Atmospheric Environment 44 (2010) 2731–2741

Contents lists available at ScienceDirect

Atmospheric Environment

journal homepage: www.elsevier.com/locate/atmosenv

Source apportionment using positive matrix factorization on daily measurements of inorganic and organic speciated $PM_{2.5}$

Steven J. Dutton [a], Sverre Vedal [b], Ricardo Piedrahita [c], Jana B. Milford [c], Shelly L. Miller [c], Michael P. Hannigan [c,*]

[a] National Center for Environmental Assessment, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA
[b] Department of Environmental and Occupational Health Sciences, School of Public Health and Community Medicine, University of Washington, Seattle, WA 98195, USA
[c] Department of Mechanical Engineering, College of Engineering and Applied Science, University of Colorado, Boulder, CO 80309, USA

ARTICLE INFO

Article history:
Received 14 September 2009
Received in revised form
20 April 2010
Accepted 21 April 2010

Keywords:
Particulate matter
$PM_{2.5}$
Organic molecular markers
Source apportionment
Positive matrix factorization
Denver

ABSTRACT

Particulate matter less than 2.5 microns in diameter ($PM_{2.5}$) has been linked with a wide range of adverse health effects. Determination of the sources of $PM_{2.5}$ most responsible for these health effects could lead to improved understanding of the mechanisms of such effects and more targeted regulation. This has provided the impetus for the Denver Aerosol Sources and Health (DASH) study, a multi-year source apportionment and health effects study relying on detailed inorganic and organic $PM_{2.5}$ speciation measurements.

In this study, $PM_{2.5}$ source apportionment is performed by coupling positive matrix factorization (PMF) with daily speciated $PM_{2.5}$ measurements including inorganic ions, elemental carbon (EC) and organic carbon (OC), and organic molecular markers. A qualitative comparison is made between two models, PMF2 and ME2, commonly used for solving the PMF problem. Many previous studies have incorporated chemical mass balance (CMB) for organic molecular marker source apportionment on limited data sets, but the DASH data set is large enough to use multivariate factor analysis techniques such as PMF.

Sensitivity of the PMF2 and ME2 models to the selection of speciated $PM_{2.5}$ components and model input parameters was investigated in depth. A combination of diagnostics was used to select an optimum, 7-factor model using one complete year of daily data with pointwise measurement uncertainties. The factors included 1) a wintertime/methoxyphenol factor, 2) an EC/sterane factor, 3) a nitrate/polycyclic aromatic hydrocarbon (PAH) factor, 4) a summertime/selective aliphatic factor, 5) an n-alkane factor, 6) a middle-oxygenated PAH/alkanoic acid factor and 7) an inorganic ion factor. These seven factors were qualitatively linked with known $PM_{2.5}$ emission sources with varying degrees of confidence. Mass apportionment using the 7-factor model revealed the contribution of each factor to the mass of OC, EC, nitrate and sulfate. On an annual basis, the majority of OC and EC mass was associated with the summertime/selective aliphatic factor and the EC/sterane factor, respectively, while nitrate and sulfate mass were both dominated by the inorganic ion factor. This apportionment was found to vary substantially by season. Several of the factors identified in this study agree well with similar assessments conducted in St. Louis, MO and Pittsburgh, PA using PMF and organic molecular markers.

1. Introduction

Recent studies have linked particulate matter (PM) with a broad array of adverse health effects (HEI, 2002; Lippmann et al., 2003). The magnitude of these effects, however, can vary greatly by geographic location (Dominici et al., 2005). One possible explanation for the observed heterogeneity of effects is regional variation in air pollution sources. Distinguishing the impact of individual PM sources on health could lead to improved understanding of the mechanisms of such effects as well as more targeted regulation. Identification of sources contributing to fine particulate matter ($PM_{2.5}$) is a primary objective of the Denver Aerosol Sources and Health (DASH) study, a multi-year source apportionment and health effects study conducted in Denver, CO (Vedal et al., 2009).

Receptor models have been widely used for source apportionment of speciated $PM_{2.5}$ (Thurston et al., 2005; Watson et al., 2008). The most common receptor models can be categorized into univariate models such as chemical mass balance (CMB) and

* Corresponding author.
E-mail address: hannigan@colorado.edu (M.P. Hannigan).

1352-2310/$ – see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2010.04.038

multivariate models such as principal components analysis (PCA), positive matrix factorization (PMF) and UNMIX. Each model has its advantages and disadvantages, but all rely on the basic principles of mass balance. In comparison to the other models, PMF is well suited for longer time series studies and incorporates non-negativity constraints on the output, frequently resulting in a more physically interpretable result. PMF also incorporates individual measurement uncertainties into the model framework. These features make PMF a good choice for source apportionment of the multi-year DASH data set containing inorganic and organic speciated $PM_{2.5}$ concentrations with pointwise measurement uncertainty described previously (Dutton et al., 2009a, 2009b).

Two programs were used to solve the PMF problem for this study, PMF2 (Paatero and Tapper, 1994) and ME2 (Paatero, 1999). In addition to requiring a long series of input data, PMF2 and ME2 also require a priori determination of the number of factors to be output by the model. While this requirement reduces the objectivity of the results, numerous diagnostic tools are available to help determine the most appropriate number of factors (Lee et al., 1999). In addition, it is common practice to run the models over a range of factor numbers to provide insight into the optimum solution (e.g., Jaeckels et al., 2007; Shrivastava et al., 2007).

Ultimately, there will be more than four years of daily data available from the DASH study. This preliminary analysis used 1.5 years of data, and subsets of the 1.5-year data set were tested to determine the sensitivity of the source apportionment results to data selection. Likewise, PMF2 and ME2 operational parameters were varied to explore the sensitivity of the results to their values. An optimum data set with optimum model parameters was selected and used for inter-model comparison and to perform a bulk species mass apportionment.

2. Methods

2.1. Ambient data collection and chemical characterization

Daily filter samples were collected from a receptor site in Denver located on a two-story elementary school building. This site is in a residential neighborhood, far from major emission point sources and heavily traveled roadways (Vedal et al., 2009). Details of the sampling equipment, protocols and analyses are discussed in Dutton et al. (2009a,b). Teflon filters were used for mass determination and analysis of inorganic ions including nitrate and sulfate. Quartz filters were used for organic analysis of bulk elemental carbon (EC), bulk organic carbon (OC) and for individual quantification of 72 non-polar and moderately polar trace organic molecular marker compounds by gas chromatography–mass spectrometry (GC–MS). Table 1 lists the species quantified during the first 1.5 years of the DASH study along with their mean concentration, mean uncertainty, and prevalence of missing and below detection limit observations. Additional figures and statistics pertaining to the bulk species and organic molecular marker species are presented in Dutton et al. (2009a) and Dutton et al. (2009b), respectively. Temporal patterns and correlations between the individual species are presented in Dutton et al. (2010).

2.2. Positive matrix factorization framework

The PMF multivariate factor analysis algorithm PMF2 (Paatero, 1998a,b) was used as the primary source apportionment tool for this study. PMF2 attempts to apportion a series of observations, x_{ij}, into p distinct factors according to the model:

Table 1
$PM_{2.5}$ species quantified during the first 1.5 years of the DASH study (July 1, 2002–December 31, 2003).

