Source apportionment of PM$_{2.5}$: Comparing PMF and CMB results for four ambient monitoring sites in the southeastern United States

Sangil Lee[a,b,*], Wei Liu[a], Yuhang Wang[a], Armistead G. Russell[b], Eric S. Edgerton[c]

[a]*Georgia Institute of Technology, School of Earth and Atmospheric Sciences, 311 Ferst Drive, Atlanta, GA 30332-0340, USA*
[b]*Georgia Institute of Technology, School of Civil and Environmental Engineering, Atlanta, GA 30332-0512, USA*
[c]*Atmospheric Research & Analysis, Inc., Cary, NC 27513, USA*

Abstract

Two commonly used receptor models, positive matrix factorization (PMF) and chemical mass balance (CMB), are applied to 3-year PM$_{2.5}$ data at two urban sites (Atlanta, GA and Birmingham, AL) and two rural sites (Yorkville, GA and Centreville, AL). Source apportionment results using the two receptor models are analyzed and compared. Both models are able to identify major sources at all sites, though the degree of agreements and correlations between source impacts estimated by PMF and CMB varies depending on sources and receptor sites. Estimated contributions of secondary inorganic particles are the most comparable and highly correlated. The lesser comparability and correlations of estimated contributions of other sources (mostly primary) may be attributed to several factors. Resolved source profiles in PMF have more processed (or aged) characteristics resulting in part from atmospheric mixing and condensation of oxidized compounds, whereas source profiles used in CMB are obtained from measurements of emission sources with minimum amount of atmospheric processing. The PMF profiles vary from site to site; both atmospheric processing and local source variability contribute. In comparison, the CMB profiles obtained from a limited number of emission measurements may not be locally representative even if they are regionally representative. The omission of possible known or unknown sources due to lack of proper source profiles or proper "marker" species may also cause the differences in the source apportionment results. In addition, the implication for PM time-series health study is discussed based on the results from this study.
© 2008 Elsevier Ltd. All rights reserved.

Keywords: PMF; CMB; Source apportionment; PM$_{2.5}$; Time-series health study

1. Introduction

Particulate matter has been linked with negative cardiovascular and respiratory health outcomes, including effects leading to premature mortality (Metzger et al., 2004; Peel et al., 2004; Peters et al., 2001; Pope et al., 2002). In addition, PM$_{2.5}$ contributes to visibility impairment and acid deposition (US-EPA,

*Corresponding author at: Georgia Institute of Technology, School of Earth and Atmospheric Sciences, 311 Ferst Drive Atlanta, GA 30332-0340, USA. Tel.: +1 404 385 4415; fax: +1 404 894 5638.

E-mail address: sangil.lee@eas.gatech.edu (S. Lee).

2002; Watson, 2002). As of December 2006, it has been reported that the 98th percentile of 24 h $PM_{2.5}$ concentrations in 12 counties in the southeastern United States (US) are exceeding the 24 h $PM_{2.5}$ National Ambient Air Quality Standards (NAAQS) of $35 \mu g\, m^{-3}$, suggesting that more than 1.5 million people in the southeast are exposed to elevated levels of $PM_{2.5}$ (www.epa.gov/air/data/). It is important to understand which emission sources contribute to the elevated daily $PM_{2.5}$ levels for developing effective control strategies. This places increased emphasis on correct daily source apportionments to such high levels vary by location and time, and epidemiologic studies suggest that such sources have different health outcomes (Laden et al., 2000; Mar et al., 2000; Sarnat et al., 2006).

Receptor models, which attribute observed concentrations to sources through statistical and/or meteorological interpretation of data, often yield useful insights on the sources of aerosols. Receptor models are based on mass conservation (or balance) of species. The general mass balance form is (Hopke, 2003)

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

where x_{ij} is the ambient concentration of species j in sample i, f_{kj} is the mass fraction (or factor loading) of species j in source k, g_{ik} is the source contribution (or factor scores) of source k in sample i, and e_{ij} is error.

Receptor modeling can be categorized into two different types based on whether $PM_{2.5}$ chemical characteristics from emission sources are required to be known prior to source apportionment (Hopke, 2003). Chemical mass balance (CMB) requires a priori knowledge of major sources (N) and their emission characteristics (f) in the study area, while others (i.e., factor analysis and positive matrix factorization (PMF)) require only ambient measurement data to perform source apportionment. Both CMB (i.e., known sources) and PMF (i.e., unknown sources) have been widely used for understanding of source impacts on ambient $PM_{2.5}$ levels. Both CMB and PMF use a least-squares weighted (by input data uncertainties) fitting through minimizing the differences between measured and estimated concentrations. With known source profiles, CMB seeks to minimize

the χ^2 value,

$$\chi^2 = \sum_{j=1}^{n} \left[\frac{(x_j - \sum_{k=1}^{N} f_{jk} g_k)^2}{\sigma_{xj}^2 + \sum_{k=1}^{N} \sigma_{f_{jk}}^2 g_k^2} \right], \tag{2}$$

where $\sigma_{xj}(\mu g\, m^{-3})$ is the uncertainty on the ambient concentration of species j, and $\sigma_{f_{jk}}$ is the uncertainty in the fraction of species j in the source k profile (Watson et al., 1984). With unknown source profiles, PMF minimizes the Q value,

$$Q = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(\frac{x_{ij} - \sum_{k=1}^{N} g_{ik} f_{jk}}{\sigma_{ij}} \right)^2, \tag{3}$$

where x_{ij} is the ambient concentration of species j in sample i, f_{kj} is the factor loading of species j in source k, g_{ik} is the factor score of source k in sample i, and σ_{ij} is the uncertainty of ambient concentration of species j in sample i (Paatero and Tapper, 1993, 1994). PMF requires a substantial number of ambient samples to resolve source factors (i.e., matrix g) and their profiles (i.e., matrix f), whereas CMB can be applied to any number of samples since the source profiles are already known.

The Southeastern Aerosol Research and Characterization (SEARCH) study provides a consistent, rich, and long-term dataset of ambient $PM_{2.5}$ chemical characteristics in the southeastern United States (Edgerton et al., 2005; Hansen et al., 2003). In order to identify major $PM_{2.5}$ sources and quantify their impacts on ambient $PM_{2.5}$ levels in both urban and rural areas of the Southeast, two different source apportionment methods (i.e., CMB and PMF) were applied to four different ambient data sets, two urban (Atlanta, GA and Birmingham, AL) and two rural (Yorkville, GA and Centreville, AL) sites from the SEARCH. Using this 3-year dataset, we compare the source apportionment results of CMB and PMF. Selection of source profiles in CMB is not always straightforward because the regional representativeness of specific emission studies is often unclear. Sensitivity studies using CMB are therefore conducted. We investigate the usefulness and limitations of these two commonly used source apportionment methods and examine how the uncertainties of these methods affect further studies that utilize source apportionment results (e.g., PM epidemiologic studies or policy analysis). In this study, we examine whether identified sources and their contribution estimates, by the two methods, are similar and how well they

are correlated, which is important for time-series PM health studies.

