

POSITIVE MATRIX FACTORIZATION OF PM<sub>2.5</sub> – IMPACTS OF SPATIAL VARIABILITY,  
SPECIATION DATA SET AND GAS/PARTICLE PARTITIONING OF SEMI-VOLATILE  
ORGANIC COMPOUNDS

written by

MINGJIE XIE

B.S., Nanjing University, 2007

M.S., Nanjing University, 2010

A thesis submitted to the  
Faculty of the Graduate School of the  
University of Colorado in partial fulfillment  
of the requirement for the degree of  
Doctor of Philosophy  
Department of Mechanical Engineering

2013

This thesis entitled:

POSITIVE MATRIX FACTORIZATION OF PM<sub>2.5</sub> – IMPACTS OF SPATIAL VARIABILITY,  
SPECIATION DATA SET AND GAS/PARTICLE PARTITIONING OF SEMI-VOLATILE  
ORGANIC COMPOUNDS

written by

Mingjie Xie

has been approved for the Department of Mechanical Engineering

---

Dr. Michael P. Hannigan, Chair

---

Dr. Jana B. Milford, Committee Member

Date \_\_\_\_\_

The final copy of this thesis has been examined by the signatories, and we  
find that both the content and the form meet acceptable presentation standards  
of scholarly work in the above mentioned discipline.

## ABSTRACT

Xie, Mingjie (Ph.D., Mechanical Engineering)

POSITIVE MATRIX FACTORIZATION OF  $PM_{2.5}$  – IMPACTS OF SPATIAL VARIABILITY,  
SPECIATION DATA SET AND GAS/PARTICLE PARTITIONING OF SEMI-VOLATILE  
ORGANIC COMPOUNDS

Thesis directed by Professor Michael P. Hannigan

The Denver Aerosol Sources and Health (DASH) study aims to identify and quantify the sources of  $PM_{2.5}$  that are related to negative health outcomes. The positive matrix factorization (PMF), a multivariate receptor model, was used as the primary tool for source apportionment of  $PM_{2.5}$  based on particulate speciation data. However, several questions need to be addressed on the receptor-based source apportionment of  $PM_{2.5}$ .

In DASH study, 24-h  $PM_{2.5}$  samples were collected at one centrally located site in Denver. This raises the question of whether the heterogeneity in  $PM_{2.5}$  sources or source contributions across the urban area might lead to biased health effects estimation. In this work,  $PM_{2.5}$  samples were collected at four urban sites in Denver for one year. The carbonaceous speciation data were used as inputs for PMF analysis. The results showed that the four sampling sites have consistent source profiles and similar source distribution of elemental carbon (EC) and organic carbon (OC).

The speciation of  $PM_{2.5}$  in the DASH study includes inorganic ions, EC and OC, organic molecular markers (OMMs) and water soluble elements (WSEs). To evaluate the utility of different speciation data sets for source apportionment of bulk  $PM_{2.5}$  species, different

combinations of source tracers with bulk  $PM_{2.5}$  species were applied for PMF analysis. The results suggested that OMMs were better source tracers for EC and OC than WSEs.

However, OMMs are mostly semi-volatile organic compounds (SVOCs), and their particle-phase fractions are impacted by gas/particle (G/P) partitioning. In this work, a 32-month series of  $PM_{2.5}$  speciation data was available for PMF analysis. The influence of G/P partitioning was identified by the comparison of PMF analysis of the full data set versus temperature-stratified sub-data sets. With the prediction of gas-phase SVOC concentrations by an equilibrium absorption model, the PMF analysis using total SVOC (gas + particle phase) data set showed consistent results between the full data set and temperature-stratified sub sets. A 1-year field study of both gas- and particle-phase SVOCs was conducted to verify the gas-phase SVOCs prediction. The observed G/P partitioning of SVOCs was reasonably consistent with that predicted by an equilibrium absorption model.

## ACKNOWLEDGEMENTS

I would like to express the deepest appreciation to my research advisor Dr. Michael Hannigan, who designed the big picture for my research, and the rest of my Ph.D committee for helping me to achieve my goal. Special thanks go to Dr. Jana Milford, Dr. Shelly Miller, Dr. Kelley Barsanti, Dr. Jenifer Peel, Dr. Steven Dutton, Dr. James Schauer and Dr. Sverre Vedal for their time and expertise. This dissertation could not be finished without substantial contributions from previous and current group members and collaborators: Teresa Coons, Brett Casso, Ricardo Piedrahita, Joshua Hemann, John Ortega, Daniel Williams, Cathy Vos, Jessica Garcia, Toni Newville, Sheila Burns, Pat McGraw, Bradley Rink and Joni Rix. I would also like to thank the faculty and staff at Palmer Elementary School for their assistance with the sampling site. Funding for this work was provided by NIEHS research grant number RO1ES010197. Additional support for curricular practical training was provided by University of Washington. Special thanks also go to my family members and friends who always encourage me to pursue my dreams.

## TABLES OF CONTENTS

<b>ABSTRACT</b>	iii
<b>ACKNOWLEDGEMENTS</b>	v
<b>TABLES OF CONTENTS</b>	vi
<b>LIST OF TABLES</b>	xiii
<b>LIST OF FIGURES</b>	xvi
<b>CHAPTER 1 INTRODUCTION</b>	1
<b>1.1 RESEARCH MOTIVATION, CONTEXT AND GOLAS</b>	1
<b>1.2 THESIS ORGANIZATION</b>	6
<b>CHAPTER 2 INTRA-URBAN SPATIAL VARIABILITY OF PM<sub>2.5</sub>-BOUND CARBONACEOUS COMPOUNDS</b>	7
<b>2.0 ABSTRACT</b>	8
<b>2.1 INTRADUCTION</b>	10
<b>2.2 METHODS</b>	12
2.2.1 PM <sub>2.5</sub> sampling sites	12
2.2.2 Sample collection and chemical characterization	14
2.2.3 Uncertainty estimation	15
2.2.4 Statistical analysis	16
<b>2.3 RESULTS AND DISCUSSION</b>	18
2.3.1 General description of measurement data	18
2.3.1.1 Spatial samples	18
2.3.1.2 Co-located samples	22
2.3.2 Analysis of spatial variability	22

2.3.2.1 Correlation coefficient ( <i>r</i> ) and coefficient of divergence (COD)	22
2.3.2.2 Spatial variability framed by variability in co-located sample results	26
2.3.3 Spatial distribution of mobile sources influences	27
<b>2.4 CONCLUSIONS</b>	<b>30</b>
<b>CHAPTER 3 INTRA-URBAN SPATIAL VARIABILITY AND UNCERTAINTY ASSESSMENT OF PM<sub>2.5</sub> SOURCES BASED ON CARBONACEOUS SPECIES</b>	<b>32</b>
<b>3.0 ABSTRACT</b>	<b>33</b>
<b>3.1 INTRODUCTION</b>	<b>35</b>
<b>3.2 METHODS</b>	<b>38</b>
3.2.1 Ambient sample collection and chemical characterization	38
3.2.2 Positive matrix factorization (PMF2) modeling	41
3.2.3 Uncertainty assessment for PMF modeling	41
3.2.4 Analysis of spatial variability	42
<b>3.3 RESULTS AND DISCUSSION</b>	<b>43</b>
3.3.1 PMF solution based on pooled data set	43
3.3.1.1 Factor profile interpretation	47
3.3.1.2 Uncertainty assessment for factor contributions	51
3.3.2 Spatial variability of factor contributions	54
3.3.3 Comparison between PMF solutions of pooled and site specific data sets	56
<b>3.4 CONCLUSIONS</b>	<b>59</b>
<b>CHAPTER 4 POSITIVE MATRIX FACTORIZATION OF PM<sub>2.5</sub>: COMPARISON AND IMPLICATIONS OF USING DIFFERENT SPECIATION DATA SETS</b>	<b>60</b>
<b>4.0 ABSTRACT</b>	<b>61</b>

<b>4.1 INTRODUCTION</b>	62
<b>4.2 METHODS</b>	65
4.2.1 Ambient sample collection and chemical characterization	65
4.2.2 Source apportionment model and uncertainty assessment	66
4.2.3 Development of speciation subsets	67
<b>4.3 RESULTS AND DISCUSSION</b>	69
4.3.1 PMF results for different data sets	69
4.3.2 Comparison between PMF estimation and observed concentration	76
4.3.3 Comparison between PMF solutions	78
<b>4.4 IMPLICATIONS</b>	82
<b>CHAPTER 5 POSITIVE MATRIX FACTORIZATION OF A 32-MONTH SERIES OF DAILY PM<sub>2.5</sub> SPECIEATION DATA WITH INCORPORATION OF TEMPERATURE STRATIFICATION</b>	84
<b>5.0 ABSTRACT</b>	85
<b>5.1 INTRODUCTION</b>	87
<b>5.2 METHODS</b>	89
5.2.1 Ambient data collection and chemical speciation	89
5.2.2 PMF analysis and uncertainty assessment	89
5.2.3 Preparation of PMF input data set	90
<b>5.3 RESULTS</b>	92
5.3.1 PMF results for the full data set	92
5.3.2 PMF results for temperature-stratified data sets	97
<b>5.4 DISCUSSION</b>	99

5.4.1 Factor interpretation for full data set solution	99
5.4.1.1 Inorganic ion factor	101
5.4.1.2 <i>n</i> -Alkane factor	101
5.4.1.3 EC/sterane factor	102
5.4.1.4 Light <i>n</i> -alkane/PAH factor	102
5.4.1.5 Medium alkane/alkanoic acid factor	103
5.4.1.6 PAH factor	103
5.4.1.7 Winter/methoxyphenol factor	104
5.4.1.8 Summer/odd <i>n</i> -alkane factor	104
5.4.1.9 Comparison to previous DASH source apportionment	105
5.4.2 Comparison to source apportionment of temperature-stratified sub-data sets	107
5.4.2.1 Cold period	107
5.4.2.2 Warm period	108
5.4.2.3 Hot period	108
<b>5.5 CONCLUSIONS</b>	110
<b>CHAPTER 6 POSITIVE MATRIX FACTORIZATION OF PM<sub>2.5</sub> – ELIMINATING THE EFFECTS OF GAS/PARTICLE PARTITIONING OF SEMI-VOLATILE ORGANIC COMPOUNDS</b>	111
<b>6.0 ABSTRACT</b>	112
<b>6.1 INTRODUCTION</b>	114
<b>6.2 METHODS</b>	117
6.2.1 Particle phase measurements	117
6.2.2 Gas phase concentration and uncertainty estimation	117

6.2.3 PMF analysis and uncertainty assessment	119
6.2.4 Preparation of PMF input data set	120
<b>6.3 RESULTS AND DISCUSSION</b>	122
6.3.1 Total SVOCs and their particle-phase fractions	122
6.3.2 Sensitivity of total SVOC estimation based on G/P partitioning theory	124
6.3.3 PMF results for the full data set	127
6.3.4 Comparison to particle only-based source apportionment	130
6.3.5 PMF results for temperature-stratified sub-data sets	133
6.3.6 Comparison to PMF results of the full data set	134
6.3.6.1 Cold period	135
6.3.6.2 Warm period	137
6.3.6.3 Hot period	137
<b>6.4 CONCLUSIONS</b>	139
<b>CHAPTER 7 CHARACTERIZATION OF <i>n</i>-ALKANES, PAHs, 2-METHYLTETROLS AND LEVOGLUCOSAN IN GAS PHASE AND PM<sub>2.5</sub> USING A MEDIUM VOLUME SAMPLER</b>	140
<b>7.0 ABSTRACT</b>	141
<b>7.1 INTRODUCTION</b>	142
<b>7.2 METHODS</b>	145
7.2.1 Sampling	145
7.2.2 Sample extraction and instrumental analysis	147
7.2.2.1 Carbon analysis	147
7.2.2.2 Non-polar SVOC analysis	148

7.2.2.3 Polar SVOCs analysis	149
7.2.2.4 Quality assurance and control	151
<b>7.3 RESULTS AND DISCUSSION</b>	<b>153</b>
7.3.1 OC and EC concentrations of QFF samples	153
7.3.2 Breakthrough experiments	154
7.3.3 Concentrations of target SVOCs on QFFs and adsorbents	158
7.3.3.1 <i>n</i> -Alkanes and PAHs	158
7.3.3.2 2-Methyltetrols and levoglucosan	160
<b>7.4 CONCLUSIONS</b>	<b>163</b>
<b>CHAPTER 8 GAS/PARTICLE PARTITIONING OF SEMI-VOLATILE ORGANIC COMPOUNDS IN DENVER-FITNESS OF ABSORPTIVE PARTITIONING THEORY</b>	<b>164</b>
<b>8.0 ABSTRACT</b>	<b>165</b>
<b>8.1 INTRODUCTION</b>	<b>167</b>
<b>8.2 METHODS</b>	<b>171</b>
8.2.1 Sampling, chemical analysis and SVOC selection	171
8.2.2 Artifact corrections using bQFF measurements	171
8.2.3 Calculation of $K_{p,OM}$	173
<b>8.3 RESULTS AND DISCUSSION</b>	<b>175</b>
8.3.1 Regression of $\log K_{p,OM}^m$ vs. $\log p_L^\circ$	175
8.3.2 Log $K_{p,OM}^m$ vs. $\log K_{p,OM}^t$	181
<b>8.4 CONCLUSIONS</b>	<b>187</b>
<b>CHAPTER 9 SUMMARY AND FUTURE RESEARCH</b>	<b>188</b>

<b>9.1 STUDY SUMMARY</b>	188
<b>9.2 RESULTS SUMMARY</b>	191
<b>9.3 FURTHER RESEARCH</b>	197
<b>CHAPTER 10 BIBLIOGRAPHY</b>	200
<b>CHAPTER 11 SUPPLEMENTAL INFORMATION</b>	225
<b>11.1 SUPPLEMENTAL TABLES</b>	225
<b>11.2 SUPPLEMENTAL FIGURES</b>	267
<b>11.3 SUPPLEMENTAL DISCUSSIONS</b>	305
11.3.1 Extreme spike during Thanksgiving holiday (for CHAPTER 2)	305
11.3.2 Impact of quantification uncertainty on correlation coefficient ( $r$ ) and COD (for CHAPTER 2)	306
11.3.3 Ratio-ratio plot analysis (for CHAPTER 3)	307
11.3.4 Comparing two correlation coefficients ( $r_1$ & $r_2$ ; for CHAPTER 8)	308
11.3.5 Compare the slope of linear regression to 1.0 (for CHAPTER 8)	309

## LIST OF TABLES

Table 2.1 Meteorological statistics based on hourly observation from nearby monitoring stations during March 1, 2008 - March 14, 2009. (PAL - CAMP; EDI - Carriage; ALS - Welby; MAS - Denver Animal Shelter)	14
Table 2.2 Studies on spatial variations of PM <sub>2.5</sub> components	21
Table 3.1 Locations and characteristics of the four sampling sites in Denver.	39
Table 3.2 Average concentration ± standard deviation and signal to noise ratios of selected species for source apportionment.	40
Table 3.3 Pearson correlation coefficients ( <i>r</i> ) and coefficients of divergence (COD) for median factor contributions obtained at the four sampling sites.	55
Table 3.4 Simulation statistics for individual data sets.	57
Table 4.1 Simulation statistics for all data sets.	71
Table 4.2 Annual average source apportionment results for PM <sub>2.5</sub> components, μg m <sup>-3</sup> (%) and median coefficient of variation (CV) of factor contributions.	71
Table 5.1 PMF simulation statistics for different data sets.	93
Table 5.2 Correlations between factor contributions and meteorological and trace gas measurements.	100
Table 5.3 PMF factor contribution regression statistics for 2003 derived from the 32-month data set versus a 1-year data set.	106
Table 5.4 PMF factor contribution regression statistics between full data set solution and temperature-stratified sub-data set solutions.	106
Table 6.1 Values of parameters used to test the sensitivity of total SVOC estimation.	125
Table 6.2 PMF simulation statistics for different data sets.	128
Table 6.3 Average factor contributions to bulk components for full data set solution and sub-data set solutions (μg m <sup>-3</sup> )	129
Table 6.4 Regression statistics of factor contributions between full data set and sub-data set solutions.	135
Table 7.1 Statistics for each carbon fraction on QFF samples ( <i>N</i> =50, μg m <sup>-3</sup> )	154

Table 7.2 Gas-phase breakthrough (%) for individual SVOCs.	155
Table 8.1 Sampling artifact corrections for all selected SVOCs.	173
Table 8.2 Regression statistics for $\log K_{p,OM}^m$ vs. $\log p_L^\circ$ with different sampling artifact corrections.	176
Table 2.S1 The fraction of the compounds (%) detected in samples of one quantification batch (sample No. = 79) falling within, below and above the calibration range.	225
Table 2.S2 Statistics for PM <sub>2.5</sub> carbonaceous components at the four sampling sites (March 1, 2008-March 14, 2009).	226
Table 2.S3 Statistics for PM <sub>2.5</sub> components derived from bi-weekly pairs of co-located samples at PAL (July 13, 2004- September 20, 2005).	228
Table 2.S4 Simulated effect of quantification uncertainty on <i>r</i> and COD statistics for side-by-side and multi-site comparisons. Asterisks denote simulated values without quantification uncertainties.	230
Table 4.S1 Statistics for each species quantified during January 27 – December 31, 2003.	231
Table 5.S1 Statistics for PM <sub>2.5</sub> species quantified from January 27, 2003 – October 2, 2005 (Full).	234
Table 5.S2 Meteorological and trace gas statistics based on hourly observations for downtown Denver (January 27, 2003 – October 2, 2005).	236
Table 5.S3 Statistics for PM <sub>2.5</sub> species quantified on sampling days with ambient temperature lower than 10°C (Cold) from January 27, 2003 – October 2, 2005.	237
Table 5.S4 Statistics for PM <sub>2.5</sub> species quantified on sampling days with ambient temperature between 10°C and 20°C (Warm) from January 27, 2003 – October 2, 2005.	239
Table 5.S5 Statistics for PM <sub>2.5</sub> species quantified on sampling days with ambient temperature higher than 20°C (Hot) from January 27, 2003 – October 2, 2005.	241
Table 5.S6 Median coefficient of variation (CV) of factor contributions for bootstrapped PMF solutions for each data set.	243
Table 6.S1 Statistics for the total (gas + particle phase) concentration of each SVOC estimated from January 27, 2003 – October 2, 2005 (Full data set).	244
Table 6.S2 Statistics for the total (gas + particle phase) concentration of each SVOC estimated for sampling days with ambient temperature lower than 10 °C (Cold period).	246

Table 6.S3 Statistics for the total (gas + particle phase) concentration of each SVOC estimated for sampling days with ambient temperature between 10 °C and 20 °C (Warm period).	248
Table 6.S4 Statistics for the total (gas + particle phase) concentration of each SVOC estimated for sampling days with ambient temperature above 20 °C (Hot period).	250
Table 6.S5 Mean particle-phase fractions of selected SVOCs and data from other field studies.	252
Table 6.S6 Correlation coefficients of factor contributions from full data set solution versus meteorological and trace gas measurements.	254
Table 6.S7 Correlation coefficients of factor contributions from sub-data set solutions versus meteorological and trace gas measurements.	255
Table 7.S1 Sampling dates and sampling matrices for gas- and particle-phase SVOCs.	256
Table 7.S2 Non-polar SVOC compounds and their referenced internal standards.	258
Table 7.S3 Recoveries of meso-erythritol and levoglucosan for all sampling matrices used in this work.	260
Table 7.S4 Concentrations of polar SVOCs in different parts of selected PUF/XAD/PUF samples.	261
Table 7.S5 Statistics of <i>n</i> -alkanes and PAHs quantified from QFFs and PUF/XAD-4/PUF samples.	262
Table 8.S1 Properties of selected SVOCs for analysis.	264
Table 8.S2 Statistics for bulk carbon ( $\mu\text{g m}^{-3}$ ) and all selected SVOCs ( $\text{ng m}^{-3}$ ) concentrations in QFFs and PUF/XAD/PUF samples.	265
Table 8.S3 Statistics of $K_{p,OM}^m$ values with different approaches for artifact corrections.	266

## LIST OF FIGURES

- Figure 2.1 Locations of the four sampling sites in Denver urban area: (1) Palmer Elementary School (PAL); (2) Edison Elementary School (EDI); (3) Alsup Elementary School (ALS); (4) Municipal Animal Shelter (MAS). 13
- Figure 2.2 Pearson correlation coefficients (bars) and coefficients of divergence (light blue circles) for organic molecular markers and bulk carbon contents obtained at the 4 sampling sites (a-f) as well as the co-located samples at PAL site (g). CODs of collocated samples were averaged in groups and overlaid on each plot. Bars are shaded to show compound class separations. 25
- Figure 2.3 Ratio-ratio plots for two polycyclic aromatic hydrocarbons (PAHs) normalized by EC (a-d) and two hopanes normalized by EC (e-f) at each of the four sites. Gasoline, diesel (L light duty, H heavy duty) and wood burning (different woods and burning conditions in fireplace and woodstove) source profiles overlaid were obtained from Cadle et al. (1999) (only for light duty diesel associated PAHs in Denver) and NFRAQS study (Zielinska et al., 1998). 28
- Figure 3.1 Box plots of normalized factor profiles derived from 769 PMF solutions out of 1000 replicate data sets. The boxes depict the median (dark line), inner quartile range (blue box), lower quartile  $- 1.5 \times \text{IQR}$  and upper quartile  $+ 1.5 \times \text{IQR}$  (whisker), base case (blank star) and outliers (red plus). Values of the medians are labeled on the top of each plot. 44
- Figure 3.2 Factor contribution time series separated by site from PMF bootstrap solutions. Each plot shows the time series of factor contribution for base case (blue diamond), median (green line), and two bands based on the empirical quantiles of the bootstrap solutions ( $25^{\text{th}} - 75^{\text{th}}$  percentiles, gray dashed line;  $5^{\text{th}} - 95^{\text{th}}$  percentiles, red solid line). 46
- Figure 3.3 Comparison of median factor contributions versus uncertainties represented by  $5^{\text{th}} - 95^{\text{th}}$  percentile range of factor contributions from bootstrap solutions for each factor. 53
- Figure 3.4 Distribution of factor contributions to bulk carbon components including (a) EC, (b) OC at each site. Apportionment results are calculated on the annual average using median daily contributions of bootstrap solutions. The overall height of each bar represents the average concentration of EC or OC estimated from the model, and can be compared to the average measured concentrations represented by the triangles. 56
- Figure 4.1 Median normalized factor profiles derived from PMF bootstrap solutions for (a) Bulk, (b) WSE+Bulk and (c) OMM+Bulk data sets. The whiskers represent one standard deviation. 72
- Figure 4.2 Correlation contour plots illustrating the Pearson correlation coefficient ( $r$ ) between median factor contribution time series of each pair of PMF solutions. 79

- Figure 5.1 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. 94
- Figure 5.2 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). 96
- Figure 5.3 PMF factor profiles of each temperature-stratified sub-data set solution, (a) cold, (b) warm, (c) hot. 98
- Figure 6.1 Mean particle-phase fractions of all SVOCs during cold, warm and hot periods. 123
- Figure 6.2 Sensitivity of the calculation of total SVOC concentration ( $S$  value), determined by  $z$  value, to ambient temperature ( $T$ ), mole fraction scale activity coefficient ( $\zeta_{OM}$ ), OM/OC ratio and mean molecular weight of absorbing OM phase ( $\overline{MW}_{OM}$ ). The  $z$  value equals to the ratio of gas phase to particle phase SVOC. 126
- Figure 6.3 Linear regressions of matched pairs of factors between particle only-based and total SVOC-based PMF analysis. 131
- Figure 6.4 Median factor contribution time series (blank circle) of (a) light SVOC factor from the total SVOC-based solution, and (b) light  $n$ -alkane/PAH factor from the particle only-based solution. The red line represents the timeseries of daily average temperature. 131
- Figure 6.5 Linear regressions of factor contributions between the full data set and sub-data set solutions, (a-c) light  $n$ -alkane/PAH factor from particle only-based analysis; (d-f) light SVOC factor from total SVOC-based analysis. 136
- Figure 7.1 Scheme of the sampler used to collect SVOCs in gaseous and particle phases. 146
- Figure 7.2 Comparisons of median concentrations of (a)  $n$ -alkanes ( $MW = 170 - 324$ ), (b) PAHs ( $MW = 128 - 202$ ) across tQFFs, bQFFs and PUF/XAD-4/PUF samples. 159
- Figure 7.3 Concentration time series of 2-methyltetrols and levoglucosan in tQFFs, bQFFs and PUF/XAD/PUF samples. 162
- Figure 8.1 Regressions of (a)  $F/M_{OM}$  vs.  $A$  for 2-methyltetrols and (b)  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  for levoglucosan. 179
- Figure 8.2 Regressions of  $\log K_{p,OM}^m$  vs.  $\log K_{p,OM}^t$  with artifact corrections C1 and C3 for (a) docosane, (b) tricosane, (c) fluoranthene, (d) pyrene, (e) acenaphthenone, (f) fluorenone. 183

- Figure 8.3 Regressions of  $\log K_{p,OM}^m$  vs.  $\log K_{p,OM}^t$  for (a) 2-methyltetrols and (b) levoglucosan. 185
- Figure 2.S1 Calibration curves for selected species from one of the three quantification batches. 267
- Figure 2.S2 Concentrations of different groups of PM<sub>2.5</sub> species on November 25<sup>th</sup> 2008 (black triangles) vs. annual averages (bar plots). 268
- Figure 2.S3 Hourly observations for (a) CO, (b) NO<sub>x</sub>, (c) CO/NO<sub>x</sub> ratios, (d) wind speed, (e) ambient temperature on November 25, 2008 (red diamond) vs. annual distribution data (box plots). The boxes depict the median (black line), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and mean (light blue circle). 269
- Figure 3.S1 Correlations of median daily contribution of Factor 2 versus ambient temperature at the four sites. 270
- Figure 3.S2 Ratio-ratio plots for two low molecular weight (LMW) PAHs (a-d) and two medium MW PAHs (e-h) normalized by EC at each of the four sites. Black and blank triangles represent gasoline and diesel (L light duty, H heavy duty) source profiles respectively. Blank diamonds represent wood burning source profiles. Source profile data are obtained from Zielinska et al. (1998). 271
- Figure 3.S3 Comparison of median factor contributions versus uncertainties represented by 5<sup>th</sup> – 95<sup>th</sup> percentile range of factor contribution from bootstrap solutions for each factor, separated by site. 272
- Figure 3.S4 Comparison of normalized factor profiles from PMF bootstrap solutions using site specific data sets versus those obtained from pooled data set (TOT). The boxes depict the median (dark line), inner quartile range (blue box), lower quartile – 1.5 × IQR and upper quartile + 1.5 × IQR (whisker), base case (blank star) and outliers (red plus). Values of the medians are labeled on the top of each plot. 273
- Figure 3.S5 Factor contribution time series from PMF bootstrap solutions for site specific data. Each plot shows the time series of daily factor contribution for base case (blue diamond), median (green line), and two bands based on the empirical quantiles of the bootstrap solutions (25<sup>th</sup> – 75<sup>th</sup> percentiles, gray dashed line; 5<sup>th</sup> – 95<sup>th</sup> percentiles, red solid line). Correlation coefficient (*r*) and coefficient of divergence (COD) listed show the relationship between median factor contributions derived from site specific data sets and those from pooled data set. 280
- Figure 4.S1 Median normalized factor profiles derived from PMF bootstrap solutions using Composite data set including all species. The whiskers represent standard deviation. 281
- Figure 4.S2 Median PMF factor contributions for the Bulk data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions (+/- 1 SD). 283

- Figure 4.S3 Median PMF factor contributions for the WSE+Bulk data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD). 284
- Figure 4.S4 Median PMF factor contributions for the OMM+Bulk data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD). 285
- Figure 4.S5 Median PMF factor contributions for the Composite data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD). 286
- Figure 4.S6 Day of the week box plots for factor contributions from the Bulk data set. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle). 287
- Figure 4.S7 Day of the week box plots for factor contributions from the WSE+Bulk data set. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle). 288
- Figure 4.S8 Day of the week box plots for factor contributions from the OMM+Bulk data set. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle). 289
- Figure 4.S9 Comparisons of median factor contribution time series between factors resolved by the Composite data set and the OMM+Bulk and WSE+Bulk data sets. 290
- Figure 4.S10 Comparisons of PMF estimated vs. observed time series of PM<sub>2.5</sub> components. 291
- Figure 5.S1 Median PMF factor profiles for the full data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions ( $\pm 1$  SD). The bars are shaded to show compound class separations. 292
- Figure 5.S2 Median PMF factor contributions for the full data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD). 293
- Figure 5.S3 Day of the week and seasonal box plots for PMF factor contributions of the full data set solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle). 294
- Figure 5.S4 Day of the week box plots for PMF factor contributions of the cold period solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle). 296

- Figure 5.S5 Day of the week box plots for PMF factor contributions of the warm period solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle). 297
- Figure 5.S6 Day of the week box plots for PMF factor contributions of the hot period solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle). 298
- Figure 6.S1 Median PMF factor profiles for the full data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations. 299
- Figure 6.S2 Median PMF factor contributions for the full data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions (+/- 1 SD). 300
- Figure 6.S3 Day of the week box plots for the factor contributions of full data set solution. The boxes depict the median (dark line), inner quartile range (shaded box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (asterisk). The dashed line across the plot is the overall median. The individual median values and the number of points contained within each box are listed below the box. 301
- Figure 6.S4 Median PMF factor profiles for the cold period sub-data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations. 302
- Figure 6.S5 Median PMF factor profiles for the warm period sub-data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations. 303
- Figure 6.S6 Median PMF factor profiles for the hot period sub-data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations. 304

## CHAPTER 1 INTRODUCTION

### 1.1 RESEARCH MOTIVATION, CONTEXT AND GOLAS

Ambient fine particles less than 2.5  $\mu\text{m}$  in diameter ( $\text{PM}_{2.5}$ ) are composed of complex chemicals from primary or secondary sources. The short- and long-term exposures to  $\text{PM}_{2.5}$  were linked to increased health risks (Dockery et al., 1993; Ware, 2000; Zanobetti and Schwartz, 2009). Several epidemiological studies have applied receptor models to identify  $\text{PM}_{2.5}$  sources and quantify source contributions, and found relationships between some  $\text{PM}_{2.5}$  sources and negative health outcomes (Laden et al., 2000; Ito et al., 2005). However, these studies usually used  $\text{PM}_{2.5}$  data from only one receptor site for a certain region. The spatial variability of  $\text{PM}_{2.5}$  compositions and sources or differences in concentrations of  $\text{PM}_{2.5}$  components across a certain area might lead to biased estimation of health risks (Zeger et al., 2001).

Receptor models (e.g., Positive Matrix Factorization, PMF; Chemical Mass Balance, CMB) have been widely used to apportion  $\text{PM}_{2.5}$  to pollution sources/factors (Jaekels et al., 2007; Lee and Russell, 2007; Shrivastava et al., 2007). Unlike the CMB model, PMF (Paatero and Tapper) does not rely on a priori source profile information, and the factor profiles and contributions are directly solved from the ambient data. Speciation data used for PMF modeling are usually a combination of bulk species (e.g., elemental carbon, EC; organic carbon, OC) and a large array of trace elements (e.g., Al, Cu, Cd) (Kim et al., 2005; Hwang et al., 2008; Mooibroek et al., 2011) or organic molecular markers (OMMs) (Jaekels et al., 2007; Shrivastava et al., 2007; Schnelle-Kreis et al., 2007). Up to now, very few studies have evaluated the consistency of source apportionment results from different speciation data sets using an identical receptor model.

A critical study design feature such as the type of chemical analysis could change the results of source apportionment.

For receptor-based source apportionment, the source profile is pre-assumed as constant over the period of ambient and source sampling (Chen et al., 2011). However, the output factors of a receptor model are not necessarily emission sources. A factor could also reflect atmospheric processes like photochemical reaction or gas/particle (G/P) partitioning. The influence of atmospheric processes on a certain output factor could change along with meteorological conditions (e.g., solar irradiance, ambient temperature). Thus, the assumption of constant source profiles should not be true for all output factors. Compositional data of semi-volatile organic compounds (SVOCs) in PM<sub>2.5</sub> are becoming more commonly used than those of elements as inputs for receptor-based source apportionment (Jaekels et al., 2007; Shrivastava et al., 2007; Dutton et al., 2010b), because a number of source emissions dominated by fine particles do not have unique elemental composition (Schauer et al., 1996). In contrast, some individual or groups of SVOCs could be uniquely associated with specific emission sources, and then commonly referred to as organic molecular markers. Examples include levoglucosan for biomass burning (Simoneit et al., 1999) and 2-methyltetrols for isoprene derived secondary organic aerosol (SOA) (Claeys et al., 2004). However, the SVOCs are mostly subject to G/P partitioning, and the weight fraction of total SVOCs in the particle phase can change with ambient temperature and the chemical composition of ambient particles. As such, the source profiles of SVOC-based source apportionment could change due to the influence of G/P partitioning, especially for those factors characterized by light SVOCs.

The Denver Aerosol Sources and Health (DASH) project aimed to relate short-term exposure to individual PM<sub>2.5</sub> components and sources to negative health effects (Vedal et al.,

2009). Daily 24-h  $PM_{2.5}$  samples were collected from mid-2002 to the end of 2008 at a single receptor site located at an elementary school in downtown Denver. Speciation of the  $PM_{2.5}$  included gravimetric mass, inorganic ionic compounds (sulfate, nitrate and ammonium), EC, OC and a large array of organic molecular markers (OMMs). One year (2003) subset of the samples were also measured for water soluble carbon, water soluble nitrogen and water soluble elements (WSE) species. Dutton et al. (2009a, b, 2010a) have discussed the chemical speciation methods, point-wise uncertainty estimation and the temporal variations in bulk species and OMMs. In addition, source apportionment was conducted using a 1-year (January 27 – December 31, 2003) data set of bulk species and OMMs (Dutton et al., 2010b). An epidemiological study based on DASH data found that the estimated short-term effects of  $PM_{2.5}$  bulk components, especially those of EC and OC, were more immediate for cardiovascular diseases and more delayed for respiratory diseases (Kim et al., 2012b).

The first goal of this work was to examine the spatial variability of source contributions to ambient  $PM_{2.5}$ , which could benefit the understanding of the limitation in representativeness of the single site in the DASH study. One year of supplemental  $PM_{2.5}$  samples were collected at four sites, and their carbonaceous components were measured. The pooled data from all sampling sites were applied for source apportionment using the PMF2 model. The spatial variability in source contributions was investigated using correlation coefficients ( $r$ ) and coefficients of divergence (COD). Factor profiles derived from the pooled data set were also compared to those from site-specific data sets, so as to ensure the validity in using the pooled data set.

Second, the PMF model was applied to four different data sets composed of (1) bulk species, (2) bulk species and WSE, (3) bulk species and OMM, and (4) combination of all

species, so as to evaluate the utility of different speciation data sets for source apportionment of PM<sub>2.5</sub>. The three types of speciation data (bulk species, WSE and OMM) were all obtained for 1 year (2003) of daily PM<sub>2.5</sub> samples. In addition, the bootstrap technique developed by Hemann et al. (2009) was coupled to the PMF2 model to assess the uncertainty due to random sampling error. This technique also provided an alternative criterion – matching rate of bootstrapped factors to base case factors – for the selection of the number of factors to include in the PMF model.

Thirdly, a 32-month series of daily speciated PM<sub>2.5</sub> data was used as input for source apportionment to obtain more reliable PM<sub>2.5</sub> source information, which was used to associate individual factor contributions and short-term adverse health effects in a further DASH study. Moreover, a temperature stratified analysis was undertaken to identify the influence of atmospheric processes on factors resolved from the 32-month data set. To eliminate the influence of G/P partitioning on source apportionment using SVOCs data, gas-phase concentrations of SVOCs were calculated using an equilibrium absorption model (Pankow, 1994a, b), and added to their particle-phase concentrations for source apportionment.

The last goal was to verify the method for gas-phase SVOCs calculation by field measurement. In this work, a 1-year field study (with sampling every sixth day) was conducted for SVOCs in both gaseous and particle phases. To ensure high collection efficiency for gas-phase SVOCs, breakthrough experiments were performed on selected sampling days throughout the year. Beside those SVOCs reported by Dutton et al. (2009b), many more volatile *n*-alkanes and PAHs were analyzed. In addition, a different method was applied to measure polar organic species in both gaseous and particle phases. The gas- and particle-phase SVOCs data were used

to calculate observationally-based G/P partitioning coefficients ( $K_{p, OM}^m$ ,  $m^3 \mu g^{-1}$ ). Finally, the  $K_{p, OM}^m$  values were compared to those predicted by absorptive partitioning theory ( $K_{p, OM}^t$ ).

## 1.2 THESIS ORGANIZATION

This thesis is primarily constituted by original written work by the author either published or in preparation for submission. Chapters 2 – 8 are taken from the manuscripts with minor revisions in text schemes. The co-authors listed at the start of each of the following chapters made contributions in providing the conceptual framework, sample collection, data analysis or draft editing. Chapters 2 – 3 focus on the spatial variability of carbonaceous components of PM<sub>2.5</sub> and their sources (Xie et al., 2012a, b). Chapter 4 examines the consistency of source apportionment results from different speciation data sets using the PMF model (Xie et al., 2012c). Chapter 5 presents the source apportionment results using a 32-month series of daily PM<sub>2.5</sub> speciation data (Xie et al., 2013b). Temperature-stratified source apportionment is also included to identify the influence from atmospheric processes. Chapter 6 predicts the gas phase SVOCs based on an equilibrium absorption model for source apportionment (Xie et al., 2013a). Chapter 7 describes the sampling and chemical analysis of non-polar and polar SVOCs in both gaseous and particle phases (Xie et al., submitted). Chapter 8 compares the observed G/P partitioning of light SVOCs with their predicted partitioning, so as to verify the prediction of gas-phase SVOCs in Chapter 6. A summary of the whole study and possible future research are given in Chapter 9.

**CHAPTER 2 INTRA-URBAN SPATIAL VARIABILITY OF PM<sub>2.5</sub>-BOUND  
CARBONACEOUS COMPOUNDS**

Mingjie Xie<sup>a</sup>, Teresa L. Coons<sup>a</sup>, Steven J. Dutton<sup>b</sup>, Jana B. Milford<sup>a</sup>, Shelly L. Miller<sup>a</sup>, Jennifer L. Peel<sup>c</sup>, Sverre Vedal<sup>d</sup>, Michael P. Hannigan<sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, College of Engineering and Applied Science,  
University of Colorado, Boulder, CO 80309, USA

<sup>b</sup> National Center for Environmental Assessment, U.S. Environmental Protection Agency,  
Research Triangle Park, NC 27711, USA

<sup>c</sup> Department of Environmental and Radiological Health Sciences, Colorado State University,  
Fort Collins, CO 80523, USA.

<sup>d</sup> Department of Environmental and Occupational Health Sciences, School of Public Health and  
Community Medicine, University of Washington, Seattle, WA 98195, USA

## 2.0 ABSTRACT

The Denver Aerosol Sources and Health (DASH) study was designed to evaluate associations between PM<sub>2.5</sub> species and sources and adverse human health effects. The DASH study generated a five-year (2003-2007) time-series of daily speciated PM<sub>2.5</sub> concentration measurements from a single, special-purpose monitoring site in Denver, CO. To evaluate the ability of this site to adequately represent the short term temporal variability of PM<sub>2.5</sub> concentrations in the five county Denver metropolitan area, a one year supplemental set of PM<sub>2.5</sub> samples was collected every sixth day at the original DASH monitoring site and concurrently at three additional sites. Two of the four sites, including the original DASH site, were located in residential areas at least 1.9 km from interstate highways. The other two sites were located within 0.3 km of interstate highways. Concentrations of elemental carbon (EC), organic carbon (OC), and 58 organic molecular markers were measured at each site. To assess spatial variability, site pairs were compared using the Pearson correlation coefficient ( $r$ ) and coefficient of divergence (COD), a statistic that provides information on the degree of uniformity between monitoring sites. Bi-weekly co-located samples collected from July 2004 to September 2005 were also analyzed and used to estimate the uncertainty associated with sampling and analytical measurement for each species. In general, the two near-highway sites exhibited higher concentrations of EC, OC, polycyclic aromatic hydrocarbons (PAHs), and steranes than did the more residential sites. Lower spatial heterogeneity based on  $r$  and COD was inferred for all carbonaceous species after considering their divergence and lack of perfect correlations in co-located samples. Ratio-ratio plots combined with available gasoline- and diesel-powered motor vehicle emissions profiles for the region suggested a greater impact to high molecular weight

(HMW) PAHs from diesel-powered vehicles at the near-highway sites and a more uniformly distributed impact to ambient hopanes from gasoline-powered motor vehicles at all four sites.

## 2.1 INTRADUCTION

Numerous studies have found that both short- and long-term exposures to ambient particles less than 2.5  $\mu\text{m}$  in diameter ( $\text{PM}_{2.5}$ ) are associated with increased risk of mortality, as well as respiratory illness, lung cancer, asthma and heart disease (e.g., Dockery et al., 1993; Pope et al., 2002; U.S. EPA, 2009; Zanobetti and Schwartz, 2009).  $\text{PM}_{2.5}$  is a complex mixture of chemicals often emitted directly from combustion sources or formed from atmospheric transformation of gas-phase precursors. To examine the associations between increased health risks and distinct sources that contribute to  $\text{PM}_{2.5}$ , several epidemiologic studies have applied receptor models to identify and quantify source impacts (e.g., Laden et al., 2000; Ito et al., 2005). Such studies have suggested links between health effects and sources in particular regions, but the collective evidence on this association from different studies remains inconclusive (Stanek et al., 2011). This variability in the evidence could be due to differences in  $\text{PM}_{2.5}$  sources and composition across studies, or to exposure estimation error resulting from spatial variability in  $\text{PM}_{2.5}$  sources and composition within individual study areas (Ito et al., 2004; Kim et al., 2005; Marmur et al., 2006). Heterogeneity in  $\text{PM}_{2.5}$  components and sources or substantial differences in the magnitude of concentrations across an urban area can lead to biased health effects estimates (Zeger et al., 2001).

The Denver Aerosol Sources and Health (DASH) study was designed to evaluate chemical components and sources of  $\text{PM}_{2.5}$  in relation to acute risk of mortality and morbidity (Vedal et al., 2009). The DASH study obtained a five-year (2002-2006) time series of daily speciated  $\text{PM}_{2.5}$  measurements from a single special-purpose monitoring site in Denver, CO. Speciation measurements included inorganic ions, elemental carbon (EC), total organic carbon (OC), and organic molecular markers. The resulting time series provide significant insight into the temporal

variability of PM<sub>2.5</sub> components and source contributions in Denver (Dutton et al., 2009a,b; Dutton et al., 2010a,b). Moreover, preliminary results for the DASH project indicated that in Denver, roadway-associated sources, especially diesel emissions characterized by EC, might be more strongly linked with respiratory hospital admissions and mortality than PM<sub>2.5</sub> mass and other PM components (Peel et al., 2009).

To understand the limitations of the single site used in the DASH study, we collected one year of supplemental PM<sub>2.5</sub> samples at three additional sites to investigate the spatial variability of PM<sub>2.5</sub> components and source contributions in Denver. To frame the spatial variability, uncertainty associated with sampling and analytical measurement needs to be addressed. For example, assume measurements taken at two sites in an urban area are correlated with an  $r$  value of 0.6. This degree of correlation across sites might be viewed as relatively high, if side-by-side measurements from a single site are only correlated with an  $r$  value of 0.7 due to uncertainty in each measurement. To assess the side-by-side uncertainty in this study, PM<sub>2.5</sub> concentrations were measured bi-weekly from mid-July 2004 to late September 2005, using two co-located samplers. Finally, spatial distributions of diesel- and gasoline-powered motor vehicle source influences are discussed based on the comparison of the ambient data to published source profiles.

## 2.2 METHODS

### 2.2.1 PM<sub>2.5</sub> sampling sites

The primary sampling site for the DASH study was located on the rooftop of Palmer Elementary School (PAL), where daily samples were collected beginning on July 1, 2002. The three additional sites began sample collection on March 1, 2008 and were located to the north, northwest and southwest of PAL, as shown in Figure 2.1. PAL is in a large residential area with relatively dense population, far from industrial point sources, and with the nearest major interstate highway (I-25) located 5.2 km to the southwest (Vedal et al., 2009). The supplemental site northwest of PAL was located on the rooftop of Edison Elementary School (EDI). EDI is also in a residential neighborhood far from industrial point sources, with the nearest major interstate highways located 1.9 km to the north (I-70) and 3.0 km to the southeast (I-25). In contrast, the other two sites were located in more industrial areas and closer to major interstate highways. The site to the north was located on the rooftop of Alsup Elementary School (ALS) in the north end of the major industrial area of Denver and 0.3 km from I-76. The site to the southwest was located on the roof of the Denver Municipal Animal Shelter (MAS), surrounded by small industrial facilities and only 0.2 km from the major north-south interstate highway (I-25). The MAS site was selected to leverage co-located multi-pollutant measurements taken as part of the State of Colorado's NCore monitoring program for Denver (U.S. EPA, 2011). These locations provide for comparisons between two comparatively residential sites (PAL and EDI) and two sites with greater expected roadway influences (MAS and ALS).

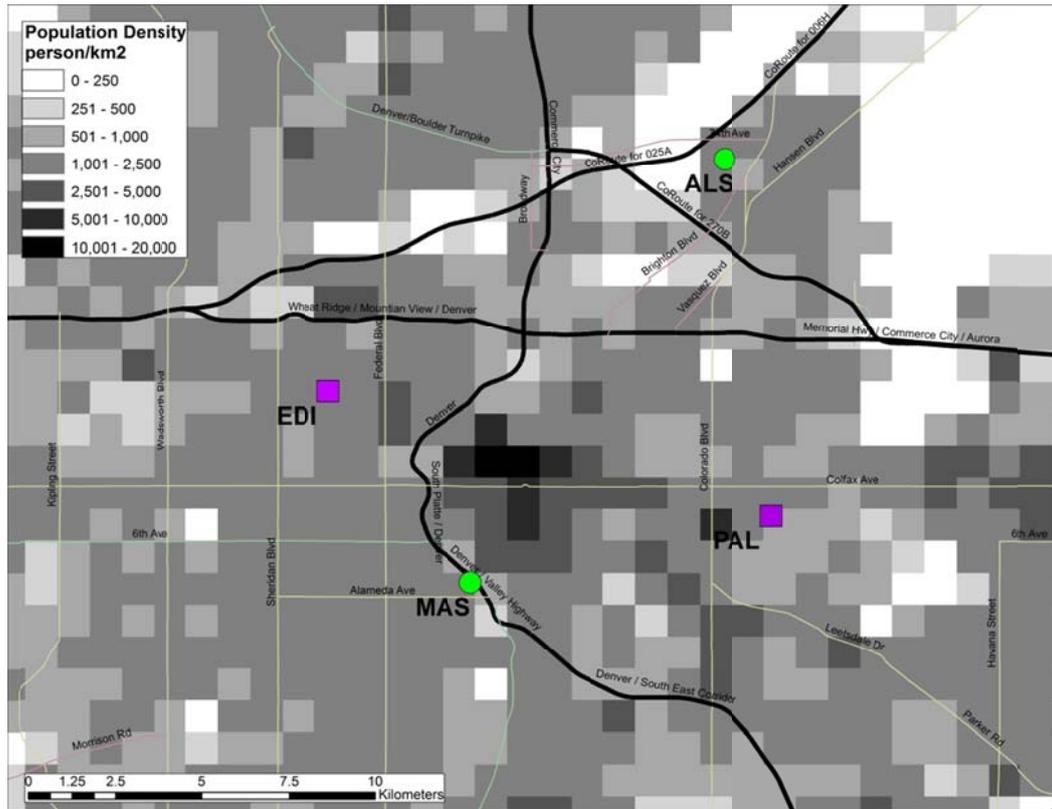


Figure 2.1 Locations of the four sampling sites in Denver urban area: (1) Palmer Elementary School (PAL); (2) Edison Elementary School (EDI); (3) Alsup Elementary School (ALS); (4) Municipal Animal Shelter (MAS).

### 2.2.2 Sample collection and chemical characterization

Twenty-four hour PM<sub>2.5</sub> filter samples were collected at the four sites every sixth day from March 1, 2008 to March 14, 2009. Hourly observed meteorological data for this period were obtained from nearby monitoring stations operated by the Colorado Department of Public Health and Environment (CDPHE) (Table 2.1). Prior to conducting multi-site sampling, bi-weekly co-located samples were collected at PAL from July 2004 to September 2005. Details of the sampling set up, protocols, and chemical analysis were discussed previously (Dutton et al., 2009a, b). Briefly, PM<sub>2.5</sub> was separated by aerodynamic diameter using a cyclone incorporating a 2.5 µm size cut at a flow rate of 92 L min<sup>-1</sup>. After the cyclone, the airstream was split with 20 L min<sup>-1</sup> passing through a 47 mm diameter, 2 µm pore size Teflon (PTFE) filter and 72 L min<sup>-1</sup> passing through a 90 mm diameter pre-baked (500 °C) quartz fiber filter. This paper focuses on the EC, OC, and organic molecular marker measurements obtained from the quartz fiber filters.

Table 2.1 Meteorological statistics based on hourly observation from nearby monitoring stations during March 1, 2008 - March 14, 2009. (PAL - CAMP; EDI - Carriage; ALS - Welby; MAS - Denver Animal Shelter)

Parameter	Statistic	PAL	EDI	ALS	MAS
Temperature (°C)	Mean	12	11	11	11
	SD <sup>a</sup>	9	9	9	10
	Mean daily max <sup>b</sup>	19	19	18	17
	Mean daily min <sup>c</sup>	6	2	4	5
Scalar Wind Speed (km h <sup>-1</sup> )	Mean	7	5	8	8
	SD	2	2	4	2
	Mean daily max	13	11	17	16
	Mean daily min	3	2	2	3
Relative Humidity (%) <sup>d</sup>	Mean	42	42		
	SD	17	16		
	Mean daily max	63	65		
	Mean daily min	25	23		

(a) Standard deviation of daily means.

(b) mean of the daily maximum observation.

(c) Mean of the daily minimum observation.

(d) Only available for two sites, data for PAL is from DESC (2.7 miles) and that for EDI is from Auraria (2.8 miles).

EC and OC were measured on 1.5 cm<sup>2</sup> punches taken from the quartz filters, using a Sunset Laboratory ECOC analyzer operated under the NIOSH 5040 thermal optical transmission (TOT) method. Organic molecular markers were extracted from the filters with reagent grade methylene chloride and analyzed using an Agilent 6890N gas chromatograph coupled with an Agilent 5975 mass spectrometer. High volume injection (50 μL) achieved by programmable temperature vaporization was applied to improve the detection limit (Dutton et al, 2009b). Fifty-eight organic molecular marker compounds were quantified for each sample, including *n*-alkanes, polycyclic aromatic hydrocarbons (PAHs), oxygenated PAHs (oxy-PAHs), steranes, *n*-alkanoic acids, sterols and methoxyphenols. Samples were run in sequences incorporating 14 samples, 2 field blanks, 6 solvent blanks and 5 dilutions of quantification standards. Quadratic calibration curves were generated for each molecular marker from all available runs of quantification standards in a given batch (4-7 sequences). Example calibration curves from one of the three batches are shown in Figure 2.S1 (supporting information); Table 2.S1 presents the fraction of the compounds detected in samples of that batch falling in the calibration range. The calibration curves were used along with the known mass of internal standards pre-spiked before filter extraction to determine the final mass amount of each molecular marker by converting peak area ratios to mass ratios. In this work, all components were field-blank corrected by subtracting off the median blank value within a given analysis batch. More details of the quantification method are given in the supporting information after Figure 2.S1. The 31 pairs of co-located samples were analyzed using the same protocols as the spatial samples. Table 2.S2 presents statistics for each species quantified at the four sites, and Table 2.S3 presents those for the co-located samples.

### 2.2.3 Uncertainty estimation

Quantification uncertainties for EC were estimated using the calculation recommended by the instrument manufacturer (Sunset Laboratory) and those for OC based on laboratory observations (Schauer et al., 2003). For the organic molecular markers, quantification uncertainty was estimated empirically from the analytical calibration curve (Dutton et al., 2009b). The root sum of squares (RSS) method (NIST, 1994) was used to propagate the uncertainty for point-wise uncertainty estimation, involving uncertainty in instrumental analysis, blank correction (standard deviation of field blanks within each batch) and sample air volume (Dutton et al., 2009b). These quantification uncertainties are reported as signal to noise (S/N, mean concentration/mean uncertainty) ratios in Tables 2.S2 and 2.S3.

#### 2.2.4 Statistical analysis

In this study, the spatial characteristics of concentrations of carbonaceous species were evaluated using Pearson correlation coefficients ( $r$ ) and coefficients of divergence (COD). Pearson correlation coefficients show the degree of correspondence of the chemical components between two sampling sites (Wongphatarakul, et al., 1998). High  $r$  values (close to unity) indicate that concentrations of the chemical component are proportional between the two sites throughout the sampling period. COD is applied to further evaluate the similarity between concentrations at two different monitoring sites and is defined as

$$\text{COD}_{fn} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{x_{if} - x_{ih}}{x_{if} + x_{ih}} \right)^2}$$

where  $x_{if}$  and  $x_{ih}$  are the concentrations of one species for the  $i$ th time period at sites  $f$  and  $h$ , respectively, and  $n$  is the number of observations. Values approaching 0 represent uniformity between pairs of samples, while values approaching 1 represent complete divergence (Wongphatarakul et al., 1998; Kim et al., 2005). From previous studies (Wongphatarakul, et al.,

1998; Kim et al., 2005; Wilson et al., 2005), Krudysz and co-workers (2008) inferred a boundary COD value of 0.2, where COD values  $> 0.20$  are defined as heterogeneous spatial distribution and values  $< 0.20$  represent spatially homogeneous air pollutants.

## 2.3 RESULTS AND DISCUSSION

### 2.3.1 General description of measurement data

#### 2.3.1.1 Spatial samples

Table 2.S2 lists statistics for concentrations of each species at each site, including mean, median, coefficient of variation (CV, standard deviation/mean concentration), S/N ratio and the percentage of observations below detection limit (BDL). The data obtained on Nov. 25, 2008 (during the Thanksgiving holiday) were not included in Table 2.S2 due to the extremely high concentrations observed on that day for all categories of carbonaceous species (Figure 2.S2). The elevated concentrations are probably attributable to heavy traffic combined with a stagnant atmosphere and cold weather (Figure 2.S3). Further discussion of the Nov. 25<sup>th</sup> data is provided in the supplemental information. During the rest of the sampling campaign, total *n*-alkanoic acids were the most abundant and evenly distributed of the molecular marker classes, followed by *n*-alkanes and methoxyphenols. PAHs, oxy-PAHs and steranes exhibited lower concentrations by 1-2 orders of magnitude than the *n*-alkanoic acids. The odd-even patterns of *n*-alkanes were similar among the four sites, peaking at C<sub>29</sub> and C<sub>31</sub> (Table 2.S2), consistent with the patterns observed in leaf abrasion products (Rogge et al., 1993c). Lower S/N ratios for some species (e.g., tridecanoic acid and heptadecanoic acid) suggest higher analytical uncertainty, and correspond to a larger number of observations below detection limits. The CV value is a robust metric used to assess the temporal variation of species concentrations (Dutton et al., 2009b; Krudysz et al., 2009). At all four sampling sites, the CVs of most organic molecular markers were close to or higher than unity. The highest CV values were observed for retene (1.2 - 1.7), *n*-alkanoic acids (1.0 - 1.7) and methoxyphenols (1.1 - 2.9). In contrast, CVs for bulk species were lower than

unity, ranging from 0.5 to 0.7. These results are consistent with the observations from Dutton et al. (2009a, b).

In Table 2.S2, the average EC concentrations at ALS ( $0.66 \mu\text{g m}^{-3}$ ) and MAS ( $0.48 \mu\text{g m}^{-3}$ ) were higher than at PAL ( $0.31 \mu\text{g m}^{-3}$ ) and EDI ( $0.32 \mu\text{g m}^{-3}$ ), likely reflecting larger contributions from vehicle emissions at the sites located closer to interstate highways. Unlike EC, which is exclusively a primary species, OC is of both primary and secondary origin (Aurela et al., 2011). However, the higher OC concentrations at ALS ( $3.56 \mu\text{g m}^{-3}$ ) and MAS ( $3.51 \mu\text{g m}^{-3}$ ) might be mostly due to primary emissions, because average OC/EC ratios at these two sites (mean  $\pm$  sd,  $9.1 \pm 5.0$  and  $6.9 \pm 4.0$ ) are lower than those at the two more residential sites (EDI  $12.2 \pm 6.5$ , PAL  $12.0 \pm 6.9$ ). The highest correlation between OC and EC concentrations was observed at ALS ( $r = 0.63$ ) while the lowest was detected at PAL ( $r = 0.41$ ). Steranes and PAHs (not including retene) exhibited higher average concentrations at ALS ( $2.03 \text{ ng m}^{-3}$  and  $3.19 \text{ ng m}^{-3}$ ) and MAS ( $1.89 \text{ ng m}^{-3}$  and  $2.91 \text{ ng m}^{-3}$ ) than at PAL ( $1.43 \text{ ng m}^{-3}$  and  $1.88 \text{ ng m}^{-3}$ ) and EDI ( $1.52 \text{ ng m}^{-3}$  and  $2.21 \text{ ng m}^{-3}$ ), consistent with the spatial variation of EC concentrations. Sources of particulate oxy-PAHs include both direct incomplete combustion (e.g., motor vehicle exhaust, wood burning) (Fitzpatrick et al., 2007) and oxidation of parent PAHs in the environment (Walgraeve et al., 2010), so oxy-PAHs species concentrations were not always higher at the two near-highway sites. Methoxyphenols and retene, which are widely used as biomass burning tracers (Schauer et al., 2001), were found in higher concentrations at PAL ( $18.3 \text{ ng m}^{-3}$  and  $0.72 \text{ ng m}^{-3}$ ) and EDI ( $29.4 \text{ ng m}^{-3}$  and  $0.64 \text{ ng m}^{-3}$ ) than ALS ( $13.7 \text{ ng m}^{-3}$  and  $0.41 \text{ ng m}^{-3}$ ) and MAS ( $17.0 \text{ ng m}^{-3}$  and  $0.43 \text{ ng m}^{-3}$ ). In addition, larger weekend increases in total concentrations of these species were observed at PAL (145% higher than the weekday average) and EDI (92%)

than at ALS (25%) and MAS (57%). These observations may be explained by the fact that EDI and PAL are closer than the other sites to residential wood combustion sources.

