Published in final edited form as:

Atmos Environ (1994). 2013 February 1; 65: 11–20. doi:10.1016/j.atmosenv.2012.09.034.

Positive matrix factorization of a 32-month series of daily PM_{2.5} speciation data with incorporation of temperature stratification

Mingjie Xie^a, Ricardo Piedrahita^a, Steven J. Dutton^b, Jana B. Milford^a, Joshua G. Hemann^a, Jennifer L. Peel^c, Shelly L. Miller^a, Sun-Young Kim^d, Sverre Vedal^d, Lianne Sheppard^d, and Michael P. Hannigan^{a,*}

^aDepartment of Mechanical Engineering, College of Engineering and Applied Science, University of Colorado, Boulder, CO 80309, USA

^bNational Center for Environmental Assessment, U.S. Environmental Protection Agency, Research Triangle Park, NC 27711, USA

^cDepartment of Environmental and Radiological Health Sciences, Colorado State University, Fort Collins, CO 80523, USA

^dDepartment of Environmental and Occupational Health Sciences, University of Washington, School of Public Health, University of Washington, Seattle, WA 98195, USA

Abstract

This study presents source apportionment results for PM_{2.5} from applying positive matrix factorization (PMF) to a 32-month series of daily PM_{2.5} compositional data from Denver, CO, including concentrations of sulfate, nitrate, bulk elemental carbon (EC) and organic carbon (OC), and 51 organic molecular markers (OMMs). An optimum 8-factor solution was determined primarily based on the interpretability of the PMF results and rate of matching factors from bootstrapped PMF solutions with those from the base case solution. These eight factors were identified as inorganic ion, *n*-alkane, EC/sterane, light *n*-alkane/polycyclic aromatic hydrocarbon (PAH), medium alkane/alkanoic acid, PAH, winter/methoxyphenol and summer/odd *n*-alkane. The inorganic ion factor dominated the reconstructed PM_{2.5} mass (sulfate + nitrate + EC + OC) in cold periods (daily average temperature < 10 °C; 43.7% of reconstructed PM_{2.5} mass) whereas the summer/odd *n*-alkane factor dominated in hot periods (> 20 °C; 53.1%). The two factors had comparable relative contributions of 26.5% and 27.1% in warm periods with temperatures between 10 °C and 20 °C. Each of the seven factors resolved in a previous study (Dutton et al., 2010b) using a 1-year data set from the same location matches one factor from the current work based on comparing factor profiles. Six out of the seven matched pairs of factors are linked to similar source classes as suggested by the strong correlations between factor contributions ($r = 0.89 - 0.98$). Temperature-stratified source apportionment was conducted for three subsets of the

© 2012 Elsevier Ltd. All rights reserved.

*Corresponding author: Prof. Michael Hannigan, Michael.Hannigan@Colorado.edu, Tel: +1303 735 5045; Fax: +1 303 492 3498, Mailing Address: Department of Mechanical Engineering, Engineering Center, 427 UCB, Boulder, CO 80309-0427.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

data in the current study, corresponding to the cold, warm and hot periods mentioned above. The cold period (7-factor) solution exhibited a similar distribution of reconstructed PM_{2.5} mass as the full data set solution. The factor contributions of the warm period (7-factor) solution were well correlated with those from the full data set solution ($r = 0.76 - 0.99$). However, the reconstructed PM_{2.5} mass was distributed more to inorganic ion, *n*-alkane and medium alkane/alkanoic acid factors in the warm period solution than in the full data set solution. For the hot period (6-factor) solution, PM_{2.5} mass distribution was quite different from that of the full data set solution, as illustrated by regression slopes as low as 0.2 and as high as 4.8 of each matched pair of factors across the two solutions.

Keywords

Source apportionment; Positive matrix factorization; Organic molecular markers; Temperature-stratified; Reconstructed PM_{2.5} mass

1. Introduction

The Denver Aerosol Sources and Health (DASH) study aims to identify and quantify the sources of ambient PM_{2.5} that are detrimental to human health given short-term exposure (Vedal et al., 2009). Daily PM_{2.5} samples were collected from a centrally located site in Denver, Colorado, over the period from July 2002 through December 2008. Speciation of the PM_{2.5} included gravimetric mass, inorganic ionic compounds (sulfate, nitrate and ammonium), elemental carbon (EC), organic carbon (OC), and a large array of organic molecular markers (OMMs). Prior DASH studies have discussed the methods used for chemical speciation and point-wise uncertainty estimation (Dutton et al., 2009a, b). The temporal patterns in bulk species (July 1, 2002 – December 31, 2006) and OMMs (July 1, 2002 – December 31, 2003) were published by Dutton et al. (2010a). Moreover, a 1-year data set (January 27, 2003 – December 31, 2003) of inorganic and organic speciated PM_{2.5} was analyzed using positive matrix factorization (PMF) for source apportionment (Dutton et al., 2010b). In that work, two out of the seven factors identified, together accounting for 16% of the annual EC mass and 24% of the annual OC mass, could not be clearly associated with PM_{2.5} sources.

In the current study, a 32-month series of daily speciated PM_{2.5} data, including the 1-year data set analyzed previously, was utilized as input for source apportionment to obtain more reliable PM_{2.5} source information. The PMF2 model, coupled with a stationary block bootstrap technique for quantifying errors due to random sampling (Hemann et al., 2009), was the primary source apportionment tool. In this work, correlations between factor contribution time series and meteorological and trace gas measurements were used to aid the interpretation of each factor. The 32-month source apportionment results were then compared to those published earlier based on a 1-year data set. In previous receptor-based source apportionment studies, source profiles have typically been assumed to be constant over the period of ambient and source sampling (Chen et al., 2011). This assumption is not true for semi-volatile and secondary species. When employing PMF to understand source contributions, care must be taken in attributing source types to output factors because a factor is not necessarily a source; a factor could be an atmospheric process like

photochemistry or gas-particle partitioning or a combination of sources and/or processes. Thus, a temperature stratified analysis was undertaken in current work. Sub-sets of the data were used as input to identify the influence of temperature on factors resolved using the full data set, which was feasible because 32 months of daily speciated data were available. The source apportionment results presented will be used to identify the associations between individual factor contributions and short-term adverse health effects in a future study.

