CHEMOMETRIC EVALUATION OF HEAVY METALS POLLUTION PATTERNS OF SOIL IN THE VICINITY OF SOME DUMPSITES IN LOKOJA, NIGERIA

Supported by tetfund TERTIARY EDUCATION TRUST FUND

A. B. Olukoya, J. Yisa and J. O. Jacob*
Department of Chemistry, Federal University of Technology, Minna, Niger State, Nigeria
*Correspondent author: john.jacob@futminna.edu.ng

Received: October 12, 2018 **Accepted:** January 28, 2019

Abstract: Multivariate methods of data analysis were used to describe patterns of soil contamination with heavy metals in Lokoja city dumpsites. Groups of metals and dumpsites were identified using cluster analysis. Score plot was used to grouped dumpsites into four; Group I consists of Pata and Cinima dumpsites, while groups II, III and IV consist of Caturgo, High Court and New Market dumpsite respectively. The loading plot shows that Cu, Co, Cr and Cd have less significant contribution than the other five elements with positive value on the principal components. The bi-plot sketch shows the correlation between site samples and metals, in which Group I sample sites, consisting of sample 1 and 3, does not really correlates with particular metal, making it exceptional in terms of contamination. It is probably, the least contaminated among the sample site groups. Group II, consisting of sample site 4, is characterized by Zn and Cu. Group III, the group of sample site 5, has Co, Fe and Mn as main metal contaminant. Ni and Pb are dominant contaminants of Group IV. Strong and positive correlation co-efficient of element pairs are indicates common source. This can be observed in the cases of Fe and Mn, Zn and Mn, Zn and Fe, Co and Mn as well as Cd and Cr, whereas, Ni and Pb, originate from different pollution source. The cluster analysis gives pattern recognition based on the similarity and closeness of dump sites and the nature of the contaminants. The result reveals the closeness of sites 1, and 2 in terms of pollution level.

Keywords: Heavy metals, dumpsite, principal components, bi-plot sketch, score plot

Introduction

Heavy metals are natural components of the earth crust. They become concentrated as a result of human activities and can enter plant, animal and human tissues through inhalation, diet and manual handling (Abosede, 2017). The term heavy metals is generally used to describe a group of metals and metalloids with an atomic density greater than 5.0 g/cm^3 and is toxic or poisonous even at low concentration. Their effects on living organisms generally results from contamination of either abiotic systems (soil, water and air) and subsequent uptake and bio-accumulation. The presence of heavy metals in the environment is of great ecological significance due to their toxicity at certain concentrations, translocation through food chains and non-biodegradability which is responsible for their accumulation in the biosphere (Ackova, 2018). In developing countries, like Nigeria, it is common to see crops grown around dumpsites due to the belief that these wastes are high in organic matter and thus can be used as manure and plants also tend to grow in these dumpsites and there is a high tendency for heavy metals to accumulate in the soil which is taken up by plants (Orji *et al.,* 2018). Theses wastes often contain heavy metals in various forms and at different contamination levels. Lead is a particularly dangerous metal that has no biological role and negatively affects children in significant ways (Renfrew, 2018). The environmental problem with heavy metals is that they are unaffected during breakdown of organic waste and have toxic effects on living organisms when they exceed a certain concentration. The high concentration of heavy metals in soils is reflected by concentrations of metals in plants, water, animal, and human bodies (Koki *et al*., 2018).

