# Multivariate statistics and contamination factor to identify trace elements pollution in soil around Gerga City, Egypt

Ibrahim Said, Salman A. Salman and Ahmed A. Elnazer[*]

## Abstract

**Background:** Gerga district contains different activities, urban, agriculture, and industry, which can impact adversely on the soil quality. Sixteen samples of the agricultural soil (0–30-cm depth) were collected to investigate the pollution of soil with trace elements (Co, Ni, Pb, and Mn). The statistical techniques were applied to discriminate the sources of these elements.

**Results:** The studied soil ranged from uncontaminated to moderately contaminated with the studied trace elements based on the contamination factor index. The statistical analyses indicated the anthropogenic source of Co, Ni, and Pb as well as the natural source of Mn.

**Conclusions:** The statistical analyses assisted in the discrimination of natural and anthropogenic sources of trace elements in the investigated soil samples. Mn is mainly of natural origin, affected by pedogenic factor, whereas traffic emissions and phosphate fertilizer, as well as domestic activities, are relevant sources of Co, Ni, and Pb elements in the studied soil. Consequently, the recommendation is periodic environmental monitoring and minimizing the fertilization rate.

**Keywords:** Trace elements, Soil contamination, Multivariate statistics, Gerga, Upper Egypt

## Introduction

Trace element contamination has received considerable attention due to their negative impact on the human health and environment (Adriano 2001). The natural and anthropogenic inputs enrich the soil with trace elements. Pollution occurs when an element quantity excess its background concentrations (Kabata-Pendias 2010). The elements that come from anthropogenic sources are generally more bioavailable than pedogenic and lithogenic ones (Kabata-Pendias 1993 and Kobierski and Dabkowska-Naskret 2012). Pedogenic metals are of lithogenic and anthropogenic origin, but their distribution in soil profiles changes due to mineral transformation and other pedogenic processes (Kobierski and Dabkowska-Naskret 2012). Understanding the pollutant's sources and distributions is among the most critical concerns for environmental management and decision-making (Sun et al. 2013).

Unfortunately, many researchers recorded soil pollution with trace elements in many parts of Egypt: Aswan (Darwish and Pöllmann 2015), Assiut (Asmoay 2017), Helwan (Said 2015), Kafr El-Sheikh (Naggar et al. 2014), Sohag (Salman 2013), and (Salman et al. 2017). Salman et al. 2017 pointed out the accumulation of trace elements in the food chain (Egyptian clover) in Sohag. The presence of such metals in the food chain can cause an adverse impact on human beings (Karim et al. 2015).

Many researchers used statistical analyses as powerful tools in geo-environmental studies (Zhang et al. 2009; Sundaray et al. 2011; Ming-Kai et al. 2013; Kelepertzis 2014; Simu et al. 2016; and Guo et al. 2017). Statistical analysis is a useful tool for assessing the possible sources of pollutants because it allows for consideration of cause-and-effect relationships, highlighting exceedances. Contamination indices also help in the understanding of ecological status. Contamination factor (CF) is employed to evaluate the level of soil contamination and to infer anthropogenic inputs from the natural one. In the present study, we applied both statistical analysis and

* Correspondence: nazer42@yahoo.com
Geological Sciences Department, National Research Centre, POB: 12622, 33 El Bohouth St. (former El Tahrir St.) - Dokki, Giza, Egypt

Said *et al. Bulletin of the National Research Centre*        (2019) 43:43

Page 2 of 6

contamination factor in an attempt to distinguish anthropic sources from the natural one and to evaluate the level of soil contamination.

## Materials and methods

Gerga district is one of the most important agricultural areas in Sohag governorate, Egypt. It contains one of the biggest sugar factories in Egypt. This industry was reported as a pollution source of soil with different chemicals (Zaki et al. 2015). It extends between longitudes 31° 46′–31° 55′ E and latitudes 26° 12′–26° 22′ N (Fig. 1). Geologically, it consists of Quaternary deposits and floodplain sediments of the River Nile (Said 1990). Due to soil degradation since the construction of the High Dam in 1968, more fertilizer has been applied to restore soil fertility. P and N fertilizers are the major used agrochemicals in the Gerga area. Roadways are crossing the cultivated lands of the study area. The traffic emissions besides application of fertilizers lead to the accumulation of trace elements which threaten agricultural soils and in turn the human health.