2. Method

2.1. Ambient data collection and chemical speciation

PM_{2.5} samples were collected daily on the roof of a two-story elementary school building in Denver. This site was located in a residential area that is roughly in the center of the metropolitan area, far (~5 km) from any major industrial emission sources and heavy traffic (Vedal et al., 2009). Details of the sampling set up, protocols, and chemical analysis were given by Dutton et al. (2009a, b). Both Teflon and quartz fiber filters were used for ambient PM_{2.5} collection from the same air stream. Teflon sample speciation included gravimetric mass, inorganic ions (sulfate, nitrate and ammonium). The quartz fiber filters were analyzed for bulk EC, OC and 73 OMMs. Table S1 (Supporting Information) lists the statistics for each species quantified from January 27, 2003 to October 2, 2005, including the mean and median concentrations, coefficient of variation (CV = standard deviation/mean concentration), signal to noise ratio (S/N = mean concentration/mean uncertainty), percent of missing values and observations below detection limit (BDL). During the same study period, meteorological and trace gas measurement data were obtained from the National Renewable Energy Laboratory (NREL; solar irradiance) and a nearby community monitor (CAMP, AQS ID: 080310002) operated by the Colorado Department of Public Health and Environment (CDPHE; temperature, relative humidity, ozone, nitrogen oxides (NO_x) and CO). Mean values and standard deviations for these variables are given in Table S2.

2.2. PMF analysis and uncertainty assessment

PMF2 was utilized as the primary tool for source apportionment in this study. The principles and application of this model have been described previously (Paatero, 1998a, b; Jaekels et al., 2007; Shrivastava et al., 2007). Briefly, PMF assumes that the ambient pollutants measured at one receptor site are linear combinations from a number of time-variant factors/sources. It uses an uncertainty-weighted least-squares fitting approach to identify distinct factor profiles and determine factor contributions from a series of observations. Importantly, the number of factors must be chosen prior to using PMF. In this study, PMF solutions using 5-9 factors were considered, with the final number chosen based on interpretability as well stability across bootstrap-replicate data sets.

The bias and variability in factor profiles and contributions due to random sampling error were estimated by applying a method from Hemann et al. (2009). Generally, 1,000 replicate data sets were generated from the original data set using a stationary block bootstrap technique (Politis and Romano, 1994) and each was analyzed with PMF. Because the ordering of factors may differ across solutions on bootstrap replicate data sets (e.g., factor *i* in one solution may correspond to factor *j* in another), a Multilayer Feed Forward Neural

Networks were trained to sort and align the factor profiles of each PMF bootstrap solution to that of the original solution based on the observed data (known as the *base case*). A PMF bootstrap solution was recorded only when each factor of that solution could be uniquely matched to a base case factor. The measurement days resampled in each recorded solution were tracked to examine the bias and variability in contributions of each factor on each day, which could then be used to assess the uncertainty of the PMF model fit. In this work, all PMF results are presented using the median factor profiles and median contribution time series derived from bootstrapped solutions.

2.3. Preparation of PMF input data set

For the whole measurement campaign, only ten days with unavailable PM_{2.5} samples or multiple missing species observations were omitted for all PMF analysis. Candidate PMF species were selected from the 79 species listed in Table S1 based on the following criteria. First, the species with more than 50% missing values or observations below the detection limit were excluded from the input data set (Dutton et al., 2010b). Second, only species that had average S/N ratios greater than 3 were included (except nitrate, an important marker for secondary inorganic aerosol in winter). In addition, after bootstrap simulation, some species were also removed from subsequent analysis due to their high variability in factor profiles (e.g., xanthone, and odd *n*-alkanoic acids), which could reduce factor-matching rates. This resulted in 55 species (Table S1) with 970 daily observations being included as the primary PMF input data set. PMF analysis requires a complete data set so all missing values of individual species were replaced by the geometric mean of the remaining observations, and their accompanying uncertainties were set to four times the geometric mean. Similarly, the BDL values were set to half the detection limit, with uncertainties set at five-sixths the detection limit (Polissar et al., 1998).

PMF analysis was also performed for three temperature-stratified subsets of the original 970 observations. The three sub-data sets consisted of sampling days with daily average ambient temperature less than 10 °C ($N = 364$), between 10 °C and 20 °C ($N = 318$), and greater than 20 °C ($N = 288$), respectively. The sampling periods of these three sub-data sets were defined as cold, warm and hot. Statistics for all species quantified during each of these three periods are shown in Tables S3 – S5. PMF input species screening and the handling of missing and BDL values for each sub-data set were conducted in the same manner as for the full data set.

3. Results

3.1. PMF results for the full data set

An 8-factor solution (inorganic ion, *n*-alkane, EC/sterane, light *n*-alkane/polycyclic aromatic hydrocarbon (PAH), medium alkane/alkanoic acid, PAH, winter/methoxyphenol and summer/odd *n*-alkane factors) was chosen for the full data set as it had the most readily interpretable resulting factors and a relatively high factor matching rate of 52.1% between bootstrapped and base case solutions (Table 1). The factor matching rates for 7- and 9-factor solutions were only 19.1% and 39.5% respectively. The 5-factor solution was difficult to interpret since different source tracers were lumped in one factor (e.g., sulfate and *n*-

alkanes). The 6-factor solution could not resolve the light *n*-alkane/PAH factor and had a slightly lower factor matching rate (49.9%) than the 8-factor solution. Figure 1a, b presents the factor profiles and monthly average factor contributions for the 8-factor solution. The factor profiles have been normalized by

$$F_{kj}^* = \frac{F_{kj}}{\sum_{k=1}^p F_{kj}}, \quad (1)$$

where F_{kj}^* is the weighting of species j in factor k , relative to all other factors. The monthly average factor contributions, shown in Figure 1b, are expressed as reconstructed PM_{2.5} mass -the sum of nitrate, sulfate, EC and OC contributed by each factor. Only five observations were included for January 2003 and two for October 2005, so the average contributions of each factor for these two months are not shown. The overall average factor contributions, represented by the height of the bars in Figure 1b, agreed with the observations (PMF simulation/Observation = 0.77 – 1.04). Monthly average temperatures are overlaid in Figure 1b to compare the monthly trend of ambient temperature and factor contributions. For example, the summer/odd *n*-alkane factor exhibits an increase in contribution when the temperature rises while the PAH factor shows an opposing temperature dependence. The variability of factor profiles and contributions are shown using one standard deviation of all bootstrapped solutions in Figures S1 and S2. The medium alkane/alkanoic acid and winter/methoxyphenol factors exhibit the highest variability in daily factor contributions, as suggested by their highest CVs (median) of factor contributions in Table S6.