Soil as a component of terrestrial ecosystem, being essential for the growth of plants is a dynamic system and subject to short term fluctuations, such as variation in moisture status, pH and release conditions and also undergoing gradual alterations in response to changes in management and environmental factors (Zauro *et al.*, 2013). Pollution of the soil and water environment by inorganic chemicals has been considered a major threat to the environment including living things. Agricultural drainage water contain pesticides, fertilizers and effluents of industrial activities plus runoffs in addition to sewage effluents supply water bodies and soils with huge quantities of inorganic anions and heavy metals (Mateo-Sagasta *et al*., 2017). Potentially contaminated soil may occur at old landfill sites, particularly those that accepted industrial wastes; old orchards that used insecticides containing arsenic as an active ingredient, and field that had past application of waste pipes and tacking, industrial areas where chemicals have been dumped on the ground (Ojiako *et al.*, 2013). Nonetheless, the most anthropogenic sources of metals which affect land and soil are industrial, petroleum contamination and sewage disposal (Abosede, 2017). The pollutants have different sources reaching the level that could be harmful to humans and other biota. Most studies done so far revealed that elevated levels of metals in soils are due to anthropogenic input (Koki *et al*., 2018).

Chemometric techniques are widely applied in atmospheric studies for identifying sources of pollutants or for determining the importance of contaminant source contributions to a particular site, have been little use in soil studies, but may be effectively employed in such investigations. Chemometric techniques such as correlation analysis, principal component analysis (PCA) and cluster analysis (CA) are employed to identify sources contributing to observing heavy metal pollution of soils (Salah *et al*., 2015).

Chemometric technique can be divided into two categories which are supervised and unsupervised methods. There are several techniques included in the two categories such as hierarchical cluster analysis (HCA), principal component analysis (PCA), k-nearest neighbor (KNN) and soft independent model of classification analogy (SIMCA). HCA and PCA are the unsupervised techniques while the KNN and SIMCA included into the supervised technique (Messai *et al.*, 2016). There also has machinery learning for computational method which is support vector machine (SVM). SVM is normally used to handle problems involving a non-linear classification. Principal component analysis (PCA) is one of the unsupervised techniques which are frequently used in recognition and finding patterns in data of high dimension. PCA is a mathematical method of rearrangement information in a data set of samples (Qureshi, 2017). The mathematical concept in PCA covers standard deviation, covariance, eigenvectors and eigenvalues. It becomes more effective with the present of huge number of variables, as in spectroscopic data. PCA is a powerful tool for analysing data (Messai *et al.*, 2016).

Furthermore, PCA can reduce the dimensionality of data without losing much information of original data which is very useful in the classification process (Olejniczak *et al.*, 2010). PCA is a technique to develop a smaller number of artificial variables based on the original variables. It is the summarization of data with many (p) variables by a smaller set of (k) derived variable. It takes a data matrix of *n* objects by *p* variables, which may be correlated, and summarizes it by uncorrelated axes that are linear combinations of the original *p* variables. The discovery of new variables is known as principle components (PCs) which consist of the majority of variability in the data. The first PC describes the most variability in the data while the second PC explains the maximum amount of the remaining variability. The second PC must be orthogonal to the previous PC which is PC1 and it is the same goes to the next PCs (Olejniczak *et al.,* 2010).

Results of PCA can be seen from the loading and scores plots. Loadings plot shows the distribution of the variables and it is important for describing the variation in the original data set. Loading that close to 1 (or – 1) and the angle between the principle component and the original variables near 0 (or 180) degrees, it indicates that variables contribute significantly to the PC. Scores plot also known as sample diagnostic. It illustrates the distribution of samples present in data. The plot reveals how the samples are related to each (Messai *et al.*, 2016).

The aim of this study is to apply the chemometric techniques of chemical analysis to describe the patterns at which soil is been polluted with inorganic contaminants such as Cu, Pb, Cd, Fe, Mn, Se, Co, Hg, Cr and Zn in the vicinity of selected dumpsites in Lokoja City, Kogi State, Nigeria.

Materials and Methods

Collection of sample

Soil samples were collected from five (5) dumpsites via; Pata dumpsite, Caturgo dumpsite, Cinema dumpsite, New Market dumpsite and High court premises dumpsite, respectively. Ten composite samples of soil were collected from each of the site at a distance of not more than 10 m to the refuse heap.

