

Multivariate receptors model for environmental source apportionment

Maurizio Caselli

Chemistry Department - University of Bari, Via Orabona, 4 -70126 Bari,
caselli@chimica.uniba.it

Eleonora Andriani

Chemistry Department - University of Bari, Via Orabona, 4 -70126 Bari,
eleonora.andriani@chimica.uniba.it

Abstract: Three receptors models(CMB, APCS, PMF) were applied to the same dataset and the results were compared. PMF gives the better reconstruction of the data matrix whichever type of error is considered. The nature of the sources determined is essentially the same for the considered models but sources composition and their contributions are somewhat different.

Keywords: PM, receptor model, source apportionment.

1. Introduction

Atmospheric pollutant dispersion models have presently reached high reliability but they are characterized by some difficulties. First of all, these models present a considerable mathematical and physical complexity, then they require a huge amount of calculation and a not negligible cost. The greater drawback is anyway related to the need of reliable data about the emitted pollutants nature and quantity; these data are usually not available using dataset stored in pollutant emission inventories.

Receptor models can be a useful tool to overcome this difficulty. These models start by collecting samples of airborne pollution in the sampling site under investigation and by determining concentration of chemical species that are concerning pollutants or can be markers of particular sources.

From a mathematical point of view, Chemical Mass Balance (CMB) is the most reliable receptor model. However, it requires an “a priori” knowledge of the number and composition of the meaningful pollutant sources; this knowledge can be obtained by experiments or assuming available literature data as valid for the sampling area. The main CMB constraint is the necessity of knowing the number and the composition of pollutant sources; other multivariate models try to overcome this difficulty by increasing the number of the samples collected in a given sampling site and by supposing that the source profiles remain constant during all sampling collection.

Some of the most employed multivariate receptor models are Target Transformation Factor Analysis (TTFA), Absolute Principal Component Scores (APCS) and Positive Matrix Factorization (PMF). In this paper, two of these models (PMF and APCS) are analyzed and compared between them and with CMB.

2. Application to real data

Three receptor models (CMB, APCS and PMF) have been applied to a set of fifty samples of atmospheric PM₁₀ particulate collected in four sites near the sea and close to a power plant, (not operating during the sampling period). 20 parameters (EC, OC, Cl⁻, NO₃⁻, SO₄²⁻, Na⁺, NH₄⁺, K⁺, Mg²⁺, Ca²⁺, Al, Si, P, Ti, V, Mn, Fe, Ni, Cu, and Zn) were determined in all samples. EC and OC were measured by thermal optical technique; Cl⁻, NO₃⁻, SO₄²⁻, Na⁺, NH₄⁺, K⁺, Mg²⁺, Ca²⁺ by ionic chromatography and all other compounds by X-Ray Fluorescence.

3. Results and discussion

3.1 Errors on data matrix reconstruction

Some results obtained by comparing different models are shown later on. First of all, principal components analysis was applied to the dataset and it was observed that the first seven components explain more than 90% of the total variance.

As concerns CMB, seven sources were considered relevant for the sampling site (Earth's Crust, Marine Aerosol, Field Burning, Ammonium Nitrate, Ammonium Sulphate, Traffic and Sulfate); US - EPA Speciate Database was employed in order to obtain the profiles of the first six sources. An additional Sulfate source was considered in order to explain the total sulfate of the samples.

In Figure 1, the total analysed mass of each sample is compared to the sum of mass contribution of the reconstructed data matrix for the three models.