Sixteen agricultural soil samples (0–30-cm depth) were collected randomly as accessible (Fig. 1). ArcGIS10.2 (Desktop 2014) was used to prepare the sample map. Total trace elements were determined by digestion with 3 HCl:1 HNO$_3$ mixture and analyzed using the atomic absorption spectrophotometer (Buck Scientific 205AA). The pH in soil was measured in 1:1 soil to water ratio by using the HANNA (HI93300) combined electrode. Calcium carbonate percentage (CaCO$_3$%) and phosphorous (P) were estimated by the titrimetric and colorimetric methods, respectively. Soil organic matter percentage (SOM %) was determined according to the modified Walkley and Black method (USDA 2004).
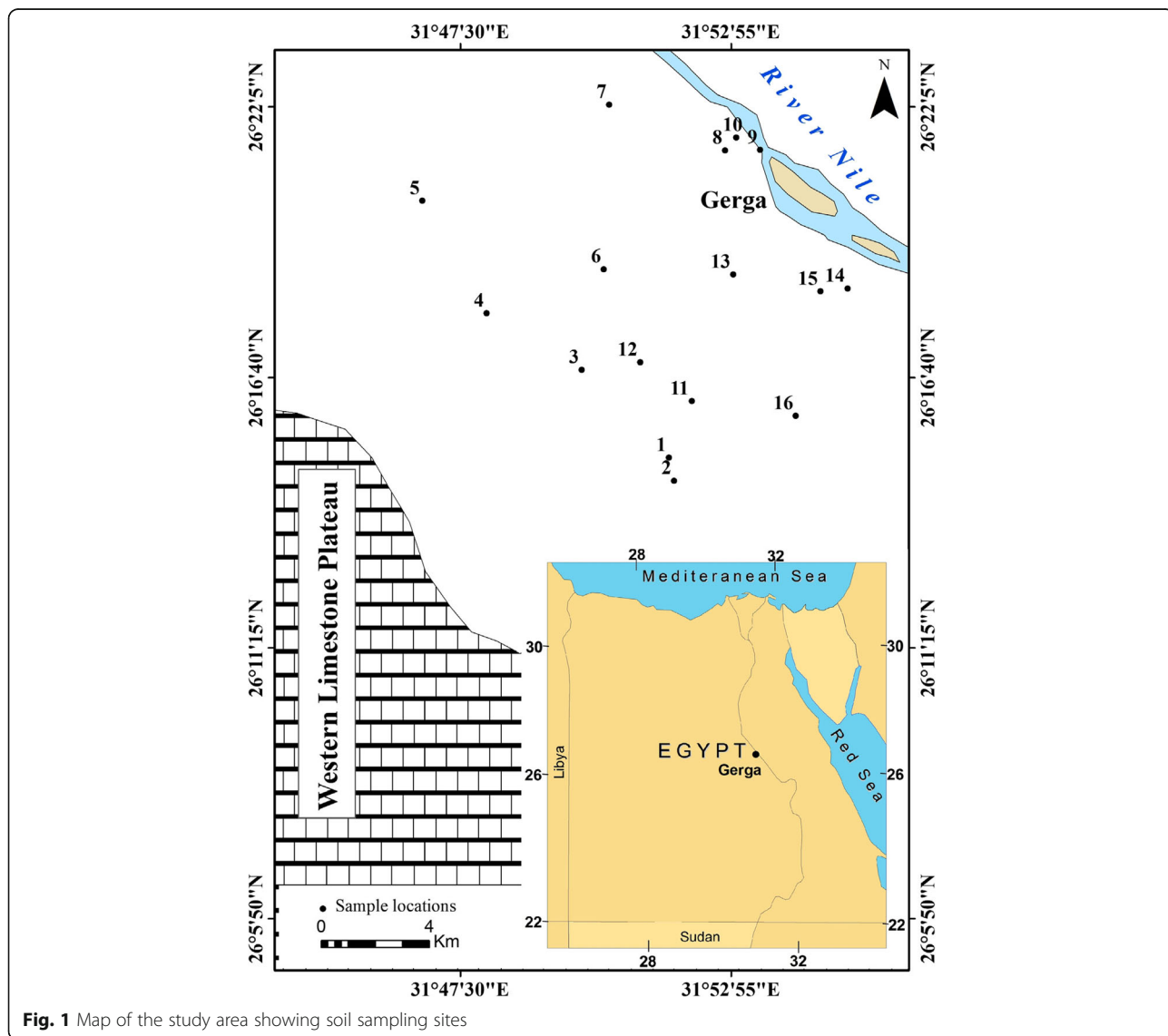