Table 2.2 compares studies that have investigated the spatial distribution of PM<sub>2.5</sub> components, including organic molecular markers. The total average concentrations of *n*-alkanes, PAHs and steranes in Denver were higher than those observed in Long Beach, CA (Krudysz et al., 2009). The average concentrations of these organic molecular markers were less varied across different sites in Denver than in Vancouver, BC (Cheng et al., 2004, 2006), Atlanta, GA (Li et al., 2009) and Hong Kong, China (Guo et al., 2003; Ho et al., 2003). Concentrations of EC and OC at the four sampling sites in Denver were lower than observed in other studies. All of the studies listed in Table 2.2 show smaller spatial variations for OC than EC.

Table 2.2 Studies on spatial variations of PM<sub>2.5</sub> components

Location	Site	Sample No.	Organic Markers (ng m <sup>-3</sup> )				Bulk species (µg m <sup>-3</sup> )		
			<i>n</i> -alkanes	PAHs	Steranes	<i>n</i> -alkanoic acids	OC	EC	PM <sub>2.5</sub> Mass
Denver <sup>a</sup>	PAL (Residential)	62	30.5	2.60	1.43	223	3.03	0.31	6.51
	EDI (Residential)	59	24.9	2.85	1.52	234	3.39	0.32	6.81
	ALS (Traffic)	63	29.4	3.60	2.03	249	3.56	0.66	9.32
	MAS (Traffic)	63	24.3	3.34	1.89	233	3.51	0.48	7.79
Long Beach <sup>b,c</sup>	NLB (Traffic)	10	16.6	1.07	0.43		4.03	1.32	12.9
	SLB (Traffic)	7	17.6	1.23	0.77		3.45	0.74	11.2
	Sutter (Traffic)	10	21.6	1.65	0.58		5.07	1.70	14.6
	Riley (Residential)	10	23.5	1.48	0.62		5.09	0.92	13.6
Lower Fraser Valley <sup>d,e</sup>	GEP (Forest)	9	10.6			60.0			2.37
	CT (Tunnel)	8	83.9			156			1.12
	SP (Urban park)	23	14.8			50.0			1.71
	SER (Urban/forest)	25	10.3			65.0			1.96
	LEL (Rural)	23	8.20			39.0			1.92
Atlanta <sup>f</sup>	High way	7	40.1	11.0	7.97	67.7	8.16	4.06	22.7
	GeoTech. campus	7	25.1	3.19	3.15	32.0	5.53	0.96	16.4
Hong Kong, China <sup>g,h</sup>	PU (Traffic)	14		41.8			9.45	5.80	50.9
	KT (Industrial)	12		27.9			10.2	5.05	57.3

(a) This study.

(b) Krudysz et al. (2008).

(c) Krudysz et al. (2009).

(d) Cheng et al. (2004).

(e) Cheng et al. (2006).

(f) Li et al. (2009).

(g) Ho et al. (2003).

(h) Guo et al. (2003).

### 2.3.1.2 Co-located samples

Statistics of species concentrations from 31 pairs of co-located samples obtained at the PAL site from July 2004 – September 2005 are given in Table 2.S3, including mean and median concentrations, standard deviations, S/N ratios and percentages of BDL measurements. Values of  $r$  and COD for each species between the co-located samples were also calculated to reflect the combined uncertainty associated with sampling and analytical measurement. Ideally, all species concentrations should be highly correlated ( $r \sim 1$ ) with relatively low CODs ( $\sim 0$ ) between the co-located samples. In fact, co-located measurements of HMW  $n$ -alkanes ( $C_{32} - C_{35}$ ), oxy-PAHs (except benzo[de]anthracene-7-one),  $n$ -alkanoic acids, cholesterol and methoxyphenols were less correlated and had higher CODs than other species (Table 2.S3). The concentrations of the species with high side-by-side divergence (e.g., dotriacontane, tridecanoic acid, sterols and most methoxyphenols) were very low in a number of samples and prone to have relatively high quantification uncertainties, as indicated by their high percentages of measurements below the detection limit and low S/N ratios (Table 2.S3). These low concentrations show large differences between co-located samples, leading to lower  $r$  and higher COD values than other species; the relatively small sample number ( $N = 31$ ) of co-located samples could amplify the impacts of those low measurement values in the comparison using  $r$  and COD. In addition, chromatogram integrations for HMW and semi-polar components have higher uncertainties than lighter and non-polar ones due to their wider peaks. Other species, including lighter  $n$ -alkanes ( $C_{22}$ - $C_{31}$ ), PAHs, steranes, EC and OC have relatively high correlations ( $r = 0.82 - 0.97$ ) between co-located samples.

### 2.3.2 Analysis of spatial variability

#### 2.3.2.1 Correlation coefficient ( $r$ ) and coefficient of divergence (COD)

Figure 2.2 (a-f) shows correlation coefficients and CODs for concentrations of organic molecular markers, EC, and OC measured at the four sites from March 2008 – March 2009. OC concentrations had higher correlation across sites ( $r = 0.80 - 0.85$ ) and lower CODs ( $0.17 - 0.20$ ) than EC, consistent with the observations of other studies (Krudysz et al., 2009; Cheung et al., 2011), which is likely caused by the smaller number of EC source types and different source strength across different sampling sites. Heavier odd *n*-alkanes ( $C_{27}, C_{29}, C_{31}, C_{33}$ ) are primarily of biogenic origin (Simoneit and Mazurek, 1982; Rogge et al., 1993c; Cass, 1998). The heavier odd *n*-alkanes were highly spatially correlated ( $r = 0.69 - 0.94$ ) with a clear odd to even carbon number predominance in  $r$  value, and showed an opposite pattern in CODs across all spatial site pairs. PAH (not including retene) concentrations had weaker correlations ( $r = 0.27 - 0.66$ ) between all site pairs, suggesting there were different types of PAH sources contributing at the four sites. However, CODs of these compounds were lower between MAS and ALS ( $0.25 - 0.45$ ) than for the other site pairs, which might be attributed to the similar effects of nearby roadway emissions at MAS and ALS. Steranes are primarily from lubricating oil use of both diesel and gasoline-powered vehicles (Minguillon et al., 2008), so the spatial variations of these components should be dominated by traffic emission distributions. In Figure 2.2a and b, steranes had higher correlations ( $r = 0.55 - 0.85$ ) and lower CODs ( $0.26 - 0.36$ ) than in the other plots, suggesting similar impacts of roadway emissions between the two residential sites (EDI and PAL), and the two near-highway sites (MAS and ALS). The three oxygenated PAHs (xanthone, 1,8-naphthalic anhydride, anthracene-9,10-dione) and *n*-alkanoic acids exhibited peak concentrations in warm seasons at all four sites, and were generally more correlated across sites than PAHs and steranes. Concentrations of the biomass burning markers (methoxyphenols and retene) were generally well correlated but had higher CODs ( $0.31 - 0.56$ ) than most of the other

species. To draw conclusions about the spatial variability of one compound class relative to another, we need to understand the uncertainties associated with sampling and analytical measurement for the different compound classes.

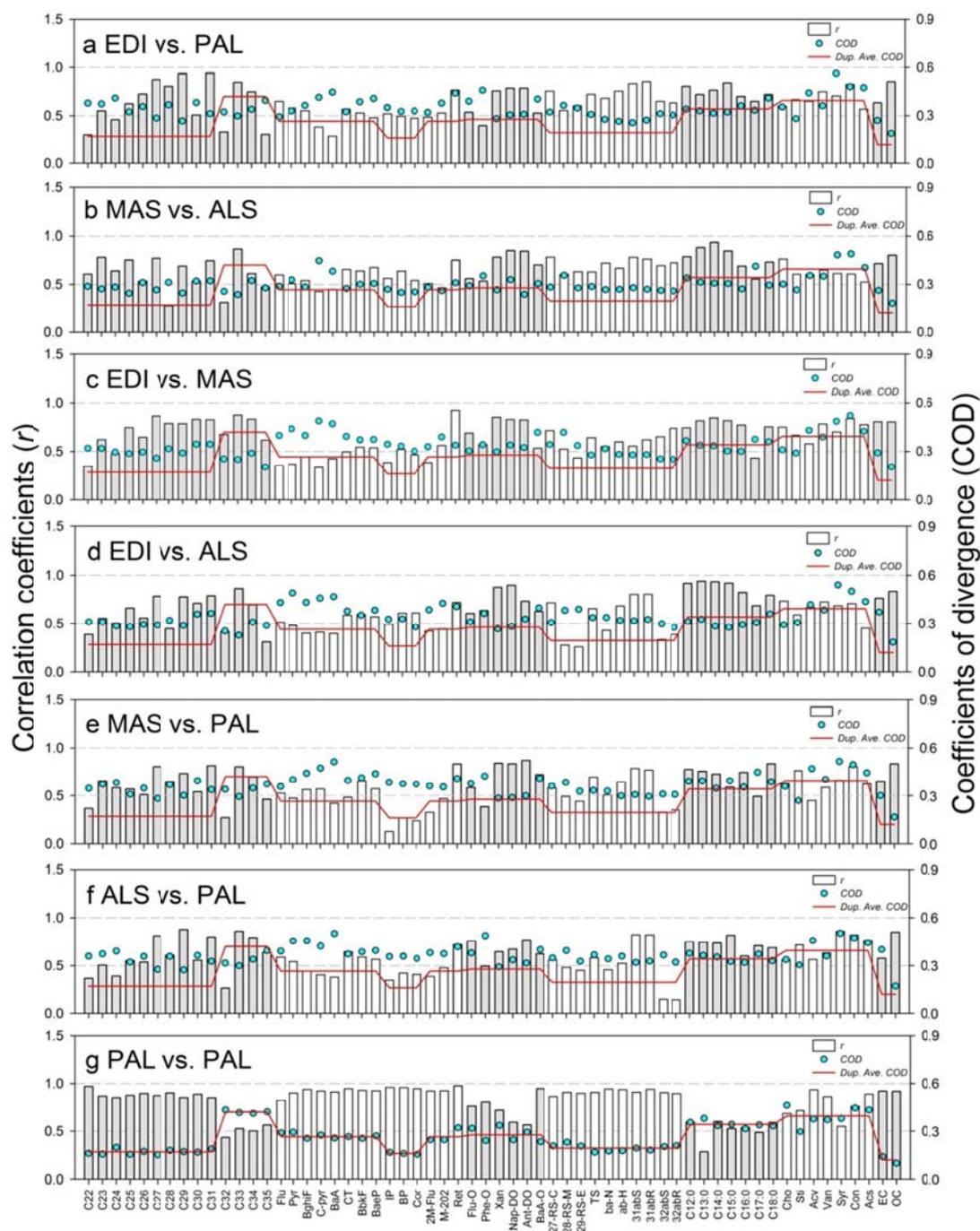


Figure 2.2 Pearson correlation coefficients (bars) and coefficients of divergence (light blue circles) for organic molecular markers and bulk carbon contents obtained at the 4 sampling sites (a-f) as well as the co-located samples at PAL site (g). CODs of collocated samples were averaged in groups and overlaid on each plot. Bars are shaded to show compound class separations.

### 2.3.2.2 Spatial variability framed by variability in co-located sample results

The statistics  $r$  and COD have been widely used to describe spatial variability (Kim et al., 2005; Hwang et al., 2008). However, very few studies have examined the point of reference for interpreting the degree to which they represent spatial variability versus uncertainty associated with sampling and analytical measurement (Krudysz et al., 2008). Given the divergence and lack of perfect correlation observed in co-located samples (Table 2.S3), it is unrealistic to expect a COD near zero or value of  $r$  near 1 in comparing observations across sites, even if atmospheric concentrations are highly uniform. For reference in this study, values of  $r$  and COD for each species from co-located samples are shown in Figure 2.2g; the COD values for different groups of molecular markers are averaged and overlaid on each plot in Figure 2.2. This comparison is rough, because the co-located sample comparisons are for a different time-period and include a smaller number of samples than the comparisons across sampling locations. Some species (e.g., sterols and methoxyphenols) exhibited large side-by-side divergence due to their high percentages of BDL measurements. Comparing the COD between each sampling site pair and the COD for the co-located pair, the heavy odd  $n$ -alkanes ( $C_{27}$ ,  $C_{29}$ ,  $C_{31}$ ), steranes and OC have low divergence between co-located samples and slightly higher divergence across sampling locations;  $n$ -alkanes with a chain length ranging from  $C_{22}$  to  $C_{30}$  (not including  $C_{27}$  and  $C_{29}$ ), HMW PAHs (MW = 276, 300) and EC have low divergence between co-located samples but high divergence across locations; low MW PAHs (MW  $\leq$  226) exhibit moderate divergence between co-located samples and higher divergence for pairs of different sites (e.g., PAL vs. MAS); other species (e.g.,  $C_{32}$ - $C_{35}$ ,  $n$ -alkanoic acids, sterols and methoxyphenols) show high divergence between co-located samples with comparable or even lower divergence across locations. Therefore, the spatial distribution of organic species in this work might be less

heterogeneous than it first appears, as some of the apparent differences across locations may be due to sampling and analytical uncertainty. The impacts of quantification uncertainty on  $r$  and COD for side-by-side and multi-site comparisons were simulated (Table 2.S4) and are discussed in the supporting information. The results suggest that uncertainties in species quantification can have a substantial influence on values of  $r$  and COD that are used to assess spatial variability.

### 2.3.3 Spatial distribution of mobile sources influences

EC in urban environments is mainly associated with diesel vehicle emissions, while large PAHs (6-8 aromatic rings) are primarily emitted from gasoline powered vehicles (Rogge et al., 1993a; Schauer et al., 1996; Riddle et al., 2007) and hopanes are from motor oil use (both diesel and gasoline; Kleeman et al., 2008). To better understand the relative influence of gasoline and diesel emissions spatially, two ratio-ratio plots including different motor vehicle emission markers were made for each of the four sites (Figure 2.3 a-h). One series of plots compares indeno[1,2,3-cd]pyrene and benzo[ghi]perylene that are normalized by EC; the other compares ba-30-norhopane and ab-hopane, again normalized by EC. Motor vehicle and wood burning profiles specifically for the Denver area (Zielinska et al., 1998; Cadle et al., 1999) were also overlaid on the plots with ambient data. If only gasoline-powered vehicle emissions or only diesel vehicle emissions are the dominant source of all three compounds, then the ambient measurements should cluster near the points of the corresponding source profiles. If both gasoline and diesel emissions are primary sources for the three compounds, then the ambient data in the plots should fall on a mixing line between the two source profiles. Finally, the ambient data can appear scattered on a ratio-ratio plot if there are three or more contributing sources (Robinson et al., 2006).

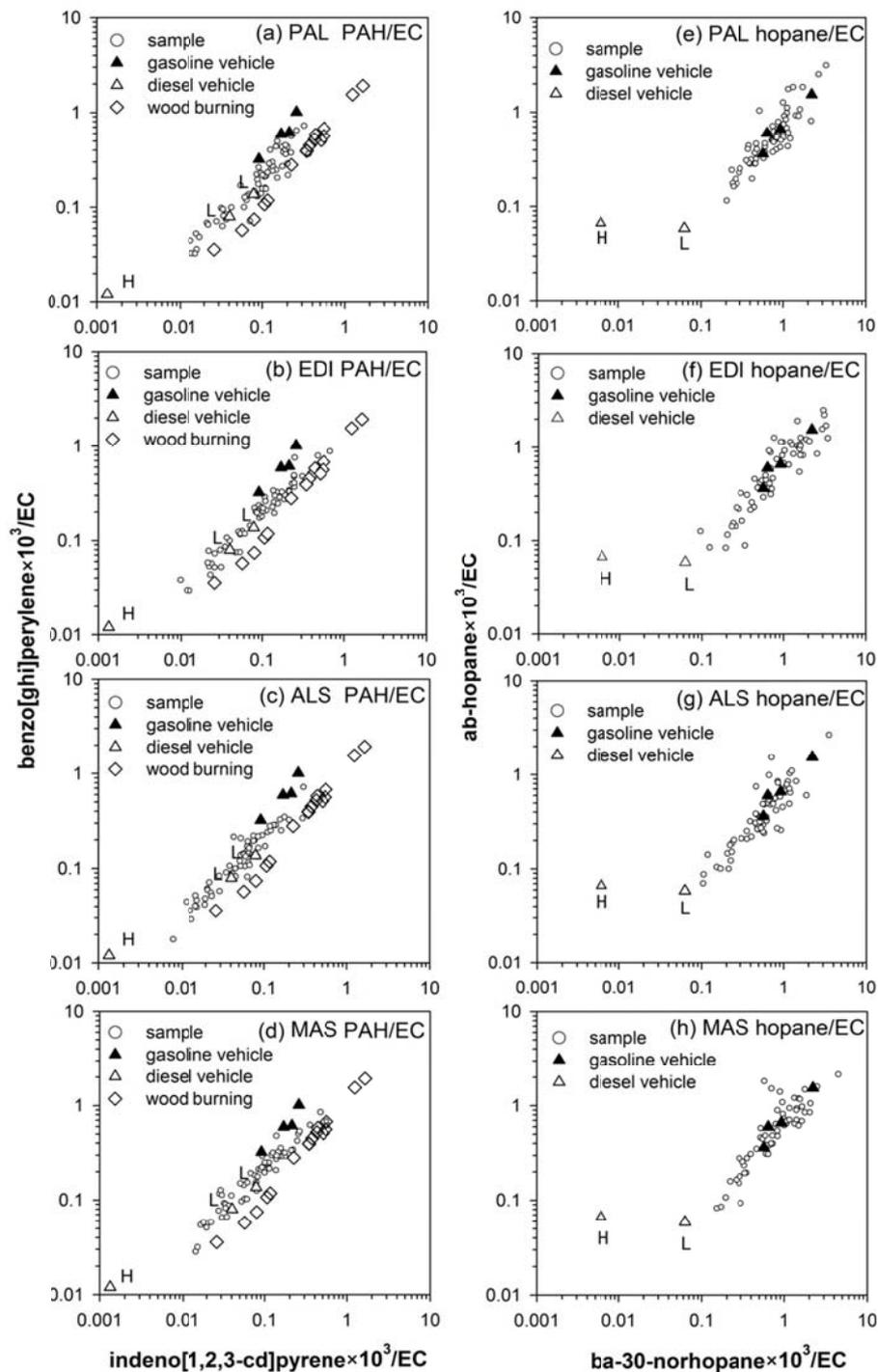


Figure 2.3 Ratio-ratio plots for two polycyclic aromatic hydrocarbons (PAHs) normalized by EC (a-d) and two hopanes normalized by EC (e-f) at each of the four sites. Gasoline, diesel (L light duty, H heavy duty) and wood burning (different woods and burning conditions in fireplace and woodstove) source profiles overlaid were obtained from Cadle et al. (1999) (only for light duty diesel associated PAHs in Denver) and NFRAQS study (Zielinska et al., 1998).

The ambient PAH data at each site fall along a relatively continuous line in all of the ratio-ratio plots (Figure 2.3a-d). The ambient PAH concentrations are more likely to be explained by a combination of motor vehicle source profiles (both gasoline and diesel) than wood burning, because the line representing wood burning sources lies below that of the ambient PAH data at all sites. More ambient PAH data at PAL and EDI sites (Figure 2.3a, b) correspond to gasoline and wood burning source profiles in the upper right extreme than at ALS and MAS sites (Figure 2.3c, d), indicating more diesel vehicle activity in the areas nearer to the interstate highways. The ambient PAH data also show significant influence from heavy duty diesel vehicles on a number of days, since a certain fraction of PAH to EC ratios fall below the source profiles of light duty diesel vehicle emissions and extend to the region of heavy duty diesel emissions in the lower left extreme. Unlike the PAHs, ambient hopane measurements mostly fall near the gasoline and light duty diesel source profiles at all four sites (Figure 2.3 e-h), supporting the conclusion that gasoline and light duty diesel emissions explain the Denver ambient hopane data (Dutton et al., 2009b).

## 2.4 CONCLUSIONS

Organic molecular markers and bulk carbon contents were analyzed in PM<sub>2.5</sub> samples collected every sixth day at four sampling sites for one year. Our results show that PAHs, steranes, methoxyphenols and EC concentrations were more affected by primary emissions and show distinct differences in concentrations between near-highway and residential sites. The HMW and semi-polar compound concentrations exhibit relatively less correlation and higher divergence between co-located samples, indicating high uncertainty from sampling and analytical measurements. The data and discussion in this paper also provide information on spatial variability of fine particulate organics, which was not examined in the previous DASH study. PAHs and steranes exhibited some degree of homogeneity for site pairs where each site is from a residential or a near-highway area, but showed less homogeneity for site pairs where one site was residential and one site was near-highway. OC was more strongly correlated and exhibited relatively lower CODs than EC. The spatial heterogeneity of organic molecular markers would be overestimated without considering their divergence between co-located samples. Values of  $r$  and CODs derived from co-located samples should be used as points of reference to analyze spatial variability of PM<sub>2.5</sub> species. In particular, the benchmark COD of 0.20 recommended by Krudysz et al. (2008) is too low to account for the influence of sampling and measurement uncertainties on CODs for many organic components in this work. The implication for DASH is that we could have high confidence in the data and representativeness from the PAL site for those compounds with low side-by-side divergence and low spatial variability, whereas compounds with either high side-by-side divergence or high spatial variability (or both) need to be treated more cautiously. From the ratio-ratio plots and overlaid source profiles, diesel vehicle emissions appear to contribute more HMW PAHs at near-highway

sites than at the more residential sites. Ambient hopane concentrations at all four sites appear to be mostly explained by a combination of gasoline and light duty diesel emissions. The data sets in this study are being used to perform source apportionment. Results from that work will be presented in a future paper, which will focus on the sources of particulate organics and the spatial variability of source contributions.

**CHAPTER 3 INTRA-URBAN SPATIAL VARIABILITY AND UNCERTAINTY**  
**ASSESSMENT OF PM<sub>2.5</sub> SOURCES BASED ON CARBONACEOUS SPECIES**

Mingjie Xie<sup>a</sup>, Teresa L. Coons<sup>a</sup>, Joshua G. Hemann<sup>a</sup>, Steven J. Dutton<sup>b</sup>, Jana B. Milford<sup>a</sup>,  
Jennifer L. Peel<sup>c</sup>, Shelly L. Miller<sup>a</sup>, Sun-Young Kim<sup>d</sup>, Sverre Vedal<sup>d</sup>, Lianne Sheppard<sup>d</sup>, Michael  
P. Hannigan<sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, College of Engineering and Applied Science,  
University of Colorado, Boulder, CO 80309, USA

<sup>b</sup> National Center for Environmental Assessment, U.S. Environmental Protection Agency,  
Research Triangle Park, NC 27711, USA

<sup>c</sup> Department of Environmental and Radiological Health Sciences, Colorado State University,  
Fort Collins, CO 80523, USA

<sup>d</sup> Department of Environmental and Occupational Health Sciences, University of Washington,  
School of Public Health, University of Washington, Seattle, WA 98195, USA

### 3.0 ABSTRACT

To identify the sources of PM<sub>2.5</sub> – bound carbonaceous species and examine the spatial variability of source contributions in the Denver metropolitan area, positive matrix factorization (PMF) was applied to one year of every sixth day ambient PM<sub>2.5</sub> compositional data, including elemental carbon (EC), organic carbon (OC), and 32 organic molecular markers, from four sites (two residential and two near-traffic). Statistics (median, inner quantiles and 5<sup>th</sup> – 95<sup>th</sup> percentiles range) of factor contributions, expressed as reconstructed carbonaceous mass (EC + OC), were estimated from PMF solutions of replicate data sets generated by using a stationary block bootstrap technique. A seven-factor solution was resolved for a set of data pooled across the four sites, as it gave the most interpretable results and had the highest rate of neural network factor matching (76.9%). Identified factors were primarily associated with high plant wax, summertime emission, diesel vehicle emission, fossil fuel combustion, motor vehicle emission, lubricating oil combustion and wood burning. Pearson correlation coefficients ( $r$ ) and coefficients of divergence (COD) were used to assess spatial variability of factor contributions. The summertime emission factor exhibited the highest spatial correlation ( $r = 0.74 - 0.88$ ) and lowest CODs (0.32 – 0.38) among all resolved factors; while the three traffic dominated factors (diesel vehicle emission, motor vehicle emission and lubricating oil combustion) showed lower correlations ( $r = 0.47 - 0.55$ ) and higher CODs (0.41 – 0.53) on average. Average total EC and OC mass were apportioned to each factor and showed a similar distribution across the four sites. Modeling uncertainties were defined as the 5<sup>th</sup> – 95<sup>th</sup> percentile range of the factor contributions derived from valid bootstrap PMF solutions, and were highly correlated with the median factor contribution in each factor ( $r = 0.77 - 0.98$ ). Source apportionment was also performed on site specific data sets; the results exhibited similar factor profiles and temporal variation in factor

contribution as those obtained for the pooled data set, indicating that the four sites are primarily influenced by similar types of sources. On the other hand, differences were observed in absolute factor contributions between PMF solutions for the pooled versus site-specific data sets, likely due to the large uncertainties in EC and OC factor profiles derived from the site specific data sets with limited numbers of observations.

### 3.1 INTRODUCTION

Numerous epidemiological studies report an association between short-term  $PM_{2.5}$  concentrations and increased incidence of various adverse health outcomes (U.S. EPA, 2009). To determine if these associations are driven by the origin of the  $PM_{2.5}$ , previous studies have used chemical speciation combined with receptor modeling to apportion PM mass to its sources, and then used the resulting source contribution time-series in health effects modeling (Laden et al., 2000; Ito et al., 2006; Stanek et al., 2011). PM toxicity likely varies by source, which drives the chemical composition. For example, a recent epidemiologic study has found a higher health risk associated with exposure to traffic PM than with exposure to sulfate PM from power plant emissions (Maynard et al., 2007). Therefore, in developing effective regulatory strategies to reduce the health impacts associated with  $PM_{2.5}$ , it might prove valuable to identify the source types and their contributions to  $PM_{2.5}$  mass concentration.

Chemical Mass Balance (CMB) and Positive Matrix Factorization (PMF) receptor models are commonly used to undertake source apportionment. CMB is not limited by the size of data set but requires quantitative information on source profiles in addition to ambient PM composition. A major limitation of CMB is that the available source profiles may not be representative of the sources impacting particular receptor sites (Jaeckels et al., 2007). Uncertainties in CMB source apportionment results are more influenced by the bias associated with source profile data than uncertainties in the ambient measurement itself (Lee and Russell, 2007). PMF is a multivariate model with non-negative factor constraints (Paatero and Tapper, 1994) that does not rely on a priori source profile information. Instead, the factor profiles and contributions are derived directly from the ambient data. The factors resolved from PMF can be

described as specific PM source types if the chemical profile of the factor matches that of emissions from a particular source type.

Very few studies have addressed uncertainties in PMF model solutions. The origins of this uncertainty include the input measurement uncertainty, the choice of input species measurement data, and the selection of PMF model parameters (e.g. factor number ( $p$ )) (Hemann et al., 2009). Christensen and Schauer (2008a) evaluated the stability of PMF solutions using perturbed measurement uncertainty matrices as inputs and resolved 10 factors with one specific factor having the most stable estimates of average source contribution. The bootstrap tool incorporated in the EPA-PMF model can be used to assess the effects of sampling error on factor profiles (Reff et al., 2007). The EPA-PMF bootstrap tool applies linear regression to match a factor from one solution to the “closest” factor in another solution, and is sensitive to outliers (Hemann et al., 2009).

In a previous Denver Aerosol Sources and Health (DASH) study, source apportionment was conducted by applying a PMF model on one-year time series of daily inorganic and organic speciated  $PM_{2.5}$  measurements at a single site (Dutton et al., 2010b). The resulting time series will be used to model associations between individual factor contributions and short term, adverse health effects. A handful of multi-site studies (Kim et al., 2005; Hwang et al., 2008; Kim and Hopke, 2008) have found heterogeneous spatial distributions of PM components and sources; thus, there is a question of the representativeness of the one site approach for estimating source contributions to area-wide exposures. In the Denver urban area, no significant  $PM_{2.5}$  point sources were identified by Dutton et al. (2010b), but lack of representativeness of a single site could still arise due to differences in the magnitude of source contributions across the metropolitan area. In order to examine the spatial variability of source contributions to ambient

PM<sub>2.5</sub> samples were collected at four sites in Denver every sixth day for one year. The carbonaceous composition of all the PM<sub>2.5</sub> samples was characterized and reported in a separate paper (Xie et al., 2012a). In this work, PMF2 was applied to the pooled data set, combining measurements of all sites, to identify and quantify PM<sub>2.5</sub> factor/source profiles and factor/source contributions. To assess the uncertainty of the PMF results, we applied a bootstrap technique described previously (Hemann et al., 2009), which aligns factors from different bootstrap solutions based on similarity between factor profiles rather than factor contribution time series. Spatial variability in source contributions was evaluated by using correlation coefficients ( $r$ ) and coefficients of divergence (COD). Finally, a test for the validity of using a pooled data set was conducted by comparing the source apportionment results with those obtained for site-specific data sets.

## 3.2 METHODS

### 3.2.1 Ambient sample collection and chemical characterization

PM<sub>2.5</sub> filter samples were obtained at four locations every sixth day from March 1, 2008 to March 14, 2009. A detailed description of the four sampling sites was given in Xie et al. (2012a); here the sites are summarized in Table 3.1. Details of the sampling equipment, protocols and chemical analysis were described previously (Dutton et al., 2009a, b; Vedal et al., 2009). Elemental carbon (EC), total organic carbon (OC) and 32 organic molecular markers were used for source apportionment in this study. Species were originally selected based on their association with specific sources, excluding species with missing values, species with more than 15% of observed concentrations below the detection limit (BDL) and species with signal to noise ratio (S/N, mean concentration/mean uncertainty) less than 5. EC (S/N, 2.2 – 4.1; Table 3.2) was retained in the data set since it is an important marker for motor vehicle emissions. Concentration values that are BDL are used as reported, because PMF itself does not require replacement of BDL values (Dutton et al., 2010b). After bootstrap runs, some species were also removed from subsequent analysis based on their effects on factor matching (e.g. using alkanolic acids resulted in very low factor matching rates (<20%) due to their high variability in factor profiles). The samples collected on November 25, 2008 at all four sites were excluded from the source apportionment analysis as concentrations of all groups of species were unusually high (1.6 – 11 times greater than the annual averages) (Xie et al., 2012a). These extreme measurements were caused by heavy traffic combined with a stagnant atmosphere and cold weather. Including these measurements for PMF analysis did not result in the identification of additional factors. Table 3.2 lists the average concentrations and S/N ratios of selected species at each site.

Table 3.1 Locations and characteristics of the four sampling sites in Denver.

Sampling sites Location	Abbr.	Latitude	Longitude	Site type	Closest Major Interstate	
					Name	Distance (km)
Palmer Elementary School	PAL	39.73	-104.92	Residential	I-25	5.2
Edison Elementary School	EDI	39.77	-105.04	Residential	I-70	1.9
Alsup Elementary School	ALS	39.83	-104.94	Traffic	I-76	0.3
Municipal Animal Shelter	MAS	39.72	-105.03	Traffic	I-25	0.2

Table 3.2 Average concentration  $\pm$  standard deviation and signal to noise ratios of selected species for source apportionment.

Species (ng m <sup>-3</sup> )	Abbr.	PAL (N=63)		EDI (N=59)		ALS (N=63)		MAS (N=62)	
		Mean $\pm$ SD <sup>a</sup>	S/N <sup>b</sup>	Mean $\pm$ SD	S/N	Mean $\pm$ SD	S/N	Mean $\pm$ SD	S/N
docosane	C22	2.19 $\pm$ 2.07	8.56	2.00 $\pm$ 2.31	8.33	2.23 $\pm$ 1.97	8.5.0	1.66 $\pm$ 1.18	8.54
tricosane	C23	4.00 $\pm$ 5.33	11.2	3.10 $\pm$ 3.25	10.7	2.75 $\pm$ 2.59	10.4	2.30 $\pm$ 2.04	10.7
tetracosane	C24	1.89 $\pm$ 1.97	13.1	1.11 $\pm$ 1.03	12.2	1.25 $\pm$ 1.04	12.6	1.05 $\pm$ 0.78	12.5
pentacosane	C25	1.88 $\pm$ 1.93	10.8	1.66 $\pm$ 1.73	10.8	1.69 $\pm$ 1.51	10.8	1.62 $\pm$ 1.37	11.2
hexacosane	C26	0.85 $\pm$ 0.86	9.43	0.56 $\pm$ 0.42	6.73	0.73 $\pm$ 0.73	8.51	0.66 $\pm$ 0.58	8.98
heptacosane	C27	1.64 $\pm$ 2.05	12.7	1.31 $\pm$ 1.23	11.2	1.62 $\pm$ 1.98	12.3	1.42 $\pm$ 1.45	12.8
octacosane	C28	0.82 $\pm$ 0.91	9.24	0.62 $\pm$ 0.64	7.08	0.89 $\pm$ 1.67	9.35	0.73 $\pm$ 0.92	9.25
nonacosane	C29	4.20 $\pm$ 7.92	13.8	3.43 $\pm$ 5.21	13.4	4.36 $\pm$ 9.86	13.1	3.51 $\pm$ 5.52	13.7
triacontane	C30	0.92 $\pm$ 1.25	9.96	0.67 $\pm$ 0.84	7.24	0.74 $\pm$ 0.93	8.03	0.82 $\pm$ 1.03	10.2
hentriacontane	C31	11.2 $\pm$ 22.8	11.9	9.66 $\pm$ 16.8	11.9	12.4 $\pm$ 31.8	11.4	9.67 $\pm$ 15.9	11.9
fluoranthene	Flu	0.21 $\pm$ 0.18	13.1	0.24 $\pm$ 0.24	14.0	0.40 $\pm$ 0.32	15.3	0.35 $\pm$ 0.24	15.6
pyrene	Pyr	0.13 $\pm$ 0.17	12.5	0.17 $\pm$ 0.25	13.9	0.29 $\pm$ 0.27	14.0	0.23 $\pm$ 0.16	13.5
benzo[ghi]fluoranthene	BghiF	0.08 $\pm$ 0.08	19.3	0.11 $\pm$ 0.13	23.6	0.17 $\pm$ 0.17	22.9	0.16 $\pm$ 0.13	24.2
cyclopenta[cd]pyrene	C-pyr	0.04 $\pm$ 0.05	10.1	0.04 $\pm$ 0.05	9.99	0.05 $\pm$ 0.06	11.4	0.06 $\pm$ 0.08	13.3
benz[a]anthracene	BaA	0.05 $\pm$ 0.07	12.1	0.07 $\pm$ 0.10	15.3	0.09 $\pm$ 0.11	15.4	0.11 $\pm$ 0.18	17.4
chrysene/triphenylene	CT	0.18 $\pm$ 0.19	16.8	0.23 $\pm$ 0.23	17.6	0.33 $\pm$ 0.29	17.1	0.33 $\pm$ 0.29	18.0
benzo[b&k]fluoranthene	BbkF	0.22 $\pm$ 0.26	14.5	0.28 $\pm$ 0.33	15.6	0.36 $\pm$ 0.40	15.5	0.37 $\pm$ 0.42	17.0
benz[a&e]pyrene	BaeP	0.18 $\pm$ 0.21	10.6	0.21 $\pm$ 0.24	11.3	0.28 $\pm$ 0.32	12.8	0.29 $\pm$ 0.32	14.6
indeno[1,2,3-cd]pyrene	IP	0.03 $\pm$ 0.03	8.26	0.03 $\pm$ 0.02	7.86	0.04 $\pm$ 0.04	9.84	0.05 $\pm$ 0.04	12.1
benzo[ghi]perylene	BP	0.06 $\pm$ 0.05	8.16	0.06 $\pm$ 0.04	7.57	0.09 $\pm$ 0.08	9.81	0.09 $\pm$ 0.06	11.3
coronene	Cor	0.03 $\pm$ 0.02	9.58	0.03 $\pm$ 0.02	9.55	0.04 $\pm$ 0.04	11.0	0.04 $\pm$ 0.03	11.1
2-methylfluoranthene	2M-Flu	0.20 $\pm$ 0.20	14.5	0.22 $\pm$ 0.26	14.6	0.29 $\pm$ 0.23	15.8	0.25 $\pm$ 0.17	16.3
methyl-202-PAH sum	M-202	0.68 $\pm$ 0.84	16.9	0.72 $\pm$ 0.77	16.8	1.05 $\pm$ 0.90	17.0	0.83 $\pm$ 0.59	17.5
retene	Ret	0.72 $\pm$ 1.22	16.3	0.64 $\pm$ 0.96	15.4	0.41 $\pm$ 0.48	13.2	0.43 $\pm$ 0.59	14.7
1,8-naphthalic anhydride	Nap-DO	0.38 $\pm$ 0.53	18.4	0.52 $\pm$ 0.78	19.3	0.60 $\pm$ 0.95	17.8	0.46 $\pm$ 0.61	18.8
anthracene-9,10-dione	Ant-DO	0.53 $\pm$ 0.66	18.9	0.57 $\pm$ 0.56	19.2	0.78 $\pm$ 0.90	17.8	0.67 $\pm$ 0.80	18.7
benz[de]anthracene-7-one	BaA-O	0.06 $\pm$ 0.07	14.9	0.08 $\pm$ 0.09	16.7	0.10 $\pm$ 0.10	16.1	0.10 $\pm$ 0.11	17.2
20R & S-abb-ethylcholestane	29-RS-E	0.11 $\pm$ 0.09	16.3	0.11 $\pm$ 0.09	16.5	0.16 $\pm$ 0.16	16.2	0.13 $\pm$ 0.10	17.0
ba-30-norhopane	ba-N	0.25 $\pm$ 0.19	18.2	0.28 $\pm$ 0.21	18.6	0.42 $\pm$ 0.38	16.9	0.38 $\pm$ 0.26	18.2
ab-hopane	ab-H	0.19 $\pm$ 0.17	14.4	0.21 $\pm$ 0.19	14.4	0.31 $\pm$ 0.31	14.3	0.28 $\pm$ 0.22	14.9
acetovanillone	Acv	0.98 $\pm$ 1.61	14.0	1.18 $\pm$ 1.30	16.0	0.65 $\pm$ 0.82	12.6	0.55 $\pm$ 0.63	13.0
vanillin	Van	6.03 $\pm$ 7.96	14.3	9.25 $\pm$ 11.7	15.8	4.84 $\pm$ 6.06	13.2	4.67 $\pm$ 6.26	14.2
EC ( $\mu\text{g m}^{-3}$ )		0.31 $\pm$ 0.19	2.20	0.32 $\pm$ 0.19	2.18	0.66 $\pm$ 0.45	4.09	0.48 $\pm$ 0.30	3.67
OC ( $\mu\text{g m}^{-3}$ )		3.03 $\pm$ 1.61	10.2	3.39 $\pm$ 1.99	10.9	3.56 $\pm$ 1.96	11.2	3.51 $\pm$ 1.63	11.9

(a) Standard deviation (SD).

(b) Signal to noise ratio (mean concentration /mean uncertainty).

### 3.2.2 Positive matrix factorization (PMF2) modeling

PMF2 (Paatero, 1998a, b), a multivariate receptor model based on a weighted least squares approach, was applied under robust mode for source apportionment in this work. PMF2 is solved by apportioning a series of observations to several distinct factors and minimizing the sum of the squared, scaled residuals ( $Q$ ). FPEAK, which controls the rotational freedom, and factor number ( $p$ ) are two important parameters adjusted to improve model fit. In this work, FPEAK was ultimately set at zero for all PMF runs since indistinguishable results were observed for factor profiles and contributions for FPEAK varying from -0.2 to 0.3. The number of factors was determined primarily on the interpretability of different solutions (5 – 8 factors) and the rate of unique factor matching between bootstrapped factors and base case factors.

### 3.2.3 Uncertainty assessment for PMF modeling

Details of the method for uncertainty assessment of model fit applied in this work have been published by Hemann et al. (2009). Briefly, a stationary block bootstrap technique (Politis and Romano, 1994) was used to generate 1000 replicate data sets from the original data set by re-sampling blocks of samples with replacement. The block size was chosen as the median ( $b$ ) of all the lag-dependence seen in the data, and then 1 is added, so that when a day was resampled, the day plus the previous  $b$  days would be included. After that, each new data set was analyzed with PMF2. Factor profiles from each PMF solution were sorted and aligned to base case factors by using Multilayer Feed Forward Neural Networks, which relies on supervised learning from given inputs. If even one factor from a solution could not be uniquely matched to a base case factor, then that solution was dropped from subsequent results. Sampled measurement days in each replicate data set were tracked to examine the bias and variability in the PMF solution for each factor on each day.

### 3.2.4 Analysis of spatial variability

The spatial variability of PM<sub>2.5</sub> source contributions were evaluated using  $r$  and COD with COD defined as:

$$\text{COD}_{fh} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{x_{if} - x_{ih}}{x_{if} + x_{ih}} \right)^2} \quad (1)$$

where  $x_{if}$  and  $x_{ih}$  are one specific factor contribution for the  $i^{\text{th}}$  time period at sites  $f$  and  $h$ , respectively, and  $n$  is the number of observations. Values approaching 0 represent uniformity between pairs of sites, while values approaching 1 represent strong divergence (Wongphatarakul et al., 1998; Kim et al., 2005).

### 3.3 RESULTS AND DISCUSSION

#### 3.3.1 PMF solution based on pooled data set

Under the assumption that the same sources, with similar composition, were present at each receptor site (Mooibroek et al., 2011), we first combined the PM<sub>2.5</sub> sample data from all sites for source apportionment analysis. A seven-factor solution was chosen as having the most physically interpretable results. Moreover, with the seven-factor solution, the neural network factor matching method uniquely matched bootstrapped factors to base case factors in 76.9% of the bootstrapped data sets, compared to 55.1% and 53.0% matching rates for six- and eight-factor solutions. The higher rate of successful bootstrap matching is evidence that the seven-factor solution is more stable under the effects of sampling error. The distributions of the factor profiles for the seven-factor solution are shown in Figure 3.1, in which all factor profiles were normalized by

$$F_{kj}^* = \frac{F_{kj}}{\sum_{k=1}^p F_{kj}} \quad (2)$$

where  $F_{kj}^*$  is the relative weighting of species  $j$  in factor  $k$  to all other factors. Figure 3.2 shows factor contributions of reconstructed carbonaceous mass (EC + OC) for all seven factors at each site, expressed in ng m<sup>-3</sup>. Each plot in Figure 3.2 presents time series of daily factor contributions for the base case, median, and two empirical quantiles (interquartile range and 5<sup>th</sup>-95<sup>th</sup> percentile range) derived from PMF bootstrap solutions.

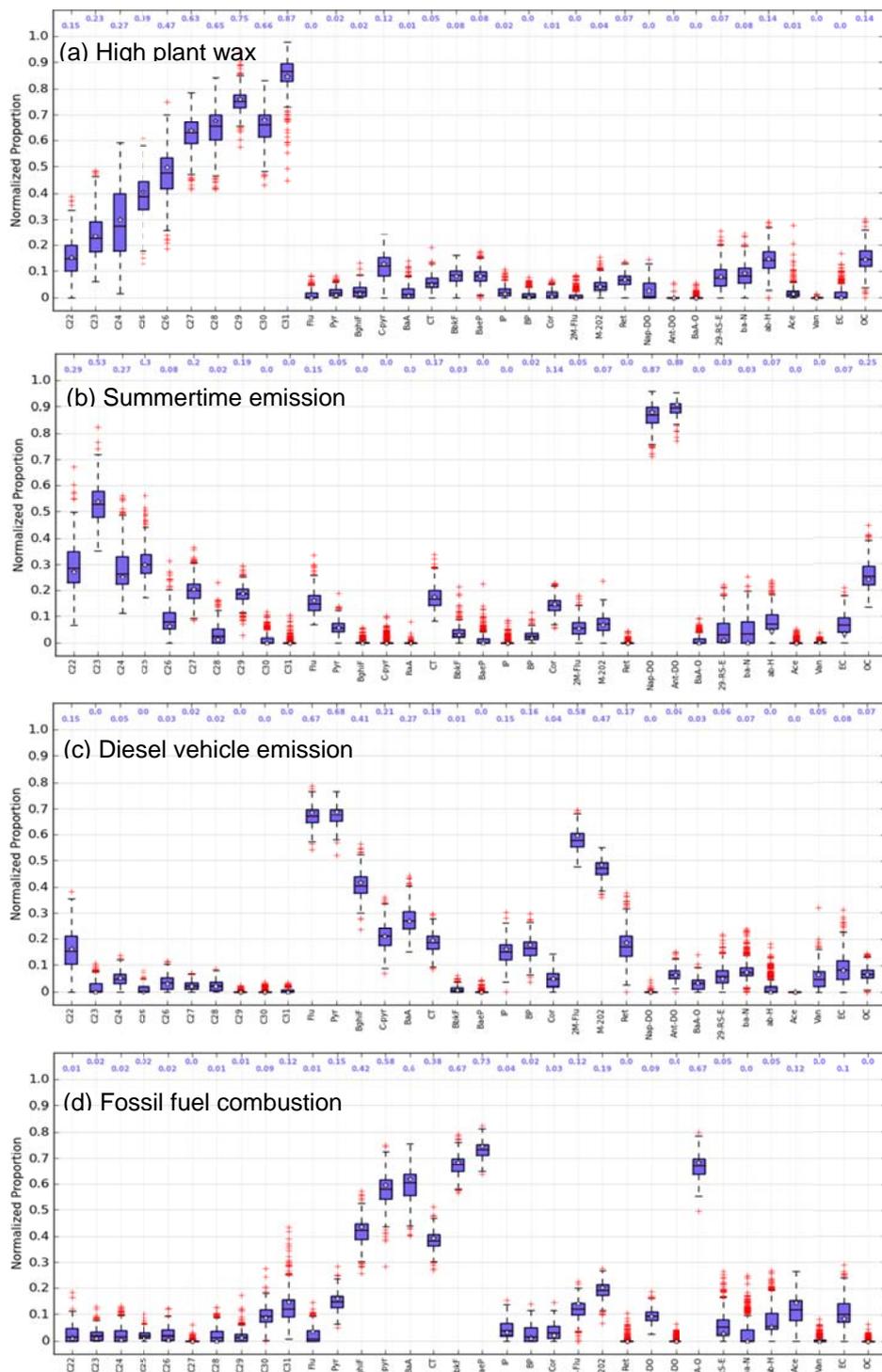


Figure 3.1 Box plots of normalized factor profiles derived from 769 PMF solutions out of 1000 replicate data sets. The boxes depict the median (dark line), inner quartile range (blue box), lower quartile  $- 1.5 \times \text{IQR}$  and upper quartile  $+ 1.5 \times \text{IQR}$  (whisker), base case (blank star) and outliers (red plus). Values of the medians are labeled on the top of each plot.

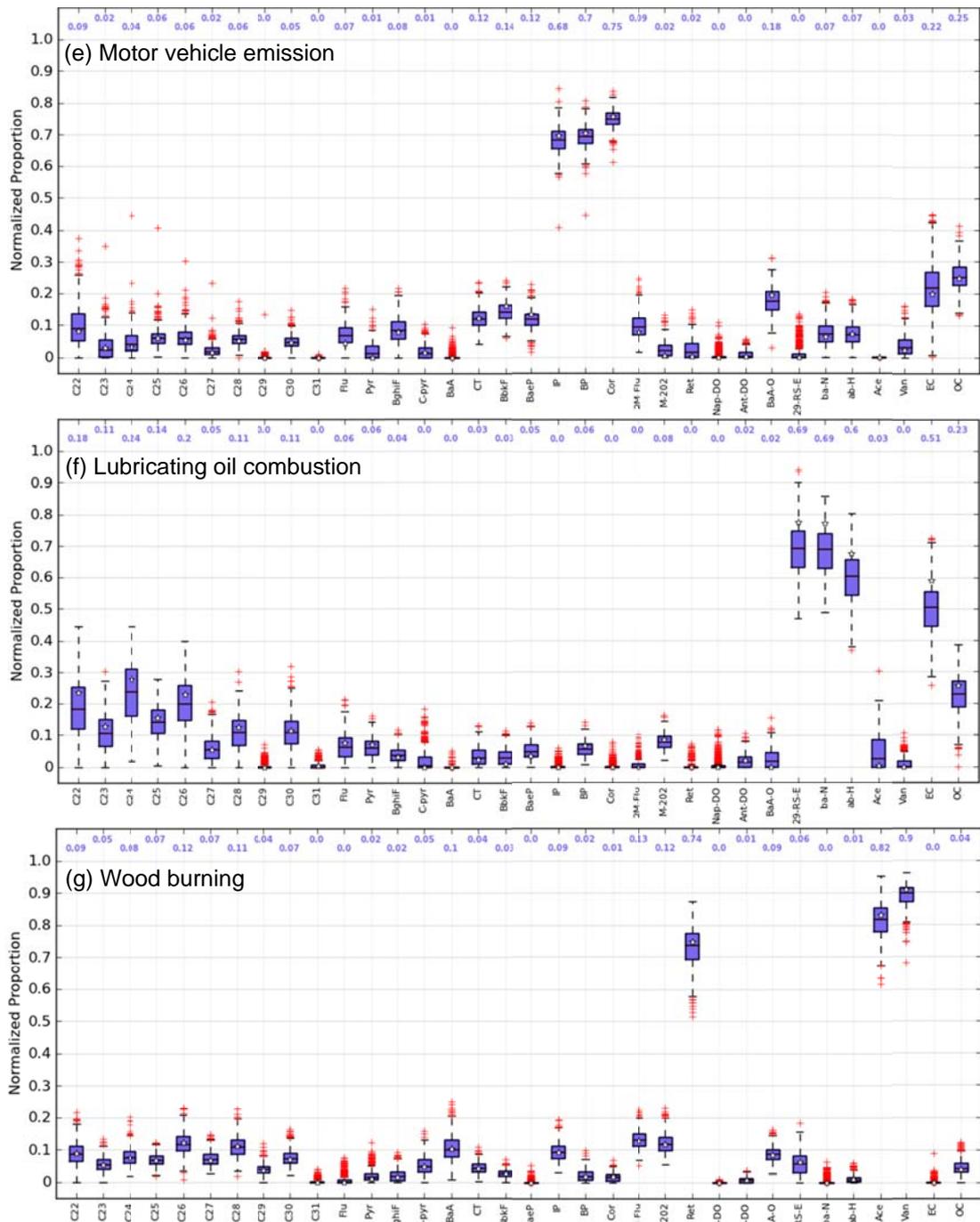


Figure 3.1 Continued.

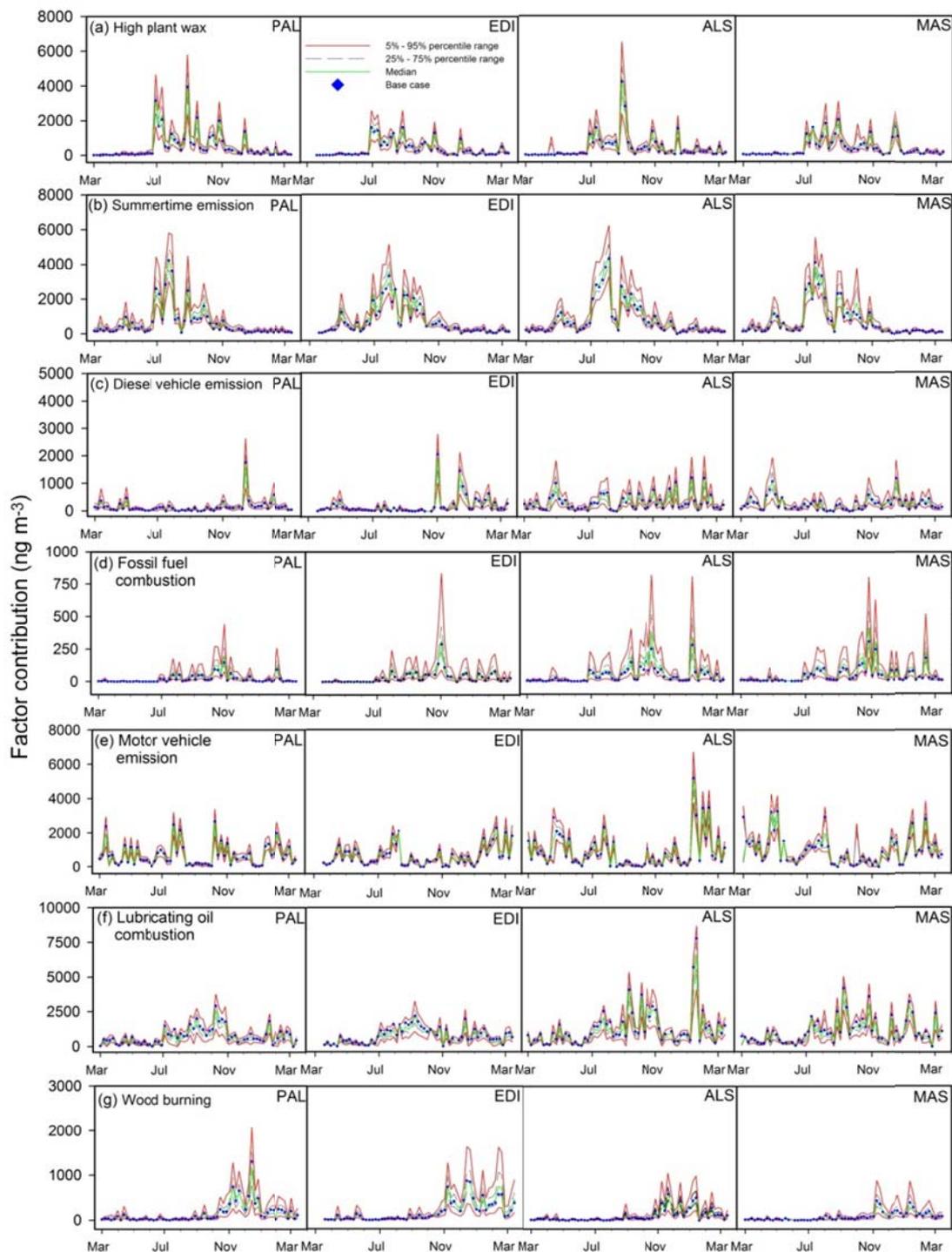


Figure 3.2 Factor contribution time series separated by site from PMF bootstrap solutions. Each plot shows the time series of factor contribution for base case (blue diamond), median (green line), and two bands based on the empirical quantiles of the bootstrap solutions (25<sup>th</sup> – 75<sup>th</sup> percentiles, gray dashed line; 5<sup>th</sup> – 95<sup>th</sup> percentiles, red solid line).

### 3.3.1.1 Factor profile interpretation

Factor 1 contains the largest fraction of *n*-alkanes with peak contributions mainly distributed in summer and autumn (Figures 3.1a, 3.2a). As summarized by Schnelle-Kreis et al. (2007), *n*-alkanes with a chain length of C<sub>14</sub> to C<sub>25</sub> could come from unburned heating oil, which has peaks in emissions mass fractions in the C<sub>19</sub>-C<sub>21</sub> range. *n*-alkanes from vehicle emissions have chain lengths ranging from C<sub>19</sub> to C<sub>32</sub> with peaks at C<sub>20</sub> and C<sub>25</sub> for gasoline vehicles and C<sub>20</sub> for diesel vehicles. High plant waxes, emitted through the abrasion of plant leaves, are dominated by odd *n*-alkanes in the range of C<sub>27</sub>-C<sub>33</sub>, peaking at C<sub>29</sub> and C<sub>31</sub>. In this work, Factor 1 accounts for the majority of the *n*-alkanes in the size range from C<sub>27</sub> to C<sub>31</sub>, but for less than half of the smaller ones with no obvious odd – even predominance. However, observed concentrations of the two highest loaded odd *n*-alkanes (C<sub>29</sub> and C<sub>31</sub>) during the growing season were 1-2 orders of magnitude higher than the adjacent even ones, similar results were also obtained for median daily factor contributions. Therefore, this factor was probably influenced by high plant wax from leaf abrasion emissions. A clear wintertime peak of factor contribution centered in December was also observed at all four sites, accompanied with peak concentrations at C<sub>29</sub> and C<sub>31</sub>, which might be caused by the high plant wax of dead leaves (Rogge et al., 1993c) enriched in the road dust.

The two oxygenated PAHs (1, 8-naphthalic anhydride and anthracene-9, 10-dione), each having two ketones, were predominately loaded in Factor 2 (Figure 3.1b). Factor 2 also contains around one-third to half of *n*-alkanes in the size range from C<sub>22</sub> to C<sub>25</sub>. Factor contributions across all of the sites showed similar seasonal patterns, mostly prominent in summer (Figure 3.2b), and were highly correlated with the ambient temperature (Supporting Information Figure 3.S1). These observations might indicate a summertime biogenic source, high temperature

dependent emissions, and potential contributions from photochemical reactions which we collectively refer to as summertime emissions. As suggested by Dutton et al. (2010b), the addition of polar species (e.g. 2-methylthreitol, pinic acid) would benefit the separation of secondary organic PM from primary PM emissions during summer for this factor. In addition, Factor 2 is responsible for the highest proportion of OC mass annually with a median of 25%; while EC mass associated with Factor 2 was the second lowest (median 7%).

Factor 3 accounts for most of the low molecular weight (LMW) PAHs ( $MW < 228$ , Figure 3.1c), which could be attributed to emissions from diesel vehicles (Miguel et al., 1998; Caricchia et al., 1999), biomass burning (Jenkins et al., 1996), unburned petroleum products (Li et al., 2009) or incineration (Ravindra et al., 2006). Factor 3 contributions at the two residential sites (EDI and PAL) were highest in winter, while the two near-road sites (MAS and ALS) exhibited intermittently high factor contributions throughout the year (Figure 3.2c). Possible explanations for the wintertime increase include enhanced domestic biomass burning and increased gas to particle phase sorption at lower temperatures (Subramanyam et al., 1994), as well as reduced photo-chemical reactions and decreased atmospheric mixing (Ravindra et al., 2006; Ravindra et al., 2008). In the ratio-ratio plots (Figure 3.S2a-d), ambient fluoranthene and pyrene (two dominant species in this factor) measurements generally fall along a line framed by wood burning, gasoline and diesel vehicle emissions profiles, and mostly lie between the source profiles of light and heavy duty diesel emissions at each site. Thus this factor could be primarily associated with the emissions from diesel-powered vehicles. Besides biomass burning, petroleum emissions (fuel leaks and evaporation) may also contribute to this factor, as the largest fractions of alkyl PAHs are apportioned to this factor (Simo et al., 1997; Yunker et al., 2002).

