APPLICATION OF MULTIVARIATE AND GEOSTATISTICAL APPROACHES IN ANALYSIS AND ASSESSMENT OF HEAVY METAL SOURCES IN SOIL OF WASTE DISPOSAL SITE AT KHULNA

Sanjida Khair[1] and Islam M. Rafizul[2]

[1]Post graduate Student, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh, email: priyo.0901091@gmail.com
[3]Associate Professor, Khulna University of Engineering & Technology, Khulna-9203, Bangladesh, email: imrafizul@yahoo.com

ABSTRACT

The main focused of the study was to identify the possible sources of contamination of metal elements in soil of waste disposal site. To theses endeavor, sixty soil samples were collected at a depth of 0-30 cm from the existing ground surface from a selected waste disposal site at Rajbandh, Khulna, Bangladesh. In the laboratory, the concentrations of metal elements of Al, As, Ba, Ca, Cd, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Ni, Pb, Sb, Sc, Sr, Ti, V and Zn in soil were measured. Descriptive and multivariate statistics including Pearson's correlation, principal component analysis (PCA) and agglomerative hierarchical clustering (AHC) were used. In addition, Inverse distance weighting (IDW) with power of 1-5 through ArcGIS were performed. Results of descriptive statistics reveals that concentrations of metal elements followed almost same pattern during both the dry and rainy seasons. Results of Pearson's correlation depicts that the sources of metal elements were almost the same and these metal elements might be derived from the waste accumulation activity. In addition, PCA reported that generation of Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V from anthropogenic activities, while, Cu, Hg, Mn, Pb, and Zn from natural sources and Cr from both sources in dry season. In rainy season, Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V generated from anthropogenic activities whereas As, Cr, Hg, Mn, K and Na from natural sources and Cd, Pb and Zn from both sources.These results is in agreement with output obtained from AHC. In addition, patial distribution of waste disposal site epitomized most contaminated hotspots were found to be the nearest soil sampling point with respect to the central point of the selected disposal site.

Keywords: *Disposal site, Multivariate analysis, spatial distribution, Inverse Distance Weighting, Khulna*

1. INTRODUCTION

In recent periods, existence of metal elements in soil has become a foremost concern that arise necessities of monitoring the endangerment of soil by contamination of metal elements. Non-biodegradability characteristics and elongated biological half-lives of metal elements for abolition, their accretion in nutrition chain will obligate a substantial effect on all living element. The huge quantities of municipal solid waste (MSW) in waste disposal site going through distinct biological, physical and chemical processes for decomposition, produced leachate and contaminated soil which creates vulnerable to the environmental components and nearby inhabitants (Nriagu and Pacyna, 1988). The metal elements in soil are either derived from natural parent rock materials or anthropogenic activities such as urban-industrial development, landfill management, vehicular emissions, fossil fuel combustion and agricultural practices (Tahir et al., 2007). Contamination of soil occurs when the presence of toxic chemicals, pollutants or contaminants from fertilizers, organic wastes, organic pesticides, with high concentrations in soil (Jia et al., 2010). Ingestion of food grown in contaminated soil and intake of water from contaminated water bodies cause great risk to plants, wildlife, humans and of course for the soil itself. The application of multivariate statistical approaches such as principal component analysis (PCA), agglomerative hierarchical clustering (AHC) permit a better technique for classification, modeling and interpretation of soil monitoring data (Stanimirova et al., 2006). In addition, spatial distribution is essential for assessing the effect of metal elements in soil and to delineate contamination zones using geographic information system (GIS) techniques (Omran and Razek, 2012).

Khulna is one of the fast growing commercial cities in Bangladesh. Most of the MSW are disposed in waste disposal site at Rajbandh, the only authorized waste disposal site of Khulna city. Due to inadequate management practices and the low standard sanitary landfill, the leachate percolates into the groundwater and contaminates the groundwater source which is a potential threat to next generation. Thus, necessities arise to take steps for proper disposal of MSW as well as maintenance of MSW disposal site in Khulna. To theses endeavour, sixty soil samples were collected at a depth of 0-30 cm from the existing ground surface and the relevant metal elements of Al, As, Ba, Ca, Cd, Co, Cr, Cu, Fe, Hg, K, Mn, Na, Ni, Pb, Sb, Sc, Sr, Ti, V and Zn were measured in the laboratory. Descriptive and multivariate statistics including Pearson's correlation, PCA and AHC were used to ascertain the possible pollutant sources of metal elements and their correlation. In addition, Inverse distance weighting (IDW) with power 1 to 5 were used to interpolate the concentrations of metal elements of unsampled locations and provide soil map to sensitize their spread over the study area using ArcGIS software. The purpose of the study are to (i) develop correlations of metal element and their possible sources of contamination in soil of waste disposal site (ii) visualize the level of contamination and distribution of metal element spatially in the soil of the waste disposal site.

2. METHODOLOGY

The sampling of soil samples, laboratory investigations and application of various multivariate and Geostatistical approaches are presented and hence discussed in the following articles.

2.1 Soil Sampling

In this study, sixty soil samples (forty for dry season and twenty for rainy season) were collected at a depth of 0-30 cm from the existing ground surface of the waste disposal site. Moreover, the sampling points were selected maintaining gradual addition of about 10 m distance from the first borehole (BH-1) located at the centre of the waste disposal site by the subsequent boreholes. In contrast, the first borehole of rainy season (BH-41) is about 30 m apart from BH-1 which is the centre of the site and maintains a gradual addition of about 15 m in selecting other following boreholes. Figure 1 depicted the soil sampling locations in waste disposal site at Rajbandh, red circles indicated sampling points in dry season and blue triangles indicated sampling points in rainy season.