**Fig. 1** Map of the study area showing soil sampling sites

Contamination level was assessed using the contamination factor (CF) recognized in (Hakanson 1980) based on the following equation:

$$CF = C_s/C_b$$

where $C_s$ is the concentration of metal in the study samples and $C_b$ is the baseline concentration. Baseline concentrations as reported by (Turekian and Wedepohl 1961) were used as $C_b$ during this study (Mn = 850 ppm, Co = 19 ppm, Ni = 68 ppm, and Pb = 20 ppm). Hakanson (1980) classified the contamination factor as follows: CF < 1 low, 1 to < 3 moderate, 3 to < 6 considerable, and > 6 high contamination.

Although the number of samples is relatively low, cluster analysis (CA) and principal component analysis (PCA) were conducted to take into account the complicated environmental situation of the study area. Several sources of contamination and several processes are influencing the occurrence of trace elements in the investigated soil. The minimum sample size recommended for conducting principle component is debatable in the literature. Generally, as the sample size increases, sampling error is reduced. Survey of literature relating to the minimum sample size used in principle component studies exhibited wide range of variation from 2 or less to 20 times the number of variables (Lingard and Rowlinson 2006). However, it is fair to say that no absolute rules can exist (Lingard and Rowlinson 2006). MacCallum et al. (1999) report that when data are strong the impact of sample size is greatly reduced. Strong data is data in which item commonalities that are consistently high factors exhibit high loadings (≥ 0.8) on a substantial number of items (at least three or four) and the number of factors is small. According to Guadagnoli and Velicer (1988), if components possess four or more variables with loadings above 0.60, the pattern may be interpreted whatever the sample size used. In a word, with high loadings, any sample size is okay. The present data possess seven variables with loadings of 0.712. The statistical analysis was performed using SPSS 16.0 software. Descriptive statistical analysis (minimum, maximum, mean, and coefficient of variation) of the soil physico-chemical characteristics and element contents was performed as a first step towards an initial understanding of their distribution. And then, multivariate statistics were performed including principal component analysis (PCA) and cluster analysis (CA). PCA and CA were employed in an attempt to identify the common sources of the trace elements in the studied soil. PCA was implemented by means of the varimax rotation method which helps reduce the number of variables in fewer high loading components and facilitates their interpretation (Chen et al. 2008). Kaiser-Meyer-Olkin measure of sampling

adequacy (KMO MSA) for the set of variables included in the analysis was 0.712. It exceeds the minimum requirement of 0.50 for overall MSA, with Bartlett's test of sphericity (0.00), be less than the level of significance (Tabachnick and Fidell 2007). PCA was developed based on the Ward method using squared Euclidean distances (z-transformation) as a measure of similarity between samples based on their element content (Co, Ni, Pb, Mn, and P). The clustering results were provided in a hierarchical cluster (dendrogram).

## Results

Table 1 illustrates the statistical summary of the analytical data. The measured sand, silt, clay, pH, $CaCO_3$, OM, P, Co, Ni, Pb, and Mn values were 68.8%, 12.4%, 18.9%, 8.5, 2.8%, 2.1%, 0.3%, 12.6 ppm, 46.8 ppm, 11.9 ppm, and 990.6 ppm, respectively. P, Co, Ni, and Pb showed marked flocculation (C.V = 92.7%, 82.9%, 52.8%, and 92.2% respectively), whereas Mn exhibited a uniform distribution pattern (C.V = 15.4%). The calculated CF for Co, Ni, Pb, and Mn were around 0.7, 0.7, 0.6, and 1.2, respectively (Table 2). Generally, the studied samples were low contaminated with Co, Ni, and Pb and moderately contaminated with Mn according to CF.

