The Nucleus, **45** (3-4) 2008 : 131-135

The Nucleus
A Quarterly Scientific Journal of Pakistan Atomic Energy Commission
NCLEAM, ISSN 0029-5698

PRINCIPAL COMPONENT ANANLYSIS OF AIR PARTICULATE DATA FROM THE INDUSTRIAL AREA OF ISLAMABAD, PAKISTAN

*S. WAHEED, N SIDDIQUE and M. DAUD

Chemistry Division, PINSTECH, P.O. Nilore, Islamabad, Pakistan

(Received October 21, 2008 and accepted in revised form December 2, 2008)

A Gent air sampler was used to collect 72 pairs of size fractionated coarse and fine (PM_{10} and $PM_{2.5}$) particulate mass samples from the industrial zone (sector 1-9) of Islamabad. These samples were analyzed for their elemental composition using Instrumental Neutron Activation Analysis (INAA). Principal component analysis (PCA), which can be used for source apportionment of quantified elemental data, was used to interpret the data. Graphical representations of loadings were used to explain the data through grouping of the elements from same source. The present work shows well defined elemental fingerprints of suspended soil and road dust, industry, motor vehicle exhaust and tyres, and coal and refuses combustions for the studied locality of Islamabad.

Keyword: Principal Component Analysis (PCA), Source apportionment, Islamabad, Gent sampler, Instrumental Neutron Activation Analysis (INAA), Air particulates

1. Introduction

Combustion of fossil fuels for power generation, rapid industrialization, vehicular and industrial exhaust, modernized farming along with household and industrial waste, contribute significantly to the amount of potentially toxic matter prevalent in the atmosphere [1,2]. Metropolitan cities with industrial areas and heavy traffic are more vulnerable to this issue. The prerequisite for healthy and safe living in these regions requires the identification and remediation of all possible sources of environmental hazards. In view of these problems, a project was initiated to study the level of air pollution in the industrial areas of Islamabad, Pakistan; i.e. sectors I-9 and I-10. Brick kilns are located in the suburbs of Islamabad and there are also substantial amounts of vehicular emissions. To estimate the pollution level and its trends, our main objective was to obtain inorganic elemental data of airborne particulate matter (PM) from this area and to assess the possible sources; whether natural or anthropogenic [3].

Modern multielement analytical techniques can generate large multivariate data sets. The magnitude and complexity of the data pose a challenge to its full interpretation. To aid in extracting information from such voluminous and complicated data sets, numerous data analysis approaches based on mathematical and statistical methods have been developed [4]. The use of modern computers has further simplified data evaluation where the complex mathematical calculations have been reduced to simple commands. Principal component analysis (PCA) is one such tool that can identify different patterns in multivariate data [5]. The current paper describes the application of PCA to extract information from the elemental data obtained using instrumental neutron activation analysis (INAA) to analyze PM samples collected in the I-9 area of Islamabad. The study was performed under the aegis of IAEA Coordinated Research Program (CRP) entitled "Better management of the environment, natural resources and industrial growth through isotope and radiation technology" for the Asia and the Pacific region.

2. Experimental

A size-fractionating aerosol sampler using Gent-type stacked filter unit (SFU) was used to collect 72 pairs of filters from the industrial area I-9, located in the vicinity of Islamabad city. The airborne particulate matter was collected in two size fractions by a sequential filtering technique; coarse (aerodynamic diameter d ~ 10-2.5 μm) and fine (d < 2.5 μm) size fractions. Sampling was carried out from October 1998 to January 1999 [3].

* Corresponding author : swaheed@pinstech.org.pk

Elemental analysis was carried out using INAA. Pakistan Research Reactor II (PARR-II) which is a 27 kW miniature neutron source reactor (MNSR) with a nominal thermal neutron flux of 1x 10^{12} cm^{-2} s^{-1} and the 9 MW Pakistan Research Reactor-I (PARR-I) with a nominal thermal neutron flux of 9 x 10^{13} cm^{-2} s^{-1} were used for the irradiations of the targets [6]. An HPGe detector (Canberra Model AL-30) connected to a PC-based Intertechnique multichannel analyzer (MCA) measuring system was used for the acquisition of spectra. The system resolution and efficiency are 1.9 keV and ~10% for the 1332.5 keV peak of ^{60}Co with peak to Compton ratio of 40:1. Intergamma, version 5.03 software is used for data acquisition. Further details about the experimental procedures are described in our previous publication[3]. The elemental data for PM was analyzed using excel Add-ins provided by the Centre for Chemometrics, School of Chemistry, University of Bristol and MATLAB v 6.0.