2.2 Laboratory Investigations

10 g of each soil sample was taken into a 100 mL conical flask washed with deionized water prepared by adding 6 mL HNO3/HClO4 acid in ratio 2:1 and left overnight. Each sample was kept into the temperature of 150°C for about 90 minutes followed by 230°C for 30 minutes. Subsequently, HCl solution was added in ratio 1:1 to the digested sample and re-digested again for another 30 minutes. The digested sample was washed into 100 mL volumetric flask and mixture obtained was cooled down to room temperature. The concentrations of metal elements in this digested solution were determined using atomic absorption spectrophotometer (AAS) and the amount of each heavy metal was deduced from the calibration graph and the concentration of the metal elements were reported in mg/kg.

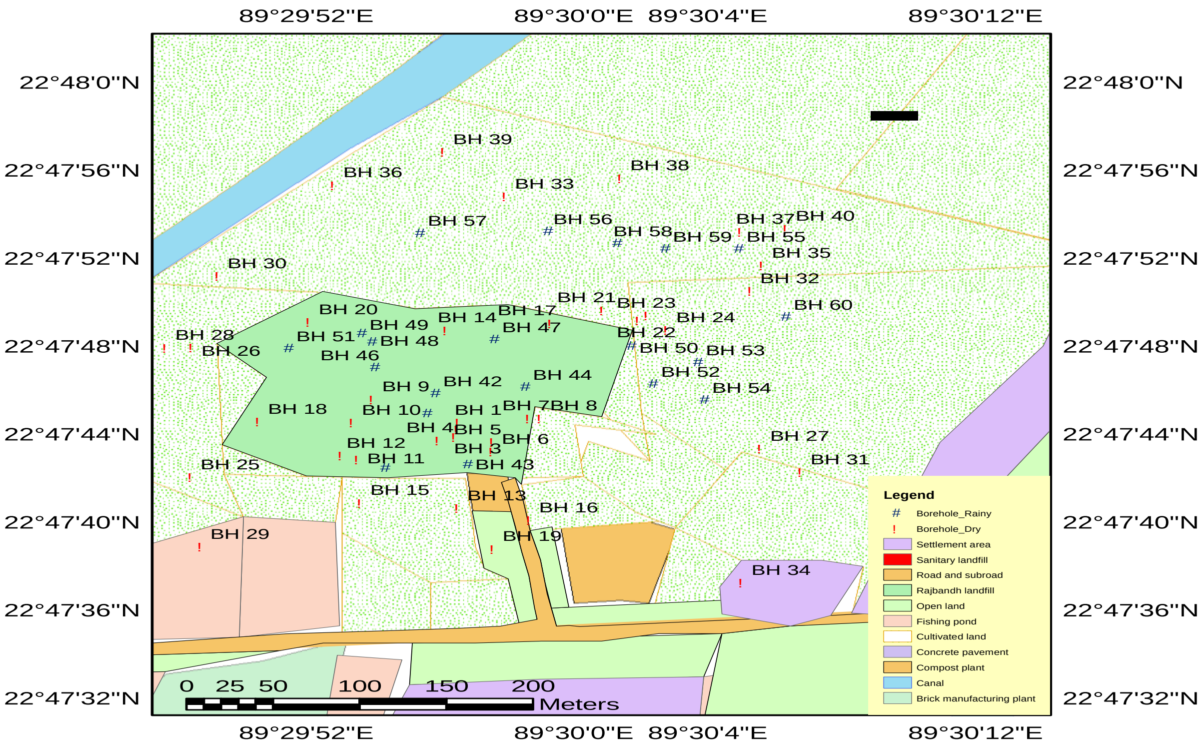


Figure 1: Map showing of soil sampling locations in waste disposal site

2.3 Descriptive Statistics

In this study, the descriptive statistics including normality test such as shapiro-wilk (S-W) test and kolmogorov–smirnov (K-S) test was performed using Statistical Package for the Social Sciences (SPSS) software. In addition, the normal quantile- quantile (QQ) plot was also schemed to check the distribution of data points more accurately. In this study, the conventional statistical parameters in terms of mean, maximum, minimum, median, SD, CV, skewness and kurtosis for two seasons (i.e. dry and rainy) was analysed to check the variability of metal elements due to anthropogenic activities as well as from natural parent materials as well as to show the seasonal variation of metal elements in soil.

2.4 Multivariate Statistics

In this study, the multivariate statistical analysis including Pearson's correlation, principal component analysis (PCA) and Agglomerative hierarchical clustering (AHC) were used and hence described in the following articles.

2.4.1 Pearson's correlation

In this study, Pearson's correlation was performed using XLSTAT to examine the association of metal elements in soil irrespective to their sources. In this study, the value of correlation coefficient, r, was computed using the following Equation 1 considering one dataset $\{x_1,...,x_n\}$ containing n values and another dataset $\{y_1,...,y_n\}$ containing n values.

$$r = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}} \quad (1)$$

In this study, the null hypothesis (H_0) and alternative hypothesis (H_1) of the significance test for correlation was expressed depending on a two-tailed test. Two-tailed significance test are as follows: H_0: r = 0 ("the correlation coefficient is 0, there is no association"); H_1: r > 0 ("the correlation coefficient is not 0, a nonzero correlation could exist").

2.4.2 Principal Component Analysis

PCA is probably the most popular multivariate technique that is used to analyse a dataset of inter-correlated quantitative dependent variables. The principal components (PCs) with variables, the high loadings (eigenvalues) depicted greater importance from the contamination sources, whereas, lower loadings (eigenvalues) point to lower importance with regards the sources of these contaminations (Lee et al., 2006; Zou et al., 2015). In this study, the PCA method was performed sequentially, first by information extraction in the input space (with n-dimensions) to determine the directions of which the input variables display the most substantial variability. The PC coefficients and the eigenvalues $(\lambda_i > 0, i = 1,2, \ldots..n)$ 0 for the correlation matrix $(C = E\{xx^T\})$ with respect to their eigenvectors $(e_i > 0, i = 1,2, \ldots..n)$ 0, is called the loadings were then calculated which gives a new set of variables that explains the variability in the original dataset; the first PCs retained a greater proportion of the total variance, consequently leading to effective and practical dimensionality reduction exercise. the following equation was used through XLSTAT to compute the variance.