Sample preparation and digestion

The composites were air dried, thoroughly mixed and sieved, using 2 mm aperture sieve. The acid digests of the samples were prepared by weighing 0.50 g of the dried powder in triplicates. In each case, this was transferred into 150 cm^3 clean borosilicate beaker and 10 cm^3 concentrated trioxonitrate (V) acid (HNO_3) was added. It was then covered with a watch glass and kept for an hour till the primary reactions subsided. It was then heated on hot a plate till all the material was completely dissolved. It was allowed to cool to room temperature and then 10 cm^3 of 60% tetraoxochlorate (VII) acid ($HClO_4$) was added, then thoroughly mixed, heated strongly on the hot plate until the solution became colourless and reduced to about 2-3 cm^3. While heating, the solution was not allowed to dry. After cooling, it was filtered through Whatman No. 44 (ashless) filter paper into a 100 cm^3 capacity volumetric flask, diluted to the mark distilled water and kept in a plastic sample bottle for analysis of inorganic constituents.

Heavy metal determination

Analysis for inorganic contaminant was done using atomic absorption spectrophotometer (AAS) (Buck Scientific 210VGB).

Chemometrics analysis

Correlation matrix was done with excel spread sheet for windows. The Chemometrics software packages: MATLAB from Mathwork Inc. and PLS-Toolbox from Eigenvectors Research Inc. were used to perform the Principal Component Analysis (PCA) and Cluster Analysis (CA).

Results and Discussion

The Principal component analysis was employed to extract information from a set of large variables to developed a reduced variable dimensions that are called principal components which still give adequate information about the variance as in the large dataset. PCA is a procedure for variable reduction. Principal Component Analysis that is used for variable reduction is shown in Table 1. The PCA1 is strongly correlated with five of the original variables. The first principal component increases with increasing Co, Fe, Zn, Cu and Mn scores. This suggests that these five criteria vary together. If one increases, then the remaining two also increases. This component can be viewed as a measure of the quality of Co, Fe, Zn and Cu. It is noted that based on the correlation of 0.852 that this principal component is primarily a measure of Co. The PC2 increases with increasing Cd and Cr. This suggests that places with high Cd also tend to have high Cr. The third PC3 increases with only one of the values decreasing with Cu. This component can be viewed as a measure of the location that is more contaminated.

Table 1: Principal component analysis

Component	PC1	PC2	PC3
Mn	0.777	0.55	0.286
Fe	0.821	0.199	0.508
Cu	0.788	-0.311	-0.496
Zn	0.806	-0.14	0.575
Co	0.852	0.46	-0.213
Cr	-0.139	0.917	-0.296
Cd	-0.107	0.985	-0.113
Pb	-0.633	0.149	0.752
Ni	-0.942	0.316	0.114

Extraction Method: Principal Component Analysis

Table 2: Test for normality

Component	Kolmogorov-Smirnova			ShapiroWilk		
	Statistic	Df	Sig.	Statistic	df	Sig.
Mn	0.195	5	.200*	0.98	5	0.935
Fe	0.187	5	.200*	0.96	5	0.808
Cu	0.345	5	0.052	0.759	5	0.036
Zn	0.169	5	.200*	0.967	5	0.853
Co	0.292	5	0.19	0.845	5	0.18
Cr	0.198	5	.200*	0.951	5	0.742
Cd	0.136	5	.200*	0.987	5	0.967
Pb	0.221	5	.200*	0.902	5	0.421
Ni	0.136	5	.200*	0.987	5	0.967

* This is a lower bound of the true significance.

Table 3: Factor analysis

Component	F1	F2	F3
Mn	0.889	0.195	0.401
Fe	0.978	0.123	-0.001
Cu	0.223	0.934	-0.203
Zn	0.931	0.124	-0.344
Co	0.62	0.629	0.451
Cr	-0.078	-0.036	0.97
Cd	0.074	-0.162	0.981
Pb	0.014	-0.994	-0.033
Ni	-0.572	-0.752	0.326

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization

Test for normality obtained from principal component analysis of each inorganic contaminant is shown in Table 2. There were two correlations between Kolmogorov-Smirnova and Shapiro Wilk. It was noted that the statistic value of Cu, Co and Pb were significantly higher compare to Mn, Fe, Zn and Ni which as lower value in Shapiro Wilk. The factor analysis which is used to extract information from set of large variables to developed a reduced variable dimensions is shown in Table 3. This Table shows a three factor analysis of the same data analyzed in Table 2. It is noted that Mn, Fe and Zn were significantly present and correlated with F1 while Pd and Ni were not significant present since it possess negative value.