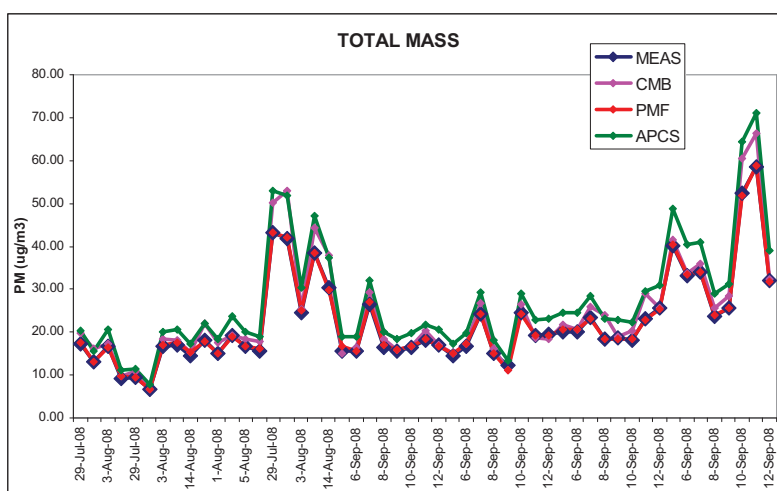


Figure 1: Samples total mass (mean values) measured and estimated by models.

As shown, PMF reproduces very well the sample mass while APCS and, in some samples CMB, overestimates the sample mass. This result is obtained also for the mean value of single parameters, as it is shown in Figure 2 for major parameters.

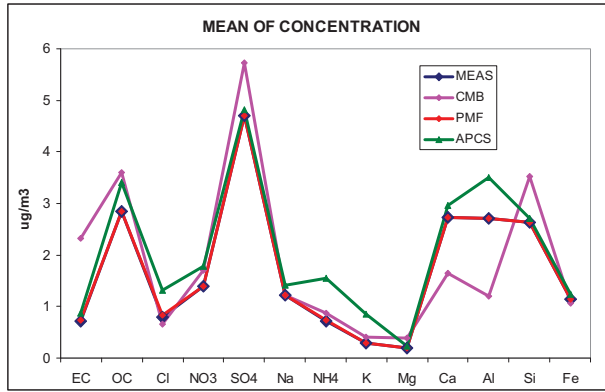


Figure 2.a: Mean concentration of major parameters estimated by models.

Different expression can be used to estimate the error on the reproduction of the data matrix as the product of profile and contributes matrices. These different types of errors are reported for the three models in the following table:

		CMB	PMF	APCS
Percent Relative Mean Square Root Error	$RSE = \frac{100}{n \cdot m} \sqrt{\sum_i \sum_j \left(\frac{X_{ij} - X_{ij}^{sp}}{X_{ij}^{sp}} \right)^2}$	8.5%	1.9%	7.2%
Absolute Mean Square Root Error	$RSE = \frac{1}{n \cdot m} \sqrt{\sum_i \sum_j (X_{ij} - X_{ij}^{sp})^2}$	0.043 (ug/m ³)	0.004 (ug/m ³)	0.014 (ug/m ³)
Relative Frobenius Norm	$RFN = \sqrt{\frac{\ X\ _{fro} - \ X^{sp}\ _{fro}}{\ X^{sp}\ _{fro}}}$	0.37	0.044	0.34
Q Error	$QE = \frac{1}{n \cdot m} \sqrt{\sum_i \sum_j \left(\frac{X_{ij} - X_{ij}^{sp}}{\sigma_{ij}^{sp}} \right)^2}$	0.47	0.079	0.35

Table 1: Errors on data matrix reconstruction.

3.2 Sources and contributions apportionment

CMB assumes number and source profile as known data. Therefore, the contributions only remain to be calculated: in this dataset, Traffic and Earth's Crust cover 72% of the total mass.

In Figure 3 the profiles and the percent mean contributions (on each bar in the figure) of the seven sources calculated by PMF are shown; the source denomination was given

taking into account the comparison with CMB, but it must be stressed out that both composition and weights are not completely similar to CMB ones. For example, PMF subdivide the CMB Traffic profile in two different sources that can be, at large, identified as Traffic and Resuspended Crust Matter.

