INSTRUMENTAL NEUTRON ACTIVATION OF TRACE ELEMENTS IN SOILS AND SEDIMENTS AROUND THE KADUNA REFINERY AND DATA EVALUATION USING CLUSTER TECHNIQUES

ONOJA, M. A.[1]; EWA, I. O. B.[2]; OLADIPO, M. O. A.[2]; & JONAH S. A.[2]
[1]Department of Physics, Ahmadu Bello University, Zaria, Nigeria
[2]Centre for Energy Research and Training, Ahmadu Bello University, Zaria, Nigeria
E-mail: gtob2014@gmail.com **Phone No:** +234-805-073-4928

Abstract
A large number of experimental objects is usually associated with large data. Meaningful interpretation of such data set is very crucial to ascertain the pattern of their existence within the host matrices. Achieving this requires some sort of groupings in order of hierarchy. Instrumental Neutron Activation Analysis (INAA) was used to determine the abundance and distribution of elemental pollutants in soils and sediments around Kaduna Refinery. Results of the activation analysis revealed that Twenty Five (25) elements including: Mg, Al, Ti, V, Mn, Na, K, As, La, Sc, Cr, Fe, Co, Zn, Rb, Sb, Cs, Ba, Eu, Yb, Lu, Hf, Ta, Th, and Dy were detected. The data set acquired from this work was subjected to cluster analysis using the method of SINGLE LINKAGE. Using the computer-aided statistical package, SAS, dendograms were generated. The resulting dendograms revealed some useful patterns which were used to segregate the different soil environments (upland soil, fadama soil, rocky soil, and sediments) in the study area. A repeat of clustering procedures without the elemental concentrations of the sediments led to regrouping of old clusters, giving rise to clearer picture of the groupings. The samples were clearly regrouped based on the soil environments (upland and fadama soils).

Keywords: Cluster technique, Elemental pollutants, Soil environment, INAA

Introduction

Cluster technique is a statistical tool used to analyze a large data suite (data set). The program is used to observe the clustering pattern of samples based on the concentrations of elemental pollutants measured. Cluster analysis becomes very useful when the data set is very large such that it becomes difficult to recognize trend or pattern (Everitt, 1979). Useful algorithms have been proposed to overcome this dilemma (Dubes & Jain, 1979). To achieve the desired clustering, the very large data suites are reduced to points in the hyperspace based on similarities and differences in the concentrations of different elements (Oladipo, 1987). Data evaluation begins first by considering each observation in itself as the smallest cluster. This is then followed by fusing of two nearest neighbours to replace the old clusters (agglomerative clustering). The procedure continues this way until only one cluster is left at the last stage (Milligan & Cooper, 1985; Ibàñez et al., 2013). The clustering pattern which is based on hierarchy completes (n – 1) fusion steps, resulting to a tree of clusters called dendogram (Jain & Dubes, 1988). Van der Sloot, (1980) classified coal by trace element analysis using INAA, cluster analysis, and leaching of precipitator ash. Hopke *et al*., (1987) also interpreted multielemental INAA data using pattern recognition methods. Ewa *et al*., (1992) used cluster technique to study the elemental concentration of cored Nigerian river sediment, and using data evaluation of trace elements, Ewa, (2004) was able to segregate the Nigerian coal.

The results of cluster analysis can be presented in form of dendograms which are graphical representations that display the groups formed by clustering of observations and their similarity levels. In the Ward method (Ward, 1963), points representing individual samples are tied together to form a cluster. A plot is made of the fusion of successive samples against fusion or similarity coefficients as the membership of a group or heterogeneity increases (Oladipo, 1987). Individual samples also exist, and are referred to as outliers. Outliers sometimes make it difficult in distinguishing clusters in residual data. The way out of this problem according to Oladipo, (1987) is to rid the data of outliers and repeat the clustering procedures. Meaningful explanations can be made of outliers, hence, they are not out-rightly discarded (Oladipo, 1987).