Table 4: Correlation matrix of inorganic elements

	Mn	Fe	Cu	Zn	Co	Cr	Cd	Pb	Ni
Mn	1.000								
Fe	0.875	1.000							
Cu	0.279	0.365	1.000						
Zn	0.715	0.924	0.391	1.000					
Co	0.868	0.661	0.608	0.502	1.000				
Cr	0.288	-0.044	-0.204	-0.414	0.336	1.000			
Cd	0.436	0.038	-0.349	-0.287	0.396	0.933	1.000		
Pb	-0.207	-0.090	-0.898	-0.101	-0.645	0.027	0.121	1.000	
Ni	-0.524	-0.655	-0.899	-0.738	-0.679	0.384	0.400	0.728	1.000

The correlation matrix Table reveals the contaminants of same source (Table 4). The strong positive correlation co-efficient among elements pairs are indication of common source. This could be observed in the cases of Fe and Mn, Zn and Mn, Zn and Fe, Co and Mn. Similarly, Cd has common origin with Cr, whereas, Ni and Pb, originate from different pollution site.

The score plot (Fig. 1), is a projection of data onto subspace and shows the patterns of soil sample sites contamination. It is used for interpreting relations among observation. It grouped the site into four groups: Group I consists of Pata and Cinima dumpsites, Group II consist of Caturgo dumpsite, Group III consists of High Court premises dumpsite and Group IV, the New Market dumpsite.

The loading plot that shows the contributions of the inorganic contaminants to the principal components is shown in Fig. 2. The loading plot is a plot of the relationship between all the inorganic contaminant and the PC3 at 18.23%. It shows that Cu, Co, Cr and Cd are not significantly present while Mn, Fe, Zn, Pb and Ni are present at high level.