2. Method

2.1. Measurement data

PM$_{2.5}$ composition data analyzed in this study consist of the measurements taken at two urban–rural pair sites in Alabama (North Birmingham [BHM] and Centreville [CTR]) and Georgia (Atlanta [JST] and Yorkville [YRK]) from the SEARCH study. Twenty-four hour integrated PM$_{2.5}$ samples were collected daily at the JST site. PM$_{2.5}$ samples were collected every third day at the other sites. Samples were collected using particulate composition monitors (PCM, Atmospheric Research and Analysis, Inc., Durham, NC) that have three sampling lines (air flow rate $16.7 \, \text{l min}^{-1}$) with inlets 5 m above ground. More detailed descriptions can be found elsewhere (Edgerton et al., 2005; Hansen et al., 2003). A total of 932 samples for the JST site, 336 samples for the BHM site, 347 samples for the YRK site and 338 samples for the CTR site were obtained and analyzed, covering the time period from January 2000 to December 2002. For each sample, concentrations of the following 19 chemical species were usually available: SO$_4^{2-}$, NO$_3^-$, NH$_4^+$, EC, OC (OC was calculated as OC1 + OC2 + OC3 + OC4 + OP and EC as EC1 + EC2 + EC3-OP), As, Ba, Br, Cu, Mn, Pb, Se, Ti, Zn, Al, Si, K, Ca, and Fe, although there are occasional "missing data" (no reported measurements) for one or more species. Total PM$_{2.5}$ mass concentrations for each day, analytical uncertainty and detection limit for each chemical species were also obtained. A detailed data description was presented by Liu et al. (2005).

2.2. PMF

PMF (Paatero and Tapper, 1994; Paatero, 1997) was used to analyze PM$_{2.5}$ data at the four sites. In this work, measurement uncertainties were used for the error estimates of the measured values; missing data were replaced by the geometric mean of corresponding species and four times of geometric mean was taken as the corresponding error estimates. Half of the detection limit was used for the values below the detection limit and 5/6 of the detection limit was used for the corresponding error estimate (Polissar et al., 1998).

With the total PM$_{2.5}$ mass concentration measured for each sample, multiple linear regression (MLR) was performed to regress the mass concentration against the factor scores obtained from PMF. PMF results always have a portion of unexplained variation. Mass concentrations excluding the unexplained variation portion from G factors (factor contributions) were used to regress the factor scores to obtain the quantitative factor contributions for each resolved factor. PMF was able to resolve eight factors for the two urban sites and seven for the two rural sites. PMF factors were identified as sulfate, nitrate, wood burning, coal combustion, motor vehicle, dust, industrial process, and industrial dust. Results from PMF and more detailed information about the analysis are also available in Liu et al. (2005).

2.3. CMB

EPA's CMB 8.0 model, using the effective variance weighted least-squares fitting, was applied to estimate source contributions to PM$_{2.5}$ for each sampling day. The same data treatments of missing and below detection limit data in PMF analysis were used in the CMB analysis. The major primary source categories included in the CMB analysis are motor vehicle, wood burning, coal combustion (Chow et al., 2004), and dust (Cooper, 1981) for all four sites, and additionally metal production only for the BHM site. Source profiles based on molecular weight fraction for ammonium bisulfate (NH$_4$HSO$_4$), ammonium sulfate ((NH$_4$)$_2$SO$_4$), ammonium nitrates (NH$_4$NO$_3$), and secondary organic carbon (SOC) were also included to address secondary particle formation.

Without additional "marker" species in the measurements, a collinearity problem arises in CMB when profiles for both gasoline and diesel vehicles are included. Thus, one motor vehicle source profile obtained from roadside tests (Chow et al., 2004), instead of two separate source profiles for gasoline and diesel vehicles, was used in CMB. To address the uncertainty of the motor vehicle source profile, sensitivity tests with JST site data were performed with four different profiles including those from dilution chamber (Zielinska et al., 1998) and roadside tests (Chow et al., 2004). Three different composite source profiles were created by weighted-averages of gasoline and diesel vehicle source profiles from dilution chamber tests. Each source profile represents a different type of motor vehicle emission.

The first profile is the diesel vehicle dominant emission (diesel:gasoline = 2:1, 67% diesel and 33% gasoline), the second one is the diesel–gasoline vehicle equivalent emission (diesel:gasoline = 1:1), and the third one is the gasoline vehicle dominant emission (diesel:gasoline = 1:2). Based on the U.S. EPA 2002 national emissions inventories, the ratio of diesel-to-gasoline vehicle emission in Fulton County, GA, where the JST site is located, is about 2, which is close to that of the first profile. The fourth source profile was attained from roadside experiments (Chow et al., 2004).

Applying different motor vehicle source profiles changes the contributions of other sources, though the most notable difference among the sensitivity tests is obviously in the motor vehicle source contribution, increasing from 13–15% (the profiles from dilution chamber experiments) to 21% (the profile from roadside experiments) of $PM_{2.5}$ mass (Fig. 1). The CMB performance parameters and ratios of calculated/measured concentrations indicate that the source apportionment with the roadside source profile has a better fit to the measurements (Fig. 2). Therefore, the source profile from the roadside experiments was used in the CMB source apportionments in this study.

The chemical species used in CMB are assumed to be primary. However, ambient OC measured at a receptor site includes primary OC directly from emission sources and secondary OC (SOC) from photochemical formation. In order to apply CMB using OC information, it is necessary to estimate SOC either by adding a SOC profile (hereafter referred to as CMB_OC) or by estimating primary OC using the EC tracer method (Turpin and Huntzicker, 1991) prior to the source apportionment (hereafter referred to as CMB_POC). Both methods were applied to estimate SOC in this study. The SOC profile is developed by assigning 1.0 as OC fraction in order to account for only OC in secondary organic aerosol (SOA). The estimated SOC is not for SOA, but for SOC in SOA instead. As organic compounds in SOA are also associated with other elements (e.g., oxygen and hydrogen), the OC fraction should be less than 1.0 for SOA. However, incomplete scientific understanding of SOA formation limits estimating the true OC fraction in SOA, which may introduce errors into SOC estimates. The EC tracer method is an indirect method, which estimates primary and secondary OC, based on the fact that EC is a good tracer for primary carbonaceous particles (i.e., OC and EC) from combustion sources. The Deming linear regression was applied for daily OC and EC data in the lowest 10% by OC/EC ratio to obtain primary OC/EC ratios for four different seasons at each site (Lim and Turpin, 2002). Primary OC was calculated by multiplying OC/EC ratio to EC assuming that all EC is primary and then SOC was estimated by subtracting primary OC from measured OC.