2.1. Data analysis

A total of 31 elements namely Ag, Al, Ba, Ca, Ce, Co, Cr, Cs, Eu, Fe, Hf, Ho, Hg, Lu, Mg, Mn, Mo, Na, Nd, Rb, Sb, Sc, Se, Sn, Ta, Tb, Th, V, Yb, Zn and Zr were analysed in coarse PM samples while a total of 26 elements namely Al, Ca, Ce, Co, Cr, Cs, Eu, Fe, Hf, Lu, Mg, Mn, Mo, Na, Nd, Rb, Sb, Sc, Se, Sn, Ta, Tb, Th, V, Yb, and Zn were analysed in fine PM samples using sequential and long irradiation protocols. More details about the elemental concentration and how they are obtained are produced in our earlier publication [3].

Prior to analysis the data for all those elements were discarded which were measured in less than 50% samples, thereby reducing the elements to Al, Ce, Co, Cr, Cs, Eu, Fe, Hg, Mn, Na, Rb, Sb, Se, Tb, Th, and Zn in coarse and Al, Ce, Co, Cr, Cs, Eu, Fe, Mn, Na, Sb, Sc, Th, and Zn in fine PM samples. Any missing values for any element were replaced by half of the minimum of elemental concentration [7]. Since the elemental data covers a huge concentration range from fraction of ng/g to μg/g levels therefore the concentrations were converted to log scale and then standardized to get values close to each other.

2.2. Principal Component Analysis (PCA)

The elemental data contain samples in rows and concentration of elements in columns. Mathematically, PCA [8-13] is defined as:

$$X = T P + E$$

where *X* is the data matrix, *T* is called the scores, *P* is the loadings and *E* is the error matrix. PCA combines original variables into a set of new variables which are orthogonal to each other called principal components (PC). Each PC is associated with an eigenvalue which represents the importance of each principal component.

In our study the first five eigenvalues for coarse PM explain 87% of the data and first six eigenvalues for the fine PM explain 85% of the total variation in the PM data. The latter eigenvalues were small and considered insignificant. In the scores matrix each row represents each sample and contains as many rows as the original data. Since in this study the individual PM samples are not compared therefore the scores have not been discussed. The loadings consist of series of row vectors and contain as many columns as the original data matrix; therefore each column represents each element [8].

3. Results and Discussion

Size fractionated PM samples were collected on filter paper around the industrial area of I-9 in Islamabad from the 2nd of October 1998 to the 15th of January 1999. These were analyzed using INAA. The original data tables of elemental concentrations have been presented elsewhere[3] and have not been incorporated in this paper due to their large size. The trace elements data was subjected to PCA using MATLAB. In the data processing steps the number of elements was reduced from 31 to 16 in coarse PM and from 26 to 13 elements in fine PM. After the application of PCA, the original variables are reduced to a number of smaller significant principal components that can be easily displayed graphically and thus facilitate in perceiving different trends in the data [14]. Therefore one PC from the loadings is plotted against the other and can provide detailed information about those elements that tend to correlate with each other. Accordingly graphical representation of the loadings for both coarse and fine elemental data was performed and is presented in Figures 1 and 2 respectively. In Figure 1 PC1 has been plotted against PC2 for coarse PM elemental data to obtain an XY scatter plot on which the points for each element are labelled. In this figure the elements Al, Mn and Na are clearly clustered together as are Eu, Th and Cs, Ce and Se and Fe, Co, Cr and Tb. From Figure 2 once again Al, Mn and Na form one cluster while Co and Cs and Sc and Th form two