Chemical species	Mean	Mean uncertainty	Percent missing	Percent BDL[a]
Bulk species (μg/m^3)				
$PM_{2.5}$ mass[b]	8.50	1.43	1%	5%
Nitrate	0.96	0.34	1%	58%
Sulfate	1.15	0.08	1%	3%
Ammonium (2003 only)	0.55	0.03	0%	0%
EC	0.65	0.10	4%	1%
OC	3.13	0.32	4%	0%
Alkanes and cycloalkanes (ng/m^3)				
Docosane	1.45	0.11	13%	0%
Tricosane	2.16	0.21	13%	0%
Tetracosane	1.13	0.16	13%	1%
Pentacosane	1.65	0.19	13%	1%
Hexacosane	0.92	0.13	13%	8%
Heptacosane	1.29	0.15	13%	5%
Octacosane	0.86	0.14	13%	17%
Nonacosane	1.91	0.20	13%	3%
Triacontane	0.70	0.12	13%	19%
Hentriacontane	1.75	0.25	13%	1%
Dotriacontane	0.47	0.17	13%	49%
Tritriacontane	0.79	0.09	13%	3%
Tetratriacontane	0.58	0.07	13%	9%
Pentatriacontane	0.45	0.06	13%	8%
Hexatriacontane	0.26	0.05	13%	22%
Heptatriacontane	0.19	0.04	13%	35%
Octatriacontane	0.15	0.04	13%	49%
Nonatriacontane[b]	0.15	0.04	13%	51%
Tetracontane[b]	0.10	0.04	20%	59%
Pentadecylcyclohexane	0.16	0.02	4%	8%
Nonadecylcyclohexane	0.15	0.02	3%	7%
PAHs (ng/m^3)				
Fluoranthene	0.19	0.03	2%	4%
Pyrene	0.16	0.02	2%	3%
Benzo[ghi]fluoranthene	0.10	0.01	2%	0%
Cyclopenta[cd]pyrene	0.04	0.00	2%	5%
Benz[a]anthracene	0.06	0.01	2%	3%
Chrysene/triphenylene	0.17	0.02	2%	0%
Benzo[b&k]fluoranthene	0.22	0.02	2%	0%
Benzo[j]fluoranthene[b]	0.01	0.00	2%	60%
Benz[a&e]pyrene	0.18	0.02	2%	4%
Perylene[b]	0.01	0.00	2%	60%
Indeno[1,2,3-cd]pyrene	0.08	0.01	2%	6%
Benzo[ghi]perylene	0.21	0.02	2%	4%
Dibenz[ah]anthracene	0.02	0.00	2%	44%
Picene	0.01	0.00	2%	44%
Coronene	0.10	0.01	2%	8%
Methyl-202-PAH sum	0.62	0.08	2%	2%
Retene	0.42	0.06	2%	23%
Methyl-228-PAH sum	0.11	0.01	3%	1%
Oxy-PAHs (ng/m^3)				
Acenaphthenone[b]	0.10	0.02	4%	50%
Fluorenone	0.36	0.06	4%	26%
1H-phenalen-1-one	0.25	0.03	4%	5%
Xanthone	0.15	0.02	16%	10%
1,8-Naphthalic anhydride	0.31	0.04	16%	0%
Anthracene-9,10-dione	0.32	0.04	16%	1%
Benz[de]anthracene-7-one	0.08	0.01	2%	2%
Steranes (ng/m^3)				
20R-abb & 20S-aaa-Cholestane	0.14	0.01	3%	2%
20R & S-abb-Methylcholestane	0.10	0.01	3%	4%
20R & S-abb-Ethylcholestane	0.11	0.01	3%	3%
a-22,29,30-Trisnorhopane	0.09	0.01	3%	3%
ba-30-Norhopane	0.35	0.03	3%	0%
ab-Hopane	0.22	0.03	3%	2%
22S-ab-30-Homohopane	0.10	0.01	3%	3%
22R-ab-30-Homohopane	0.07	0.01	3%	4%
22S-ab-30-Bishomohopane	0.06	0.01	3%	5%
22R-ab-30-Bishomohopane	0.05	0.01	3%	7%
Fatty acids (ng/m^3)				
Dodecanoic acid	3.08	0.44	5%	5%
Tridecanoic acid[b]	0.22	0.15	5%	79%

Table 1 (*continued*)

Chemical species	Mean	Mean uncertainty	Percent missing	Percent BDL[a]
Tetradecanoic acid	4.58	0.66	4%	4%
Pentadecanoic acid	0.99	0.18	4%	11%
Hexadecanoic acid	18.08	3.28	4%	6%
Heptadecanoic acid	0.62	0.27	4%	38%
Octadecanoic acid	11.28	3.37	4%	16%
Oleic acid[b]	2.04	0.40	4%	69%
Sterols and methoxyphenols (ng/m^3)				
Cholesterol[b]	0.24	0.21	3%	73%
Stigmasterol	0.35	0.07	3%	29%
Vanillin	2.44	0.38	3%	21%
Acetovanillone	0.58	0.08	3%	34%
Coniferaldehyde	1.21	0.17	2%	37%
Syringaldehyde	1.20	0.16	5%	47%
Acetosyringone[b]	0.29	0.05	2%	56%

[a] Percent of observations below detection limit (BDL).
[b] Not used for source apportionment.

$$x_{ij} = \sum_{k=1}^{p} g_{ik} f_{kj} + \varepsilon_{ij}, \quad i = 1,\ldots,m \; j = 1,\ldots,n \tag{1}$$

In Equation (1), g_{ik} represents the contribution (mass/volume) of factor k to sample i, f_{kj} represents the relative fraction (mass/mass) of species j in factor k, and ε_{ij} is the residual concentration (mass/volume) not explained by the model. In our application, the observations, x_{ij} ($i = 1,\ldots,m$, $j = 1,\ldots,n$), represent m daily concentration measurements of n speciated components of $PM_{2.5}$.

ME2 (Paatero, 1999) was used to further explore the data as it allows greater freedom in specifying the model to be fit. ME2 uses a conjugant gradient algorithm rather than PMF2's Gauss–Newton/Newton–Raphson algorithm to iteratively perform the factorizations. Three ME2 models were explored in this analysis: 1) a two-way model, also specified by Equation (1) above, 2) an enhanced model incorporating two temporal indices, one for season and one for weekday/weekend, and 3) an enhanced model incorporating hourly wind speed and direction data in addition to the season and weekday/weekend information. The enhanced models use the bilinear Equation (1) augmented by a second multilinear equation incorporating the extra explanatory variables (Kim et al., 2003).

In the case of the enhanced model with two temporal indices, the augmented equation can be written as:

$$x_{ij} = \sum_{k=1}^{p} S(\eta_i, k) W(\omega_i, k) f_{kj} + \varepsilon'_{ij}, \quad i = 1,\ldots,m \; j = 1,\ldots,n \tag{2}$$

where S represents the seasonal dependence for the factor contributions and η_i identifies the seasonal classification for day i. Six 2-month long seasonal classifications were used, starting with January/February. Similarly, W represents the weekend/weekday dependence for the factor contributions and ω_i classifies each day i as a weekday or weekend. For this study, major holidays were classified as weekends regardless of the day they fell on since source emissions on these days were expected to more closely resemble weekends. In Equation (2), the indices for the auxiliary factors are shown in parentheses for clarity and ε'_{ij} is the residual concentration (mass/volume) not explained by the enhanced model.

The enhanced model with meteorological data contains the same seasonal and weekday/weekend parameters but is further supplemented with a product of parameters representing the influence of the hourly ($h = 1,\ldots,24$) meteorological data, including wind direction (D), wind speed (V), and calm wind (R):

$$x_{ij} = \sum_{k=1}^{p} S(\eta_i, k) W(\omega_i, k) \sum_{h=1}^{24} D(\delta_{ih}, k) V(\nu_{ih}, k) R(c_{ih}, k) f_{kj} + \varepsilon'_{ij} \quad i = 1,\ldots,m \; j = 1,\ldots,n \tag{3}$$

Eighteen wind directions (δ_{ih}), five wind speed ranges (ν_{ih}) and a binary indicator for calm wind conditions (c_{ih}) were used. More specifics on the enhanced models can be found in the Supplemental information and further explanation of these models is given by Kim et al. (2003).

PMF2 and the ME2 two-way models are solved by minimizing Q, the sum of the squared, scaled residuals, defined by:

$$Q = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\frac{\varepsilon_{ij}}{s_{ij}} \right)^2 \quad i = 1,\ldots,m \; j = 1,\ldots,n \tag{4}$$

where the scaling value s_{ij} is the pointwise observation uncertainty. Dutton et al. (2009a) provide details of the pointwise uncertainty specification, and average uncertainty values are given in Table 1. The enhanced ME2 models are solved by minimizing a modified Q, with the second term coming from the augmented equation, and using increased uncertainty scaling values (s'_{ij}) to place primary emphasis on the basic model.

$$Q = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\frac{\varepsilon_{ij}}{s_{ij}} \right)^2 + \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\frac{\varepsilon'_{ij}}{s'_{ij}} \right)^2 \quad i = 1,\ldots,m \; j = 1,\ldots,n \tag{5}$$

In this study the pointwise uncertainty for each observation was increased by a factor of eight (Paatero et al., 2002) for the augmented equation. More discussions about the model solutions have been presented previously (Paatero et al., 2005; Paatero, 1997) and details of specifics for the work presented here are given in the Supplemental information.