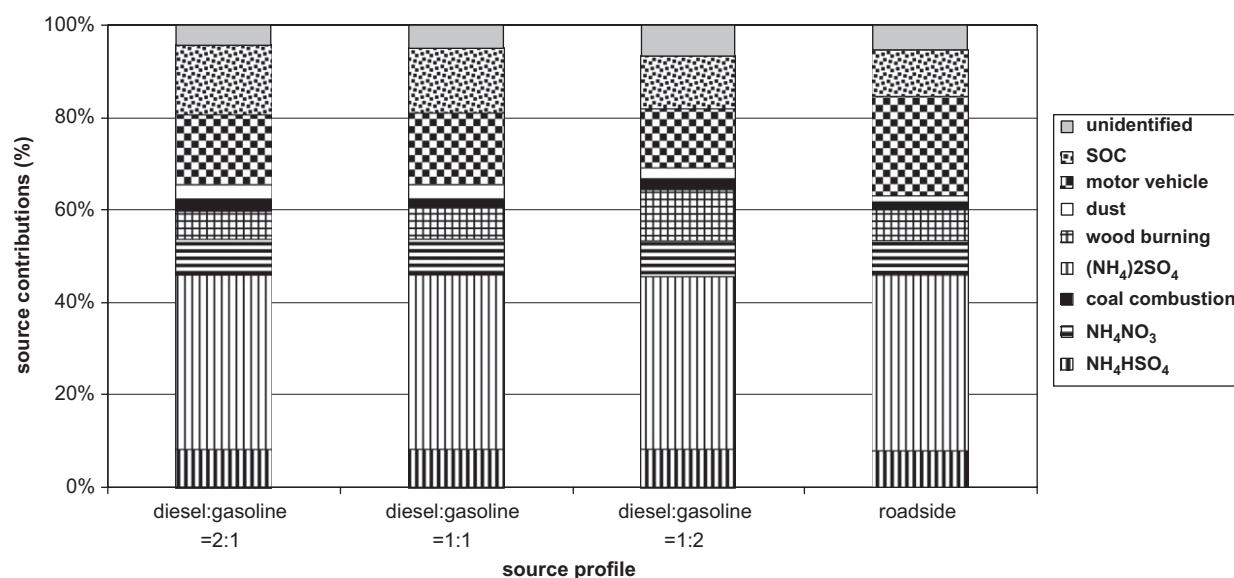Prior to comparing with PMF results, the performance of CMB source apportionment was

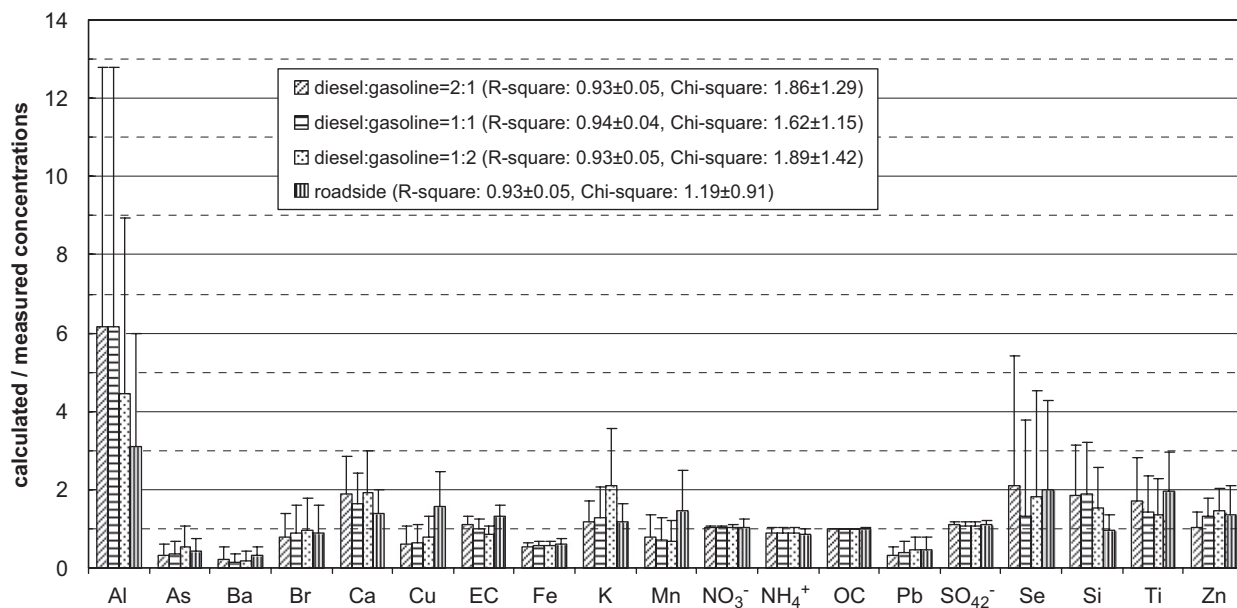Fig. 1. Average source contributions (%) to $PM_{2.5}$ mass for four sensitivity tests at JST.

Fig. 2. Ratios of calculated-to-measured concentrations of fitting species for four sensitivity tests at JST.

3. Results

validated by performance measures. The averaged performance measures for CMB (i.e., r^2, χ^2, and predicted percent mass) are 0.94 ± 0.05, 1.05 ± 0.91, 90 ± 14 (CMB_OC), and 0.92 ± 0.06, 1.21 ± 0.93, 90 ± 13 (CMB_POC) at JST, 0.93 ± 0.09, 1.60 ± 2.11, 91 ± 15 (CMB_OC), and 0.92 ± 0.09, 1.56 ± 2.04, 91 ± 16 (CMB_POC) at BHM, and 0.95 ± 0.05, 1.13 ± 1.13, 85 ± 13 (CMB_OC), and 0.94 ± 0.05, 1.05 ± 1.04, 86 ± 13 (CMB_POC) at YRK, and 0.97 ± 0.04, 0.61 ± 0.76, 78 ± 10 (CMB_OC), and 0.96 ± 0.04, 0.58 ± 0.67, 78 ± 10 (CMB_POC) at CTR. The CMB performance measures were satisfied for the recommended targets (Watson et al., 1998).

3. Results

Both CMB and PMF results show that secondary sulfate is the dominant contributor (33–45%) of ambient $PM_{2.5}$ mass in urban and rural areas (Figs. 3 and 4). Motor vehicle and wood burning are the two major primary sources. The former contributes 17–25% in urban areas and 7–9% in rural areas; the latter contributes 6–13% in urban areas and 6–30% in rural areas. Coal combustion, dust, and industrial/metal process are minor contributors (2–13% of $PM_{2.5}$ mass).