Factor 4 contains large fractions of medium molecular weight PAHs (MW = 226, 228, 230, 252) (Figure 3.1d) and shows a similar seasonal pattern in factor contributions across all sites (Figure 3.2d). In urban areas, the majority of PAHs could be attributed to anthropogenic emissions, like motor vehicle emissions and industrial activities (Schauer et al., 1996; Ravindra et al., 2008; Li et al., 2009). In this work, measurements of ambient benzo[b&k]fluoranthene and benzo[a&e]pyrene (two dominant species in this factor) in ratio-ratio plots at each site also fall on a line, which is below that framed by wood burning source profiles and has more scattered points at the two residential sites (PAL and EDI) (Figure 3.S2 e-h). Moreover, some ambient data deviates or exceeds the line framed by a continuum of motor vehicle source profiles. Therefore, these medium MW PAHs are likely not only contributed by motor vehicle emissions. Fossil fuel combustions of industrial facilities could be another primary source for this factor. The 4- and 5-ring PAHs (BaA, CT, BghiF, C-pyr, BbkF, BaeP) have been identified as indicator markers from industrial stacks (Yang et al., 1998), and the similar seasonal pattern in factor contributions also suggests significant contribution from an area source like multi-flue gas stacks within a certain industrial area. As a result, the higher factor contributions at the two near-road sites (ALS and MAS) might be attributed to both the nearby traffic and industrial emissions.

Factor 5 contains the highest loadings of three heavy PAHs (MW = 276, 300), and shows an enhanced fraction of EC (Figure 3.1e). All these carbonaceous tracers are primarily associated with motor vehicle emissions (Rogge et al., 1993a; Schauer et al., 1996; Riddle et al., 2007). The identification of this factor with motor vehicle emissions is also suggested by the higher factor contributions observed at the two near-road sites (Figure 3.2e). However, gasoline versus diesel vehicle contributions to these heavy PAHs could not be separated in this work, likely due to their similar emission factors (Phuleria et al., 2007). In a separate paper, Xie et al. (2012a) presented a

ratio-ratio plot comparing indeno[1,2,3-cd]pyrene and benzo[ghi]perylene normalized by EC, which also suggested that gasoline and diesel were both significant sources.

Factor 6 includes the highest percentage of the three steranes and EC (median 51%, Figure 3.1f). Steranes are reliable organic markers for motor oil combustion (Kleeman et al., 2008), while EC is often but not exclusively associated with diesel vehicle emissions (Schauer, 2003; Diaz-Robles et al., 2008; Krudysz et al., 2009). This association is supported by observation that EC concentrations show a greater proportional reduction on the weekend than that in overall traffic counts (Dutton et al., 2010a). In ratio-ratio plots, the ambient sterane data at the four sites were mostly explained by a continuum of gasoline vehicle source profiles, and partially ascribed to light duty diesel vehicles (Xie et al., 2012a). As a result, this sterane/EC factor was categorized as lubricating oil combustion, which is also obtained by Dutton et al., (2010b).

Retene and the two methoxyphenols were the most characteristic compounds in factor 7 (Figure 3.1g). Factor contributions exhibited a similar and strong seasonality across the four sites with a wintertime maximum (Figure 3.2g), corresponding to an enhanced residential wood combustion in winter (Dutton et al., 2010a). As expected, the contributions from this wood burning factor are higher in residential areas (EDI and PAL) than at sites more impacted by traffic. Identification of this factor with residential wood combustion is also consistent with the significant weekend increases in total concentrations of these species at PAL (69%) and EDI (33%), while ALS and MAS had no weekend increase.

In prior work, Dutton et al. (2010b) obtained a 7-factor PMF solution for one year (2003) of daily observations at one of the four sites included in this study (PAL). Their speciated PM<sub>2.5</sub> data included EC, OC, nitrate, sulfate, ammonium and 62 organic molecular markers. Five of the

factors could be associated with distinct pollution sources, including wood combustion, motor oil combustion, motor vehicle emission, biogenic emission and secondary inorganic compounds. In that study, the wood combustion factor also contained a substantial fraction of the lighter PAHs that were apportioned to a distinct factor in the current work. In the previous study, the rest of the PAHs (medium and high MW PAHs) were incorporated in only one factor; such a combined factor was also observed for the data in this work when inorganic ions (only available for spring samples) were included in the model.

### 3.3.1.2 Uncertainty assessment for factor contributions

The source apportionment method applied in this study yields distributions of daily contributions for each factor (Hemann et al., 2009). The medians from the bootstrapped PMF solutions are treated as the best estimate of the true factor contribution time series, and the intervals from 5<sup>th</sup> to 95<sup>th</sup> percentile represent the variability in daily factor contribution. In Figure 3, the median factor contributions were highly correlated with the variability ( $r = 0.77 - 0.98$ ), indicating a stable relative uncertainty of daily factor contribution for each factor. The data were also separated by site and compared in Figure 3.S3; the linear regression equations for each factor were similar across sites. Slopes of the linear equations in Figure 3 ranged from 0.52 to 1.84 with the highest value for fossil fuel combustion (1.84) followed by wood burning (1.47) and diesel vehicle emission (1.13) factors. Thus, these three factors have larger relative uncertainties in daily factor contributions than the others, which might be related to their low factor contributions. The bootstrap approach used in this work could be applied to other source apportionment methods, since the analyses mainly reveal uncertainties in factor contribution estimates that result from random resampling error, without considering errors in measurement,

uncertainties associated with different source apportionment methods, factor number and pollutant species selection.

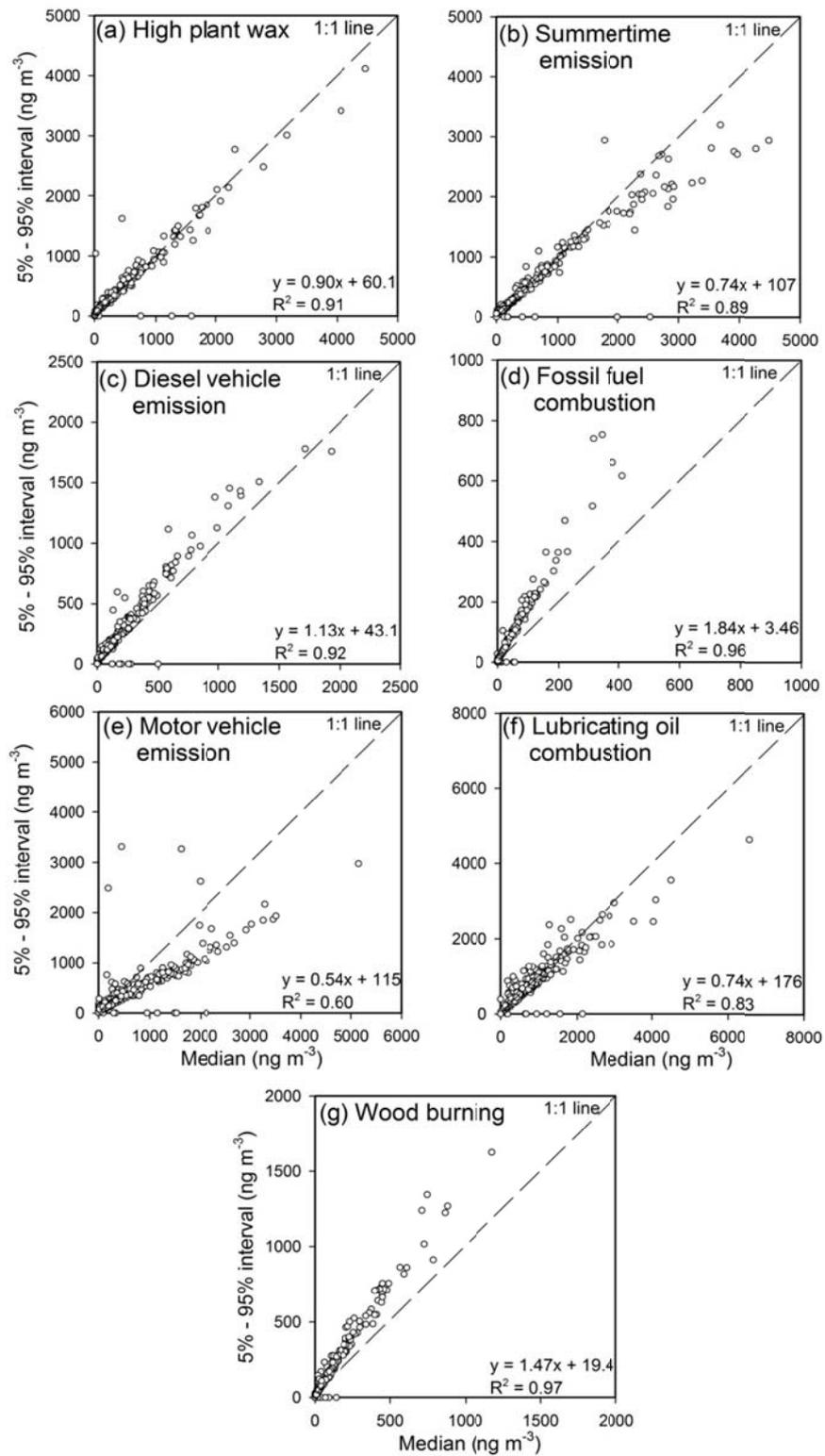


Figure 3.3 Comparison of median factor contributions versus uncertainties represented by 5<sup>th</sup> – 95<sup>th</sup> percentile range of factor contributions from bootstrap solutions for each factor.

### 3.3.2 Spatial variability of factor contributions

The spatial divergences of source contributions were assessed using  $r$  and COD. The median contribution derived from bootstrapped PMF solutions were used instead of base case for a more robust comparison. In Table 3.3, the contributions from the summertime emission factor exhibited the highest correlations ( $r = 0.74 - 0.88$ ) and the lowest spatial divergence (COD = 0.32 - 0.38), consistent with the behavior of the secondary sulfate contributions observed in other studies (Kim et al., 2005; Kim and Hopke, 2008; Mooibroek et al., 2011). The contributions from the high plant wax factor and the wood burning factor showed the next two highest correlations ( $r = 0.70 - 0.87$  and  $0.65 - 0.78$ ). However, the wood burning contributions demonstrated higher divergence (COD = 0.44 - 0.58) than the high plant wax contributions (COD = 0.41 - 0.43), with the lowest divergence for wood burning contributions between two residential sites closer to wood burning sources. Contributions from the fossil fuel combustion factor were reasonably correlated ( $r = 0.54 - 0.76$ ) but exhibited the highest spatial divergence (COD = 0.45 - 0.62), suggesting a similar emission schedule coupled with different transport conditions across the four sites. The other three traffic related sources were moderately correlated ( $r = 0.47 - 0.55$  on average) in factor contribution time series. The lowest CODs were all observed between the two near-traffic sites (ALS and MAS). The highest correlation for the diesel vehicle emission factor contribution, which was observed between the two residential sites (PAL and EDI), might be driven by the wintertime peak contributions. Contributions from the lubricating oil combustion factor exhibited higher correlations on average ( $r = 0.55$ ) and lower spatial divergence (COD = 0.35 - 0.45) than the other two traffic related source categories, because the sources associated with lubricating oil combustion are more similar across the four sites (mostly gasoline vehicles). However, COD values could be affected by the uncertainty

associated the median daily factor contributions derived from bootstrapped PMF solutions. In Figures 3.2 and 3.3, fossil fuel combustion, wood burning and diesel vehicle emission had lower absolute factor contributions and higher variability in median factor contribution (depicted by regression slopes) than others, which might be another explanation for the higher CODs.

Table 3.3 Pearson correlation coefficients ( $r$ ) and coefficients of divergence (COD) for median factor contributions obtained at the four sampling sites.

Factors	EDI vs. PAL		MAS vs. ALS		EDI vs. MAS		EDI vs. ALS		MAS vs. PAL		ALS vs. PAL	
	$r$	COD										
High plant wax	0.87	0.43	0.72	0.41	0.79	0.43	0.70	0.42	0.76	0.42	0.78	0.42
Summertime emission	0.85	0.33	0.85	0.33	0.88	0.38	0.86	0.32	0.87	0.36	0.74	0.32
Diesel vehicle emission	0.68	0.48	0.38	0.43	0.40	0.60	0.47	0.59	0.38	0.53	0.50	0.55
Fossil fuel combustion	0.62	0.62	0.72	0.45	0.54	0.58	0.62	0.59	0.76	0.60	0.69	0.62
Motor vehicle emission	0.44	0.51	0.65	0.42	0.54	0.46	0.62	0.52	0.28	0.48	0.41	0.48
Lubricating oil combustion	0.60	0.42	0.71	0.35	0.58	0.41	0.41	0.41	0.52	0.40	0.50	0.45
Wood burning	0.77	0.44	0.71	0.53	0.78	0.53	0.68	0.58	0.65	0.54	0.67	0.53

The apportionment of average total EC and OC mass to identified factors are shown in Figure 3.4. In general, all sites showed a consistent distribution of factor contributions to EC and OC mass. Most of the EC mass was apportioned to the lubricating oil combustion factor (48.7% – 53.3%) and motor vehicle emission (19.3% – 24.1%) with small difference across the four sites. Similar results were also obtained by Dutton et al. (2010b) at PAL. Those two factors also contributed significantly to OC concentrations (43.2% – 52.5%). The other two dominant OC factors are summertime emissions and high plant wax emissions; together they accounted for a large percentage of OC mass (38.1% – 46.9%). Very few studies have examined the spatial variability of source contributions to organic particulate matter. Von Schneidmesser et al. (2010b) compared carbonaceous aerosol concentrations and sources at two sites in East and West Jerusalem. They found similar temporal trends in source contribution of secondary organic aerosol and vegetative detritus between the two sampling sites, but large spatial difference in

those of fossil fuel combustions, consistent with the observations in this work. Unlike those studies in the Middle East (Von Schneidemesser et al., 2010a, b), biomass burning in Denver did not have a significant contribution to carbonaceous aerosols.

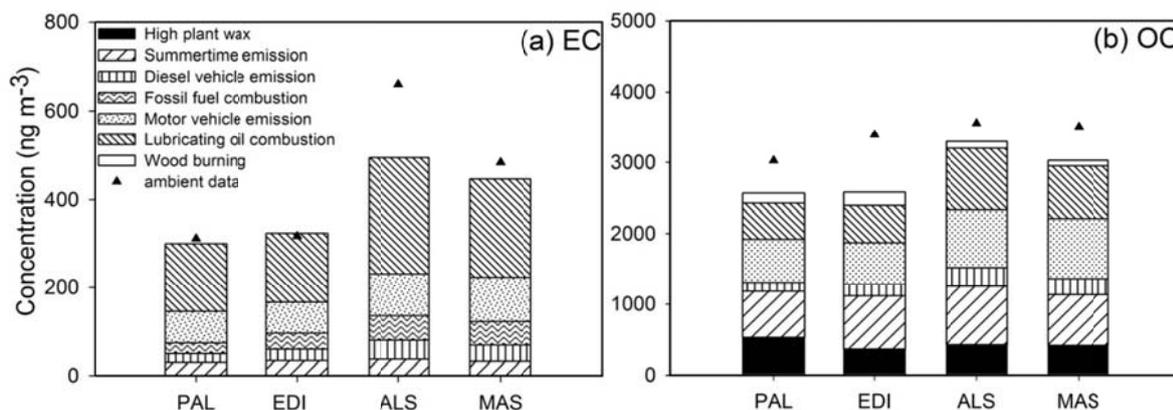


Figure 3.4 Distribution of factor contributions to bulk carbon components including (a) EC, (b) OC at each site. Apportionment results are calculated on the annual average using median daily contributions of bootstrap solutions. The overall height of each bar represents the average concentration of EC or OC estimated from the model, and can be compared to the average measured concentrations represented by the triangles.

### 3.3.3 Comparison between PMF solutions of pooled and site specific data sets

In order to evaluate the influence of differences in source types between receptor sites, we performed source apportionment on samples at each individual receptor site and compared the PMF solutions to those for the pooled data set. As with the pooled data set, an optimal seven-factor simulation was determined for each site; however, the success rates (18.5% – 26.6%) of neural network factor matching decreased significantly, likely due to the limited number of samples available for individual sites (Table 3.4).

Table 3.4 Simulation statistics for individual data sets.

Sampling Sites	PAL	EDI	ALS	MAS
No. of samples	63	59	63	62
No. of factors	7	7	7	7
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 not uniquely matched	815	784	768	734

The distributions of factor profiles and contributions (expressed as reconstructed carbonaceous mass (EC + OC)) at each site, derived from PMF using the bootstrapping approach, are summarized in Figures 3.S4-1 to 3.S4-7 and 3.S5. The correlation coefficients ( $r$ ) and coefficients of divergence (COD) in Figure 3.S5 demonstrate the relationship between the median factor contributions, derived from the site-specific data sets and those from the pooled data set. Generally, all sampling sites exhibited source profiles that were similar to those derived from the pooled data set (Figure 3.S4), and high correlations ( $r = 0.79 - 1.00$ ) between factor contributions derived from the site-specific data sets and those from the pooled data set. These results suggest that the assumption of similar source types across the four sites is applicable for this work.

However, the factor contributions of reconstructed carbonaceous mass show high divergence between site specific and pooled data set solutions, especially at the two residential sites (COD = 0.30 – 0.97 for PAL and 0.25 – 0.63 for EDI). One explanation is that unlike organic molecular markers, EC and OC could be related and apportioned to multiple factors, and a small difference in factor profile distributions can correspond to significant divergence in factor contributions. As reported by Zhang et al. (2009), molecular marker-based PMF results obtained from small subsets ( $N \sim 50 - 60$ ) of a larger set of data were similar to those from the full data set, but the use of the subset of data led to large uncertainties in OC factor profiles.

Moreover, the median values of factor profiles and contributions for the pooled data set in this work were derived from 769 bootstrapped PMF solutions with successful factor matching to the base case solution, while the site-specific data sets had fewer matching solutions (185 – 266 PMF solutions out of 1000 trials using replicate data sets) (Table 3.4). Thus the PMF solutions for the site-specific data sets have greater uncertainties than those for the pooled data set. Therefore, the high divergences in factor contributions could be ascribed to the large uncertainties in EC and OC apportionment for the site-specific PMF solutions.

### 3.4 CONCLUSIONS

Source apportionment was performed on the daily PM<sub>2.5</sub> carbonaceous speciation data collected during March 1, 2008 – March 14, 2009 from four sites. A seven-factor solution was identified using PMF2, coupled with a bootstrap technique to assess uncertainty associated with daily factor contributions. The seven factors could be mainly associated with high plant wax, summertime emission, diesel vehicle emission, fossil fuel combustion, motor vehicle emission, lubricating oil combustion and wood burning. Variability of daily factor contributions derived from bootstrapped PMF solutions was highly correlated with the corresponding median value ( $r = 0.77 - 0.98$ ), which indicated that the relative uncertainty for a factor was relatively stable. Factor contribution of summertime emission exhibited the highest correlation ( $r = 0.74 - 0.88$ ) and lowest CODs (CODs = 0.32 – 0.38) between each site pair among all resolved factors, while higher spatial variability was observed for those traffic related factors. Source contributions to average total EC and OC mass were similarly distributed across the four sites. Moreover, consistent PMF source profiles were obtained when we performed source apportionment for site specific data sets, and the median factor contribution of each factor at each site were highly correlated with that derived from the pooled data set solution ( $r = 0.79 - 1.00$ ). These results suggest that characterizing organic PM<sub>2.5</sub> concentrations at a single well-chosen site in Denver is adequate to reasonably assess relative source contributions for the urban area.

## **CHAPTER 4 POSITIVE MATRIX FACTORIZATION OF PM<sub>2.5</sub>: COMPARISON AND IMPLICATIONS OF USING DIFFERENT SPECIATION DATA SETS**

Mingjie Xie<sup>a</sup>, Michael P. Hannigan<sup>a</sup>, Steven J. Dutton<sup>b</sup>, Jana B. Milford<sup>a</sup>, Joshua G. Hemann<sup>a</sup>,  
Shelly L. Miller<sup>a</sup>, James J. Schauer<sup>c</sup>, Jennifer L. Peel<sup>d</sup>, Sverre Vedal<sup>e</sup>

<sup>a</sup> Department of Mechanical Engineering, College of Engineering and Applied Science,  
University of Colorado, Boulder, CO 80309, USA

<sup>b</sup> National Center for Environmental Assessment, U.S. Environmental Protection Agency,  
Research Triangle Park, NC 27711, USA

<sup>c</sup> Environmental Chemistry and Technology Program, College of Engineering, University of  
Wisconsin-Madison, WI 53706, USA

<sup>d</sup> Department of Environmental and Radiological Health Sciences, Colorado State University,  
Fort Collins, CO 80523, USA

<sup>e</sup> Department of Environmental and Occupational Health Sciences, School of Public Health,  
University of Washington, Seattle, WA 98195, USA

#### 4.0 ABSTRACT

To evaluate the utility and consistency of different speciation data sets in source apportionment of PM<sub>2.5</sub>, positive matrix factorization (PMF) coupled with a bootstrap technique for uncertainty assessment, was applied to four different 1-year data sets composed of bulk species, bulk species and water soluble elements (WSE), bulk species and organic molecular markers (OMM) and all species. The five factors resolved by using only the bulk species best reproduced the observed concentrations of PM<sub>2.5</sub> components. Combining WSE with bulk species as PMF inputs also produced five factors. Three of them were linked to soil, road dust and processed dust, and together contributed 26.0% of reconstructed PM<sub>2.5</sub> mass. A 7-factor PMF solution was identified using speciated OMM and bulk species. The EC/sterane and summertime/selective aliphatic factors had the highest contributions to EC (39.0%) and OC (53.8%) respectively. The nine factors resolved by including all species as input data are consistent with those from the previous two solutions (WSE and bulk species, OMM and bulk species) in both factor profiles and contributions ( $r = 0.88 - 1.00$ ). The comparisons across different solutions indicate that the selection of input data set may depend on the PM components or sources of interest for specific source-oriented health study.

#### 4.1 INTRODUCTION

Airborne particulate matter with a diameter less than 2.5  $\mu\text{m}$  ( $\text{PM}_{2.5}$ ) has been associated with increased rates of daily mortality and morbidity for cardiovascular and respiratory disease (U.S. EPA, 2009). To develop control strategies for  $\text{PM}_{2.5}$ , receptor models have been applied to quantitatively apportion  $\text{PM}_{2.5}$  to sources using ambient speciation measurements taken from one or more receptor sites. Commonly used receptor-based source apportionment models include Principal Component Analysis (PCA)/Multiple Linear Regression (MLR) (Thurston and Spengler, 1985), UNMIX (Ronald, 2003), Positive matrix factorization (PMF) (Paatero and Tapper, 1994) and Chemical Mass Balance (CMB) (Watson et al., 1990). Comparisons of source apportionment results using different receptor models suggest that major  $\text{PM}_{2.5}$  sources could be identified consistently regardless of the model, while the degree of agreement in source contribution estimates depended on the source type, receptor sites and input data (Shrivastava et al., 2007; Bullock et al., 2008; Lee et al., 2008). As such, when designing a source apportionment study, the selection of appropriate receptor model depends on the available ambient data and source information (Larson and Baker, 2003; Lee et al., 2008).

Determining the type of chemical analysis to undertake is a critical decision in study design. However, very few studies have examined the consistency of source apportionment results from different speciation data sets using a single receptor model. Shrivastava et al. (2007) used PMF to investigate source apportionment results derived from 21 different combinations of input species, including both elemental and organic compounds. Each of the 21 data sets was developed by adding or deleting just a few species. Most PMF solutions in that study contained similar source-related factors; six of the factors were associated with primary emissions and one with biogenic secondary organic aerosol (SOA) that contributed more than 50% of the

summertime OC. The similarity between factors across PMF solutions might be caused by the fact that most data sets included pivotal species in common for each of the factors. On the other hand, some modifications were more influential. For example, including nitrate in the PMF analysis resulted in mixing of emissions from multiple source classes, as nitrate co-varied with hardwood markers.

The Denver Aerosol Sources and Health (DASH) study was designed to investigate associations between human health effects and sources of PM<sub>2.5</sub> in Denver, CO (Vedal et al., 2009). Daily filter samples were collected at one centrally located site from July 2002 through October 2008. The collected samples were speciated for inorganic ions, elemental carbon (EC), organic carbon (OC), and organic molecular markers (OMM) (Dutton et al., 2009a, b). A one year subset of the samples collected in 2003 was also analyzed for water soluble organic carbon, water soluble nitrogen and water soluble elements (WSE). The time series of chemical component concentrations have been used previously for source apportionment (Zhang et al., 2008; Dutton et al., 2010b).

Previous studies have used bulk species (e.g., nitrate, sulfate, EC and OC) and elements (Kim et al., 2005; Liu et al., 2006; Hwang et al., 2008; Mooibroek et al., 2011) or OMM (Jaeckels et al., 2007; Schnelle-Kreis et al., 2007; Shrivastava et al., 2007) as inputs to receptor models to determine the source contributions to the ambient PM<sub>2.5</sub>. In this work, all three types of speciation data were obtained for one year (2003) of daily PM<sub>2.5</sub> samples. To evaluate the utility of different speciation data sets, we applied the PMF2 model, coupled with a bootstrap technique for uncertainty assessment developed by Hemann et al. (2009), to four different data sets composed of (1) bulk species, (2) bulk species and WSE, (3) bulk species and OMM and (4) all species combined. Since these speciation data sets are obtained from the application of

different chemical analysis tools that are typically chosen at the study design phase, the comparison of the PMF solutions from these different data sets provides insight for the design of future source-health studies.

## 4.2 METHODS

### 4.2.1 Ambient sample collection and chemical characterization

Daily filter samples were collected at one centrally-located site on the roof of a two-story elementary school located in a residential neighborhood of Denver, far (~ 5 km) from any industrial facilities and major interstate highways (Vedal et al., 2009). Details of the sampling set up, protocols, and chemical analysis were discussed by Dutton et al. (2009a, b). Briefly, PM<sub>2.5</sub> was separated by passing the airstream through a cyclone with 2.5 μm size cut. After the cyclone, the flow was split into two channels, with 20 L min<sup>-1</sup> passing through a Teflon (PTFE) filter and 72 L min<sup>-1</sup> passing through a quartz fiber filter. Teflon filters were used for analysis of gravimetric mass (microbalance), inorganic ions (ion chromatography, IC) and WSE (inductively coupled plasma/mass spectrometry, ICP-MS). Bulk EC, OC (Sunset Laboratories Thermal Optical Transmission Carbon analyzer) and OMM (gas chromatography/mass spectrometry, GC-MS) were measured from the quartz fiber filters. In the current work, the total organic carbon was divided into five parts (OC\_PK1, OC\_PK2, OC\_PK3, OC\_PK4, PC), representing the carbon measured at four distinct temperature steps (340, 500, 615 and 900 °C) with a pyrolyzed carbon (PC) adjustment in the first heating cycle of the NIOSH 5040 thermal optical transmission (TOT) method (NOISH, 2003; Schauer et al., 2003). The root sum of squares (RSS) method was used for uncertainty propagation to derive the final pointwise, blank corrected species concentration uncertainties that were input to PMF (Dutton et al., 2009a, b). Hourly carbon monoxide (CO) concentrations (ng/m<sup>3</sup>) measured at a community monitoring station (AQS ID: 080013001) were obtained from the Colorado Department of Public Health and Environment (CDPHE). One standard deviation of CO concentrations during the whole period was used as the uncertainty for each measurement. Table 4.S1 (Supporting Information) lists the

statistics for each species quantified from January 27, 2003 to December 31, 2003, including the number of valid measurements, mean and median concentrations, coefficient of variation (CV = standard deviation/mean concentration), signal to noise ratio (S/N = mean concentration/mean uncertainty) and the percent of missing measurements and observations below detection limit (BDL).

#### 4.2.2 Source apportionment model and uncertainty assessment

PMF2 (Paatero, 1998a, b) was used in the robust mode as the source apportionment tool for this study. PMF model does not require quantitative information on source profiles and is well suited for long time-series measurements (Dutton et al., 2010b). Positive matrix factorization is a multivariate receptor model resolved by minimizing the sum of the squared, scaled residuals (Q). The behavior of the Q value as a function of the rotational parameter FPEAK ( $\phi$ ) has been used to provide insight into the rotational stability of modeling results, with a lower Q value corresponding to a more stable PMF solution (Paatero et al., 2002). In the current work, PMF solutions for different data sets with optimum-selected factor number showed indistinguishable Q values for  $\phi$  ranging from -0.2 to 0.3; no significant effects on factor profiles and contributions of PMF solutions were observed with  $\phi$  in the same range. Thus  $\phi$  was set at zero for all PMF runs reported here.

The method for uncertainty assessment of PMF solutions was applied as described by Hemann et al. (2009) and as applied in a previous DASH study (Xie et al., 2012b). Briefly, 1000 replicate data sets were generated from the original data set by resampling blocks of samples (block size is chosen algorithmically) with replacement using a stationary block bootstrap technique (Politis and Romano, 1994), and then each was independently analyzed with PMF2. Multilayer Feed Forward Neural Networks were trained to sort and align the factors from each

PMF bootstrap solution to the factors found in the base case, by matching factor profiles. The PMF bootstrap solution was collected for use in subsequent analysis only when each factor could be uniquely matched to a base case factor. The measurement days resampled in each bootstrapped data set were tracked to examine the bias and variability in the PMF solution for each factor on each day. The factor number ( $p$ ) was determined using two criteria: the interpretability of resultant PMF factor profiles and the success rate in factor matching for the bootstrap runs. This bootstrap uncertainty assessment method does not account for the uncertainty in the results due to model assumptions, like factor number, errors in measurement and pollutant species selection (Hemann et al., 2009).

#### 4.2.3 Development of speciation subsets

In this work, PM<sub>2.5</sub> speciation data include bulk carbon (EC, OC), inorganic ions (nitrate, sulfate and ammonium), WSE (30 species), and OMM (70 species). The first data set (labeled “Bulk”) for source apportionment was comprised of bulk species including EC, 4 OC fractions (OC\_PK1, OC\_PK2, OC\_PK3 and PC), nitrate, and sulfate, and thus would represent a study design in which an EC/OC analyzer and an ion chromatograph were used as the chemical analysis tools. In addition, the readily available trace gas CO measurement was included in the Bulk data set. The second data set (labeled “WSE+Bulk”) was constructed by adding the 30 WSE species to the first data set, which requires the addition of ICP-MS analysis. The third data set (labeled “OMM+Bulk”) combined the first data set with 50 selected OMM, which requires the addition of GC-MS analysis for trace organics. A combined data set (labeled “Composite”) including all species was also analyzed to verify a consistency of PMF results between the Composite data set and the WSE+Bulk and OMM+Bulk data sets. Species were selected based on their relationship with specific sources, percentage of missing values and measurements

below detection limit (BDL) (> 50% removed), and signal to noise ratios (S/N < 3 removed). Nitrate, PC, Ti, Zn, syringaldehyde and CO are exceptions that were included despite falling outside our standard criteria because they are important markers for secondary and primary emissions. After preliminary bootstrap runs, some species were also removed from subsequent analysis based on their effects on factor matching rates (e.g., including odd *n*-alkanoic acids and xanthone resulted in much lower factor matching rates (~50%) due to their high variability in factor profiles).

Two days (February 18 and March 22, 2003) without PM<sub>2.5</sub> samples and two days with many missing species observations (September 25 and November 9, 2003) were omitted for all data sets from PMF analysis. Three days (July 3-5, 2003) with extremely high values of several metals due to heavy firework activity were omitted from the WSE+Bulk and Composite data sets, but not from the Bulk and OMM+Bulk data sets. In contrast, samples with a small number of missing species observations and individual species with BDL values were handled using an approach suggested by Polissar et al. (1998). Missing species concentrations were replaced by the geometric mean of the remaining observations, and their accompanying uncertainties were assigned values four times the geometric mean. The species BDL concentrations were replaced by half the detection limit, with uncertainties set at five-sixths of the detection limit.

## 4.3 RESULTS AND DISCUSSION

### 4.3.1 PMF results for different data sets

Parameters of PMF source apportionment for each data set are listed in Table 4.1. Five factors were identified for both the Bulk and WSE+Bulk data sets with factor matching rates of 98.0% and 96.7%. A seven-factor solution was determined for the OMM+Bulk data set with a factor matching rate of 75.6%. The Composite data set solution had the biggest factor number ( $N = 9$ ) but lowest factor matching rates (30.7%). The number of factors was selected to ensure high factor matching rates as well as yielding the most physically interpretable results. For example, the factor matching rates of 6- and 8- factor solutions for the OMM+Bulk data set were 50.5% and 27.6% respectively, much lower than that for the 7-factor solution. Factor profiles obtained for each data set were normalized by

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

where  $F_{kj}^*$  is the relative weighting of species  $j$  in factor  $k$  to all other factors, and are shown in Figure 4.1 (median  $\pm$  standard deviation) for the Bulk, WSE+Bulk and OMM+Bulk data sets. The normalized factor profiles of the Composite data set solution are shown in Figure 4.S1, because they are similar to those derived from WSE+Bulk and OMM+Bulk data sets but with a much lower factor matching rate. Factor contribution time series (median  $\pm$  standard deviation) derived from PMF bootstrap solutions are shown in Figures 4.S2 – S5 in the Supporting Information. Figures 4.S6 – S8 present the factor contributions by day of the week for Bulk, WSE+Bulk and OMM+Bulk data set solutions. Table 4.2 shows annual average factor contributions to major  $PM_{2.5}$  components that can be compared to the observed average concentrations, not including the results derived from the Composite data set. The median CV

values were estimated by dividing the standard deviation of factor contribution derived from bootstrapped PMF solutions by median factor contribution of each factor, representing the variability due to random sampling error in factor contribution estimates.

Table 4.1 Simulation statistics for all data sets.

	Data sets			
	Bulk	WSE+Bulk	OMM+Bulk	Composite
Species	EC, OC fractions, ions, CO	30 water soluble elements + Bulk data set	50 organic molecular markers + Bulk data set	WSE + OMM + Bulk data set
No. of species	8	38	58	88
No. of samples	335	332	335	332
No. of factor	5	5	7	9
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 not uniquely matched	20	33	244	693

Table 4.2 Annual average source apportionment results for PM<sub>2.5</sub> components,  $\mu\text{g m}^{-3}$  (%) and median coefficient of variation (CV) of factor contributions.

Factors	PM <sub>2.5</sub> mass <sup>a</sup>	EC	OC <sup>b</sup>	Nitrate	Sulfate	WSE	OMM	CV <sup>c</sup>
Bulk data set								
Nitrate	0.93 (17.1)	0.022 (4.34)	0.037 (1.32)	0.77 (83.8)	0.097 (8.58)			0.21
Sulfate	1.1 (20.6)	0.010 (2.00)	0.087 (3.08)	0.029 (3.20)	0.99 (87.1)			0.038
EC	0.56 (10.4)	0.36 (70.8)	0.17 (5.95)	0.021 (2.29)	0.010 (0.92)			0.14
OC1	1.2 (22.2)	0.044 (8.64)	1.1 (37.6)	0.083 (9.05)	0.0059 (0.52)			0.10
OC2	1.6 (29.6)	0.073 (14.3)	1.5 (52.1)	0.016 (1.69)	0.033 (2.90)			0.10
Subtotal	5.4 (100)	0.51 (100)	2.8 (100)	0.92 (100)	1.1 (100)			
WSE+Bulk data set								
Soil	0.61 (12.6)	0.0083 (1.78)	0.38 (17.1)	0.00 (0.00)	0.15 (16.4)	0.069 (45.0)		0.29
Road dust	0.25 (5.08)	0.091 (19.5)	0.053 (2.37)	0.081 (7.55)	0.00 (0.00)	0.020 (13.2)		0.59
Processed dust	0.40 (8.31)	0.016 (3.45)	0.25 (11.2)	0.00 (0.00)	0.13 (14.4)	0.0049 (3.17)		0.39
Inorganic ion	1.4 (28.2)	0.034 (7.34)	0.041 (1.85)	0.64 (59.1)	0.63 (69.2)	0.023 (14.8)		0.31
Carbonaceous	2.2 (45.8)	0.32 (67.9)	1.5 (67.4)	0.36 (33.3)	0.00 (0.00)	0.036 (23.8)		0.15
Subtotal	4.8 (100)	0.47 (100)	2.2 (100)	1.1 (100)	0.91 (100)	0.15 (100)		
OMM+Bulk data set								
n-Alkane	0.30 (6.03)	0.015 (3.31)	0.28 (12.0)	0.00 (0.00)	0.0015 (0.14)		0.0098 (19.0)	0.35
Nitrate/PAH	0.33 (6.75)	0.072 (15.4)	0.032 (1.35)	0.23 (21.9)	0.00 (0.00)		0.0043 (8.38)	0.44
Winter/methoxy phenol	0.12 (2.52)	0.00 (0.00)	0.12 (5.30)	0.0001 (0.01)	0.0001 (0.01)		0.0061 (11.8)	0.40
LMW PAH	0.18 (3.66)	0.050 (10.6)	0.13 (5.60)	0.00 (0.00)	0.00 (0.00)		0.0030 (5.79)	1.0
Summer/selective aliphatic	1.5 (30.6)	0.12 (26.1)	1.3 (53.8)	0.00 (0.00)	0.13 (12.0)		0.020 (38.0)	0.22
EC/sterane	0.72 (14.5)	0.19 (39.9)	0.38 (16.1)	0.15 (14.6)	0.00 (0.00)		0.0064 (12.4)	0.22
Inorganic ion	1.8 (35.9)	0.022 (4.64)	0.14 (5.92)	0.67 (63.5)	0.94 (87.8)		0.0024 (4.62)	0.12
Subtotal	4.9 (100)	0.47 (100)	2.3 (100)	1.0 (100)	1.1 (0.00)		0.052 (100)	
Observed	5.6 <sup>d</sup>	0.52	2.9	0.88	1.1	0.15	0.062	

(a) Reconstructed PM<sub>2.5</sub> mass: sum of all species included for source apportionment except CO for the Bulk and WSE+Bulk data sets; not include CO and OMM for the OMM+Bulk data set.

(b) Sum of OC fractions for source apportionment, OC<sub>-</sub> (PK1 + PK2 + PK3) + PC.

(c) CV = standard deviation/median factor contribution.

(d) Sum of EC, OC fractions, nitrate, sulfate and elements concentrations.

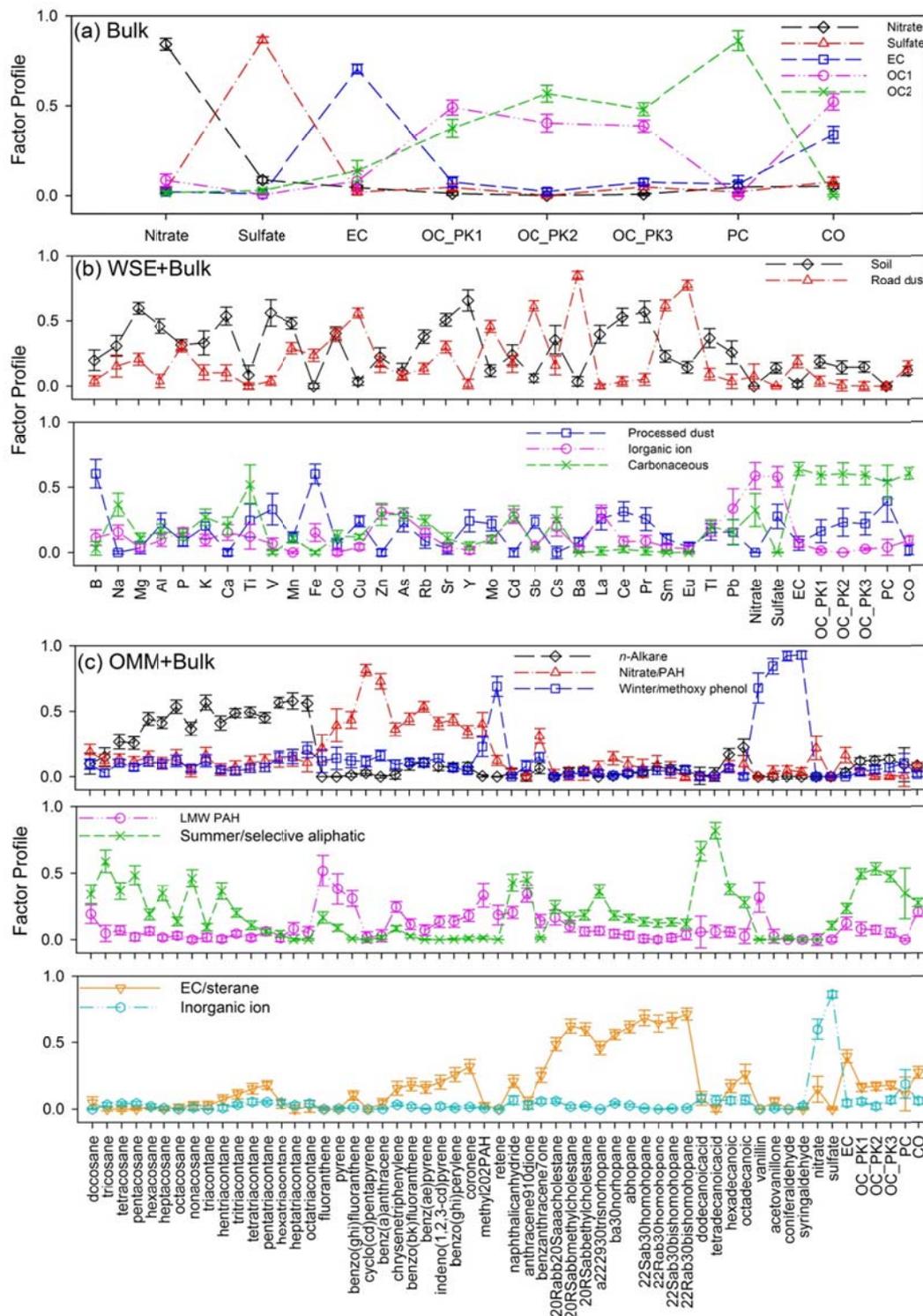


Figure 4.1 Median normalized factor profiles derived from PMF bootstrap solutions for (a) Bulk, (b) WSE+Bulk and (c) OMM+Bulk data sets. The whiskers represent one standard deviation.

For the Bulk data set, five factors were identified and named based on the predominantly loaded species (Figure 4.1a). The nitrate and sulfate factors contributed 83.8% and 87.1% to annual average nitrate and sulfate concentrations, respectively (Table 4.2). The nitrate factor had prominent contributions in winter (Figure 4.S2a), reflecting the effects of temperature on the formation of particulate nitrate (Schaap et al., 2004). The contributions of the sulfate factor were generally higher during warm periods (Figure 4.S2b), due to faster oxidation of sulfur dioxide gas through photochemical mechanisms (Hidy, 1994). However, two spikes of sulfate contributions were observed in early and late winter, driven by short-term atmospheric temperature inversions during cold periods (Neff, 1997). Besides bulk EC mass, the EC factor also contained a significant fraction of CO (Figure 4.1a). EC is formed during the thermal degradation of organic materials (Schauer et al., 2003) and reported to be more associated with diesel- than gasoline-power vehicle emissions (Schauer, 2003; Diaz-Robles et al., 2008). Thus, the CO loaded on the EC factor could be related to incomplete combustion from motor vehicles. The other two factors (OC1 and OC2) approximately split the three OC (OC\_PK1, OC\_PK2 and OC\_PK3) fractions. However, OC1 contained the highest percentage of CO and showed higher factor contributions in fall and winter (Figure 4.S2d), suggesting that the OC fractions in this factor could be more related to combustion. The pyrolyzed OC (PC) was loaded primarily on the OC2 factor, which had prominent contributions in summer (Figure 4.S2e). Therefore, the OC2 factor might be primarily associated with summertime emission sources (e.g., high plant wax, biogenic SOA) (Shrivastava et al., 2007). The OC1 and OC2 factors together accounted for 89.7% of the average total OC concentration.

A 5-factor PMF solution was also obtained by PMF using the WSE+Bulk data set even though more species were added for this analysis. The factors were linked with soil, road dust,

processed dust, an inorganic ion factor and a carbonaceous factor based on their factor profiles. The airborne soil factor was represented by the highest loadings of Mg, Al, Ca, and K (Figure 4.1b), and exhibited higher contributions in warm than in cold seasons (Figure 4.S3a). The significant Pb fraction (25.8%) attributed to this factor is consistent with the fact that lead deposited during years of leaded gasoline use was resuspended (Harris and Davidson, 2005). In Table 4.2, the airborne soil factor contributed only 12.6% to the average reconstructed PM<sub>2.5</sub> mass, similar to the observed airborne soil contributions (6% – 14%) in St. Louis (Kim et al., 2005). However, airborne soil was the primary contributor to total WSE concentration (45.0%). The road dust factor was identified by the large fractions of Co, Cu, Mo, Sb, Ba, Sm and Eu. Copper (Cu) and barium (Ba) are linked with mechanical abrasion of brake and tire wear (McKenzie et al., 2009). Molybdenum (Mo) is used as an additive in lubricant and could be released from motor vehicles (Demydov et al., 2010). Moreover, the road dust factor also contributed 19.5% of the average total EC concentration (Table 4.2). Both the soil and road dust factors exhibited higher contributions on weekdays than weekends (Figure 4.S7), suggesting that human activity on weekdays was the major source of suspended soil and road dust (Hwang et al., 2008). The processed dust factor contained the largest fractions of boron (B) and iron (Fe), and exhibited seasonality with a summertime maximum (Figure 4.S3c). Anthropogenic boron emissions are mainly in the vapor form of boric acid and B halogens (BF<sub>3</sub>, BCl<sub>3</sub> and BBr<sub>3</sub>) (Kot, 2009). The organic components associated with this factor, probably containing more –O<sup>-</sup> and –OH functional groups due to photochemical reactions in summer, could be potential B binders in PM<sub>2.5</sub> (Murphy et al., 1998). Majestic et al. (2007) found that the PM soluble iron fraction was primarily Fe(II), and the soluble Fe(II) in fine particles increased during a short-term aging process (1-3 days); while Fe(III) remained relatively constant. Therefore, the processed dust

factor could be categorized as “aged aerosol”. Similar to the soil and road dust factors, the processed dust factor had low contributions to reconstructed PM<sub>2.5</sub> mass (8.31%, Table 4.2). This might be due to the fact that dust resuspension from anthropogenic sources yields primarily coarse particles and contributes only a small part to the fine modes (Kim et al., 2005; Hwang et al., 2008). Another possible explanation is that only WSE, not the total elements were used for PMF analysis. The inorganic ion factor consisted of secondary nitrate and sulfate. The loadings of Pb, Zn, Cd and As in this factor might be attributed to the co-emission of these elements and SO<sub>2</sub> from industrial facilities (e.g., power plants). The carbonaceous factor accounted for the majority of EC, OC fractions and CO, which could not be linked with any pollution sources due to the lack of organic tracers.

A 7-factor PMF solution was found for the OMM+Bulk data set, including factors labeled *n*-alkane, nitrate/PAH, winter/methoxyphenol, low molecular weight (LMW) PAH, summer/selective aliphatic, EC/sterane and inorganic ion (Figure 4.1c). As shown in Table 4.2, the summertime/selective aliphatic factor contributed the largest portion of the total OC (53.8%) and OMM (38.0%) concentrations. This factor and the inorganic ion factor together accounted for 66.5% of the reconstructed PM<sub>2.5</sub> mass. EC concentration had the highest contribution from the EC/sterane factor (39.0%), followed by the summertime/selective aliphatic (26.1%) and nitrate/PAH (15.4%) factors. In prior work, Dutton et al. (2010b) also obtained a 7-factor solution using a similar data set, including all the species in the OMM+Bulk data set used in this work except CO, with the same 2003 time series. In contrast to Dutton et al. (2010b), the current work excluded the OMM used in the previous study that had high variability in factor profiles, and used OC fractions instead of total OC. Six out of the seven factors in the previous study, not including the middle oxy-PAH/alkanoic acid factor (Dutton et al., 2010b), had similar factor

profiles and contribution time series with the corresponding factors in the current work. The wintertime/methoxyphenol factor in that study also contained several of the LMW PAHs and oxy-PAHs. In this study, LMW PAHs and oxy-PAH were apportioned to a distinct factor (LMW PAH).

A 9-factor solution was determined for the Composite data set. As shown in Figure 4.S1, eight out of the nine factors exhibited similar factor profiles as those resolved by using the WSE+Bulk and OMM+Bulk data sets; the remaining factor profile appears to be a combination of the processed dust and inorganic ion factors derived from the WSE+Bulk data set. The median factor contribution time series of the nine factors were strongly correlated ( $r = 0.88 - 1.00$ ) with the corresponding factors resolved by using the WSE+Bulk and OMM+Bulk data sets (Figure 4.S9); and the regression slopes were generally close to unity ( $0.62 - 1.35$ ). Therefore, combining all speciation data gave similar results as those obtained by using subsets of species individually. Christensen and Schauer (2008b) found that organic species could be overshadowed by elements unless their influence is manually inflated by adapting uncertainties. This is not observed for the PMF solution of Composite data set in current work, even for solutions with smaller factor number ( $p = 7$  or  $8$ ), which might be attributed to the different method used for uncertainty estimation. However, the factor matching rate for the Composite data set solution was only 30.7%, much lower than those of the WSE+Bulk or OMM+Bulk data sets. In addition, the processed dust factor found in the WSE+Bulk data set solution could not be resolved using the Composite data set. This Composite data set solution was not used in subsequent analysis, because it did not provide extra factor/source information and had higher variability.

#### 4.3.2 Comparison between PMF estimation and observed concentration

In Table 4.2, the PMF estimation of reconstructed PM<sub>2.5</sub> average mass and bulk components from the Bulk data set are closer to observed values than those from the WSE+Bulk and OMM+Bulk data sets. Average nitrate concentrations were overestimated by PMF using all data sets; however the difference was small, for example only 4.1% for the Bulk data set. This overestimation is likely an artifact of the use of 0.5×DL for all BDL measurements as well as a result of using an F-Peak of 0, which shifts the PMF solutions away from solutions with frequent null contributions from specific sources.

Comparisons between PMF estimated and observed time series of PM<sub>2.5</sub> components are shown in Figure 4.S10. Pearson correlation coefficients (*r*) and coefficients of divergence (COD) were also calculated to assess the resemblance between estimated and observed time series. The COD is defined as:

$$\text{COD} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left( \frac{x_i^{\text{estimated}} - x_i^{\text{observed}}}{x_i^{\text{estimated}} + x_i^{\text{observed}}} \right)^2} \quad (2)$$

where  $x_i^{\text{estimated}}$  and  $x_i^{\text{observed}}$  are estimated and observed concentrations of a specific component for the  $i^{\text{th}}$  time period, and  $n$  is the number of observations. Values approaching 0 represent uniformity between estimated and observed concentrations, while values approaching 1 represent strong divergence (Wongphatarakul et al., 1998; Kim et al., 2005). Among all PM<sub>2.5</sub> components, reconstructed PM<sub>2.5</sub> mass exhibited the greatest similarity between estimated and observed time series ( $r = 0.90 - 1.00$ ,  $\text{COD} = 0.02 - 0.11$ ), followed by sulfate ( $r = 0.93 - 1.00$ ,  $\text{COD} = 0.02 - 0.15$ ) and EC ( $r = 0.83 - 1.00$ ,  $\text{COD} = 0.02 - 0.17$ ). The least similarity was obtained for nitrate concentrations from the WSE+Bulk ( $r = 0.81$ ,  $\text{COD} = 0.41$ ) and OMM+Bulk ( $r = 0.81$ ,  $\text{COD} = 0.40$ ) data sets. In the current work, the PMF model tended to overestimate low concentrations and underestimate high concentrations of nitrate when using the WSE+Bulk and OMM+Bulk

data sets. Such a tendency could be associated with all linear models and exacerbated by the non-negativity constraint of the PMF model (Henry and Christensen, 2010).

Comparing the PMF estimation across the three data sets, the Bulk data set has the closest correspondence to the observed time series of all bulk species, with the highest  $r$  (0.99 – 1.00) and lowest COD (0.02 – 0.05, Figure 4.S10). Moreover, the median ratio of standard deviation to median factor contribution of each factor, derived from bootstrapped PMF solutions, was lowest for the Bulk data set (CV = 0.038 – 0.21, Table 4.2). These results suggest that the PMF source apportionment using a data set composed of a small number of bulk species can obtain more stable and accurate estimates of observations. However, due to the lack of source markers (e.g., elements, OMM), we cannot match the factors resolved using the Bulk data set to PM<sub>2.5</sub> source classes. The WSE+Bulk and OMM+Bulk data sets had similar degree of correspondence to the contribution time series and average concentrations of bulk species, but the average concentration of total WSE was better estimated than that of total OMM (Table 4.2).

#### 4.3.3 Comparison between PMF solutions

To compare time series of factor contributions resulting from the use of different data sets in PMF, Figure 4.2 presents correlation contour plots, illustrating the Pearson correlation coefficient ( $r$ ) between median factor contribution time series from pairs of PMF solutions. In general, the inorganic ion related factors always show high correlations ( $r = 0.62 – 0.92$ ), due to their dominant contributions to nitrate and/or sulfate. In Figure 4.2a, the carbonaceous factor of the WSE+Bulk data set is much better correlated with the EC factor ( $r = 0.79$ ) than the OC1 ( $r = 0.41$ ) and OC2 ( $r = 0.44$ ) factors of the Bulk data set; both the carbonaceous and EC factors have dominant contributions to average EC mass (Table 4.2), suggesting that the carbonaceous factor contains most of the carbon content from motor vehicle emissions, because motor vehicles are

the primary emission sources for ambient EC (Schauer et al., 2003). Moderately strong correlation ( $r = 0.66$ ) is also observed between the processed dust factor from the WSE+Bulk data set and the OC2 factor from the Bulk data set; however, the OC mass accounted for by the processed dust factor is much less than that of the OC2 factor. Except for the processed dust factor, no other factor resolved by using the WSE+Bulk data set is highly correlated with the OC1 or OC2 factor of the Bulk data set. Similar to EC mass apportionment, most of the OC in the WSE+Bulk data set was also apportioned to the carbonaceous factor. As such, using WSE and bulk species data as inputs to PMF results in limited ability to resolve carbonaceous sources of  $PM_{2.5}$ .

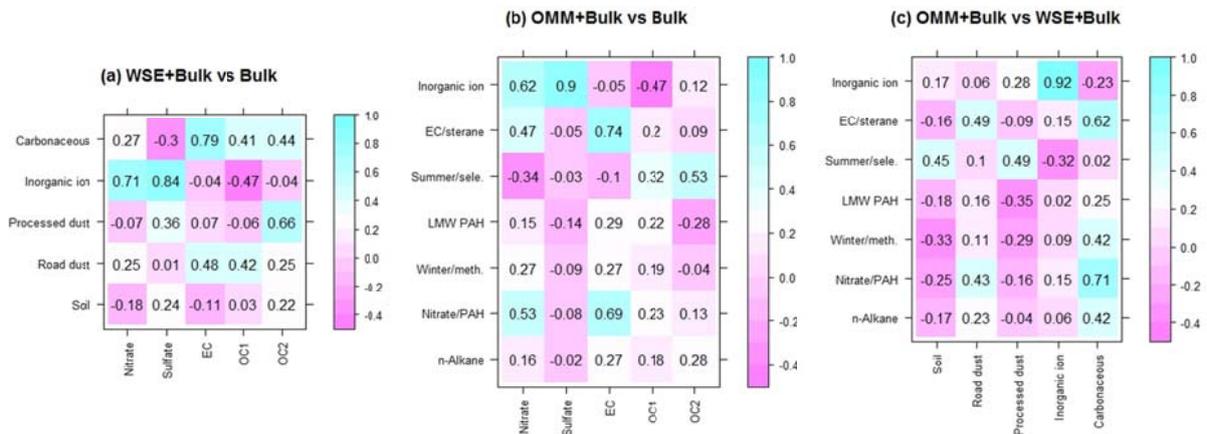


Figure 4.2 Correlation contour plots illustrating the Pearson correlation coefficient ( $r$ ) between median factor contribution time series of each pair of PMF solutions.

In Figure 4.2b, the nitrate factor of the Bulk data set is moderately correlated with the two traffic dominated factors (nitrate/PAH,  $r = 0.53$ ; EC/sterane,  $r = 0.47$ ) of the OMM+Bulk data set, suggesting the ambient nitrate is from the oxidation of  $\text{NO}_2$  primarily emitted by motor vehicles (Mooibroek et al., 2011). The EC factor of the Bulk data set exhibits moderately strong correlations with both nitrate/PAH ( $r = 0.69$ ) and EC/sterane ( $r = 0.74$ ) factors of the OMM+Bulk data set, consistent with the expectation that EC, PAHs and steranes are good markers for motor vehicle emissions in urban areas (Phuleria et al., 2006, 2007). A moderate correlation exists between the OC2 factor of the Bulk data set and the summer/selective aliphatic factor of the OMM+Bulk data set ( $r = 0.53$ ), possibly due to their having summertime biogenic emission and photochemical influences in common. However, none of the factors resolved by the OMM+Bulk data set is strongly correlated with the OC1 or OC2 factor of the Bulk data set. One plausible explanation is that OC mass is apportioned to more specific source-related factors by using the OMM+Bulk data set. Thus, the  $\text{PM}_{2.5}$  associated bulk species, especially the carbonaceous components (EC and OC), could be more effectively resolved with OMM.

In Figure 4.2c, except the inorganic ion factor, good correlation was only observed between the carbonaceous factor of the WSE+Bulk data set and the two traffic dominated factors (nitrate/PAH,  $r = 0.71$ ; EC/sterane,  $r = 0.62$ ) of the OMM+Bulk data set, supporting the dominance of motor vehicle emissions in the carbonaceous factor of the WSE+Bulk data set. Due to the traffic influence, the road dust factor of the WSE+Bulk data set is moderately correlated with the nitrate/PAH ( $r = 0.43$ ) and EC/sterane ( $r = 0.49$ ) factors. Similar correlations also exist between the summer/selective aliphatic factor of the OMM+Bulk data set and the soil ( $r = 0.45$ ) and processed dust ( $r = 0.49$ ) factors of the WSE+Bulk data set, suggesting these three

factors have some sources or processing impacts in common, such as biogenic material or atmospheric photochemistry.

#### 4.4 IMPLICATIONS

In this work, four different PM<sub>2.5</sub> speciation data sets were analyzed using the PMF model. The Bulk, WSE+Bulk and OMM+Bulk data sets correspond to the use of different chemical analysis tools. Some of the factors resolved by these three data sets were consistent as inferred from the moderate to strong correlations in Figure 4.2. In comparing results from the three data sets, the effect of random resampling error on the PMF solutions was accounted for by replacing the base case factor contribution with the median values derived from PMF bootstrap solutions.