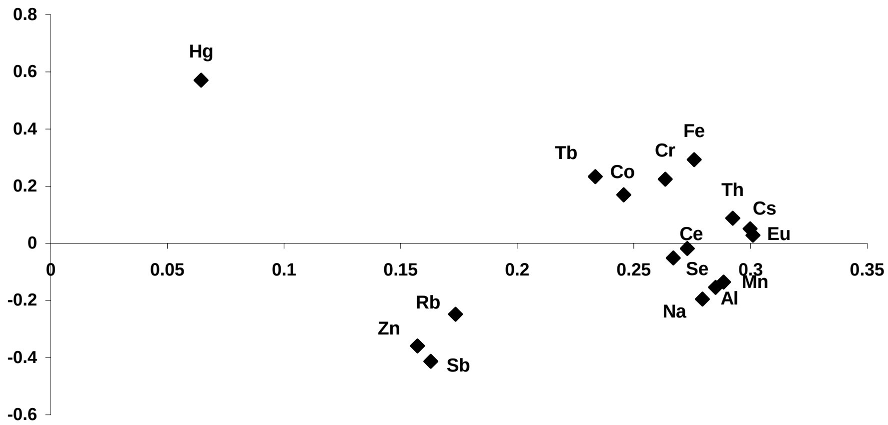

Figure 1. PC1 Vs PC2 loadings for trace elements in coarse PM of I-9.

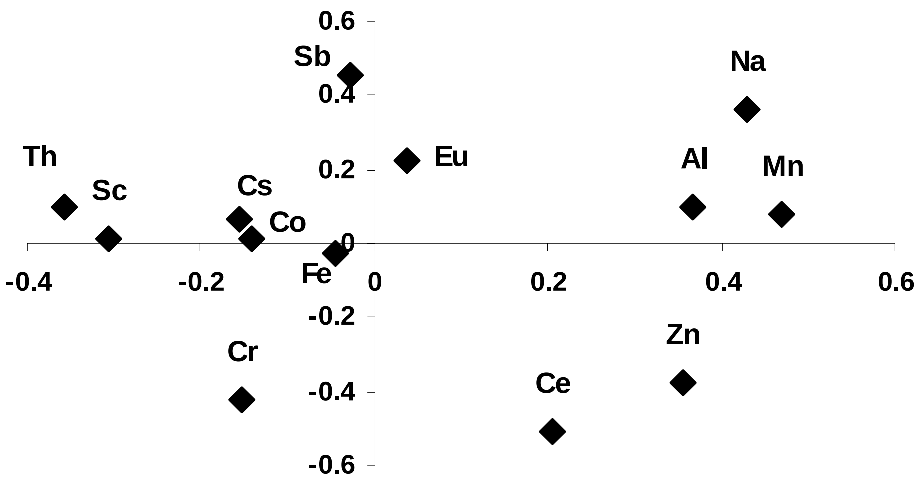

Figure 2. PC2 Vs PC4 loadings for trace elements in fine PM of I-9.

more. This indicates that these PC plots have one common set of elements (Al, Mn and Na) with good correlation while the rest of the elements in each satellite cluster are not in good correlation. As Al, Mn and Na have a strong relationship with each other therefore they originate from the same source. Since these elements are present in soil they may be representative of the suspended particles of soil dust [15,16]. These plots also show that Hg and Zn are present in the particulate matter without forming any relationship with any other element and they have their independent source.

In order to further categorize the elemental clustering in the particulate matter three-dimensional PC plot using MATLAB were also constructed. The idea behind this is to see the best possible clustering of the elements via the option of rotation of plot in the MATLAB software. The MATLAB was used to plot PC1, PC2 and PC3 of coarse and fine PM data loadings and then each element was labelled. These plots for the coarse and fine PM are shown in Figures 3 and 4 respectively. Through the rotation of these plots elements which cluster together can be identified. In Figure 3 the plot for coarse PM the elements Al, Mn, Na and Co, Ce, Sb and Se and Eu, Fe, Cr, Tb and Th were found to form three groups while in Figure 4 the plot for fine PM the elements Al, Mn and Na, Co, Cr, Cs, Eu and Fe and Th and Sc form three groups. From these MATLAB plots it is clear that Hg, Rb and Zn do not form any relationship with any other element and all three of them fall on different faces of the plots. Sb also seems to behave in an isolated manner as supported by two dimensional PC plots. The 2 and 3 dimensional (2-D and 3-D) PC plots predict correlation between Al, Mn and Na for both mass fractions implying suspended road dust or soil as their source. For the coarse mass fraction these plots further show correlation between Ce and Se and Cr, Fe and Tb while for the fine fraction good correlation is found between Th and Sc and Co and Cs. Therefore