Secondary sulfate is a dominant source factor in urban and rural areas. CMB has two source profiles to represent secondary sulfate (NH_4HSO_4 and

$(NH_4)_2SO_4$). PMF resolved factor profiles are comparable with the CMB source sulfate profiles at all four sites (Fig. 5a) except that OC and small amounts of EC are associated with this factor in the PMF results. In PMF, the OC and EC components reflect mixing of primary and secondary pollutants while CMB is not affected by mixing. The resolved PMF factor therefore does not represent a single pure source. The OC association implies that SOA formation coincides with the secondary sulfate formation, while the small EC content likely reflects an increase of sulfate and EC concentrations during stagnant conditions (see Figs. S1 and S2 in the supporting information). In the PMF sulfate factor, molar ratios of ammonium to sulfate are 2.3, 2.0, 2.1, and 1.6 for the JST, YRK, BHM, and CTR site, respectively. The ratios suggest that sulfate is present primarily as ammonium sulfate at these four receptor sites, although sulfate at CTR is probably not fully neutralized. CMB results show that NH_4HSO_4 contributes 16% (40% of secondary sulfate) to $PM_{2.5}$ at CTR, which is much larger than the other sites, indicating a lower degree of sulfate neutralization. Both PMF and CMB results show partial sulfate neutralization at CTR, which is consistent with the result by Edgerton et al. (2005), implying less NH_3 availability at CRT than other sites. The estimated sulfate source contributions from PMF and CMB are well correlated
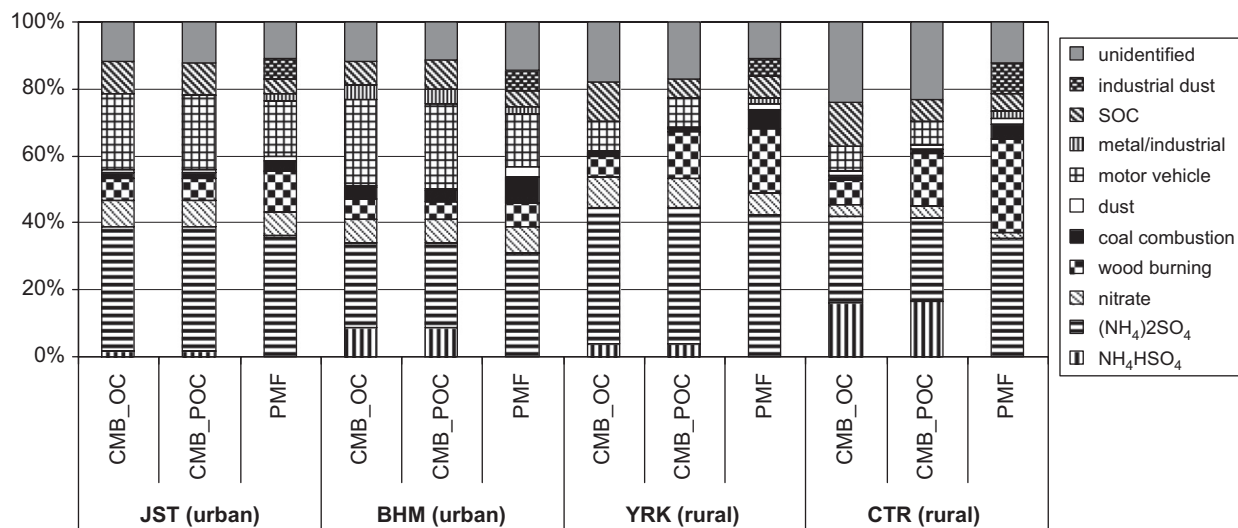
Fig. 3. Averaged source contributions (%) to PM$_{2.5}$ mass. Note sulfate factors from PMF are labeled as (NH$_4$)$_2$SO$_4$ just for display purpose.
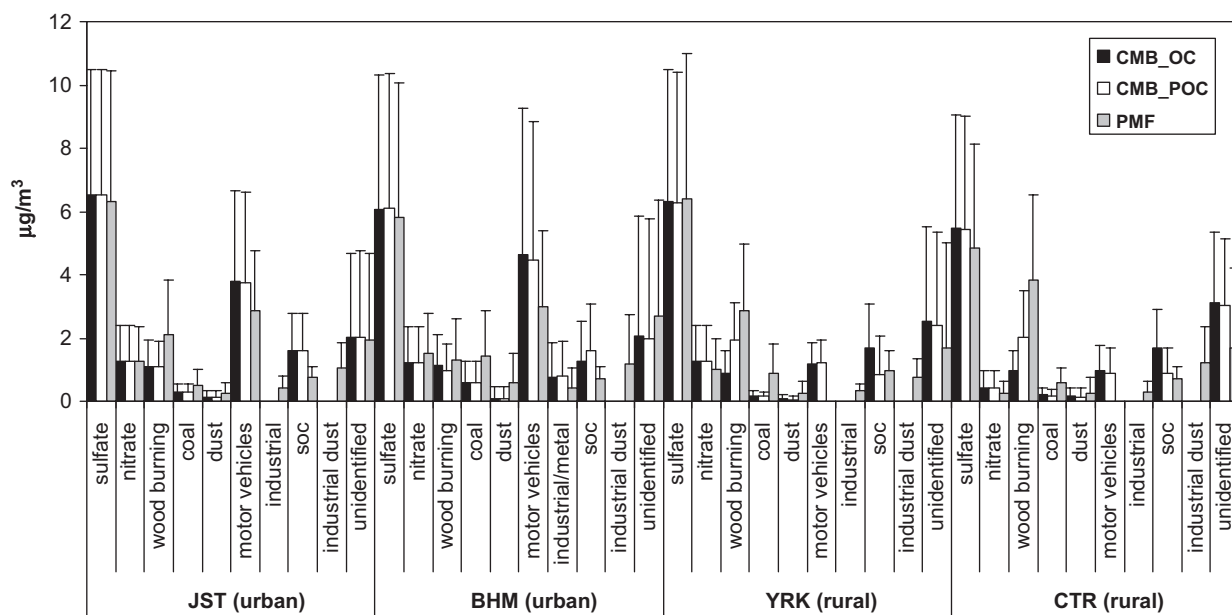

Fig. 4. Averaged source contributions (μg m^{-3}). Note the bars represent standard deviations. Note NH$_4$HSO$_4$ and (NH$_4$)$_2$SO$_4$ from CMB are lumped together as sulfate for comparison purpose.

(correlation coefficient $(r) = 0.94$–0.99 in Table 1) and the average source contributions agree well (Fig. 4; CMB [6.53 ± 3.96], PMF [6.32 ± 4.14]).

PMF resolved a nitrate factor, which corresponds to the NH$_4$NO$_3$ profile in CMB (Fig. 5b). Sulfate and some OC are mixed in the PMF profiles, likely arising from concurrent oxidations of NO$_2$, SO$_2$,