The bootstrap approach applied here aids in interpretation of PMF solutions as it provides estimates of variability in factor profiles and contributions. Moreover, this approach also provides an alternative criterion – matching rate of bootstrapped factors to base case factors – for the selection of number of factors. Physical interpretability of PMF solutions is a widely accepted criterion for factor number determination (Kim et al., 2005; Hwang et al., 2008; Mooibroek et al., 2011), which might be difficult when knowledge of PM<sub>2.5</sub> sources is limited or when apportionment of trace-level species yields the only clues that link factors to sources. The underlying principle of the bootstrap approach is that if more factors than necessary are selected, then a bootstrapped factor might be matched with more than one base case factor due to similar factor profiles. In this case, the bootstrapped solution cannot be retained, resulting in a higher fail rate in factor matching. Thus, a high factor matching rate reflects the uniqueness of base case factors, robustness of this solution to the input data set and corresponds to distinguishable PMF results.

Based on the comparisons across PMF solutions using different speciation data sets from Denver, we draw some cautious conclusions here. PMF analysis of a small number of bulk species data provided the closest estimates of observed species concentrations. But none of the

factors resolved could be described as specific PM<sub>2.5</sub> pollution sources due to the lack of unique source markers. Adding speciated WSE to the bulk species data set for PMF modeling led to the identification of three specific source related factors, including soil, road dust and processed dust. However, most of the bulk carbon components (EC and OC) were lumped into one factor. Still, some of the bulk carbon components are attributed to road dust, which makes sense but complicates interpretation of the factors. If road dust was thought to represent the entire contribution from the traffic and roads, then it would be an underestimate and the temporal pattern would be erroneous. If the carbonaceous factor was assumed to represent all the motor vehicle emissions, as well as other carbon sources, then some portion of the traffic source signal contained in the road dust factor would be missed. This problem might only be solved appropriately by combining speciated OMM with bulk species data, as OMM are better source tracers for OC than WSE. On the other hand, if there is a need to resolve the origins of the WSE or to nail down the contribution of soil, then the OMM data set results in poor resolution. Therefore, both speciation data sets can assist in source apportionment, and the choice of input data set may depend on the PM components or sources of interest for a particular source-oriented health study.

**CHAPTER 5 POSITIVE MATRIX FACTORIZATION OF A 32-MONTH SERIES OF  
DAILY PM<sub>2.5</sub> SPECIEATION DATA WITH INCORPORATION OF TEMPERATURE  
STRATIFICATION**

Mingjie Xie<sup>a</sup>, Ricardo Piedrahita<sup>a</sup>, Steven J. Dutton<sup>b</sup>, Jana B. Milford<sup>a</sup>, Joshua G. Hemann<sup>a</sup>,  
Jennifer L. Peel<sup>c</sup>, Shelly L. Miller<sup>a</sup>, Sun-Young Kim<sup>d</sup>, Sverre Vedal<sup>d</sup>, Lianne Sheppard<sup>d</sup>, Michael  
P. Hannigan<sup>a</sup>

<sup>a</sup> Department of Mechanical Engineering, College of Engineering and Applied Science,  
University of Colorado, Boulder, CO 80309, USA

<sup>b</sup> National Center for Environmental Assessment, U.S. Environmental Protection Agency,  
Research Triangle Park, NC 27711, USA

<sup>c</sup> Department of Environmental and Radiological Health Sciences, Colorado State University,  
Fort Collins, CO 80523, USA

<sup>d</sup> Department of Environmental and Occupational Health Sciences, University of Washington,  
School of Public Health, University of Washington, Seattle, WA 98195, USA

## 5.0 ABSTRACT

This study presents source apportionment results for PM<sub>2.5</sub> from applying positive matrix factorization (PMF) to a 32-month series of daily PM<sub>2.5</sub> 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<sub>2.5</sub> mass (sulfate + nitrate + EC + OC) in cold periods (daily average temperature < 10 °C; 43.7% of reconstructed PM<sub>2.5</sub> 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 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<sub>2.5</sub> 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<sub>2.5</sub> 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.

## 5.1 INTRODUCTION

The Denver Aerosol Sources and Health (DASH) study aims to identify and quantify the sources of ambient PM<sub>2.5</sub> that are detrimental to human health given short-term exposure (Vedal et al., 2009). Daily PM<sub>2.5</sub> samples were collected from a centrally located site in Denver, Colorado, over the period from July 2002 through December 2008. Speciation of the PM<sub>2.5</sub> 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<sub>2.5</sub> 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<sub>2.5</sub> sources.

In the current study, a 32-month series of daily speciated PM<sub>2.5</sub> data, including the 1-year data set analyzed previously, was utilized as input for source apportionment to obtain more reliable PM<sub>2.5</sub> 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. Subsets 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.

## 5.2 METHODS

### 5.2.1 Ambient data collection and chemical speciation

PM<sub>2.5</sub> 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<sub>2.5</sub> 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 5.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<sub>x</sub>) and CO). Mean values and standard deviations for these variables are given in Table 5.S2.

### 5.2.2 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.

### 5.2.3 Preparation of PMF input data set

For the whole measurement campaign, only ten days with unavailable PM<sub>2.5</sub> samples or multiple missing species observations were omitted for all PMF analysis. Candidate PMF species were selected from the 79 species listed in Table 5.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 5.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 5.S3 – 5.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.

## 5.3 RESULTS

### 5.3.1 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 5.1). 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 5.1a, 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 5.1b, are expressed as reconstructed PM<sub>2.5</sub> 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 5.1b, agreed with the observations (PMF simulation/Observation = 0.77 – 1.04). Monthly average temperatures are overlaid in Figure 5.1b 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 5.S1 and 5.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 5.S6.

Table 5.1 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

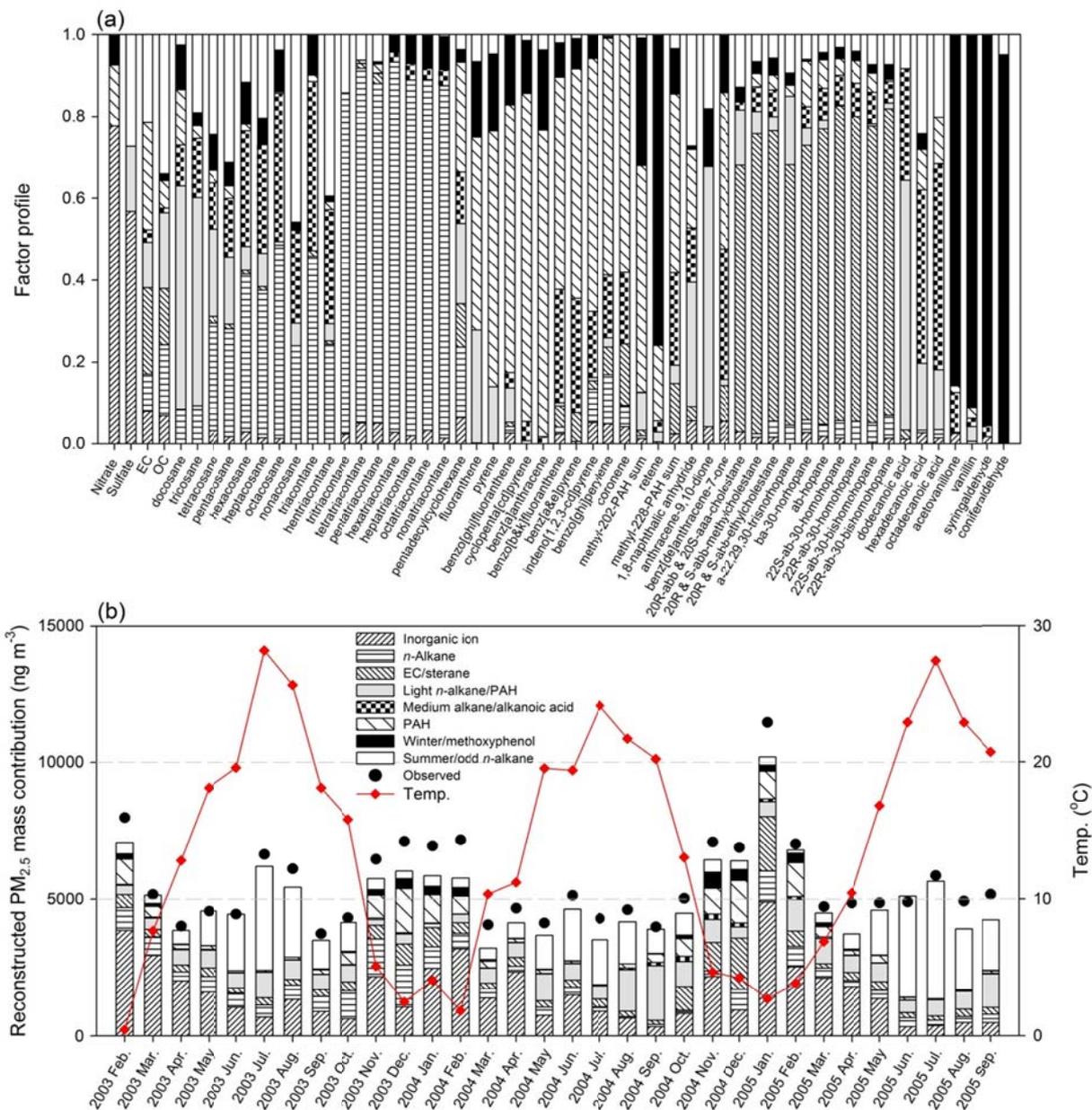


Figure 5.1 Results of PMF analysis for full data set: (a) distributions of species across the 8 factors; (b) monthly average factor contribution to reconstructed PM<sub>2.5</sub> mass overlaid with observed values and ambient temperature. Median factor profiles and contributions were used from bootstrapped PMF solutions.

The factor contribution time series from the full data set were divided into three periods (cold, warm and hot as defined in Section 5.2.3). In each period, the contribution of each factor was averaged and expressed in percentage of reconstructed PM<sub>2.5</sub> mass (Figure 5.2a-c). In cold and hot periods, respectively, the reconstructed PM<sub>2.5</sub> 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 5.S3), so as to illustrate more detailed temporal patterns and assist in the interpretation of each factor as explained in the Discussion section.

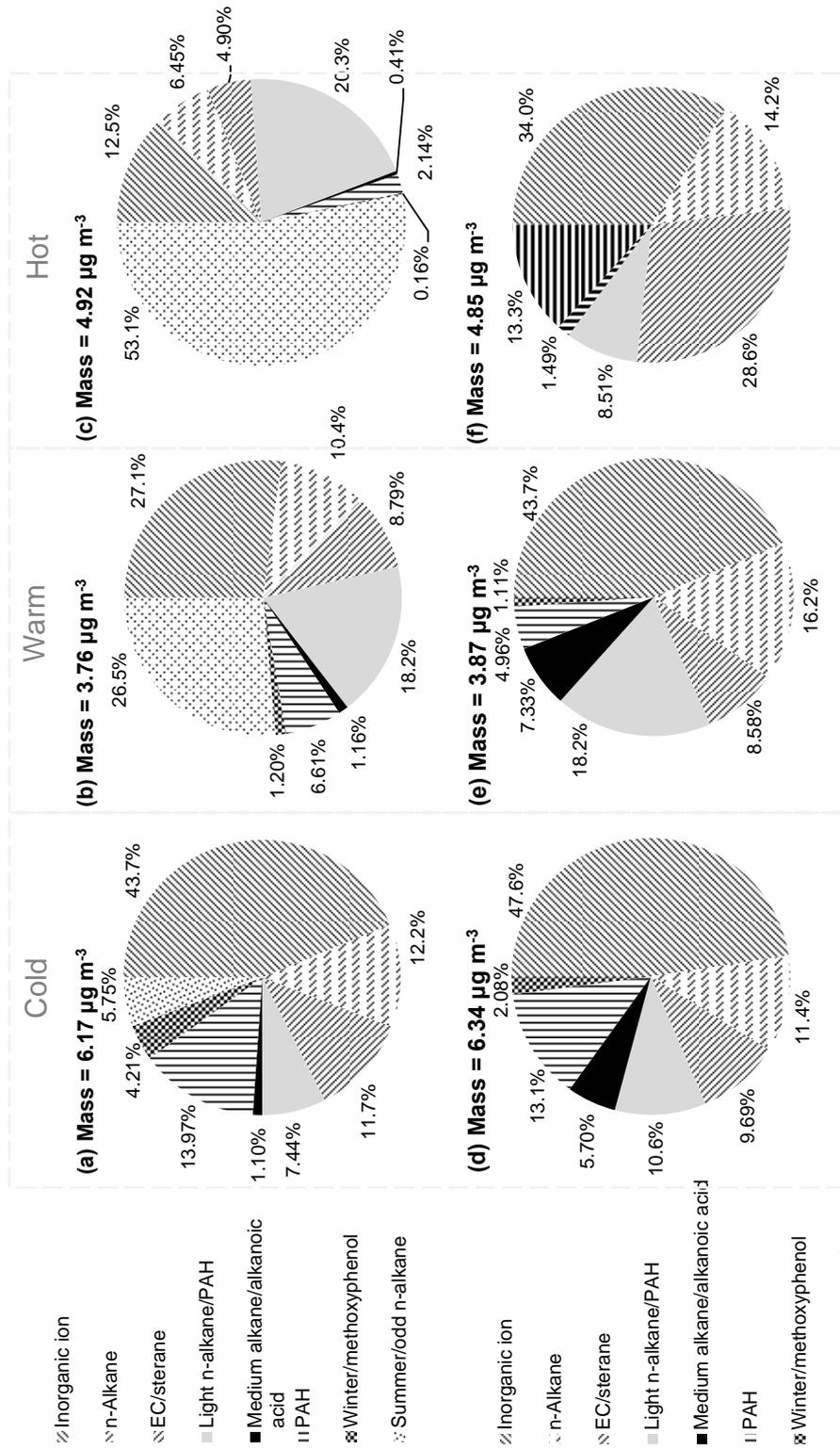


Figure 5.2 Comparison of relative contribution to reconstructed  $\text{PM}_{2.5}$  mass between full data set solution in (a) cold, (b) warm, (c) hot periods and the corresponded sub-data set solutions (d-f).

### 5.3.2 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 5.1. Compared with the full data set, fewer species were included for PMF analysis based on the same screening criteria (Tables 5.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 5.3a-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 5.2d-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 5.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 5.S4 – 5.S6). The sum of factor contributions is also close to that for the full data set solution in each period (Figure 5.2). 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.

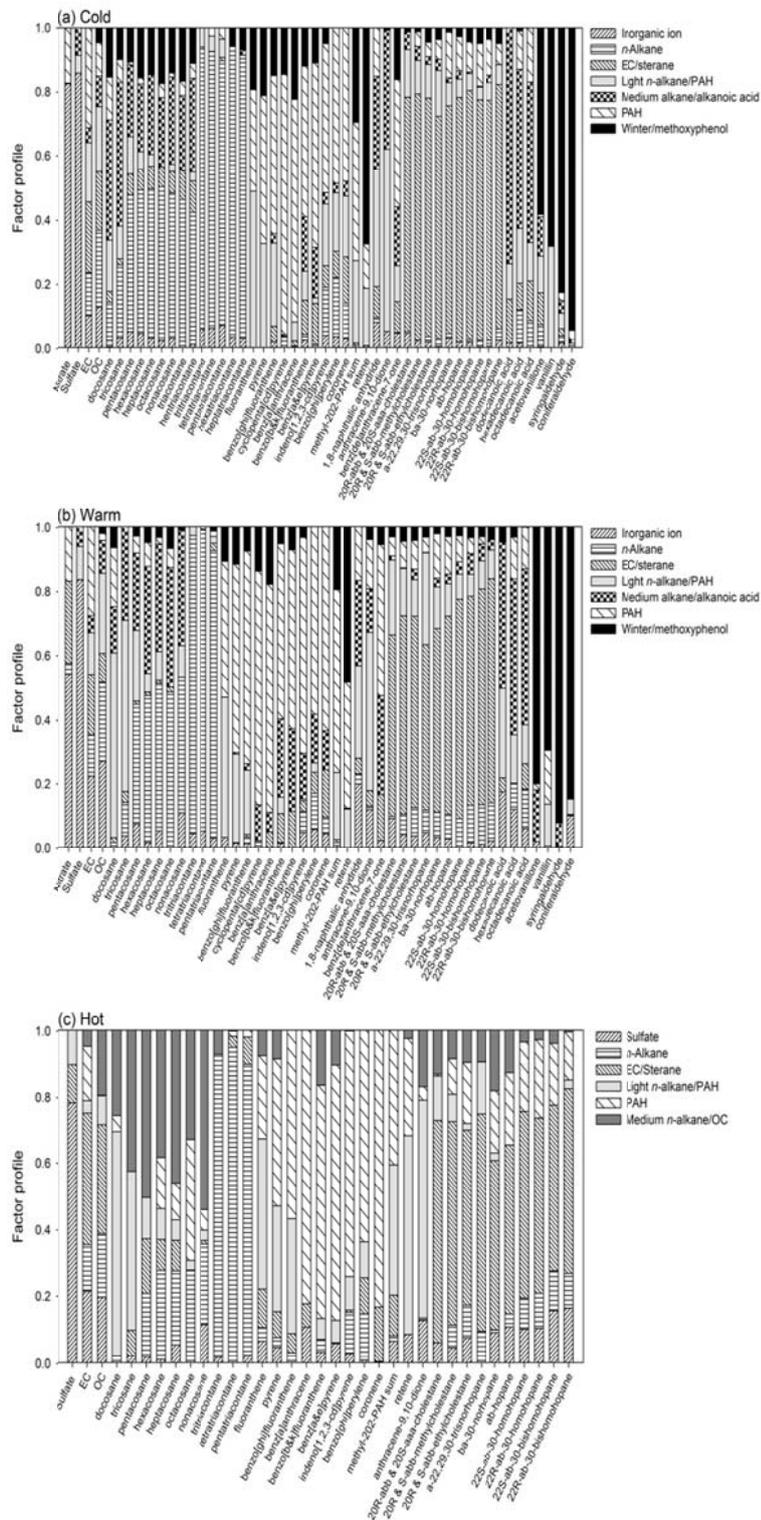


Figure 5.3 PMF factor profiles of each temperature-stratified sub-data set solution, (a) cold, (b) warm, (c) hot.

## 5.4 DISCUSSION

### 5.4.1 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 5.1a), the temporal trends of factor contributions (Figure 5.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 ( $O_3$ ,  $NO_x$  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 5.2 contains Pearson correlation coefficients ( $r$ ) of factor contributions versus all six variables for the full and temperature-stratified time series.

Table 5.2 Correlations between factor contributions and meteorological and trace gas measurements.

Factors	Temperature (°C)	Radiance (KW-hr m <sup>-2</sup> )	RH (%)	Ozone (ppm)	NO <sub>x</sub> (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*
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.

#### 5.4.1.1 Inorganic ion factor

This factor contained the largest fractions of secondary nitrate and sulfate (Figure 5.1a), and dominated the reconstructed PM<sub>2.5</sub> mass on cold days (43.7%, Figure 5.2a). 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 5.2). 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 (Vukmirović, 1986).

#### 5.4.1.2 *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 5.1a). 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<sub>x</sub> and CO further support the association of this factor with motor vehicle traffic (Table 5.2). 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 5.S3b). In addition, the factor contributions were highest in cold months with an enhancement in May and June of the year (Figure 5.S3b). Therefore, this

factor might contain influences from other PM<sub>2.5</sub> contributors besides the road dust, such as meat cooking or primary biogenic emissions.

#### 5.4.1.3 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 5.1a). 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 5.S3c) and the correlations between the factor contributions and NO<sub>x</sub> and CO concentrations in each period (Table 5.2). The relative contribution of this factor to reconstructed PM<sub>2.5</sub> mass in cold periods was more than twice that in hot periods (Figure 5.2a, 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 5.2).

#### 5.4.1.4 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 5.1a). Figure 5.S3d presents the seasonality of this factor contribution with a maximum in late summer and no weekly trend. In Table 5.2, 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.

#### 5.4.1.5 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 5.1a). Similar to the *n*-alkane factor, it was most prominent in winter and exhibited a weekend enhancement (Figure 5.S3e). The factor contribution was positively correlated with NO<sub>x</sub> and CO concentrations, but anti-correlated with ozone (Table 5.2). 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<sub>2.5</sub> mass (Figure 5.2a-c). Therefore, it might be linked with un-apportioned molecular markers from primary sources.

#### 5.4.1.6 PAH factor

This factor was characterized by the presence of EC and the highest loadings of all PAHs except retene (Figure 5.1a). The temporal pattern of factor contributions presented in Figure 5.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<sub>x</sub> and CO concentrations in each period (Table 5.2) 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.

#### 5.4.1.7 Winter/methoxyphenol factor

This factor was characterized by the highest loadings of retene and all four methoxyphenols (Figure 5.1a). These species are typical markers for wood burning (McDonald et al., 2000; Bari et al., 2009). The factor contribution showed wintertime dominance (Figure 5.53g) and a weekend enhancement, and was significantly anti-correlated with ambient temperature (Table 5.2), consistent with residential wood-burning stove use patterns.

#### 5.4.1.8 Summer/odd *n*-alkane factor

This factor contains mainly nonacosane (C<sub>29</sub>) and hentriacontane (C<sub>31</sub>), two characteristic markers from high plant waxes (Rogge et al., 1993c), and the largest fraction of OC (Figure 5.1a). In hot periods, the factor contribution dominated the reconstructed PM<sub>2.5</sub> mass (53.1%) (Figure 5.2c) 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<sub>2.5</sub> (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.

#### 5.4.1.9 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<sub>2.5</sub> (0.54% for 2003). Table 5.3 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 5.3). 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<sub>2.5</sub> speciation data in the current work and different input data can change the output (Zhang et al., 2009). In addition, source patterns of PM<sub>2.5</sub> might change over time, particularly for sources like motor vehicles where the fleet is evolving to have improved combustion and better maintenance.

Table 5.3 PMF factor contribution regression statistics for 2003 derived from the 32-month data set versus a 1-year data set.

Factor		Regression statistics		
Current <sup>a</sup>	Previous <sup>b</sup>	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 <sup>c</sup>	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., 2010b).

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

Table 5.4 PMF factor contribution regression statistics between full data set solution and temperature-stratified sub-data set solutions.

Factor		Cold			Warm			Hot		
Full <sup>a</sup>	Sub <sup>b</sup>	Slope	Intercept	<i>r</i>	Slope	Intercept	<i>r</i>	Slope	Intercept	<i>r</i>
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> -alkane/PAH	Light <i>n</i> -alkane/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 <sup>c</sup>	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.

#### 5.4.2 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 5.1 and 5.3) and weekly patterns of factor contributions (Figures 5.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 5.4), so as to evaluate the dependence of PMF results on temperature.

##### 5.4.2.1 Cold period

For the cold period, the distribution of factor contributions, expressed as reconstructed PM<sub>2.5</sub> mass, for the sub-data set was similar to that for the full data set during the same period (Figure 5.2a, 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 5.4, 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<sub>2.5</sub> mass (Figure 5.2a, d) and good correlations ( $r = 0.76, 0.95$ ) between the full and sub-data set solutions (Table 5.4). 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 5.4). 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 5.S1), and the gas-to-particle partitioning should be enhanced by the lower temperatures. Figure 5.3a 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.

#### 5.4.2.2 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 5.4). However, Figures 5.2b and 5.2e show that the inorganic ion, *n*-alkane, and medium alkane/alkanoic acid factors had larger relative contributions to reconstructed PM<sub>2.5</sub> 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 5.4. 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 5.3b) than in the full data set solution (Figure 5.1a). 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 5.3a, 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.

#### 5.4.2.3 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 5.2c, these two factors had relatively minor contributions to PM<sub>2.5</sub> 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 5.3c). 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 5.2c, 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 5.4, 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.

## 5.5 CONCLUSIONS

This study leveraged a longer time series (32 months) of daily PM<sub>2.5</sub> 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 alkanolic 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<sub>2.5</sub> 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.

**CHAPTER 6 POSITIVE MATRIX FACTORIZATION OF PM<sub>2.5</sub> – ELIMINATING THE  
EFFECTS OF GAS/PARTICLE PARTITIONING OF SEMI-VOLATILE ORGANIC  
COMPOUNDS**

Mingjie Xie<sup>a</sup>, Kelley C. Barsanti<sup>b</sup>, Michael P. Hannigan<sup>a</sup>, Steven J. Dutton<sup>c</sup>, Sverre Vedal<sup>d</sup>

<sup>a</sup> Department of Mechanical Engineering, College of Engineering and Applied Science,  
University of Colorado, Boulder, CO 80309, USA

<sup>b</sup> Department of Civil and Environmental Engineering, Portland State University, PO Box 751,  
Portland, OR 97207, USA

<sup>c</sup> National Center for Environmental Assessment, U.S. Environmental Protection Agency,  
Research Triangle Park, NC 27711, USA

<sup>d</sup> Department of Environmental and Occupational Health Sciences, School of Public Health,  
University of Washington, Seattle, WA 98195, USA

## 6.0 ABSTRACT

Gas-phase concentrations of semi-volatile organic compounds (SVOCs) were calculated from gas/particle (G/P) partitioning theory using their measured particle-phase concentrations. The particle-phase data were obtained from an existing filter measurement campaign (January 27, 2003 – October 2, 2005) as a part of the Denver Aerosol Sources and Health (DASH) study, including 970 observations of 71 SVOCs (Xie et al., 2013b). In each compound class of SVOCs, the lighter species (e.g., docosane in *n*-alkanes, fluoranthene in PAHs) had higher total concentrations (gas + particle phase) and lower particle-phase fractions. The total SVOC concentrations were analyzed using positive matrix factorization (PMF). Then the results were compared with source apportionment results where only particle-phase SVOC concentrations were used (particle only-based study; Xie et al., 2013b). For the particle only-based PMF analysis, the factors primarily associated with primary or secondary sources (*n*-alkane, EC/sterane and inorganic ion factors) exhibit similar contribution time series ( $r = 0.92 - 0.98$ ) with their corresponding factors (*n*-alkane, sterane and nitrate+sulfate factors) in the current work. Three other factors (light *n*-alkane/PAH, PAH and summer/odd *n*-alkane factors) are linked with pollution sources influenced by atmospheric processes (e.g. G/P partitioning, photochemical reaction), and were less correlated ( $r = 0.69 - 0.84$ ) with their corresponding factors (light SVOC, PAH and bulk carbon factors) in the current work, suggesting that the source apportionment results derived from particle-only SVOC data could be affected by atmospheric processes. PMF analysis was also performed on three temperature-stratified subsets of the total SVOC data, representing ambient sampling during cold (daily average temperature  $< 10$  °C), warm ( $\geq 10$  °C and  $\leq 20$  °C) and hot ( $> 20$  °C) periods. Unlike the particle only-based study, in this work the factor characterized by the low molecular weight (MW) compounds (light

SVOC factor) exhibited strong correlations ( $r = 0.82 - 0.98$ ) between the full data set and each sub-data set solution, indicating that the impacts of G/P partitioning on receptor-based source apportionment could be eliminated by using total SVOC concentrations.

## 6.1 INTRODUCTION

The Denver Aerosol Sources and Health (DASH) study was designed to explore the associations between short-term exposure to individual PM<sub>2.5</sub> components, sources and negative health effects (Vedal et al., 2009). Daily 24-h PM<sub>2.5</sub> sampling was conducted from mid-2002 to the end of 2008. Speciation of PM<sub>2.5</sub> has been carried out for gravimetric mass, inorganic ionic compounds (sulfate, nitrate and ammonium) and carbonaceous components, including elemental carbon (EC), organic carbon (OC) and a large array of semi-volatile organic compounds (SVOCs). Kim et al. (2012b) have investigated the lag structure of the association between PM<sub>2.5</sub> constituents and hospital admissions by disease using the 5-year bulk speciation data set of DASH study (nitrate, sulfate, EC and OC). They found that the estimated short-term effects of PM<sub>2.5</sub> bulk components, especially those of EC and OC, were more immediate for cardiovascular diseases and more delayed for respiratory diseases. Future work will focus on the association between specific PM<sub>2.5</sub> sources and health outcomes.

To develop control strategies for PM<sub>2.5</sub>, receptor-based models (e.g. Positive Matrix Factorization, Chemical Mass Balance) have been applied to quantitatively apportion PM<sub>2.5</sub> to sources that are detrimental to human health (Laden et al., 2000; Mar et al., 2005; Ito et al., 2006). One basic assumption of receptor-based models is that source profiles are constant over the period of ambient and source sampling (Chen et al., 2011). However, the output factors of a receptor model are not necessarily emission sources, and could be affected by atmospheric processes like photochemical reaction or gas/particle (G/P) partitioning (May et al., 2012). The influence of atmospheric processes on certain output factors can change with meteorological conditions (e.g. solar irradiance, ambient temperature). Thus, the assumption of constant source profiles does not hold for all output factors, especially for long time series studies.

PM<sub>2.5</sub> associated SVOCs data have been used as inputs for receptor models in many studies (Jaekels et al., 2007; Schnelle-Kreis et al., 2007; Shrivastava et al., 2007; Dutton et al., 2010b). All SVOCs are subject to G/P partitioning and thus partly distributed in the gas phase. According to the G/P partitioning theory developed by Pankow (1994a, b), which has been applied to the predictions of particulate matter (PM) formation (Liang and Pankow, 1996; Liang et al., 1997; Mader and Pankow, 2002), the partitioning of each individual compound is governed by its absorptive G/P partitioning coefficient,  $K_{p,OM}$ , which can either be measured directly (Eq. 1) or calculated from theory (Eq. 2):

$$K_{p,OM} = \frac{K_p}{f_{OM}} = \frac{F / M_{OM}}{A} \quad (1)$$

$$K_{p,OM} = \frac{RT}{10^6 \overline{MW}_{OM} \zeta_{OM} p_L^0} \quad (2)$$

where it is assumed that particle-phase organic material (OM) is primarily responsible for the absorptive uptake. Thus, it is meaningful to normalize the G/P partitioning constant ( $K_p$ ,  $m^3 \mu g^{-1}$ ) by the weight fraction of the absorptive OM phase ( $f_{OM}$ ) in the total PM phase (Eq. 1), so as to obtain  $K_{p,OM}$  ( $m^3 \mu g^{-1}$ ).  $F$  ( $ng\ m^{-3}$ ) is the mass concentration of each compound associated with the particle phase;  $A$  ( $ng\ m^{-3}$ ) is the mass concentration of each compound in the gas phase;  $M_{OM}$  ( $\mu g\ m^{-3}$ ) is the mass concentration of the particle-phase OM;  $R$  ( $m^3\ atm\ K^{-1}\ mol^{-1}$ ) is the ideal gas constant;  $T$  (K) is the ambient temperature;  $\overline{MW}_{OM}$  ( $g\ mol^{-1}$ ) is the mean molecular weight (MW) of the absorbing OM phase;  $\zeta_{OM}$  is the mole fraction scale activity coefficient of each compound in the absorbing OM phase; and  $p_L^0$  (atm) is the vapor pressure of each pure compound. For a given SVOC and a single OM phase, the G/P partitioning is only controlled by ambient temperature (Eq. 2). The mass fraction of the total SVOC in the atmosphere that contributes to the particle phase thus can change with ambient temperature. As such, the source profiles of

particle-phase SVOCs are expected to vary due to the influence of G/P partitioning, especially for those sources primarily contributing light SVOCs (e.g. docosane, fluoranthene). Therefore, when using a long time series of speciated PM<sub>2.5</sub> data as input for receptor model analysis, the light SVOC related sources/factors for a sub period of observation might be obscured by the influence of G/P partitioning, which will subsequently affect the health effect estimation of specific PM<sub>2.5</sub> sources.

In this study, gas-phase SVOC concentrations were estimated using their particle-phase concentrations based on absorptive mechanism (Eq. 1). The adsorption of SVOC onto particle surfaces (e.g. soot surface) was not considered in this work. The particle-phase concentrations of SVOCs were obtained from an existing 32-month series of daily PM<sub>2.5</sub> speciation, which has been used for source apportionment in a previous study (Xie et al., 2013b). In order to eliminate the influence of G/P partitioning on source apportionment, the total concentrations of gas- and particle-phase SVOCs were used as inputs for PMF analysis. The PMF2 model (Paatero, 1998a, b), coupled with a stationary block bootstrap technique quantifying errors due to random sampling (Hemann et al., 2009), was the primary source apportionment tool. Moreover, the 32-month data set of total SVOCs was divided into three sub-data sets by daily average temperature for source apportionment using the identical method. The use of smaller sub-data sets as inputs is to verify the elimination of G/P partitioning influence from the total SVOC-based PMF analysis.

## 6.2 METHODS

### 6.2.1 Particle phase measurements

Daily PM<sub>2.5</sub> samples were collected on the top of a two-story elementary school building in urban Denver. Details of the sampling site, set up, protocols and chemical analysis have been published by Vedal et al. (2009) and Dutton et al. (2009a, b). Daily average particle-phase SVOCs concentrations were obtained from existing PM<sub>2.5</sub> measurements, including 970 observations of 71 species (January 27, 2003 – October 2, 2005). Concentrations of inorganic ions, bulk elemental carbon (EC) and organic carbon (OC) were also measured for the same study period. The pointwise, blank corrected concentration uncertainties of each species were estimated by using the root sum of squares (RSS) method (Dutton et al., 2009a, b). The concentration and uncertainty data sets have been used as inputs for a particle only-based source apportionment in a previous study (Xie et al., 2013b). The meteorological (temperature, relative humidity and solar irradiance) and trace gas (ozone, nitrogen oxides (NO<sub>x</sub>) and CO) data used in this study were also obtained from Xie et al. (2013b).

### 6.2.2 Gas phase concentration and uncertainty estimation

The  $K_{p,OM}$  value for each species on each day was calculated by Eq. (2). Here four parameters are required, including  $T$ ,  $\overline{MW}_{OM}$ ,  $\zeta_{OM}$  and  $p_L^o$ . For this application  $T$  is the measured daily average temperature. Based on smog chamber and ambient studies (Odum et al., 1996; Hallquist et al., 2009), 150–250 g mol<sup>-1</sup> is a reasonable range for the average MW of the particulate OM phase; here we assume the  $\overline{MW}_{OM}$  to be 200 g mol<sup>-1</sup> for all samples, as is used in previous work (Barsanti and Pankow, 2004; Williams et al., 2010). Values of  $\zeta_{OM}$  were assumed to be unity for all species in each sample. Values of  $p_L^o$  were estimated using the group contribution methods (GCMs) SPARC (Hilal et al., 1995; <http://archemcalc.com/sparc/test/>) and

SIMPOL (Pankow and Asher, 2008). The  $p_L^{\circ}$  value for each species on each day was adjusted by daily average temperature:

$$p_L^{\circ} = p_L^{\circ,*} \exp \left[ \frac{\Delta H_{\text{vap}}^*}{R} \left( \frac{1}{298.15} - \frac{1}{T} \right) \right] \quad (3)$$

where  $p_L^{\circ,*}$  is the vapor pressure of each pure compound at 298.15 K;  $\Delta H_{\text{vap}}^*$  is the enthalpy of vaporization of the liquid ( $\text{kJ mol}^{-1}$ ) at 298.15 K. The  $p_L^{\circ,*}$ ,  $\Delta H_{\text{vap}}^*$  and average  $K_{\text{p,OM}}$  value for each species are given in Table 6.S1.

Gas-phase concentrations of each SVOC were calculated by Eq. (1). The values of  $F$  for each SVOC in Eq. (1) were obtained from existing  $\text{PM}_{2.5}$  measurements (Xie et al., 2013b);  $M_{\text{OM}}$  was estimated by multiplying the OC concentrations by a scaling factor of 1.53, which resulted in optimum mass closure of  $\text{PM}_{2.5}$  in a previous DASH study (Dutton et al., 2009a). The total concentration of each SVOC ( $S$ , gas + particle phase) on each day is then obtained by Eq. (4),

$$S = F + A = \frac{1 + K_{\text{p,OM}} M_{\text{OM}}}{K_{\text{p,OM}} M_{\text{OM}}} F \quad (4)$$

The uncertainty associated with  $S$  estimation was also calculated using the RSS method,

$$\delta S = \sqrt{\left( \frac{\partial S}{\partial F} \delta F \right)^2 + \left( \frac{\partial S}{\partial M_{\text{OM}}} \delta M_{\text{OM}} \right)^2} \quad (5)$$

where  $\delta S$  is the propagated uncertainty in  $S$ ;  $\delta F$  and  $\delta M_{\text{OM}}$  are the propagated uncertainties associated with particle-phase SVOC and  $M_{\text{OM}}$  measurements, and could be obtained from the uncertainty data sets introduced in section 6.2.1. The  $K_{\text{p,OM}}$  value uncertainty was not estimated in the current work. Statistics for the total concentration of each SVOC from January 27, 2003 to October 2, 2005 are listed in Table 6.S1, including the mean and median concentrations, mean particle-phase fractions, signal to noise ratios ( $S/N = \text{mean concentration}/\text{mean uncertainty}$ ) and coefficients of variation ( $CV = \text{standard deviation}/\text{mean concentration}$ ). Table 6.S1 also lists

statistics of particulate bulk components (mass, nitrate, sulfate, ammonium, EC and OC). The OC concentrations are shown in 5 fractions (OC1 – 4 and PC), representing the carbon measured at four distinct temperature steps (340, 500, 615 and 900 °C) with a pyrolyzed carbon adjustment in the first heating cycle of NOISH 5040 thermal optical transmission (TOT) method (NOISH, 2003; Schauer et al., 2003).

### 6.2.3 PMF analysis and uncertainty assessment

PMF2 (Paatero, 1998a, b), a multivariate receptor model, was used for source apportionment in this study. It is the primary source apportionment tool applied in the DASH project, and is discussed in detail by Dutton et al. (2010b). PMF uses an uncertainty-weighted least-squares fitting approach to identify distinct factor profiles and quantify factor contributions from a time series of observations. The bias and variability in factor profile and contribution due to random sampling error were estimated by applying a method from Hemann et al. (2009). 1,000 replicate data sets were generated from the original data set using a stationary block bootstrap technique 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), the Multilayer Feed Forward Neural Networks were trained to sort and align the factor profiles from each PMF bootstrap solution to that of the base case solution derived from the original data set. 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 contribution of each factor on each day, which could then be used to assess the variability of the PMF model fit. In this work, the factor number was determined based on the interpretability of different PMF

solutions (5 – 9 factors) as well stability across bootstrap-replicate data sets as represented by factor matching rate.

#### 6.2.4 Preparation of PMF input data set

Fifty one SVOCs and four bulk species (sulfate, nitrate, EC and total OC) were selected from all species with 970 daily observations for particle only-based PM<sub>2.5</sub> source apportionment (Xie et al., 2013b). The species screening was based on the percentage of missing values and observations below detection limit (BDL), S/N ratios and the stability of PMF solution. In this work, the candidate SVOCs for source apportionment were selected from the fifty one species used in the previous study. Bulk species were selected from nitrate, sulfate, EC and the five OC fractions. Interpretability and factor matching rate (> 50%) of the PMF solution were criteria for species screening. Among the five OC fractions, the OC1 concentration was measured under the lowest temperature (340 °C) and most likely influenced by G/P partitioning. The gas-phase concentrations of OC1 (or total OC) could not be estimated by using G/P partitioning theory in this work. Using particle-only OC1 (or total OC) as input for PMF will lead to biased source apportionment results, since all the SVOC compounds were adjusted by adding their estimated gas phase concentrations to measured particle phase concentrations. The OC4 concentration was very low with low S/N ratio. Thus OC1 and OC4 were excluded for PMF analysis. The other three fractions (OC2, OC3, PC) were assumed to be less or non-volatile and were included for PMF analysis. Finally, the six bulk species with 970 daily observations and forty six SVOCs with 970 estimated total concentrations constituted the primary PMF input data set.

Similarly to the previous Xie et al. (2013b) study, PMF analysis was also performed for three temperature-stratified subsets of the original 970 samples. The three sub-data sets consisted of sampling days with daily average 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. The statistics of total SVOCs during each of these three periods are shown in Tables 6.S2 – S4. PMF input species screening for each sub-data set was conducted in the same manner as for the full data set.

## 6.3 RESULTS AND DISCUSSION

### 6.3.1 Total SVOCs and their particle-phase fractions

Except steranes, the low MW species have the highest total concentrations and the lowest particle-phase fractions in each class of SVOCs (Table 6.S1). For example, docosane and fluoranthene are the most abundant species in *n*-alkanes and PAHs with mean concentrations of 32.8 ng m<sup>-3</sup> and 11.2 ng m<sup>-3</sup> respectively, one to two orders of magnitudes higher than those of high MW species in their chemical classes. In this study, the total concentrations of light *n*-alkane (e.g. docosane – pentacosane) and PAH (e.g. MW = 202) species increased by more than 100% from the cold to the hot periods (Tables 6.S2 – S4), possibly due to the evaporation of fossil fuels (Nahir, 1999) and increases in biogenic VOC emissions with increasing temperature.

The average particle-phase fraction of each SVOC was calculated for the cold, warm and hot periods and shown in Figure 6.1. All SVOCs exhibit the highest particle-phase fractions in cold periods and the lowest in hot periods, especially for those light SVOCs (e.g. docosane, fluoranthene), indicating a change in G/P partitioning behavior across different temperatures. Long chain *n*-alkanes (chain length > 27), heavy PAHs (MW > 252), steranes, hopanes, and sterols are mostly in the particle phase (> 75%) for all periods and less subject to evaporation (or partitioning to the gas phase) under higher temperatures. In Table 6.S5, the estimated particle-phase fractions of selected SVOCs (*n*-alkanes, PAHs, sterane and hopanes) in hot periods are more comparable with those observed by Fraser et al. (1997, 1998) in summer Los Angeles than in summer Athens (Greece) (Mandalakis et al., 2002). Average fractions of particulate PAHs for the whole period are similar to those annual averages measured by Tsapkis and Stephanou (2005) in Heraklion (Greece). While large differences were observed for the particle-phase fractions of light PAHs (MW < 252) in cold and hot periods compared with those measured in urban

Chicago (Simcik et al., 1997, 1998). These comparisons indicate that the estimations of G/P distributions of the SVOCs in this work are reasonable. Keep in mind that these differences may be influenced by parameters other than T, like  $\overline{MW}_{OM}$ ,  $\zeta_{OM}$  and  $M_{OM}$  in Eqs. (1) and (2).

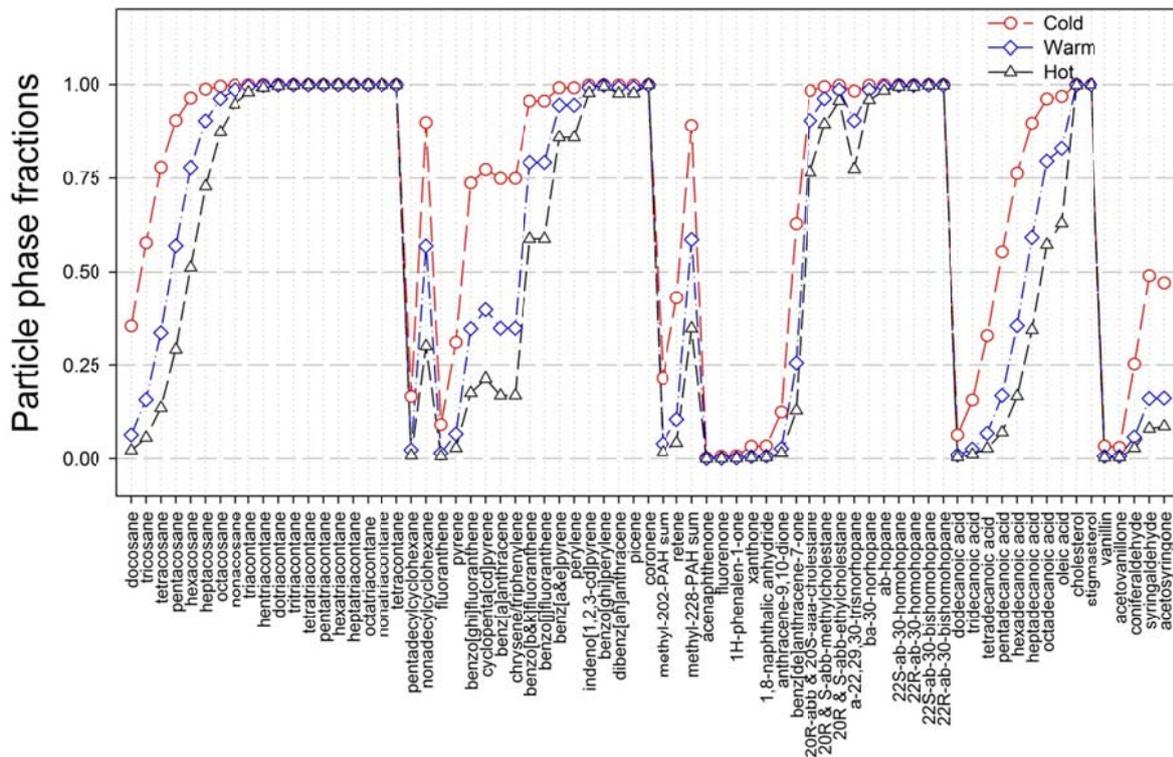


Figure 6.1 Mean particle-phase fractions of all SVOCs during cold, warm and hot periods.

### 6.3.2 Sensitivity of total SVOC estimation based on G/P partitioning theory

Based on G/P partitioning theory, changes in ambient temperature lead to the evaporation or condensation of SVOCs; the extent of such changes with temperature depend in part on values of  $\overline{MW}_{OM}$  and  $\zeta_{OM}$ , here assumed to be  $200 \text{ g mol}^{-1}$  and unity respectively. However,  $\overline{MW}_{OM}$  and  $\zeta_{OM}$  are highly dependent on the composition of PM, which is complex in an urban area and mostly unknown. The  $\overline{MW}_{OM}$  values are typically based on MW of organic compounds detected in laboratory and field studies, but in some cases (e.g. under high relative humidity (RH)) need to be adjusted downward for the presence of water in the particulate OM phase (Pankow and Chang, 2008; Chang and Pankow, 2010). The  $\zeta_{OM}$  values for organic compounds in atmospheric applications are not necessarily unity for different SVOCs in varied PM composition (e.g., varied amounts of polar and non-polar organic compounds and water) (Pankow and Chang, 2008; Pun, 2008). The uncertainties in these two parameters, as well as the OM/OC ratio, could affect the estimation of total SVOC concentration as described in section 6.3.1.

Combining Eqs. (2) and (4), the equation for total SVOC calculation can be re-written as:

$$S = F + A = \left(1 + \frac{10^6 p_L^0 \overline{MW}_{OM} \zeta_{OM}}{RTM_{OM}}\right) F \quad (6)$$

from which we can infer that the estimation of total concentration ( $S$  value) for specific SVOC is primarily determined by the following term:

$$z = \frac{10^6 p_L^0 \overline{MW}_{OM} \zeta_{OM}}{RTM_{OM}} \quad (7)$$

if  $z$  is close to 0, then most of the target SVOC is in particle phase; if  $z$  is close to or higher than 1, then the target SVOC is strongly subject to G/P partitioning. The sensitivity of total SVOC estimation ( $S$  value) to  $T$ ,  $\zeta_{OM}$ , OM/OC ratio,  $\overline{MW}_{OM}$  can be evaluated as the changes of  $z$  value

to these uncertain parameters in Eq. (7). To test the sensitivity, the average temperatures and OC concentrations during the cold, warm and hot periods (defined in section 6.2.4) were investigated; docosane was selected as an example to represent SVOCs with similar pure vapor pressure and G/P partitioning behavior. Three  $\zeta_{OM}$  (0.5, 1.0, 1.5) and four  $\overline{MW}_{OM}$  (50, 150, 200, 300 g mole<sup>-1</sup>) values, based on Pankow and Chang (2008) and four OM/OC (1.3, 1.4, 1.5, 1.6) ratios, based on Bae et al. (2006), were used to test the sensitivity of  $z$  value (or  $S$  value) calculation. The values of the above parameters investigated were listed in Table 6.1.

Table 6.1 Values of parameters used to test the sensitivity of total SVOC estimation.

Parameters	Cold	Warm	Hot
T (K)	276.5	288.5	297.6
$M_{oc}$ ( $\mu\text{g m}^{-3}$ ) <sup>a</sup>	2.78	2.39	3.45
$p^{\circ}L$ (atm) <sup>b</sup>	8.52E-10	6.80E-09	2.96E-08
$\zeta_{OM}$	0.5, 1, 1.5		
$\overline{MW}_{OM}$ (g mol <sup>-1</sup> ) <sup>c</sup>	50, 150, 200, 300		
OM/OC	1.3, 1.4, 1.5, 1.6		

(a) Mean organic carbon concentrations during different periods.

(b) Vapor pressures of docosane at different temperatures.

(c) Mean molecular weight of absorbing organic material.

In Figure 6.2, the sensitivity of  $z$  value to T,  $\zeta_{OM}$ , OM/OC ratio and  $\overline{MW}_{OM}$  are shown in nine mesh plots. Each mesh plot exhibits the changes of  $z$  value to varied  $M_{OM}$  and  $\overline{MW}_{OM}$  for a given T and  $\zeta_{OM}$ . From the left to the right in Figure 6.2,  $z$  values are increased by 1-2 times as  $\zeta_{OM}$  increases, which can be expected from Eq. (7); while from the top to the bottom,  $z$  values are increased by more than one order of magnitude when the ambient temperature increases by 21 K. Thus, for docosane, the calculation of  $z$  value (or  $S$  value) is more sensitive to the changes in ambient temperature than the prescribed changes in activity coefficient. This is largely due to the exponential increase in vapor pressure with temperature of docosane and other SVOCs (Eq. 3).

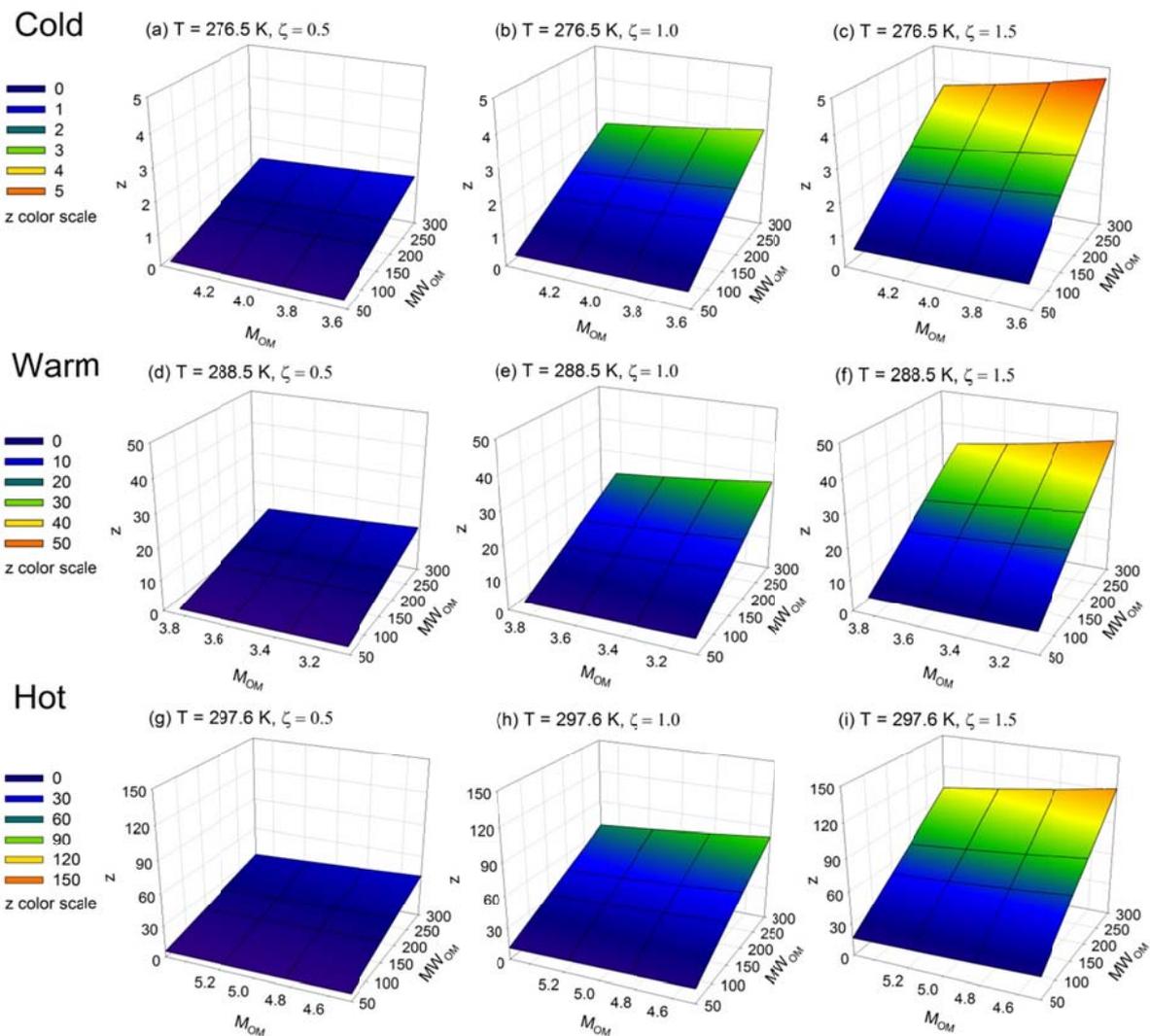


Figure 6.2 Sensitivity of the calculation of total SVOC concentration ( $S$  value), determined by  $z$  value, to ambient temperature ( $T$ ), mole fraction scale activity coefficient ( $\zeta_{OM}$ ), OM/OC ratio and mean molecular weight of absorbing OM phase ( $\overline{MW}_{OM}$ ). The  $z$  value equals to the ratio of gas phase to particle phase SVOC.

Within each mesh plot,  $z$  value has a linear and reciprocal relationship with  $\overline{MW}_{OM}$  and  $M_{OM}$  respectively, which can also be expected from Eq. (7). The maximum  $z$  value is 7.4 times as the minimum  $z$  value in each mesh plot. In this test, the variations of  $\overline{MW}_{OM}$  are much larger than those of  $M_{OM}$ , so the effects of  $\overline{MW}_{OM}$  to the calculation of  $z$  value seems more important than that of OM/OC ratio. However, if  $M_{OM}$  and  $\overline{MW}_{OM}$  have similar variations (e.g. OM/OC ranges from 1.2 to 2.0, and  $\overline{MW}_{OM}$  ranges from 150 to 250 g mol<sup>-1</sup>), then these two parameters should have similar effects on the calculation of  $z$  value (or  $S$  value).

As demonstrated by the sensitivity study, the estimation of total SVOC concentration is mostly sensitive to ambient temperature. In this work, the sensitivity of G/P partitioning to ambient temperature is largely accounted for by adjusting the vapor pressure of each SVOC according to the daily average temperature. However, the total SVOC concentration estimated in the current work might be subject to considerable uncertainty due to the variations of  $\zeta_{OM}$ ,  $\overline{MW}_{OM}$  and OM/OC ratio across the sampling period.

### 6.3.3 PMF results for the full data set

A 7-factor solution was determined for the full data set using total SVOC concentration due to the most readily interpretable resulting factors and a relatively high factor matching rate of 79.9% between bootstrapped and base case solutions (Table 6.2). These seven factors are identified as nitrate, sulfate, *n*-alkane, sterane, light SVOC, PAH and bulk carbon. Figures 6.S1 and 6.S2 present the median factor profiles and contributions with one standard deviation from bootstrapped PMF solutions, which represent the variability of PMF solution due to random sampling error. The factor contributions are also summarized by day of the week in boxplots (Figure 6.S3). The factor profiles have been normalized by

$$F_{kj}^* = \frac{F_{kj}}{\sum_{k=1}^p F_{kj}} \quad (8)$$

where  $F_{kj}^*$  is the relative weighting of species  $j$  in factor  $k$  to all other factors. The median factor contributions in Figure 6.S2 are expressed as reconstructed PM<sub>2.5</sub> mass – the sum of nitrate, sulfate, EC and straight OC fractions (OC2, OC3 and PC) contributed by each factor. The contribution time series were divided into three periods (cold, warm and hot) and shown as the average contributions to major PM<sub>2.5</sub> components (nitrate, sulfate, EC and OC; Table 6.3). The sum of factor contributions to each component can be compared with the observed average concentration (Table 6.3). The sampling variability of factor contributions are represented by the median CVs (CV = standard deviation/median factor contribution). In addition, the factor contributions during each period were linearly regressed to meteorological and trace gas measurements in the same manner as discussed in the previous Xie et al. (2013b) study, so as to understand the association between each factor and pollution sources/processes. The resulting correlation coefficients are given in Table 6.S6.

Table 6.2 PMF simulation statistics for different data sets.

Parameters	Data sets			
	Full	Cold	Warm	Hot
No. of species	52	52	52	37
No. of samples	970	364	318	288
No. of factors	7	7	7	7
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	799	886	772	619

Table 6.3 Average factor contributions to bulk components for full data set solution and sub-data set solutions ( $\mu\text{g m}^{-3}$ )

Factors	Full data set solution					Sub-data set solution				
	Nitrate	Sulfate	EC	OC <sup>a</sup>	CV <sup>b</sup>	Nitrate	Sulfate	EC	OC	CV
	Cold period					Cold period				
Nitrate	2.2	0.24	0.060	0.076	0.036	2.1	0.14	0.031	0.14	0.074
Sulfate	0.035	1.0	0.0026	0.022	0.060	0.12	1.1	0.015	0.015	0.11
<i>n</i> -Alkane	0.0004	0.0079	0.0003	0.26	0.35	0.0007	0.0023	0.00	0.25	0.27
Sterane	0.0008	0.0079	0.13	0.17	0.52	0.012	0.025	0.070	0.10	0.52
Light SVOC	0.0009	0.0013	0.0012	0.027	0.22	0.0040	0.0045	0.030	0.18	0.14
PAH	0.0003	0.0010	0.21	0.15	0.31	0.0005	0.0030	0.057	0.019	0.84
Bulk carbon	0.0081	0.0052	0.12	0.41	0.33	0.0009	0.0095	0.37	0.47	0.23
Subtotal	2.2	1.3	0.54	1.1		2.2	1.3	0.58	1.2	
Observed Conc.	2.2	1.3	0.61	1.4						
	Warm period					Warm period				
Nitrate	0.32	0.036	0.0089	0.011	0.23	0.37	0.10	0.028	0.021	0.44
Sulfate	0.032	0.93	0.0023	0.020	0.031	0.011	0.86	0.00	0.12	0.11
<i>n</i> -Alkane	0.0002	0.0038	0.0001	0.12	0.39	0.0026	0.0034	0.00	0.16	0.44
Sterane	0.0003	0.0031	0.053	0.069	0.61	0.0007	0.0090	0.069	0.068	0.68
Light SVOC	0.0041	0.0061	0.0056	0.12	0.15	0.0012	0.0069	0.012	0.14	0.15
PAH	0.0002	0.0005	0.11	0.082	0.33	0.0001	0.0003	0.091	0.057	0.41
Bulk carbon	0.014	0.0089	0.21	0.70	0.13	0.0050	0.0010	0.19	0.58	0.21
Subtotal	0.37	0.99	0.39	1.1		0.39	0.98	0.39	1.1	
Observed Conc.	0.40	1.0	0.43	1.2						
	Hot period					Hot period				
Nitrate	0.11	0.012	0.0030	0.0038	0.35	/	/	/	/	/
Sulfate	0.040	1.2	0.0029	0.025	0.037	/	1.0	0.035	0.13	0.14
<i>n</i> -Alkane	0.0002	0.0031	0.0001	0.10	0.46	/	0.0001	0.051	0.46	0.50
Sterane	0.0002	0.0020	0.035	0.045	0.73	/	0.035	0.077	0.24	0.52
Light SVOC	0.011	0.016	0.015	0.33	0.15	/	0.079	0.012	0.11	0.30
PAH	0.0001	0.0002	0.051	0.037	0.37	/	0.0005	0.039	0.0070	0.74
Bulk carbon	0.023	0.015	0.35	1.2	0.14	/	0.056	0.22	0.55	0.39
Median <i>n</i> -alkane	/	/	/	/	/	/	0.0026	0.0070	0.17	0.56
Subtotal	0.18	1.2	0.45	1.7		/	1.2	0.44	1.7	
Observed Conc.	0.19	1.2	0.46	1.8						

(a) Sum of contributions to OC2, OC3 and PC fractions.

