHEs and terrain factor.

Terrain factor	Cr	Cu	Ni	Pb	Zn	Cd	As	Hg
Elevation	0.492**	0.113	0.388**	0.121	0.098	0.347**	0.321**	-0.413**
Slope	-0.190	0.050	-0.194	-0.033	0.045	0.004	-0.117	0.002

** Correlation is significant at the 0.01 level

spatial dependence, and greater than 75%, the variable shows only weak spatial dependence [17, 27-30]. In this study, the optimal spatial variation model was selected based on the principle of the maximum value of R^2 and the minimum value of RSS.

The premise of variogram fitting and kriging interpolation requires data to be normally distributed. We used the Kolmogorov-Smirnov (K-S) test to assess whether the data conformed to a normal distribution. The K-S test indicated that only Cr conformed to a normal distribution, and could be directly fitted and interpolated by kriging interpolation. Therefore, Cu, Ni, Pb, Zn, Cd, As, and Hg concentrations were normalized after logarithmic processing to enable interpolation analysis. According to the statistical results of the theoretical model of the optimal variation function (Table 5), the RSS of each optimal model was smaller and R^2 was more than 0.65, showing an ideal effect. The theoretical models of Cu, Pb, As, Zn, and Cr were spherical, Hg was exponential, Ni and Cd were Gaussian. The Co/Co+C of Pb, Zn, and Cd were between 0.25 and 0.75 with moderate spatial autocorrelation, indicating that their concentrations were affected by both natural and human factors, which is consistent with the source interpretation analysis. The Co/Co+C for Ni, Cr, and As were less than 0.25. This indicated that spatial structural variation was dominant, the spatial correlation of each sampling site was strong, and the concentrations of As, Ni, and Cr in soil were controlled by the parent material and other internal factors. The nugget values of Hg and Cu were 0.7903 and 0.8535, respectively, indicating that they were dominated by spatial variations caused by random factors, and the spatial distribution was greatly affected by human factors. This was consistent with the multivariate analysis results. The variation range values of each variable were between 3,476 and 7,213 m, which were larger than the actual maximum sampling interval of this study, indicating that all sites were within the range of spatial variation and could reflect the real variation characteristics of the PHEs.

Spatial Distribution and Variation Analysis

The correlations among elevation data, slope data, and PHE concentration data were analyzed (Table 6). There was no significant correlation between the PHE concentration and slope, indicating that the distribution of PHEs was not affected by slope, and that spatial interpolation analysis could be performed. Based on the optimal variogram fitting results (Table 5), spatial

interpolation was performed using the ordinary kriging method (Fig. 5).

The spatial distributions of Cr, Ni, and As were similar, with lower concentrations in the northwest and higher concentrations in the northeast and south (Fig. 5). According to the correlation analysis of terrain factors and PHE concentrations (Table 6), the correlation coefficients of Cr, Ni, and As with elevation were 0.492, 0.388, and 0.321, respectively ($p < 0.01$). Combined with Fig. 2d), the spatial distribution characteristics of Cr, Ni, and As were similar to that of elevation. This indicated that elevation affected the spatial distributions of Cr, Ni, and As. In combination with Fig. 2b), the high-concentration areas of Cr, Ni, and As were generally consistent with the distribution of granite. From the ANOVA, the concentrations of Cr, Ni, and As in granite were higher than those in other parent materials.

Hg accumulation was observed in the central urban area and north and parts of the west. In the correlation analysis between elevation and Hg, the correlation coefficient reached -0.413 ($p < 0.01$), and elevation was strongly negatively correlated with Hg. Most of the central area was construction land (Fig. 2a). Meanwhile, the ANOVA of PHE concentrations of different land-use types revealed markedly higher Hg concentrations in construction land than in other land-use types. Therefore, areas with heavy industrial and transportation activities in the urban areas of the central region and lower elevations in the north and west led to the accumulation of Hg.