Methodology

The Study Area

Plate I. The study area indicating sampled locations

Kaduna Refining and Petrochemical Company (KRPC) and its environs which is partly hosted by Rido village lies in the southern area of Kaduna city, occupying an area between latitude 10^o, 24' 22.32" - 10^o 25' 18.8" N and longitude 7^o, 29' 9.6" - 7^o 29' 55.2" E**.** The major occupation of the inhabitants of the host communities is farming. In addition to this, petroleum industry-related activities take place on daily basis in the area.

Sample Collection

Fifteen (15) soil samples within the top 0 – 20 cm were collected from farmlands around the Refinery Complex (including Rido village) using an Auger and labeled S01, S02,…, S15. Similarly, five (5) sediment samples were collected from a river and a stream that traverse the study area and labeled SD1, SD2, .., SD5. The samples were thereafter dried in an oven at 100℃ for 48 hours.

Sample preparation, irradiation, and counting

The sample preparation for INAA determination of elemental pollutants commenced in the NAA preparation room with the measurement of a quantity of finely crushed samples to be irradiated using the four-digit weighing balance (Metler AE 166, Delta range).

The INAA was carried out using a 31 kW miniature neutron source reactor with a neutron flux of 5 x 10^{11} n/cm^2/s. Crushed samples of soils and sediments were weighed into pre-treated polyethylene bags and heat - sealed in 7 cm^3 rabbit capsules which were packaged into bigger vials in readiness for irradiation. The vials containing the rabbit capsules were sent to the reactor irradiation sites using the pneumatic transfer system. The Nigeria Research Reactor – 1 (NIRR – 1) at the Centre for Energy Research and Training (CERT), Ahmadu Bello University, Zaria, contains four irradiation channels in which short and long irradiations were done. This was followed by measurements of radioactivity of induced radionuclides in a process known as counting.

The concentration c_x of each element in the sample is expressed as:

$$c_x = c_s \frac{w_s}{w_x} * \frac{A_x}{A_s} \tag{1.1}$$

where c_x and c_s are the concentrations of sample and standard respectively.

Table 1: Typical irradiation and counting schemes (Jonah *et al.*, 2006)

Neutron flux /irradiation channel	Procedure	T_{irr}	T_d	T_m	Activation products
$1x10^{11} n/cm^2$ s /outer irradiation channels (B4, A2)	S1	2 min	2-15 min	10 min	^{28}Al, ^{27}Mg, ^{38}Cl, ^{49}Ca, ^{66}Cu, ^{51}Ti, ^{52}V, ^{116m}In
	S2	2 min	3-4 h	10 min	^{24}Na, ^{42}K, ^{165}Dy, ^{56}Mn, ^{152m}Eu
$5\times10^{11} n/cm^2 s$ /inner irradiation Channels (B1, B2, B3, and A1)	L1	6 h	4-5 d	30 min	^{24}Na, ^{42}K, ^{76}As, ^{82}Br, ^{140}La, ^{153}Sm, ^{198}Au,
	L2	6 h	10-15 d	60 min	$^{139}Np(U)$, ^{72}Ga, ^{122}Sb ^{46}Sc, ^{141}Ce, ^{60}Co, ^{51}Cr, ^{134}Cs, ^{152}Eu, ^{177}Lu, ^{131}Ba, ^{86}Rb, ^{182}Ta, ^{160}Tb, ^{175}Yb, $^{233}Pa(Th)$, ^{65}Zn, ^{59}Fe, ^{181}Hf.

Table 1 gives the details of both the irradiation and counting regimes used. Finally, the analysis software, WINSPAN was used for quantitative spectrum analysis.

A very sensitive (high-resolution) gamma-ray spectrometer was used to register the delayed gamma, and the computer-coupled multichannel analyzer sorted the delayed gamma according to their energies. Finally, the gamma-ray acquisition system used the computer software (MAESTRO – 32) to measure the associated activity.