$$\left.\begin{aligned} PC1 &= a_1x_1 + a_2x_2 + \cdots \ldots.. + a_nx_n \\ PCn &= \sum_{j=1}^{n} a_{1j}x_j \end{aligned}\right\} \qquad \ldots\ldots\ldots.. (2)$$

where; a_{1j} = eigenvectors obtained from the correlation matrix; x_j = input variables

2.4.3 Agglomerative Hierarchical Clustering

AHC accomplishes successive fusions of data into clusters where each object initially starts out as its own cluster. It also differs to the extent that different measures are employed to measure the distance between clusters. In this study, dendrograms were the output of AHC which display the cluster hierarchy and the distances at which the clusters were joined helpful to select an appropriate number of clusters for the dataset using XLSTAT. Cluster was selected by cutting the dendrogram where there is a significant jump in the distance of the cluster joins which is equivalent to selecting the knee point in a k-Means curve.

2.5 Spatial Distribution of Metal elements

The interpolated map using geostatistics provides the best and simpler way to comprehend spatial distribution of metal elements to recognize the risk of contamination zone depending on the concentrations plotted with the optimal interpolation model. Soil maps convey information about the soil to land users. In the present investigation, the selection of deterministic interpolation technique such as IDW was made based on the extent of similarity with a known scattered set of points. The assigned values to unknown points are calculated with a weighted average of the values available at the known points (Pebesma et al., 2007). With the increase in distance, the weight of interpolation decreases (Gotway et al., 1996) and weighting power that decides how the weight decreases as the distance increases. The value of the assigned power is one of the important factor of IDW interpolation method. In this study, the technique of IDW with integer power of 1 to 5 was performed to give more weight to the closest sampled points. Additionally, the size and number of neighborhood also affect the accuracy of the interpolation techniques (Isaake and Srivastava, 1989). The prediction of unsampled points was measured using the following Equation 3 by ArcGIS software.

$$Z_0 = \frac{\sum_{i=1}^{N} z_i d_i^{-n}}{\sum_{i=1}^{n} d_i^{-n}} \qquad (3)$$

Where, Z_0 is the estimation value of variable z in point i, Z_i is the sample value in point i, di is the distance of sample point to estimated point, N is the coefficient that determines weigh based on a distance., n is the total number of predictions for each validation case. In this study, estimations were made using different integer powers of 1 to 5.

3. RESULTS AND DISCUSSION

The findings of the study are illustrated in the following sections.

3.1 Analysis of Heavy Metal Concentration in Soil

For better result, it was found from normal QQ plot that almost all the metal elements in soil were distributed normally except As for both the dry and rainy seasons. Thus, log transformation was applied to As for normal distribution. The descriptive statistics of heavy metal concentrations in soil of waste disposal site in the dry season is provided in Table 1.

Table 1: Descriptive statistics of metal elements in soil in dry season (n=40)

Metal	Min	Max	Median	Mean	CV (%)	SD	Skewness	Kurtosis
Al	158.35	874.78	458.46	490.25	40.31	197.61	0.303	-0.727
Ca	100.20	318.00	173.19	183.80	33.55	61.67	0.577	-0.793
Cd	2.55	7.03	4.46	4.55	24.99	1.14	0.387	-0.530
Cu	2.92	16.45	4.82	6.20	59.41	3.68	1.541	1.110
Fe	733.19	1987.76	1386.50	1363.94	25.67	350.15	-0.081	-1.199
Hg	1.98	9.20	4.01	4.63	44.63	2.07	0.797	-0.460
K	104.88	460.33	316.37	292.00	35.98	105.06	-0.416	-0.937
Ni	2.56	8.06	4.71	4.83	33.02	1.60	0.409	-0.984
Pb	21.29	90.55	33.94	37.61	36.34	13.67	1.840	4.393
Ti	643.33	1937.3	1223.8	1221.2	33.27	406.26	0.160	-1.198
Zn	22.79	50.76	34.64	34.57	22.11	7.65	0.612	-0.116

Result from descriptive statistics reveals that the CV varies 22.11% of Zn to 59.41% of Cu in dry season as well as 18.25% of Zn to 77.25% of Mn in rainy season, respectively, which indicated a great degree of variability. The contours rising of CV values reflected the non-homogeneous distribution of concentrations of anthropogenically emitted metal elements (Li et al., 2012). The greatest and the smallest SD were detected for metal element of Ti (406.2571) and Cd (1.137265) in the dry season. Similarly, the greatest and the smallest SD were detected for metal element of Ti (297.8519) and Cd (0.736449) in the rainy season. Additionally for dry season, the metal elements of Al, Cd, Co, K, Na, Ni, Sc and Ti in soil were fairly symmetrical (skewness -0.5 to 0.5), whereas, As, Ba, Ca, Cr, Hg, Sb, Sr, V and Zn in soil indicated the data points were moderately skewed as the skewness value varies from -0.1 to -0.5 and 0.5 to 0.1. Moreover, metal elements of Cu, Fe, Mn and Pb in soil were highly skewed exhibited skewness value of <-1 and >1. In addition, the metal elements of Al, As, Ba, Ca, Cd, Co, Cr, Fe, Hg, K, Na, Ni, Sc, Ti and Zn exhibited platykurtic distribution (Kurtosis<0), whereas, Cu, Mn, Pb, Sb, Sr and V exhibited leptokurtic distribution (Kurtosis >0). Moreover in rainy season, the skewness of metal elements of Al, Ba, Co, Fe, Ni, Pb, Sb, V and Zn were fairly symmetrical; however the metal elements of Ca, K, Mn, Na, Sc, Sr and Ti indicated the data points were moderately and metal elements of As, Cd, Cr, Cu and Hg were highly skewed. Furthermore, the metal elements of Al, Ba, Ca, Co, Fe, K, Mn, Na, Ni, Pb, Sb, Sc, Sr, V and Zn exhibited platykurtic distribution, whereas the metal elements of As, Cd, Cr, Cu, Hg and Ti exhibited leptokurtic distribution.