Areas with high Zn concentrations were mainly distributed in the northwest and south. The use of fertilizers and pesticides in agricultural production, as well as the discharge of domestic sewage and domestic waste, enriched Zn concentrations in these areas. Moreover, the PCA revealed that soil parent materials and industrial and agricultural activities impacted Zn enrichment. However, the correlation between elevation and PHE concentrations were low (Table 6), and Zn concentrations were not affected by elevation.

Pb accumulation was observed in the southwest and northeast. Field investigations revealed many electroplating and metallurgical industries in the northeast, which emitted large amounts of Pb. Meanwhile, farmland was widely distributed in southwestern area, with substantial sewage irrigation in this area (Fig. 2a). Therefore, accumulation of Pb was caused by the direct irrigation of farmland with treated industrial sewage. In addition, alluvial sediments were the parent material in the southwest, and the

ANOVA of parent materials showed that Pb levels were significantly higher in alluvial sediments than in other parent materials.

Cd was mainly distributed in the central and southern parts of the study area. The mean Cd concentration (Table 2) was more than twice the BV of eastern Shandong. The central part of the study area was urban construction land, containing various industrial parks and traffic pollution that caused the accumulation of Cd. From the distribution pattern of parent materials, granite was distributed in the south. Meanwhile, from the ANOVA of different parent materials, the concentration of Cd in granite was higher than that of other parent materials, indicating that the distribution of Cd was affected by industrial activities in the middle and granite in the south of the study area. Finally, the correlation coefficient between the concentration of Cd and elevation was 0.347 ($p < 0.01$) (Table 5), which indicated that the distribution of Cd was also affected by elevation.

Finally, Cu was mainly distributed in the northwest and southeast of the study area. The land-use map revealed many orchards in the southeast part of the study area. Areas with high Cu concentrations were mainly associated with grape, apple, and other fruit tree cultivation, because local farmers overuse pesticides and fertilizers, resulting in Cu enrichment. In the northwest area of the study area, there were many coal mines. Dust, coal gangue, and coal washing water from coal mining cause PHE pollution, especially highly migratory elements such as Cu.

Conclusions

This study conducted an in-depth analysis of the sources, distributions and spatial variations of eight PHEs in Longkou city, China, using multivariate statistics and geostatistics methods based on multisource data. The results revealed that Cr, Ni, and As were derived from natural sources, Hg originated from transportation and industrial activity, Cu was mainly derived from human activities such as industry and agriculture, Pb and Cd were derived from the natural geological background and the influence of industrial activities, and Zn had a mixed source and was controlled both by the geological background and human activities. The spatial variation of PHE was caused by terrain factors, the distribution of soil parent material, and industrial and agricultural activities within different land-use types. The spatial variations of Cr, Ni, and As were mainly affected by the presence of granite in the south and northeast, as well as elevation. Cu was mainly distributed near coal mines in the northwest and orchards in the southeast. Zn was affected by industry, agriculture, and alluvial sediments. The spatial variation of Cd was mainly influenced by industrial activities in the central urban area, the geological background, and the elevation of the southern region. Pb was mainly affected by smelting,

gold plating, and other industries in the northeast area and alluvial sediments in the southwest. The spatial variation of Hg was mainly affected by industry, transportation, and elevation.

As evidenced by this study, the use of multivariate statistics and geostatistical analysis combined with multisource data can be used as a powerful tool to study the source and spatial variation of PHEs. This study explored a geographic statistical method combining geographic information science and remote sensing, which provided scientific basis and support for land-use planning and soil environment management.

Acknowledgments

This work was supported by the National Natural Science Foundation of China [Grant No. 41371395] and Shandong Province Key Research and Development Plan [Grant No. 2019GSF109034].

Conflict of Interest

The authors declare no conflict of interest.