and VOCs. Average source contributions from PMF and CMB agree well with a good correlation (Table 1 and Fig. 4). The nitrate source contributions are considerably lower at the CTR site, less than a half of those at the other sites, reflecting that sulfate is not fully neutralized limiting the availability of free ammonia at this site.
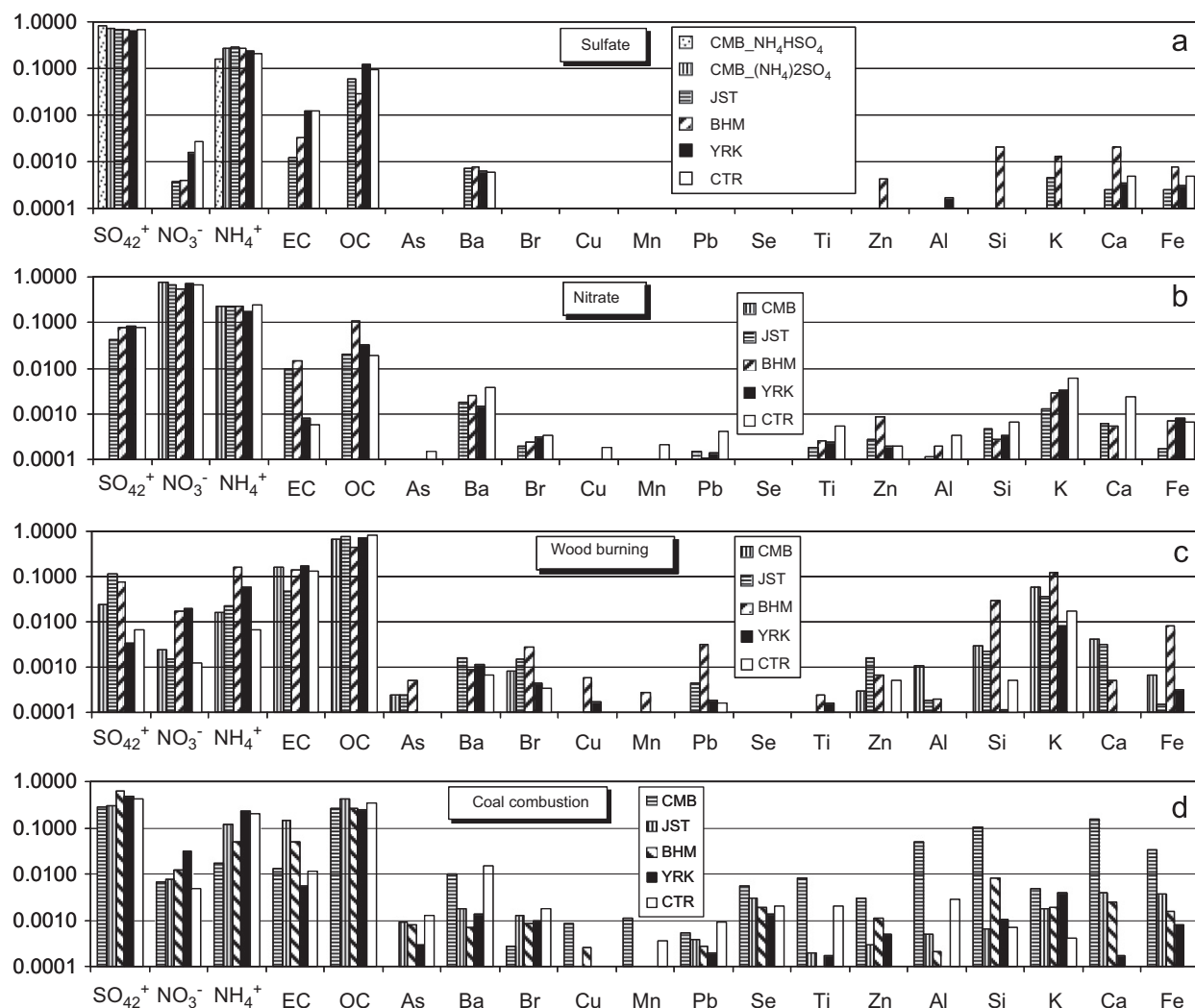
Fig. 5. Source (or factor) profiles for CMB and PMF: (a) sulfate, (b) nitrate, (c) wood burning, and (d) coal combustion. One profile is used for each source in CMB. PMF resolved factor profiles are site specific.

The CMB source profile and the PMF factors associated with wood burning are in relatively good agreement, with high concentrations of OC, EC, and K, although PMF factor profiles are mixed with some sulfate at the urban sites (Fig. 5c). In general, wood burning contributions from the two methods are correlated well at the urban sites, but slightly lower correlations are observed at the rural sites. The weaker correlations at the rural sites are likely because PMF cannot resolve a motor vehicle factor for the rural sites so that OC and EC become more influential in identifying wood burning rather than K. For the urban sites, OC and EC are the main species resolving the motor vehicle factors and K

is the most influential species for wood burning (Table S.1.c and d). An analogous case is found for CMB analysis. In the urban areas, average source contributions by CMB are comparable or less than those by PMF, while in the rural areas, source contributions by PMF are much larger than those by CMB.

The PMF coal combustion factors have strong signals of sulfate, ammonium, EC, OC, and Se at all sites (Fig. 5d). The CMB source profile has higher levels of elements associated with dust. This difference may be due, in part, to a lack of locally appropriate source profile for used in CMB, or PMF is "losing" the other elements to dust and

Table 1
Correlation coefficients between CMB and PMF source contributions

	Sulfate	Nitrate	Wood burning	Coal combustion	Dust[a]	Motor vehicle	SOC	Metal/ industrial
JST[b]								
CMB_OC vs. CMB_POC	1.00	1.00	0.91	1.00	1.00	1.00	0.84	
CMB_OC vs. PMF	0.99	1.00	0.81	0.23	0.85	0.85	0.58	
CMB_POC vs. PMF	0.99	1.00	0.76	0.24	0.85	0.86	0.58	
BHM								
CMB_OC vs. CMB_POC	1.00	1.00	0.97	1.00	1.00	1.00	0.84	1.00
CMB_OC vs. PMF	0.94	0.98	0.83	0.26	0.82	0.94	0.60	0.86
CMB_POC vs. PMF	0.94	0.98	0.81	0.26	0.83	0.95	0.78	0.86
YRK								
CMB_OC vs. CMB_POC	1.00	1.00	0.52	0.98	0.99	0.97	0.67	
CMB_OC vs. PMF	0.97	0.97	0.70	0.09	0.88		0.68	
CMB_POC vs. PMF	0.97	0.97	0.59	0.09	0.88		0.40	
CTR								
CMB_OC vs. CMB_POC	1.00	1.00	0.69	0.99	0.99	0.95	0.76	
CMB_OC vs. PMF	0.99	0.98	0.77	−0.12	0.89		0.74	
CMB_POC vs. PMF	0.99	0.98	0.76	−0.13	0.90		0.50	

[a] Note that PMF resolve two dust factors (dust and industrial dust). Industrial dust is not used for the calculation.

[b] The sample size of JST is three times larger than that of other sites because of different sampling frequency (everyday at JST vs. every third day at the other sites). The larger sample size effect at JST on the correlation coefficients was investigated by recalculating the correlation coefficients for every third day source apportionment results. We found that the larger sampling size had little effect on the correlation coefficient estimates.

industrial dust factors. In PMF, Se is the primary driving species for coal combustion, whereas in CMB, Ca is the most influential species along with Fe, Si, and Al. As a result, the correlations between CMB and PMF are poor at all sites. Since Se is a well-known marker for coal combustion, the poor correlations between Se and coal combustion source impacts estimated by CMB suggest that the CMB analysis of this source needs to be improved either by developing local source profiles for coal combustion or by incorporating other information (e.g., gas-to-particle ratios as done in Marmur et al., 2005).