(b) Median coefficient of variation (CV) of factor contributions, CV = standard deviation/median factor contribution.

In Table 6.3, the nitrate and sulfate concentrations are dominated by the nitrate (average 59.4% – 97.4%) and sulfate (79.5% – 96.0%) factors in all periods. In cold periods, the PAH factor (39.9%) had the highest contribution to EC concentrations, followed by the sterane (25.2%) and bulk carbon (23.0%) factors; while in warm and hot periods, the bulk carbon factor contributed the most of the EC concentrations (warm, 53.3%; hot, 76.5%). The bulk carbon

factor also has the highest contribution to OC (36.6% – 67.9%) in all periods. Here the OC consists of the three less or non-volatile OC fractions (OC2, OC3 and PC) that were used for source apportionment. The factors with small contributions to reconstructed PM<sub>2.5</sub> are prone to having high variability, as shown by their higher CVs (e.g., *n*-alkane, sterane and PAH factors). In each period, the sum of factor contributions to each major PM<sub>2.5</sub> component is close to the observed average concentration.

#### 6.3.4 Comparison to particle only-based source apportionment

In the previous Xie et al. (2013b) study, an 8-factor solution was determined with factors labeled as inorganic ion, *n*-alkanes, EC/sterane, light *n*-alkane/PAH, medium alkane/alkanoic acid, PAH, winter/methoxyphenol and summer/odd *n*-alkane. The medium alkane/alkanoic acid and winter/methoxyphenol factors only contributed a small part (0.41% – 1.10%; 0.16% – 4.21%) of reconstructed PM<sub>2.5</sub> mass and were not resolved in this study. The 7 factors resolved in the current work could be matched with the remaining 6 factors in the particle only-based solution after combining the nitrate and sulfate factors. According to the previous Xie et al. (2013b) study, the 7 factors in the current work could be primarily or partly related to secondary ion formation (nitrate and sulfate factors), road dust (*n*-alkane factor), lubricating oil combustion (sterane factor), fossil fuel evaporation and biogenic emissions that subject to atmospheric processes (light SVOC factor), motor vehicle emissions (PAH factor) and secondary organic aerosol (SOA) formation (Bulk carbon factor). Correlations of factor contributions between the matched pairs of factors are shown in Figure 6.3.

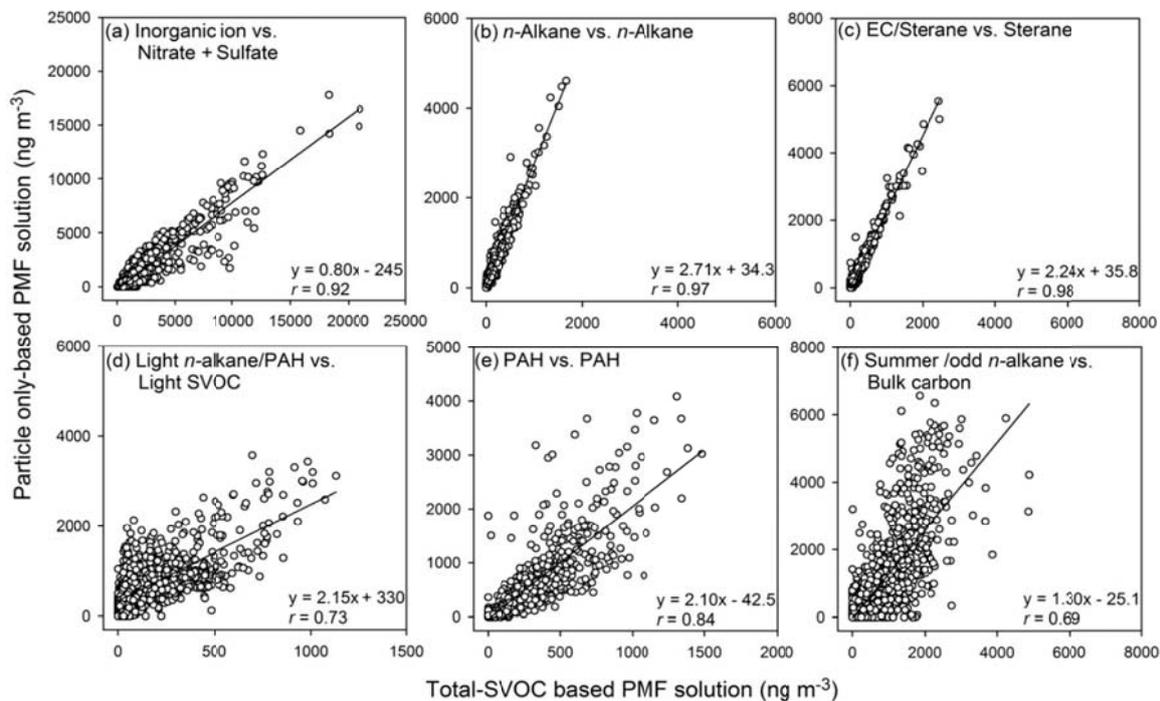


Figure 6.3 Linear regressions of matched pairs of factors between particle only-based and total SVOC-based PMF analysis.

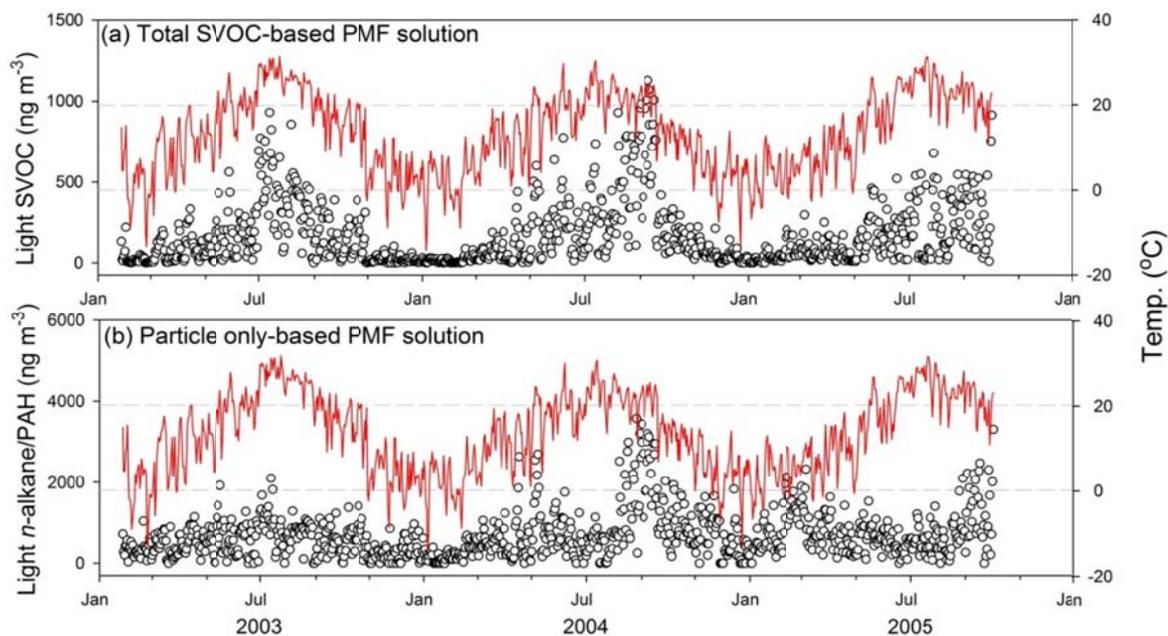


Figure 6.4 Median factor contribution time series (blank circle) of (a) light SVOC factor from the total SVOC-based solution, and (b) light *n*-alkane/PAH factor from the particle only-based solution. The red line represents the timeseries of daily average temperature.

The factors characterized by inorganic ions, heavy *n*-alkanes and steranes exhibit strong correlations ( $r = 0.92 - 0.98$ ) between the particle only-based and total SVOC-based PMF solutions (Figure 6.3). This strong correlation is because these factors are primarily linked with secondary formation or primary emission, and the heavy *n*-alkanes and steranes are mostly distributed in particle phase (Figure 6.1). The light *n*-alkane/PAH and PAH factors from the particle only-based solution are less correlated with the light SVOC ( $r = 0.73$ ) and PAH ( $r = 0.84$ ) factors from the total SVOC-based solution (Figure 6.3). This is because these factors contain a significant fraction of light organic compounds, being subject more strongly to G/P partitioning. In Figure 6.4a, the light SVOC factor shows an increase in contribution when the temperature rises, supporting the association of this factor with fossil fuel evaporation and biogenic emissions. In contrast, the light *n*-alkane/PAH factor from the particle only-based solution exhibits low contributions in mid-summer when the temperature is the highest of the year and small peaks in winter when the temperature is low (Figure 6.4b). The high temperatures in mid-summer keep light organic compounds in the gas phase, while the low temperatures in winter benefit the partitioning of gas-phase organics to the particle phase. In addition, the high ozone concentrations in mid-summer could also be responsible for the decrease in factor contribution, since negative correlations have been observed between ozone concentration and the two matched factors (Light SVOC: -0.48, Table 6.S6; light *n*-alkane/PAH: -0.52, (Xie et al., 2013b)) from both solutions during hot periods. No obvious difference in contribution time series was observed for the PAH factor between the particle only-based and total SVOC-based PMF solutions, since the PAH factor was mostly characterized by medium and high MW PAHs (MW  $\geq 226$ ; Figure 6.S1f).

The bulk carbon factor in the current work contains the largest percentages of EC and OC fractions (Figure 6.S1g), and has maximum contributions in summer (Figure 6.S2g). This factor should be influenced by both SOA formation, as supported by the correlation between the factor contribution and ozone concentrations in hot periods ( $r = 0.36$ ; Table 6.S6), and primary emissions from motor vehicles, as supported by the weekend decrease in factor contribution (Figure 6.S3g) and the correlations between the factor contribution and  $\text{NO}_x$  and CO concentrations (Table 6.S6). The summer/odd *n*-alkane factor from the particle only-based solution was primarily associated with SOA formation, which lead to a moderate correlation ( $r = 0.69$ ; Figure 6.3f) with the bulk carbon factor in the current work. Except the inorganic ion factors, all other carbonaceous factors from the particle only-based solution show higher contributions than their matched factors from the total SVOC-based solution, as illustrated by the regression slopes ranging from 1.3 to 2.7 (Figure 6.3). This can mostly be attributed to the fact that the OC1 fraction was not included for source apportionment in the current study, which accounted for 47.6% of the total OC on average. While the particle only-based study used total OC for PMF analysis.

#### 6.3.5 PMF results for temperature-stratified sub-data sets

Statistics of PMF simulations for the three temperature-stratified sub-data sets are given in Table 6.2. Comparing to the full data set, the same species and factor number were chosen for PMF analysis of the cold and warm period sub-data sets. The factor matching rates are 88.6% and 77.2%, respectively (Table 6.2). For the hot period sub-data set, fewer species were used to obtain physically meaningful solution with high factor matching rate. Finally, a 7-factor solution was chosen with a factor matching rate of 61.9% (Table 6.2). Figures 6.S4 – 6.S6 show the normalized factor profiles for each sub-data set solution with one standard deviation. The median

factor contributions to major PM<sub>2.5</sub> components during each period were averaged and presented in Table 6.3, and can be compared to those from full data set solution. Median CVs of factor contributions are also included in Table 6.3 to reflect the variability from random sampling error. In addition, the correlations between factor contributions and meteorological and trace gas measurements are given in Table 6.S7. Similarly to the full data set solution, the nitrate and sulfate concentrations are mostly accounted for by the nitrate (average 93.9% – 94.7%) and sulfate (85.2% – 87.9%) factors (Table 6.3). The EC and OC concentrations are highest apportioned to the bulk carbon factor (EC, 48.9% – 64.9%; OC, 32.9% – 50.7%) for all periods.

#### 6.3.6 Comparison to PMF results of the full data set

The factors from the analysis of each temperature-stratified sub-data set were matched to those from the full data set based on factor profiles. The linear regressions of factor contributions between matched pairs of factors are given in Table 6.4, so as to verify that the influence of G/P partitioning was eliminated from the PMF analysis by using the total SVOC data set. However, we cannot rule out the impacts of other atmospheric processes like photochemical reactions, which is partly dependent on ambient temperature (Crouse et al., 2011; Pathak et al., 2007) and not considered in this work.

Table 6.4 Regression statistics of factor contributions between full data set and sub-data set solutions.

Factor		Cold			Warm			Hot		
Full <sup>a</sup>	Sub <sup>b</sup>	Slope	Intercept	<i>r</i>	Slope	Intercept	<i>r</i>	Slope	Intercept	<i>r</i>
Nitrate	Nitrate	0.94	-49.7	1.00	1.20	56.7	0.98	/	/	/
Sulfate	Sulfate	1.12	33.2	1.00	1.02	-20.9	0.99	1.12	-219	0.99
<i>n</i> -alkane	<i>n</i> -alkane	0.98	-4.18	0.98	1.17	14.5	0.99	3.37	162	0.79
Sterane	Sterane	0.70	12.8	0.98	1.19	-2.65	0.99	3.45	71.8	0.81
Light SVOC	Light SVOC	5.34	50.2	0.96	1.30	-21.2	0.98	0.80	-102	0.82
PAH	PAH	0.24	-10.9	0.97	0.73	5.33	0.99	0.39	12.6	0.91
Bulk carbon	Bulk carbon	1.12	236	0.54	0.96	-118	0.96	0.59	-80.5	0.81
Sum <sup>c</sup>	Sum	1.02	-34.3	0.99	1.00	17.9	0.99	0.74	153	0.89

(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 factor contributions.

### 6.3.6.1 Cold period

All the factors resolved by using the cold period sub-data set show similar factor profiles as their corresponding factors from the full data set solution (Figure 6.S1 and 6.S4). The EC concentration is more strongly apportioned to the bulk carbon factor from the cold period solution (average 63.8%) than that from the full data set solution (22.2%; Table 6.3). Moreover, strong correlations were observed between the bulk carbon factor from the cold period solution and NO<sub>x</sub> ( $r = 0.76$ ) and CO ( $r = 0.76$ ; Table 6.S7) concentrations. As such, the bulk carbon factor from the cold period solution should be mainly associated with primary emissions (e.g., gasoline and diesel vehicles). The full data set solution assumes constant co-influence of primary and secondary sources throughout the sampling period, which leads to a moderate correlation ( $r = 0.54$ ; Table 6.4) of the bulk carbon factor between the full data set and cold period solutions. For other factors, relatively strong correlations ( $r = 0.96 - 1.00$ ; Table 6.4) were observed between the two solutions, indicating that these matched pairs of factors could be linked to

similar pollution sources/processes. Among all the factors, the light SVOC factor is most likely influenced by G/P partitioning when we only use the filter measurement data for source apportionment. The influence of G/P partitioning should be different across different periods due to the distinct temperature ranges, while the particle-only full data set solution assumes constant G/P partitioning influence. In Figure 6.5a,d, the light *n*-alkane/PAH factor from the particle only-based PMF analysis was more poorly correlated ( $r = 0.41$ ) between the cold period and the full data set solutions (Xie et al., 2013b) than the light SVOC factor from the total SVOC-based PMF analysis ( $r = 0.96$ ). These results suggested that the G/P partitioning influence was removed from PMF analysis by using the total SVOC data set as input.

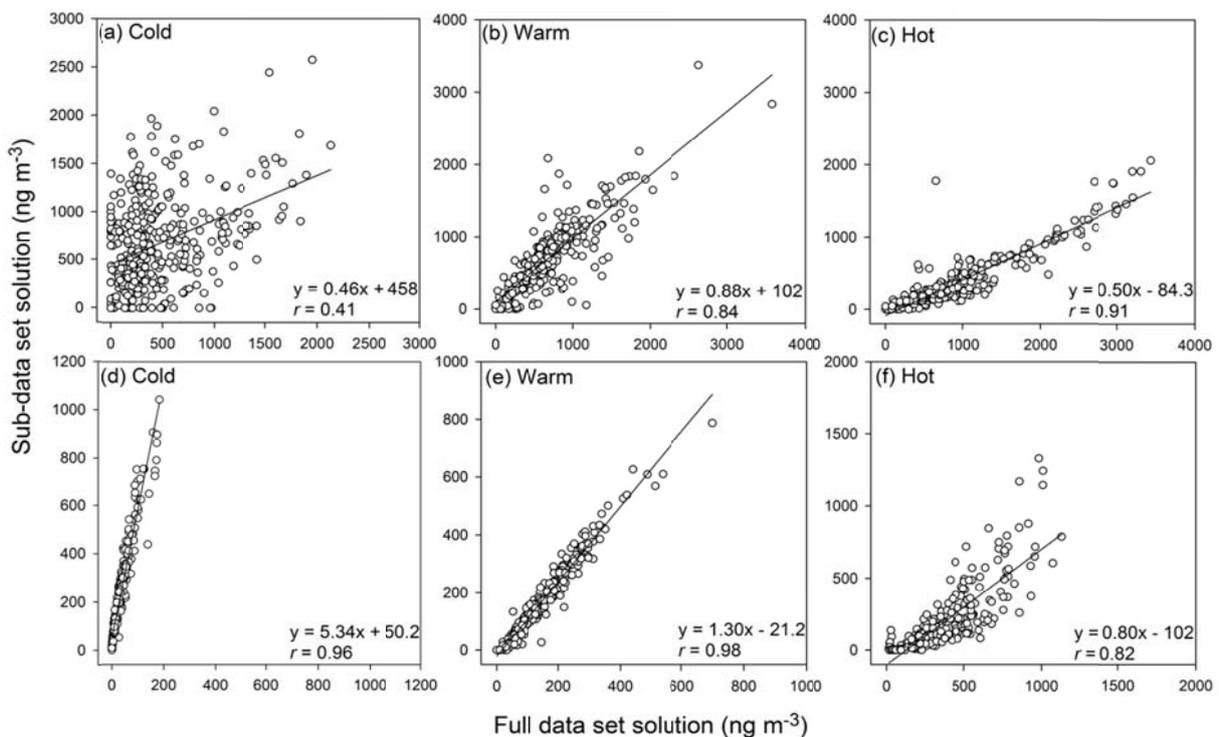


Figure 6.5 Linear regressions of factor contributions between the full data set and sub-data set solutions, (a-c) light *n*-alkane/PAH factor from particle only-based analysis; (d-f) light SVOC factor from total SVOC-based analysis.

### 6.3.6.2 Warm period

The factors resolved by using the warm period sub-data set are also similar as those from the full data set solution on factor profiles (Figure 6.S1 and 6.S5). Moreover, the factor contributions of the warm period and full data set solutions are relatively strongly correlated ( $r = 0.96 - 0.99$ ) with regression slopes close to unity ( $0.73 - 1.30$ ; Table 6.4). Such consistency between the warm period and full data set solutions was also observed in the previous Xie et al. (2013b) study. One explanation is that the PMF model is solved by minimizing the sum of the squared, scaled residues, and then requires the mean concentrations of most species to be fit well. The average concentrations of most SVOCs in warm periods are closer to the averages of the whole period than those during cold and hot periods. Thus, the factor contributions of the warm period solution are more consistent with those of the full data set solution.

### 6.3.6.3 Hot period

For the hot period, the nitrate measurements were not included for source apportionment due to the high percentages of missing and BDL observations, resulting in the omission of the nitrate factor. Meanwhile, a new factor was resolved and labeled as median *n*-alkane. It contains significant fraction of *n*-alkane with a chain length ranging from 22 to 29 (Figure 6.S6g). The factor contribution was moderately correlated with ambient temperature ( $r = 0.59$ ) and anti-correlated with relative humidity ( $r = -0.45$ ; Table 6.S7). So the median *n*-alkane factor might be linked with temperature-dependent summertime emissions with contribution time series opposing to that of relative humidity. The median *n*-alkane factor was also identified by using the particle-only sub-data set for hot periods (Xie et al., 2013b), and well correlated ( $r = 0.80$ ) with that identified in this work. The other factors were matched to those from the full data set solution with strong correlations ( $r = 0.79 - 0.99$ ; Table 6.4). However, the regression plot for

the light SVOC factor in hot periods (Figure 6.5f) is more scattered than those in cold and warm periods (Figure 6.5d,e); and from the cold to hot periods, the light SVOC factor becomes less correlated with ambient temperature ( $r$ , 0.61 – 0.07; Table 6.S7). These could be caused by the increased photochemical reactions during hot periods, supported by the negative correlation ( $r = -0.46$ ) between the light SVOC factor and ozone concentration.

## 6. 4 CONCLUSIONS

The gas-phase concentrations of 71 SVOCs were estimated using particle-phase measurements by G/P partitioning theory. In order to eliminate the impacts of G/P partitioning on PMF analysis, the gas-phase concentrations of all SVOCs were added to their particle-phase concentrations as inputs for source apportionment. Seven factors were identified from the full data set, including the nitrate, sulfate, *n*-alkane, sterane, light SVOC, PAH and bulk carbon factors, and could be matched to those from a previous particle only-based PMF study (Xie et al., 2013b) with reasonable ( $r = 0.69$ ) to excellent ( $r = 0.98$ ) correlations. Three temperature-stratified sub-data sets, representing ambient sampling during the cold, warm and hot periods, were also analyzed using PMF. Unlike the light *n*-alkane/PAH factor from the particle only-based study, the light SVOC factor from the total-SVOC based PMF solution exhibited strong correlations ( $r = 0.82 - 0.98$ ) between the full data set and each sub-data set solutions. These results suggested that the influences of G/P partitioning on PMF analysis could be removed by using total SVOC (gas + particle phase) data. However, the impact of photochemical process has not been ruled out in this work, as illustrated by the moderate correlation ( $r = 0.54$ ) between the bulk carbon factor of the full data set solution and that of the cold period solution.

This study is our first step in improving SVOC-based PMF analysis by removing the impacts of G/P partitioning. However, the pre-assumptions (e.g. absorptive partitioning,  $\overline{MW}_{OM}$  and  $\zeta_{OM}$  values) made for the calculation of gas-phase SVOC concentrations need to be verified, and if necessary refined, by comparing with field measurements. Additionally, more source markers are required to further apportion the bulk carbon factor. Finally, total and speciated gas-phase SVOCs (e.g. *n*-alkanes, PAHs) data are needed to further understand the ambient OC sources. All of the above will be considered in our subsequent work.

**CHAPTER 7 CHARACTERIZATION OF *n*-ALKANES, PAHs, 2-METHYLTETROLS  
AND LEVOGLUCOSAN IN GAS PHASE AND PM<sub>2.5</sub> USING A MEDIUM VOLUME  
SAMPLER**

Mingjie Xie, Michael P. Hannigan

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

## 7.0 ABSTRACT

In this study, fifty pairs of 24-h gas- and particle-phase ( $PM_{2.5}$ ) samples of semi-volatile organic compounds (SVOCs) were collected from August 2012 to July 2013. A medium volume sampler was used incorporating quartz fiber filter (QFF) and polyurethane foam (PUF)/XAD/PUF sandwich. Non-polar (*n*-alkanes and PAHs) and polar (2-methylterols and levoglucosan) SVOCs were extracted and measured separately for all samples. A backup QFF (bQFF) was used to estimate the positive sampling artifact of particulate organics due to gas-phase adsorption. The organic carbon (OC) detected in the bQFF was OC1 fraction that evolved off the QFF in the first temperature step (340 °C). The breakthrough experiments showed that the PUF/XAD/PUF sandwich could collect gas-phase *n*-alkanes, PAHs, 2-methyltetrols and levoglucosan with low breakthrough, even for volatile species (e.g., dodecane, naphthalene); however, the recoveries of levoglucosan in PUF/XAD/PUF sandwiches were lower than 70% (51.9 – 63.3%). Comparing species concentrations across different sampling matrices, the light *n*-alkanes (MW < 282) and PAHs (MW < 192) in the top QFF (tQFF, where  $PM_{2.5}$  loaded) had comparable median concentrations as those in the bQFF, and much lower median concentrations than those in PUF/XAD-4/PUF samples. The relative abundance of the heavier *n*-alkanes (MW  $\geq$  282) and PAHs (MW  $\geq$  192) in the tQFF to those in the bQFF and PUF/XAD-4/PUF samples increased with molecular weight (MW). The concentration ratios of 2-methyltetrols and levoglucosan in the bQFF or PUF/XAD/PUF samples to those in the tQFF changed across the sampling period, and were often close to or higher than unity in summer.

## 7.1 INTRODUCTION

Compositional data of semi-volatile organic compounds (SVOCs) are becoming more commonly used than those of elements as inputs for receptor models (e.g., Positive Matrix Factorization, PMF; Chemical Mass Balance, CMB) to apportion particulate organics (organic carbon, OC; and elemental carbon, EC) into sources (Jaekels et al., 2007; Shrivastava et al., 2007). Xie et al. (2012c) applied PMF model to evaluate the utility and consistency of different speciated data sets in source apportionment of PM<sub>2.5</sub>, and found that SVOCs were more effective than water soluble elements in resolving organic sources. The Denver Aerosol Sources and Health (DASH) study applied a 32-month data set of daily particle-phase SVOCs for source apportionment using PMF model (Xie et al., 2013b), with the goal of relating short term exposure to individual PM<sub>2.5</sub> sources to negative health effects. However, the output factors derived from receptor-based source apportionment using particle-phase SVOCs data are not necessarily pollution sources. A factor could also reflect an atmospheric process (e.g., gas/particle partitioning, photochemical reactions) or a combination of sources and/or processes (Xie et al., 2013b). To eliminate the effects of one process, gas/particle (G/P) partitioning, on receptor-based source apportionment using SVOCs data, Xie et al. (2013a) added estimated gas-phase concentrations of SVOCs, derived from particle-phase concentrations by equilibrium G/P partitioning model (Pankow, 1994a, b), to their particle-phase concentrations for source apportionment. Unlike the particle-only based source apportionment (Xie et al., 2013b), the source that was most likely subject to G/P partitioning (light SVOC factor) exhibited consistent contributions between full data set solution and temperature-stratified sub-data set solutions, suggesting that using total SVOCs (gas + particle phase) data for receptor-based source apportionment could remove the influences of G/P partitioning. To use this technique routinely,

the estimation of gas-phase SVOCs needs to be verified, and if necessary refined, by field measurement.

High volume air samplers equipped with glass fiber (GFF) or quartz fiber (QFF) filters followed by polyurethane foam (PUF) have been widely used for monitoring G/P partitioning of SVOCs (Fraser et al., 1997; Simcik, et al., 1997; Chen et al., 2006; Yang et al., 2013). The sampling artifacts that might result in biased G/P partitioning include the re-evaporation of particle-phase SVOCs (“blow-off”, negative artifact) and adsorption of gas-phase SVOCs onto filter media (“blow-on”, positive artifact). Schauer et al. (1999) compared different sampling equipment and found that the “blow-on” effect dominated the “blow-off” effect, leading to positive sampling artifacts for particulate OC. In addition, high breakthrough of light SVOCs (e.g., naphthalene) has been reported as a limiting factor on the application of PUF adsorbent (Hart et al., 1992; Peters et al., 2000). XAD-coated annular denuder followed by filter pack in Integrated Organic Gas and Particle Sampler (IOGAPS) system has low breakthrough in collecting gas-phase per- and poly-fluorinated compounds (PFCs) (Ahrens et al., 2011). This sampling method can also reduce the positive artifacts in particle-phase SVOCs collection. However, considerable particle loss (10-24%) has been observed by Zhang et al. (2012) when applying the IOGAPS system for the measurements of diesel engine emissions.

Most previous G/P partitioning studies focused on non-polar or semi-polar SVOCs like *n*-alkanes and polycyclic aromatic hydrocarbons (PAHs) (Fraser et al., 1997; Eiguren-Fernandez et al., 2004), polychlorinated biphenyls (PCBs) (Simcik et al., 1997) and polybrominated diphenyl ethers (PBDEs) (Chen et al., 2006). A few recent studies have investigated the G/P partitioning of secondary organic aerosols (SOA, e.g., carbonyls) (Healy et al., 2008; Perraud et al, 2012; Kawamura et al., 2013; Zhao et al., 2013), and found that the equilibrium G/P partitioning model

could not apply to all observations. However, very few field measurements have been conducted on the G/P partitioning of specific source related polar SVOCs. Specifically, no study has looked at 2-methyltetrols for isoprene derived SOA (Claeys et al., 2004), and levoglucosan for biomass burning (Simoneit et al., 1999). Additionally, field measurements are in need to verify the satisfaction of equilibrium G/P partitioning for those species.

In this work, we investigated the concentrations of non-polar (*n*-alkanes and PAHs) and polar (levoglucosan and 2-methylterols) SVOCs in gaseous and particle (PM<sub>2.5</sub>) phases using a medium volume sampler quipped with two sampling trains. One sampling train consisted of a QFF pack followed by a PUF/XAD/PUF cartridge; the other train was composed of a Teflon membrane filter (TMF) and another PUF/XAD/PUF cartridge. A backup QFF (bQFF) was used in the QFF pack to evaluate the adsorption of gas-phase organics onto filter media. In addition, the breakthrough of gas-phase SVOCs was measured on selected sampling days to evaluate the performance of our sampling method.

## 7.2 METHODS

### 7.2.1 Sampling

Air samples were collected on the top of a two-story elementary school building in Denver. This site was located in a residential area 5.3 km east of downtown Denver. More details about this sampling site have been included in Vedel et al. (2009). Fifty pairs of 24-h gas- and particle-phase SVOC samples were collected from August 28, 2012 to July 25, 2013.

A medium volume sampler with two sampling trains was used to collect SVOCs in gaseous and particle phases (Figure 7.1). The sampler was equipped with a 2.5  $\mu\text{m}$  cut cyclone (University Research Glassware, URG) at a flow rate of 92 L  $\text{min}^{-1}$ . After the cyclone, the stream was split with 72 L  $\text{min}^{-1}$  passing through a QFF pack composed of two 90 mm diameter QFFs (Pall Gelman Tissuequartz<sup>TM</sup>) in tandem, followed by a cartridge composed of 5 g XAD-4 resin (Amberlite) “sandwiched” between two PUFs (50 mm diameter and 40 mm length each, URG), and 20 L  $\text{min}^{-1}$  passing through a 47 mm diameter, 2  $\mu\text{m}$  pore size TMF (Pall Gelman Teflo<sup>TM</sup>), followed by a similar PUF/XAD/PUF sandwich using XAD-7 resin (Amberlite) instead of XAD-4 resin. The top QFF (tQFF) in the filter pack was used for the analysis of all particle-phase organics (OC, EC and all target SVOCs). The bQFF was used to examine possible sampling artifacts due to gas sorption to filters. The PUF/XAD-4/PUF sandwich was used for the analysis of gas-phase *n*-alkanes and PAHs. The TMF was used for gravimetric ( $\text{PM}_{2.5}$  mass) analysis, and the PUF/XAD-7/PUF sandwich was used for the analysis of gas-phase 2-methylterols and levoglucosan.

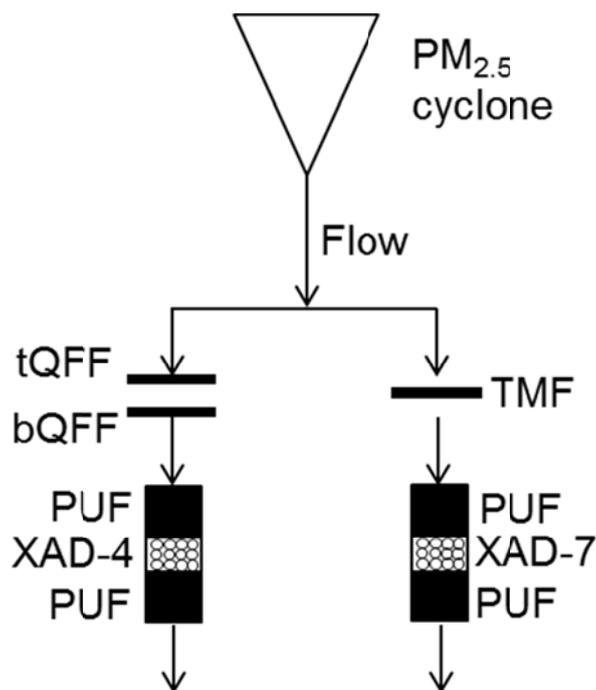


Figure 7.1 Scheme of the sampler used to collect SVOCs in gaseous and particle phases.

To evaluate the performance of this sampling system for gas-phase SVOCs collection, breakthrough experiments were conducted on eight sampling days. Two PUF (50 mm diameter and 20 mm length)/XAD-4 (5 g)/PUF and two PUF/XAD-7/PUF sandwiches in tandem were installed in corresponding cartridges to measure the breakthrough of gas-phase SVOCs. After February 12, 2013, the XAD-7 resin was replaced by XAD-4 resin for sampling, so as to understand which XAD resin is preferable in collecting gas-phase 2-methylterols and levoglucosan. The details on sampling dates, matrices used for gas- and particle-phase SVOCs collection are listed in Table 7.S1 in the supporting information. Field blank filters and PUF/XAD/PUF sandwiches were collected every 5<sup>th</sup>-6<sup>th</sup> samples to address contamination concerns.

Prior to sampling, QFFs were baked for 12 h at 500 °C. The PUF adsorbent was cleaned by soap water, tap water and deionized water, and then Soxhlet extracted for 24 h using a mixture of methanol and methylene chloride (1:1, v/v). XAD-4 and XAD-7 resin were cleaned with triplicate rinses of tap water, Mili-Q water and methanol, respectively, followed by Soxhlet extraction using methanol for 24 h and methylene chloride for 48 h. After cleaning, the PUF adsorbent was air dried for 3-4 h, and XAD resin were placed in a vacuum evaporator at 40-50 °C until no odor was detected. All prepared sampling matrices (QFFs, PUF adsorbent and XAD resin) were stored in pre-baked glass jars. Once the samples were collected and returned to the laboratory, they were stored in pre-baked glass jars at -20 °C until analysis.

## 7.2.2 Sample extraction and instrumental analysis

### 7.2.2.1 Carbon analysis

Details of the bulk OC and EC analysis were discussed by Dutton et al. (2009a). Briefly, a 1.5 cm<sup>2</sup> punch taken from each QFF sample was analyzed using the NIOSH 5450 thermal optical transmission (TOT) method (NIOSH, 2003; Schauer et al., 2003) on a Sunset Laboratory ECOC analyzer. The total OC carbon was composed of OC1, OC2, OC3, OC4 and PC, representing the carbon measured at four distinct temperature steps (340, 500, 615 and 900 °C) with a pyrolyzed carbon (PC) adjustment in the first heating cycle of the TOT method. The carbon measured during the second heating cycle with a final temperature of 910 °C, minus PC, made up EC. The OC and EC amount, ratio of the punch area (1.5 cm<sup>2</sup>) to total deposit area of the QFF and sample volume were used to obtain the final concentration. Concentration uncertainties were calculated using root sum of squares (RSS) method (NIST, 1994), incorporating the analytical uncertainties, standard deviation of the field blanks and the sampler volume uncertainties (Dutton et al., 2009a).

#### 7.2.2.2 Non-polar SVOC analysis

Details of QFF extraction and analysis for non-polar SVOCs were provided by Dutton et al. (2009b). Prior to extraction, a half of each QFF sample was spiked with an internal standard mixture containing isotopically labeled standards, which have similar structure to our target compounds. Table 7.S2 lists the internal standard compounds utilized in the current work. Methylene chloride was used to extract target compounds from QFFs ultrasonically. Then the extracts were filtered and concentrated to 150  $\mu\text{L}$  for analysis using an Agilent 6890N gas chromatograph (GC) coupled with an Agilent 5975 mass spectrometer (MS). Programmable temperature vaporization was used to improve the detection limit by injecting large volume (50  $\mu\text{L}$ ) extracts; and the GC was equipped with a HP-5ms capillary column (30 m  $\times$  0.25 mm  $\times$  0.25  $\mu\text{m}$ , Agilent). Samples were analyzed by GC-MS in autosampler sequences along with 5 dilutions of quantification standards. Quadratic calibration curves were generated for each target compound from all available runs of quantification standards in a given batch (3-4 sequences). The final mass amount of each target compound on QFF was determined by converting peak area ratios to mass ratios using calibration curves and known mass of pre-spiked internal standards. Quantification uncertainties derived from calibration curves were estimated empirically within each batch (Dutton et al., 2009b). *N*-Alkanes, PAHs, steranes, fatty acids, sterols and methoxyphenols were measured for each QFF sample. In this work, only *n*-alkanes and PAHs concentrations were presented, as most other species were not detected in gas phase, or had large quantification uncertainties due to blank contamination for PUF/XAD-4/PUF sandwiches.

For the analysis of non-polar SVOCs in PUF/XAD-4/PUF sandwiches, an internal standard mixture with the same species, but ten times concentrated as that for QFF analysis, was

injected to each sample before extraction. The PUF/XAD-4/PUF sandwiches were Soxhlet extracted using 250 mL methylene chloride for 24 h. After that, the extracts were rotary evaporated, filtered and blown down with N<sub>2</sub> to a final volume of 1.5 mL for GC-MS analysis. The quantification procedures for each non-polar SVOC were the same as that for QFF analysis. The two PUF/XAD-4/PUF sandwiches in tandem used for breakthrough experiment were extracted and quantified separately using the identical method as above. The final concentration of each non-polar SVOC in the air was obtained by the total amount of each compound quantified from the calibration curve and sample volume; and the associated uncertainty was calculated in a same manner as that described above.

#### 7.2.2.3 Polar SVOCs analysis

The second half of each QFF filter with a 1.5 cm<sup>2</sup> hole (for carbon analysis) was sonicated by 20 mL of methanol and methylene chloride (1:1, v/v) 2 times for 15 min each. The total extracts for each QFF sample were filtered through a pre-baked GFF (Pall Gelman Type A/E) to a 100 mL round flask, and concentrated to ~ 0.5 mL by rotary evaporator. After that, the extracts were transferred to a 2 mL glass vial, combined with 3 rinses of the round flask using a mixture of methanol and methylene chloride (1:2, v/v). The extracts were then blown to dryness under a gentle steam of ultrapure N<sub>2</sub> and reacted with 50 µL of *N*, *O*-bis (trimethylsilyl) trifluoroacetamide (BSTFA) containing 1% trimethylchloride and 10 µL of pyridine for 3 h at 70 °C, so as to convert COOH and OH groups into corresponding trimethylsilyl (TMS) esters and ethers. When the derivatives cooled down to room temperature, 80 µL of internal standard mixture (dodecane-*d*<sub>26</sub> 8.64 ng µL<sup>-1</sup>, hexadecane-*d*<sub>34</sub> 8.87 ng µL<sup>-1</sup> and tetracosane-*d*<sub>50</sub> 9.91 ng µL<sup>-1</sup> mixed in hexane) and 260 µL pure hexane were added before instrumental analysis. The resulting solution was analyzed by GC-MS operated in the electron ionization mode (70 eV). An

aliquot of 2  $\mu\text{L}$  of each sample was injected under splitless mode. The GC separation was carried out with a DB-5ms capillary column (30 m  $\times$  0.25 mm  $\times$  0.25  $\mu\text{m}$ , Agilent). The GC oven temperature was programmed from 80  $^{\circ}\text{C}$  (hold for 5 min) to 200  $^{\circ}\text{C}$  at 3  $^{\circ}\text{C min}^{-1}$ , then increased to a final temperature of 300  $^{\circ}\text{C}$  (hold for 10 min) at 15  $^{\circ}\text{C min}^{-1}$ . Linear calibration curves were derived from 5 dilutions of quantification standards. Dicarboxylic acids and saccharides were quantified by authentic standards; 2-methyltetros (2-methylthreitol and 2-methylerythritol) were quantified using meso-erythritol (Hu et al., 2008; Ding et al., 2012); other SOA tracers (e.g., 3-hydroxyglutaric acid,  $\beta$ -caryophyllinic acid) were quantified using *cis*-ketopinic acid (KPA). Those species that were not quantified using authentic standards were identified by the comparison of mass spectra with previous reported data (Claeys et al., 2004, 2007; Kleindienst et al., 2007). In this work, only 2-methyltetros and levoglucosan concentrations were shown and discussed, because other species were not detected in gas phase, or had very low recoveries for the analysis of PUF/XAD/PUF sandwiches. The slope of calibration curve for each compound changed across different sequences, so the calibration curve of each compound was only used to analyze samples in the same sequence (12 – 16 samples), and the quantification uncertainties derived from calibration curves could not be estimated.

For the analysis of gas-phase polar SVOCs in the PUF/XAD-7/PUF or PUF/XAD-4/PUF sandwiches, each sample was Soxhlet extracted using a mixture of 230 mL methylene chloride and 20 mL of methanol. Then the extracts were rotary evaporated, filtered and concentrated to a final volume of 1.5 – 2 mL. After that, the extracts were blown down to dryness and derivatized with 100  $\mu\text{L}$  BSTFA (1% trimethylchloride) and 20  $\mu\text{L}$  pyridine at 70  $^{\circ}\text{C}$  for 3 h. Finally, 80  $\mu\text{L}$  of internal standard mixture as that for QFF analysis and 200  $\mu\text{L}$  of hexane were added to the resulting derivatives before GC-MS analysis. The quantification procedures of each polar SVOC

were the same as that for QFF analysis. The tandem PUF/XAD/PUF sandwiches of each pair of breakthrough samples were extracted separately and quantified using the identical method as above.

In addition, the top PUF (tPUF), middle XAD resin and bottom PUF (bPUF) of selected PUF/XAD/PUF samples ( $N = 8$ ) were extracted and quantified separately, so as to understand which part of the sandwich was mostly enriched with the compounds of interest. The PUF adsorbent (top + bottom) and XAD resin of two pairs of breakthrough samples (Sample ID: S5 and S21; Table 7.S1) were also analyzed separately. The PUF adsorbent was analyzed using the same method as that for regular PUF/XAD/PUF samples; while the XAD resin were analyzed using the same method for QFF samples, but extracted with more solvents ( $40 \text{ mL} \times 2$  times).

#### 7.2.2.4 Quality assurance and control

The variability in recoveries for non-polar SVOCs could be accounted for by those pre-spiked internal standards, so no additional recovery experiment was conducted. Recoveries of polar SVOCs were obtained by spiking standards on to prebaked blank QFF, blank PUF adsorbent and XAD resin, and blank PUF/XAD/PUF sandwiches for regular and breakthrough sampling, followed by extraction and quantification in the same way as that for collected samples. Details of the recoveries for the target compounds in this work are listed and discussed in the supporting information. Briefly, the recoveries of meso-erythritol were higher than 70% (79.3 – 103.8% on average) for all collected media (Table 7.S3). The recoveries of levoglucosan were higher than 70% (70.7 – 95.2%) in QFF, XAD resin and PUF adsorbent, but lower than 70% (51.9 – 63.3%) in PUF/XAD/PUF sandwiches for regular or breakthrough sampling (Table 7.S3). The concentrations of 2-methyltetrols and levoglucosan quantified from different sampling matrices in the current work were adjusted by their corresponding recoveries, so as to obtain

correct G/P distribution of these compounds. All measurements in this study were field-blank corrected if necessary. To reduce the sensitivity to occasional outliers in field blanks, the median rather than the mean value of field blanks was subtracted from all observations.

## 7.3 RESULTS AND DISCUSSION

### 7.3.1 OC and EC concentrations of QFF samples

In Table 7.1, the EC concentrations quantified from the tQFF, where  $PM_{2.5}$  was loaded, ranged from 0.06 to 1.83  $\mu\text{g m}^{-3}$  with an average of 0.40  $\mu\text{g m}^{-3}$ ; while no EC was observed on the bQFF. Among the five OC fractions on the tQFF, OC1 has the highest concentration ranging from 0.75 to 3.23  $\mu\text{g m}^{-3}$  with an average of 1.58  $\mu\text{g m}^{-3}$ , followed by OC2 (average, 0.60  $\mu\text{g m}^{-3}$ ) and OC4 (0.61  $\mu\text{g m}^{-3}$ ) (Table 7.1). Only OC1 has been observed on bQFFs with an average concentration of 0.48  $\mu\text{g m}^{-3}$ , accounting for 28% (range, 15 – 56%) of that on tQFFs. The concentrations of OC2, OC3, OC4 and PC that evolved off the bQFF at higher temperatures ( $>340\text{ }^{\circ}\text{C}$ ) were comparable or lower than field blanks. The OC1 concentrations from bQFFs were significantly correlated with those from tQFF ( $r = 0.67, p < 0.01$ ) and ambient temperatures ( $r = 0.41, p < 0.01$ ), consistent with the fact that the artifacts for particulate organic sampling using QFF was caused by volatile or semi-volatile organics. Subramanian et al. (2004) estimated the positive and negative artifacts in particulate organic sampling by comparing two different sampling approaches (tQFF or TMF/bQFF, denuder/QFF/carbon-impregnated GFF). They found that using a bQFF after the tQFF could reasonably estimate the positive artifact due to gas-phase sorption for 24-h samples; while the negative artifact from volatilization loss on the tQFF was small and negligible. As a result, future work should subtract OC1 concentrations on bQFFs from those on tQFFs to study G/P partitioning of SVOCs.

Table 7.1 Statistics for each carbon fraction on QFF samples ( $N=50$ ,  $\mu\text{g m}^{-3}$ )

Carbon fractions	No. of Obs. <sup>a</sup>	Median	Mean	Range	S/N <sup>b</sup>
tQFF					
EC	48	0.28	0.40	0.06 - 1.83	3.2
OC1	48	1.46	1.58	0.75 - 3.23	3.2
OC2	46	0.55	0.60	0.01 - 2.48	2.1
OC3	48	0.35	0.41	0.07 - 1.33	2.5
OC4	48	0.51	0.61	0.06 - 2.37	0.8
PC	48	0.29	0.39	0.06 - 2.44	2.5
bQFF					
OC1	49	0.45	0.48	0.18 - 1.13	3.1

(a) Number of observations out of the 50 samples.

(b) Signal to noise ratio (mean concentration/mean uncertainty).

### 7.3.2 Breakthrough experiments

Eight pairs of breakthrough samples were collected to evaluate the efficiency of gas-phase SVOCs collection. The breakthrough value for each SVOC is calculated from

$$B = \frac{[\text{PXP}]_{\text{back}}}{[\text{PXP}]_{\text{front}} + [\text{PXP}]_{\text{back}}} \times 100\% \quad (1)$$

where B (%) is the gas-phase breakthrough value and [PXP] is the concentration ( $\text{ng m}^{-3}$ ) of specific SVOC in the front or back PUF/XAD/PUF sandwiches (Hart et al., 1992). 50% or higher for B value indicate complete breakthrough, and small B values (e.g., <10%) suggest high collecting efficiency. A value of 33% was used as critical value to indicate excessive breakthrough (Peters et al., 2000; Ahrens et al., 2011).

Table 7.2 Gas-phase breakthrough (%) for individual SVOCs.

Compounds	MW	No of Obs. <sup>a</sup>	Median	Mean	Range
<i>n-Alkanes</i>					
dodecane	170	8	1.2	2.4	0 – 8.4
tridecane	184	8	1.4	2.0	0 – 8.0
tetradecane	198	8	0.9	1.6	0 – 8.2
pentadecane	212	8	0.3	1.6	0 – 7.9
hexadecane	226	8	0.5	2.4	0 – 9.4
heptadecane	240	8	0.3	2.8	0 – 9.2
octadecane	254	7	0	2.9	0 – 10
nonadecane	268	6	0	0	/
eicosane	282	4	0	0	/
heneicosane	296	3	0	0	/
docosane	310	3	0	0	/
tricosane	324	3	0	0	/
tetracosane	338	2	0	0	/
pentacosane	352	2	0	0	/
hexacosane	366	2	0	0	/
heptacosane	380	2	0	0	/
<i>PAH</i>					
naphthalene	128	8	1.7	3.0	0 – 12
2-methylnaphthalene	142	8	1.2	1.8	0 – 8.2
1-methylnaphthalene	142	8	0	0.7	0 – 5.1
acenaphthylene	152	8	0	0.6	0 – 3.8
acenaphthene	154	8	0.6	1.2	0 – 4.2
fluorene	166	8	0	0.5	0 – 2.4
2-methylfluorene	180	8	0	2.3	0 – 14
phenanthrene	178	8	0	0.9	0 – 4.7
anthracene	178	5	0	0	/
methyl-178-PAH	192	5	0	0	/
fluoranthene	202	6	0	0	/
pyrene	202	3	0	0	/
methyl-202-PAH	216	3	0	0	/
<i>Polar SVOCs</i>					
2-Methylthreitol	136	3	0	0	/
2-Methylerythritol	136	3	0	0	/
Levoglucosan	162	5	0	11	0 - 31

(a) Number of observations out of 8 samples.

The statistics of breakthrough values for *n*-alkanes and PAH with molecular weight (MW) ranging from 170 to 380 and from 128 to 216, respectively, are listed in Table 7.2. Other *n*-alkanes and PAHs with higher MW were not observed in breakthrough samples. In some front and back PUF/XAD-4/PUF sandwiches, *n*-alkanes heavier than octadecane (MW = 254) and PAHs heavier than anthranthene (MW = 202) were not detected or comparable to field blanks before correction. Missing values and measurements below detection limit (BDL), defined as two times as the uncertainty of final concentration, for those compounds were not included for breakthrough calculation. In Table 7.2, the average breakthrough values for all *n*-alkanes and PAHs are close or equal to 0. Hart et al. (1992) and Peters et al. (2000) collected gas-phase PAHs using only PUF adsorbent, and concluded that the PUF adsorbent was unsuitable for sampling volatile PAHs (e.g., naphthalene) due to high breakthrough (> 33%). In this work, the low breakthrough values suggested that a combination of PUF adsorbent and XAD-4 resin could provide very high efficiency (~ 100%) in collecting volatile *n*-alkanes (e.g., dodecane) and PAHs (e.g., naphthalene).

Seven out of the eight pairs of breakthrough samples were analyzed for 2-methyltetrols (2-methylthreitol and 2-methylerythritol) and levoglucosan. The first pair of breakthrough samples collected on September 9, 2012 was used to test the extraction and quantification methods, and not included for data analysis. 2-Methylterols were only observed in three front PUF/XAD/PUF sandwiches of breakthrough samples (2-methylthreitol, 1.06 – 1.71 ng m<sup>-3</sup>; 2-methylerythritol, 2.79 – 4.87 ng m<sup>-3</sup>) collected on summer days (September 15 and 21, 2012 and July 8, 2013) with ambient temperatures from 19.5 to 29.4 °C. This is consistent with the fact that 2-methyltetrols were isoprene derived SOA tracers mostly observed in summer (Kleindienst et al., 2007). No 2-methylterols was observed on the other four breakthrough sampling days with

ambient temperature ranging from -12.6 to 10.8 °C. Levoglucosan was observed in five front PUF/XAD/PUF sandwiches of breakthrough samples (2.16 – 33.9 ng m<sup>-3</sup>) (Sample ID: S5, S6, S14, S21 and S47; Table 7.S1) and two back PUF/XAD/PUF sandwiches of samples S6 and S21. The breakthrough of levoglucosan derived from sample S6 (25%) might be partly attributed to the uncertainties in quantification, since low concentrations were observed in the front (5.60 ng m<sup>-3</sup>) and back (1.91 ng m<sup>-3</sup>) PUF/XAD/PUF sandwiches. The PUF adsorbent and XAD resin in the front and back PUF/XAD/PUF sandwiches of sample S21 were analyzed separately, and the levoglucosan was detected for each of the four individual analyses (front PUF 16.0 ng m<sup>-3</sup>, front XAD 17.9 ng m<sup>-3</sup>; back PUF 5.96 ng m<sup>-3</sup>, back XAD 9.22 ng m<sup>-3</sup>), resulting in a breakthrough value of 31%. The high breakthrough observed for sample S21 might be related to the high gas-phase concentration of levoglucosan on that sampling day, and the front PUF/XAD/PUF sandwich could not capture all of them. The breakthrough values summarized in Table 7.2 suggested that the PUF/XAD/PUF sandwich was suitable for sampling gas-phase 2-methyltetrols and levoglucosan.

The tPUF, middle XAD resin and bPUF of selected PUF/XAD/PUF samples ( $N = 8$ ) were analyzed separately, and the results were shown in Table 7.S4 in the supporting information. XAD-4 resin was applied instead of XAD-7 resin for sampling after sample S26, and each type of XAD resin was used in four selected samples presented in Table 7.S4. 2-Methyltetrols and levoglucosan were only observed in the tPUF of those selected samples, but not in the XAD resin or bPUF. The concentration of gas-phase levoglucosan observed in sample S44 (44.1 ng m<sup>-3</sup>, Table 7.S4) was comparable to that of sample S21 (49.1 ng m<sup>-3</sup>). But unlike sample S21, all gas-phase levoglucosan was captured by the tPUF. One explanation is because the length of PUF adsorbent in PUF/XAD/PUF sandwiches (20 mm) used for breakthrough sampling is half of that

for regular sampling (Table 7.S1) due to the limited size of the adsorbent holder, resulting in a reduction in retention time. These results suggested that the PUF adsorbent could be applied in collecting gas-phase 2-methyltetrols and levoglucosan. Unfortunately, we could not infer which XAD resin (XAD-4 or 7) was preferable for sampling.

### 7.3.3 Concentrations of target SVOCs on QFFs and adsorbents

#### 7.3.3.1 *n*-Alkanes and PAHs

The statistics for concentrations of each species were listed in Table 7.S5 in the supporting information, including the number of observations, median and mean concentrations, concentration range and signal to noise ratio (S/N). The median, mean concentrations and S/N ratios were calculated from observed values, including those BDL measurements. The first two PUF/XAD-4/PUF samples collected on August 22 and 28, 2012 were used to test the methods and not included for data analysis. In Table 7.S5, *n*-alkanes heavier than tricosane ( $C_{23}$ , MW = 324) observed in the PUF/XAD-4/PUF samples have low S/N ratios ( $< 2$ ), suggesting that most of the observations were below detection limit. Additionally, those species are not observed for more than half of all PUF/XAD-4/PUF samples. In the bQFF, those same *n*-alkanes have much lower (one order of magnitude) median concentrations when compared to those in the tQFF or have S/N ratios lower than 2. Thus, we compared the median concentrations of *n*-alkanes with a chain length from 12 to 23 across all sampling matrices in Figure 7.2a. For *n*-alkanes with MW lower than eicosane, the median concentrations of the tQFF were comparable or lower than those of the bQFF, and at least 5 times lower than those of the PUF/XAD-4/PUF samples, suggesting that these species were primarily existing in gas phase, and the concentrations of these species quantified from the tQFF were likely caused by gas-phase adsorption to filter media. As the MW of *n*-alkanes increased, the median concentrations of the tQFF became higher than those of the

bQFF and PUF/XAD-4/PUF samples, indicating that the MW, corresponding to vapor pressure, was an internal factor determining the G/P partitioning of *n*-alkanes.

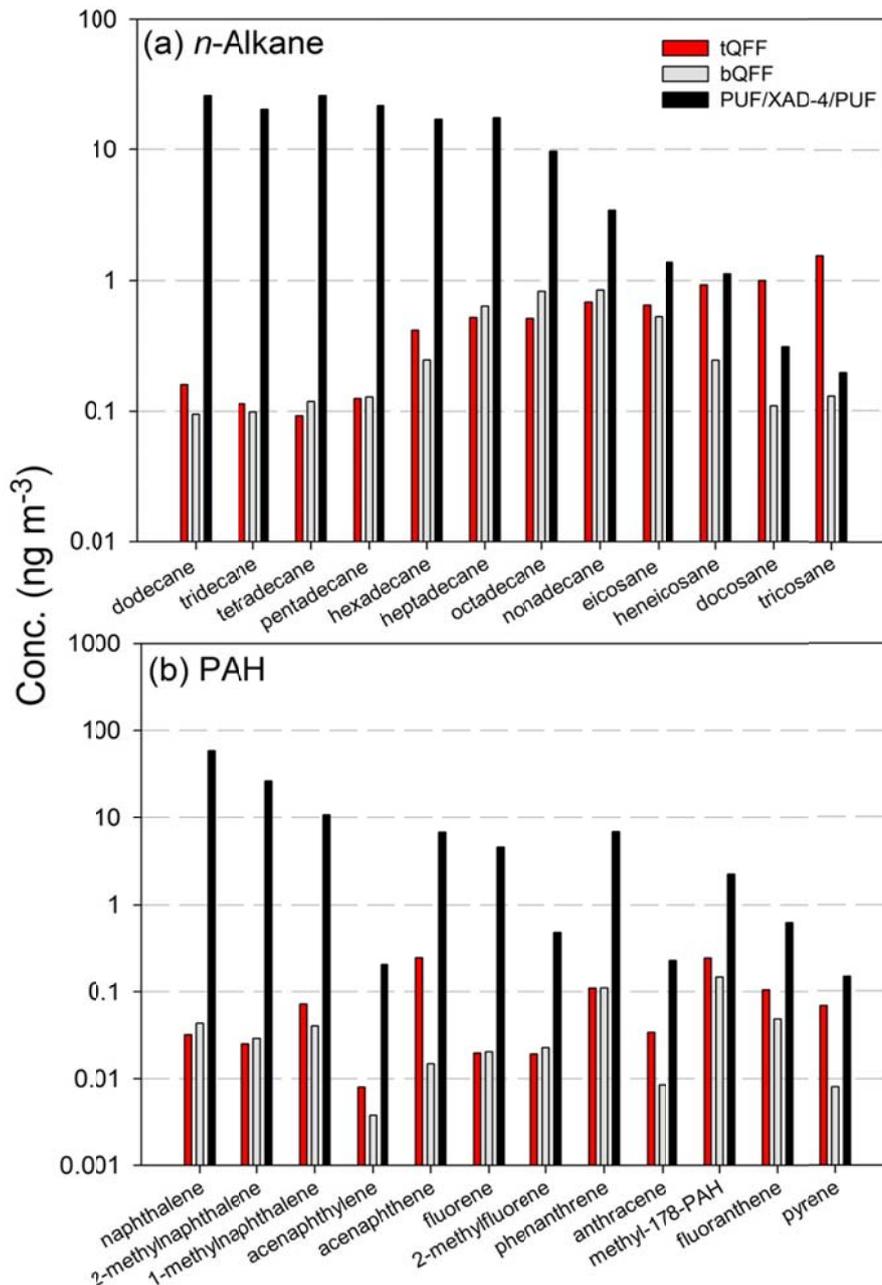


Figure 7.2 Comparisons of median concentrations of (a) *n*-alkanes (MW = 170 – 324), (b) PAHs (MW = 128 – 202) across tQFFs, bQFFs and PUF/XAD-4/PUF samples.