2.3. Handling of missing values and values below detection limit

Factor analysis does not require a continuous data set in time and therefore days with many missing species observations or days with suspect speciation data were omitted from the source apportionment analysis. These included three missing days in 2002, four missing days in 2003, one week in 2002 (August 26–September 1) without organics measurements and eight weeks in 2002/2003 (December 2–January 26) with high alkane contamination on the filters. These eight weeks were part of an early pilot analysis before refinements in filter handling and extraction techniques were employed.

In contrast, the data on days with a small number of missing species observations were handled in a manner similar to that proposed by Polissar et al. (1998). Missing species concentrations were replaced by the geometric mean of the remaining concentrations of the same species and assigned an uncertainty equal to four times the geometric mean. In addition, individual species measured below detection limits (bdl) were replaced by 1/2 the detection limit (dl) and assigned an uncertainty equal to 5/6 the dl as proposed by Polissar et al. (1998). The dl for each observation was calculated as discussed in Dutton et al. (2009a,b). While replacing bdl and negative-valued measurements comports with physical interpretation of the data set as representing positive pollutant concentrations and with the inability to statistically distinguish these low values from one another, it should be noted that PMF itself does not require replacement of bdl values. The

model can factor data sets that include negative-valued measurements.

2.4. Selection of input data and model operating parameters

The robustness of the source apportionment techniques was tested by varying the input data and model operating parameters and observing the effect on the model output. Input data were varied by altering the number of included days and by adjusting the list of included species. PMF2 operating conditions were varied by modifying the treatment of outliers, exploring the rotational freedom of the results, and varying the number of factors included in the model. The results from this sensitivity analysis are summarized below and additional details were presented in Dutton (2008). Once the final data set and number of factors were set for PMF2, the same configuration was used with the ME2 models.

3. Results

3.1. Data set sensitivity analysis

A higher frequency of missing data and elevated uncertainties during the first six months of the study in 2002 prompted us to limit the source apportionment to the 2003 data where filter collection and analytical procedures were more refined. Subsequent inclusion of the 2002 data in the source apportionment produced similar results, but they were more sensitive to the specified outlier distance. This is likely a result of the higher frequency of imputed values present in the 2002 data. In contrast, running PMF2 on the 2003 data alone produced results that were virtually unaffected by outlier distance and were stable even without using robust mode.

With the exception of nitrate, species containing more than 50% of their observations bdl had little effect on the apportionment. Ten organic molecular marker compounds had more than half their observations bdl (see Table 1) and were thus excluded from the data set. Nitrate was bdl 58% of the time, but it was retained in the data set since it is an important marker for secondary inorganic aerosol during winter months. A 1-year data set created from all available 2003 data including sulfate, nitrate, ammonium, EC, OC and 62 non-polar organic compounds (those shown in Table 1 with fewer than 50% of observations bdl) was used for the source apportionment work reported here. Further details of the data set sensitivity analysis are presented by Dutton (2008).

3.2. PMF2 results

The behavior of Q and the maximum value in T ($\max(T)$) as a function of φ and p have been used to make inferences about the rotational stability of the PMF2 results and the optimum number of factors supported by the model (Lee et al., 1999; Reff et al., 2007). For this study, we varied φ and p to determine their influence on the apportionment results. The findings from this sensitivity analysis are summarized below with additional details presented in the Supplemental information.

Since a lower value for Q implies a better fitting model, the behavior of Q as a function of φ can provide insight into the rotational stability of the model (Paatero et al., 2002). Our analyses of a 7-factor PMF2 model showed a minimum in Q at $\varphi = 0.1$ with indistinguishable effects on the resulting factor profiles and contributions with φ varying from -0.2 to 0.3. Results were similar for solutions to the model with differing number of factors. Therefore, we selected $\varphi = 0$, the most central solution.

A stepwise approach was used to select the number of factors for PMF2, beginning with a 2-factor model and successively adding factors up to a maximum of twelve. As each factor was added, changes in Q and $\max(T)$ and the scaled residuals were examined to attempt to identify the optimum number of factors. Additionally, the factor profiles and contributions were examined to see if physical interpretability was enhanced. Our analysis revealed large values in T for an eleven-factor solution, suggesting a non-unique solution which may be improved by selecting fewer factors (Paatero, 1998a). However, below eleven, the values of Q and $\max(T)$ alone revealed no clear distinction in the appropriate number of factors. Other investigators using PMF2 have similarly not found a definitive choice in p based on these overly-simplified diagnostic parameters (Brinkman et al., 2006; Ulbrich et al., 2008). The scaled residuals gradually improved as the number of factors increased, but there was no one point where they got significantly smaller (with the exception of the residuals for nitrate, sulfate, and ammonium, which were large up until the point when more than three factors were used and a secondary inorganic ion factor was produced by the model).

Interpretability of the resulting factor profiles and contributions was ultimately used as the primary tool for selecting the optimum number of factors. Beyond a 7-factor model, the factors produced by PMF2 began splitting into similar, highly correlated factors that did not match any known source profiles. Therefore, a 7-factor solution was chosen as most interpretable. A similar approach to finding the optimum number of factors was used by Jaeckels et al. (2007) and Shrivastava et al. (2007).

The PMF2 factor profiles and contributions for the 7-factor model are presented in Figs. 1 and 2, respectively. The factor profiles have been normalized so each species sums vertically to 100%. $PM_{2.5}$ mass apportionment was not attempted in this work since our $PM_{2.5}$ mass measurements exhibited a large degree of uncertainty (Dutton et al., 2009a). Furthermore, we could not rely on mass closure in the model since metals were not included in the analysis and since conversion from OC to organic carbon mass (OCM) is not well constrained (Pang et al., 2006). As a result, the factor contributions are not scaled to $PM_{2.5}$ mass and are displayed in arbitrary units in the figures. Contributions from each factor can be compared from day to day, but the magnitude of the contributions should not be compared across factors. The factors have been labeled based on the prominent compound class or classes present in the factor profiles and/or the seasonal characteristics of the factor contributions. Fig. 3 contains box plots for each of the factors stratified by both day of the week and season to illustrate the temporal patterns present in the factor contributions. For example, the EC/sterane factor in Fig. 3b shows a decrease in contribution on the weekends with higher values in the wintertime. Interpretations of these factors are provided in the Discussion section.

Although $PM_{2.5}$ mass apportionment was not performed in this study, bulk species mass apportionment was performed and Fig. 4 illustrates the influence of each factor on total nitrate, sulfate, EC and OC mass. For each species, the apportionment is provided a) on an annual basis, b) limited to days having the highest 10% of bulk species mass concentration, and c) for each individual season. The average observed ambient data stratified in the same manner are represented by triangles in Fig. 4 for comparison with the model output.