A motor vehicle factor is resolved only at the two urban sites by PMF. Motor vehicle contributions estimated by CMB at the two rural sites are much smaller ($<1/3$) than the urban sites. This factor has large contributions of EC and OC (Fig. 6a). However, the OC/EC ratios in the source profile assigned in CMB and those in the factors calculated by PMF are different. The ratios are about 2 in PMF as compared to 1 in CMB. The profiles from PMF are more close to the emission characteristics of gasoline vehicles, whereas that from CMB represents mixed emissions from both gasoline and

diesel vehicles. The larger OC contributions in the PMF factors could arise from condensation of SOC onto the particles, or that the ratio in the profile used in the CMB analysis is low. Despite the difference in the OC/EC ratios, source contributions from both models are well correlated for the urban sites since OC and EC are the main species driving motor vehicle source impacts in both methods (Table S.1.a and b). Motor vehicle sources have been identified at the two rural sites in other studies (Zheng et al., 2002a, 2006). However, they are not resolved in PMF in this study. The unreserved motor vehicle source in PMF affects source apportionment results (e.g., wood burning impacts), and may introduce errors into the outcomes of health studies since a known source is not incorporated into the health studies. It is necessary to utilize more information (e.g., temperature-resolved carbon Liu et al., 2006) to resolve the motor vehicle source.

PMF resolved two dust factors, dust and industrial dust, while CMB has only one dust profile (Fig. 6b and c). The dust factors are rich in Si, Al, K, Ca, and Fe, and are associated with some OC and sulfate. The Al/Si ratio (0.47) in the dust profile
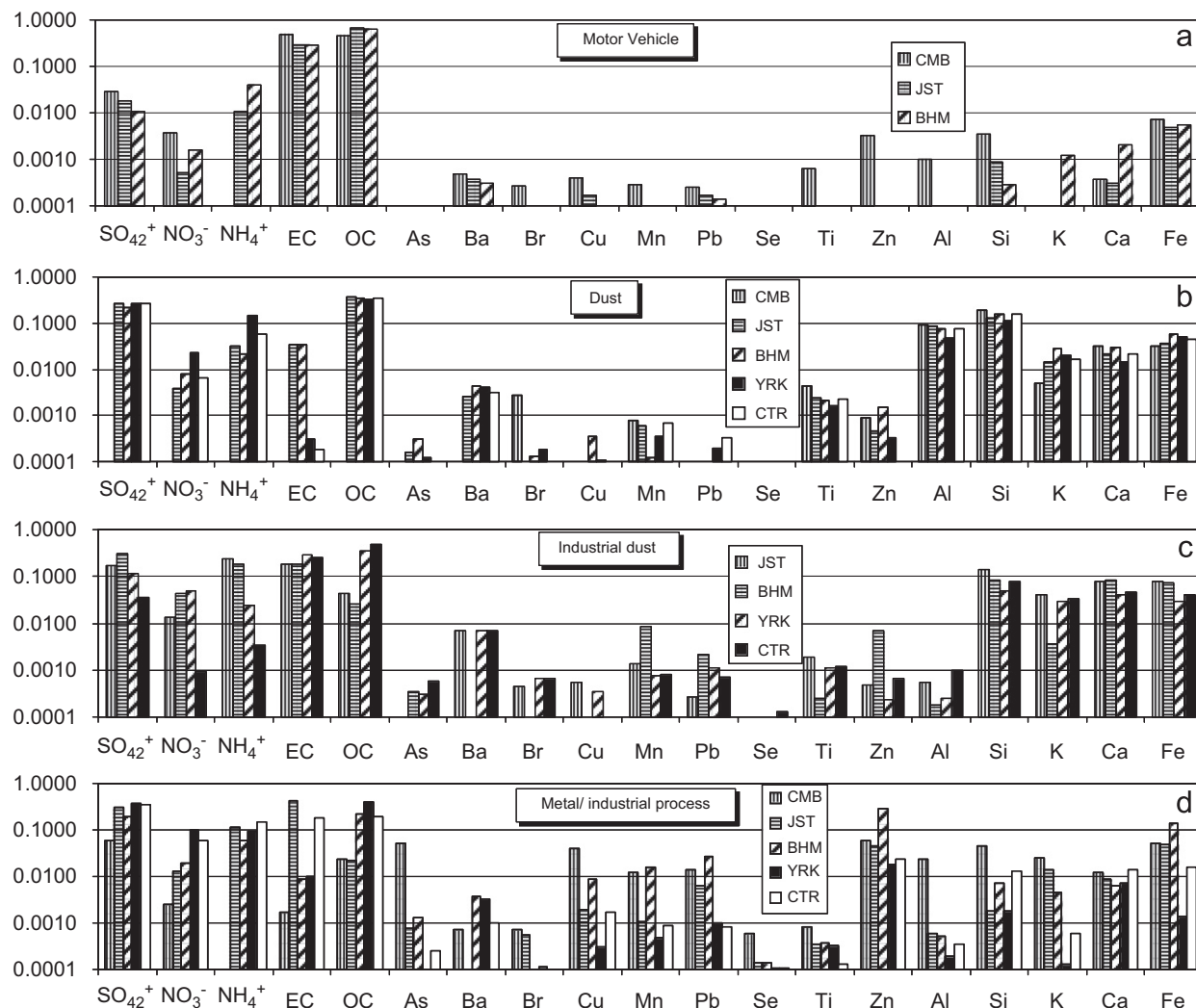
Fig. 6. Source (or factor) profiles for CMB and PMF: (a) motor vehicles, (b) dust, (c) dust 2, and (c) metal and industrial process. One profile is used for each source in CMB. PMF resolved factor profiles are site specific.

used in CMB is similar with those (0.67, 0.49, 0.44, and 0.48) in the PMF dust factors. In both models, Al and Si are the primary species driving dust source contributions (Table S.1). As a result of that, dust source contributions are well correlated ($r = 0.82$–0.90). Inspection of the time series of the dust factor reveals that the factor is episodic in nature; these episodes correspond most likely to long-range transport from Asian and Saharan deserts (Liu et al., 2005). The industrial dust factors only resolved by PMF are rich in Si, Ca, Fe, and K coupled with EC, OC, and sulfate. Calcium (Ca) is the most influential species of the industrial dust along with Si and Fe, which are the

primary species impacting coal combustion in CMB (Table S.1).

SOC was estimated by adding a SOC profile in CMB in CMB_OC and by using the EC tracer method prior to CMB source apportionment in CMB_POC. In PMF, SOC was calculated by summing OC fractions mixed in the PMF sulfate and nitrate factors, and unexplained variations. SOC levels from CMB_OC are comparable to those from CMB_POC at urban sites, but they are larger than those from CMB_POC at rural sites (Fig. 4). In general, SOC concentrations from PMF are lower than those from CMB_OC or CMB_POC except CMB_POC at the YRK site. Yuan et al. (2005) also

found that SOC estimated from PMF is lower than that from the EC tracer method. SOC estimates of both CMB_POC and CMB_OC are correlated well ($r = 0.67$–0.84, Table 1). Yet, the correlations between SOC estimated by CMB and PMF are generally lower than those between SOC estimates of two CMB except the correlations between PMF and CMB_OC at the rural sites. Overall, the correlations between SOC estimated by CMB and PMF are not as high as observed in secondary inorganic particles (i.e., sulfate and nitrate).