3.2 Seasonal Variation of the Concentration of Metal Elements

Based on descriptive statistical analysis, large SD was found for all metal elements, especially for Fe, Al, K and Ca in soil for both the dry and rainy seasons indicated wide variation of their concentrations in soil. Thus, contamination of soil from metal elements by anthropogenic activities as well as natural soil parent materials. The highest mean concentration of Fe and Ti both the dry and rainy season (Table 1). Moreover, based on mean concentration, the level of metal elements can be ordered as Fe> Al> K> Ca> Ba> Na> P B> V> Ti> Sr> Zn> Mn> Sc> Cu > Sb >Co >Cr >Hg> As >Ni> Cd in dry season and Fe>Al> K> Ca >Ba> Na >V> Ti> Sr >Zn >Pb> Mn> Sc >Co> Cu> Sb>Cr>Hg>As>Ni>Cd in rainy season. The concentrations of metal elements for rainy reason were relatively lower as compared to the dry season and the magnitude of concentrations followed almost same pattern for both the seasons.

Table 2: Correlation between metal elements in dry season

	Ca	Al	Ti	Sb	Sc	Sr	V	Ba
Ca	1.000							
Al	0.979	1.000						
Ti	0.978	0.976	1.000					
Sb	0.972	0.965	0.966	1.000				
Sc	0.985	0.979	0.987	0.986	1.000			
Sr	0.974	0.967	0.958	0.987	0.985	1.000		
V	0.970	0.974	0.956	0.984	0.980	0.984	1.000	
Ba	0.992	0.963	0.975	0.975	0.983	0.973	0.967	1.000

3.3 Correlation between Heavy Metals

In this study, Pearson's correlation coefficients were calculated to measure the intensity degree of association between metal elements. The Pearson's correlation matrix of metal elements of Ca, Al, Ti, Sb, Sc, Sr, V and Ba in soil for the dry season is provided in Table 2. The most significant correlation was observed for Ca and Ba (0.992) in dry season and Ti and Sr (0.991) in rainy season indicating same source of pollution. In contrary, the concentration of Cr showed very weak correlations with Mn, indicated Mn is from different sources than Cr in dry season. However, the concentrations of Hg showed very weak correlations with Zn (0.569) in rainy season. High positively correlations were observed between all metal elements, such as, Sc and Ti (0.987), Sb and Sc (0.986), Sb and V (0.984), Sc and Ba (0.983), Al and Sc (0.979), Ti and Ba (0.975), As and Al (0.974), V and Ca (0.974) , As and Ba (0.970) Ti and Sb (0.966), Ni and Fe(0.927), Zn and V (0.906), Fe and Cd (0.888), Al and Mn (0.645),Cr and Hg (0.575) as well as Hg and V (0.876) (Table 2). In addition, the some parameters in on soil of waste disposal site in rainy season showed also high positively correlated as Sr and Ti (0.991), Al and V (0.986), V and Ba (0.982), Ca and Sc (0.985), Ti and V (0.978) and Co and Ca (0.957), V and Co (0.961), Fe and V (0.948), Fe and Na (0.938), Cd and Sr (0.933), Ni and Sc (0.918), Zn znd Ti (0.861), Zn and Sb (0.861), Hg and Fe (0.733), Fe and Mn (0.947) and As and Ca (0.691). Here it should be noted that concentrations of Ca, Al, Ti, Sb, Sc, Sr, V, Ba showed strong correlation with each other in both the dry and rainy season, which indicated same sources of contamination for these metals.

3.4 Principal Component Analysis

In this study, variability calculation based on eigenvector and factor loadings leads to identify distinct source of generation of metal elements in soil of the studied area. PCs of 21 for dry and 19 for rainy were considered for the metal elements and results provided in Table 3. In addition, the larger eigenvalue obtained for F1 (18.5331) that indicated large proportion of variability (88.2530%) for rainy season (Table 3). The percentage contribution of the 1st to 2nd PCs for the metal elements in dry season as represented in Equation 4 was 92.105%.

$$\left(\left\{\sum_{i=1}^{2}\lambda_i\right\}/\left\{\sum_{i=1}^{21}\lambda_i\right\}\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(4)$$

The selection of the first two parameters as the PCs since there was significant evidence of high enough total variance from the percentage contributions. The eigenvalues λ_i $(\lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{15}, \lambda_{16}, \lambda_{17}, \lambda_{18}, \lambda_{19}, \lambda_{20}, \lambda_{21})$, had little contributions to the total structure of the data under study. The percentage contribution of the 3th to 21st components (Equation 5) is 7.895% (Olawoyin, 2012). This suggested that very little information, which can be considered negligible, will be lost.

$$\left(\left\{\sum_{i=3}^{21}\lambda_i\right\}/\left\{\sum_{i=1}^{21}\lambda_i\right\}\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(5)$$

However, the percentage contribution of the 1st PC for the metal elements in rainy season as represented in Equation 6 was 88.25%

$$\left(\left\{\sum_{i=1}^{1}\lambda_i\right\}/\left\{\sum_{i=1}^{18}\lambda_i\right\}\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(6)$$

Which verbalized the selection of the first three parameters as the PCs and the eigenvalues λ_i $(\lambda_3, \lambda_4, \lambda_5, \lambda_6, \lambda_7, \lambda_8, \lambda_9, \lambda_{10}, \lambda_{11}, \lambda_{12}, \lambda_{13}, \lambda_{14}, \lambda_{15}, \lambda_{16}, \lambda_{17}, \lambda_{18}, \lambda_{19})$, had minimal contributions to the nature of the general data. The % contribution of the 2nd to 19th PCs as illustrated in Equation 7 was 11.75% for rainy season.

$$\left(\left\{\sum_{i=2}^{19}\lambda_i\right\}/\left\{\sum_{i=1}^{19}\lambda_i\right\}\right)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(7)$$