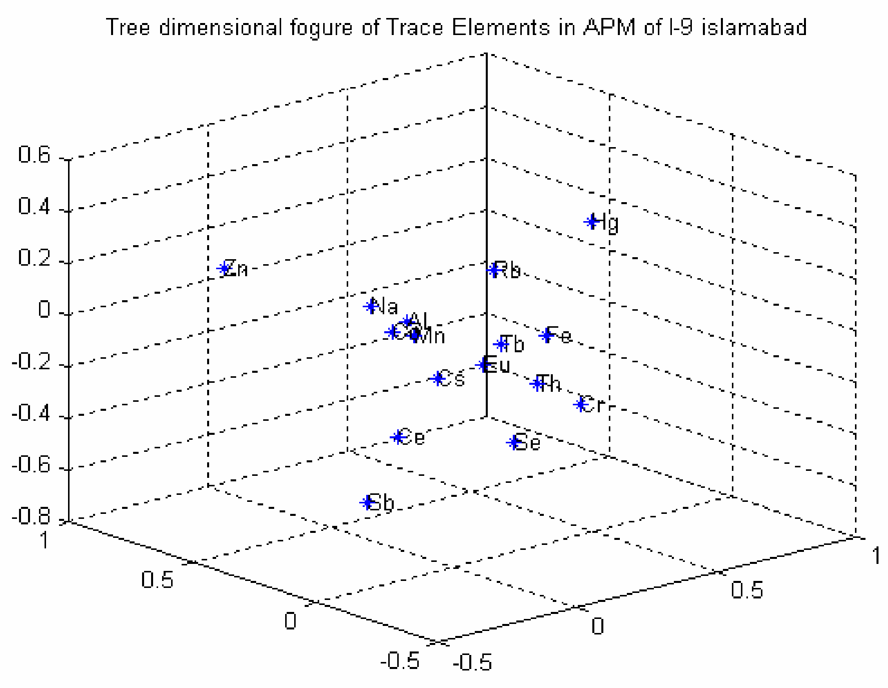

Figure 3. Three dimensional plot of loadings vectors (PC1, PC2 & PC3) for trace elements in coarse PM of I-9, Islamabad.

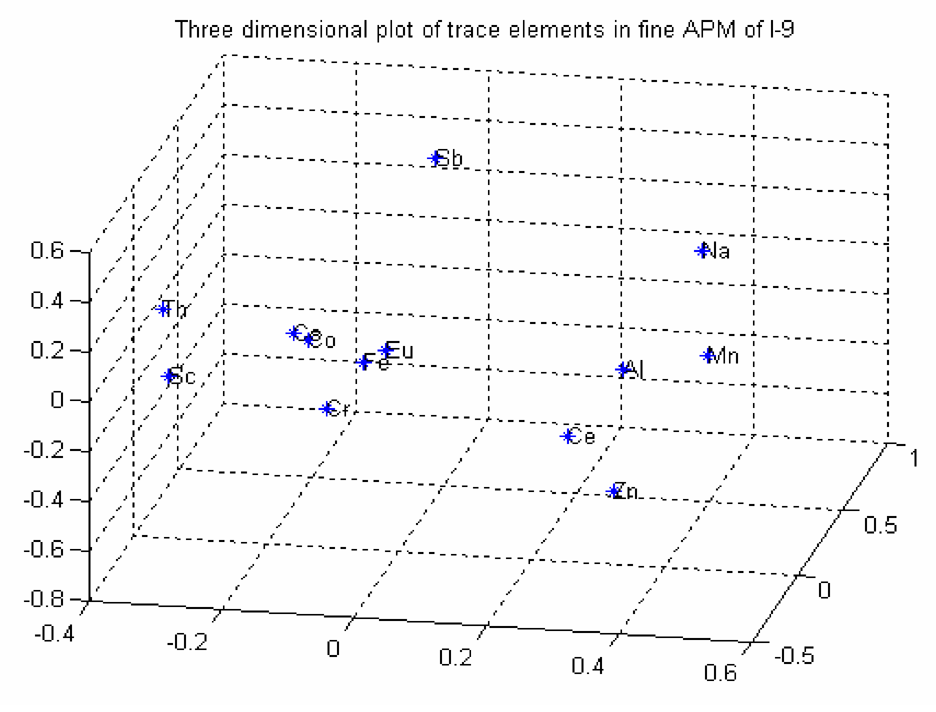

Figure 3. Three dimensional plot of loadings vectors (PC1, PC2 & PC3) for trace elements in fine PM of I-9, Islamabad.

these elements may originate from industrial sources. Similarly the remaining elements especially Eu, Rb, Sb and Zn have their independent sources. Eu is a rare earth element and its presence may be due to 3-way catalyst used in vehicles. Sb occurs in the atmosphere mainly through the combustion of refuse and plastics while Zn could be due to the motor vehicle exhaust and from the wearing of the tyres [17,18].