The large uncertainty in the SOC estimates of CMB_OC and CMB_POC arises from the fact that OC dominant emission sources (e.g., meat cooking and natural gas combustion) may be aggregated into SOC, leading to overestimation, and the assumption that a single seasonal OC/EC ratio can represent a mixture of primary sources varying in time and space, respectively, while SOC mixed into the primary source factors in PMF may lead to underestimation. There is also evidence that meat cooking and natural gas combustion contribute significantly to OC in the southeast (Baek et al., 2005; Zheng et al., 2002b). However, neither PMF nor CMB can resolve such sources because of the lack of additional markers and source profiles. These source impacts must have been distributed among the sources/factors resolved above to an unknown degree. This may also cause some difference between the two models.

4. Discussion

Two different receptor models (PMF and CMB) were applied to estimate $PM_{2.5}$ source impacts at four different receptor sites in the southeastern US. Secondary sulfate and, to a much lesser extent, SOC and nitrate account for a large portion of ambient $PM_{2.5}$ mass in the Southeast. Among primary sources, motor vehicles and wood burning are suggested as the major primary sources at the urban and/or rural sites.

Sulfate and nitrate contributions estimated by the two methods are the most comparable. For some other sources/factors, even for those highly correlated, there are differences in estimated source contributions. Several factors may contribute to the differences in the source apportionment results. The source profiles of primary emission sources in CMB may not be representative due to a lack of locally available source profiles, while PMF may provide information towards developing more

locally representative source profiles, though atmospheric processing and limitations in the PMF method also affect profiles. One assumption in CMB is that source profiles (i.e., source characteristics) do not change after emissions. In comparison, more processed (or aged) factors are found using PMF. Therefore, factors resolved using PMF are more likely mixed source profiles of primary and secondary compounds, and such factors can have contributions from multiple sources. The omission of possible sources and lack of proper markers for other primary OC sources (e.g., meat cooking and natural gas combustion) may also lead to differences as shown in coal combustion, major OC sources (e.g., motor vehicles and wood burning), and SOC. In addition, neither PMF nor CMB, as applied, accounts for the seasonal variations of source profiles.

The degree of correlation between sources apportionment results of CMB and PMF varies depending on sources and receptor sites. Secondary inorganic particles (i.e., sulfate and nitrate) and dust source contributions are highly correlated ($r = 0.97$–1.0 and $r = 0.82$–0.90, respectively) at all four sites, whereas little correlation is found for coal combustion at all four sites. Motor vehicle source contributions are correlated well ($r = 0.85$–0.95) at urban sites, though motor vehicle sources are not resolved by PMF at rural sites. For wood burning, better correlations are observed at the urban ($r = 0.76$–0.83) than the rural sites ($r = 0.59$–0.77). SOC is not correlated as highly as other secondary sulfate and nitrate particles ($r = 0.40$–0.74). The degree of the correlations is important for PM epidemiologic studies (i.e., time-series health studies), in which source apportionment results are associated with adverse health outcomes. The results from this study suggest that the selection of either method may not change the outcomes of time-series health studies for highly correlated sources/factors, whereas different results may be produced for less correlated sources/factors. Recent intercomparison studies (Ito et al., 2006; Mar et al., 2006; Thurston et al., 2005) of various source apportionment results for PM time-series study have shown that changes in health risk ratio estimates are small even for weakly correlated sources (median $r \approx 0.4$, interquartile range of r: 0.3–0.6) and uncertainties introduced by different source apportionment methods are significantly less than the overall uncertainty of PM time-series health study (e.g., mortality regression process).

Changes in the outcomes of PM time-series health study due to using different (or less correlated) source apportionment results may not be significant even for less correlated sources except motor vehicles at rural sites and coal combustion at all sites since the calculated correlation coefficients in this study are well above or within the range of those in the intercomparison study (Thurston et al., 2005). Source apportionments by either PMF or CMB of motor vehicles at rural sites and coal combustion need further improvements, which may be possible by incorporating more information (e.g., temperature-resolved carbon or trace gases).

Based on the results from this study, we draw some general conclusions here. While the source types resolved by PMF and CMB are similar, the estimated contributions of primary sources can be quite different depending on the method used. The implications for optimizing the emission controls, for example, can therefore be quite different. In occasions without adequate knowledge of emission characteristics (i.e., source profiles) of local sources, PMF is a more appropriate method for source apportionment than CMB since CMB source apportionment results are sensitive to source profiles used (Lee et al., 2005; Lee and Russell, 2007). However, it is difficult to assert one method is superior over the other given the uncertainties discussed previously without elaborate studies and a priori knowledge of the emission characteristics of all major local sources at a site. We suggest that multiple apportionment methods are used when possible to assess the uncertainties in the source apportionment results. For time-series health studies, which are more dependent on the correlations with factors/sources, outcomes using highly correlated source apportionment results of PMF and CMB are more compatible. We expect that the same finding will be shown in the regions other than the Southeast since the correlation agreement is due to similarities in the profiles between PMF and CMB, which is dictated qualitatively by the general characteristics of the sources/factors.

Acknowledgment

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.atmosenv.2008.01.025.

References

Baek, J., Park, S.K., Hu, Y., Russell, A.G., 2005. Source apportionment of fine organic aerosol using CMAQ tracers. In: Models-3 Conference, Research Triangle Park, NC.

Chow, J.C., Watson, J.G., Kuhns, H., Etyemezian, V., Lowenthal, D.H., Crow, D., Kohl, S.D., Engelbrecht, J.P., Green, M.C., 2004. Source profiles for industrial, mobile, and area sources in the big bend regional aerosol visibility and observational study. Chemosphere 54 (2), 185–208.

Cooper, J.A., 1981. Determination of source contributions to fine and coarse suspended particulate levels in Petersville, Alabama. Report to Tennessee Valley Authority by NEA, Inc.

Edgerton, E.S., Hartsell, B.E., Saylor, R.D., Jansen, J.J., Hansen, D.A., Hidy, G.M., 2005. The southeastern aerosol research and characterization study: Part II. Filter-based measurements of fine and coarse particulate matter mass and composition. Journal of Air and Waste Management Association 55 (10), 1527–1542.

Hansen, D.A., Edgerton, E.S., Hartsell, B.E., Jansen, J.J., Kandasamy, N., Hidy, G.M., Blanchard, C.L., 2003. The southeastern aerosol research and characterization study: Part 1—Overview. Journal of Air and Waste Management Association 53 (12), 1460–1471.

Hopke, P.K., 2003. Recent developments in receptor modeling. Journal of Chemometrics 17 (5), 255–265.

Ito, K., Christensen, W.F., Eatough, D.J., Henry, R.C., Kim, E., Laden, F., Lall, R., Larson, T.V., Neas, L., Hopke, P.K., Thurston, G.D., 2006. PM source apportionment and health effects: 2. An investigation of intermethod variability in associations between source-apportioned fine particle mass and daily mortality in Washington, DC. Journal of Exposure Science and Environmental Epidemiology 16 (4), 300–310.