References

1. CHEN H.Y., TENG Y.G., LU S.J., WANG Y.Y., WANG J.S. Contamination features and health risk of soil heavy metals in China. *Science of the Total Environment*, **512-513**, 143, **2015**.
2. BHATIA A., SINGH S., KUMAR A. Heavy metal contamination of soil, irrigation water and vegetables in peri-urban agricultural areas and markets of delhi. *Water Environment Research*, **87** (11), 2027, **2015**.
3. LIN Y., HAN P., HUANG Y., YUAN G.L., GUO J.X., LI J. Source identification of potentially hazardous elements and their relationships with soil properties in agricultural soil of the Pinggu district of Beijing, China: Multivariate statistical analysis and redundancy analysis. *Journal of Geochemical Exploration*, **173** (2), 110, **2017**.
4. GIRI, S., SINGH A.K. Risk assessment, statistical source identification and seasonal fluctuation of dissolved metals in the subarnarekha river, india. *Journal of Hazardous Materials*, **265** (jan.30), 305, **2014**.
5. LU A.X., WANG J.H., Qin X.Y., WANG K.Y., HAN P., ZHANG S.Z. Multivariate and geostatistical analyses of the spatial distribution and origin of heavy metals in the agricultural soils in Shunyi, Beijing, China. *Science of the Total Environment*, **425**, 66, **2012**.
6. LUO X.S., XUE Y., WANG Y.L., CANG L., XU B., DING J. Source identification and apportionment of heavy metals in urban soil profiles. *Chemosphere*, **127** (may), 152, **2015**.
7. GHREFAT H., YUSUF N. Assessing Mn, Fe, Cu, Zn, and Cd pollution in bottom sediments of Wadi Al-Arab Dam, Jordan. *Chemosphere*, **65** (11), 2114, **2006**.
8. LV J.S., LIU Y., ZHANG Z.L., DAI J.R. Factorial kriging and stepwise regression approach to identify environmental factors influencing spatial multi-scale variability of heavy