4. Conclusions

Principal component analysis was applied to multielement data obtained using INAA for coarse and fine air particulate matter collected in the industrial area of Islamabad. Significant principal components were selected using eigen values. The loadings provided different elemental patterns corresponding to different sources. Further authenticity of the elemental patterns was confirmed through the use of 3-D PC plots using MATLAB. The two schemes used in this work showed different elemental patterns as fingerprints of the suspended soil and road dust, industry, motor vehicle exhaust and tyres, and coal and refuse combustions.

Acknowledgements

We gratefully acknowledge the IAEA for awarding a fellowship under the IAEA Coordinated Research Program (CRP) to perform this study. Furthermore, this work would not have been possible without the software provided by Professor Dr. R.G. Brerton, School of Chemistry, University of Bristol, UK.

References

[1] R. M. Davidson and L. B. Clarke, Report IEAPER/**21,** IEA Coal Research, London, (1996).

[2] H. Lachas, R. Rachaud, K. E. Jarvis, A. A. Herod, D. R. Dugwel and R. Kandiyoti, Analyst **124** (1999) 177.

[3] M. Wasim, A. Rahman, S. Waheed, M. Daud and S. Ahmad, J. Radioanal. Nucl. Chem. **258** (2003) 397.

[4] E. J. Baum, Chemom. Intell. Lab. Sys. **3** (1988) 91.

[5] S. Wold, C. Albano, W. J. Dunn, K. Esbensen, P. Geladi, S. Hellberg, E. Johanasson, W. Lindberg, M. Sjostrom, B. Skagerberg, C. Wikstrom and J. Ohman, Multivariate data analysis: Converting chemical data tables to plots, VII Internat. Conf. Computer Chme. Res. Educ. Garmisch-Partenkirchen, June 10-14, (1985) 166.

[6] S. A. Chaudhary, S. Waheed and S. Ahmad, The Nucleus **36** (1999) 13.

[7] B. Grung and R. Manne, Chemom. Intell. Lab. Sys. **42** (1998) 125.

[8] R.G. Brerton, Chemometrics: Data Analysis for the laboratory and chemical plant, Wiley, Chischester, 2003.

[9] P. V. Espen and F. Adams, Anal. Chim. Acta **150**, (1983) 153.

[10] V. Zitko, Data evaluation in environmental research, Envchem/0205001 (2002).

[11] C. Chatfield and A. J. Collins, Introduction to multivariate analysis, Chapman and Hall, London, 1980

[12] B. G. M. Vandeginste, D. L. Massaert, L. M. C. Buydens, S. Dedjong, P. J. Lewi and J. Smeyers-Verbeke, Handbook of Chemometrics and Qualimetrics, Part A&B, Elsevier, Amsterdam, 1998.

[13] M. Wasim, M. S. Hassan and R. G. Brereton, Analyst **128** (2003) 1082.

[14] M. E. Chase, S. H. Jones, P. Hennigar, J. Sowles, G. C. H. Harding, K. Freeman, P. G. Wells, C. Krahforst, K. Coombs, R. Crawford, J. Pederson and D. Taylor, Mar. Pollut. Bull. **42** (2001) 491.

[15] W. Liua, P. K. Hopke, Y. Hanb, S. M. Yib, M. T. Holsenb, S. Cybartc, K. Kozlowskic, and M. Milliganc, Atmos. Environ. **37** (2003) 4997.

[16] D. D. Cohen, D. Garton and E. Stelcer, Nucl. Instr. Meth. Phys. Res. B **161** (2000) 775.

[17] T. Weizhi, Workbook on reactor neutron activation analysis (NAA) on airborne particulate matter (APM). IAEA/RCA Handbook, NAHRES-53, Vienna (2000).

[18] E. P. Trepata, E. Kimb, P. Paateroc and P.K. Hopke, Atmos. Environ. **41** (2007) 5921.