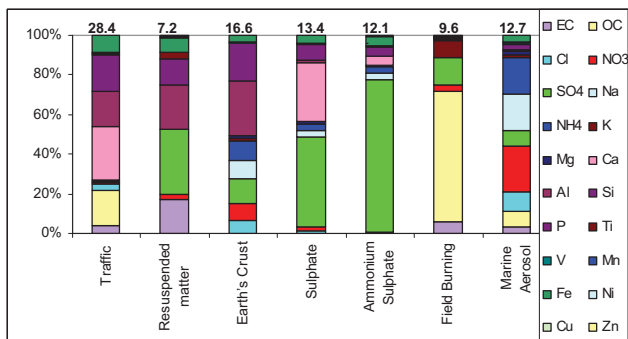


Figure 3: Profiles and percent mean contributions of PMF sources.

In Figure 4 the profiles and percent mean contributions (on each bar in the figure) of the seven sources calculated by APCS are shown. APCS profiles have a greater simplicity since the matrix presents a large number of zeros; besides, different parameters tend to concentrate on particular sources as an effect of the Varimax rotation of the loading matrix.

Using APCS it is more difficult to identify the nature of some sources such as, for example, the seventh source that is simply indicated as “Calcium” source.

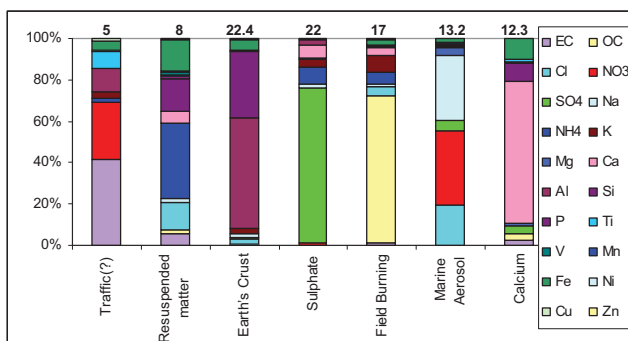


Figure 4: Profiles and percent mean contributions of APCS sources.

4. Conclusions

Among the considered source apportionment models, PMF better reconstructs the data matrix as the product of profile and contribute matrices. However, this does not assure that the source composition and the related weights are more realistic than those used in CMB model. More accurate emission inventories specific for the investigated area would give more reliable data. However, source apportion models are useful to give a general idea of the contribution of well characterized sources to the particulate composition.

References

- Henry R.C., Lewis C.W., Hopke P.K., Williamson H.J. (1984) Review of receptor model fundaments, *Atmos. Environ.*, 18(8), 1507-1515.
- Hopke P.K. (2003) Recent developments in receptor modeling, *J. Chemometr.*, 17(5), 255-265.
- Thurston G.D., Spengler J.D. (1985) A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston, *Atmos. Environ.*, 19(1), 9-25.
- Guo H., Wang T., Louie P.K.K. (2004) Source apportionment of ambient non-methane hydrocarbons in Hong Kong: application of a principal component analysis/absolute principal component scores (PCA/APCS) receptor model, *Environ. Pollut.*, 129(3), 489-498.
- Bruno P., Caselli M., de Gennaro G., Traini A. (2001) Source apportionment of gaseous atmospheric pollutants by means of an absolute principal component scores (APCS) receptor models, *Fresenius J. Anal. Chem.*, 371(8), 1119-1123.
- Marcazzan G.M., Ceriani M., Valli G., Vecchi R. (2003) Source apportionment of PM₁₀ and PM_{2.5} in Milan (Italy) using receptor modelling, *Sci. Total Environ.*, 317(1-3), 137-147.
- Lee S., Liu W., Wang Y., Russel A.G., Edgerton E.S. (2008) Source apportionment of PM_{2.5}: Comparing PMF and CMB results for four ambient monitoring sites in the southeastern United States, *Atmos. Environ.*, 42(18), 4126-4137.
- Caselli M., de Gennaro G., Ielpo P. (2006) A comparison between two receptor models to determine the source apportionment, *Environmetrics*, 17(5), 507-516.
- Chow J.C., Watson J.G., Kuhns H., Etyemezian V., Lowenthal D.H., Crow D. (2004) Source profiles for industrial, mobile, and area sources in the big bend regional aerosol visibility and observational study, *Chemosphere*, 54(2), 185-208.