3.3. ME2 results

ME2 results are presented in detail in the Supplemental Information. The factor profiles and contributions from all ME2 models were very similar to those of PMF2 for model runs with up to nine factors. For the nine-factor run, the enhanced ME2 models showed a different ninth factor than the PMF2 and ME2 two-way models. For the 7-factor solution in particular, the ME2 results closely

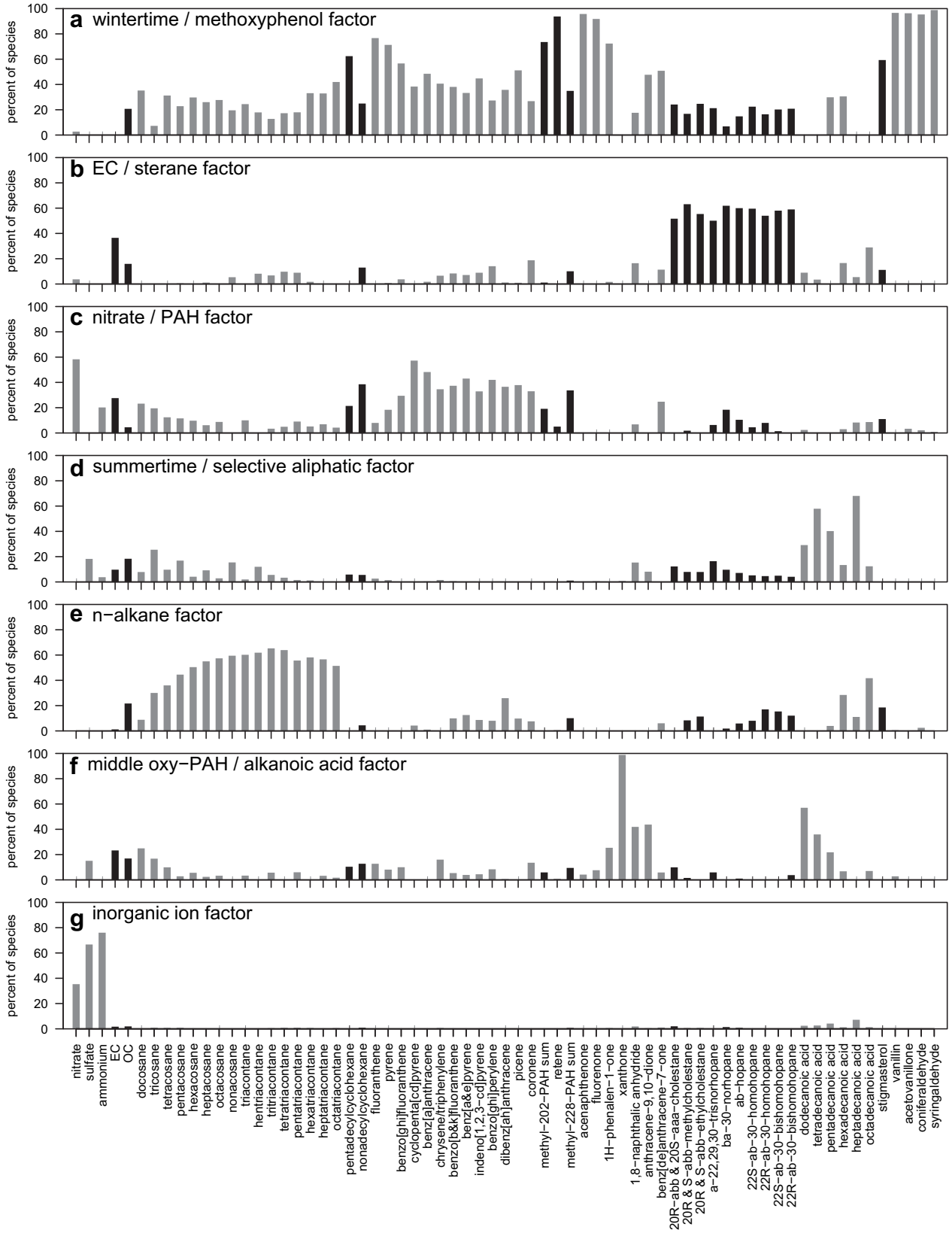

Fig. 1. PMF2 factor profiles for a 7-factor solution. The bars are shaded to show compound class separations. The *y*-axis represents the percent of each species present in each factor (i.e., each compound sums vertically to 100%).

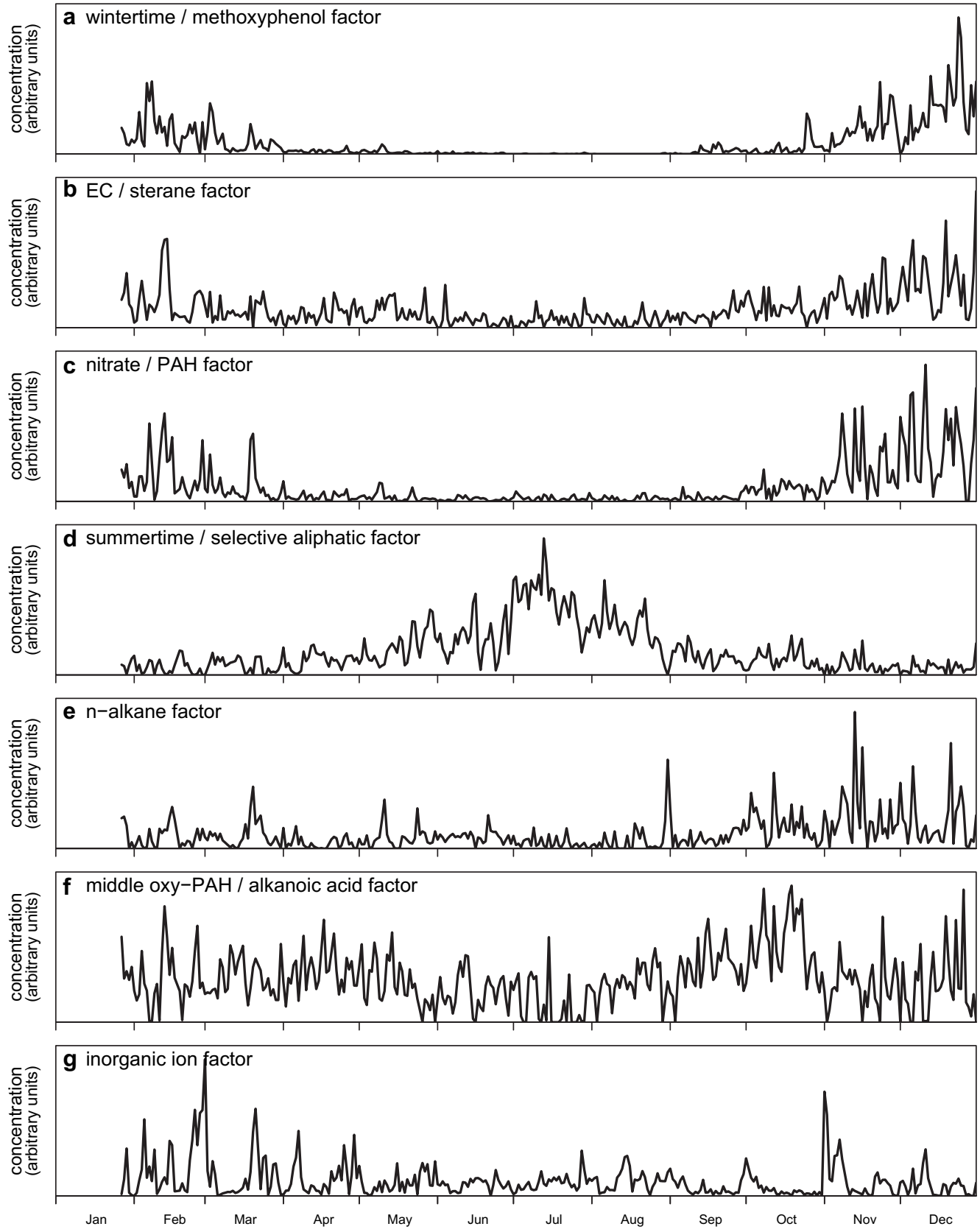

Fig. 2. PMF2 factor contributions for a 7-factor solution. The *y*-axis is in arbitrary units as described in the text and therefore the magnitude of the contributions should not be compared between factors.