Cluster Technique

The method of SINGLE LINKAGE (SAS, 1995) and the Squared Euclidian distance were chosen to determine the similarities in the clustering as a measure of minimum distance between clusters. The technique (METHOD = SINGLE LINKAGE) was employed according to Ewa, (2004) as introduced by Florek *et al.*, (1951) and Sneath, (1957) to observe the relationship between the concentration of each element in one sample and another.

Results and Discussion

INAA

Results of instrumental neutron activation analysis showed that Twenty Five (25) elements (Mg, Al, Ti, V, Mn, Na, K, As, La, Sc, Cr, Fe, Co, Zn, Rb, Sb, Cs, Ba, Eu, Yb, Lu, Hf, Ta, Th, and Dy) were detected. These elements which cut across major, minor and trace elements were detected in varying concentrations in both soils and sediments.

Cluster technique

The results of data evaluation yielded two dendograms as presented in Figures 1 and 2. The dendograms resulting shows that there are four clusters (1, 2, 3 and 4).

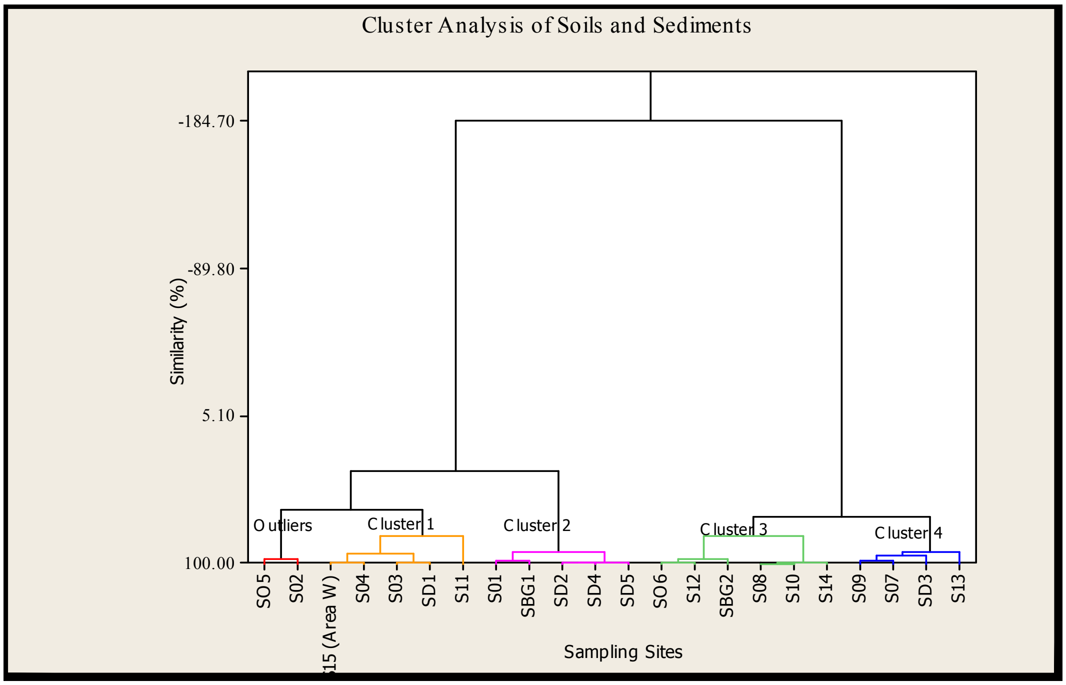

Figure 1: Cluster dendogram for soils and sediments

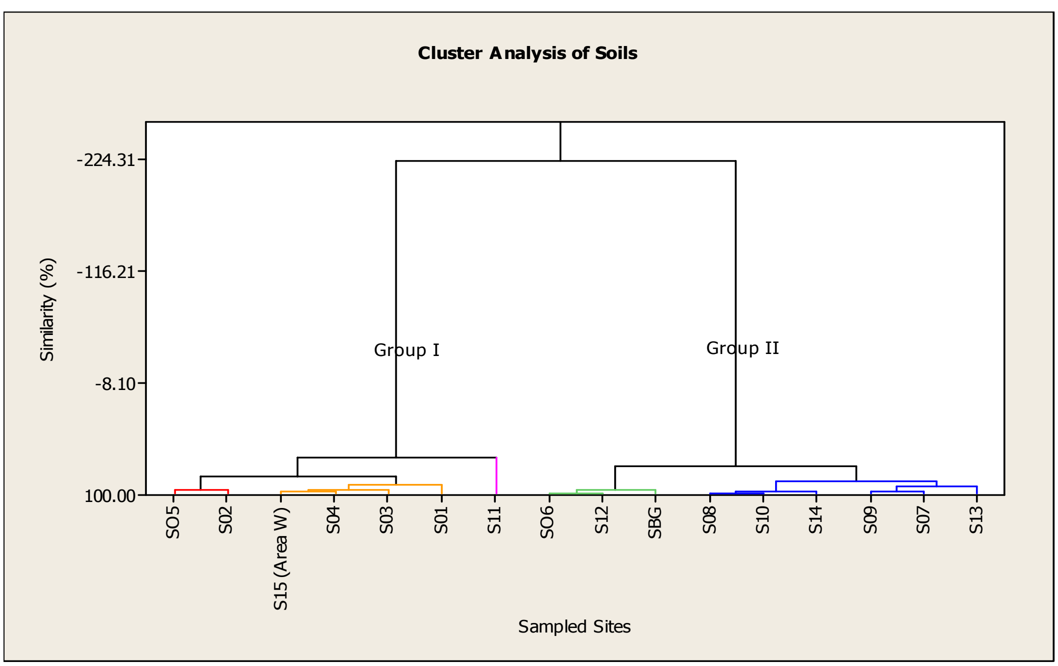

Figure 2: Cluster dendogram for soils only

Before one can start the grouping of samples, a measure of similarity must be defined (Ibàñez et al., 2013). Each sample must be compared with every other sample and those that are very "similar" in chemical composition will then be subsequently found in the same cluster, while "dissimilar" ones will be found in different clusters (Ibàñez et al., 2013).

The similarity between two elements can be measured in many different ways but in this work, "distance" has been chosen. In most cases, the "similarity" between the samples is rather hard to define, while the "distance" that is inversely proportional to it is much easier to visualize and handle.