The factor contribution time series from the full data set were divided into three periods (cold, warm and hot as defined in Section 2.3.). In each period, the contribution of each factor was averaged and expressed in percentage of reconstructed PM_{2.5} mass (Figure 2a-c). In cold and hot periods, respectively, the reconstructed PM_{2.5} mass was dominated by contributions from the inorganic ion (43.7%) and summer/odd *n*-alkane (53.1%) factors. These two factors had comparable relative contributions in the warm period (26.5% and 27.1%). Moreover, the contributions from each factor are also summarized by both day of the week and month of the year in box plots (Figure S3), so as to illustrate more detailed temporal patterns and assist in the interpretation of each factor as explained in the Discussion section.

3.2. PMF results for temperature-stratified data sets

Statistics of PMF results computed separately for the three temperature-stratified sub-data sets are listed in Table 1. Compared with the full data set, fewer species were included for PMF analysis based on the same screening criteria (Tables S3-S5). In order to obtain physically meaningful factors with high factor matching rates, 7 PMF factors were selected to model the cold and warm sub-data sets with factor matching rates of 77.9% and 69.3%; a 6-factor solution was chosen for the hot sub-data set with a factor matching rate of 66.9%. Figure 3a-c shows the relatively weighted factor profiles for each sub-data set solution. The average factor contributions were visualized in percentages of reconstructed PM_{2.5} mass (Figure 2d-f). The sampling variability of factor contributions derived from bootstrapped PMF solutions for each sub-data set is assessed by using CV values in Table S6. The factors

with low factor contributions to PM_{2.5} have high sampling variability in daily factor contributions, as illustrated by the winter/methoxyphenol factor for the cold (CV = 0.48) and warm (CV = 1.05) period solutions. The factor profiles resolved by each sub-data set could match those derived from the full data set; and the weekly patterns of factor contributions for each sub-data set solution are similar to the corresponding factors derived from the full data set (Figure S4 – S6). The sum of factor contributions is also close to that for the full data set solution in each period (Figure 2). However, the distribution of factor contributions to the reconstructed PM_{2.5} mass for the warm and hot period solutions is considerably different from that for the full data set solution.

4. Discussion

4.1. Factor interpretation for full data set solution

To link each PMF factor resolved by the full data set to specific source types, we used the characteristic species with large fractions in the factor profiles (Figure 1a), the temporal trends of factor contributions (Figure S3), and explored the relationship with other environmental parameters by regressing the factor contribution time series to meteorological (temperature, solar irradiance and relative humidity) and trace gas measurements (O₃, NO_x and CO). The log values of factor contributions were used for the regression with temperature and solar irradiance, because the gas-particle partitioning, thermal decomposition and photochemical reactions of ambient pollutants are logarithmically related to temperature and solar irradiance (Yamasaki et al., 1982; Seinfeld and Pandis, 1998; Finlayson-Pitts and Pitts Jr., 2000). The influence of ambient temperature was decreased by separating the factor contributions into three different periods (cold, warm and hot). Table 2 contains Pearson correlation coefficients (*r*) of factor contributions versus all six variables for the full and temperature-stratified time series.

4.1.1. Inorganic ion factor—This factor contained the largest fractions of secondary nitrate and sulfate (Figure 1a), and dominated the reconstructed PM_{2.5} mass on cold days (43.7%, Figure 2a). Low temperature and high relative humidity (RH) are favorable ambient conditions for nitrate formation (Stelson and Seinfeld, 1982; Kim and Hopke, 2008). In addition, Kadowaki (1986) found a good correlation (*r* = 0.64) between sulfate conversion ratio and RH, and assumed that droplet-phase reactions were important for sulfate formation. Similarly, in this study, a more significant correlation was observed between the contribution of the inorganic ion factor and RH than with the other meteorological parameters (Table 2). Thus, the high concentrations of sulfate in cold periods might be ascribed to short term temperature inversions (Neff, 1997) and high RH (average 56.7%), while the enhancement in sulfate concentration observed during the hot period likely resulted from increased photochemical reactions (Zorka, 1986).

4.1.2. *n*-Alkane factor—This factor was characterized by the highest percentages of high molecular weight (HMW) *n*-alkanes (trtriacontane – nonatriacontane), and also contained around one-fourth to half of the tetracosane to hentriacontane (Figure 1a). Such an *n*-alkane pattern is similar to that of tire wear debris in road dust (Rogge et al., 1993b). Relatively strong correlations with NO_x and CO further support the association of this factor with

motor vehicle traffic (Table 2). However, the traffic count in Denver was reported to decrease on weekends, which should result in less re-suspended road dust (Dutton et al., 2010a); in contrast, this factor shows higher contributions on weekends (Figure S3b). In addition, the factor contributions were highest in cold months with an enhancement in May and June of the year (Figure S3b). Therefore, this factor might contain influences from other PM_{2.5} contributors besides the road dust, such as meat cooking or primary biogenic emissions.