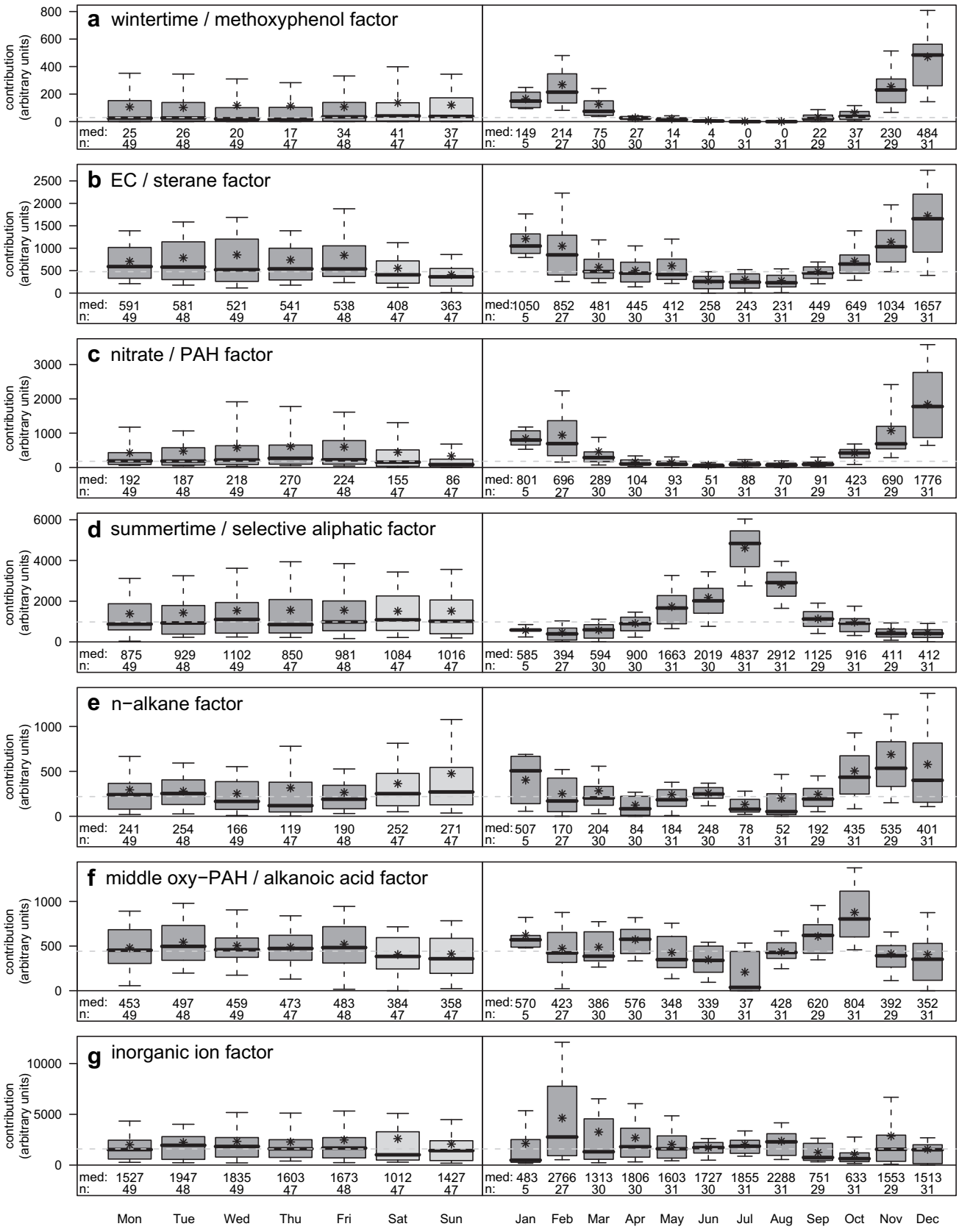

Fig. 3. Day of the week and seasonal box plots for the PMF2 factor contributions. The boxes depict the median (dark line), inner quartile range (shaded box), 10th and 90th percentiles (whiskers) and the mean (asterisk). The dashed line across the plot is the overall median. The individual median values and the number of points contained within each box are listed below the boxes.

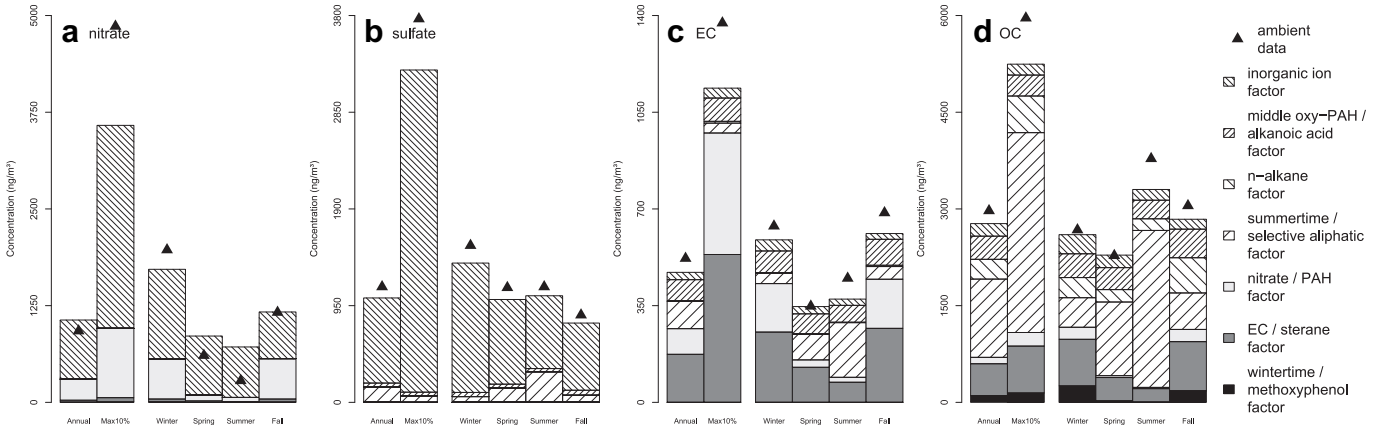

Fig. 4. PMF2 bulk species apportionment showing the distribution of the factors contributing to the total mass of a) nitrate, b) sulfate, c) EC and d) OC. Apportionment results are shown on an annual basis; limited to the top 10% highest bulk species concentration days (Max 10%); and for the four individual seasons (winter = December 22–March 21, spring = March 22–June 20, summer = June 21–September 20, fall = September 21–December 21). The overall height of each bar represents the average concentration of each bulk species apportioned in the model, which can be compared to the average of the measured concentrations represented by the triangles.

matched those from PMF2 with correlations between the factor contributions output by each model greater than 0.98. Table 2 includes Pearson correlation coefficients (r) and normalized gross errors (NGE) of the factor contributions output by the different models for a 7-factor solution. The NGE was calculated using the average factor contribution from PMF2 for normalization to avoid having zeros in the denominator:

$$\mathrm{NGE}_{p,\mathrm{ave}} = \frac{1}{m}\sum_{i=1}^{m}\left|\frac{g_{p,i}^{\mathrm{PMF2}} - g_{p,i}^{\mathrm{ME2}}}{g_{p,\mathrm{ave}}^{\mathrm{PMF2}}}\right| \quad (6)$$

The weekend and seasonal factors from the ME2 enhanced models mirrored the temporal analysis shown in Fig. 3 performed on the PMF2 source contributions. The extra factors from the enhanced model with meteorological data were able to explain some trends, but overall it was difficult to confidently distinguish transport effects from seasonality.

4. Discussion

4.1. Dependence on input data and model operating parameters

PMF2 model robustness was adversely affected by the first six-month period containing a substantial proportion of missing and high uncertainty observations. It is common practice in source apportionment to replace missing values with some statistic derived from the remaining observations (Polissar et al., 1998). However, if there is a high degree of seasonal or weekly temporal structure present in the data, this practice may produce erroneous results. In our case, when we included the first six months of data, PMF2 produced artifact factors easily identified by having contributions only present during periods with imputed or high uncertainty values. This demonstrates the importance of consistent measurement procedures.

Factors used in determining the optimum model specification included consistency in model output, detailed inspection of the residuals, behavior of the diagnostic parameters (Q and max(T)), correlation between the factor contributions, and comparison with eigenvectors obtained using principal components analysis (see Dutton, 2008). The final selection of a 7-factor PMF2 model was ultimately made based on a combination of model diagnostics and factor interpretability. Robust solutions using the two-way and enhanced ME2 models were also achieved for the 7-factor solution with further details provided in the Supplemental Information.