Five samples (S15, S04, S03, SD1, and S11) clustered together to form a group referred to as cluster 1, with the clustered samples having a similarity level of 82.19 %, five other samples (S01, SBG1, SD2, SD4, and SD5) formed cluster 2 at 92.09 % similarity level, six samples (S06, S12, SBG2, S08, S10, and S14) formed cluster 3 with member samples having a similarity level of 81.87 %, while four samples (S09, S07, SD3, and S13) formed cluster 4 at 92.34 % similarity level. The deviating elements responsible for the separation of the samples into clusters include Ta, Lu, Eu, Sb, Cs, and Yb, having grand centroids of 1.6, 5.1, 5.1, 1.5, 5.4, and 9.4 respectively. The lower the value of the grand centroid, the better the discriminating power between the samples. Cluster 1 is made up of members from different environments (fadama soil, sediments, sludge, and upland soils); cluster 2, though mixed consists mainly of sediment samples; cluster 3 is made up, mainly of upland soils while cluster 4 is also mainly upland but rocky soils with a member from the sediments.

When the elemental concentrations of the sediments were removed from the data set, the resulting dendogram is as given in Figure 2. As a result of this process, the clusters again regrouped to form two main clusters or groups (group I and group II), giving rise to a clearer picture of groupings. The samples were clearly divided along the lines of the soil environments. Cluster group I which was made up of soil samples of fadama origin (S01, S02, S03, S04, S05, S11, and S15) clustered at a similarity level of 65.75% and consists mostly of samples collected from farmlands around Rido village used mainly for cultivation of rice. The sample tagged S15 (Area W) was taken from the refinery sludge pit which has water environment similar to the fadama soils. Cluster group II (which comprised of samples S06, S07, S08, S09, S10, S12, S13, and SBG that clustered at 73.66 % similarity level) on the other hand consists of upland soils from the farmlands bordering the northern part of the refinery complex and part of Rido village that occupies higher elevations than the fadama environments. These upland regions are characterized by hard-pan concretions in both the dry and rainy seasons where metals are highly immobile.