4.1.3. EC/sterane factor—This factor is characterized by having a substantial fraction of the EC (20%) and more than 60% of all cholestanes (e.g., 20R-abb & 20S-aaa-cholestane; $n = 3$) and hopanes (e.g., a-22,29,30-trisnorhopane; $n = 7$) (Figure 1a). Such a factor was also observed in a previous DASH study (Dutton et al., 2010b), and could be associated with the combustion of lubricating oil from motor vehicles. This inference is supported by the drop in factor contribution on weekend (Figure S3c) and the correlations between the factor contributions and NO_x and CO concentrations in each period (Table 2). The relative contribution of this factor to reconstructed PM_{2.5} mass in cold periods was more than twice that in hot periods (Figure 2a, c). This difference was most likely caused by reduced wintertime atmospheric mixing rather than enhanced summertime oxidation, because no significant correlation existed between the factor contribution and the ozone concentration in hot periods (Table 2).

4.1.4. Light *n*-alkane/PAH factor—This factor consisted mainly of low molecular weight (LMW) *n*-alkanes (docosane, tricosane) and PAHs (fluoranthene, pyrene), 3-ring oxy-PAHs (1,8-naphthalic anhydride, anthracene-9,10-dione) and light *n*-alkanoic acid (dodecanoic acid) (Figure 1a). Figure S3d presents the seasonality of this factor contribution with a maximum in late summer and no weekly trend. In Table 2, this factor was anti-correlated with RH in each period ($r = -0.46 - -0.31$), and moderately anti-correlated with ozone concentration during hot periods only ($r = -0.52$). The two oxy-PAHs in this factor, each having two ketones, presented the highest concentrations in hot periods, in contrast to parent PAHs and benz[de]anthracene-7-one (one ketone oxy-PAH) with maximum concentrations in winter, indicating an influence of atmospheric oxidation on the oxy-PAHs (Walgraeve et al., 2010). Moreover, increased RH was found to have a negative effect on gas-to-particle partitioning of semivolatile organics (Pankow et al., 1993). As a result, we inferred that this factor might be associated with the evaporation of unburned petroleum fuels subject to atmospheric processing. Schauer et al. (1999) found that unburned diesel fuel and motor oil are major components of emissions from diesel engines, and enriched with LMW organics. The low factor contribution in mid-summer could be explained as more destruction in hot periods with the highest ozone concentration and high temperatures keeping organic compounds in the gas phase; while the low factor contribution in winter was likely due to less influence from atmospheric conversion.

4.1.5. Medium alkane/alkanoic acid factor—This factor contained mainly *n*-alkanoic acids, medium PAHs (MW = 230, 252) and a series of *n*-alkanes (hexacosane – hentriacontane) (Figure 1a). Similar to the *n*-alkane factor, it was most prominent in winter and exhibited a weekend enhancement (Figure S3e). The factor contribution was positively

correlated with NO_x and CO concentrations, but anti-correlated with ozone (Table 2). These results suggested an origin of primary sources (e.g., motor vehicle emission, cooking) with a seasonality opposing that of ozone. However, this factor only contributed a tiny part (0.41% – 1.10%) of the reconstructed PM_{2.5} mass (Figure 2a-c). Therefore, it might be linked with un-apportioned molecular markers from primary sources.

4.1.6. PAH factor—This factor was characterized by the presence of EC and the highest loadings of all PAHs except retene (Figure 1a). The temporal pattern of factor contributions presented in Figure S3f – wintertime peak and weekend decrease – was similar to that of the EC/sterane factor. In urban areas, the majority of PAHs could be attributed to the influence of domestic heating, mobile emission and industrial activities (Schauer et al., 1996; Ravindra et al., 2008; Li et al., 2009). For the current study, the correlation of factor contribution versus NO_x and CO concentrations in each period (Table 2) supported a preferential influence from motor vehicle emissions. Many studies have reported that gasoline-powered vehicles emitted more HMW PAHs and less LMW PAHs than diesel-powered vehicles (Rogge et al., 1993a; Miguel et al., 1998). However, the PAH patterns emitted by different vehicles were also dependent on driving conditions (e.g., speed, driving mode, start temperature) (Devos et al., 2005; Riddle et al., 2007; Bergvall and Westerholm, 2009). In this work, most of the PAHs included for PMF analysis are prominent in one factor. As such, this factor appears to link to all motor vehicle emissions regardless of fuel.

4.1.7. Winter/methoxyphenol factor—This factor was characterized by the highest loadings of retene and all four methoxyphenols (Figure 1a). These species are typical markers for wood burning (McDonald et al., 2000; Bari et al., 2009). The factor contribution showed wintertime dominance (Figure S3g) and a weekend enhancement, and was significantly anti-correlated with ambient temperature (Table 2), consistent with residential wood-burning stove use patterns.

4.1.8. Summer/odd n-alkane factor—This factor contains mainly nonacosane (C₂₉) and hentriacontane (C₃₁), two characteristic markers from high plant waxes (Rogge et al., 1993c), and the largest fraction of OC (Figure 1a). In hot periods, the factor contribution dominated the reconstructed PM_{2.5} mass (53.1%) (Figure 2c) and correlated with ambient temperature ($r = 0.35$), irradiance ($r = 0.34$) and ozone concentration ($r = 0.44$). In contrast, during the cold periods, no significant correlation could be observed and only a small part of reconstructed PM_{2.5} (5.8%) was apportioned to this factor. These observations indicated a summertime biogenic emission, followed by photochemical reactions. The minor contribution in cold periods was likely caused by the dormant living biomass, low temperatures reducing gas-phase emissions via evaporation, low solar irradiance and reduced photochemistry. We cannot rule out anthropogenic impacts for this factor as we do not necessarily have the most effective tracer species to observe anthropogenically emitted semi-volatiles and their photochemical reaction products.