As shown in Table 7.S5, PAHs heavier than pyrene (MW = 202) were rarely (< 50% of sample number) observed in PUF/XAD-4/PUF samples; likewise, in the bQFF, those species had either low S/N ratios (< 2) or much lower concentrations compared to those in the tQFF. Thus only PAHs with MW from 128 to 202 were compared across tQFFs, bQFFs and PUF/XAD-4/PUF samples in Figure 7.2b. The higher median concentrations of acenaphthylene and acenaphthene in the tQFF compared to those in the bQFF were caused by the co-elution of these two species with other compounds. The S/N ratio of 2-methylflorene, phenanthrene and anthracene in the tQFF were lower than 2 (Table 7.S5). Except those species, other PAHs with MW lower than 192 (methyl-178-PAH) in the bQFF have comparable or higher median concentrations than those in the tQFF; and the median concentrations of these compounds in PUF/XAD-4/PUF samples were at least 10 times higher than those in QFF samples. These results indicated that PAHs lighter than methyl-178-PAH were mostly in gas phase. Similar as *n*-alkanes, in the tQFF, the median concentrations of heavier PAHs (methyl-178-PAH, fluoranthene and pyrene) became higher than those in the bQFF; and the relative abundance of those PAHs in the tQFF to the PUF/XAD-4/PUF increased with MW.

#### 7.3.3.2 2-Methyltetrols and levoglucosan

2-Methylthreitol and 2-methylerythritol were both isoprene derived SOA tracers (Claeys et al., 2004) and strongly correlated ( $r = 0.86 - 1.00$ ) in tQFFs, bQFFs and PUF/XAD/PUF samples. Thus these two species were added together as 2-methyltetrols, and the concentrations in different sampling matrices were shown in Figure 7.3a. The 2-methyltetrols were mostly observed in summer for all sampling matrices. The highest concentrations were  $27.5 \text{ ng m}^{-3}$  and  $17.0 \text{ ng m}^{-3}$  for the tQFF and bQFF collected on June 26, 2013 (Sample ID: S45). However, the PUF/XAD/PUF samples collected on the same day and June 3, 2013 (Sample ID: S43) were not

extracted successfully, and no data were obtained. In summer periods, the concentrations of 2-methyltetrols from bQFFs accounted for up to 64.2% of those from tQFFs; and the PUF/XAD/PUF samples for which 2-methyltetrols were detected ( $N = 14$ ), had concentrations 0.44 – 2.06 times as those for tQFFs. As a result, the summertime 2-methyltetrols in the air should be subject to G/P partitioning.

Levoglucosan has been identified and used as a biomass burning tracer (Simoneit et al., 1999; Jeffrey et al., 2007). In Figure 7.3b, the tQFF associated levoglucosan exhibited high concentrations mostly in winter. The two extreme concentrations of levoglucosan in the tQFF were observed on January 13, 2013 ( $424 \text{ ng m}^{-3}$ ; Sample ID: S21) and June 20, 2013 ( $439 \text{ ng m}^{-3}$ ; Sample ID: S44), respectively. The high concentrations of particle-phase levoglucosan in wintertime could be ascribed to more biomass burning due to cold weather and the lower inversion layer height. While the summertime extreme concentration observed in this work might be caused by unexpected biomass burning close to the sampling site, since other biomass burning tracers (e.g., retene, methoxyphenols) also exhibited sudden increases on that sampling day. As shown in Figure 7.3b, levoglucosan was only detected in three bQFFs with low concentrations ( $0.09 - 5.22 \text{ ng m}^{-3}$ ), and the concentrations in corresponding tQFFs ( $179 - 434 \text{ ng m}^{-3}$ ) and PUF/XAD/PUF samples ( $26.6 - 49.1 \text{ ng m}^{-3}$ ) were the highest. So the positive sampling artifact might have negligible impact on the measurement of gas- or particle-phase levoglucosan. The concentration ratios of levoglucosan in PUF/XAD/PUF samples for which levoglucosan was detected ( $N = 39$ ) to those in tQFFs ranged from 0.02 to 3.04 with a median value of 0.23. Thus, the airborne levoglucosan was also subject to G/P partitioning across the sampling period.

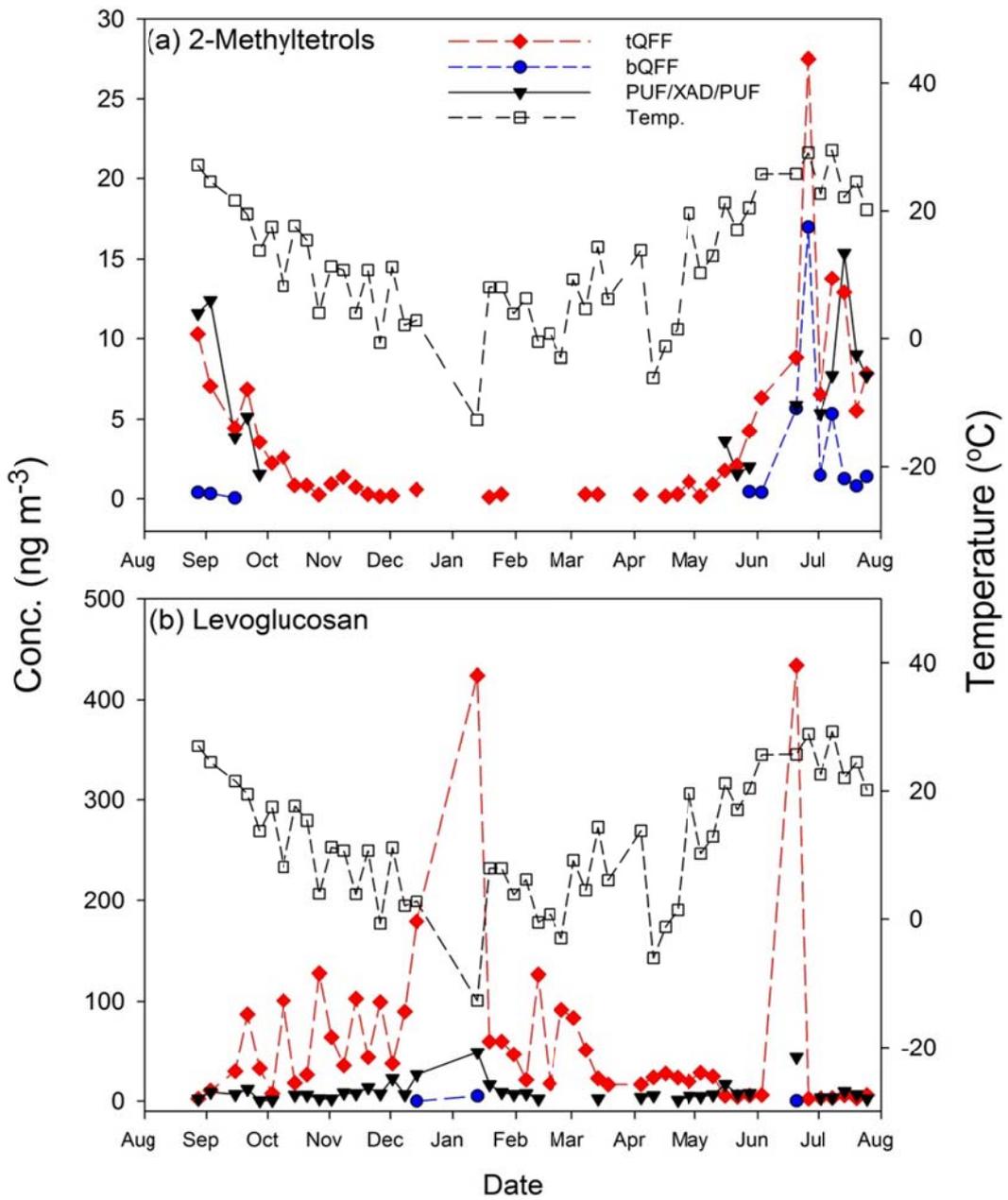


Figure 7.3 Concentration time series of 2-methyltetrols and levoglucosan in tQFFs, bQFFs and PUF/XAD/PUF samples.

## 7.4 CONCLUSIONS

In this study, QFFs and PUF/XAD/PUF sandwiches were coupled in a medium volume sampler to collect gas- and particle-phase SVOCs. The positive sampling artifacts of particulate OC primarily biased the measurement of OC1 fraction. The PUF/XAD-4/PUF sandwich has been demonstrated as high efficient (~ 100%) adsorbent in collecting gas-phase *n*-alkanes and PAHs (e.g., dodecane, naphthalene), and should be applied in field sampling instead of using PUF adsorbent only. Based on breakthrough experiments, PUF/XAD-7/PUF or PUF/XAD-4/PUF sandwich can collect most of the gas-phase 2-methyltetrols and levoglucosan in the sampled air. However, the analysis method needs to be refined to increase the recoveries for levoglucosan in PUF/XAD/PUF sandwiches. Individual analysis for different parts of PUF/XAD/PUF sandwiches showed that using PUF adsorbent only could also collect gas-phase 2-methyltetrols and levoglucosan with high efficiency. Unfortunately, no data were obtained to support which XAD resin was more suitable for sampling those polar species. The comparison of species concentrations across tQFFs, bQFFs and PUF/XAD/PUF samples suggested that the light *n*-alkanes (MW < 282) and PAHs (MW < 192) were prone to exist mostly in gas phase; while the heavier *n*-alkanes (MW > 324) and PAHs (MW > 202) were primarily in particle phase. Like many other SVOCs, 2-methyltetrols and levoglucosan are also subject to G/P partitioning in the air, which needs to be considered when only using particle-phase data for receptor-based source apportionment. Future work will apply the gas- and particle-phase concentrations of selected SVOCs to fit the absorptive G/P partitioning model, so as to verify the estimation of gas-phase SVOCs in the Xie et al. (2013a) study.

**CHAPTER 8 GAS/PARTICLE PARTITIONING OF SEMI-VOLATILE ORGANIC  
COMPOUNDS IN DENVER-FITNESS OF ABSORPTIVE PARTITIONING THEORY**

Mingjie Xie <sup>a</sup>, Michael P. Hannigan<sup>a</sup>, Kelly C. Barsanti<sup>b</sup>

<sup>a</sup> Department of Mechanical Engineering, College of Engineering and Applied Science,  
University of Colorado, Boulder, CO 80309, USA

<sup>b</sup> Department of Civil and Environmental Engineering, Portland State University, PO Box 751,  
Portland, OR 97207, USA

## 8.0 ABSTRACT

To verify the estimation of gas-phase semi-volatile organic compounds (SVOCs) using an equilibrium absorption model in the Xie et al., (2013a) study, this work compared observationally-based gas/particle (G/P) partitioning coefficients ( $K_{p,OM}^m$ ) of selected SVOCs with their predicted vapor pressures ( $p_L^\circ$ ) and theoretically-based partitioning coefficients ( $K_{p,OM}^t$ ). Measured concentrations of those SVOCs (two *n*-alkanes, two PAHs, two oxy-PAHs, 2-methyltetrols and levoglucosan) in different sampling media were from Chapter 7. Gas- and particle-phase concentrations of *n*-alkanes, PAHs and oxy-PAHs were calculated with three approaches for artifact corrections based on the measurement of backup quartz fiber filter (bQFF), and used to calculate their  $K_{p,OM}^m$ . If the bQFF associated *n*-alkanes, PAHs and oxy-PAHs of interest were dominated by positive artifact (vapor phase adsorption) or equally contributed by positive and negative (particle phase evaporation) artifacts, the correlations between  $\log K_{p,OM}^m$  and  $\log p_L^\circ$  were all significant ( $p < 0.05$ ). However, the regression slopes ( $m$ ) deviating from -1 and their wide 95% confidence interval (CI) suggested deviations from true equilibrium. For less volatile 2-methyltetrols and levoglucosan, the bQFF concentrations were assumed to be dominated by positive sampling artifacts and subtracted from the top QFF concentrations as the particle-phase; while the gas-phase concentrations were corrected with two different assumptions about the adsorption of Teflon membrane filter (TMF). The constant like  $K_{p,OM}^m$  value for 2-methyltetrols (no TMF adsorption,  $0.19 \pm 0.06$ ; with TMF adsorption,  $0.16 \pm 0.05$ ) indicated that the variation in particulate organic material dominated the G/P partitioning of 2-methyltetrols in summer Denver. The regression of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  ( $m = -0.91 \pm 0.30$ ,  $r = 0.72$ ) for levoglucosan indicated a G/P partitioning close to equilibrium. Comparing the values of  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$ , the two PAHs and levoglucosan had the best agreement. The

median  $K_{p,OM}^t$  values of the two oxy-PAHs deviated from (2-3 orders of magnitudes lower) their  $K_{p,OM}^m$  values most, which might be partly due to the ignored mechanism of strong surface interaction other than simple physical adsorption.

## 8.1 INTRODUCTION

Gas/particle (G/P) partitioning is a key process affecting the environmental fate, exposure to wild life and humans, and long range transport of semi-volatile organic compounds (SVOCs) (Liang and Pankow, 1997). This process has been studied for decades, and successfully described by an equilibrium partitioning coefficient  $K_p$  ( $\text{m}^3 \mu\text{g}^{-1}$ ) (Yamasaki et al., 1982; Pankow, 1991, 1992):

$$K_p = \frac{F / TSP}{A} \quad (1)$$

where  $F$  ( $\text{ng m}^{-3}$ ) and  $A$  ( $\text{ng m}^{-3}$ ) are the particle- and gas-phase concentrations of each SVOC; and TSP ( $\mu\text{g m}^{-3}$ ) is the concentration of total suspended particulate matter in the air. When plotting  $K_p$  value versus the corresponding liquid saturation vapor pressure ( $p_L^\circ$ , atm) of the target SVOC for a given compound class and particle type, a linear correlation in the form of

$$\log K_p = m \log p_L^\circ + b \quad (2)$$

has been found (Pankow, 1987, 1994a, b). The slope  $m$  is usually close to  $-1$ , but may deviate significantly from  $-1$  for equilibrium partitioning (Goss and Schwarzenbach, 1998). The underlying mechanisms include simple physical adsorption to particle surfaces and absorptive partitioning to particulate organics (Pankow, 1994a, b). Absorptive partitioning has been demonstrated as the dominant mechanism for the G/P partitioning of SVOCs (Liang and Pankow et al., 1997; Goss and Schwarzenbach, 1998; Mader and Pankow, 2002). Thus,  $K_p$  could be normalized by the weight fraction of the absorptive organic material ( $f_{\text{OM}}$ ) in the total PM phase to obtain absorptive G/P partitioning coefficient ( $K_{\text{p,OM}}$ ,  $\text{m}^3 \mu\text{g}^{-1}$ ), which could either be measured directly ( $K_{\text{p,OM}}^{\text{m}}$ , Eq. 3) or calculated from theory ( $K_{\text{p,OM}}^{\text{t}}$ , Eq. 4):

$$K_{\text{p,OM}}^{\text{m}} = \frac{K_p^{\text{m}}}{f_{\text{OM}}} = \frac{F / M_{\text{OM}}}{A} \quad (3)$$

$$K_{p,OM}^t = \frac{RT}{10^6 \overline{MW}_{OM} \zeta_{OM} p_L^o} \quad (4)$$

where  $K_{p,OM}^m$  in Eq.(3) is observationally-based  $K_p$ , and  $M_{OM}$  ( $\mu\text{g m}^{-3}$ ) is the mass concentration of the particle-phase organic material (OM). In Eq. (4),  $R$  ( $\text{m}^3 \text{atm K}^{-1} \text{mol}^{-1}$ ) is the ideal gas constant;  $T$  (K) is the ambient temperature;  $\overline{MW}_{OM}$  ( $\text{g mol}^{-1}$ ) is the mean molecular weight (MW) of the absorbing OM phase;  $\zeta_{OM}$  is the mole fraction scale activity coefficient of each compound in the absorbing OM phase; and  $p_L^o$  (atm) is the vapor pressure of each pure compound.

Receptor-based source apportionment of  $\text{PM}_{2.5}$  usually uses particle-phase SVOCs data as inputs (Jaekels et al., 2007; Shrivastava et al., 2007), and the source/factor profile is pre-assumed as constant over the period of ambient and source sampling (Chen et al., 2011). However, all SVOCs are subject to G/P partitioning, and the particle-phase fraction can change with ambient temperature for a given SVOC and OM phase. As such, the output factors are not necessarily pollution sources, and could also reflect the influences from G/P partitioning of SVOCs, especially those factors characterized by light SVOCs (Xie et al., 2013b). To eliminate the influences from G/P partitioning, Xie et al (2013a) calculated gas-phase concentrations of SVOCs from their particle-phase concentrations by absorptive G/P partitioning theory (Eqs. 3, 4), and added the gas- and particle-phase SVOCs together for source apportionment. The results showed consistent factor contributions between full data set (32-month series) solution and temperature stratified sub-data sets solutions, suggesting that using total SVOCs (gas + particle phase) data for source apportionment could eliminate the influences of G/P partitioning. To verify the estimation of gas-phase SVOCs in Xie et al. (2013a), the author collected fifty pairs of gas- and particle-phase samples of SVOCs from August 2012 to July 2013 in urban Denver (Chapter 7). A medium volume sampler incorporating quartz fiber filter (QFF) and polyurethane

foam (PUF)/XAD/PUF sandwich was used throughout the sampling campaign. A backup QFF (bQFF) was used to estimate possible sampling artifacts of particulate organics. The breakthrough experiments suggested that the PUF/XAD/PUF sandwich (PXP) had very high efficiency in sampling gas-phase *n*-alkanes, PAHs, 2-methylterols and levoglucosan.

Most previous G/P partitioning studies investigated *n*-alkanes and PAHs using PUF adsorbent only (Fraser et al., 1997; Kim et al., 2012; Wang et al., 2013), which could not collect volatile *n*-alkanes and PAHs efficiently (Hart et al., 1992; Peters et al., 2000). To minimize the positive sampling artifact of particulate organics (gas-phase organics adsorbed on QFF media), denuders have been used to remove gaseous organics before particle collection on QFF (Schauer et al., 1999; Peters et al., 2000; Ahrens et al., 2011, 2012). But considerable particle loss has been observed by Zhang et al. (2012) when applying denuder for the measurements of diesel engine emissions. The negative sampling artifacts of particulate organics (particle-phase organics evaporation from QFF media) were found to be negligible by Schauer et al. (1999) and Subramanian et al. (2004). In this work, potential sampling artifacts (positive and negative) were evaluated based on the measurement of bQFF samples. Additionally, a few recent studies have investigated the G/P partitioning of secondary organic aerosols (e.g., carbonyls) (Healy et al., 2008; Perraud et al., 2012; Kawamura et al., 2013; Zhao et al., 2013), but very few studies looked at those polar SVOCs (e.g., levoglucosan) commonly referred to as organic molecular markers for source apportionment (Bao et al., 2012; Zhao et al., 2013).

In this work, the gas- and particle-phase concentrations of selected SVOCs (two *n*-alkanes, two PAHs, two oxy-PAHs, 2-methyltetrols and levoglucosan) were calculated based on the field measurement from Chapter 7. The measurements of bQFF associated organics were used for artifact corrections in three different ways for *n*-alkanes, PAHs and oxy-PAHs, and two

different ways for 2-methyltetrols and levoglucosan. Then the corrected gas- and particle-phase SVOCs concentrations were used to calculate  $K_{p,OM}^m$  according to Eq. (3). The  $\log K_{p,OM}^m$  values of each SVOC were regressed to the log values of predicted vapor pressure ( $\log p_L^0$ ; Eq. 2), so as to understand if those selected SVOCs satisfy equilibrium G/P partitioning. Finally, the values of  $\log K_{p,OM}^m$  were compared to those calculated theoretically ( $\log K_{p,OM}^t$ , Eq. 4).

## 8.2 METHODS

### 8.2.1 Sampling, chemical analysis and SVOC selection

Fifty 24-h gas- and particle-phase ( $PM_{2.5}$ ) samples of SVOCs were collected from August 2012 to July 2013 on the top of a two-story elementary school building in urban Denver. Details of the sampler set up, sampling protocols and chemical analysis were provided by Chapter 7. Besides those speciated SVOCs (*n*-alkanes, PAHs, oxy-PAHs, 2-methyltetrols and levoglucosan) measured in gaseous and particle phases, concentrations of bulk organic carbon (OC) and elemental carbon (EC) were measured for the top QFF (tQFF, where  $PM_{2.5}$  was loaded) and bQFF samples.

The light SVOC factor in the Xie et al. (2013b) study is characterized by docosane, tricosane, fluoranthene, pyrene, 1,8-naphthalic anhydride, anthracene-9,10-dione and dodecanoic acid. The two oxy-PAHs and dodecanoic acid could not be quantified for gas-phase samples in Chapter 7. So the two light *n*-alkanes (docosane, tricosane) and two light PAHs (fluoranthene, pyrene) were selected for G/P partitioning analysis. Moreover, two lighter oxy-PAHs (acenaphthenone and fluorenone) that quantified in most QFF and PUF/XAD/PUF samples, but not presented in Chapter 7, were included for analysis. 2-Methyltetrols and levoglucosan were also included, because they were observed in both gaseous and particle phases (Chapter 7) and their G/P partitioning had not been studied yet. The properties and concentrations of those selected SVOCs are listed in Tables 8.S1 and 8.S2 in the supporting information.

### 8.2.2 Artifact corrections using bQFF measurements

In Table 8.1, we listed three assumptions on the origin of *n*-alkanes, PAHs and oxy-PAHs associated with the bQFF, and corresponding approaches for gas- and particle-phase corrections. If the SVOCs observed on bQFF are dominated by positive artifact, then the particle-phase

concentration ( $F$ ) of each SVOC on each day will be calculated as the tQFF concentration minus bQFF concentration if the tQFF concentration is higher, and the gas-phase concentration ( $A$ ) will be the PXP concentration adds two times the bQFF concentration; or the  $F$  value will be 0 if the tQFF concentration is lower than that of bQFF, and the  $A$  value will be the sum of PXP, tQFF and bQFF concentrations. If the bQFF associated SVOCs are dominated by negative artifact, then the  $F$  value will be the sum of tQFF and bQFF concentrations, and the  $A$  value will be the PXP concentration. If the SVOCs on bQFF are equally contributed by positive and negative artifacts (positive offsets negative), then the  $F$  value will be the tQFF concentration, and the  $A$  values will be the sum of the PXP and bQFF concentrations. Under the third assumption, the calculated  $F$  and  $A$  values should be similar as those from traditional sampling without bQFF. The three different approaches of artifact corrections are named as C1, C2 and C3 in Table 8.1. In Table 8.S1, the vapor pressures of 2-methyltetrols and levoglucosan are more than one order of magnitudes lower than other species, so the concentrations on the bQFF are assumed to come from positive artifacts. However, the PXP samples collected for the analysis of 2-methyltetrols and levoglucosan were installed in a different sampling train, right below the Teflon membrane filter (TMF; Chapter 7). We considered two conditions of TMF adsorption and approaches for corrections (C'1 and C'2) in Table 8.1. For both conditions, the  $F$  value was calculated as the tQFF concentration minus bQFF concentration. If there is no TMF adsorption, then the  $A$  value will be the PXP concentration. If the TMF adsorption is similar as that of QFF, then the  $A$  value will be the sum of PXP and QFF concentrations. In Table 8.1, PXP represents the concentration of SVOCs observed in PUF/XAD/PUF samples; tQFF and bQFF represent concentrations of SVOCs associated with top QFF and backup QFF. Particulate OC concentration was adjusted in

the same manner as those SVOCs, and the corresponding  $M_{OM}$  was estimated by multiplying the OC concentration by a scaling factor of 1.53 (Dutton et al., 2009a; Xie et al., 2013a).

**Table 8.1 Sampling artifact corrections for all selected SVOCs.**

Species	Sources of artifacts	Particle phase	Gas phase	Corrections
	gas-phase adsorption to bQFF (positive artifact)	$F = tQFF - bQFF$ , or 0	$A = PXP + (2 \times bQFF)$ or $tQFF + bQFF$	<b>C1:</b> correct the sorption to the top QFF by subtracting bQFF from tQFF; if $tQFF < bQFF$ , particle phase concentrations were assumed to be 0, and the gas phase concentration were the sum of tQFF, bQFF and PXP.
<i>n</i> -Alkanes				
PAHs	evaporation of SVOCs from tQFF (negative artifact)	$F = tQFF + bQFF$	$A = PXP$	<b>C2:</b> correct the evaporation from the top QFF by adding bQFF to tQFF;
Oxy-PAHs	positive artifact offsets negative artifact	$F = tQFF$	$A = PXP + bQFF$	<b>C3:</b> no correction was made for the tQFF; bQFF was added to PXP (similar as traditional sampling – no backup filter was used).
	no TMF adsorption	$F = tQFF - bQFF$	$A = PXP$	<b>C'1:</b> correct the sorption to the top QFF by subtracting bQFF from tQFF;
2-Methyltetrols				
Levoglucosan	with TMF absorption similar as QFF	$F = tQFF - bQFF$	$A = PXP + bQFF$	<b>C'2:</b> besides the correction of the tQFF, PXP was corrected by adding bQFF

### 8.2.3 Calculation of $K_{p,OM}$

Here observationally-based  $K_{p,OM}$  ( $K_{p,OM}^m$ ) was calculated according to Eq. (3), where  $F$  and  $A$  values were corrected gas- and particle-phase concentrations of each SVOC;  $M_{OM}$  was corrected particle-phase OM concentration (as described in the above section). Statistics of  $K_{p,OM}^m$  values calculated using Eq. (3) with different approaches for artifact corrections were listed in Table 8.S3. The theoretically-based  $K_{p,OM}$  ( $K_{p,OM}^t$ ) was calculated according to Eq. (4), where  $T$  was the measured daily average temperature;  $\overline{MW}_{OM}$  of  $200 \text{ g mol}^{-1}$  was assumed for all samples (Barsanti and Pankow, 2004; Williams et al., 2010);  $\zeta_{OM}$  was assumed to be unity for all

species in each sample;  $p_L^\circ$  values for *n*-alkanes, PAHs and oxy-PAHs were estimated using the group contribution methods (GCMs) SIMPOL (Pankow and Asher, 2008); while those for 2-methyltetrols and levoglucosan were obtained from Couvidat and Seigneur (2011) and Booth et al. (2011). The  $p_L^\circ$  value for each species on each sampling day was adjusted by daily average temperature:

$$p_L^\circ = p_L^{o,*} \exp\left[\frac{\Delta H_{vap}^*}{R} \left(\frac{1}{298.15} - \frac{1}{T}\right)\right] \quad (5)$$

where  $p_L^{o,*}$  is the vapor pressure of each pure compound at 298.15 K;  $\Delta H_{vap}^*$  is the enthalpy of vaporization of the liquid (kJ mol<sup>-1</sup>) at 298.15 K. The  $p_L^{o,*}$ ,  $\Delta H_{vap}^*$  and median  $K_{p,OM}^t$  values are given in Table 8.S1.

## 8.3 RESULTS AND DISCUSSION

### 8.3.1 Regression of $\log K_{p,OM}^m$ vs. $\log p_L^\circ$

The statistics for regressions of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  (Eq. 2) with artifact corrections C1, C2 and C3 (Table 8.1) are given in Table 8.2 for *n*-alkanes, PAHs and oxy-PAHs. The method for correlation coefficient (*r*) comparison is provided in supporting information. If the artifact correction was conducted as C1,  $\log K_{p,OM}^m$  and  $\log p_L^\circ$  values were significantly ( $p < 0.05$ ) correlated ( $r = -0.47 - -0.80$ ) for all the six species, and the regression slopes (*m*) ranged from -0.69 to -1.28. If the artifact was corrected as C2, no significant correlation ( $p > 0.05$ ) of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  could be observed for dococane and tricosane, suggesting that the concentrations of these two species observed on the bQFF should not be dominated by negative artifact. While a significant increase ( $p < 0.05$ ) in correlation coefficient was obtained for fluorenone, and the fluorenone was much more volatile than those *n*-alkanes and PAHs (Table 8.S1). This might indicate that the negative sampling artifact could not be ignored for more volatile SVOCs. In this work, the dominance of negative artifact on the bQFF might not be a real condition. The SVOCs with high volatility are prone to exist in gas phase, and if they could not be adsorbed on QFF, then the negative artifact will not be observed. Thus, the bQFF associated SVOCs could not be totally attributed to negative sampling artifact. If the artifact correction was done as C3, the correlations between  $\log K_{p,OM}^m$  and  $\log p_L^\circ$  were improved for all six species compared to those with artifact correction C1. This might support that the negative sampling artifact should be considered for light SVOCs.

Table 8.2 Regression statistics for  $\log K_{p,OM}^m$  vs.  $\log p_L$  with different sampling artifact corrections.

Species	C1 (positive artifact)			C2 (negative artifact)			C3 (positive offsets negative)		
	No. of obs.	m	r	No. of obs.	m	r	No. of obs.	m	r
dodecane	38	-0.73 ± 0.41	-0.51	27	-0.44 ± 0.58	-0.30*	43	-0.81 ± 0.32	-0.62
tricosane	36	-0.73 ± 0.45	-0.49	26	-0.21 ± 0.38	-0.22*	36	-0.64 ± 0.37	-0.51
fluoranthene	43	-0.79 ± 0.31	-0.63	47	-0.81 ± 0.30	-0.62	48	-0.80 ± 0.27	-0.66
pyrene	43	-0.69 ± 0.35	-0.52	44	-0.71 ± 0.32	-0.57	45	-0.74 ± 0.34	-0.55
acenaphthenone	24	-1.28 ± 0.43	-0.80	34	-1.38 ± 0.43	-0.80	44	-1.32 ± 0.30	-0.81
fluorenone	23	-1.20 ± 1.02	-0.47	44	-1.59 ± 0.36	-0.81	48	-1.49 ± 0.37	-0.77

\* Not significant correlation ( $p > 0.05$ )

In Table 8.2, the slope  $m$  for  $n$ -alkanes and PAHs with significant correlations lie in the range of -1.0 – -0.6, which is consistent with other studies (Hart and Pankow, 1994; Cotham and Bidleman, 1995; Wang et al., 2013); while the  $m$  values for the two oxy-PAHs are significantly ( $p < 0.05$ ) lower than -1 for artifact corrections C2 and C3. As discussed by Pankow and Bidleman (1992), the variability in  $m$  values could be attributed to several thermodynamic and non-thermodynamic sources of variability. Then the deviations of  $m$  values from -1 could indicate deviations from true equilibrium partitioning. However, Goss and Schwarzenbach (1998) found that the slope  $m$  could deviate significantly from -1 for true thermodynamic equilibrium adsorption and absorption. In this work, the  $K_{p,OM}^m$  values were calculated based on daily average concentrations of gas- and particle-phase SVOCs, which did not represent true atmospheric concentrations. The regressions of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  for  $n$ -alkanes and PAHs had less strong correlations ( $r < 0.70$ ) and slope  $m$  with wide 95% confidence interval (CI) (Table 8.2), which might suggest deviations from equilibrium partitioning. The regressions for the two oxy-PAHs with artifact corrections C2 and C3 exhibited strong correlations ( $r > 0.70$ ) (Table 8.2), and the slope  $m$  steeper than -1 ( $p < 0.05$ ) were not commonly observed in previous studies for  $n$ -alkanes and PAHs. One explanation for the steeper slope  $m$  is that we assumed constant relative contributions of positive and negative sampling artifacts during artifact corrections (C1, C2 and C3 in Table 8.1). The two oxy-PAHs are much more volatile than other species in Table 8.1. It is possible that we underestimate the relative contribution of negative sampling artifacts in summer when the temperature is high, and the bQFF associated SVOCs cannot cover all evaporation from particles on tQFF. The method to test that the slope is significantly different from unity is given in the supporting information. In this work, we compared the absolute  $m$  values with unity.

For 2-methyltetrols and levoglucosan, the data collected on January 13 and June 20, 2013 were excluded for analysis, since huge spikes in concentration were observed for levoglucosan and other biomass burning tracers (e.g., retene, methoxyphenols). Besides photochemical reactions, biomass burning might also generate 2-methyltetrols (Schkolnik et al., 2005). The fresh emitted 2-methyltetrols and levoglucosan could be collected before achieving equilibrium G/P partitioning. The regression of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  for 2-methyltetrols had correlations of -0.46 ( $p > 0.05$ ;  $m = -1.12 \pm 1.42$ ) and -0.62 ( $p < 0.05$ ;  $m = -1.50 \pm 1.25$ ) for artifact corrections C'1 and C'2 (Table 8.1), respectively. While a much stronger correlation ( $r = 0.86$ ) has been observed if we regress  $F/M_{OM}$  vs.  $A$  (Eq. 3). In Figure 8.1a, the  $K_{p,OM}^m$  value is more likely a constant, as reflected by the regression slope. The values are 0.19 (95% CI, 0.13 – 0.25) and 0.16 (95% CI, 0.11 – 0.22) for artifact corrections C'1 and C'2. The data for Figure 8.1a were obtained on summer sampling days ( $N = 13$ ) with ambient temperatures mostly higher than 20 °C (average  $21.8 \pm 4.05$  °C) and relative humidity (RH) mostly lower than 50% ( $46.0 \pm 14.7\%$ ).

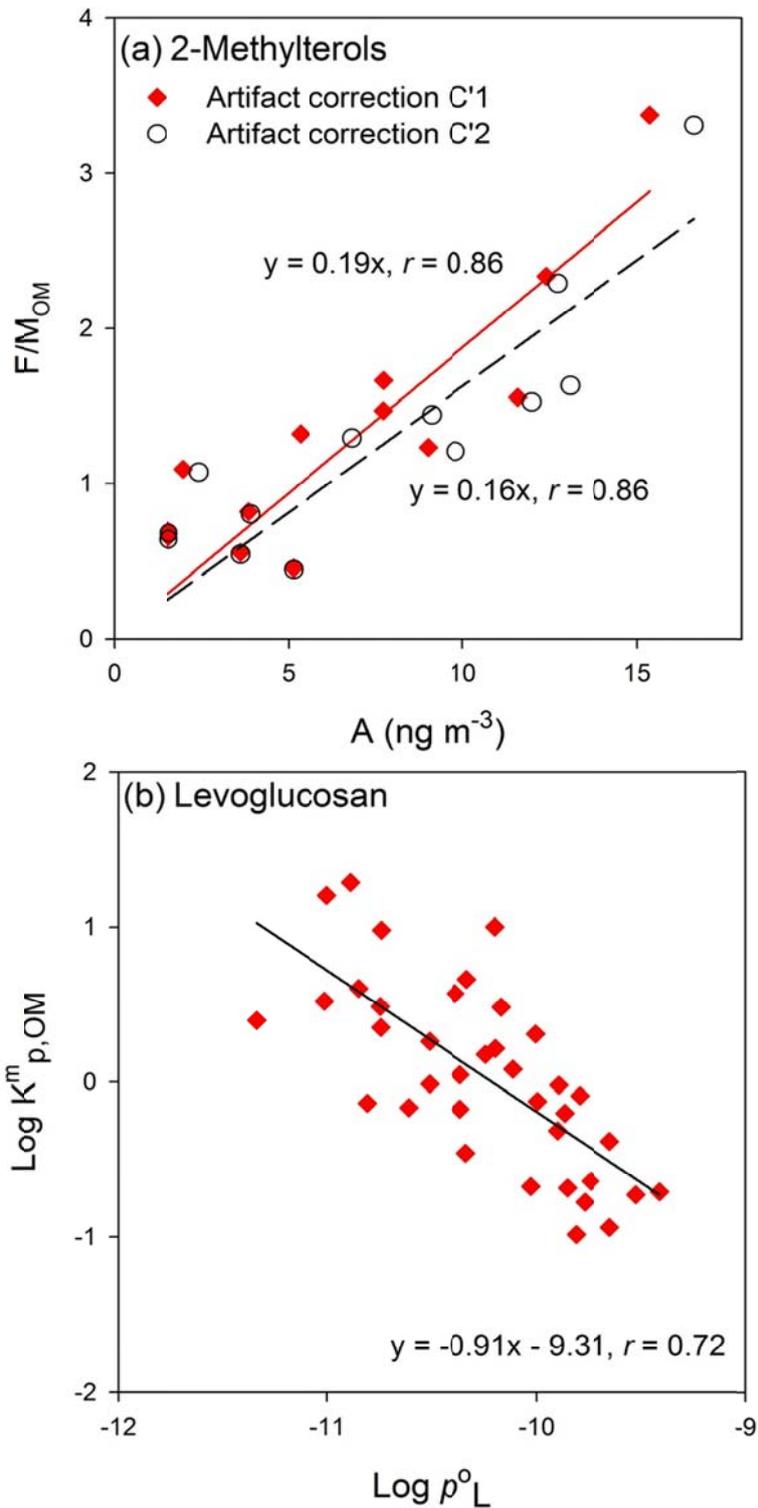


Figure 8.1 Regressions of (a)  $F/M_{OM}$  vs.  $A$  for 2-methyltetrols and (b)  $\log K^m_{p,OM}$  vs.  $\log p^o_L$  for levoglucosan.

2-Methyltetrols in the aerosol were generated after the reactive uptake of epoxydiols of isoprene (IEPOX) through the addition of H<sub>2</sub>O; while the hydroxyl sulfate or nitrate ester could also be generated through the addition of inorganic sulfate or nitrate (Surratt et al., 2010). During the analysis of 2-methyltetrols in filter samples, the derivatization step using *N*, *O*-bis(trimethylsilyl) trifluoroacetamide (BSTFA) containing 1% trimethylchlorosilane (TMCS) might convert those hydroxyl sulfate and nitrate esters back to 2-methyltetrols and form trimethylsilyl esters. As such, the filter analysis might overestimate the particle-phase 2-methyltetrols in the ambient air. The hydroxyl sulfate or nitrate ester forms of 2-methyltetrols are less volatile and more prone to exist in particle phase than 2-methyltetrols. If the mole concentrations of 2-methyltetrols in aerosols are dominated by the forms of hydroxyl sulfate or nitrate esters, then the linear relationship between  $F/M_{OM}$  and  $A$  in Figure 8.1a might not be observed. Laboratory data showed that branching ratio of IEPOX-derived organosulfates to 2-methyltetrols ranged from 5% to 40% for 0.1 to 3 M sulfate solutions (Eddingsaas et al., 2010). In addition, no significant correlation ( $p > 0.05$ ) has been observed between  $\log K_{p,OM}^m$  and RH. In this work, the linear relationship between  $F/M_{OM}$  and  $A$  suggested that the G/P partitioning of 2-methyltetrols was more likely driven by particulate OM phase rather than ambient temperature in summer Denver, and the mole concentrations of 2-methyltetrols should dominate their hydroxyl sulfate and nitrate esters.

$\log K_{p,OM}^m$  and  $\log p_L^\circ$  values of levoglucosna are regressed in Figure 8.1b using the data only corrected as C'1, because the levoglucosan was only observed on three bQFF samples with low concentrations (0.09 – 5.22 ng m<sup>-3</sup>). The regression of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  using the data corrected as C'2 ( $m = -0.91 \pm 0.30$ ,  $b = -9.31 \pm 3.06$ ,  $r = 0.72$ ) was very similar as that in Figure 8.1b. In Figure 8.1b, the regression has an  $m$  value close to -1 (-0.91, 95% CI -1.21 – -0.61) with

strong correlation ( $r = 0.72$ ), indicating that the observed G/P partitioning for levoglucosan is closer to equilibrium than those species in Table 8.2. Previous studies have observed the heterogeneous oxidation of levoglucosan with hydroxyl radicals ( $\cdot\text{OH}$ ) in chambers (Hennigan et al., 2010; Kessler et al., 2010). If the time scale for heterogeneous oxidation of levoglucosan is comparable or shorter than that for G/P partitioning equilibration, then the linear relationship between  $\log K_{\text{p,OM}}^{\text{m}}$  and  $\log p_{\text{L}}^{\circ}$  for levoglucosan might not be obtained. Thus, the results shown in Figure 1b indicated that the time scale for gas-particle partitioning equilibration of levoglucosan should be shorter than that for heterogeneous oxidation.

### 8.3.2 Log $K_{\text{p,OM}}^{\text{m}}$ vs. $\log K_{\text{p,OM}}^{\text{t}}$

In Figure 8.2, log values of observationally-based  $K_{\text{p,OM}}$  calculated using Eq. (3) ( $\log K_{\text{p,OM}}^{\text{m}}$ ) with artifact corrections C1 and C3 (Table 8.1) are regressed to those of theoretically-based ones calculated using Eq. (4) ( $\log K_{\text{p,OM}}^{\text{t}}$ ) for *n*-alkanes, PAHs and oxy-PAHs. From Eq. (4), we can infer that the  $\log K_{\text{p,OM}}^{\text{t}}$  values are negatively correlated with  $\log p_{\text{L}}^{\circ}$ . So the regression of  $\log K_{\text{p,OM}}^{\text{m}}$  vs.  $\log K_{\text{p,OM}}^{\text{t}}$  (Figure 8.2) had very similar correlation coefficients as that of  $\log K_{\text{p,OM}}^{\text{m}}$  vs.  $\log p_{\text{L}}^{\circ}$  (Table 8.2); and the absolute slope values were also very similar between the two regressions. In Figure 8.2, the regression slopes for the six species are generally in the range of 0.5 – 1.5. Then the intercept of the regression could be more indicative in reflecting the differences between  $K_{\text{p,OM}}^{\text{m}}$  and  $K_{\text{p,OM}}^{\text{t}}$ . In Figure 8.2, the regressions of  $\log K_{\text{p,OM}}^{\text{m}}$  vs.  $\log K_{\text{p,OM}}^{\text{t}}$  for the two oxy-PAHs have the largest intercepts (2.95 – 4.12). The median ratios of  $K_{\text{p,OM}}^{\text{m}}$  to  $K_{\text{p,OM}}^{\text{t}}$  are 309 and 162 for acenaphthenone and fluorenone with artifact correction C1, and 499 and 247 with artifact correction C3. Healy et al. (2008) observed similar deviations from theoretical partitioning coefficients for photo-oxidation products of isoprene and 1,3,5-trimethylbenzene with more than one carbonyl, and attributed the deviation to the reactive uptake

by organic aerosols. In this work, reactive uptake might not be the reason. Because acenaphthenone and fluorenone are mono-functional carbonyls and do not possess aldehyde functionality. The reactive uptake of such SVOC was expected to be negligible (Healy et al., 2008). Moreover, even if the heterogeneous reaction plays a role in the reactive uptake of acenaphthenone and fluorenone, the newly generated forms of the two species (e.g., acetal, dimer) might not convert back to their original forms during the analysis, since we extract all sampling media only using methylene chloride without derivatization. The linear relationship between  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$  (Figure 8.2e, f) or  $\log p_L^\circ$  (Table 8.2) also suggested temperature driven G/P partitioning for the two oxy-PAHs. Unlike typical alkyl groups which interact with a given surface by van der Waals forces only, groups containing O or N atoms or aromatic rings can also interact by Lewis acid-base interactions (Goss and Schwarzenbach, 1998). As such, besides the variability in  $\zeta_{OM}$ ,  $\overline{MW}_{OM}$  and possible overestimation on  $p_L^\circ$  (Eq. 4), the missed mechanism of surface interaction different from simple physical adsorption might also contribute to the large difference between  $K_{p,OM}^m$  and  $K_{p,OM}^t$  for acenaphthenone and fluorenone.

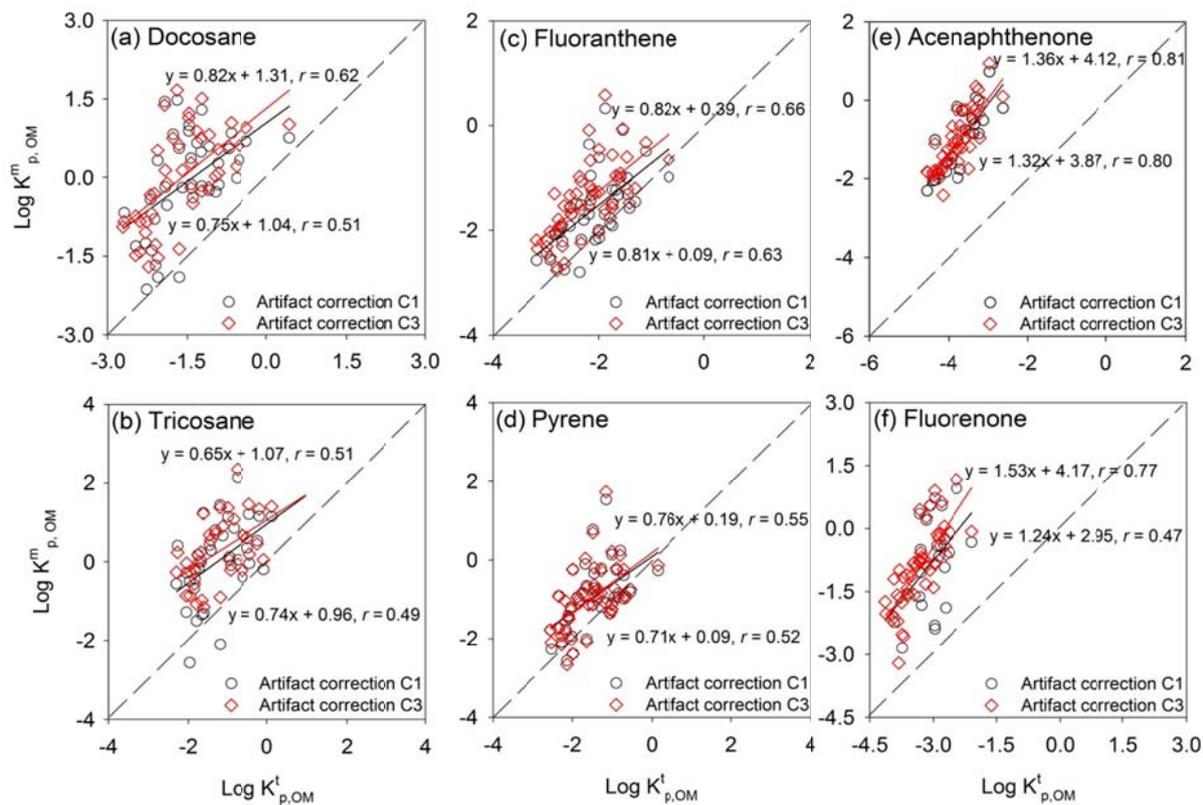


Figure 8.2 Regressions of  $\log K^m_{p,OM}$  vs.  $\log K^t_{p,OM}$  with artifact corrections C1 and C3 for (a) docosane, (b) tricosane, (c) fluoranthene, (d) pyrene, (e) acenaphthenone, (f) fluorenone.

Among the six species in Figure 8.2, the two PAHs show the best agreement between  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$  with the smallest intercepts (0.09 – 0.39). The median ratios of  $K_{p,OM}^m$  and  $K_{p,OM}^t$  are 3.12 and 3.22 for fluoranthene and pyrene with artifact correction C1, and 6.07 and 3.21 with artifact correction C3. Such differences could be accounted for by the small variability in values of  $\zeta_{OM}$  and  $\overline{MW}_{OM}$ , and/or the prediction of  $p_L^\circ$ . The median  $K_{p,OM}^m$  values of the two *n*-alkanes were a bit more than one order of magnitudes higher than those of  $K_{p,OM}^t$ , which might be more related with the overestimation of  $p_L^\circ$  than the small variability of  $\zeta_{OM}$  and  $\overline{MW}_{OM}$ . In Figure 8.2, the  $K_{p,OM}^m$  values calculated with artifact correction C1 are smaller than those with C3, and closer to the values of  $K_{p,OM}^t$ . The median ratios of  $K_{p,OM}^m$  values calculated with artifact correction C1 to those with C3 for the six species ranged from 0.99 (pyrene) to 2.02 (fluorenone), which could not explain the big gap between  $K_{p,OM}^m$  and  $K_{p,OM}^t$ . However, the values of  $K_{p,OM}^m$  and  $K_{p,OM}^t$  are significantly correlated ( $p < 0.05$ ) for all the six species in Figure 8.2, suggesting that the calculation of gas-phase concentrations of these species from their particle phases using  $K_{p,OM}^t$  can reflect reasonably correct time series.

Regressions of  $\log K_{p,OM}^m$  vs.  $\log K_{p,OM}^t$  for 2-methyltetrols with artifact corrections C'1 and C'2 are shown in Figure 8.3a. The correlations ( $r = 0.46$  and  $0.62$ ) are weaker than those between  $F/M_{OM}$  and  $A$  ( $r = 0.86$ ), and the slopes have relatively wider 95% CI. However, the ratios of  $K_{p,OM}^m$  and  $K_{p,OM}^t$  ranged from 0.81 to 5.30 (median 2.23) and 0.80 to 4.26 (2.05) for artifact corrections C'1 and C'2, respectively, much more converged than those of the six species in Figure 8.2. As shown in the previous section, the  $K_{p,OM}^m$  value of 2-methyltetrols is more likely a constant and less sensitive to small temperature changes in summer Denver. If we assume that TMF does not adsorb 2-methyltetrols, the value of 0.19 (95% CI, 0.13 – 0.25) could be used instead of  $K_{p,OM}^t$  to calculate gas-phase 2-methyltetrols in summer Denver.

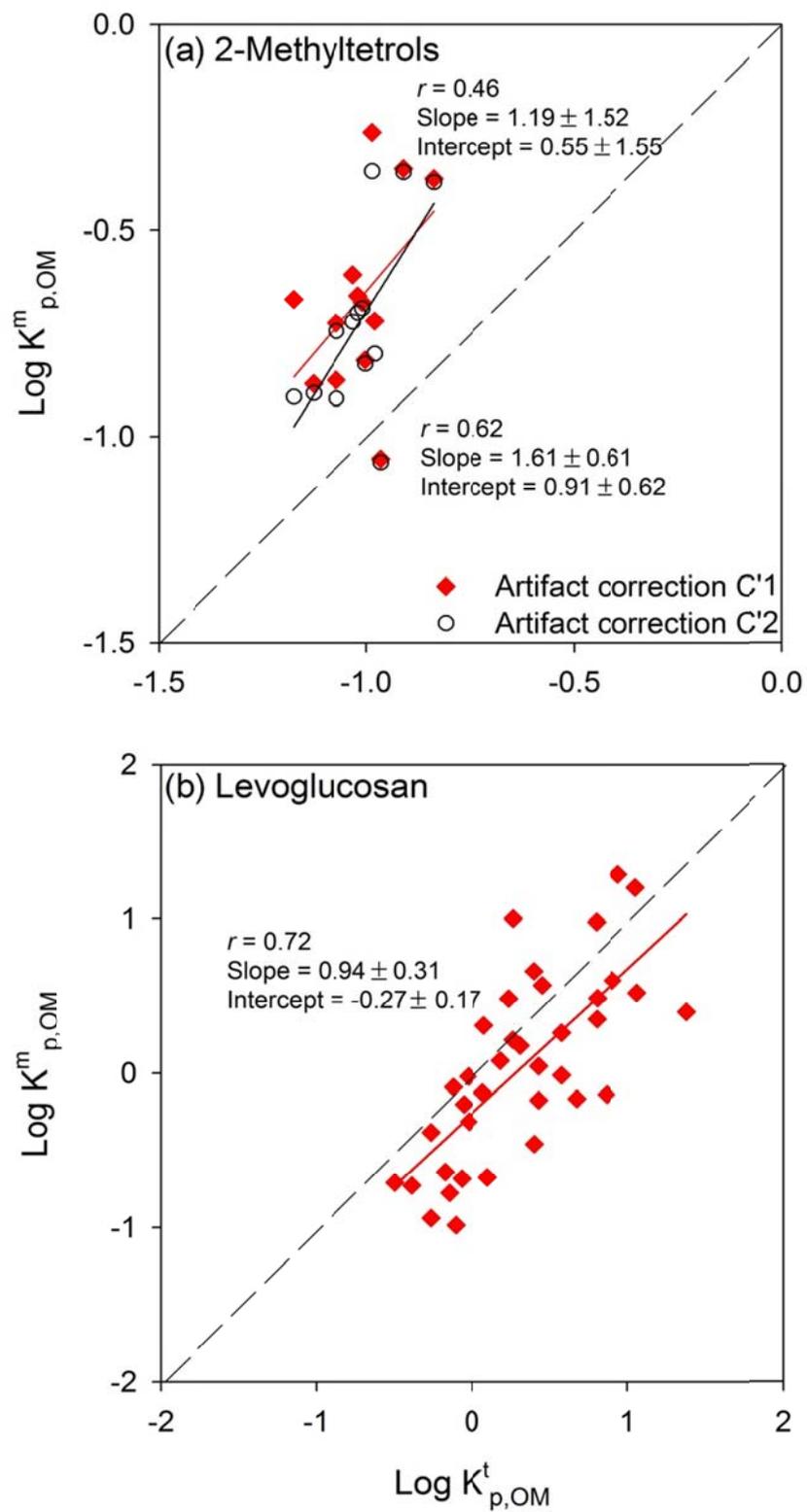


Figure 8.3 Regressions of  $\log K^m_{p,OM}$  vs.  $\log K^t_{p,OM}$  for (a) 2-methyltetrols and (b) levoglucosan.

Unlike 2-methyltetrols, the regression of  $\log K_{p,OM}^m$  vs.  $\log K_{p,OM}^t$  for levoglucosan has stronger correlation ( $r = 0.72$ ) and more stable slope ( $0.94 \pm 0.31$ ) and intercept ( $-0.27 \pm 0.17$ ; Figure 8.3b). The ratios of  $K_{p,OM}^m$  and  $K_{p,OM}^t$  ranged from 0.10 to 5.40 with a median value of 0.50. Besides the variability of  $\zeta_{OM}$ ,  $\overline{MW}_{OM}$  and  $p_L^\circ$ , the uncertainties in measurements might also contribute to the differences between  $K_{p,OM}^m$  and  $K_{p,OM}^t$ . The regression plot could be less scattered if we can improve the recoveries of analysis for levoglucosan in the PUF/XAD/PUF sandwich (51.9 – 63.3%; Chapter 7). These results suggested that the absorptive partitioning theory can reasonably reproduce the G/P partitioning of levoglucosan.

## 8.4 CONCLUSIONS

In this work, the G/P partitioning of selected SVOCs (two *n*-alkanes, two PAHs, two oxy-PAHs, 2-methyltetrols and levoglucosan) were investigated. The gas- and particle-phase concentrations of those SVOCs were derived from field measurement (Chapter 7) with different artifact corrections, and used to calculate observationally-based  $K_{p,OM}$  ( $K_{p,OM}^m$ ). The theoretically-based  $K_{p,OM}$  ( $K_{p,OM}^t$ ) were calculated using predicted values of  $\zeta_{OM}$ ,  $\overline{MW}_{OM}$  and  $p_L^\circ$  for each species. The regressions of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  for *n*-alkanes, PAHs, oxy-PAHs and levoglucosan had significant correlations ( $p < 0.05$ ). The deviation of regression slopes from -1 and the scattered regression indicated G/P partitioning deviating from true equilibrium. The  $K_{p,OM}^m$  value for 2-methyltetrols is more likely a constant in summer Denver, suggesting that the G/P partitioning of 2-methyltetrols is more sensitive to the variation in particulate OM phase than small changes in ambient temperature. The comparisons between  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$  suggested that the calculation of gas-phase concentrations of light SVOCs using  $K_{p,OM}^t$  could reflect a reasonably correct time series. The two PAHs and levoglucosan had the best agreement between  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$ . The median  $K_{p,OM}^t$  values of the two oxy-PAHs were 2-3 orders of magnitudes lower than the corresponding  $K_{p,OM}^m$ , which might not only be attributed to the variability of  $\zeta_{OM}$ ,  $\overline{MW}_{OM}$  and  $p_L^\circ$ , but also the missed mechanism of strong surface interaction other than simple physical adsorption.

## CHAPTER 9 SUMMARY AND FUTURE RESEARCH

### 9.1 STUDY SUMMARY

One objective of the Denver Aerosol Sources and Health (DASH) study is to collect a continuous time series of speciated  $PM_{2.5}$  measurements at a receptor site in Denver for source and health effect analysis (Vedal et al., 2009). To understand the limitations of the single site used in the DASH study and examine the spatial variability of source contributions to  $PM_{2.5}$ , one year of supplemental  $PM_{2.5}$  samples were collected every sixth day at four sites in Denver after the DASH sampling campaign. The carbonaceous compositions (elemental carbon (EC), organic carbon (OC) and an array of organic molecular markers (OMMs)) of all  $PM_{2.5}$  samples were characterized and used as inputs for source apportionment. In this study, positive matrix factorization (PMF2), coupled with a bootstrap technique for uncertainty assessment developed by Hemann et al. (2009), was used as the primary tool for source apportionment of  $PM_{2.5}$ . Spatial variability of  $PM_{2.5}$  components and source contributions were evaluated by using correlation coefficients ( $r$ ) and coefficients of divergence (COD).

In the DASH study, collected  $PM_{2.5}$  samples were not only speciated for carbonaceous species, but also inorganic species (sulfate, nitrate and an array of water soluble elements (WSEs)). Most previous source apportionment studies used compositional data of elements (Kim et al., 2005; Hwang et al., 2008; Mooibroek et al., 2011) or OMMs (Jaekels et al., 2007; Shrivastava et al., 2007) to apportion bulk  $PM_{2.5}$  to pollution sources. To evaluate the utility of different speciation data sets, the PMF2 model was applied to four different 1-year (2003) data sets composed of (1) bulk species, (2) bulk species and WSE, (3) bulk species and OMM, and (4) combination of all species.