Laden, F., Neas, L.M., Dockery, D.W., Schwartz, J., 2000. Association of fine particulate matter from different sources with daily mortality in six US cities. Environmental Health Perspectives 108 (10), 941–947.

Lee, S., Russell, A.G., 2007. Estimating uncertainties and uncertainty contributors of CMB PM$_{2.5}$ source apportionment results. Atmospheric Environment 41 (40), 9616–9624.

Lee, S., Baumann, K., Schauer, J.J., Sheesley, R.J., Naeher, L.P., Meinardi, S., Blake, D.R., Edgerton, E.S., Russell, A.G., Clements, M., 2005. Gaseous and particulate emissions from prescribed burning in Georgia. Environmental Science and Technology 39 (23), 9049–9056.

Lim, H.J., Turpin, B.J., 2002. Origins of primary and secondary organic aerosol in Atlanta: results of time-resolved measurements during the Atlanta supersite experiment. Environmental Science and Technology 36 (21), 4489–4496.

Liu, W., Wang, Y.H., Russell, A., Edgerton, E.S., 2005. Atmospheric aerosol over two urban–rural pairs in the

southeastern United States: chemical composition and possible sources. Atmospheric Environment 39 (25), 4453–4470.

Liu, W., Wang, Y.H., Russell, A., Edgerton, E.S., 2006. Enhanced source identification of southeast aerosols using temperature-resolved carbon fractions and gas phase components. Atmospheric Environment 40, S445–S466.

Mar, T.F., Norris, G.A., Koenig, J.Q., Larson, T.V., 2000. Associations between air pollution and mortality in Phoenix, 1995–1997. Environmental Health Perspectives 108 (4), 347–353.

Mar, T.F., Ito, K., Koenig, J.Q., Larson, T.V., Eatough, D.J., Henry, R.C., Kim, E., Laden, F., Lall, R., Neas, L., Stolzel, M., Paatero, P., Hopke, P.K., Thurston, G.D., 2006. PM source apportionment and health effects. 3. Investigation of inter-method variations in associations between estimated source contributions Of $PM_{2.5}$ and daily mortality in Phoenix, AZ. Journal of Exposure Science and Environmental Epidemiology 16 (4), 311–320.

Marmur, A., Unal, A., Mulholland, J.A., Russell, A.G., 2005. Optimization-based source apportionment of $PM_{2.5}$ incorporating gas-to-particle ratios. Environmental Science and Technology 39 (9), 3245–3254.

Metzger, K.B., Tolbert, P.E., Klein, M., Peel, J.L., Flanders, W.D., Todd, K., Mulholland, J.A., Ryan, P.B., Frumkin, H., 2004. Ambient air pollution and cardiovascular emergency department visits. Epidemiology 15 (1), 46–56.

Paatero, P., 1997. Least squares formulation of robust non-negative factor analysis. Chemometrics and Intelligent Laboratory Systems 37 (1), 23–35.

Paatero, P., Tapper, U., 1993. Analysis of different modes of factor-analysis as least-squares fit problems. Chemometrics and Intelligent Laboratory Systems 18 (2), 183–194.

Paatero, P., Tapper, U., 1994. Positive matrix factorization—a nonnegative factor model with optimal utilization of error—estimates of data values. Environmetrics 5 (2), 111–126.

Peel, J.L., Tolbert, P.E., Klein, M., Metzger, K.B., Flanders, W.D., Mulholland, J.A., Todd, K., 2004. Ambient air pollution and cardiorespiratory emergency department visits, 1993–2002. Epidemiology 15 (4), S57.

Peters, A., Dockery, D.W., Muller, J.E., Mittleman, M.A., 2001. Increased particulate air pollution and the triggering of myocardial infarction. Circulation 103 (23), 2810–2815.

Polissar, A.V., Hopke, P.K., Paatero, P., 1998. Atmospheric aerosol over Alaska-2 elemental composition and sources. Journal of Geophysical Research—Atmospheres 103 (D15), 19045–19057.

Pope, C.A., Burnett, R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K., Thurston, G.D., 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Journal of the American Medical Association 287 (9), 1132–1141.

Sarnat, J.A., Marmur, A., Klein, M., Kim, E., Russell, A.G., Mulholland, J.A., Hopke, P.K., Sarnat, S.E., Peel, J.L., Tolbert, P.E., 2006. Associations between source-resolved particulate matter and cardiorespiratory emergency department visits. Epidemiology 17 (6), S267–S268.

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 particulate matter health effects—intercomparison of results and implications. Environmental Health Perspectives 113 (12), 1768–1774.

Turpin, B.J., Huntzicker, J.J., 1991. Secondary formation of organic aerosol in the Los-Angeles basin—a descriptive analysis of organic and elemental carbon concentrations. Atmospheric Environment Part A—General Topics 25 (2), 207–215.

US-EPA, 2002. EPA Acid Rain Program 2001 Annual Progress Report. Clean air Markets Program. Office of Air and Radiation, Washington, DC.

Watson, J.G., 2002. Visibility: science and regulation. Journal of Air & Waste Management Association 52 (6), 628–713.

Watson, J.G., Cooper, J.A., Huntzicker, J.J., 1984. The effective variance weighting for least-squares calculations applied to the mass balance receptor model. Atmospheric Environment 18 (7), 1347–1355.

Watson, J.G., Robinson, N.F., Fujita, E.M., Chow, J.C., Pace, T.G., Lewis, C., Coulter, T., 1998. CMB8 Applications and Validation Protocol for $PM_{2.5}$ and VOCs. Desert Research Institute, Reno, NV.

www.epa.gov/air/data/ The US EPA AirData Web site.

Yuan, Z.B., Yu, J.Z., Lau, A.K.H., Louie, P.K.K., Fung, J.C.H., 2005. Application of positive matrix factorization in estimating aerosol secondary organic carbon in Hong Kong and insights into the formation mechanisms. Atmospheric Chemistry and Physics Discussions 5 (4), 5299–5324.

Zheng, M., Cass, G.R., Schauer, J.J., Edgerton, E.S., 2002a. Source apportionment of $PM_{2.5}$ in the southeastern United States using solvent-extractable organic compounds as tracers. Environmental Science and Technology 36 (11), 2361–2371.

Zheng, M., Cass, G.R., Schauer, J.J., Edgerton, E.S., 2002b. Source apportionment of $PM_{2.5}$ in the southeastern United States using solvent-extractable organic compounds as tracers. Environmental Science and Technology 36 (11), 2361–2371.

Zheng, M., Ke, L., Edgerton, E.S., Schauer, J.J., Dong, M.Y., Russell, A.G., 2006. Spatial distribution of carbonaceous aerosol in the southeastern United States using molecular markers and carbon isotope data. Journal of Geophysical Research—Atmospheres 111 (D10).

Zielinska, B., McDonald, J.D., Hayes, T., Chow, J.C., Fujita, E.M., Watson, J.G., 1998. Northern Front Range Air Qaulity Study Final Report. Volume B: Source Measurements. Desert Research Institute, Reno, NV.