However, the larger eigenvalue obtained for F1 (18.267) indicated large proportion of variability (86.987%) for dry season as well as for rainy season F1 was found to be 18.5331 that indicated large proportion of variability (88.253%) for rainy season. Based on the results of PCA for metal elements of dry season, the eigenvalues upto the second extracted components (F2) were found greater than 1.0 for the dry season and first extracted components (F2) were found greater than 1.0 for the rainy season. Moreover, the eigenvalues for the PCs (F3 to F21) as well as (F2 to F19) were found less than 1 can be neglected for both the dry and rainy season, respectively. Thus, variables could be reduced to 2 components model (dry season) with 92.105% variation as well as 1 component model (rainy season) that accounts for 88.253% variation (Table 3). Varimax rotation was applied to simplify the factor interpretation by reducing total number of variables that exhibit high loadings per factor. Moreover, some previous investigations indicated first principal component (PC1) and second component (PC2) refers to the contamination of soil due to anthropogenic or human activities and natural parent materials, respectively (Tahir et al., 2007). In this study, in case of dry season, factor analysis revealed that metal elements of Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V were closely related to PC1 indicated derived from anthropogenic activities and rests of the metal elements of Cu, Hg, Mn, Pb and Zn in soil were related to PC2 indicated derived from natural parent materials.

In addition, the metal element of Cr was closed to PC1 and PC2 indicated derived from both the anthropogenic activities and natural parent materials (Table 4). In addition, the factor analysis of PCA for rainy season before and after the varimax rotation is provided in Table 4. It can be estimated that As, Cr, Hg, K, Mn and Na in soil were related to PC2 indicating derived from natural parent materials. Other metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V were related to PC1 indicating derived from anthropogenic activities. Moreover, as the metal elements of Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V in soil showed closed to PC1 indicated derived from anthropogenic activities. Besides, the metal elements of Na, Pb, Cu, K, Ni, Co, Hg, Fe, As, Zn and Cd, which also positively correlated with less stronger impact because they were comparative in shorter distance from origin than that of Ca, Al, Ti, Sb, Sc, Sr, V and Ba.

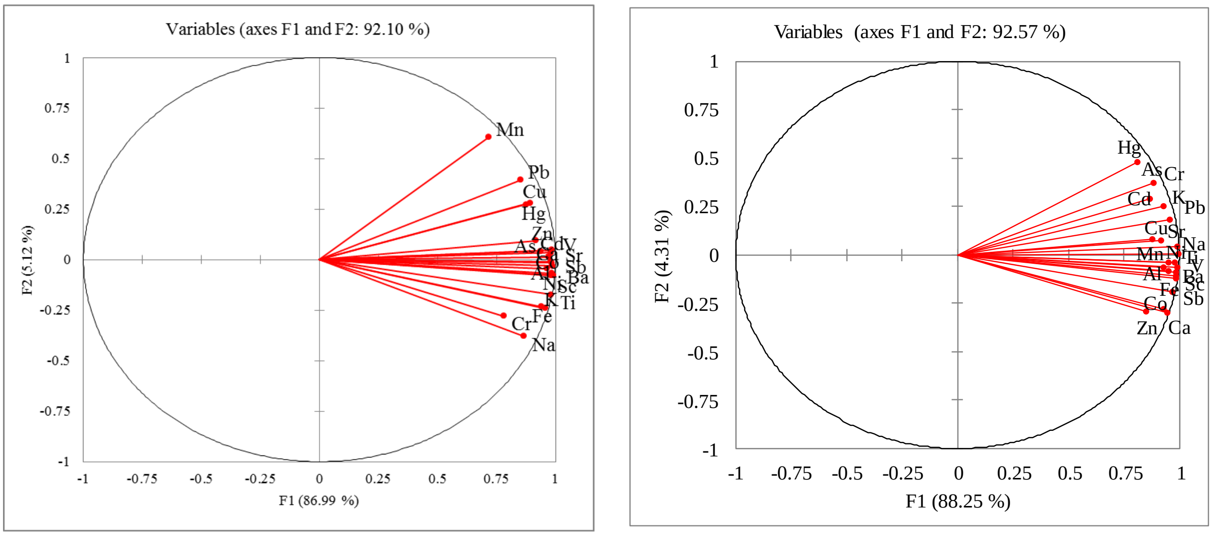
Table 3: PCA of metal elements in soil for dry and rainy seasons

PCs	**Dry season**			**Rainy season**		
	Eigenvalue	**Variability**	**Cumulative**	**Eigenvalue**	**Variability**	**Cumulative**
F1	18.267	86.987	86.987	18.533	88.253	88.253
F2	1.07	5.118	92.105	0.906	4.315	92.568
F3	0.416	1.98	94.085	0.398	1.896	94.464
F4	0.391	1.861	95.946	0.335	1.594	96.058
F5	0.23	1.095	97.041	0.263	1.252	97.310
F6	0.162	0.774	97.814	0.180	0.858	98.168
F7	0.119	0.566	98.38	0.151	0.718	98.886
F8	0.076	0.362	98.742	0.080	0.383	99.269
F9	0.058	0.277	99.019	0.055	0.26	99.529
F10	0.051	0.244	99.263	0.043	0.207	99.736
F11	0.042	0.202	99.465	0.016	0.077	99.813
F12	0.035	0.167	99.633	0.015	0.073	99.886
F13	0.026	0.122	99.755	0.009	0.045	99.931
F14	0.016	0.077	99.832	0.007	0.031	99.963
F15	0.013	0.063	99.895	0.004	0.017	99.979
F16	0.008	0.036	99.931	0.003	0.012	99.992
F17	0.004	0.02	99.952	0.001	0.005	99.997
F18	0.004	0.018	99.97	0.0005	0.003	99.999
F19	0.003	0.013	99.983	0.0002	0.001	100
F20	0.002	0.01	99.993			
F21	0.001	0.007	100			

Moreover, the metal elements of Mn and Cr having the least impact on the PCA model because they were far from each other in the same quadrant. It was noticed that Cu, Hg, Mn, Pb and Zn were located at a distance from origin of circle than that of other metal elements. This indicated the origin of these metal elements was differing from other metal elements. In a similar manner, it was clearly illustrated that; Ca, Al, Ti, Sb, Sc, Sr, V and Ba showed clear positive correlation but they have stronger impact on the PCA model than that of Na, Pb, Cu, K, Ni, Co, Mn, Fe, As, Cr and Cd, which also correlates positively with 8 variables, whereas Hg and Zn having the least impact on the PCA model. It was also found from PCA analysis that Mn, Hg, As, Cr, Na and K were located far from other metal elements indicating different origin from other metal elements.