Since the studied soil was found to be relatively enriched in some trace elements compared to those reported in Turekian and Wedepohl )1961(, PCA was performed to identify natural and anthropogenic sources. Two principal components were extracted from the investigated soil data, explaining 75.94% of the variance (Table 3). The first component (PC1) explains 48.74% and accounts for the majority of the variance in the dataset and includes the elements (Co, Ni, Pb, and P) of high variation coefficients. The second component (PC2) is responsible for 27.16% of the total variance and shows significant positive loadings for Mn, carbonate, and pH.

**Table 1** Descriptive statistical analysis of the studied soil data

| Soil parameters | Min | Max | Mean | Std. deviation | Std. error | CV% |
|---|---|---|---|---|---|---|
| Sand% | 60.0 | 76.7 | 68.8 | 4.9 | 1.2 | 7.2 |
| Silt% | 8.0 | 16.0 | 12.4 | 2.6 | 0.6 | 20.6 |
| Clay% | 15.2 | 23.9 | 18.9 | 2.6 | 0.65 | 13.8 |
| pH | 8.1 | 8.8 | 8.5 | 0.23 | 0.06 | 2.7 |
| $CaCO_3$% | 1.3 | 5.4 | 2.8 | 2.5 | 0.3 | 41.5 |
| OM% | 1.2 | 3.0 | 2.1 | 0.6 | 0.2 | 29.0 |
| P% | 0.1 | 0.9 | 0.3 | 0.2 | 0.1 | 92.7 |
| Co (ppm) | 0.01 | 28.1 | 12.6 | 10.5 | 2.6 | 82.9 |
| Ni (ppm) | 0.01 | 81.2 | 46.8 | 24.8.8 | 6.2 | 52.9 |
| Pb (ppm) | 0.001 | 29.2 | 11.9 | 11.0 | 2.8 | 92.2 |
| Mn (ppm) | 685.0 | 1317.0 | 990.6 | 153.0 | 38.2 | 15.4 |

Said et al. Bulletin of the National Research Centre     (2019) 43:43

Page 4 of 6

**Table 2** Contamination factor (CF) of the studied elements (italic numbers refer to the moderate contaminated samples)

| Sample no. | Co | Ni | Pb | Mn |
|---|---|---|---|---|
| 1 | 0.00 | 0.07 | 0.00 | 1.46 |
| 2 | *1.11* | 0.96 | 0.96 | 1.25 |
| 3 | *1.48* | *1.15* | *1.37* | 1.10 |
| 4 | *1.35* | *1.09* | *1.46* | 1.33 |
| 5 | 0.55 | 0.65 | 0.11 | 1.16 |
| 6 | 0.56 | 0.78 | 0.42 | 1.20 |
| 7 | 0.97 | 0.89 | 0.74 | 1.09 |
| 8 | 0.76 | 0.71 | 0.46 | 1.03 |
| 9 | 0.76 | 0.70 | 0.73 | 0.81 |
| 10 | 0.00 | 0.37 | 0.00 | 1.08 |
| 11 | 0.00 | 0.00 | 0.00 | 1.07 |
| 12 | 0.00 | 0.36 | 0.42 | 1.55 |
| 13 | *1.43* | *1.19* | *1.45* | 1.11 |
| 14 | 0.23 | 0.53 | 0.08 | 1.23 |
| 15 | 0.19 | 0.51 | 0.13 | 1.20 |
| 16 | *1.23* | *1.03* | *1.25* | 0.98 |
| Average | 0.66 | 0.69 | 0.60 | 1.17 |

## Discussion

The texture of the studied soil was found to be muddy sand. Carbonate and sand content increased westwards near the limestone plateau. The soil contains low OM% as a result of study area aridity and agriculture practices (seasonal tilling) that oxidize the organic matter. The calculated contamination factor (CF) indicated that the investigated soil ranged from uncontaminated to moderately contaminated with the studied metals (Table 2). Moderately contaminated sites (2, 3, 4, 13, and 16) were most affected by anthropogenic inputs. Such an anthropic portion can cause serious environmental