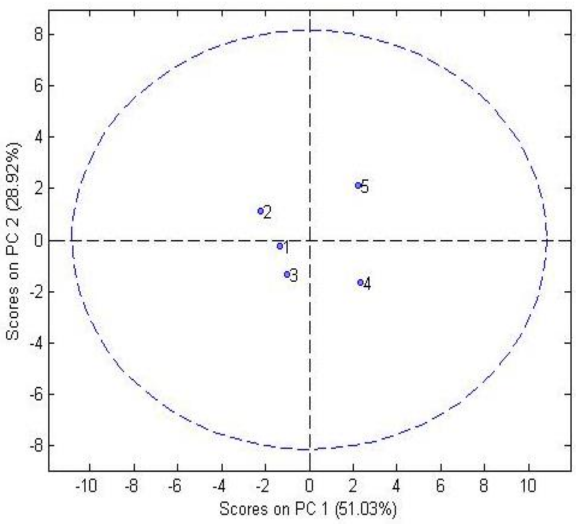

Fig. 1: Score plot

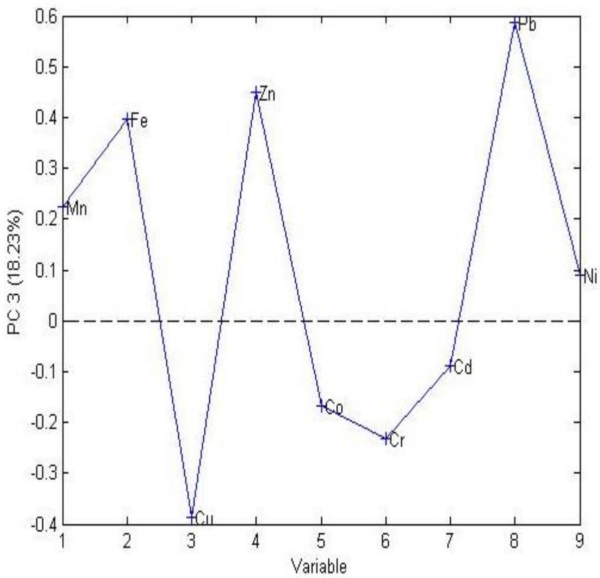

Fig. 2: Loading plot

The bi-plot sketch that represents the correlation between site samples and inorganic elements is shown in Fig. 3. The Bi-plot shows both the loadings and the scores for two selected components in parallel at 28.92% PC2 and 51.03% PC1. Bi-plot shows that the four sample groups earlier observed on score plot in Fig. 1 have their peculiar inorganic element contaminants. Group I sample sites consisting of sample 1 and 3 does not really correlates with particular contaminant. Group I is therefore exceptional in terms of contamination. It is probably, the least contaminated among the sample site groups. Group II, consisting of sample site 4, is characterized by Zn and Cu elements. Group III which is where sample site 5 belongs, has Co, Fe and Mn as main inorganic contaminants. Ni and Pb are dominating contaminants of Group IV (sample 2). The contaminants Cr and Cd are not significantly present in any of the four groupings.

The cluster analysis gives pattern recognition based on the similarity and closeness in the soil sites and the nature of the contaminants (Fig. 4). It appears that sample sites 1 and 2 are closest in terms of pollution level. Pollution type in sample site 3 appears unique that is not really comparable to others.