4.1.9. Comparison to previous DASH source apportionment—In the previous source apportionment study using a 1-year (2003) data set (Dutton et al., 2010b), 7 factors were identified and each can be matched with one factor from the current work due to their

similar factor profiles. The remaining factor observed here is the medium alkane/alkanoic acid factor, which has a relatively small contribution to PM_{2.5} (0.54% for 2003). Table 3 shows results of regressing daily factor contributions from the 1-year study to those of the current study for the same time series (January 27, 2003 – December 31, 2003). Strong correlations ($r = 0.89 - 0.98$) were observed for 6 out of the 7 common factors, indicating that these factors were related to similar source classes in both studies. However, biases in contributions existed for several of the factors, as shown by the slopes of the linear regressions (Table 3). The comparison yields one weak correlation ($r = 0.36$) between the light *n*-alkane/PAH and middle oxy-PAH/alkanoic acid factors. These differences could be ascribed to the fact that PMF fitted a longer time series of PM_{2.5} speciation data in the current work and different input data can change the output (Zhang et al., 2009). In addition, source patterns of PM_{2.5} might change over time, particularly for sources like motor vehicles where the fleet is evolving to have improved combustion and better maintenance.

4.2. Comparison to source apportionment of temperature-stratified sub-data sets

The factors resolved from the analysis of each temperature-stratified sub-data set were matched to those from the full data set based on factor profiles (Figures 1 and 3) and weekly patterns of factor contributions (Figures S3-S6). The factor contributions of each sub-data set solution were linearly regressed to the corresponding PMF factors of the full data set solution (Table 4), so as to evaluate the dependence of PMF results on temperature.

4.2.1. Cold period—For the cold period, the distribution of factor contributions, expressed as reconstructed PM_{2.5} mass, for the sub-data set was similar to that for the full data set during the same period (Figure 2a, d), even though the analysis of this sub-data set yielded one less factor. This similarity is likely due to the negligible wintertime contribution from the summer/odd *n*-alkane factor from the full analysis. In Table 4, obvious biases existed for the light *n*-alkane/PAH, medium alkane/alkanoic acid and winter/methoxyphenol factors, since their regression slopes were far from unity. The latter two factors had small contributions to reconstructed PM_{2.5} mass (Figure 2a, d) and good correlations ($r = 0.76, 0.95$) between the full and sub-data set solutions (Table 4). So these two factors should represent similar source classes in the sub-data set solution as those in the full data set solution. Weaker correlation ($r = 0.41$) was observed between the light *n*-alkane/PAH factors (Table 4). This occurs because PMF assumes constant source profiles over time. But in the cold period, the influence of atmospheric conversion should be decreased, which could be seen with the lower irradiance and ozone concentration (Table S1), and the gas-to-particle partitioning should be enhanced by the lower temperatures. Figure 3a showed that the light *n*-alkane/PAH factor contained much higher fractions of light PAHs (fluoranthene and pyrene) and lower percentiles of light *n*-alkanes (docosane and tricosane) and alkanoic acid (dodecanoic acid) in the sub-data set solution than in the full data set solution. Therefore, this factor might be mostly influenced by primary emissions in cold periods; as mentioned above this factor is likely associated with unburned petroleum emissions.

4.2.2. Warm period—High correlations ($r = 0.76 - 0.99$) were observed between PMF factors resolved for the warm period sub-data set and their corresponding factors in the full data set solution (Table 4). However, Figures 2b and 2e show that the inorganic ion, *n*-

alkane, and medium alkane/alkanoic acid factors had larger relative contributions to reconstructed $PM_{2.5}$ mass in the sub-data set solution than in the full data set solution, which is also apparent from their regression slopes and intercepts in Table 4. These differences could be attributed to the omission of the summertime/odd *n*-alkane factor in the sub-data set solution. The bulk species associated with that unresolved factor were re-distributed to other factors. We can see that the inorganic ion, *n*-alkane, and medium alkane/alkanoic acid factors all contained larger fractions of the EC and OC in the sub-data set solution (Figure 3b) than in the full data set solution (Figure 1a). Compared to the sub-data set solution for the cold periods, the light *n*-alkane factor of the warm period had higher fractions of light *n*-alkanes and alkanolic acid (Figure 3a, b). This change in factor profile might be caused by more evaporative emission of light compounds followed by gas-to-particle partitioning and photochemistry processes in the warm periods.

4.2.3. Hot period—Due to the species selection constraints mentioned previously, methoxyphenols and *n*-alkanoic acids were not included for PMF analysis of this period, resulting in the loss of winter/methoxyphenol and medium alkane/alkanoic acid factors for this sub-data set analysis. As shown in Figure 2c, these two factors had relatively minor contributions to $PM_{2.5}$ during the hot period of the full data set analysis, so their loss should not significantly affect the distribution of reconstructed $PM_{2.5}$ mass. The medium *n*-alkane/OC factor was a new factor resolved for only this period. It is likely related to the summer/odd *n*-alkane factor of the full data set solution, because this factor contained a large fraction of OC and medium *n*-alkanes exhibiting an odd carbon number preference (Figure 3c). However, the relative contribution of the medium *n*-alkane/OC factor to reconstructed $PM_{2.5}$ (13.3%) was much lower than that of the summer/odd *n*-alkane factor from the full data set solution, which dominated the reconstructed $PM_{2.5}$ mass with a 53.1% contribution during the hot periods (Figure 2c, f). One plausible explanation is that separation of the summer/odd *n*-alkane factor, representing summertime biogenic emission and SOA formation, is mainly dependent on the seasonality of those species with a summertime maximum (e.g., OC, C_{29} and C_{31}), when the temperature was also the highest. With temperature-stratified input data, PMF could not distinguish the factor dominated by summertime emissions due to the lack of specific molecular markers, and redistributed the associated bulk $PM_{2.5}$ contents (sulfate, EC and OC). For example, EC and OC were mostly accounted for by the sulfate and EC/sterane factors of the PMF solution for hot periods. From the regression parameters in Table 4, all factors of the sub-data set solution for hot periods differed from those of the full data set solution in their factor contributions. As a result, the seasonality of species concentrations might assist in the separation of summertime biogenic/SOA emission from anthropogenic and secondary inorganic ion sources when unique tracers for those sources and processes are not available.