4.2. Interpretation of factor profiles

4.2.1. Wintertime/methoxyphenol factor

This factor contained more than 80% of the retene, light oxy-PAH and methoxyphenol concentrations. It also contained a high percentage of the n-alkanes, PAHs and heavier oxy-PAHs. Retene and the methoxyphenols are markers for wood combustion (Hays et al., 2002; Schauer et al., 2001). The factor contribution shows strong seasonality with a wintertime maximum, consistent with residential wood combustion for home heating in the winter. Investigation of the median contribution by day of the week revealed an increase on Friday, Saturday and Sunday relative to the rest of the week, also consistent with residential fireplace use patterns (Chinkin et al., 2003).

Table 2
Statistics comparing the factor contributions from the PMF2 model with the three ME2 models (two-way, enhanced, and enhanced with wind).

Factor	Pearson correlation coefficient (r)			Normalized gross error (NGE)[a]		
	Two-way model	Enhanced model	Enhanced model with wind	Two-way model	Enhanced model	Enhanced model with wind
Wintertime/methoxyphenol	1.000	0.995	0.997	0.051	0.059	0.083
EC/sterane	1.000	0.993	0.996	0.012	0.069	0.072
Nitrate/PAH	1.000	0.998	0.998	0.002	0.222	0.224
Summertime/selective aliphatic	1.000	0.996	0.996	0.009	0.076	0.077
n-Alkane	1.000	0.999	0.996	0.010	0.063	0.076
Middle-oxy-PAH/alkanoic acid	1.000	0.988	0.988	0.022	0.063	0.067
Inorganic ion	1.000	0.994	0.994	0.005	0.067	0.067

[a] As defined in Equation (6).

The high abundance of other organic molecular marker species present in this factor was likely a result of temporal correlation with meteorology-driven effects. Increased concentrations during the winter across all measured compounds are expected as a direct result of diminished mixing conditions driven by frequent wintertime atmospheric inversions in the region (Neff, 1997). A 9-factor PMF2 solution included further separation of this factor into two wintertime-dominated factors: one which contained all the methoxyphenol wood smoke markers, retene, and 60 percent of the vanillin, while the other contained several of the lighter-weight PAHs and oxy-PAHs and the remainder of the vanillin. This second wintertime factor could not be linked with any known source profiles and was correlated in time with the wood smoke marker dominated factor ($r = 0.51$), so the 7-factor solution containing a combined wintertime and wood smoke source was deemed the most justifiable model. Inspection of Fig. 4 shows that the wintertime/methoxyphenol factor did not contribute substantially to nitrate, sulfate or EC mass. It did, however, have an appreciable contribution to OC mass during the fall and winter (7% and 11%, respectively).

4.2.2. EC/sterane factor

This factor was identified by the presence of EC, all three cholestanes and all seven hopanes. Cholestanes and hopanes are markers for motor vehicle emissions (Schauer et al., 2002b) and are present specifically in heavy petroleum distillates such as motor oil (Simoneit, 1999). Furthermore, Schauer et al. (2002b) show that they are emitted preferentially by older, non-catalyst controlled motor vehicles. Therefore, this factor likely represents motor oil combustion in the crankcases of motor vehicles with possibly larger contributions coming from older vehicles. Fig. 3b shows a dramatic drop in the contribution from this factor on Saturday and Sunday, consistent with expected reductions in vehicle miles traveled on weekends (Chinkin et al., 2003; Dutton et al., 2010; Harley et al., 2005). The seasonality present in this factor is less than that observed for the wintertime/methoxyphenol factor and may reflect general concentration increases resulting from reduced wintertime atmospheric mixing. Steranes have been shown to decay in the presence of OH radical concentrations typical of summer with lifetimes as short as 1–4 days, depending on the reaction pathway with gas-phase OH oxidation of the lighter steranes resulting in the shorter lifetimes (Weitkamp et al., 2008; Lambe et al., 2009). This is not likely causing seasonality of this factor as there is no decrease in this factor contribution associated with elevated ozone concentration, from winter to summer there is no preferential lose of a lighter sterane, norhopane, relative to a heavier sterane, hopane, and finally the seasonality of this factor is similar to the seasonality of EC which is a key component of this factor and unlikely to photochemically decay. This EC/sterane factor was a major contributor year-round to EC and OC mass and was the largest contributor (47%) to EC mass when the data was limited to the highest EC concentration days.

4.2.3. Nitrate/PAH factor

This factor contained 60% of the nitrate and many of the PAHs and was most prominent in the wintertime. Like the EC/sterane factor, it exhibited a dramatic decrease on weekends. Schauer et al. (1996) show that the majority of PAHs in the urban atmosphere are attributed to motor vehicle emissions. Recent work on the dependence of motor vehicle emissions on engine-temperature revealed that three of the heavier PAHs we quantified increase significantly during cold operation of test vehicles (Schauer et al., 2008). Given the strong seasonality and clear weekday preference, this factor was likely dominated by motor vehicle exhaust with a possible link to cold-start and cold-operation emissions. A similar light-duty gasoline vehicle cold-start source was identified for Denver using CMB analysis during the Northern Front Range Air Quality Study (Watson et al., 1998). That study found this source to be responsible for 21% of the OC during winter in Denver, second only to light-duty gasoline high-emitter influences (40% of OC). The presence of nitrate in this factor likely results from the temporal correlation between ammonium nitrate aerosol formation and motor vehicle PAH emission rates during colder temperatures. This nitrate/PAH factor contributed appreciable amounts to nitrate and EC mass, specifically during the fall and winter, and was responsible for 39% of EC mass when the apportionment was restricted to the highest EC concentration days.

4.2.4. Summertime/selective aliphatic factor

This factor was identified by the pattern of n-alkanes and n-alkanoic acids shown in Fig. 1d. The C22–C32 n-alkanes show a clear odd carbon number preference which has been observed in ambient measurements (Brown et al., 2002; Gogou et al., 1996) and emission source profiles for leaf abrasion (Rogge et al., 1993b) and plant wax (Simoneit and Mazurek, 1982). The n-alkanoic acids also revealed a characteristic pattern with peaks at C13 and C17. This factor is most prominent in summer, peaking in July (the warmest month of the study). These observations suggest a summertime biogenic emission source, with photochemistry potentially playing a role in addition. It is difficult to untangle photochemically produced PM mass from increases in a direct emission of PM mass for which the emission rate is a function of ambient temperature as these two sources are likely temporally correlated. In an attempt to explore this factor, the factor contributions were sorted by both ambient temperature and peak 8 h ozone concentration for the day. There was a stronger increase in the factor contributions associated with the highest 10% of ambient temperatures than the increase associated with the highest 10% of ozone concentrations. This simplistic analysis does not mean that photochemistry is unimportant but demonstrates that it is likely more than just photochemistry. The addition of polar species would aid in the separation of secondary organic PM and primary high temperature emissions PM, and is recommended for future studies exploring apportionment of secondary organic PM. The summertime/selective aliphatic factor was dominant in OC mass apportionment, specifically during the summertime when it was responsible for 74% of OC mass. It was responsible for 59% of OC mass when the apportionment was restricted to the highest OC concentration days.

4.2.5. n-Alkane factor

This factor was dominated by n-alkanes and also contained the largest percentage of C16 and C18 alkanoic acids. The weekly pattern in Fig. 3e reveals an increase on the weekends relative to the weekdays. Both C16 and C18 acids are elevated in source profiles for meat cooking (Schauer et al., 1999) and seed oil cooking (Schauer et al., 2002a). This is consistent with a weekend increase in outdoor barbequing (Chinkin et al., 2003) and restaurant patronage (Liu et al., 2001). The preferential presence of even n-alkanoic acids is also consistent with primary biogenic emissions (Hays et al., 2002; Rogge et al., 1993b; Simoneit and Mazurek, 1982). However, the predominance of n-alkanes in this factor suggests it was coming from more than just cooking or primary biogenic emissions. Similar n-alkane patterns enriched in C16 and C18 alkanoic acids are present in road dust originating from tire wear (Rogge et al., 1993a). In contrast to the weekly patterns in Fig. 3e, however, road dust is expected to decrease on the weekends as a result of fewer cars emitting and re-suspending particles on roadways. Therefore, this n-alkane factor was likely originating

from a combination of multiple primary sources. In the bulk species mass apportionment, it only contributed to OC where it was responsible for 11% of the annual OC mass.