**Table 3** Principal component loadings of soil data including variance % and cumulative %

| Variable | Component | |
|---|---|---|
| | PC1 | PC2 |
| Co | *0.974* | − 0.028 |
| Ni | *0.949* | − 0.018 |
| Pb | *0.951* | 0.111 |
| Mn | − 0.339 | *0.811* |
| P | *0.619* | 0.103 |
| CaCO₃ | 0.383 | *0.732* |
| pH | 0.117 | *0.829* |
| Variance % | 48.746 | 27.169 |
| Cumulative % | 48.746 | 75.942 |

Extraction method: principal component analysis. Rotation method: varimax with Kaiser-Meyer-Olkin normalization (KMO = 0.712 (middling)). Rotation converged in three iterations

hazards. Now, low contaminated sites are not a serious environmental concern. Mn is derived mainly from the parent rocks of the Ethiopian plateau (Omer 1996).

The statistical analyses indicated the presence of two sources of metals in the studied soils: natural and anthropogenic. The high variance of elements loaded in PC1 and the anthropogenic marker of Pb can denote the anthropogenic source of PC1. Pb is usually accounted as a marker element of traffic activities; it is originated from leaded gasoline (Elnazer et al. 2015 and Wang et al. 2017). Co comes from tires, and Ni results from brake wear, engine oil leakage, tire wear, and road abrasion (Winther and Slentø 2010). Also, the high positive loadings of Co, Ni, and Pb with P imply the role of fertilization in the distribution of these elements in the soil. The field investigation indicated the uncontrolled application of NP fertilizers, herbicides, and pesticides in the study area. These substances are considered the principal source of trace elements worldwide. The applied P fertilizers in the study area contain about 74.9 and 15.4 ppm of Pb and Co, respectively (Salman et al. 2017). On the other hand, the load of Mn (with uniform distribution pattern), CaCO₃, and pH in PC2 indicated this component is of a natural origin influenced by pedogenic factor. Organic matter content and texture had no significant influence on trace element distribution.

The hierarchical cluster divided the studied samples into two groups (A and B) based on the levels and sources of contamination (Fig. 2). Group A includes samples 2, 3, 4, 7, 8, 9, 13, and 16 (Fig. 2). These samples were relatively had higher contamination level with the studied elements (Table 2) with P content averaging 0.32%. Group B includes the remaining samples of lower trace elements with lower P content averaging 0.19%. The link between P content and levels of elements indicated the influence of fertilization process. The role of traffic activities on the pollution of soil in the study area is evidenced by the occurrence of contaminated sites adjacent to the roadway sides (group A samples). Sample no. 4 in group A is located near a residential area and is affected by domestic activities rather than road traffic emissions (Fig. 1). This is supported by the occurrence of sample no. 4 far from the remaining samples controlled by traffic emissions (Fig. 2). The hierarchical cluster confirms the previously suggested anthropogenic factor (PC1).

## Conclusion

Like all Egyptian cities, roadways passing through the farmland of Gerga area, Upper Egypt, and large amounts of agrochemicals are applied to restore soil fertility. Such activities have led to the accumulation of trace elements and threat the agricultural soils. The studied soil varied from uncontaminated to moderately contaminated with