In the current study, a 32-month (January 2003 – October 2005) series of daily speciated  $PM_{2.5}$  data set, including the 1-year (2003) sub-data set analyzed previously (Dutton et al., 2010b), was used as inputs for source apportionment to obtain more reliable  $PM_{2.5}$  source information. Moreover, temperature-stratified PMF analysis is conducted for the long time series speciation data set, which could help to identify the influence of atmospheric processes (e.g., gas/particle (G/P) partitioning, photochemical reactions) on source apportionment. The source profiles were typically assumed to be constant over the period of ambient and source sampling (Chen et al., 2011). However, OMMs are mostly semi-volatile organic compounds (SVOCs), and their mass fractions in particle phase are influenced by atmospheric processes. As such, the output factors of a receptor model are not necessarily pollution sources, and could reflect the influence of atmospheric processes. To eliminate the influence of G/P partitioning on source apportionment, gas-phase concentrations of SVOCs were predicted using their 32-month series of particle-phase concentrations based on an equilibrium absorption model (Pankow, 1994a, b), and added to the particle phase for PMF analysis.

In order to verify the prediction of gas-phase SVOCs by the equilibrium absorption model, fifty gas- and particle-phase ( $PM_{2.5}$ ) SVOCs samples were collected from August 2012 to July 2013. A medium volume sampler incorporating quartz fiber filter (QFF) and polyurethane foam (PUF)/XAD/PUF sandwich was used for sample collection throughout the sampling period. A backup QFF (bQFF) was installed to evaluate sampling artifacts of particulate organics. Speciated SVOCs include *n*-alkanes, PAHs, oxy-PAHs, 2-methyltetrols and levoglucosan. Gas- and particle-phase concentrations of selected SVOCs (two *n*-alkanes, two PAHs, two oxy-PAHs, 2-methyltetrols and levoglucosan) were calculated with different artifact corrections based on

bQFF measurements. Finally, the observed G/P partitioning of those SVOCs were compared to those predicted by equilibrium partitioning model.

## 9.2 RESULTS SUMMARY

In Chapter 2, OMMs and bulk carbon contents (EC and OC) were analyzed in PM<sub>2.5</sub> samples collected every sixth day at four sampling sites for one year. Our results show that PAHs, steranes, methoxyphenols and EC concentrations were more affected by primary emissions and show distinct differences in concentrations between near-highway and residential sites. PAHs and steranes exhibited some degree of homogeneity for site pairs where each site was from a residential or a near-highway area, but showed less homogeneity for site pairs where one site was residential and one site was near-highway. OC was more strongly correlated and exhibited relatively lower CODs than EC. The spatial heterogeneity of OMMs would be overestimated without considering their divergence between co-located samples. Values of  $r$  and CODs derived from co-located samples should be used as points of reference to analyze spatial variability of PM<sub>2.5</sub> species. The carbonaceous speciation data collected at the four sites were pooled as inputs for source apportionment in Chapter 3. A seven-factor solution was identified and associated with high plant wax, summertime emission, diesel vehicle emission, fossil fuel combustion, motor vehicle emission, lubricating oil combustion and wood burning. The summertime emission factor exhibited the highest correlation ( $r = 0.74 - 0.88$ ) and lowest CODs (CODs = 0.32 - 0.38) between each site pair among all resolved factors, while higher spatial variability was observed for those traffic related factors. Source contributions to average total EC and OC mass were similarly distributed across the four sites. Moreover, consistent PMF source profiles were obtained when we performed source apportionment for site specific data sets. These results suggest that characterizing organic PM<sub>2.5</sub> concentrations at a single well-chosen site in Denver is adequate to reasonably assess relative source contributions for the urban area.

In Chapter 4, four different 1-year PM<sub>2.5</sub> speciation data sets were analyzed using the PMF model. The Bulk, WSE+Bulk and OMM+Bulk data sets correspond to the use of different chemical analysis tools. The bootstrap approach applied aids in interpretation of PMF solutions as it provides estimates of variability in factor profiles and contributions. Moreover, this approach also provides an alternative criterion – matching rate of bootstrapped factors to base case factors – for the selection of number of factors. A high factor matching rate reflects the uniqueness of base case factors, robustness of this solution to the input data set and corresponds to distinguishable PMF results. The five factors resolved by using only the bulk species best reproduced the observed concentrations of PM<sub>2.5</sub> components. Combining WSE with bulk species as PMF inputs also produced five factors. Three of them were linked to soil, road dust and processed dust, and together contributed 26.0% of reconstructed PM<sub>2.5</sub> mass. A 7-factor PMF solution was identified using speciated OMM and bulk species. The EC/sterane and summertime/selective aliphatic factors had the highest contributions to EC (39.0%) and OC (53.8%) respectively. The nine factors resolved by including all species as input data are consistent with those from the previous two solutions (WSE and bulk species, OMM and bulk species) in both factor profiles and contributions ( $r = 0.88 - 1.00$ ). The comparisons across different solutions indicate that the selection of input data set may depend on the PM components or sources of interest for specific source-oriented health study.

As shown in Chapter 4, OMMs are better tracers than WSEs in apportion bulk organics (EC and OC) of PM<sub>2.5</sub> into pollution sources. However, the factor from OMM-based source apportionment could be influenced by atmospheric processes. In Chapter 5, a long time series (32 months) of daily PM<sub>2.5</sub> speciation data was used 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. 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 needs to be considered in epidemiological studies of the health effects of chronic exposure to source contributions.

In order to eliminate the impacts of G/P partitioning on PMF analysis, the gas-phase concentrations of 71 SVOCs were predicted using particle-phase measurements (Chapter 5) by equilibrium G/P partitioning theory. In Chapter 6, the gas-phase concentrations of all SVOCs were added to their particle-phase concentrations as inputs for source apportionment. Seven factors were identified from the full data set, including the nitrate, sulfate, *n*-alkane, sterane, light SVOC, PAH and bulk carbon factors, and could be matched to those from a particle only-based PMF solution (Chapter 5) with reasonable ( $r = 0.69$ ) to excellent ( $r = 0.98$ ) correlations. Three temperature-stratified sub-data sets, representing ambient sampling during the cold, warm and hot periods, were also analyzed using PMF. Unlike the light *n*-alkane/PAH factor from the particle only-based PMF analysis, the light SVOC factor from the total-SVOC based PMF solution exhibited strong correlations ( $r = 0.82 - 0.98$ ) between the full data set and each sub-

data set solutions. These results suggested that the influences of G/P partitioning on PMF analysis could be removed by using total SVOC (gas + particle phase) data. However, the impact of photochemical process has not been ruled out in this work, as illustrated by the moderate correlation ( $r = 0.54$ ) between the bulk carbon factor of the full data set solution and that of the cold period solution.

In Chapter 6, several pre-assumptions (e.g., equilibrium absorptive partitioning) were made for the calculation of gas-phase SVOCs, and need to be verified, and if necessary refined, by comparing with field measurements. In Chapter 7, a medium volume sampler incorporating QFF and PUF/XAD/PUF sandwich, was used to collect SVOCs in both gaseous and particle ( $PM_{2.5}$ ) phases. A bQFF was used to evaluate possible sampling artifact of particulate organics. The breakthrough experiments showed that the PUF/XAD/PUF sandwich could collect gas-phase *n*-alkanes, PAHs, 2-methyltetrols and levoglucosan with low breakthrough, even for low molecular weight (MW) species (e.g., naphthalene, 3.0%); however, the recoveries of levoglucosan in PUF/XAD/PUF sandwiches were lower than 70% (average, 51.9 – 63.3%). Comparing species concentrations across different sampling matrices, the light *n*-alkanes ( $MW < 282$ ) and PAHs ( $MW < 192$ ) were mostly distributed into the gas phase in urban Denver; while the relative abundance of heavier *n*-alkanes ( $MW \geq 282$ ) and PAHs ( $MW \geq 192$ ) in the particle phase to those in the gas phase increased with MW. The concentration ratios of 2-methyltetrols and levoglucosan in the gas phase to those in the particle phase were often close to or higher than unity in summer, indicating that these polar species could be volatile and their gas/particle partitioning should be considered when applying their particle-phase data for source apportionment.

Chapter 8 compared observationally-based G/P partitioning coefficients ( $K_{p,OM}^m$ ) of selected SVOCs with their predicted vapor pressures ( $p_L^\circ$ ) and theoretically-based partitioning coefficients ( $K_{p,OM}^t$ ). Measured concentrations of those SVOCs (two *n*-alkanes, two PAHs, two oxy-PAHs, 2-methyltetrols and levoglucosan) in different sampling media were obtained in Chapter 7. Gas- and particle-phase concentrations of *n*-alkanes, PAHs and oxy-PAHs were calculated with three approaches for artifact corrections based on the measurement of bQFF, and used to calculate their  $K_{p,OM}^m$ . If the bQFF associated *n*-alkanes, PAHs and oxy-PAHs of interest were dominated by positive artifact (vapor phase adsorption) or equally contributed by positive and negative (particle phase evaporation) artifacts, the correlations between  $\log K_{p,OM}^m$  and  $\log p_L^\circ$  were all significant ( $p < 0.05$ ). However, the regression slopes ( $m$ ) deviating from -1 and their wide 95% confidence interval (CI) might suggest deviations from true equilibrium. For less volatile 2-methyltetrols and levoglucosan, the bQFF concentrations were assumed to be dominated by positive sampling artifacts and subtracted from the top QFF concentrations as the particle-phase; while the gas-phase concentrations were corrected with two different assumptions about the adsorption of Teflon membrane filter (TMF). The constant like  $K_{p,OM}^m$  value for 2-methyltetrols (no TMF adsorption,  $0.19 \pm 0.06$ ; with TMF adsorption,  $0.16 \pm 0.05$ ) indicated that the particulate organic material dominated the G/P partitioning of 2-methyltetrols in summer Denver. The regression of  $\log K_{p,OM}^m$  vs.  $\log p_L^\circ$  ( $m = -0.91 \pm 0.30$ ,  $r = 0.72$ ) for levoglucosan indicated a G/P partitioning close to equilibrium. The comparisons between  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$  suggested that the calculation of gas-phase concentrations of light SVOCs using  $K_{p,OM}^t$  could reflect a reasonably correct time series. The two PAHs and levoglucosan had the best agreement between  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$ . The median  $K_{p,OM}^t$  values of the two oxy-PAHs were 2-3 orders of magnitudes lower than the corresponding  $K_{p,OM}^m$ , which might not only be

attributed to the variability of  $\zeta_{OM}$ ,  $\overline{MW}_{OM}$  and  $p_L^\circ$ , but also the missed mechanism of strong surface interaction other than simple physical adsorption.

### 9.3 FURTHER RESEARCH

In Chapter 2, the spatial variability of PM<sub>2.5</sub> components were framed by co-located sampling. The side-by-side measurements assessed the uncertainty associated with sampling and analytical measurement. Values of *r* and CODs derived from co-located samples should be used as points of reference to analyze the spatial variability of PM<sub>2.5</sub> components. However, the limited sample number (31) and low S/N ratios of certain classes of organic compounds might lead to an overestimation of uncertainty in sampling and analytical measurement. In further studies on spatial variability of PM<sub>2.5</sub> components, co-located samples should be collected simultaneously as those spatial samples, so as to develop more robust reference.

Chapters 3 – 5 have done source apportionment of PM<sub>2.5</sub> based on particle-phase SVOCs. All PMF solutions have a common factor characterized by high loadings of OC, *n*-alkanes with odd carbon number preference. This factor was associated with summertime biogenic emissions, followed by photochemical reactions. However, we cannot rule out the contributions from anthropogenic emissions and their photochemical reaction products. In further study, effective organic tracers of secondary organic aerosols (SOA) derived from biogenic (e.g., 2-methyltetrols) and anthropogenic (2,3-dihydroxy-4-oxopentanoic acid) precursors could be used as inputs for source apportionment, so as to distinguish source contributions of biogenic SOA from anthropogenic SOA.

A medium sampler incorporating two sampling chains was used to collect gas- and particle-phase SVOCs (Chapter 7). The breakthrough experiments showed that the PUF/XAD/PUF sandwich was highly efficient in sampling gas-phase light *n*-alkanes, PAHs, 2-methyltetrols and levoglucosan. However, the recoveries of levoglucosan analysis in PUF/XAD/PUF sandwich did not exceed 70% (51.9 – 63.3%). The low recovery could be the

primary reason for the scattered regression between  $\log K_{p,OM}^m$  and  $\log p_L^\circ$  (Chapter 8). The recovery analysis also suggested a decrease in levoglucosan recovery as the usage of PUF increased. Therefore, further research could try to increase the recovery of levoglucosan by using less PUF for sampling.

Subramanian et al. (2004) has demonstrated that using a bQFF could provide a reasonable estimate of the positive artifact (vapor phase adsorption) of particulate OC for 24-h samples. But for more volatile SVOCs (e.g., dodecane), negative artifact (particulate evaporation) should be considered and could not be estimated by the measurement of bQFF. To obtain less biased G/P partitioning of those light SVOCs, the negative sampling artifact needs to be measured. One possible design could be denuder-filter-denuder, and the first denuder should have ~100% collection efficiency for gas-phase light SVOCs.

In Chapter 7, SVOCs in gaseous and particle phases were both measured. The total SVOCs (gas + particle phase) data could be used as inputs for source apportionment, and compared to the results derived from particle-phase data. Then we can evaluate the influence of G/P partitioning on receptor-based source apportionment more accurately. Predicted gas-phase SVOCs will also be added to their particle-phase as inputs for source apportionment, so as to know how well the absorptive partitioning theory can help us to improve receptor-based source apportionment by predicting gas-phase SVOCs.

In Chapter 8, the G/P partitioning of 2-methyltetrols was more likely driven by the variation of particulate OM phase than temperature changes in summer Denver. This was derived from the measurement of only 13 pairs of gas- and particle-phase samples, since the gas-phase 2-methyltetrols could only be observed in hot periods. Future study can focus on summertime sampling to obtain more observations of those tracers in both gas- and particle-phase. A big gap

(2-3 orders of magnitudes) between  $\log K_{p,OM}^m$  and  $\log K_{p,OM}^t$  for acenaphthenone and fluorenone was observed, and the correlations between their  $\log K_{p,OM}^m$  and  $\log p_L^\circ$  indicated that the ambient temperature was the driven force for G/P partitioning. The variability in predicted species activity coefficient, vapor pressure and average molecular weight of OM could not account for the large deviations. Therefore, it is necessary to identify and quantify the contributions of other mechanisms (besides OM phase absorption) to the deviations between observed and predicted G/P partitioning.

## CHAPTER 10 BIBLIOGRAPHY

- Ahrens, L., Harner, T., Shoeib, M., Lane, D.A., Murphy, J.G., 2012. Improved characterization of gas–particle partitioning for per- and polyfluoroalkyl substances in the atmosphere using annular diffusion denuder samplers. *Environmental Science & Technology* 46, 7199-7206.
- Ahrens, L., Shoeib, M., Harner, T., Lane, D.A., Guo, R., Reiner, E.J., 2011. Comparison of annular diffusion denuder and high volume air samplers for measuring per- and polyfluoroalkyl substances in the atmosphere. *Analytical Chemistry* 83, 9622-9628.
- Aurela, M., Saarikoski, S., Timonen, H., Aalto, P., Keronen, P., Saarnio, K., Teinilä, K., Kulmala, M., Hillamo, R., 2011. Carbonaceous aerosol at a forested and an urban background sites in Southern Finland. *Atmospheric Environment* 45, 1394-1401.
- Bae, M.S., Demerjian, K.L., Schwab, J.J., 2006. Seasonal estimation of organic mass to organic carbon in PM<sub>2.5</sub> at rural and urban locations in New York state. *Atmospheric Environment* 40, 7467-7479.
- Bao, L., Matsumoto, M., Kubota, T., Sekiguchi, K., Wang, Q., Sakamoto, K., 2012. Gas/particle partitioning of low-molecular-weight dicarboxylic acids at a suburban site in Saitama, Japan. *Atmospheric Environment* 47, 546-553.
- Bari, M.A., Baumbach, G., Kuch, B., Scheffknecht, G., 2009. Wood smoke as a source of particle-phase organic compounds in residential areas. *Atmospheric Environment* 43, 4722-4732.
- Barsanti, K.C., Pankow, J.F., 2004. Thermodynamics of the formation of atmospheric organic particulate matter by accretion reactions—Part 1: aldehydes and ketones. *Atmospheric Environment* 38, 4371-4382.

- Bergvall, C., Westerholm, R., 2009. Determination of highly carcinogenic dibenzopyrene isomers in particulate emissions from two diesel- and two gasoline-fuelled light-duty vehicles. *Atmospheric Environment* 43, 3883-3890.
- Booth, A.M., Montague, W.J., Barley, M.H., Topping, D.O., McFiggans, G., Garforth, A., Percival, C.J., 2011. Solid state and sub-cooled liquid vapour pressures of cyclic aliphatic dicarboxylic acids. *Atmos. Chem. Phys.* 11, 655-665.
- Bullock, K.R., Duvall, R.M., Norris, G.A., McDow, S.R., Hays, M.D., 2008. Evaluation of the CMB and PMF models using organic molecular markers in fine particulate matter collected during the Pittsburgh Air Quality Study. *Atmospheric Environment* 42, 6897-6904.
- Cadle, S.H., Mulawa, P.A., Hunsanger, E.C., Nelson, K., Ragazzi, R.A., Barrett, R., Gallagher, G.L., Lawson, D.R., Knapp, K.T., Snow, R., 1999. Composition of light-duty motor vehicle exhaust particulate matter in the Denver, Colorado area. *Environmental Science & Technology* 33, 2328-2339.
- Caricchia, A.M., Chiavarini, S., Pezza, M., 1999. Polycyclic aromatic hydrocarbons in the urban atmospheric particulate matter in the city of Naples (Italy). *Atmospheric Environment* 33, 3731-3738.
- Cass, G.R., 1998. Organic molecular tracers for particulate air pollution sources. *Trac-Trends in Analytical Chemistry* 17, 356-366.
- Chang, E.I., Pankow, J.F., 2010. Organic particulate matter formation at varying relative humidity using surrogate secondary and primary organic compounds with activity corrections in the condensed phase obtained using a method based on the Wilson equation. *Atmospheric Chemistry and Physics* 10, 5475-5490.

- Chen, L.-G., Mai, B.-X., Bi, X.-H., Chen, S.-J., Wang, X.-M., Ran, Y., Luo, X.-J., Sheng, G.-Y., Fu, J.-M., Zeng, E.Y., 2006. Concentration levels, compositional profiles, and gas-particle partitioning of polybrominated diphenyl ethers in the atmosphere of an urban city in south China. *Environmental Science & Technology* 40, 1190-1196.
- Chen, L.W.A., Watson, J.G., Chow, J.C., DuBois, D.W., Herschberger, L., 2011. PM<sub>2.5</sub> Source Apportionment: Reconciling Receptor Models for U.S. Nonurban and Urban Long-Term Networks. *Journal of the Air & Waste Management Association* 61, 1204-1217.
- Cheng, Y., Li, S.-M., Leithead, A., Brickell, P.C., Leaitch, W.R., 2004. Characterizations of cis-pinonic acid and n-fatty acids on fine aerosols in the Lower Fraser Valley during Pacific 2001 Air Quality Study. *Atmospheric Environment* 38, 5789-5800.
- Cheng, Y., Li, S.-M., Leithead, A., Brook, J.R., 2006. Spatial and diurnal distributions of n-alkanes and n-alkan-2-ones on PM<sub>2.5</sub> aerosols in the Lower Fraser Valley, Canada. *Atmospheric Environment* 40, 2706-2720.
- Cheung, K., Daher, N., Kam, W., Shafer, M.M., Ning, Z., Schauer, J.J., Sioutas, C., 2011. Spatial and temporal variation of chemical composition and mass closure of ambient coarse particulate matter (PM<sub>10-2.5</sub>) in the Los Angeles area. *Atmospheric Environment* 45, 2651-2662.
- Christensen, W.F., Schauer, J.J., 2008a. Impact of species uncertainty perturbation on the solution stability of positive matrix factorization of atmospheric particulate matter data. *Environmental Science & Technology* 42, 6015-6021.
- Christensen, W.F., Schauer, J.J., 2008b. Quantifying and manipulating species influence in positive matrix factorization. *Chemometrics and Intelligent Laboratory Systems* 94, 140-148.

- Claeys, M., Graham, B., Vas, G., Wang, W., Vermeylen, R., Pashynska, V., Cafmeyer, J., Guyon, P., Andreae, M.O., Artaxo, P., Maenhaut, W., 2004. Formation of secondary organic aerosols through photooxidation of isoprene. *Science* 303, 1173-1176.
- Claeys, M., Szmigielski, R., Kourtshev, I., Van der Veken, P., Vermeylen, R., Maenhaut, W., Jaoui, M., Kleindienst, T.E., Lewandowski, M., Offenberg, J.H., Edney, E.O., 2007. Hydroxydicarboxylic acids: Markers for secondary organic aerosol from the photooxidation of  $\alpha$ -pinene. *Environmental Science & Technology* 41, 1628-1634.
- Cotham, W.E., Bidleman, T.F., 1995. Polycyclic aromatic hydrocarbons and polychlorinated biphenyls in air at an urban and a rural site near Lake Michigan. *Environmental Science & Technology* 29, 2782-2789.
- Couvidat, F., Seigneur, C., 2011. Modeling secondary organic aerosol formation from isoprene oxidation under dry and humid conditions. *Atmospheric Chemistry and Physics*. 11, 893-909.
- Crouse, J.D., Paulot, F., Kjaergaard, H.G., Wennberg, P.O., 2011. Peroxy radical isomerization in the oxidation of isoprene. *Physical Chemistry Chemical Physics* 13, 13607-13613.
- Demydov, D., Adhvaryu, A., McCluskey, P., Malshe Ajay, P., 2010. Advanced lubricant additives of dialkyldithiophosphate (DDP)-functionalized molybdenum sulfide nanoparticles and their tribological performance for boundary lubrication. *Nanoscale Materials in Chemistry: Environmental Applications*. American Chemical Society, pp. 137-163.
- Devos, O., Combet, E., Tassel, P., Paturel, L., 2005. Exhaust emissions of PAHs of passenger cars. *Polycyclic Aromatic Compounds* 26, 69-78.
- Diaz-Robles, L.A., Fu, J.S., Reed, G.D., 2008. Modeling and source apportionment of diesel particulate matter. *Environment International* 34, 1-11.

- Ding, X., Wang, X.-M., Gao, B., Fu, X.-X., He, Q.-F., Zhao, X.-Y., Yu, J.-Z., Zheng, M., 2012. Tracer-based estimation of secondary organic carbon in the Pearl River Delta, south China. *Journal of Geophysical Research: Atmospheres* 117, D05313.
- Dockery, D.W., Pope, C.A., Xu, X., Spengler, J.D., Ware, J.H., Fay, M.E., Ferris, B.G., Speizer, F.E., 1993. An association between air pollution and mortality in six U.S. cities. *New England Journal of Medicine* 329, 1753-1759.
- Dutton, S.J., Rajagopalan, B., Vedal, S., Hannigan, M.P., 2010a. Temporal patterns in daily measurements of inorganic and organic speciated PM<sub>2.5</sub> in Denver. *Atmospheric Environment* 44, 987-998.
- Dutton, S.J., Schauer, J.J., Vedal, S., Hannigan, M.P., 2009a. PM<sub>2.5</sub> characterization for time series studies: Pointwise uncertainty estimation and bulk speciation methods applied in Denver. *Atmospheric Environment* 43, 1136-1146.
- Dutton, S.J., Vedal, S., Piedrahita, R., Milford, J.B., Miller, S.L., Hannigan, M.P., 2010b. Source apportionment using positive matrix factorization on daily measurements of inorganic and organic speciated PM<sub>2.5</sub>. *Atmospheric Environment* 44, 2731-2741.
- Dutton, S.J., Williams, D.E., Garcia, J.K., Vedal, S., Hannigan, M.P., 2009b. PM<sub>2.5</sub> characterization for time series studies: Organic molecular marker speciation methods and observations from daily measurements in Denver. *Atmospheric Environment* 43, 2018-2030.
- Eddingsaas, N.C., VanderVelde, D.G., Wennberg, P.O., 2010. Kinetics and products of the acid-catalyzed ring-opening of atmospherically relevant butyl epoxy alcohols. *The Journal of Physical Chemistry A* 114, 8106-8113.
- Eiguren-Fernandez, A., Miguel, A.H., Froines, J.R., Thurairatnam, S., Avol, E.L., 2004. Seasonal and spatial variation of polycyclic aromatic hydrocarbons in vapor-phase and

- PM2.5 in southern California urban and rural communities. *Aerosol Science and Technology* 38, 447-455.
- Finlayson-Pitts, B. J., Pitts Jr., J.N., 2000. *Chemistry of the upper and lower atmosphere: theory experiments and applications*. Academic Press, p219.
- Fitzpatrick, E.M., Ross, A.B., Bates, J., Andrews, G., Jones, J.M., Phylaktou, H., Pourkashanian, M., Williams, A., 2007. Emission of oxygenated species from the combustion of pine wood and its relation to soot formation. *Process Safety and Environmental Protection* 85, 430-440.
- Fraser, M.P., Cass, G.R., Simoneit, B.R.T., Rasmussen, R.A., 1997. Air quality model evaluation data for organics. 4. C2–C36 non-aromatic hydrocarbons. *Environmental Science & Technology* 31, 2356-2367.
- Fraser, M.P., Cass, G.R., Simoneit, B.R.T., Rasmussen, R.A., 1998. Air quality model evaluation data for organics. 5. C6–C22 nonpolar and semipolar aromatic compounds. *Environmental Science & Technology* 32, 1760-1770.
- Goss, K.-U., Schwarzenbach, R.P., 1998. Gas/solid and gas/liquid partitioning of organic compounds: Critical evaluation of the interpretation of equilibrium constants. *Environmental Science & Technology* 32, 2025-2032.
- Guo, H., Lee, S.C., Ho, K.F., Wang, X.M., Zou, S.C., 2003. Particle-associated polycyclic aromatic hydrocarbons in urban air of Hong Kong. *Atmospheric Environment* 37, 5307-5317.
- Hallquist, M., Wenger, J.C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N.M., George, C., Goldstein, A.H., Hamilton, J.F., Herrmann, H., Hoffmann, T., Iinuma, Y., Jang, M., Jenkin, M.E., Jimenez, J.L., Kiendler-Scharr, A., Maenhaut, W., McFiggans, G., Mentel, T.F., Monod, A., Prevot, A.S.H., Seinfeld, J.H., Surratt, J.D.,

- Szmigielski, R., Wildt, J., 2009. The formation, properties and impact of secondary organic aerosol: current and emerging issues. *Atmospheric Chemistry and Physics* 9, 5155-5236.
- Harris, A.R., Davidson, C.I., 2005. The role of resuspended soil in lead flows in the California south coast air basin. *Environmental Science & Technology* 39, 7410-7415.
- Hart, K.M., Isabelle, L.M., Pankow, J.F., 1992. High-volume air sampler for particle and gas sampling. 1. Design and gas sampling performance. *Environmental Science & Technology* 26, 1048-1052.
- Hart, K.M., Pankow, J.F., 1994. High-volume air sampler for particle and gas sampling. 2. Use of backup filters to correct for the adsorption of gas-phase polycyclic aromatic hydrocarbons to the front filter. *Environmental Science & Technology* 28, 655-661.
- Healy, R.M., Wenger, J.C., Metzger, A., Duplissy, J., Kalberer, M., Dommen, J., 2008. Gas/particle partitioning of carbonyls in the photooxidation of isoprene and 1,3,5-trimethylbenzene. *Atmospheric Chemistry and Physics* 8, 3215-3230.
- Hemann, J.G., Brinkman, G.L., Dutton, S.J., Hannigan, M.P., Milford, J.B., Miller, S.L., 2009. 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* 9, 497-513.
- Hennigan, C.J., Sullivan, A.P., Collett, J.L., Robinson, A.L., 2010. Levoglucosan stability in biomass burning particles exposed to hydroxyl radicals. *Geophysical Research Letters* 37, L09806.
- Henry, R.C., Christensen, E.R., 2010. Selecting an appropriate multivariate source apportionment model result. *Environmental Science & Technology* 44, 2474-2481.

- Hidy, G.M. *Atmospheric Sulphur and Nitrogen Oxides*; Academic Press: San Diego, Clifornia,1994.
- Hilal, S.H., Karickhoff, S.W., Carreira, L.A., 1995. A rigorous test for SPARC's chemical reactivity models: Estimation of more than 4300 ionization pKas. *Quantitative Structure-Activity Relationships* 14, 348-355.
- Ho, K.F., Lee, S.C., Chan, C.K., Yu, J.C., Chow, J.C., Yao, X.H., 2003. Characterization of chemical species in PM<sub>2.5</sub> and PM<sub>10</sub> aerosols in Hong Kong. *Atmospheric Environment* 37, 31-39.
- Hu, D., Bian, Q., Li, T.W.Y., Lau, A.K.H., Yu, J.Z., 2008. Contributions of isoprene, monoterpenes,  $\beta$ -caryophyllene, and toluene to secondary organic aerosols in Hong Kong during the summer of 2006. *Journal of Geophysical Research: Atmospheres* 113, D22206.
- Hwang, I., Hopke, P.K., Pinto, J.P., 2008. Source apportionment and spatial distributions of coarse particles during the regional air pollution study. *Environmental Science & Technology* 42, 3524-3530.
- 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., 2005. 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, 300-310.
- 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, 300-310.
- Ito, K., Xue, N., Thurston, G., 2004. Spatial variation of PM<sub>2.5</sub> chemical species and source-apportioned mass concentrations in New York City. *Atmospheric Environment* 38, 5269-5282.
- Jaekels, J.M., Bae, M.S., Schauer, J.J., 2007. Positive matrix factorization (PMF) analysis of molecular marker measurements to quantify the sources of organic aerosols. *Environmental Science & Technology* 41, 5763-5769.
- Jenkins, B.M., Jones, A.D., Turn, S.Q., Williams, R.B., 1996. Emission factors for polycyclic aromatic hydrocarbons from biomass burning. *Environmental Science & Technology* 30, 2462-2469.
- Kadowaki, S., 1986. 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* 20, 1249-1253.
- Kawamura, K., Okuzawa, K., Aggarwal, S.G., Irie, H., Kanaya, Y., Wang, Z., 2013. Determination of gaseous and particulate carbonyls (glycolaldehyde, hydroxyacetone, glyoxal, methylglyoxal, nonanal and decanal) in the atmosphere at Mt. Tai. *Atmospheric Chemistry and Physics* 13, 5369-5380.
- Kessler, S.H., Smith, J.D., Che, D.L., Worsnop, D.R., Wilson, K.R., Kroll, J.H., 2010. Chemical sinks of organic aerosol: Kinetics and products of the heterogeneous oxidation of erythritol and levoglucosan. *Environmental Science & Technology* 44, 7005-7010.
- Kim, E., Hopke, P.K., 2008. Source characterization of ambient fine particles at multiple sites in the Seattle area. *Atmospheric Environment* 42, 6047-6056.

- Kim, E., Hopke, P.K., Pinto, J.P., Wilson, W.E., 2005. Spatial variability of fine particle mass, components, and source contributions during the regional air pollution study in St. Louis. *Environmental Science & Technology* 39, 4172-4179.
- Kim, J.Y., Lee, J.Y., Choi, S.D., Kim, Y.P., Ghim, Y.S., 2012a. Gaseous and particulate polycyclic aromatic hydrocarbons at the Gosan background site in East Asia. *Atmospheric Environment* 49, 311-319.
- Kim, S.-Y., Peel, J.L., Hannigan, M.P., Dutton, S.J., Sheppard, L., Clark, M.L., Vedal, S., 2012b. The Temporal Lag Structure of Short-term Associations of Fine Particulate Matter Chemical Constituents and Cardiovascular and Respiratory Hospitalizations. *Environmental Health Perspectives* 120, doi: 10.1289/ehp.1104721.
- Kleeman, M.J., Riddle, S.G., Robert, M.A., Jakober, C.A., 2008. Lubricating oil and fuel contributions to particulate matter emissions from light-duty gasoline and heavy-duty diesel vehicles. *Environmental Science & Technology* 42, 235-242.
- Kleindienst, T.E., Jaoui, M., Lewandowski, M., Offenberg, J.H., Lewis, C.W., Bhave, P.V., Edney, E.O., 2007. Estimates of the contributions of biogenic and anthropogenic hydrocarbons to secondary organic aerosol at a southeastern US location. *Atmospheric Environment* 41, 8288-8300.
- Kot, F., 2009. Boron sources, speciation and its potential impact on health. *Reviews in Environmental Science and Biotechnology* 8, 3-28.
- Krudysz, M., Dutton, S., Brinkman, G., Hannigan, M., Fine, P., Sioutas, C., Froines, J., 2009. Intra-community spatial variation of size-fractionated organic compounds in Long Beach, California. *Air Quality, Atmosphere & Health* 2, 69-88.

- Krudysz, M.A., Froines, J.R., Fine, P.M., Sioutas, C., 2008. Intra-community spatial variation of size-fractionated PM mass, OC, EC, and trace elements in the Long Beach, CA area. *Atmospheric Environment* 42, 5374-5389.
- Laden, F., Neas, L.M., Dockery, D.W., Schwartz, J., 2000. Association of fine particulate matter from different sources with daily mortality in six U.S. cities. *Environmental Health Perspectives* 108, 941-947.
- Larsen, R.K., Baker, J.E., 2003. Source apportionment of polycyclic aromatic hydrocarbons in the urban atmosphere: A comparison of three methods. *Environmental Science & Technology* 37, 1873-1881.
- Lee, S., Liu, W., Wang, Y., Russell, A.G., Edgerton, E.S., 2008. Source apportionment of PM<sub>2.5</sub>: Comparing PMF and CMB results for four ambient monitoring sites in the southeastern United States. *Atmospheric Environment* 42, 4126-4137.
- Lee, S., Russell, A.G., 2007. Estimating uncertainties and uncertainty contributors of CMB PM<sub>2.5</sub> source apportionment results. *Atmospheric Environment* 41, 9616-9624.
- Li, Z., Porter, E.N., Sjödin, A., Needham, L.L., Lee, S., Russell, A.G., Mulholland, J.A., 2009. Characterization of PM<sub>2.5</sub>-bound polycyclic aromatic hydrocarbons in Atlanta--Seasonal variations at urban, suburban, and rural ambient air monitoring sites. *Atmospheric Environment* 43, 4187-4193.
- Liang, C., Pankow, J.F., 1996. Gas/particle partitioning of organic compounds to environmental tobacco smoke: Partition coefficient measurements by desorption and comparison to urban particulate material. *Environmental Science & Technology* 30, 2800-2805.

- Liang, C., Pankow, J.F., Odum, J.R., Seinfeld, J.H., 1997. Gas/particle partitioning of semivolatile organic compounds to model inorganic, organic, and ambient smog aerosols. *Environmental Science & Technology* 31, 3086-3092.
- Liu, W., Wang, Y., Russell, A., Edgerton, E.S., 2006. Enhanced source identification of southeast aerosols using temperature-resolved carbon fractions and gas phase components. *Atmospheric Environment* 40, Supplement 2, 445-466.
- Mader, B.T., Pankow, J.F., 2002. Study of the effects of particle-phase carbon on the gas/particle partitioning of semivolatile organic compounds in the atmosphere using controlled field experiments. *Environmental Science & Technology* 36, 5218-5228.
- Majestic, B.J., Schauer, J.J., Shafer, M.M., 2007. Application of synchrotron radiation for measurement of iron red-ox speciation in atmospherically processed aerosols. *Atmospheric Chemistry and Physics* 7, 2475-2487.
- Mandalakis, M., Tsapakis, M., Tsoga, A., Stephanou, E.G., 2002. Gas-particle concentrations and distribution of aliphatic hydrocarbons, PAHs, PCBs and PCDD/Fs in the atmosphere of Athens (Greece). *Atmospheric Environment* 36, 4023-4035.
- 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., 2005. PM source apportionment and health effects. 3. Investigation of inter-method variations in associations between estimated source contributions of PM<sub>2.5</sub> and daily mortality in Phoenix, AZ. *Journal of Exposure Science and Environmental Epidemiology* 16, 311-320.
- Marmur, A., Park, S.K., Mulholland, J.A., Tolbert, P.E., Russell, A.G., 2006. Source apportionment of PM<sub>2.5</sub> in the southeastern United States using receptor and emissions-

- based models: Conceptual differences and implications for time-series health studies. *Atmospheric Environment* 40, 2533-2551.
- May, A.A., Saleh, R., Hennigan, C.J., Donahue, N.M., Robinson, A.L., 2012. Volatility of organic molecular markers used for source apportionment analysis: measurements and implications for atmospheric lifetime. *Environmental Science & Technology* 46, 12435-12444.
- Maynard, D., Coull, B.A., Gryparis, A., Schwartz, J., 2007. Mortality risk associated with short-term exposure to traffic particles and sulfates. *Environmental Health Perspectives* 115, 751-755.
- McDonald, J.D., Zielinska, B., Fujita, E.M., Sagebiel, J.C., Chow, J.C., Watson, J.G., 2000. Fine particle and gaseous emission rates from residential wood combustion. *Environmental Science & Technology* 34, 2080-2091.
- McKenzie, E.R., Money, J.E., Green, P.G., Young, T.M., 2009. Metals associated with stormwater-relevant brake and tire samples. *Science of The Total Environment* 407, 5855-5860.
- Miguel, A.H., Kirchstetter, T.W., Harley, R.A., Hering, S.V., 1998. On-road emissions of particulate polycyclic aromatic hydrocarbons and black carbon from gasoline and diesel vehicles. *Environmental Science & Technology* 32, 450-455.
- Minguillon, M.C., Arhami, M., Schauer, J.J., Sioutas, C., 2008. Seasonal and spatial variations of sources of fine and quasi-ultrafine particulate matter in neighborhoods near the Los Angeles-Long Beach harbor. *Atmospheric Environment* 42, 7317-7328.

- Mooibroek, D., Schaap, M., Weijers, E.P., Hoogerbrugge, R., 2011. Source apportionment and spatial variability of PM(2.5) using measurements at five sites in the Netherlands. *Atmospheric Environment* 45, 4180-4191.
- Murphy, D.M., Thomson, D.S., Mahoney, M.J., 1998. In situ measurements of organics, meteoritic material, mercury, and other elements in aerosols at 5 to 19 kilometers. *Science* 282, 1664-1669.
- Nahir, T.M., 1999. Analysis of semivolatile organic compounds in fuels using gas chromatography-mass spectrometry. *Journal of Chemical Education* 76, 1695.
- Neff, W.D., 1997. The Denver brown cloud studies from the perspective of model assessment needs and the role of meteorology. *Journal of the Air & Waste Management Association* 47, 269-285.
- NIST, 1994. Technical Note 1297, 1994 Edition: Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results. National Institute of Standards and Technology. Available from: <http://physics.nist.gov/Pubs/guidelines/appa.html>.
- NIOSH. Method 5040, Issue 3: Diesel Particulate Matter (as Elemental Carbon). NIOSH Manual of Analytical Methods (NMAM), 4<sup>th</sup>, ed.; National Institute of Occupational Safety and Health: Cincinnati, OH, 2003.
- Odum, J.R., Hoffmann, T., Bowman, F., Collins, D., Flagan, R.C., Seinfeld, J.H., 1996. Gas/particle partitioning and secondary organic aerosol yields. *Environmental Science & Technology* 30, 2580-2585.
- Paatero, P., Hopke, P.K., Song, X.-H., Ramadan, Z., 2002. Understanding and controlling rotations in factor analytic models. *Chemometrics and Intelligent Laboratory Systems* 60, 253-264.

- Paatero, P., 1998a. User's Guide for Positive Matrix Factorization Program PMF2 and PMF3, Part 1: Tutorial. University of Helsinki, Helsinki, Finland.
- Paatero, P., 1998b. User's Guide for Positive Matrix Factorization Program PMF2 and PMF3, Part 2: Reference. University of Helsinki, Helsinki, Finland.
- Paatero, P., Tapper, U., 1994. Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values. *Environmetrics* 5, 111-126.
- Pankow, J.F., 1987. Review and comparative analysis of the theories on partitioning between the gas and aerosol particulate phases in the atmosphere. *Atmospheric Environment* (1967) 21, 2275-2283.
- Pankow, J.F., 1991. Common y-intercept and single compound regressions of gas-particle partitioning data vs 1/T. *Atmospheric Environment. Part A. General Topics* 25, 2229-2239.
- Pankow, J.F., 1992. Application of common y-intercept regression parameters for log K<sub>p</sub> vs 1/T for predicting gas-particle partitioning in the urban environment. *Atmospheric Environment. Part A. General Topics* 26, 2489-2497.
- Pankow, J.F., 1994a. An absorption model of gas/particle partitioning of organic compounds in the atmosphere. *Atmospheric Environment* 28, 185-188.
- Pankow, J.F., 1994b. An absorption model of the gas/aerosol partitioning involved in the formation of secondary organic aerosol. *Atmospheric Environment* 28, 189-193.
- Pankow, J.F., Asher, W.E., 2008. SIMPOL.1: a simple group contribution method for predicting vapor pressures and enthalpies of vaporization of multifunctional organic compounds. *Atmospheric Chemistry and Physics* 8, 2773-2796.

- Pankow, J.F., Bidleman, T.F., 1992. Interdependence of the slopes and intercepts from log-log correlations of measured gas-particle partitioning and vapor pressure—I. theory and analysis of available data. *Atmospheric Environment. Part A. General Topics* 26, 1071-1080.
- Pankow, J.F., Chang, E.I., 2008. Variation in the sensitivity of predicted levels of atmospheric organic particulate matter (OPM). *Environmental Science & Technology* 42, 7321-7329.
- Pankow, J.F., Storey, J.M.E., Yamasaki, H., 1993. Effects of relative humidity on gas/particle partitioning of semivolatile organic compounds to urban particulate matter. *Environmental Science & Technology* 27, 2220-2226.
- Pathak, R.K., Stanier, C.O., Donahue, N.M., Pandis, S.N., 2007. Ozonolysis of  $\alpha$ -pinene at atmospherically relevant concentrations: Temperature dependence of aerosol mass fractions (yields). *Journal of Geophysical Research: Atmospheres* 112, D03201.
- Peel, J., Kim, S., Hannigan, M., Dutton, S., Clark, M., Vedal, S., 2009. Fine Particulate Matter Chemical Speciation and Hospital Admissions in Denver: The Denver Aerosol Sources and Health Study (DASH). *American Association for Aerosol Research* 295.
- Perraud, V., Bruns, E.A., Ezell, M.J., Johnson, S.N., Yu, Y., Alexander, M.L., Zelenyuk, A., Imre, D., Chang, W.L., Dabdub, D., Pankow, J.F., Finlayson-Pitts, B.J., 2012. Nonequilibrium atmospheric secondary organic aerosol formation and growth. *Proceedings of the National Academy of Sciences*. doi:10.1073/pnas.1119909109
- Peters, A.J., Lane, D.A., Gundel, L.A., Northcott, G.L., Jones, K.C., 2000. A comparison of high volume and diffusion denuder samplers for measuring semivolatile organic compounds in the atmosphere. *Environmental Science & Technology* 34, 5001-5006.

- Phuleria, H.C., Geller, M.D., Fine, P.M., Sioutas, C., 2006. Size-resolved emissions of organic tracers from light- and heavy-duty vehicles measured in a California roadway tunnel. *Environmental Science & Technology* 40, 4109-4118.
- Phuleria, H.C., Sheesley, R.J., Schauer, J.J., Fine, P.M., Sioutas, C., 2007. Roadside measurements of size-segregated particulate organic compounds near gasoline and diesel-dominated freeways in Los Angeles, CA. *Atmospheric Environment* 41, 4653-4671.
- Polissar, A.V., Hopke, P.K., Paatero, P., 1998. Atmospheric aerosol over Alaska - 2. Elemental composition and sources. *Journal of Geophysical Research-Atmospheres* 103, 19045-19057.
- Politis, D.N., Romano, J.P., 1994. The Stationary Bootstrap. *Journal of the American Statistical Association* 89, 1303-1313.
- 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. *JAMA: The Journal of the American Medical Association* 287, 1132-1141.
- Pun, B.K., 2008. Development and initial application of the sesquiversion of MADRID. *Journal of Geophysical Research: Atmospheres* 113, n/a-n/a.
- Ravindra, K., Bencs, L., Wauters, E., de Hoog, J., Deutsch, F., Roekens, E., Bleux, N., Berghmans, P., Van Grieken, R., 2006. Seasonal and site-specific variation in vapour and aerosol phase PAHs over Flanders (Belgium) and their relation with anthropogenic activities. *Atmospheric Environment* 40, 771-785.
- Ravindra, K., Sokhi, R., Van Grieken, R., 2008. Atmospheric polycyclic aromatic hydrocarbons: Source attribution, emission factors and regulation. *Atmospheric Environment* 42, 2895-2921.

- Reff, A., Eberly, S.I., Bhave, P.V., 2007. Receptor modeling of ambient particulate matter data using positive matrix factorization: Review of existing methods. *Journal of the Air & Waste Management Association* 57, 146-154.
- Riddle, S.G., Jakober, C.A., Robert, M.A., Cahill, T.M., Charles, M.J., Kleeman, M.J., 2007. Large PAHs detected in fine particulate matter emitted from light-duty gasoline vehicles. *Atmospheric Environment* 41, 8658-8668.
- Robinson, A.L., Subramanian, R., Donahue, N.M., Bernardo-Bricker, A., Rogge, W.F., 2006. Source apportionment of molecular markers and organic aerosols-1. Polycyclic aromatic hydrocarbons and methodology for data visualization. *Environmental Science & Technology* 40, 7803-7810.
- Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., Simoneit, B.R.T., 1993a. Sources of fine organic aerosol .2. Noncatalyst and catalyst-equipped automobiles and heavy-duty diesel trucks. *Environmental Science & Technology* 27, 636-651.
- Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., Simoneit, B.R.T., 1993b. Sources of fine organic aerosol .3. Road dust, tire debris, and organometallic brake lining dust - roads as sources and sinks. *Environmental Science & Technology* 27, 1892-1904.
- Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., Simoneit, B.R.T., 1993c. Sources of fine organic aerosol .4. Particulate abrasion products from leaf surfaces of urban plants. *Environmental Science & Technology* 27, 2700-2711.
- Ronald C, H., 2003. Multivariate receptor modeling by N-dimensional edge detection. *Chemometrics and Intelligent Laboratory Systems* 65, 179-189.
- Schaap, M., Spindler, G., Schulz, M., Acker, K., Maenhaut, W., Berner, A., Wieprecht, W., Streit, N., Müller, K., Brüggemann, E., Chi, X., Putaud, J.P., Hitzenberger, R., Puxbaum, H.,

- Baltensperger, U., ten Brink, H., 2004. Artefacts in the sampling of nitrate studied in the "INTERCOMP" campaigns of EUROTRAC-AEROSOL. *Atmospheric Environment* 38, 6487-6496.
- Schauer, J.J., 2003. Evaluation of elemental carbon as a marker for diesel particulate matter. *Journal of Exposure Analysis and Environmental Epidemiology* 13, 443-453.
- Schauer, J.J., Kleeman, M.J., Cass, G.R., Simoneit, B.R.T., 1999. Measurement of emissions from air pollution sources. 2. C-1 through C-30 organic compounds from medium duty diesel trucks. *Environmental Science & Technology* 33, 1578-1587.
- Schauer, J.J., Kleeman, M.J., Cass, G.R., Simoneit, B.R.T., 2001. Measurement of emissions from air pollution sources. 3. C-1-C-29 organic compounds from fireplace combustion of wood. *Environmental Science & Technology* 35, 1716-1728.
- Schauer, J.J., Mader, B.T., Deminter, J.T., Heidemann, G., Bae, M.S., Seinfeld, J.H., Flagan, R.C., Cary, R.A., Smith, D., Huebert, B.J., Bertram, T., Howell, S., Kline, J.T., Quinn, P., Bates, T., Turpin, B., Lim, H.J., Yu, J.Z., Yang, H., Keywood, M.D., 2003. ACE-Asia intercomparison of a thermal-optical method for the determination of particle-phase organic and elemental carbon. *Environmental Science & Technology* 37, 993-1001.
- Schauer, J.J., Rogge, W.F., Hildemann, L.M., Mazurek, M.A., Cass, G.R., 1996. Source apportionment of airborne particulate matter using organic compounds as tracers. *Atmospheric Environment* 30, 3837-3855.
- Schkolnik, G., Falkovich, A.H., Rudich, Y., Maenhaut, W., Artaxo, P., 2005. New analytical method for the determination of levoglucosan, polyhydroxy compounds, and 2-methylerythritol and its application to smoke and rainwater samples. *Environmental Science & Technology* 39, 2744-2752.

- Schnelle-Kreis, J., Sklorz, M., Orasche, J., Stölzel, M., Peters, A., Zimmermann, R., 2007. Semi volatile organic compounds in ambient PM<sub>2.5</sub>. Seasonal trends and daily resolved source contributions. *Environmental Science & Technology* 41, 3821-3828.
- Seinfeld, J., Pandis, S., 1998. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change*. Wiley-Interscience, p149-153.
- Shrivastava, M.K., Subramanian, R., Rogge, W.F., Robinson, A.L., 2007. Sources of organic aerosol: Positive matrix factorization of molecular marker data and comparison of results from different source apportionment models. *Atmospheric Environment* 41, 9353-9369.
- Simcik, M.F., Franz, T.P., Zhang, H., Eisenreich, S.J., 1998. Gas-particle partitioning of PCBs and PAHs in the Chicago urban and adjacent coastal atmosphere: States of equilibrium. *Environmental Science & Technology* 32, 251-257.
- Simcik, M.F., Zhang, H., Eisenreich, S.J., Franz, T.P., 1997. Urban contamination of the Chicago/Coastal Lake Michigan atmosphere by PCBs and PAHs during AEOLOS. *Environmental Science & Technology* 31, 2141-2147.
- Simo, R., Grimalt, J.O., Albaiges, J., 1997. Loss of unburned-fuel hydrocarbons from combustion aerosols during atmospheric transport. *Environmental Science & Technology* 31, 2697-2700.
- Simoneit, B.R.T., Mazurek, M.A., 1982. Organic-matter of the troposphere .2. Natural background of biogenic lipid matter in aerosols over the rural western United States. *Atmospheric Environment* 16, 2139-2159.
- Simoneit, B.R.T., Schauer, J.J., Nolte, C.G., Oros, D.R., Elias, V.O., Fraser, M.P., Rogge, W.F., Cass, G.R., 1999. Levoglucosan, a tracer for cellulose in biomass burning and atmospheric particles. *Atmospheric Environment* 33, 173-182.

- Stanek, L.W., Sacks, J.D., Dutton, S.J., Dubois, J.-J.B., 2011. Attributing health effects to apportioned components and sources of particulate matter: An evaluation of collective results. *Atmospheric Environment* 45, 5655-5663.
- Stelson, A.W., Seinfeld, J.H., 1982. Relative humidity and temperature dependence of the ammonium nitrate dissociation constant. *Atmospheric Environment* (1967) 16, 983-992.
- Subramanian, R., Khlystov, A.Y., Cabada, J.C., Robinson, A.L., 2004. Positive and negative artifacts in particulate organic carbon measurements with denuded and undenuded sampler configurations special issue of aerosol science and technology on findings from the fine particulate matter supersites program. *Aerosol Science and Technology* 38, 27-48.
- Subramanyam, V., Valsaraj, K.T., Thibodeaux, L.J., Reible, D.D., 1994. Gas-to-particle partitioning of polycyclic aromatic hydrocarbons in an urban atmosphere. *Atmospheric Environment* 28, 3083-3091.
- Surratt, J.D., Chan, A.W.H., Eddingsaas, N.C., Chan, M., Loza, C.L., Kwan, A.J., Hersey, S.P., Flagan, R.C., Wennberg, P.O., Seinfeld, J.H., 2010. Reactive intermediates revealed in secondary organic aerosol formation from isoprene. *Proceedings of the National Academy of Sciences* 107, 6640-6645.
- Thurston, G.D., Spengler, J.D., 1985. A quantitative assessment of source contributions to inhalable particulate matter pollution in metropolitan Boston. *Atmospheric Environment* 19, 9-25.
- Tsapakis, M., Stephanou, E.G., 2005. Occurrence of gaseous and particulate polycyclic aromatic hydrocarbons in the urban atmosphere: study of sources and ambient temperature effect on the gas/particle concentration and distribution. *Environmental Pollution* 133, 147-156.

- U.S. EPA, 2009. Integrated Science Assessment for Particulate Matter (Final Report). U.S. Environmental Protection Agency, Washington, DC, EPA/600/R-08/139F.
- U.S. EPA, 2011. NCore Multipollutant Monitoring Network. Retrieved September 12, 2011 from <http://www.epa.gov/ttnamti1/ncore/>.
- Vedal, S., Hannigan, M.P., Dutton, S.J., Miller, S.L., Milford, J.B., Rabinovitch, N., Kim, S.Y., Sheppard, L., 2009. The Denver Aerosol Sources and Health (DASH) study: Overview and early findings. *Atmospheric Environment* 43, 1666-1673.
- Von Schneidmesser, E., Zhou, J.B., Stone, E.A., Schauer, J.J., Qasrawi, R., Abdeen, Z., Shpund, J., Vanger, A., Sharf, G., Moise, T., Brenner, S., Nassar, K., Saleh, R., Al-Mahasneh, Q.M., Sarnat, J.A., 2010a. Seasonal and spatial trends in the sources of fine particle organic carbon in Israel, Jordan, and Palestine. *Atmospheric Environment* 44, 3669-3678.
- Von Schneidmesser, E., Zhou, J.B., Stone, E.A., Schauer, J.J., Shpund, J., Brenner, S., Qasrawi, R., Abdeen, Z., Sarnat, J.A., 2010b. Spatial Variability of Carbonaceous Aerosol Concentrations in East and West Jerusalem. *Environmental Science & Technology* 44, 1911-1917.
- Vukmirović, Z.B., 1986. Increase of particulate sulphate content in the atmosphere of Belgrade in the summer. *Atmospheric Environment* (1967) 20, 985-988.
- Walgraeve, C., Demeestere, K., Dewulf, J., Zimmermann, R., Van Langenhove, H., 2010. Oxygenated polycyclic aromatic hydrocarbons in atmospheric particulate matter: Molecular characterization and occurrence. *Atmospheric Environment* 44, 1831-1846.
- Wang, Z., Na, G., Ma, X., Fang, X., Ge, L., Gao, H., Yao, Z., 2013. Occurrence and gas/particle partitioning of PAHs in the atmosphere from the North Pacific to the Arctic Ocean. *Atmospheric Environment* 77, 640-646.

- Watson, J.G., Robinson, N.F., Chow, J.C., Henry, R.C., Kim, B.M., Pace, T.G., Meyer, E.L., Nguyen, Q., 1990. The USEPA/DRI chemical mass balance receptor model, CMB 7.0. *Environmental Software* 5, 38-49.
- Williams, B.J., Goldstein, A.H., Kreisberg, N.M., Hering, S.V., 2010. In situ measurements of gas/particle-phase transitions for atmospheric semivolatile organic compounds. *Proceedings of the National Academy of Sciences* 107, 6676-6681.
- Wilson, J.G., Kingham, S., Pearce, J., Sturman, A.P., 2005. A review of intraurban variations in particulate air pollution: Implications for epidemiological research. *Atmospheric Environment* 39, 6444-6462.
- Wongphatarakul, V., Friedlander, S.K., Pinto, J.P., 1998. A comparative study of PM<sub>2.5</sub> ambient aerosol chemical databases. *Environmental Science & Technology* 32, 3926-3934.
- Xie, M., Barsanti, K.C., Hannigan, M.P., Dutton, S.J., Vedal, S., 2013a. Positive matrix factorization of PM<sub>2.5</sub> - eliminating the effects of gas/particle partitioning of semivolatile organic compounds. *Atmospheric Chemistry and Physics* 13, 7381-7393.
- Xie, M., Coons, T.L., Dutton, S.J., Milford, J.B., Miller, S.L., Peel, J.L., Vedal, S., Hannigan, M.P., 2012a. Intra-urban spatial variability of PM<sub>2.5</sub>-bound carbonaceous components. *Atmospheric Environment* 60, 486-494.
- Xie, M., Coons, T.L., Hemann, J.G., Dutton, S.J., Milford, J.B., Peel, J.L., Miller, S.L., Kim, S.-Y., Vedal, S., Sheppard, L., Hannigan, M.P., 2012b. Intra-urban spatial variability and uncertainty assessment of PM<sub>2.5</sub> sources based on carbonaceous species. *Atmospheric Environment* 60, 305-315.
- Xie, M., Hannigan, M.P., Dutton, S.J., Milford, J.B., Hemann, J.G., Miller, S.L., Schauer, J.J., Peel, J.L., Vedal, S., 2012c. Positive Matrix Factorization of PM<sub>2.5</sub>: Comparison and

- Implications of Using Different Speciation Data Sets. *Environmental Science & Technology* 46, 11962-11970.
- Xie, M., Piedrahita, R., Dutton, S.J., Milford, J.B., Hemann, J.G., Peel, J.L., Miller, S.L., Kim, S.-Y., Vedal, S., Sheppard, L., Hannigan, M.P., 2013b. Positive matrix factorization of a 32-month series of daily PM<sub>2.5</sub> speciation data with incorporation of temperature stratification. *Atmospheric Environment* 65, 11-20.
- Yamasaki, H., Kuwata, K., Miyamoto, H., 1982. Effects of ambient temperature on aspects of airborne polycyclic aromatic hydrocarbons. *Environmental Science & Technology* 16, 189-194.
- Yang, H.H., Lee, W.J., Chen, S.J., Lai, S.O., 1998. PAH emission from various industrial stacks. *Journal of Hazardous Materials* 60, 159-174.
- Yang, M., Qi, H., Jia, H.-L., Ren, N.-Q., Ding, Y.-S., Ma, W.-L., Liu, L.-Y., Hung, H., Sverko, E., Li, Y.-F., 2013. Polybrominated diphenyl ethers in air across China: Levels, compositions, and gas-particle partitioning. *Environmental Science & Technology* 47, 8978-8984.
- Yunker, M.B., Macdonald, R.W., Vingarzan, R., Mitchell, R.H., Goyette, D., Sylvestre, S., 2002. PAHs in the Fraser River basin: a critical appraisal of PAH ratios as indicators of PAH source and composition. *Organic Geochemistry* 33, 489-515.
- Zanobetti, A., Schwartz, J., 2009. The effect of fine and coarse particulate air pollution on mortality: A national analysis. *Environmental Health Perspectives* 117, 898-903.
- Zeger, S.L., Thomas, D., Dominici, F., Samet, J.M., Schwarz, J., Dockery, D., Cohen, A., 2001. Exposure measurement error in time-series studies of air pollution: Concepts and consequences (vol 108, pg 419, 2000). *Environmental Health Perspectives* 109, A517-A517.