4.2.6. Middle-oxy-PAH/alkanoic acid factor

This factor was characterized by xanthone, two other middle-weight oxy-PAHs, and the light n-alkanoic acids. It exhibited contributions year-round with a slight increase in the fall and winter. It also had a slight decrease on the weekends. The oxy-PAHs are not quantified in most source emission studies so identification of the sources responsible for this middle-oxy-PAH/alkanoic acid factor was not possible. However, it was a relatively important year-round contributor to both EC mass (16%) and OC mass (13%). To explore the importance of xanthone in the apportionment, we repeated the analysis with xanthone removed. This exercise resulted in similar factor profiles and contributions. The one large difference was a redistribution in the wintertime of this factor's contribution to the EC/sterane and nitrate/PAH factors.

4.2.7. Inorganic ion factor

This factor was present in the source apportionment results for all data sets containing the inorganic ions. It contains nitrate, sulfate, ammonium and trace amounts of EC, OC and some organic compounds. This factor clearly represents secondary ammonium nitrate and ammonium sulfate, and dominates both nitrate (73%) and sulfate (97%) mass apportionment. The ambient data overlaid in Fig. 4, however, reveal that nitrate is not being fit well by the PMF model, particularly in summer and for the highest nitrate concentration days. This poor fit was also manifest in the nitrate residuals for all model configurations investigated. The extreme seasonality and daily variability in nitrate (driven by meteorology) is not well captured by PMF for the current data set. The result was an under-apportionment of nitrate in the winter and on the highest concentration days and an over-apportionment in the spring and summer.

4.3. Comparison with other studies

Few studies have performed multivariate factor analysis incorporating organic molecular marker data due to the time and effort required to collect a long enough time series of detailed measurements. Jaeckels et al. (2007) and Shrivastava et al. (2007) recently published source apportionment results using PMF and organic molecular marker data collected in St. Louis, MO and Pittsburgh, PA, respectively. The St. Louis study incorporated 105 organic molecular markers, EC, OC, silicone, and aluminum for 120 samples collected every 6th day over a two year period (Bae et al., 2004) and the Pittsburgh study incorporated 54 organic molecular marker species in addition to EC, OC, potassium, calcium, titanium, iron and zinc for 99 samples collected primarily every 6th day over a one year period, with two shorter periods of intensive daily sampling in January and July (Subramanian et al., 2004).

Jaeckels et al. (2007) identified eight factors contributing to the organic aerosol in St. Louis and attributed them to secondary organic aerosol (SOA), wood combustion, a mobile source factor, two point source factors, two winter combustion factors, and resuspended soil. Shrivastava et al. (2007) identified seven primary factors contributing to the organic aerosol in Pittsburgh, linking them with SOA, wood combustion, a mobile source factor, a point source factor, meat cooking, open burning/primary biogenic emissions, and an unknown factor coming from a mixture of primary sources. In St. Louis, Pittsburgh and Denver, three common factors were identified: a SOA/summertime factor, a mobile source factor, and a wintertime/wood smoke factor. Qualitative differences in the sources identified are likely due in part to differences in local emissions and in part to differences in the species quantified in each study.

5. Conclusions

This study demonstrates the ability of factor analysis coupled with organic molecular marker speciation to robustly apportion $PM_{2.5}$. Consistent source apportionment results were obtained for the 1-year data set using PMF2 and three alternative ME2 models. Five of the factors obtained from the optimum 7-factor PMF2 model could be associated primarily with wood combustion, motor oil combustion, motor vehicle emissions, biogenic emissions and secondary inorganic compounds. The remaining two factors were less easily linked with sources: one potentially incorporated contributions from multiple primary emission sources such as food cooking and road dust while the other was associated with an unknown source of oxy-PAHs. The full DASH data set will contain more than four years of daily nitrate, sulfate, EC, OC and organic molecular marker data that will allow for source apportionment stratified by season or temperature. This will allow us to address some of the issues relating to meteorological confounding encountered using the 1-year data set presented here.

Acknowledgments

This research is supported by NIEHS research grant number RO1 ES010197. Additional support for student assistance was provided by NSF Research Experience for Undergraduates award number EEC 0552895. We would like to thank Greg Brinkman, Ingrid Ulbrich and Josh Hemann for their roles in developing the source apportionment tools used in this study.

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use. The views expressed in this article are those of the authors and do not necessarily reflect the views or policies of the U.S. Environmental Protection Agency.

Appendix. Supplementary information

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.atmosenv.2010.04.038.

References

Bae, M.S., Schauer, J.J., DeMinter, J.T., Turner, J.R., 2004. Hourly and daily patterns of particle-phase organic and elemental carbon concentrations in the urban atmosphere. Journal of the Air & Waste Management Association 54, 823–833.

Brinkman, G., Vance, G., Hannigan, M.P., Milford, J.B., 2006. Use of synthetic data to evaluate positive matrix factorization as a source apportionment tool for $PM_{2.5}$ exposure data. Environmental Science & Technology 40, 1892–1901.

Brown, S.G., Herckes, P., Ashbaugh, L., Hannigan, M.P., Kreidenweis, S.M., Collett, J.L., 2002. Characterization of organic aerosol in Big Bend National Park, Texas. Atmospheric Environment 36, 5807–5818.

Chinkin, L.R., Coe, D.L., Funk, T.H., Hafner, H.R., Roberts, P.T., Ryan, P.A., Lawson, D.R., 2003. Weekday versus weekend activity patterns for ozone precursor emissions in California's South Coast Air Basin. Journal of the Air & Waste Management Association 53, 829–843.

Dominici, F., McDermott, A., Daniels, M., Zeger, S.L., Samet, J.M., 2005. Revised analyses of the National Morbidity, Mortality, and Air Pollution Study: mortality among residents of 90 cities. Journal of Toxicology and Environmental Health – Part A – Current Issues 68, 1071–1092.

Dutton, S.J., 2008. Measurement, time series analysis and source apportionment of inorganic and organic speciated $PM_{2.5}$ air pollution in Denver. Ph.D. thesis, University of Colorado.

Dutton, S.J., Schauer, J.J., Vedal, S., Hannigan, M.P., 2009a. $PM_{2.5}$ characterization for time series studies: pointwise uncertainty estimation and bulk speciation methods applied in Denver. Atmospheric Environment 43, 1136–1146.

Dutton, S.J., Williams, D.E., Garcia, J.K., Vedal, S., Hannigan, M.P., 2009b. $PM_{2.5}$ characterization for time series studies: organic molecular marker speciation methods and daily measurements in Denver. Atmospheric Environment 43, 2018–2030.

Dutton, S.J., Vedal, S., Hannigan, M.P., 2010. Long-term and short-term temporal patterns in daily measurements of inorganic and organic speciated $PM_{2.5}$ in Denver. Atmospheric Environment 44, 987–998.

Gogou, A., Stratigakis, N., Kanakidou, M., Stephanou, E.G., 1996. Organic aerosols in Eastern Mediterranean: components source reconciliation by using molecular markers and atmospheric back trajectories. Organic Geochemistry 25, 79–96.

Harley, R.A., Marr, L.C., Lehner, J.K., Giddings, S.N., 2005. Changes in motor vehicle emissions on diurnal to decadal time scales and effects on atmospheric composition. Environmental Science & Technology 39, 5356–5362.

Hays, M.D., Geron, C.D., Linna, K.J., Smith, N.D., Schauer, J.J., 2002. Speciation of gas-phase and fine particle emissions from burning of foliar fuels. Environmental Science & Technology 36, 2281–2295.

HEI, 2002. Understanding the Health Effects of Components of the Particulate Matter Mix: Progress and Next Steps. Health Effects Institute Perspectives, Cambridge, MA, pp. 1–20.

Jaeckels, J.M., Bae, M.S., Schauer, J.J., 2007. Positive matrix factorization (PMF) analysis of molecular marker measurements to quantify the sources of organic aerosols. Environmental Science & Technology 41, 5763–5769.

Kim, E., Hopke, P.K., Paatero, P., Edgerton, E.S., 2003. Incorporation of parametric factors into multilinear receptor model studies of Atlanta aerosol. Atmospheric Environment 37, 5009–5021.