Table 4: All explained variables and factors derived using the orthogonal varimax rotation method of dry season

Metal	**Before rotation**																					**After rotation**	
	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	D1	D2
Fe	0.96	-0.24	0.02	-0.11	0.01	0.01	-0.01	-0.01	0.05	0.01	-0.01	-0.04	-0.05	-0.08	-0.04	0.00	-0.03	0.00	0.00	0.00	0.00	0.91	0.37
Mn	0.72	0.61	0.29	-0.07	0.00	0.10	-0.05	-0.11	0.02	-0.07	0.02	0.02	-0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.22	0.91
Cr	0.78	-0.28	0.32	0.44	-0.08	-0.05	0.08	-0.03	0.01	-0.01	0.02	-0.03	-0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.79	0.24
Cu	0.89	0.28	-0.11	0.25	0.09	0.08	-0.10	0.13	-0.04	-0.08	-0.06	-0.03	0.01	-0.01	-0.01	-0.01	-0.01	0.00	0.00	0.00	0.00	0.55	0.75
Pb	0.85	0.40	0.18	-0.14	0.00	-0.16	0.13	0.13	0.06	0.04	-0.04	-0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.45	0.82
Zn	0.92	0.09	0.02	0.10	0.35	0.01	-0.02	-0.02	-0.01	0.12	0.04	0.03	-0.01	0.01	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.68	0.62
Ni	0.96	-0.04	-0.13	0.08	0.01	-0.01	-0.07	-0.09	0.15	0.02	-0.07	-0.04	0.06	0.00	0.02	0.01	0.01	0.00	0.00	0.00	0.00	0.80	0.53
Cd	0.94	0.04	0.14	-0.04	-0.23	0.04	-0.16	0.02	-0.07	0.12	-0.03	0.01	0.03	0.00	-0.01	0.00	0.00	0.00	0.00	-0.01	0.00	0.73	0.59
As	0.97	0.01	-0.15	0.05	-0.07	-0.08	0.03	-0.04	0.02	-0.02	-0.05	0.11	-0.03	0.01	-0.02	-0.04	0.00	-0.02	-0.01	0.00	0.00	0.77	0.59
Hg	0.88	0.27	-0.25	0.07	-0.13	0.20	0.17	-0.01	0.02	0.05	0.05	-0.03	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.54	0.74
Co	0.98	-0.08	0.01	0.02	-0.03	-0.02	-0.02	0.08	0.05	-0.03	0.08	0.10	0.06	-0.04	0.02	0.01	0.00	0.01	0.01	-0.01	0.00	0.84	0.52
Na	0.87	-0.38	0.15	-0.18	0.04	0.19	0.00	0.07	0.05	-0.04	0.01	0.00	0.02	0.04	-0.03	0.00	0.01	-0.02	0.00	0.00	0.00	0.92	0.21
K	0.94	-0.23	0.10	-0.12	0.09	0.08	0.09	-0.01	-0.06	0.00	-0.07	0.01	0.01	-0.01	0.07	-0.02	-0.01	0.00	0.00	-0.01	0.00	0.90	0.37
Ca	0.99	-0.03	-0.10	0.00	-0.05	-0.01	-0.02	0.04	-0.02	-0.01	0.02	-0.01	-0.06	0.03	0.02	0.05	-0.01	-0.01	-0.02	-0.01	-0.01	0.81	0.56
Al	0.98	-0.07	-0.07	0.01	0.00	0.01	0.07	-0.03	-0.06	-0.03	-0.08	0.03	-0.01	-0.01	-0.03	0.04	0.02	0.01	0.02	0.00	0.01	0.83	0.53
Ti	0.98	-0.17	-0.02	-0.02	0.00	0.01	-0.02	-0.01	0.00	0.01	0.01	0.02	-0.04	0.00	0.02	-0.01	0.02	0.02	-0.01	0.03	-0.02	0.89	0.44
Sb	0.98	-0.01	-0.04	-0.10	0.05	-0.10	-0.03	-0.03	-0.03	-0.06	0.04	-0.06	0.00	-0.02	-0.01	-0.02	0.03	0.00	0.00	-0.02	0.00	0.80	0.57
Sc	0.99	-0.08	-0.03	-0.05	0.00	-0.05	-0.02	0.00	-0.02	-0.01	0.02	-0.02	0.02	0.04	-0.01	-0.01	-0.01	0.04	-0.02	0.00	0.02	0.85	0.52
Sr	0.98	0.03	-0.06	-0.11	-0.01	-0.10	0.02	-0.04	0.00	-0.02	0.02	-0.01	0.02	0.05	-0.02	-0.01	-0.03	0.00	0.02	0.00	-0.02	0.77	0.61
V	0.99	0.05	-0.02	-0.03	0.01	-0.05	0.04	-0.06	-0.09	-0.02	0.04	-0.03	0.06	-0.03	0.00	0.01	-0.01	-0.02	-0.02	0.02	0.00	0.77	0.62
Ba	0.99	-0.02	-0.09	0.02	-0.04	-0.05	-0.08	0.03	0.01	0.01	0.05	-0.02	-0.05	0.01	0.04	-0.01	0.00	-0.02	0.02	0.01	0.02	0.80	0.57



All sampling points and metal elements were overlapped, indicating high correlation with each other except Mn, Pb, Hg, Cu and Zn.