**Fig. 2** Dendrogram providing a graphic summary of the clustering processes
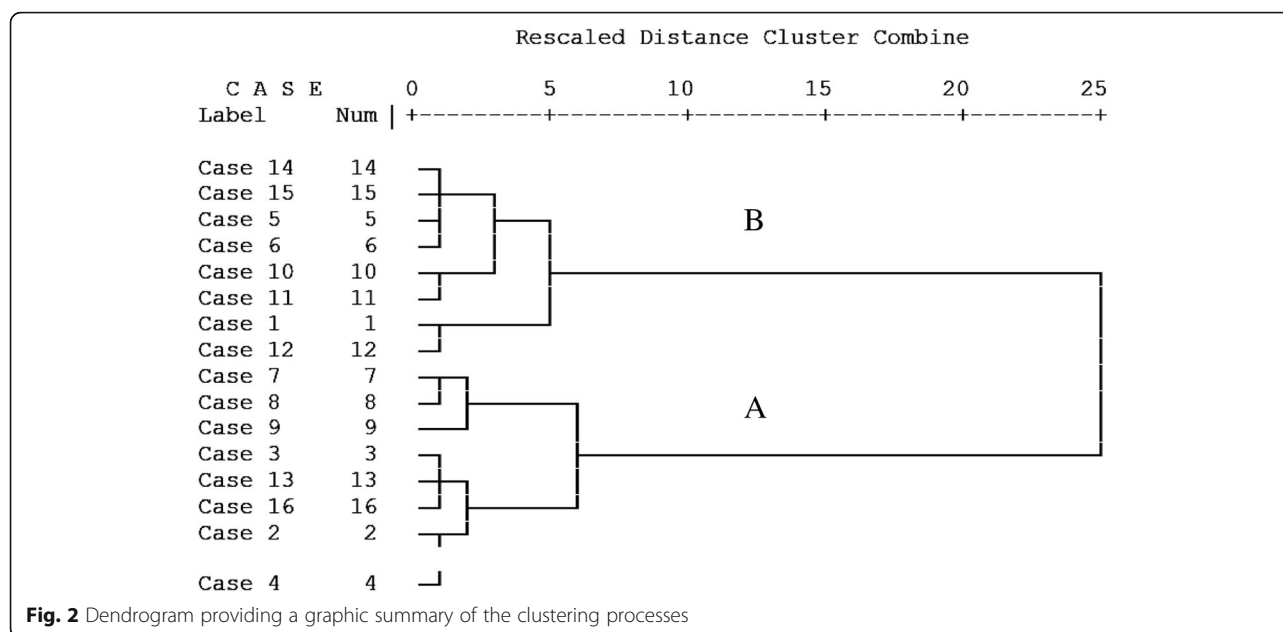
trace elements. Low contaminated soils are not currently a serious environmental concern. Moderately contaminated soil indicates a presence of slightly non-residual portion having a tendency to become bio-available. Mn was statistically proven as from natural origin affected by pedogenic processes, indicated by its homogeneous distribution and high loading with pH. On the other hand, traffic emissions and phosphate fertilizers were labeled as important sources of Pb, Co, and Ni elements in the soil. The matter was evidenced by the occurrence of contaminated sites adjacent to the roadway sides with their relative higher phosphorus content. Hence, periodic environmental monitoring is recommended; planting of crops sensitive to Pb, Co, and Ni must be avoided; and fertilization rate should be minimized.

### Availability of data and materials
Not applicable.

### Authors' contributions
Dr. Ibrahim Said has a 50% contribution to the manuscript. Dr. Salman A. Salman has a 25% contribution to the manuscript. Dr. Ahmed A. Elnazer has a 25% contribution to the manuscript. All authors read and approved the final manuscript.

### Authors' information
Not applicable.

### Ethics approval and consent to participate
Accepted.

### Consent for publication
Accepted.

### Competing interests
The authors declare that they have no competing interests.

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

### References
Adriano DC (2001) Trace elements in the terrestrial environment Springer-Verlag New York Google Scholar

Al NY, Naiem E, Mona M et al (2014) Metals in agricultural soils and plants in Egypt. Toxicol Environ Chem 96:730–742

Asmoay ASA (2017) Hydrogeochemical studies on the water resources and soil characteristics in the western bank of the River Nile between Abu Qurqas and Dayr Mawas, El Minya Governorate, Egypt. Al-Azhar University

Chen T, Liu X, Zhu M et al (2008) Identification of trace element sources and associated risk assessment in vegetable soils of the urban–rural transitional area of Hangzhou, China. Environ Pollut 151:67–78

Darwish MAG, Pöllmann H (2015) Trace elements assessment in agricultural and desert soils of Aswan area, south Egypt: geochemical characteristics and environmental impacts. J African Earth Sci 112:358–373

Desktop EA (2014) Release 10.2. 2. Redlands, CA, USA Environ Syst Res Inst

Elnazer AA, Salman SA, Seleem EM, Abu El Ella EM (2015) Assessment of some heavy metals pollution and bioavailability in roadside soil of Alexandria-Marsa Matruh Highway, Egypt. Int J Ecol 2015:7 https://doi.org/10.1155/2015/689420

Guadagnoli E, Velicer WF (1988) Relation of sample size to the stability of component patterns. Psychol Bull 103:265