5. Conclusions

This study leveraged a longer time series (32 months) of daily $PM_{2.5}$ speciation data to investigate the impact of temperature stratification on PMF source apportionment results. For the full data set, an 8-factor solution was selected based on the most interpretable factors and the highest factor matching rate between bootstrapped PMF solutions and the base case solution. Seven out of the 8 factors are likely associated with secondary inorganic ions, tire

wear debris, lubricating oil combustion, diesel vehicle emission/atmospheric processing, motor vehicle emission, wood burning, and biogenic emission/atmospheric processing. The remaining factor (medium alkane/alkanoic acid) was likely a residual factor containing un-apportioned OMMs from primary emissions. The comparisons of source apportionment results between the current 32-month data set and a previously available 1-year data set indicate that similar source-related factors were observed, except the factor characterized by oxy-PAHs and light alkanoic acids. Temperature-stratified PMF source apportionment was performed by dividing the full data set into three sub-data sets, representing cold, warm and hot periods, respectively. The source contributions were not necessarily consistent between the full data set and the temperature-stratified sub-data sets, especially for those sources subject to seasonal atmospheric processing. As a result, the influences of atmospheric processes on source apportionment of long time-series of PM_{2.5} speciation data could be decreased by using temperature-stratified data sets, and needs to be considered in epidemiological studies of the health effects of chronic exposure to source contributions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work has supported by NIEHS research grant RO1 ES010197. We would thank Sheila Burns, Pat McGraw and Bradley Rink (CDPHE) for their help with STN data retrieval, and Joni Rix (Denver Public School District) and Palmer Elementary School faculty and staff for their assistance with the sampling site. We also thank the many people that helped to develop the organics data set: John Ortega, Dan Williams, Brett Casso, Teresa Coons, Cathy Vos, and Jessica Garcia.

References

- Bari MA, Baumbach G, Kuch B, Scheffknecht G. Wood smoke as a source of particle-phase organic compounds in residential areas. *Atmospheric Environment*. 2009; 43:4722–4732.
- Bergvall C, Westerholm R. Determination of highly carcinogenic dibenzopyrene isomers in particulate emissions from two diesel- and two gasoline-fuelled light-duty vehicles. *Atmospheric Environment*. 2009; 43:3883–3890.
- Chen LWA, Watson JG, Chow JC, DuBois DW, Herschberger L. PM_{2.5} Source apportionment: Reconciling receptor models for U.S. nonurban and urban long-term networks. *Journal of the Air & Waste Management Association*. 2011; 61:1204–1217. [PubMed: 22168104]
- Devos O, Combet E, Tassel P, Paturel L. Exhaust emissions of PAHs of passenger cars. *Polycyclic Aromatic Compounds*. 2005; 26:69–78.
- Dutton SJ, Rajagopalan B, Vedal S, Hannigan MP. Temporal patterns in daily measurements of inorganic and organic speciated PM_{2.5} in Denver. *Atmospheric Environment*. 2010a; 44:987–998. [PubMed: 23486844]
- Dutton SJ, Schauer JJ, Vedal S, Hannigan MP. PM_{2.5} characterization for time series studies: Pointwise uncertainty estimation and bulk speciation methods applied in Denver. *Atmospheric Environment*. 2009a; 43:1136–1146. [PubMed: 20126292]
- Dutton SJ, Vedal S, Piedrahita R, Milford JB, Miller SL, Hannigan MP. Source apportionment using positive matrix factorization on daily measurements of inorganic and organic speciated PM_{2.5}. *Atmospheric Environment*. 2010b; 44:2731–2741. [PubMed: 22768005]
- Dutton SJ, Williams DE, Garcia JK, Vedal S, Hannigan MP. PM_{2.5} characterization for time series studies: Organic molecular marker speciation methods and observations from daily measurements in Denver. *Atmospheric Environment*. 2009b; 43:2018–2030. [PubMed: 20161318]

- Finlayson-Pitts, BJ.; Pitts, JN, Jr. Chemistry of the upper and lower atmosphere: theory experiments and applications. Academic Press; 2000. p. 219
- Hemann JG, Brinkman GL, Dutton SJ, Hannigan MP, Milford JB, Miller SL. Assessing positive matrix factorization model fit: a new method to estimate uncertainty and bias in factor contributions at the measurement time scale. *Atmospheric Chemistry and Physics*. 2009; 9:497–513.
- Jaekels JM, Bae MS, Schauer JJ. Positive matrix factorization (PMF) analysis of molecular marker measurements to quantify the sources of organic aerosols. *Environmental Science & Technology*. 2007; 41:5763–5769. [PubMed: 17874784]
- Kadowaki S. On the nature of atmospheric oxidation processes of sulfur dioxide to sulfate and of nitrogen dioxide to nitrate on the basis of diurnal variations of sulfate, nitrate, and other pollutants in an urban area. *Environmental Science & Technology*. 1986; 20:1249–1253.
- Kim E, Hopke PK. Source characterization of ambient fine particles at multiple sites in the Seattle area. *Atmospheric Environment*. 2008; 42:6047–6056.
- Li Z, Porter EN, Sjödin A, Needham LL, Lee S, Russell AG, Mulholland JA. Characterization of PM_{2.5}-bound polycyclic aromatic hydrocarbons in Atlanta--Seasonal variations at urban, suburban, and rural ambient air monitoring sites. *Atmospheric Environment*. 2009; 43:4187–4193.
- McDonald JD, Zielinska B, Fujita EM, Sagebiel JC, Chow JC, Watson JG. Fine particle and gaseous emission rates from residential wood combustion. *Environmental Science & Technology*. 2000; 34:2080–2091.
- Miguel AH, Kirchstetter TW, Harley RA, Hering SV. On-road emissions of particulate polycyclic aromatic hydrocarbons and black carbon from gasoline and diesel vehicles. *Environmental Science & Technology*. 1998; 32:450–455.
- Neff WD. The Denver Brown Cloud studies from the perspective of model assessment needs and the role of meteorology. *Journal of the Air & Waste Management Association*. 1997; 47:269–285.
- Paatero, P. User's Guide for Positive Matrix Factorization Program PMF2 and PMF3, Part 1: Tutorial. University of Helsinki; Helsinki, Finland: 1998a.
- Paatero, P. User's Guide for Positive Matrix Factorization Program PMF2 and PMF3, Part 2: Reference. University of Helsinki; Helsinki, Finland: 1998b.
- Pankow JF, Storey JME, Yamasaki H. Effects of relative humidity on gas/particle partitioning of semivolatile organic compounds to urban particulate matter. *Environmental Science & Technology*. 1993; 27:2220–2226.
- Polissar AV, Hopke PK, Paatero P. Atmospheric aerosol over Alaska - 2. Elemental composition and sources. *Journal of Geophysical Research-Atmospheres*. 1998; 103:19045–19057.
- Politis DN, Romano JP. The Stationary Bootstrap. *Journal of the American Statistical Association*. 1994; 89:1303–1313.
- Ravindra K, Sokhi R, Van Grieken R. Atmospheric polycyclic aromatic hydrocarbons: Source attribution, emission factors and regulation. *Atmospheric Environment*. 2008; 42:2895–2921.
- Riddle SG, Jakober CA, Robert MA, Cahill TM, Charles MJ, Kleeman MJ. Large PAHs detected in fine particulate matter emitted from light-duty gasoline vehicles. *Atmospheric Environment*. 2007; 41:8658–8668.
- Rogge WF, Hildemann LM, Mazurek MA, Cass GR, Simoneit BRT. Sources of fine organic aerosol . 2. Noncatalyst and catalyst-equipped automobiles and heavy-duty diesel trucks. *Environmental Science & Technology*. 1993a; 27:636–651.
- Rogge WF, Hildemann LM, Mazurek MA, Cass GR, Simoneit BRT. Sources of fine organic aerosol . 3. Road dust, tire debris, and organometallic brake lining dust-roads as sources and sinks. *Environmental Science & Technology*. 1993b; 27:1892–1904.
- Rogge WF, Hildemann LM, Mazurek MA, Cass GR, Simoneit BRT. Sources of fine organic aerosol . 4. Particulate abrasion products from leaf surfaces of urban plants. *Environmental Science & Technology*. 1993c; 27:2700–2711.
- Schauer JJ, Kleeman MJ, Cass GR, Simoneit BRT. Measurement of emissions from air pollution sources. 2. C-1 through C-30 organic compounds from medium duty diesel trucks. *Environmental Science & Technology*. 1999; 33:1578–1587.