The deviation of the elements resulting into clusters and the interstitial nature of the membership of some clusters (especially Figure 1) may be attributed to the complex nature of some elements that enable them to bind with many other elements and react differently. The reaction of a particular element depends on its solubility, physico-chemical properties of water, water content of soil, and other physical/chemical properties which make predictions of their existence in a particular environment rather difficult. This probably explains the reason for lack of systematic trend in the membership of the clusters.

Conclusion

Twenty five (25) elements (major, minor, and trace) were detected in soils and sediments around the refinery complex. Subjecting the data to clustering procedures revealed that the behaviour of elements in a water environment differs from that in dry soil. The cluster technique has therefore proved to be a veritable tool in handling large data suites based on similarities or dissimilarities of members of a set of experimental objects. In the final analysis, the soil samples were segregated into two large groups/clusters with memberships belonging to either fadama soil or upland soil.

Acknowledgements
The authors wish to acknowledge the Management of the Kaduna Refining and Petrochemical Company (KRPC) for granting unrestricted access to their areas of operation and facilities. We also wish to specially thank Mr. Solomon Olorunselu for the valuable assistance. Appreciation goes to the staff of the Chemical Laboratory for their supports towards the successful completion of the measurements carried out in the laboratory.

References

Dube, S. R., & Jain, A. K. (1979). Validity studies in clustering methodologies. *Pattern Recognition*, 11, 235 – 254.

Everitt, B. S. (1979). Unresolved problem in cluster analysis. *Biometrics,* 21, 362 – 375.

Ewa, I. O. B., Oladipo, M. O. A., & Dim, L. A. (1992). Cluster analysis of elemental concentration of cored Nigerian river sediments. *Journal of Environmental Science and Health*, A27, 1 – 11.

Ewa, I. O. B. (2004). Data evaluation of trace elements determined in Nigerian coal using cluster procedures. *Applied Radiation and Isotopes*, 60, 751 – 758.

Florek, K., Lukaszewicz, J., Perkal, J., Steinhans, H., & Zubrzycki, S. (1957). *Colloq. Math*., 2, 282.

Hopke, P. K., Martin, R. C., & Evins, M. A. (1987). The interpretation of multielemental INAA data using pattern recognition methods. *Journal of Radioanalytical and Nuclear Chemistry*, 112(1), 215 – 222.

Ibàñez, A., Larrañaga, P., & Bielza, C. (2013). Cluster methods for assessing research performance: Exploring Spanish computer science. *Scientometrics*, Springer DOI: 10.1007/s11192-013-0985-9.

Jain, A., & Dubes, R. (1988). *Algithorms for clustering data.* Englewood Cliffs: Prentice-hall.

Jonah, S. A., Umar, I. M., Oladipo, M. O. A., Balogun, G. I., & Adeyemo, D. J. (2006). Standardization of NIRR-1 irradiation and counting facilities for instrumental neutron activation analysis. *Applied Radiation and Isotopes*, 64, 818-822.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50(2), 159 – 179.

Oladipo, M. O. A. (1987). Trace element analysis of Corinthian pottery and related clays. Unpublished Ph.D Dissertation, Victoria University, Manchester.

SAS (1985). *The cluster procedure, SAS users' guide statistical version, 5th Edition.* SAS Institute Inc., Cary NC, 255 – 315.

Sneath, P. H. A. (1957). Application of computers to taxonomy. *Journal of General Microbiology*, 17, 201 – 226.

Van der Sloot, H. A. (1980). Classification of coal by trace analysis using INAA, cluster analysis and leaching of precipitator ash. Proceedings of the Fourth International Conference on Nuclear Methods in Environmental and Energy Research, University of Missouri, Colombia, USA.

Ward, J. A. (1963). Hierarchical Grouping to optimize an objective function. *Journal of American Statistical Association*, 58, 236.