(a)

All sampling points and metal elements were overlapped except Hg, As, Cr, Cd and K.

(b)

Figure 2: Correlation circle for metal elements in soil (a) dry and (b) rainy season

Additionally, plotted correlation circle exhibited a graphical representation of loading vectors for metal elements in order to determine the most influential variables interaction in this model. Variables that showed longer distances from origin of circle have larger impacts on the general architecture of the model than variables with shorter distances (Olawoyin, 2012). In the loading plot corresponding to the first two PCs (Figure 2); the eight metal elements of Ca, Al, Ti, Sb, Sc, Sr, V and Ba showed clear positive correlation because they were in the same quadrant. In addition, these metal elements have stronger impact on the PCA model because these metal elements were at a longer distance from origin.

3.5 Agglomerative Hierarchical Clustering

In case of dry season, cluster 1 comprises with metal elements of Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V in soil which indicated these metal elements were generated from anthropogenic activities . In addition, cluster 2 comprises with Cu, Hg, Mn, Pb, and Zn, indicating origination from natural sources and Cluster 3 comprises with Cr which derived from both the natural parent materials and anthropogenic sources. Table 5 showed the results of cluster analysis by class for both the dry and rainy season, respectively. In case of dry season, maximum distance to centroid was found for cluster 2 of 170.279 between three clusters, indicating generation of metal elements from natural sources (Table 5). Moreover, maximum distance to centroid for cluster 1 was found comparatively smaller of 122.5230 than that of cluster 2, indicating generation of metal elements from anthropogenic activities. Cluster 3 showed maximum distances to centroid was zero, indicating it was closed to both the clusters, consequently generated from both natural sources and anthropogenic activities. In addition, for rainy season, cluster 1 comprises with Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V which indicated these metal were generated from anthropogenic activities (Figure 3). Cluster 2 comprises with As, Cr, Hg, Mn, K and Na indicating origination from natural sources and Cluster 3 comprises with Cd, Pb and Zn which derived from both natural parent materials and anthropogenic sources. Similarly, it was found that metal elements of cluster 2 indicating generation of metal elements from natural sources; whereas cluster 1 indicating generation of metal elements from anthropogenic activities and cluster 3 consequently generated from both natural and anthropogenic sources (Table 5).

Table 5: Results of cluster analysis

Season	Dry			Rainy		
Class	1	2	3	1	2	3
Objects	15	5	1	12	6	3
Sum of weights	15	5	1	12	6	3
Within-class variance	4511.6859	18600.6792	0.0000	3042.8452	12931.7	4559.68
Maximum distance to centroid	122.5230	170.2796	0.0000	89.0645	148.207	64.2498
Metal elements	Fe, Ni, Cd, As, Co, Na, K, Ca, Al,Ti, Sb, Sc, Sr, V, Ba	Mn, Cu, Pb Zn, Hg	Cr	Fe, Cu, Ni, Co, Ca, Al, Ti, Sb, Sc, Sr, V, Ba	Mn, Cr, As, Hg, Na, K	Pb, Zn, Cd

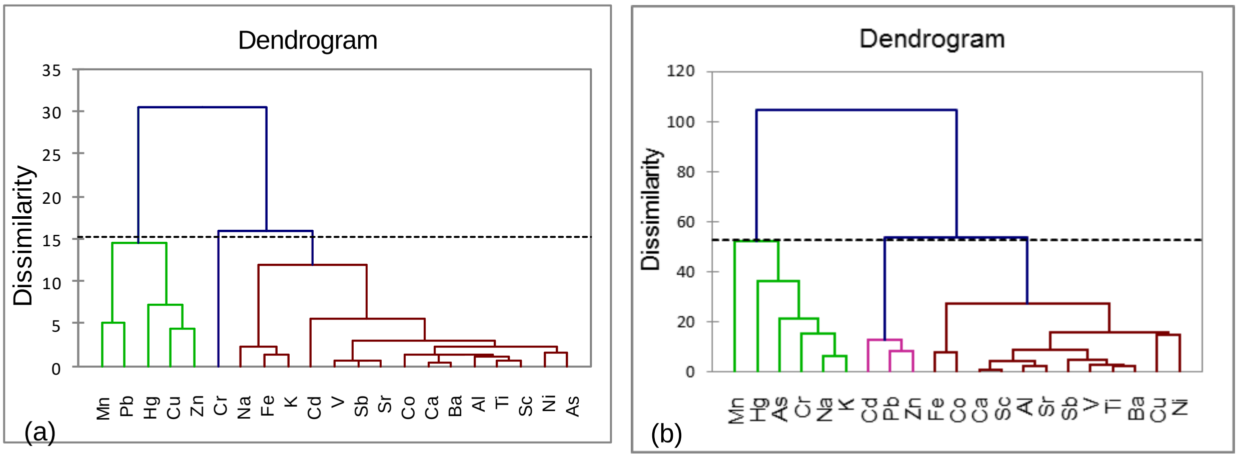


Figure 3: Dendrogram for metal elements in soil during (a) dry and (b) rainy season

3.6 Spatial distribution of metal elements

The spatial distribution of metal concentrations is a convenient tool to identify the sources of generation of metal elements as well as contamination hotspots with high metal concentrations in a visual form. The predicted map produced from different interpolation techniques for metal elements provide a field scale contamination of soil by metal elements present in waste disposal site. The predicted map of almost all the metal elements showed almost similar pattern of contamination indicated same sources of generation indicating anthropogenic activities accompanied by particular natural soil materials.