- Schauer JJ, Rogge WF, Hildemann LM, Mazurek MA, Cass GR. Source apportionment of airborne particulate matter using organic compounds as tracers. *Atmospheric Environment*. 1996; 30:3837–3855.
- Seinfeld J, Pandis S. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley-Interscience. 1998:149–153.
- Shrivastava MK, Subramanian R, Rogge WF, Robinson AL. Sources of organic aerosol: Positive matrix factorization of molecular marker data and comparison of results from different source apportionment models. *Atmospheric Environment*. 2007; 41:9353–9369.
- Stelson AW, Seinfeld JH. Relative humidity and temperature dependence of the ammonium nitrate dissociation constant. *Atmospheric Environment (1967)*. 1982; 16:983–992.
- Vedal S, Hannigan MP, Dutton SJ, Miller SL, Milford JB, Rabinovitch N, Kim SY, Sheppard L. The Denver Aerosol Sources and Health (DASH) study: Overview and early findings. *Atmospheric Environment*. 2009; 43:1666–1673. [PubMed: 22723735]
- Walgraeve C, Demeestere K, Dewulf J, Zimmermann R, Van Langenhove H. Oxygenated polycyclic aromatic hydrocarbons in atmospheric particulate matter: Molecular characterization and occurrence. *Atmospheric Environment*. 2010; 44:1831–1846.
- Yamasaki H, Kuwata K, Miyamoto H. Effects of ambient temperature on aspects of airborne polycyclic aromatic hydrocarbons. *Environmental Science & Technology*. 1982; 16:189–194.
- Zhang YX, Sheesley RJ, Bae MS, Schauer JJ. Sensitivity of a molecular marker based positive matrix factorization model to the number of receptor observations. *Atmospheric Environment*. 2009; 43:4951–4958.
- Zorka BV. Increase of particulate sulphate content in the atmosphere of Belgrade in the summer. *Atmospheric Environment (1967)*. 1986; 20:985–988.

High lights

- >Positive Matrix Factorization (PMF) was applied to a 32-month series of PM_{2.5} data.
- >Temperature stratified subsets of the data were also analyzed by using PMF.
- >Environmental parameters were used to help the identification of resolved factors.

Fig.1. Results of PMF analysis for full data set: (a) distributions of species across the 8 factors; (b) monthly average factor contribution to reconstructed PM_{2.5} mass overlaid with observed values and ambient temperature. Median factor profiles and contributions were used from bootstrapped PMF solutions.

Fig.2.

Comparison of relative contribution to reconstructed $PM_{2.5}$ mass between full data set solution in (a) cold, (b) warm, (c) hot periods and the corresponded sub-data set solutions (d-f).

Fig.3. PMF factor profiles of each temperature-stratified sub-data set solution, (a) cold, (b) warm, (c) hot.

Table 1
PMF simulation statistics for different data sets

Parameters	Data sets			
	Full	Cold	Warm	Hot
No. of species	55	50	46	35
No. of samples	970	364	318	288
No. of factors	8	7	7	6
No. of bootstrap replicate data sets	1000	1000	1000	1000
No. of data sets for which PMF did not converge to a solution	0	0	0	0
No. of data sets for which factors were uniquely matched	521	779	693	669

Table 2
Correlations between factor contributions and meteorological and trace gas measurements