- metals in soils. *Journal of Hazardous Materials*, **261**, 387, **2013**.
9. QIN F., JI H.B., LI Q., GUO X.Y., TANG L., FENG J.G. Evaluation of trace elements and identification of pollution sources in particle size fractions of soil from iron ore areas along the Chao River. *Journal of Geochemical Exploration*, **138**, 33, **2014**.
 10. DUODU G.O., GOONETILLEKE A., AYOKO G.A. Comparison of pollution indices for the assessment of heavy metal in Brisbane River sediment. *Environmental Pollution*, **219** (dec), 1077, **2016**.
 11. KAPUSTA P., SOBCZYK L. Effects of heavy metal pollution from mining and smelting on enchytraeid communities under different land management and soil conditions. *Science of the Total Environment*, **536** (dec.1), 517, **2015**.
 12. LV J.S., LIU Y., ZHANG Z.L., DAI J.R., DAI B., ZHU Y.C. Identifying the origins and spatial distributions of heavy metals in soils of Ju country (Eastern China) using multivariate and geostatistical approach. *Journal of Soils and Sediments*, **15** (1), 163, **2015**.
 13. KOTTEK M., GRIESER J., BECK C., RUDOLF B., RUBEL F. World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift*, **15** (3), 259, **2006**.
 14. OLIVER M.A., WEBSTER R. Kriging: A method of interpolation for geographical information systems. *International Journal of Geographical Information Systems*, **4** (3), 313, **1990**.
 15. WILDING L.P. Spatial variability: its documentation, accomodation and implication to soil surveys. *Soil spatial variability*, Wageningen Academic Publishers, Wageningen, The Netherlands, pp. 166-194, **1985**.
 16. DUDHAGARA D.R., RAJPARA R.K., BHATT J.K., GOSAI H.B., SACHANIYA B.K., DAVE B.P. Distribution, sources and ecological risk assessment of PAHs in historically contaminated surface sediments at Bhavnagar coast, Gujarat, India. *Environmental Pollution*, **213** (jun), 338, **2016**.
 17. WEI J., CEN K. Assessment of human health risk based on characteristics of potential toxic elements (PTEs) contents in foods sold in Beijing, China. *Science of the Total Environment*, **703**, 134747, **2020**.
 18. JAN F.A., ISHAQ M., IHSANULLAH I., ASIM S.M. Multivariate statistical analysis of heavy metals pollution in industrial area and its comparison with relatively less polluted area: A case study from the City of Peshawar and district Dir Lower. *Journal of Hazardous Materials*, **176** (1-3), 609, **2010**.
 19. WANG Y., GUO G., ZHANG D., LEI M. An integrated method for source apportionment of heavy metal(loid)s in agricultural soils and model uncertainty analysis. *Environmental Pollution*, **276**, 116666, **2021**.
 20. ZHANG Y.R., GUI H.R., HUANG Y.H., YU H., LI J., WANG M.C., FANG H.X., JIANG Y.Q., WANG C.L., CHEN C. Characteristics of soil heavy metal contents and its source analysis in affected areas of luning coal mine in huaibei coalfield. *Polish Journal of Environmental Studies*, **30** (2), 1465, **2021**.
 21. OTHMAN F., CHOWDHURY M.S.U., JAAFAR W., FARESH E., SHIRAZI S.M. Assessing risk and sources of heavy metals in a tropical river basin: A case study of the Selangor river, Malaysia. *Polish Journal of Environmental Studies*, **27** (4), 1659, **2018**.
 22. CAI L.M., XU Z.C., REN M.Z., GUO Q.W., HU X.B., HU G.C., WAN H.F., PENG P.G. Source identification of eight hazardous heavy metals in agricultural soils of Huizhou, Guangdong Province, China. *Ecotoxicology and Environmental Safety*, **78** (3), 2, **2012**.
 23. JIRIES A.G., NASIR, F.M., BEESE F. Pesticide and heavy metals residue in wastewater, soil and plants in wastewater disposal site near Al-Lajoun Valley, Karak/Jordan. *Water, Air, and Soil Pollution*, **133** (1-4), 97, **2002**.
 24. ZHAO Y.C., WANG Z.G., SUN W.X., HUANG B., SHI X.Z., JI J.F. Spatial interrelations and multi-scale sources of soil heavy metal variability in a typical urban-rural transition area in Yangtze River Delta region of China. *Geoderma*, **156** (3-4), 216, **2010**.
 25. CHEN W.P., LU S.D., PENG C., JIAO W.T., WANG M. Accumulation of Cd in agricultural soil under long-term reclaimed water irrigation. *Environmental Pollution*, **178** (jul), 294, **2013**.
 26. ZARE-MEHRJARDI M., TAGHIZADEH-MEHRJARDI R., AKBARZADEH A. Evaluation of Geostatistical Techniques for Mapping Spatial Distribution of Soil PH, Salinity and Plant Cover Affected by Environmental Factors in Southern Iran. *Notulae Scientia Biologicae*, **2** (4), 92, **2010**.
 27. LIU Z.P., SHAO M.A., WANG Y.Q. Spatial patterns of soil total nitrogen and soil total phosphorus across the entire Loess Plateau region of China. *Geoderma*, **197-198**, 67, **2013**.
 28. LIU M.X., YANG Y.Y., YUN X.Y., ZHANG M.M., WANG J. Concentrations, distribution, sources, and ecological risk assessment of heavy metals in agricultural topsoil of the Three Gorges Dam region, China. *Environmental Monitoring and Assessment*, **187** (3), 147, **2015**.
 29. SUN L.H., FENG S.B., F. Heavy metals in the surface soil around a coalmine: Pollution assessment and source identification. *Polish Journal of Environmental Studies*, **28** (4), 2717, **2019**.
 30. LEE H.G., KIM H.K., NOH H.J., BYUN Y.J., CHUNG H.M., KIM J.I. Source identification and assessment of heavy metal contamination in urban soils based on cluster analysis and multiple pollution indices. *Journal of Soils and Sediments*, **20** (12), 1, **2020**.