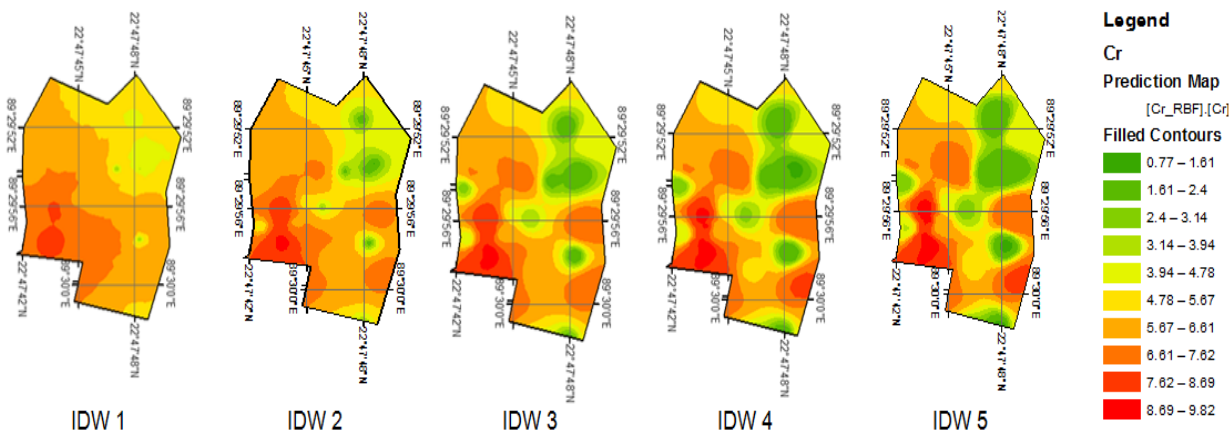


Figure 4: Spatial distribution of Cr in soil using IDW with power of 1- 5.

This provide a refinement and reconfirmation of the result obtained from statistical analyses of Pearson's correlation and AHC. Figure 4 showed spatial distribution of metal element of Cr generated from both natural and anthropogenic activities to visualize the extent up to which the severities of contamination of this metal element exist. Moreover, from produced prediction surface, it was found that for all the metal elements most of the contaminated hotspots were found near the central point of the disposal site for all studied metal elements.

4. CONCLUSION

Results of Pearson's correlation reveals that almost all the metal elements were strongly correlated with each other indicating these metal elements were derived from the same generation sources. In addition, results of PCA depicts that Cu, Hg, Mn, Pb, and Zn in soil derived from natural parent materials, while, Al, As, Ba, Ca, Cd, Co, Fe, K, Na, Ni, Sb, Sc, Sr, Ti and V in soil from anthropogenic activities and Cr from both the natural and anthropogenic sources in dry season. It can be demonstrated that As, Cr, Hg, K, Na and Mn in soil derived from natural parent materials; Al, Ba, Ca, Co, Cu, Fe, Ni, Sb, Sc, Sr, Ti and V from anthropogenic activities as well as Cd, Pb and Zn from both the natural parent and anthropogenic sources in rainy season. Results from AHC also proved the same generation sources of metal elements as similar of PCA in soil for both the dry and rainy seasons. Produced prediction surface for all the interpolation techniques showed most of the contaminated hotspots was found near the central point of the disposal site for all studied metal elements. Here, it can be concluded that contamination of soil by heavy metals can be minimized by controlling anthropogenic activities. Moreover, inhabitants nearby the waste disposal site should be shifted to keep safe from the adverse effects of intake foods growing in this contaminated soil as well as water which is also contaminated by heavy metals in soil.

REFERENCES

Nriagu, J.O. and Pacyna, J.M. (1988). Quantitative assessment of worldwide contamination of air, water and soils by trace metals, Nature Vol. 333, 134–139.

Tahir, N.M., Chee, P.S. and Maisarah, J. (2007). Determination of heavy metals content in soils and indoor dusts from nurseries in dungun, Terengganu, The Malay J. Anal. Sci., Vol. 11, 280-286.

Jia, L., W. Wang, Y. Li and Yang, L. (2010). Heavy metals in spol and crops of an intensively farmed area: A case study in Yucheng City, Shandong Province, China, International J. of Environmental Research and Public Health, vol. 7 (2), 395-412.

Stanimirova, I., Zehl, K., Massart, D.L., Vander Heyden, Y. and Einax, J.W. (2006). Chemometric analysis of soil pollution data using the Tucker N-way method, Analytical and Bioanalytical Chemistry, Vol. 385, 771–779.

Omran, E.S. and El Razek, A. (2012). Mapping and screening risk assessment of heavy metal concentrations in soils of the Bahr El-Baker Region, J. of Soil Sci. and En. Man. Vol. 6(7), 182–195.

Lee, C.S.L., Li, X. and Wenzhong Shi, Cheung, S.C. and Thornton, I. (2006). Metal contamination in urban, suburban, and country park soils of Hong Kong: A study based on GIS and multivariate statistics, Sci of the Total Environ, Vol. 356, 45 – 61.

Li, J., Xie, Z.M., Xu, J.M. and Sun, Y.F. (2006). Risk assessment for safety of soils and vegetables around a lead/zinc mine, Environ Geochem Health, Vol. 28, 37–44.

Zou, J., Dai, W., Gong, S. and Ma, Z. (2015). Analysis of Spatial Variations and Sources of Heavy Metals in Farmland Soils of Beijing Suburbs, Vol. 10(2): DOI:10.1371/journal.pone.0118082.

Pebesma, E.J., de Jong, K., Briggs and D.J. (2007). Visualizing uncertain spatial and spatio-temporal data under different scenarios: an air quality example. J. of Geo. Infor Sci., Vol. 21 (5): 515-527.

Gotway, C.A., Ferguson, R.B., Hergert, G.W. and Peterson, T.A. (1996). Comparison of kriging and IDW methods for mapping soil parameters. Soil Sci. Soc. Am. J., Vol. 60, pp. 1237–1247.

Isaake, E.H. and Srivastava, R.M. (1989). An Introduction to Applied Geostatistics, Oxford Univ., YK.

Olawoyin, R., Oyewole, S.A. and Grayson, R. L. (2012). Potential risk effect from elevated levels of soil heavy metals on human health in the Niger delta, Ecotoxicol. Environ. Saf., Vol.85(1), 120–130.