Factors	Temperature (°C)	Radiance (KW-hr m ⁻²)	RH (%)	Ozone (ppm)	NO _x (ppm)	CO (ppm)
Full						
Inorganic ion	-0.32	-0.17	0.55	-0.33	0.23	0.18
<i>n</i> -Alkane	-0.23	-0.17	0.07*	-0.34	0.54	0.48
EC/sterane	-0.22	-0.22	0.16	-0.44	0.64	0.59
Light <i>n</i> -alkane/PAH	0.30	0.28	-0.38	0.11	-0.15	-0.08
Medium alkane/ alkanolic acid	-0.28	-0.30	0.19	-0.49	0.39	0.38
PAH	-0.47	-0.37	0.11	-0.60	0.79	0.74
Winter/methoxyphenol	-0.65	-0.49	0.18	-0.49	0.36	0.28
Summer/odd <i>n</i> -alkane	0.54	0.43	-0.28	0.58	-0.32	-0.19
Cold						
Inorganic ion	-0.33	0.04*	0.51	-0.19	0.07*	0.05*
<i>n</i> -Alkane	-0.15	-0.08*	-0.11*	-0.35	0.52	0.49
EC/sterane	-0.08*	-0.21	0.00*	-0.49	0.59	0.56
Light <i>n</i> -alkane/PAH	0.23	0.16	-0.18	0.06*	-0.05*	0.03*
Medium alkane/ alkanolic acid	0.17	-0.08*	0.02*	-0.39	0.29	0.30
PAH	-0.12*	-0.05*	-0.21	-0.57	0.76	0.76
Winter/methoxyphenol	-0.14	-0.15	-0.13*	-0.27	0.15	0.14
Summer/odd <i>n</i> -alkane	0.07*	0.08*	0.03*	0.07*	-0.14	-0.12*
Warm						
Inorganic ion	-0.03*	0.05*	0.44	-0.04*	0.02*	0.08*
<i>n</i> -Alkane	-0.03*	0.04*	-0.10*	-0.04*	0.28	0.22
EC/sterane	-0.16	-0.07*	0.02*	-0.20	0.60	0.55
Light <i>n</i> -alkane/PAH	-0.01*	0.24	-0.32	0.07*	0.05*	0.05*
Medium alkane/ alkanolic acid	-0.24	-0.20	0.14*	-0.42	0.43	0.43
PAH	-0.23	-0.08*	-0.17	-0.37	0.71	0.68
Winter/methoxyphenol	-0.38	-0.17	-0.16	-0.28	0.16	0.12*

Factors	Temperature (°C)	Radiance (KW-hr m ⁻²)	RH (%)	Ozone (ppm)	NO _x (ppm)	CO (ppm)
Summer/odd <i>n</i> -alkane	0.35	0.17	0.31	0.12 *	-0.11 *	-0.01 *
Hot						
Inorganic ion	-0.15	-0.04 *	0.36	0.13 *	0.09 *	0.06 *
<i>n</i> -Alkane	0.02 *	0.00 *	-0.08 *	0.03 *	0.38	0.25
EC/sterane	0.05 *	-0.02 *	0.06 *	0.04 *	0.35	0.34
Light <i>n</i> -alkane/PAH	0.02 *	-0.08 *	-0.37	-0.52	0.12 *	0.01 *
Medium alkane/ alkanolic acid	-0.10 *	-0.13 *	-0.04 *	-0.56	0.22	0.15
PAH	-0.26	-0.16	-0.13 *	-0.22	0.40	0.27
Winter/methoxyphenol	-0.16	0.00 *	-0.07 *	0.01 *	0.03 *	-0.10 *
Summer/odd <i>n</i> -alkane	0.35	0.34 *	-0.13 *	0.44	-0.19	0.02 *

* Correlation is not significant at the 0.01 level (2-tailed); absolute *r* values equal or higher than 0.40 are in bold.

Table 3
PMF factor contribution regression statistics for 2003 derived from the 32-month data set versus a 1-year data set

Factor		Regression statistics		
Current ^a	Previous ^b	Slope	Intercept	<i>r</i>
Inorganic ion	Inorganic ion	0.92	-92.4	0.95
<i>n</i> -Alkane	<i>n</i> -Alkane	1.71	183	0.98
EC/sterane	EC/sterane	0.46	42.0	0.98
Light <i>n</i> -alkane/PAH	Middle oxy-PAH/alkanoic acid	0.55	338	0.36
Medium alkane/alkanoic acid	/	/	/	/
PAH	Nitrate/PAH	0.88	42.8	0.97
Winter/methoxyphenol	Wintertime/methoxyphenol	1.18	-1.19	0.96
Summer/odd <i>n</i> -alkane	Summertime/selective aliphatic	0.80	-1.58	0.89
Sum ^c	Sum	0.90	253	0.96

^a Full data set solution in this study, of which the factor contributions in 2003 were regressed as dependent;

^b PMF solution of previous 1-year DASH study (Dutton et al., 2010);

^c Sum of all factor contributions except the median alkane/alkanoic acid factor for current study.

Table 4
PMF factor contribution regression statistics between full data set solution and temperature-stratified sub-data set solutions

Factor		Cold			Warm			Hot		
Full ^a	Sub ^b	Slope	Intercept	r	Slope	Intercept	r	Slope	Intercept	r
Inorganic ion	Inorganic ion/sulfate	1.06	140	0.98	0.82	853	0.76	0.85	1129	0.70
<i>n</i> -Alkane	<i>n</i> -Alkane	0.98	-16.5	0.99	1.37	86.7	0.97	1.60	181	0.88
EC/sterane	EC/sterane	0.87	-13.7	1.00	1.15	-46.8	0.99	4.80	227	0.92
Light <i>n</i> -alkae/PAH	Light <i>n</i> -alkae/PAH	0.46	-458	0.41	0.88	102	0.84	0.50	-84.3	0.91
Medium alkane/alkanoic acid	Medium alkane/alkanoic acid	3.57	119	0.76	5.59	40.1	0.96	/	/	/
PAH	PAH	1.09	-116	0.98	0.76	4.08	0.99	0.37	33.3	0.80
Winter/methoxyphenol	Winter/methoxyphenol	0.57	-16.8	0.95	1.04	-4.28	0.99	/	/	/
Summer/odd <i>n</i> -alkane	Medium <i>n</i> -alkane/OC	/	/	/	/	/	/	0.19	153	0.55
Sum ^c	Sum	1.04	-113	0.99	0.82	779	0.90	0.75	1185	0.86

^a Full data set solution, of which the factor contributions were regarded as independent variables for regression;

^b Temperature-stratified sub-data set solutions;

^c Sum of all factor contributions.