Lambe, A.T., Miracolo, M.A., Hennigan, C.J., Robinson, A.L., Donahue, N.M., 2009. Effective rate constants for the reactions of organic molecular markers (n-alkanes, hopanes and steranes) in motor oil and diesel primary organic aerosols with hydroxyl radicals. Environmental Science & Technology 43, 8794–8800.

Lee, E., Chan, C.K., Paatero, P., 1999. Application of positive matrix factorization in source apportionment of particulate pollutants in Hong Kong. Atmospheric Environment 33, 3201–3212.

Lippmann, M., Frampton, M., Schwartz, J., Dockery, D., Schlesinger, R., Koutrakis, P., Froines, J., Nel, A., Finkelstein, J., Godleski, J., Kaufman, J., Koenig, J., Larson, T., Luchtel, D., Liu, L.J.S., Oberdorster, G., Peters, A., Sarnat, J., Sioutas, C., Suh, H., Sullivan, J., Utell, M., Wichmann, E., Zelikoff, J., 2003. The US Environmental Protection Agency particulate matter health effects research centers program: a midcourse report of status, progress, and plans. Environmental Health Perspectives 111, 1074–1092.

Liu, L.-M., Bhattacharyya, S., Sclove, S.L., Chen, R., Lattyak, W.J., 2001. Data mining on time series: an illustration using fast-food restaurant franchise data. Computational Statistics & Data Analysis 37, 455–476.

Neff, W.D., 1997. The Denver Brown Cloud studies from the perspective of model assessment needs and the role of meteorology. Journal of the Air & Waste Management Association 47, 269–285.

Paatero, P., Hopke, P.K., Begum, B.A., Biswas, S.K., 2005. A graphical diagnostic method for assessing the rotation in factor analytical models of atmospheric pollution. Atmospheric Environment 39, 193–201.

Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. Chemometrics and Intelligent Laboratory Systems 37, 23–35.

Paatero, P., 1998a. User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3, Part 1: Tutorial. University of Helsinki, Helsinki, Finland.

Paatero, P., 1998b. User's Guide for Positive Matrix Factorization Programs PMF2 and PMF3, Part 2: Reference. University of Helsinki, Helsinki, Finland.

Paatero, P., 1999. The multilinear engineer – a table-driven least squares program for solving multilinear problems, including the n-way parallel factor analysis model. Journal of Computational and Graphical Statistics 8 (4), 854–888.

Paatero, P., Hopke, P.K., Song, X.H., Ramadan, Z., 2002. Understanding and controlling rotations in factor analytic models. Chemometrics and Intelligent Laboratory Systems 60, 253–264.

Paatero, P., Tapper, U., 1994. Positive matrix factorization – a nonnegative factor model with optimal utilization of error-estimates of data values. Environmetrics 5, 111–126.

Pang, Y., Turpin, B.J., Gundel, L.A., 2006. On the importance of organic oxygen for understanding organic aerosol particles. Aerosol Science and Technology 40, 128–133.

Polissar, A.V., Hopke, P.K., Paatero, P., 1998. Atmospheric aerosol over Alaska – 2. Elemental composition and sources. Journal of Geophysical Research – Atmospheres 103, 19045–19057.

Reff, A., Eberly, S.I., Bhave, P.V., 2007. Receptor modeling of ambient particulate matter data using positive matrix factorization: review of existing methods. Journal of the Air & Waste Management Association, 146–154.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., Simoneit, B.R.T., 1993a. Sources of fine organic aerosol. 3. Road dust, tire debris, and organometallic brake lining dust – roads as sources and sinks. Environmental Science & Technology 27, 1892–1904.

Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., Simoneit, B.R.T., 1993b. Sources of fine organic aerosol. 4. Particulate abrasion products from leaf surfaces of urban plants. Environmental Science & Technology 27, 2700–2711.

Schauer, J.J., Christensen, C.G., Kittelson, D.B., Johnson, J.P., Watts, W.F., 2008. Impact of ambient temperatures and driving conditions on the chemical composition of particulate matter emissions from non-smoking gasoline-powered motor vehicles. Aerosol Science and Technology 42, 210–223.

Schauer, J.J., Kleeman, M.J., Cass, G.R., Simoneit, B.R.T., 1999. Measurement of emissions from air pollution sources. 1. C-1 through C-29 organic compounds from meat charbroiling. Environmental Science & Technology 33, 1566–1577.

Schauer, J.J., Kleeman, M.J., Cass, G.R., Simoneit, B.R.T., 2001. Measurement of emissions from air pollution sources. 3. C-1–C-29 organic compounds from fireplace combustion of wood. Environmental Science & Technology 35, 1716–1728.

Schauer, J.J., Kleeman, M.J., Cass, G.R., Simoneit, B.R.T., 2002a. Measurement of emissions from air pollution sources. 4. C-1–C-27 organic compounds from cooking with seed oils. Environmental Science & Technology 36, 567–575.

Schauer, J.J., Kleeman, M.J., Cass, G.R., Simoneit, B.R.T., 2002b. Measurement of emissions from air pollution sources. 5. C-1–C-32 organic compounds from gasoline-powered motor vehicles. Environmental Science & Technology 36, 1169–1180.

Schauer, J.J., Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., 1996. Source apportionment of airborne particulate matter using organic compounds as tracers. Atmospheric Environment 30, 3837–3855.

Shrivastava, M.K., Subramanian, R., Rogge, W.F., Robinson, A.L., 2007. Sources of organic aerosol: positive matrix factorization of molecular marker data and comparison of results from different source apportionment models. Atmospheric Environment 41, 9353–9369.

Simoneit, B.R.T., 1999. A review of biomarker compounds as source indicators and tracers for air pollution. Environmental Science and Pollution Research 6, 159–169.

Simoneit, B.R.T., Mazurek, M.A., 1982. Organic-matter of the troposphere. 2. Natural background of biogenic lipid matter in aerosols over the rural western United States. Atmospheric Environment 16, 2139–2159.

Subramanian, R., Khlystov, A.Y., Cabada, J.C., Robinson, A.L., 2004. Positive and negative artifacts in particulate organic carbon measurements with denuded and undenuded sampler configurations. Aerosol Science and Technology 38, 27–48.

Thurston, G.D., Ito, K., Mar, T., Christensen, W.F., Eatough, D.J., Henry, R.C., Kim, E., Laden, F., Lall, R., Larson, T.V., Liu, H., Neas, L., Pinto, J., Stolzel, M., Suh, H., Hopke, P.K., 2005. Workgroup report: workshop on source apportionment of PM health effects – intercomparison of results and implications. Environmental Health Perspectives 113, 1768–1774.

Ulbrich, I.M., Canagaratna, M.R., Zhang, Q., Worsnop, D.R., Jimenez, J.L., 2008. Interpretation of organic components from positive matrix factorization of aerosol mass spectrometric data. Atmospheric Chemistry and Physics Discussions 8, 6729–6791.

Vedal, S., Hannigan, M.P., Dutton, S.J., Miller, S.L., Milford, J.B., Rabinovitch, N., Kim, S.Y., Sheppard, L., 2009. The Denver Aerosol Sources and Health (DASH) study: overview and early findings. Atmospheric Environment 43, 1666–1673.

Watson, J.G., Chen, L.W.A., Chow, J.C., Doraiswamy, P., Lowenthal, D.H., 2008. Source apportionment: findings from the US Supersites Program. Journal of the Air & Waste Management Association 58, 265–288.

Watson, J.G., Fujita, E., Chow, J., Zielinska, B., Richards, L.W., Neff, W., Dietrich, D., 1998. Northern Front Range Air Quality Study Final Report. Prepared for Colorado State University, Fort Collins, CO and Electric Power Research Institute, Palo Alto, CA. DRI Document No. 6580-685-8750-2F2. Desert Research Institute, Reno, NV.

Weitkamp, E.A., Lambe, A.T., Donahue, N.M., Robinson, A.L., 2008. Laboratory measurements of the heterogeneous oxidation of condensed-phase organic molecule markers for motor vehicle exhaust. Environmental Science & Technology 42, 7950–7956.