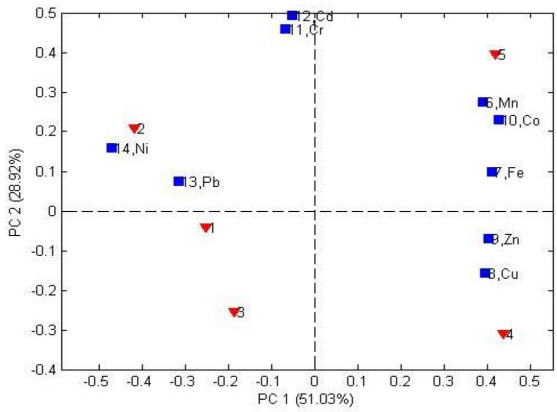

Fig. 3: Bi-plot

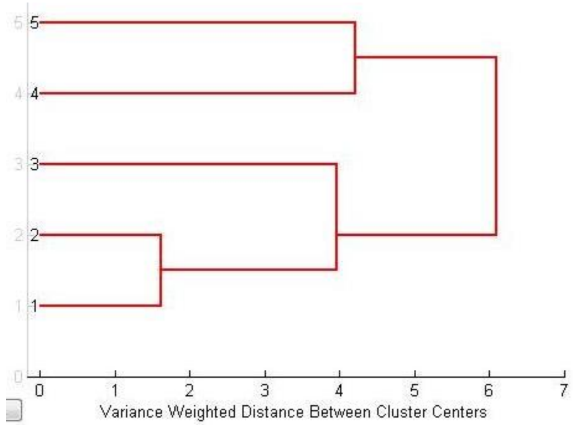

Fig. 4: Dendrogram of cluster analysis

The score plot shows the pattern of soil sample sites' contamination. Based on this plot, the soil can be grouped into 4 based on the inorganic contaminants. Group I consists of Pata dumpsite and Cinima dumpsite, while Groups II, III and IV consist of Caturgo dumpsite, High Court premises dumpsite and New Market dumpsite, respectively. The loading plot shows the contributions of the inorganic contaminants to the principal components. It shows that elements Cu, Co, Cr and Cd have much less significant contribution than the other five elements with positive value on the principal components. The bi-plot sketch represents the correlation between site samples and inorganic elements. It can be observed that the four sample groups earlier observed on score plot in Fig. 1, have their peculiar inorganic element contaminants. Group I sample sites consisting of sample 1 and 3 does not really correlates with particular contaminant. Group I is therefore exceptional in terms of contamination. It is probably, the least contaminated among the sample site groups. Group II, consisting of sample site 4, is characterized by zinc and copper elements. Group III which is where sample site 5 belongs, has Co, Fe and Mn as main inorganic contaminants. Ni and Pb are dominating contaminants of Group IV (sample 2). The contaminants Cr and Cd are not significantly present in any of the four groupings. The correlation matrix table reveals the contaminants of same source. The high positive correlation co-efficient among elements pairs are indication of common source. This can be observed in the cases of Fe and Mn, Zn and Mn, Zn and Fe, Co and Mn. Similarly, Cd has common origin with Cr, whereas, Ni and Pb, originate from different pollution site. The cluster analysis gives pattern recognition based on the similarity and closeness is the soil sites and the nature of the contaminants. This neighbourhood closeness is measured by variance weighted distance between cluster centers, using Ward's procedure. It appears that Sample sites 1 and 2 are closest in terms of pollution level. Pollution type in Sample site 3 appear unique that is not really comparable to others.

Conclusion

Chemometrics methods of chemical data analysis were able to describe pattern of soil pollution with inorganic contaminants in different dumpsites in Lokoja, Kogi State. The study was able to reveals the contaminants of same source using correlation matrix table. Further study could be carried out to ascertain the main emission sources because when the sources have been identified and their contributions estimated, it may be possible to eliminate or to modify the source, so that emission of pollutants is minimized.

References

Abosede OA 2017. Review on heavy metals contamination in the environment. *Eur. J. Earth and Envt.,* 4(1): 1-6.

Ackova DG 2018. Heavy metals and their general toxicity on plants. *Plant Science Today,* 5(1): 14 - 18.

Koki IB, Taqui SN & Dhaif-Allah MAH 2018. A chemometric approach for the distribution and source identification of heavy metals in tannery contaminated soil. *Int. J. Theoretical & Appl. Sci.,* 10(1): 01-08.

Mateo-Sagasta J, Zadeh SM & Turral H 2017. *Water Pollution from Agriculture: A Global Review*. FAO, United Nations, Rome.

Messai H, Farman M, Sarraj-Laabidi A, Hammami-Semmar A & Semmar N 2016. Chemometrics methods for specificity, authenticity and traceability analysis of olive oils: Principles, classification and applications. *Food,* 5: 77.

Olejniczak A, Chostenko AG & Fall J 2010. Discrimination of base oils and semi-products using principal component analysis and self organizing maps. *Fuel,* 89: 1150-1155.

Ojiako EN & Okonkwo MN 2013. Analysis of heavy metals in soil of mechanic workshops in Onitsha metropolis. *Advances in Appl. Sci. Res.,* 4(1): 79-81.

Orji OA, Edosa IVO, Ogbemudia I, Omokaro SO & Elenwo CE 2018. Assessment of some heavy metals in soils and plants growth in dumpsites. *IOSR J. Envtal. Sci., Toxicol. and Food Techn.,* 12(3): 17 – 23.

Qureshi NA, Suthur V, Magsi H, Sheikh MJ, Pathan M & Qureshi B 2017. Application of principal component analysis (PAC) to medical data. *Indian J. Sci. and Techn.*, 10(20): 10 – 19.

Renfrew D 2018. Lead poisoning and the dangers of pragmatism. *Int. J. Envtal. Research and Public Health,* 15: 1 – 4.

Salah EA, Turki AM & Mahal SN 2015. Chemometric evaluation of the heavy metals in urban soil of Fallujah City, Iraq. *J. Envtal. Protection,* 6: 1279 - 1292.

Zauro SA, Dabai MU, Tsafe AI, Umar KJ & Lawal AM 2013. Extent of some heavy metals contamination in soil of farmlands around Sokoto metropolis. *European Scientific J.,* 9(3): 30-37.