ELSEVIER

Fine particulate source apportionment using data from the USEPA speciation trends network in Chicago, Illinois: Comparison of two source apportionment models

Michael J. Rizzo[a,*], Peter A. Scheff[b]

[a]USEPA, Office of Air Quality Planning and Standards, Air Quality Analysis Division, Air Quality Analysis Group, Research Triangle Park, NC 27711, USA
[b]University of Illinois – Chicago, School of Public Health, 2121 W. Taylor, Chicago, IL 60612, USA

Abstract

Data from two of the United States Environmental Protection Agency's speciation trends network fine particulate matter sites within Chicago, Illinois were analyzed using the chemical mass balance (CMB) and positive matrix factorization (PMF) models to determine source contributions to the ambient fine particulate concentrations. The results from the two models were compared to determine the similarities and differences in the source contributions. This included examining the differences in the magnitude of the individual source contributions as well as the correlation between the contribution values from the two methods. The results showed that both models predicted sulfates, nitrates and motor vehicles as the three highest fine particle contributors for the two sites accounting for approximately 80% of the total. The PMF model attributed a slightly greater amount of fine particulate to the road salt, steel and soil sources while vegetative burning contributed more in the CMB results. Correlations between the contribution results from the two models were high for sulfates, nitrates and road salt with very good correlations existing for motor vehicles and petroleum refineries. The predicted PMF profiles agreed well with measured source profiles for the major species associated with each source.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Source apportionment; Positive matrix factorization; Chemical mass balance; $PM_{2.5}$; Speciation trends network (STN); Receptor modeling

1. Introduction

In 2000, the United States Environmental Protection Agency (USEPA) established the fine particulate speciation trends network (STN) to expand on its existing $PM_{2.5}$ monitoring activities. The purpose of the network is to characterize individual species which compose the total fine particulate measured at the Agency's federal reference method (FRM) $PM_{2.5}$ monitoring sites. The data from the speciation network serve an important role in aiding the agency in determining which species are the most prevalent in areas of the nation and, thus, allowing for the formulation of control strategies. Secondary

*Corresponding author. Tel.: +1 919 541 3604;
fax: +1 919 541 3613.

E-mail address: rizzo.michael@epa.gov (M.J. Rizzo).

sulfates comprise a large part of the fine particulate in the Eastern part of the United States while secondary nitrates dominate the total $PM_{2.5}$ in parts of California and secondary organics are predominant in the Western United States. The Midwestern section of the country is dominated by both secondary sulfates and nitrates. All areas of the country have been shown to have a large portion of the total fine particulate comprised of organic matter (http://www.epa.gov/airtrends/pm.html).

Another use for the speciation is to determine possible fine particulate sources. Traditional source apportionment techniques have centered around the use of the chemical mass balance (CMB) model which utilizes source profiles and speciated data to determine source contributions for either gaseous or particulate compounds or a combination of both (Li et al., 2003; Watson et al., 1984, 1990, 2002a; Scheff et al., 1994). There have been many analyses conducted regarding volatile organic compounds (Scheff et al., 1996; Kenski et al., 1994, 1995; Scheff and Wadden, 1991, 1993; Chung et al., 1994; Fujita et al., 1995). Recently, more studies have focused on implementing the CMB technique for fine particulate matter. However, source profiles for many of the primary $PM_{2.5}$ sources need to be further developed to yield better results. Positive matrix factorization (PMF) is a technique that has already been used for a variety of source apportionment and spatial analyses (Kim et al., 2003, 2004, 2005a,b; Kim and Hopke, 2004a–c; Rizzo and Scheff, 2004; Hopke et al., 2003; Song et al., 2001; Paatero et al., 2003). It is related to factor analysis where the underlying covariability of many variables is analyzed so that the original data can be described by a smaller set of factors to which the original variables are related. However, PMF differs from traditional factor analysis in the weighting placed on the certainty of individual species. PMF is advantageous in that one does not require profiles to determine the possible source contributions as with the CMB model. However, when using PMF, it can be difficult to identify potential sources without some sort of profile to which to compare the final results. Both models assume that none of the fitting species used in the analysis are reactive or reacts significantly in the atmosphere between the point of emission and the receptor location. This work will examine the relationship between the PMF and CMB receptor models to determine their similarities and differences using data from two sites within Chicago, Illinois metropolitan area.

2. Methodology

2.1. PMF analysis

PMF is a factor analytic tool that is able to provide non-negative solutions for a variety of uses. PMF iteratively solves the following equation:

$$X = GF + E, \tag{1}$$

where X is a matrix of observed fine particulate species concentrations with the dimensions of number of observations by the number of species, G a matrix of source contributions by observation day whose sum is normalized to the total number of observations in the analysis with the dimensions of number of observations by the number of factors, F a matrix of source profiles normalized to the total fine particulate with the dimensions of number of factors by the number of species, and E a matrix of random errors with the dimensions of number of observations by number of species.

Multivariate factor analytic techniques have been shown to be sensitive to variables with a high proportion of data less than the minimum detectable limit (MDL). Thus, the uncertainty for each measured concentration in the analysis was based on the frequency of an individual species' measurements that were greater than the method detection limit. It has been shown that species which are consistently below the detection limit and constitute mostly noise greatly influence the final result of a PMF analysis (Paatero and Hopke, 2003). For the purpose of this work, the following equation was used to assess the signal to noise ratio of specific species to be included or excluded from the model (Paatero and Hopke, 2003).

$$S/N = \frac{\sum_{\{i|x_{ij} > \delta_j\}} x_{ij}}{\delta_j m_{DL_j}}, \tag{2}$$

where x_{ij} is the value of a specific variable j collected at time i which is greater than the method detection limit, δ_j the method detection limit for variable j, and m_{DL_j} the number of values less than the method detection limit.

If the value from Eq. (2) was greater than 2, then the variable was considered "good". If the value was between 0.2 and 2, the variable was considered "weak" and "bad" if it was less than 0.2. These categories were used to develop the uncertainties associated with each value used in the analysis (Paatero and Hopke, 2003). Uncertainties for "good" variables were either the MDL or the root