Guo L, Zhao W, Gu X et al (2017) Risk assessment and source identification of 17 metals and metalloids on soils from the half-century old tungsten mining areas in Lianhuashan, Southern China. Int J Environ Res Public Health 14:1475

Hakanson L (1980) An ecological risk index for aquatic pollution control. A sedimentological approach. Water Res 14:975–1001

Kabata-Pendias A (1993) Behavioural properties of trace metals in soils. Appl Geochem 8:3–9

Kabata-Pendias A (2010) Trace elements in soils and plants. CRC press

Said *et al. Bulletin of the National Research Centre*        (2019) 43:43

Page 6 of 6

Karim Z, Qureshi BA, Mumtaz M (2015) Geochemical baseline determination and pollution assessment of heavy metals in urban soils of Karachi, Pakistan. Ecol Indic 48:358–364

Kelepertzis E (2014) Accumulation of heavy metals in agricultural soils of Mediterranean: insights from Argolida basin, Peloponnese, Greece. Geoderma 221:82–90

Kobierski M, Dabkowska-Naskret H (2012) Local background concentration of heavy metals in various soil types formed from glacial till of the Inowroclawska Plain. J Elem 17

Lingard HC, Rowlinson S (2006) Sample size in factor analysis: why size matters. Hong Kong University, Hong Kong

MacCallum RC, Widaman KF, Zhang S, Hong S (1999) Sample size in factor analysis. Psychol Methods 4:84

Ming-Kai QU, Wei-Dong LI, ZHANG C-R et al (2013) Source apportionment of heavy metals in soils using multivariate statistics and geostatistics. Pedosphere 23:437–444

Omer AA (1996) Geological, mineralogical and geochemical studies on the Neogene and Quaternary Nile basin deposits, Qena-Assiut stretch, Egypt. Ph. D. thesis, Geology Dept. Faculty of Science, Sohag, South Valley University

Said I (2015) Geochemical speciation and enrichment of toxic heavy metals in Nile sediments and soils in El-Tebbin area, Egypt. Ain Shams University, Egypt

Said R (1990) The geology of Egypt: AA Balkema. Rotterdam/Brookfield

Salman SA (2013) Geochemical and environmental studies on the territories West River Nile, Sohag Governorate- Egypt. Al-Azhar University

Salman SA, Elnazer AA, El Nazer HA (2017) Integrated mass balance of some heavy metals fluxes in Yaakob village, south Sohag, Egypt. Int J Environ Sci Technol 14:1011–1018

Simu SA, Uddin MJ, Majumder RK et al (2016) Multivariate statistical analysis of trace elements in soil of Gazipur industrial area, Bangladesh. Univers J Environ Res Technol 6

Sun G, Chen Y, Bi X et al (2013) Geochemical assessment of agricultural soil: a case study in Songnen-Plain (Northeastern China). Catena 111:56–63

Sundaray SK, Nayak BB, Lin S, Bhatta D (2011) Geochemical speciation and risk assessment of heavy metals in the river estuarine sediments—a case study: Mahanadi basin, India. J Hazard Mater 186:1837–1846

Tabachnick BG, Fidell LS (2007) Using multivariate statistics. Allyn & Bacon/Pearson Education

Turekian KK, Wedepohl KH (1961) Distribution of the elements in some major units of the earth's crust. Geol Soc Am Bull 72:175–192

USDA N (2004) Soil survey laboratory methods manual, vol 42. Soil Survey Investigations Report

Wang G, Zeng C, Zhang F, Zhang Y, Scott CA, Yan X (2017) Traffic-related trace elements in soils along six highway segments on the Tibetan Plateau: influence factors and spatial variation. Sci Total Environ 581:811–821

Winther M, Slentø E (2010) Heavy metal emissions for Danish road transport. National Environmental Research Institute Denmark

Zaki R, Ismail EA, Mohamed WS, Ali AK (2015) Impact of surface water and groundwater pollutions on irrigated soil, El Minia Province, northern Upper Egypt. J Water Resour Prot 7:1467

Zhang XY, Lin FF, Wong MTF et al (2009) Identification of soil heavy metal sources from anthropogenic activities and pollution assessment of Fuyang County, China. Environ Monit Assess 154:439