- Zhang, J., Fan, X., Graham, L., Chan, T.W., Brook, J.R., 2012. Evaluation of an annular denuder system for carbonaceous aerosol sampling of diesel engine emissions. *Journal of the Air & Waste Management Association* 63, 87-99.
- Zhang, Y., Schauer, J.J., Shafer, M.M., Hannigan, M.P., Dutton, S.J., 2008. Source apportionment of in vitro reactive oxygen species bioassay activity from atmospheric particulate matter. *Environmental Science & Technology* 42, 7502-7509.
- Zhang, Y.X., Sheesley, R.J., Bae, M.S., Schauer, J.J., 2009. Sensitivity of a molecular marker based positive matrix factorization model to the number of receptor observations. *Atmospheric Environment* 43, 4951-4958.
- Zhao, Y., Kreisberg, N.M., Worton, D.R., Isaacman, G., Weber, R.J., Liu, S., Day, D.A., Russell, L.M., Markovic, M.Z., VandenBoer, T.C., Murphy, J.G., Hering, S.V., Goldstein, A.H., 2013. Insights into secondary organic aerosol formation mechanisms from measured gas/particle partitioning of specific organic tracer compounds. *Environmental Science & Technology* 47, 3781-3787.
- Zielinska, B., McDonald, J., Hayes, T., Chow, J.C., Fujita, E.M., Watson, J.G., 1998. Northern Front Range Air Quality Study final report, volume b: source measurements. Prepared for Colorado State University, Fort Collins, CO. Desert Research Institute, Reno, NV.

## CHAPTER 11 SUPPLEMENTAL INFORMATION

### 11.1 SUPPLEMENTAL TABLES

Table 2.S1 The fraction of the compounds (%) detected in samples of one quantification batch (sample No. = 79) falling within, below and above the calibration range.

Molecular markers	Within	Below	Above	Molecular markers	Within	Below	Above
<i>n-Alkanes</i>				<i>Steranes</i>			
docosane	62.0	1.27	36.7	20R-abb & 20S-aaa-cholestane	84.8	3.80	11.4
tricosane	64.6	1.27	34.2	20R & S-abb-methylcholestane	79.8	13.9	6.33
tetracosane	86.1	1.27	12.7	20R & S-abb-ethylcholestane	86.1	2.53	11.4
pentacosane	81.0	1.27	17.7	a-22,29,30-trisnorhopane	81.0	2.53	16.5
hexacosane	88.6	10.1	1.27	ba-30-norhopane	35.4	1.27	63.3
heptacosane	83.5	10.1	6.33	ab-hopane	58.2	1.27	40.5
octacosane	84.8	15.2	0.00	22S-ab-30-homohopane	79.8	5.06	15.2
nonacosane	76.0	5.06	19.0	22R-ab-30-homohopane	83.5	6.33	10.1
triacontane	68.4	31.7	0.00	22S-ab-30-bishomohopane	81.0	13.9	5.06
hentriacontane	76.0	6.33	17.7	22R-ab-30-bishomohopane	83.5	13.9	2.53
dotriacontane	62.0	38.0	0.00	<i>n-Alkanoic acids</i>			
tritriacontane	86.1	13.9	0.00	dodecanoic acid	39.2	0.00	60.8
tetratriacontane	69.6	30.4	0.00	tridecanoic acid	70.9	20.3	8.86
pentatriacontane	54.4	45.6	0.00	tetradecanoic acid	30.4	0.00	69.6
<i>PAHs</i>				pentadecanoic acid	92.4	1.27	6.33
fluoranthene	53.2	0.00	46.8	hexadecanoic acid	1.27	0.00	98.7
pyrene	55.7	1.27	43.0	heptadecanoic acid	60.8	39.2	0.00
benzo[ghi]fluoranthene	68.4	2.53	29.1	octadecanoic acid	10.1	0.00	89.9
cyclopenta[cd]pyrene	58.2	39.2	2.53	<i>Sterols and methoxyphenols</i>			
benz[a]anthracene	63.3	27.9	8.86	cholesterol	91.1	3.80	5.06
chrysene/triphenylene	50.6	1.27	48.1	stigmasterol	69.6	25.3	5.06
benzo[b&k]fluoranthene	84.8	10.1	5.06	acetovanillone	41.8	58.2	0.00
benz[a&e]pyrene	82.3	12.7	5.06	vanillin	69.6	1.27	29.1
indeno[1,2,3-cd]pyrene	82.3	8.86	8.86	syringaldehyde	58.2	10.1	31.7
benzo[ghi]perylene	84.8	3.80	11.4	coniferaldehyde	65.8	32.9	1.27
coronene	91.1	3.80	5.06	acetosyringone	15.2	84.8	0.00
2-methylfluoranthene	70.9	1.27	27.9				
methyl-202-PAH sum	34.2	1.27	64.6				
retene	67.1	5.06	27.9				
fluorenone	39.2	0.00	60.8				
1H-phenalen-1-one	50.6	1.27	48.1				
xanthone	68.4	1.27	30.4				
1,8-naphthalic anhydride	53.2	1.27	45.6				
anthracene-9,10-dione	27.9	1.27	70.9				
benz[de]anthracene-7-one	78.5	15.2	6.33				

Table 2.S2 Statistics for PM<sub>2.5</sub> carbonaceous components at the four sampling sites (March 1, 2008-March 14, 2009).

Molecular Marker (ng m <sup>-3</sup> )	PAL (N=63)						EDI (N=59)						ALS (N=63)						MAS (N=62)					
	Mean	Median	CV <sup>a</sup>	S/N <sup>b</sup>	BDL (%) <sup>c</sup>	Mean	Median	CV	S/N	BDL (%)	Mean	Median	CV	S/N	BDL (%)	Mean	Median	CV	S/N	BDL (%)				
<i>n-Alkanes</i>																								
docosane	2.19	1.62	0.94	8.56	2	2.00	1.50	1.15	8.33	0	2.23	1.85	0.88	8.50	0	1.66	1.47	0.71	8.54	0				
tricosane	4.00	1.98	1.33	11.2	0	3.10	1.77	1.05	10.7	0	2.75	1.78	0.94	10.4	0	2.30	1.58	0.88	10.7	0				
tetracosane	1.89	1.17	1.04	13.1	0	1.11	0.83	0.92	12.2	0	1.25	1.00	0.83	12.6	0	1.05	0.76	0.75	12.5	0				
pentacosane	1.88	1.13	1.03	10.8	0	1.66	0.98	1.05	10.8	0	1.69	1.17	0.89	10.8	0	1.62	1.11	0.85	11.2	0				
hexacosane	0.85	0.59	1.01	9.43	2	0.56	0.43	0.76	6.73	2	0.73	0.52	1.00	8.51	0	0.66	0.50	0.87	8.98	0				
heptacosane	1.64	0.87	1.25	12.7	0	1.31	0.68	0.94	11.2	0	1.62	0.93	1.22	12.3	0	1.42	0.83	1.02	12.8	0				
octacosane	0.82	0.48	1.12	9.24	6	0.62	0.37	1.04	7.08	5	0.89	0.44	1.87	9.35	3	0.73	0.38	1.26	9.25	5				
nonacosane	4.20	1.60	1.89	13.8	0	3.43	1.54	1.52	13.4	0	4.36	1.49	2.26	13.1	0	3.51	1.56	1.57	13.7	0				
triacontane	0.92	0.47	1.36	9.96	5	0.67	0.39	1.26	7.24	12	0.74	0.35	1.26	8.03	13	0.82	0.41	1.26	10.2	5				
hentriacontane	11.2	2.39	2.04	11.9	0	9.66	2.50	1.74	11.9	0	12.4	2.54	2.57	11.4	0	9.67	3.93	1.64	11.9	0				
dotriacontane	0.20	0.14	1.00	4.37	17	0.15	0.12	1.12	3.60	15	0.16	0.12	1.28	3.75	17	0.17	0.12	1.27	4.21	11				
triacontane	0.36	0.22	1.22	7.14	0	0.32	0.18	1.22	6.32	0	0.32	0.20	1.31	6.73	0	0.35	0.21	1.31	7.57	0				
tetracontane	0.22	0.13	1.24	7.32	8	0.18	0.10	1.18	6.18	3	0.21	0.14	1.23	6.91	8	0.20	0.11	1.23	7.08	5				
pentatriacontane	0.16	0.09	1.27	7.47	16	0.09	0.07	1.25	4.75	24	0.12	0.08	1.04	5.51	22	0.11	0.08	0.95	5.93	16				
Subtotal	30.5	12.7	1.41			24.9	12.9	1.17			29.4	12.0	1.66			24.3	15.3	1.16						
<i>PAHs</i>																								
fluoranthene	0.21	0.16	0.86	13.1	0	0.24	0.17	1.00	14.0	0	0.40	0.30	0.81	15.3	0	0.35	0.31	0.68	15.6	0				
pyrene	0.13	0.11	1.31	12.5	0	0.17	0.08	1.50	13.9	0	0.29	0.20	0.94	14.0	0	0.23	0.20	0.71	13.6	0				
benz[ghi]fluoranthene	0.08	0.06	1.03	19.3	0	0.11	0.08	1.19	23.6	0	0.17	0.12	0.97	22.9	0	0.16	0.14	0.80	24.2	0				
C-pyr	0.04	0.02	1.30	10.1	5	0.04	0.02	1.31	9.99	12	0.05	0.04	1.09	11.4	3	0.06	0.03	1.39	13.3	3				
BaA	0.05	0.02	1.50	12.1	8	0.07	0.04	1.34	15.3	3	0.09	0.06	1.13	15.4	0	0.11	0.06	1.57	17.4	2				
Chry	0.18	0.13	1.03	16.7	0	0.23	0.19	0.99	17.6	0	0.33	0.24	0.89	17.1	0	0.33	0.24	0.87	18.0	0				
BbF	0.22	0.10	1.21	14.5	0	0.28	0.17	1.15	15.6	0	0.36	0.17	1.12	15.5	0	0.37	0.24	1.13	17.0	0				
BaP	0.18	0.09	1.15	10.6	3	0.21	0.15	1.13	11.3	0	0.28	0.15	1.12	12.8	0	0.29	0.19	1.11	14.6	0				
Indeno[1,2,3-cd]pyrene	0.03	0.02	0.98	8.26	6	0.03	0.02	0.79	7.86	7	0.04	0.02	0.99	9.84	6	0.05	0.03	0.97	12.1	5				
benz[ghi]perylene	0.06	0.04	0.89	8.16	3	0.06	0.05	0.72	7.57	3	0.09	0.06	0.91	9.81	3	0.09	0.08	0.71	11.3	0				
coronene	0.03	0.02	0.82	9.58	2	0.03	0.02	0.76	9.55	2	0.04	0.03	0.92	11.0	0	0.04	0.03	0.80	11.1	0				
2-methylfluoranthene	0.20	0.14	1.01	14.5	2	0.22	0.13	1.21	14.6	0	0.29	0.24	0.79	15.8	0	0.25	0.22	0.67	16.3	0				
Methyl-202-PAH sum	0.68	0.47	1.24	16.9	0	0.72	0.52	1.06	16.8	0	1.05	0.84	0.86	17.1	0	0.83	0.77	0.71	17.5	0				
retene	0.72	0.28	1.69	16.3	6	0.64	0.24	1.50	15.4	5	0.41	0.20	1.16	13.2	8	0.43	0.20	1.37	14.7	5				
Subtotal	2.60	1.94	1.05			2.85	2.24	1.01			3.60	3.10	0.78			3.34	2.84	0.74						
<i>Oxy-PAHs</i>																								
fluorenone	1.08	0.57	1.14	17.0	0	1.36	0.79	1.12	17.8	0	1.24	0.74	1.13	16.7	0	1.22	0.56	1.31	18.4	0				
1H-phenalen-1-one	0.45	0.26	1.19	18.4	0	0.66	0.50	0.97	18.8	2	0.93	0.58	0.98	17.9	0	0.65	0.38	1.32	18.8	0				
xanthone	0.25	0.11	1.26	16.8	0	0.27	0.13	1.22	17.4	2	0.26	0.13	1.15	15.8	0	0.21	0.12	1.09	16.5	2				
1,8-naphthalic anhydride	0.38	0.16	1.39	18.4	0	0.52	0.21	1.50	19.3	0	0.60	0.27	1.57	17.8	0	0.46	0.21	1.32	18.8	0				
anthracene-9,10-dione	0.53	0.23	1.26	18.9	0	0.57	0.35	0.98	19.2	0	0.78	0.43	1.16	17.8	0	0.67	0.34	1.19	18.7	0				
benz[de]anthracene-7-one	0.06	0.04	1.12	14.9	2	0.08	0.05	1.11	16.7	0	0.10	0.06	1.00	16.1	0	0.10	0.06	1.11	17.2	0				
Subtotal	2.75	2.34	0.71			3.47	3.56	0.61			3.92	3.28	0.71			3.32	2.71	0.70						

(a) Coefficient of variation (CV) = standard deviation/mean concentration.

(b) Signal to noise ratio (mean concentration/mean uncertainty).

(c) Percent of observations not significantly different from zero using a p-value of 0.05.

Table 2.S2 Continued

Molecular Marker ( $\mu\text{g m}^{-3}$ )	Abbr.	PAL (N=63)					EDI (N=59)					ALS (N=63)					MAS (N=62)				
		Mean	Median	CV	S/N	BDL (%)	Mean	Median	CV	S/N	BDL (%)	Mean	Median	CV	S/N	BDL (%)	Mean	Median	CV	S/N	BDL (%)
<i>Steranes</i>																					
20R-abb & 20S-aaa-cholestane	27-RS-C	0.19	0.11	1.03	21.5	0	0.21	0.14	0.91	22.2	0	0.23	0.18	0.82	20.5	0	0.27	0.17	0.96	23.8	2
20R & S-abb-methylcholestane	28-RS-M	0.10	0.04	1.32	19.8	16	0.10	0.06	1.11	19.2	5	0.15	0.08	1.13	15.8	5	0.14	0.09	1.05	18.7	2
20R & S-abb-ethylcholestane	29-RS-E	0.11	0.09	0.79	16.3	2	0.11	0.09	0.85	16.5	2	0.16	0.11	1.05	16.2	0	0.13	0.10	0.78	17.0	2
a-22,29,30-trisnorhopane	TS	0.14	0.08	1.04	17.2	3	0.15	0.10	1.03	17.8	0	0.16	0.11	0.94	17.6	0	0.16	0.12	0.88	19.2	0
ba-30-norhopane	ba-N	0.25	0.21	0.74	18.2	0	0.28	0.24	0.73	18.6	0	0.42	0.31	0.91	16.9	0	0.38	0.32	0.68	18.2	0
ab-hopane	ab-H	0.19	0.13	0.90	14.4	0	0.21	0.15	0.91	14.4	2	0.31	0.18	1.00	14.3	0	0.28	0.21	0.80	14.9	0
22S-ab-30-homohopane	31abS	0.15	0.09	1.18	13.4	3	0.14	0.10	1.35	13.4	7	0.19	0.12	1.23	13.3	5	0.17	0.11	1.01	14.4	2
22R-ab-30-homohopane	31abR	0.13	0.08	1.24	12.9	6	0.12	0.08	1.25	12.8	12	0.16	0.12	1.18	13.0	5	0.14	0.09	0.92	13.9	3
22S-ab-30-bisnorhopane	32abS	0.09	0.05	1.43	10.8	22	0.11	0.09	1.16	12.0	12	0.14	0.09	1.05	12.6	11	0.11	0.08	1.14	13.5	10
22R-ab-30-bisnorhopane	32abR	0.08	0.04	1.37	10.6	24	0.10	0.08	1.20	11.5	15	0.12	0.08	1.20	12.3	13	0.11	0.07	1.30	13.3	10
Subtotal		1.43	1.01	0.90			1.52	1.19	0.86			2.03	1.66	0.86			1.89	1.37	0.79		
<i>n-alkanoic acids</i>																					
dodecanoic acid	C12:0	30.9	6.62	1.39	10.5	2	30.3	6.95	1.33	10.4	3	31.9	6.36	1.39	10.2	2	24.3	5.65	1.45	10.5	2
tridecanoic acid	C13:0	2.98	0.71	1.47	8.05	13	3.35	0.86	1.46	8.05	14	3.82	0.61	1.52	8.06	24	3.35	0.75	1.55	8.19	13
tetradecanoic acid	C14:0	18.0	4.19	1.40	9.54	2	23.8	6.12	1.43	9.51	0	26.9	5.05	1.66	9.40	2	21.4	6.02	1.65	9.57	2
pentadecanoic acid	C15:0	2.98	1.21	1.33	3.31	6	3.34	1.18	1.32	3.35	8	3.10	0.96	1.35	3.31	10	3.02	1.26	1.48	2.75	11
hexadecanoic acid	C16:0	106	62.9	1.41	7.67	0	109	73.1	0.97	7.68	0	103	72.5	1.00	7.65	0	108	77.3	1.09	7.72	0
heptadecanoic acid	C17:0	3.03	0.82	1.43	2.49	29	2.06	0.84	1.15	2.46	25	2.20	0.98	1.29	2.48	30	2.51	0.91	1.49	2.48	26
octadecanoic acid	C18:0	84.8	24.5	1.69	14.5	0	62.6	32.1	1.16	14.1	2	62.2	40.7	1.05	13.4	0	59.9	29.4	1.22	14.1	0
Subtotal		249	112	1.30			234	121	1.00			233	153	1.03			223	128	1.06		
<i>Sterols and methoxyphenols</i>																					
cholesterol	Cho	0.73	0.58	0.98	3.19	27	0.77	0.59	0.79	3.78	31	0.78	0.57	0.91	3.85	33	0.65	0.53	0.76	3.76	26
stigmasterol	Sti	0.84	0.79	0.78	16.2	14	1.00	0.78	0.83	19.0	12	0.87	0.76	0.85	16.0	16	0.86	0.75	0.71	19.4	6
acetovanillone	Ace	0.98	0.27	1.65	14.0	13	1.18	0.88	1.10	16.0	12	0.65	0.32	1.26	12.6	14	0.55	0.35	1.15	13.0	11
vanillin	Van	6.03	2.35	1.32	14.3	0	9.25	3.00	1.27	15.8	2	4.84	1.95	1.25	13.2	0	4.67	2.10	1.34	14.2	2
syringaldehyde	Syr	3.55	1.21	2.41	11.9	5	8.68	2.59	2.21	11.1	2	2.69	1.09	2.10	11.5	3	4.64	1.89	1.82	9.33	2
coniferaldehyde	Con	5.91	1.10	2.86	10.6	16	7.73	1.85	2.05	10.8	12	3.60	1.05	2.08	10.1	14	5.20	1.24	2.22	9.92	13
acetosyringone	Ace	0.30	0.10	2.13	6.02	25	0.77	0.18	2.44	8.75	14	0.30	0.11	1.72	6.26	22	0.45	0.13	2.32	5.65	19
Subtotal		18.3	8.36	1.88			29.4	12.2	1.54			13.7	7.15	1.36			17.0	8.15	1.47		
<i>Bulk species (<math>\mu\text{g m}^{-3}</math>)</i>																					
EC		0.31	0.27	0.61	2.20	21	0.32	0.29	0.61	2.18	12	0.66	0.54	0.69	4.09	3	0.48	0.44	0.62	3.67	3
OC		3.03	2.87	0.53	10.2	0	3.39	3.10	0.59	10.9	0	3.56	3.34	0.55	11.2	0	3.51	3.40	0.47	11.9	0
PM2.5		6.51	6.20	0.57	3.90	8	6.81	6.61	0.53	4.27	7	9.32	8.39	0.50	5.55	3	7.79	7.48	0.47	6.37	0

Table 2.S3 Statistics for PM<sub>2.5</sub> components derived from bi-weekly pairs of co-located samples at PAL (July 13, 2004- September 20, 2005).

Molecular marker (ng m <sup>-3</sup> )	Primary (N=31)					Duplicate (N=31)					<i>r</i> <sup>b</sup>	COD <sup>c</sup>
	Mean	SD <sup>a</sup>	Median	S/N	BDL %	Mean	SD	Median	S/N	BDL %		
<i>n-Alkanes</i>												
C22	1.86	1.69	1.32	14.8	0	1.4.0	1.27	1.1.0	14.0	0	0.97	0.16
C23	2.22	1.60	1.71	12.2	0	1.91	1.67	1.55	12.3	0	0.87	0.16
C24	1.32	0.97	1.12	8.22	0	1.23	0.84	1.04	8.07	0	0.85	0.20
C25	1.57	0.96	1.38	10.4	0	1.46	0.84	1.21	10.4	0	0.88	0.16
C26	0.90	0.60	0.83	12.4	0	0.81	0.52	0.63	11.7	3	0.90	0.17
C27	1.10	0.65	0.96	17.6	0	1.01	0.54	0.82	17.0	0	0.87	0.15
C28	0.78	0.64	0.45	18.5	0	0.7	0.54	0.43	16.5	0	0.90	0.18
C29	1.60	1.01	1.48	11.9	0	1.56	1.00	1.38	11.2	0	0.85	0.17
C30	0.68	0.59	0.36	6.80	10	0.63	0.52	0.39	5.87	13	0.89	0.17
C31	2.25	2.42	1.65	1.45	29	2.26	2.35	1.42	1.37	29	0.85	0.19
C32	0.16	0.15	0.09	0.14	94	0.40	0.32	0.42	0.40	45	0.43	0.43
C33	0.43	0.37	0.32	5.36	0	0.79	0.52	0.84	7.21	0	0.53	0.42
C34	0.27	0.21	0.21	5.98	3	0.54	0.40	0.51	8.18	3	0.50	0.41
C35	0.28	0.22	0.22	7.69	6	0.56	0.42	0.56	9.07	0	0.56	0.42
Subtotal	15.4	8.79	14.0			15.2	7.63	15.0				
<i>PAHs</i>												
Flu	0.22	0.22	0.13	5.43	0	0.16	0.13	0.10	4.88	0	0.82	0.29
Pyr	0.18	0.20	0.07	6.02	0	0.13	0.13	0.06	5.36	0	0.90	0.29
BghiF	0.13	0.16	0.05	9.52	0	0.09	0.10	0.05	9.43	0	0.94	0.25
C-pyr	0.06	0.09	0.02	9.44	6	0.04	0.04	0.02	8.95	23	0.92	0.28
BaA	0.11	0.15	0.03	11.7	10	0.07	0.08	0.04	11.0	6	0.91	0.26
CT	0.28	0.35	0.13	20.5	0	0.22	0.24	0.11	19.9	0	0.94	0.27
BbkF	0.34	0.42	0.14	10.6	0	0.24	0.25	0.13	9.63	0	0.93	0.25
BaeP	0.31	0.41	0.11	12.0	3	0.21	0.25	0.09	10.7	0	0.92	0.27
IP	0.09	0.09	0.04	11.4	0	0.08	0.09	0.04	12.0	3	0.96	0.17
BP	0.21	0.21	0.09	19.7	0	0.19	0.2	0.12	19.7	0	0.96	0.16
Cor	0.15	0.17	0.05	3.27	3	0.15	0.16	0.07	3.36	3	0.95	0.16
2M-Flu	0.27	0.35	0.12	6.74	0	0.18	0.21	0.09	5.51	0	0.92	0.25
M-202	0.86	1.15	0.28	6.18	0	0.59	0.72	0.23	5.97	0	0.92	0.25
Ret	0.57	0.92	0.11	5.61	19	0.39	0.64	0.11	5.17	13	0.97	0.32
Sub total	3.48	4.36	1.18			2.55	2.87	1.07				
<i>Oxy-PAHs</i>												
Flu-O	0.69	0.87	0.38	1.68	39	0.32	0.28	0.23	0.87	45	0.76	0.32
Phe-O	0.73	0.62	0.55	7.07	3	0.46	0.48	0.36	6.62	13	0.80	0.24
Xan	0.26	0.16	0.20	8.35	10	0.24	0.15	0.19	8.61	13	0.72	0.34
Nap-DO	0.49	0.30	0.43	7.46	0	0.44	0.29	0.38	7.35	0	0.59	0.25
Ant-DO	0.53	0.34	0.44	7.41	3	0.51	0.32	0.42	8.19	6	0.56	0.29
BaA-O	0.12	0.15	0.05	18.4	3	0.09	0.09	0.05	17.9	3	0.95	0.23
Subtotal	2.73	1.69	2.36			1.93	0.92	1.91				

(a) Standard deviation.

(b) Correlation coefficient.

(c) Coefficient of divergence.

Table 2.S3 Continued

Molecular marker (ng m <sup>-3</sup> )	Primary (N=31)					Duplicate (N=31)					<i>r</i>	COD
	Mean	SD	Median	S/N	BDL %	Mean	SD	Median	S/N	BDL %		
<i>Steranes</i>												
27-RS-C	0.30	0.38	0.15	12.3	0	0.33	0.41	0.15	12.1	0	0.86	0.21
28-RS-M	0.24	0.35	0.11	10.8	0	0.24	0.33	0.09	10.8	0	0.91	0.23
29-RS-E	0.24	0.33	0.11	16.7	0	0.25	0.30	0.09	16.2	0	0.90	0.21
TS	0.15	0.16	0.08	17.5	0	0.16	0.16	0.08	16.9	0	0.91	0.17
ba-N	0.65	0.89	0.33	13.8	0	0.65	0.84	0.26	13.6	0	0.94	0.18
ab-H	0.48	0.71	0.17	13.8	0	0.50	0.67	0.20	13.6	0	0.94	0.18
31abS	0.21	0.29	0.10	14.2	0	0.21	0.28	0.08	13.8	0	0.91	0.19
31abR	0.15	0.23	0.06	14.0	0	0.16	0.21	0.06	13.5	0	0.94	0.18
32abS	0.12	0.17	0.05	13.6	0	0.13	0.17	0.06	13.4	0	0.91	0.20
32abR	0.09	0.12	0.05	13.0	0	0.10	0.12	0.04	13.0	0	0.90	0.21
subtotal	2.64	3.62	1.28			2.71	3.46	1.05				
<i>n-alkanoic acids</i>												
C12:0	8.92	12.0	4.47	6.50	0	6.47	5.27	4.73	6.27	3	0.59	0.35
C13:0	1.43	1.30	0.94	3.10	35	1.14	1.07	0.64	2.65	55	0.28	0.38
C14:0	9.9	8.39	7.52	6.25	0	7.14	5.14	6.40	5.32	10	0.60	0.33
C15:0	2.26	1.70	1.81	6.01	0	1.43	0.84	1.34	4.67	10	0.53	0.34
C16:0	35.2	27.7	28.0	6.10	3	23.3	14.6	20.5	4.98	3	0.54	0.31
C17:0	1.81	1.70	1.40	2.98	3	1.09	0.71	0.91	2.78	10	0.49	0.34
C18:0	18.7	17.7	16.4	5.25	0	11.4	8.37	8.85	4.34	6	0.59	0.33
Subtotal	77.7	62.6	62.6			51.3	31.0	42.6				
<i>Sterols and methoxyphenols</i>												
Cho	0.84	0.99	0.42	3.62	48	0.92	1.16	0.27	3.75	61	0.68	0.46
Sti	1.02	1.11	0.43	2.43	32	1.06	1.10	0.30	2.57	35	0.71	0.30
Acv	1.67	2.23	0.95	3.98	48	1.35	1.49	0.82	4.07	45	0.93	0.38
Van	4.14	5.20	1.24	3.58	16	2.62	3.47	1.47	3.29	16	0.86	0.37
Syr	4.53	5.61	2.25	8.10	42	3.43	3.23	2.35	8.53	45	0.55	0.38
Con	4.52	5.08	1.89	8.68	29	3.36	3.46	2.25	8.59	39	0.76	0.44
Acs	1.17	1.49	0.52	3.02	45	0.78	0.92	0.28	2.22	52	0.89	0.43
Subtotal	13.7	17.7	3.50			9.74	12.4	3.20				
<i>Bulk species</i> (µg m <sup>-3</sup> )												
EC	0.69	0.42	0.53	7.88	0	0.68	0.41	0.56	7.80	0	0.92	0.14
OC	3.07	1.17	3.04	9.14	0	3.09	1.22	2.98	9.17	0	0.92	0.10

Table 2.S4 Simulated effect of quantification uncertainty on  $r$  and COD statistics for side-by-side and multi-site comparisons. Asterisks denote simulated values without quantification uncertainties.

Species	Side-by-side					PAL vs. EDI				
	$r$		COD		Average S/N <sup>a</sup>	$r$		COD		Average S/N <sup>b</sup>
	( $X_f^*$ , $X_h^*$ )	( $X_f$ , $X_h$ )	( $X_f^*$ , $X_h^*$ )	( $X_f$ , $X_h$ )		( $X_f^*$ , $X_h^*$ )	( $X_f$ , $X_h$ )	( $X_f^*$ , $X_h^*$ )	( $X_f$ , $X_h$ )	
C29	0.87	0.85	0.14	0.16	11.6	0.94	0.93	0.16	0.17	13.6
C32	N.A. <sup>c</sup>	N.A.	N.A.	N.A.	0.27	0.37	0.33	0.38	0.40	3.99
BbkF	0.94	0.93	0.18	0.19	10.1	0.53	0.53	0.40	0.41	15.1
IP	0.97	0.96	0.11	0.13	11.7	0.54	0.52	0.34	0.34	8.06
Nap-DO	0.63	0.60	0.23	0.25	7.40	0.79	0.79	0.28	0.28	18.8
Ant-DO	0.60	0.57	0.25	0.27	7.80	0.79	0.79	0.26	0.26	19.0
ba-N	0.95	0.94	0.13	0.14	13.7	0.69	0.68	0.28	0.29	18.4
ab-H	0.94	0.94	0.14	0.15	13.7	0.76	0.76	0.26	0.26	14.4
C12:0	0.62	0.59	0.32	0.34	6.39	0.81	0.80	0.29	0.30	10.4
C13:0	0.36	0.28	0.40	0.67	2.88	0.72	0.72	0.33	0.33	8.05
C16:0	0.58	0.54	0.32	0.34	5.54	0.72	0.70	0.34	0.35	7.67
C17:0 <sup>d</sup>	0.63	0.49	0.33	0.82	2.88	0.72	0.65	0.34	0.47	2.48
Cho	0.74	0.68	0.34	0.48	3.69	0.67	0.60	0.36	0.41	3.49
Van	0.95	0.86	0.21	0.30	3.44	0.76	0.75	0.33	0.33	15.0
EC <sup>d</sup>	0.96	0.92	0.08	0.12	7.84	0.83	0.63	0.17	0.31	2.19
OC	0.99	0.92	0.04	0.09	9.16	0.88	0.85	0.13	0.15	10.5

(a) Average signal to noise ratios of side-by-side samples.

(b) Average signal to noise ratios of the PAL and EDI sites.

(c) Not available, simulated  $r$  and COD without quantification uncertainties cannot be obtained due to the extremely low S/N ratio.

(d) Standard deviation of  $m_i$  (random number from a normal distribution with a mean of 0 and standard deviation of 1) was decreased to 0.7 to obtain simulated  $r$  and COD without quantification uncertainties between PAL and EDI.

Table 4.S1 Statistics for each species quantified during January 27 – December 31, 2003.

Chemical Species	No. of measurements	Mean	Median	CV <sup>a</sup>	S/N <sup>b</sup>	Missing +BDL (%) <sup>c</sup>
Bulk PM <sub>2.5</sub> species (µg m <sup>-3</sup> )						
Mass <sup>d</sup>	337	8.53	7.54	0.57	5.51	7
Nitrate	337	0.88	0.20	1.86	3.25	72
Sulfate	337	1.14	0.84	1.04	14.6	4
Ammonium <sup>d</sup>	337	0.55	0.36	1.18	20.4	0
EC	337	0.52	0.43	0.70	6.38	1
OC_PK1	337	1.28	1.22	0.44	4.90	0
OC_PK2	337	0.68	0.62	0.59	7.68	0
OC_PK3	337	0.47	0.45	0.42	4.63	0
OC_PK4 <sup>d</sup>	337	0.05	0.01	2.70	5.00	0
PC	337	0.48	0.40	0.81	2.69	0
Total OC <sup>d</sup>	337	2.98	2.75	0.49	10.3	0
Metals (ng m <sup>-3</sup> )						
B	337	2.55	1.89	0.92	7.98	15
Na	337	24.8	18.9	1.11	8.10	7
Mg	337	7.71	6.18	0.95	9.30	5
Al	337	2.89	2.42	0.91	9.73	3
P	337	3.02	2.56	0.75	7.15	6
K	337	25.9	17.1	1.79	12.1	3
Ca	337	68.9	49.1	1.07	4.64	23
Ti	337	0.10	0.03	2.57	3.08	70
V	337	0.05	0.03	1.11	8.58	4
Mn	337	0.75	0.57	0.82	14.2	0
Fe	337	4.28	3.28	0.88	14.8	1
Co	337	0.01	0.01	0.77	4.81	8
Cu	337	0.95	0.74	0.96	13.2	0
Zn	337	4.62	3.03	1.26	2.50	54
As	337	0.12	0.10	0.91	3.19	35
Rb	337	0.03	0.02	0.75	6.97	5
Sr	337	0.43	0.30	1.97	6.47	16
Y	337	0.003	0.002	0.75	3.33	30
Mo	337	0.06	0.05	0.81	4.53	14
Cd	337	0.03	0.03	0.89	8.19	2
Sb	337	0.22	0.18	0.87	16.4	1
Cs	337	0.003	0.002	1.61	7.04	9
Ba	337	2.39	1.51	1.43	17.4	1
La	337	0.005	0.003	1.09	9.18	4
Ce	337	0.006	0.005	0.93	9.37	3
Pr	337	0.001	0.0007	1.03	4.78	13
Sm	337	0.003	0.002	0.91	5.07	16
Eu	337	0.002	0.001	1.03	5.51	7
Tl	337	0.003	0.003	0.93	6.22	8
Pb	337	1.20	0.44	2.48	16.7	4

(a) Coefficient of variation (CV) = standard deviation/mean concentration.

(b) Signal to noise ratio (S/N) = mean concentration/mean uncertainty.

(c) Percent of missing measurements and observations not significantly different from zero using a p-value of 0.05. (d) Not used for source apportionment.

Table 2.S3 (continued)

Table 4.S1 Continued

Chemical Species	No. of measurements	Mean	Median	CV	S/N	Missing + BDL (%)
Organic molecular markers (ng m <sup>-3</sup> )						
Alkanes and cycloalkanes						
docosane (C22)	335	1.29	1.10	0.66	15.2	0
tricosane (C23)	335	1.90	1.55	0.67	10.4	0
tetracosane (C24)	334	1.08	0.92	0.68	7.13	1
pentacosane (C25)	335	1.56	1.30	0.66	8.68	1
hexacosane (C26)	335	0.93	0.66	0.82	6.09	10
heptacosane (C27)	335	1.27	1.04	0.71	7.21	7
octacosane (C28)	335	0.87	0.58	0.95	5.23	24
nonacosane (C29)	335	1.84	1.49	0.76	8.75	4
triacontane (C30)	335	0.71	0.45	1.04	5.03	26
hentriacontane (C31)	335	1.74	1.40	0.88	6.66	2
dotriacontane (C32) <sup>d</sup>	335	0.52	0.28	1.24	2.61	57
tritriacontane (C33)	335	0.90	0.72	0.73	7.98	4
tetatriacontane (C34)	335	0.68	0.52	0.84	7.93	5
pentatriacontane (C35)	335	0.51	0.36	0.85	7.31	4
hexatriacontane (C36)	335	0.29	0.18	0.99	6.16	21
heptatriacontane (C37)	335	0.21	0.12	1.17	4.99	36
octatriacontane (C38)	335	0.17	0.09	1.21	4.05	49
nonatriacontane (C39) <sup>d</sup>	335	0.16	0.09	1.11	3.81	51
tetracontane (C40) <sup>d</sup>	335	0.12	0.06	1.22	3.10	57
pentadecylcyclohexane (C21) <sup>d</sup>	331	0.15	0.09	1.31	8.13	9
nonadecylcyclohexane (C25) <sup>d</sup>	335	0.15	0.11	0.88	7.14	2
PAHs						
fluoranthene	336	0.17	0.10	0.98	6.46	2
pyrene	336	0.13	0.07	1.19	8.23	1
benzo[ghi]fluoranthene	336	0.09	0.05	1.13	12.9	0
cyclopenta[cd]pyrene	336	0.04	0.01	1.71	13.5	8
benz[a]anthracene	336	0.06	0.02	1.34	13.2	3
chrysene/triphenylene	336	0.15	0.11	0.90	11.4	0
benzo[b&k]fluoranthene	336	0.20	0.12	1.05	10.0	0
benzo[j]fluoranthene <sup>d</sup>	336	0.01	0.00	2.25	2.83	65
benz[a&e]pyrene	336	0.16	0.08	1.28	8.55	4
perylene <sup>d</sup>	336	0.01	0.00	1.85	3.13	64
indeno[1,2,3-cd]pyrene	336	0.08	0.04	1.12	9.66	7
benzo[ghi]perylene	336	0.20	0.12	1.12	8.91	7
dibenz[ah]anthracene <sup>d</sup>	336	0.02	0.01	1.48	5.07	48
picene <sup>d</sup>	336	0.01	0.00	1.57	4.74	49
coronene	336	0.10	0.06	1.10	5.69	12
methyl-202-PAH sum	336	0.52	0.25	1.21	7.41	4
retene	336	0.41	0.10	2.14	7.11	28
methyl-228-PAH sum <sup>d</sup>	336	0.10	0.07	0.99	7.51	1

Table 4.S1 Continued

Chemical Species	No. of measurements	Mean	Median	CV	S/N	Missing + BDL (%)
oxy-PAHs						
acenaphthenone <sup>d</sup>	336	0.09	0.03	1.85	4.82	49
fluorenone <sup>d</sup>	336	0.30	0.14	1.49	5.75	29
1H-phenalen-1-one <sup>d</sup>	336	0.23	0.14	1.31	7.79	6
xanthone <sup>d</sup>	336	0.17	0.16	0.60	8.26	8
1,8-naphthalic anhydride	336	0.35	0.31	0.53	8.08	0
anthracene-9,10-dione	336	0.35	0.33	0.40	8.01	0
benz[de]anthracene-7-one	336	0.07	0.04	0.99	11.4	1
steranes						
20R-abb & 20S-aaa-cholestane	336	0.13	0.11	0.59	10.7	0
20R & S-abb-methylcholestane	336	0.09	0.07	0.73	15.7	1
20R & S-abb-ethylcholestane	336	0.11	0.09	0.67	10.3	0
a-22,29,30-trisnorhopane	336	0.09	0.08	0.64	10.2	0
ba-30-norhopane	336	0.32	0.25	0.80	12.8	0
ab-hopane	336	0.21	0.16	0.78	10.1	0
22S-ab-30-homohopane	336	0.09	0.07	0.83	9.93	1
22R-ab-30-homohopane	336	0.07	0.05	0.87	8.61	3
22S-ab-30-bishomohopane	336	0.06	0.04	0.80	8.44	3
22R-ab-30-bishomohopane	336	0.05	0.04	0.80	7.96	6
alkanoic acids						
dodecanoic acid (C12:0)	332	3.14	2.46	0.71	7.53	6
tridecanoic acid (C13:0) <sup>d</sup>	332	0.21	0.00	1.37	1.42	80
tetradecanoic acid (C14:0)	335	4.38	3.00	0.91	7.40	2
pentadecanoic acid (C15:0) <sup>d</sup>	335	0.91	0.70	0.88	5.90	6
hexadecanoic acid (C16:0)	335	18.3	14.5	0.77	6.28	1
heptadecanoic acid (C17:0) <sup>d</sup>	335	0.61	0.47	0.86	2.14	38
octadecanoic acid (C18:0)	335	10.8	8.33	0.86	4.49	11
sterols and methoxyphenols						
cholesterol <sup>d</sup>	336	0.28	0.12	1.96	2.14	67
stigmasterol <sup>d</sup>	335	0.29	0.14	1.31	5.88	29
vanillin	335	2.46	0.85	2.03	5.49	39
acetovanillone	333	0.58	0.17	1.60	6.55	24
syringaldehyde	324	1.17	0.10	1.94	5.69	52
coniferaldehyde	336	1.16	0.20	1.91	6.32	45
acetosyringone <sup>d</sup>	336	0.29	0.00	2.33	5.62	61
Trace gas (ng m <sup>-3</sup> )						
CO	339	0.72	0.66	0.39	2.49	33

Table 5.S1 Statistics for PM<sub>2.5</sub> species quantified from January 27, 2003 – October 2, 2005 (Full).

	Abbreviation	Mean	Median	CV <sup>a</sup>	S/N <sup>b</sup>	Missing (%)	BDL <sup>c</sup> (%)
<b>Bulk species</b>							
(µg m <sup>-3</sup> )							
Mass*		8.05	7.03	0.61	3.89	0.71	17.5
Nitrate	Nitr	1.00	0.22	1.73	6.53	0.71	40.5
Sulfate	Sulf	1.18	0.97	0.88	13.3	0.71	1.43
Ammonium*	Ammo	0.56	0.43	1.04	14.2	56.5	0.10
EC		0.51	0.43	0.69	6.53	0.71	0.92
OC		2.85	2.64	0.45	9.42	0.71	0.10
<b>Organic molecular</b>							
makers (ng m <sup>-3</sup> )							
<i>Alkanes</i>							
docosane	C22	1.55	1.18	0.78	15.0	1.33	1.22
tricosane	C23	2.03	1.70	0.73	11.8	1.33	0.31
tetracosane	C24	1.19	0.97	0.73	5.54	1.43	10.6
pentacosane	C25	1.52	1.25	0.68	9.93	1.33	0.41
hexacosane	C26	0.86	0.65	0.78	9.00	1.33	4.38
heptacosane	C27	1.11	0.91	0.69	11.2	1.33	2.55
octacosane	C28	0.77	0.51	0.94	8.99	1.33	7.94
nonacosane	C29	1.68	1.36	0.76	10.7	1.33	1.32
triacontane	C30	0.64	0.39	1.02	6.55	1.33	13.9
hentriacontane	C31	2.07	1.32	1.50	3.18	1.33	15.2
dotriacontane*	C32	0.30	0.15	1.52	0.86	1.33	70.4
tritriacontane	C33	0.62	0.47	0.88	7.11	1.33	1.32
tetracontane	C34	0.45	0.31	1.00	7.56	1.33	4.28
pentatriacontane	C35	0.38	0.27	0.96	7.57	1.33	4.38
hexatriacontane	C36	0.20	0.12	1.18	5.77	1.33	17.7
heptatriacontane	C37	0.16	0.08	1.29	4.62	1.33	33.3
octatriacontane	C38	0.13	0.06	1.32	3.74	1.33	44.5
nonatriacontane	C39	0.13	0.07	1.22	3.60	1.33	40.6
tetracontane*	C40	0.09	0.05	1.33	2.81	1.33	53.8
pentadecylcyclohexane	cycC21	0.17	0.11	1.11	7.38	1.73	10.6
nonadecylcyclohexane*	cycC25	0.15	0.11	0.82	7.23	1.33	2.14
<i>PAHs</i>							
fluoranthene	Flu	0.19	0.13	0.98	5.90	1.33	3.77
pyrene	Pyr	0.14	0.07	1.20	6.77	1.33	2.44
benzo[ghi]fluoranthene	BghiF	0.09	0.05	1.17	10.3	1.33	0.81
cyclopenta[cd]pyrene	C-pyr	0.04	0.01	1.68	10.4	1.33	13.0
benz[a]anthracene	BaA	0.06	0.02	1.46	11.1	1.33	11.3
chrysene/triphenylene*	CT	0.20	0.12	1.11	14.9	1.33	0.00
benzo[b&k]fluoranthene	BbkF	0.24	0.13	1.21	9.77	1.33	0.20
benzo[j]fluoranthene*	BjF	0.01	0.00	5.20	1.91	1.33	72.7
benz[a&e]pyrene	BaeP	0.19	0.09	1.37	9.88	1.33	3.97
perylene*	Per	0.02	0.00	1.80	3.39	1.33	50.5
indeno[1,2,3-cd]pyrene	IP	0.07	0.04	1.18	11.2	1.33	3.36
benzo[ghi]perylene	BP	0.18	0.11	1.13	12.4	1.33	3.05
dibenz[ah]anthracene*	DahA	0.02	0.01	1.56	5.50	1.33	40.8
picene*	Pic	0.02	0.01	3.92	3.82	1.33	48.2
coronene	Cor	0.11	0.06	1.26	4.42	1.33	6.21
2-methylfluoranthene*	2M-Flu	0.21	0.10	1.25	6.33	1.33	2.44
methyl-202-PAH sum	M-202	0.62	0.27	1.32	6.39	1.33	1.63
retene	Ret	0.49	0.13	1.88	6.13	1.33	19.7
1-methylchrysene*	1M-Chr	0.03	0.01	1.50	4.98	1.33	43.2
methyl-228-PAH sum	M-228	0.14	0.07	1.28	10.7	1.33	0.61

Table 5.S1 Continued

	Abbreviation	Mean	Median	CV	S/N	Missing (%)	BDL (%)
<i>Oxy-PAHs</i>							
acenaphthenone*	Ace-O	0.26	0.04	6.67	2.12	1.73	63.0
fluorenone*	Flu-O	0.89	0.24	5.98	3.35	1.73	38.9
1H-phenalen-1-one*	Phe-O	0.63	0.27	2.79	7.37	1.73	8.15
xanthone*	Xan	0.21	0.18	0.70	8.01	1.33	5.60
1,8-naphthalic anhydride	Nap-DO	0.39	0.33	0.64	7.79	1.33	0.10
anthracene-9,10-dione	Ant-DO	0.46	0.37	0.70	7.85	1.33	0.92
benz[de]anthracene-7-one	BaA-O	0.08	0.04	1.13	13.5	1.33	1.63
<i>Steranes</i>							
20R-abb & 20S-aaa-cholestane	27-RS-C	0.18	0.13	1.02	12.0	1.43	0.41
20R & S-abb-methylcholestane	28-RS-M	0.13	0.08	1.29	12.2	1.43	0.71
20R & S-abb-ethylcholestane	29-RS-E	0.14	0.09	1.14	13.5	1.43	0.10
a-22,29,30-trisnorhopane	TS	0.10	0.08	0.91	14.0	1.43	0.71
ba-30-norhopane	ba-N	0.38	0.24	1.18	12.6	1.43	0.00
ab-hopane	ab-H	0.27	0.16	1.31	11.7	1.43	0.00
22S-ab-30-homohopane	31abS	0.12	0.07	1.31	11.9	1.43	0.51
22R-ab-30-homohopane	31abR	0.09	0.05	1.30	11.1	1.43	1.32
22S-ab-30-bishomohopane	32abS	0.07	0.05	1.24	11.0	1.43	1.32
22R-ab-30-bishomohopane	32abR	0.06	0.04	1.17	10.2	1.43	2.75
<i>Alkanoic acids</i>							
dodecanoic acid	C12:0	4.74	3.11	1.14	7.41	4.69	3.16
tridecanoic acid*	C13:0	0.47	0.21	1.53	2.14	4.69	56.2
tetradecanoic acid*	C14:0	6.44	4.73	0.90	6.30	2.45	1.02
pentadecanoic acid*	C15:0	1.48	1.11	0.87	5.88	2.45	2.95
hexadecanoic acid	C16:0	25.7	17.6	0.97	6.26	2.45	3.67
heptadecanoic acid*	C17:0	1.08	0.76	0.97	2.64	2.45	20.4
octadecanoic acid	C18:0	13.8	8.52	1.14	5.15	2.45	7.13
oleic acid*	C18:1	2.38	0.00	2.86	7.73	2.45	70.7
<i>Sterols and methoxyphenols</i>							
cholesterol*	Cho	0.36	0.15	1.73	2.54	1.33	62.7
stigmasterol*	Sti	0.43	0.20	1.59	3.13	1.43	33.1
acetovanillone	Acv	0.82	0.17	1.68	4.43	1.63	41.8
vanillin	Van	3.62	1.08	1.76	4.25	1.43	23.5
syringaldehyde	Syr	1.87	0.31	2.05	5.96	2.55	41.6
coniferaldehyde	Con	1.72	0.34	1.97	6.11	1.33	38.9
acetosyringone*	Ace	0.44	0.03	2.27	3.09	1.33	57.3

(a) Coefficient of variation = standard deviation/mean concentration.

(b) Signal to noise ratio = mean concentration/mean uncertainty.

(c) percent of observations not significantly different from zero using a *p*-value of 0.05.

\* Species not included for source apportionment.

Table 5.S2 Meteorological and trace gas statistics based on hourly observations for downtown Denver (January 27, 2003 – October 2, 2005).

Time period	Statistics	Temperature <sup>a</sup> (°C)	Irradiance <sup>a</sup> (kW-hr m <sup>-2</sup> )	Relative humidity <sup>a</sup> (%)	Ozone <sup>a</sup> (ppm)	NO <sub>x</sub> <sup>a</sup> (ppm)	CO <sup>a</sup> (ppm)
Full	Mean	13.5	4.71	46.9	0.046	0.090	0.72
	SD <sup>b</sup>	9.39	2.12	17.7	0.017	0.091	0.31
Cold	Mean	3.34	3.02	56.7	0.034	0.14	0.85
	SD	4.54	1.47	17.8	0.013	0.12	0.40
Warm	Mean	15.3	5.11	44.6	0.048	0.074	0.66
	SD	2.94	1.82	16.9	0.011	0.063	0.23
Hot	Mean	24.4	6.49	36.5	0.060	0.048	0.64
	SD	2.81	1.32	9.98	0.014	0.028	0.17

(a) 24-h average.

(b) Standard deviation.

Table 5.S3 Statistics for PM<sub>2.5</sub> species quantified on sampling days with ambient temperature lower than 10°C (Cold) from January 27, 2003 – October 2, 2005.

	Abbreviation.	Mean	Median	CV	S/N	Missing (%)	BDL (%)
<b>Bulk species</b>							
<b>(µg m<sup>-3</sup>)</b>							
Mass*		9.57	7.64	0.70	4.13	2.15	22.0
Nitrate	Nitr	2.23	1.41	1.04	12.2	2.15	15.1
Sulfate	Sulf	1.30	0.79	1.13	13.5	2.15	1.34
Ammonium*	Ammo	0.88	0.56	1.02	21.0	67.5	0.00
EC		0.61	0.48	0.75	7.20	2.15	1.34
OC		2.78	2.38	0.56	9.75	2.15	0.27
<b>Organic molecular makers (ng m<sup>-3</sup>)</b>							
<i>Alkanes</i>							
docosane	C22	1.90	1.56	0.75	17.1	2.15	0.00
tricosane	C23	1.72	1.36	0.71	12.2	2.15	0.00
tetracosane*	C24	1.24	1.12	0.62	4.83	2.15	15.3
pentacosane	C25	1.46	1.30	0.63	10.7	2.15	0.27
hexacosane	C26	1.19	1.01	0.67	11.4	2.15	2.15
heptacosane	C27	1.37	1.17	0.69	12.1	2.15	3.23
octacosane	C28	1.17	1.00	0.73	12.2	2.15	5.38
nonacosane	C29	1.41	1.21	0.69	12.0	2.15	1.34
triacontane	C30	0.99	0.89	0.77	8.94	2.15	12.4
hentriacontane	C31	1.43	1.23	0.76	8.51	2.15	3.49
dotriacontane*	C32	0.47	0.31	1.23	2.55	2.15	52.4
tritriacontane	C33	0.73	0.55	0.90	7.86	2.15	1.34
tetatriacontane	C34	0.58	0.43	0.91	8.49	2.15	2.15
pentatriacontane	C35	0.55	0.42	0.85	9.18	2.15	2.69
hexatriacontane	C36	0.33	0.22	0.95	7.55	2.15	10.8
heptatriacontane	C37	0.27	0.18	0.97	6.41	2.15	19.1
octatriacontane*	C38	0.22	0.15	0.98	5.44	2.15	24.7
nonatriacontane*	C39	0.21	0.14	0.95	4.94	2.15	24.2
tetracontane*	C40	0.15	0.10	1.01	3.86	2.15	32.5
pentadecylcyclohexane*	cycC21	0.24	0.18	1.00	6.91	2.15	6.18
nonadecylcyclohexane*	cycC25	0.20	0.17	0.75	7.72	2.15	0.54
<i>PAHs</i>							
fluoranthene	Flu	0.28	0.22	0.82	5.79	2.15	4.57
pyrene	Pyr	0.25	0.19	0.89	7.02	2.15	2.42
benzo[ghi]fluoranthene	BghiF	0.17	0.15	0.77	11.3	2.15	0.00
cyclopenta[cd]pyrene	C-pyr	0.08	0.05	1.11	13.2	2.15	2.69
benz[a]anthracene	BaA	0.12	0.09	0.97	13.3	2.15	1.88
chrysene/triphenylene*	CT	0.33	0.26	0.84	16.1	2.15	0.00
benzo[b&k]fluoranthene	BbkF	0.41	0.30	0.79	11.6	2.15	0.00
benzo[j]fluoranthene*	BjF	0.02	0.00	3.74	3.98	2.15	43.0
benz[a&e]pyrene	BaeP	0.36	0.24	0.94	11.8	2.15	0.27
perylene*	Per	0.03	0.02	1.18	5.16	2.15	28.8
indeno[1,2,3-cd]pyrene	IP	0.13	0.11	0.76	13.0	2.15	0.00
benzo[ghi]perylene	BP	0.32	0.25	0.80	15.3	2.15	0.00
dibenz[ah]anthracene*	DahA	0.04	0.03	0.92	9.48	2.15	7.53
picene*	Pic	0.03	0.02	0.95	7.75	2.15	15.9
coronene	Cor	0.18	0.13	0.91	4.22	2.15	1.88
2-methylfluoranthene*	2M-Flu	0.39	0.32	0.84	6.64	2.15	0.00
methyl-202-PAH sum	M-202	1.17	0.89	0.88	6.64	2.15	0.27
retene	Ret	1.07	0.71	1.16	6.63	2.15	1.34
1-methylchrysene*	1M-Chr	0.05	0.04	0.99	7.41	2.15	21.5
methyl-228-PAH sum*	M-228	0.24	0.16	0.95	12.1	2.15	0.27

Table 5.S3 Continued

	Abbreviation.	Mean	Median	CV	S/N	Missing (%)	BDL (%)
<i>Oxy-PAHs</i>							
acenaphthenone*	Ace-O	0.64	0.20	4.36	3.72	2.42	40.1
fluorenone*	Flu-O	1.99	0.60	4.32	5.18	2.42	21.0
1H-phenalen-1-one*	Phe-O	1.06	0.49	2.57	7.66	2.42	2.96
xanthone*	Xan	0.17	0.14	0.77	7.50	2.15	10.2
1,8-naphthalic anhydride	Nap-DO	0.37	0.31	0.67	7.46	2.15	0.00
anthracene-9,10-dione	Ant-DO	0.38	0.34	0.56	7.51	2.15	2.42
benz[de]anthracene-7-one	BaA-O	0.15	0.13	0.74	15.3	2.15	0.27
<i>Steranes</i>							
20R-abb & 20S-aaa-cholestane	27-RS-C	0.24	0.14	1.08	13.6	2.42	0.00
20R & S-abb-methylcholestane	28-RS-M	0.20	0.11	1.18	11.6	2.42	0.54
20R & S-abb-ethylcholestane	29-RS-E	0.22	0.14	1.02	16.9	2.42	0.27
a-22,29,30-trisnorhopane	TS	0.13	0.09	0.98	17.2	2.42	0.27
ba-30-norhopane	ba-N	0.59	0.37	1.07	14.4	2.42	0.00
ab-hopane	ab-H	0.43	0.27	1.12	13.7	2.42	0.00
22S-ab-30-homohopane	31abS	0.19	0.12	1.11	14.0	2.42	0.27
22R-ab-30-homohopane	31abR	0.14	0.09	1.12	13.8	2.42	0.81
22S-ab-30-bishomohopane	32abS	0.11	0.07	1.12	13.6	2.42	0.81
22R-ab-30-bishomohopane	32abR	0.09	0.05	1.05	13.1	2.42	1.88
<i>Alkanoic acids</i>							
dodecanoic acid	C12:0	3.97	2.26	1.25	8.92	3.23	6.72
tridecanoic acid*	C13:0	0.24	0.00	1.83	1.25	3.23	76.6
tetradecanoic acid*	C14:0	4.60	2.74	0.92	4.72	2.69	2.15
pentadecanoic acid*	C15:0	1.16	0.78	0.83	4.74	2.69	3.76
hexadecanoic acid	C16:0	25.8	20.1	0.89	7.95	2.69	0.54
heptadecanoic acid*	C17:0	1.09	0.67	0.97	2.51	2.69	26.9
octadecanoic acid	C18:0	15.1	10.1	0.96	6.02	2.69	5.11
oleic acid*	C18:1	5.28	0.57	1.89	10.6	2.69	46.5
<i>Sterols and methoxyphenols</i>							
cholesterol*	Cho	0.58	0.33	1.42	4.00	2.15	43.8
stigmasterol*	Sti	0.77	0.49	1.12	9.05	2.42	11.8
acetovanillone	Acv	1.77	1.31	0.95	6.98	2.15	2.69
vanillin	Van	7.69	5.34	1.07	4.59	2.15	2.69
syringaldehyde	Syr	4.01	2.52	1.24	6.83	2.15	6.18
coniferaldehyde	Con	3.84	2.45	1.16	6.81	2.15	7.53
acetosyringone*	Ace	1.06	0.60	1.32	7.91	2.15	16.4

\* Species not included for source apportionment.

Table 5.S4 Statistics for PM<sub>2.5</sub> species quantified on sampling days with ambient temperature between 10°C and 20°C (Warm) from January 27, 2003 – October 2, 2005.

	Abbreviation	Mean	Median	CV	S/N	Missing (%)	BDL (%)
<b>Bulk species</b>							
(µg m <sup>-3</sup> )							
Mass*		6.54	6.18	0.44	3.26	0.31	20.4
Nitrate	Nitr	0.37	0.17	1.58	2.56	0.31	53.6
Sulfate	Sulf	1.01	0.90	0.67	12.4	0.31	1.88
Ammonium*	Ammo	0.32	0.23	0.96	14.0	58.9	0.00
EC		0.43	0.36	0.62	6.35	0.31	0.94
OC		2.39	2.32	0.36	8.78	0.31	0.00
<b>Organic molecular makers (ng m<sup>-3</sup>)</b>							
<i>Alkanes</i>							
docosane	C22	1.24	1.06	0.73	13.5	0.94	3.45
tricosane	C23	1.60	1.31	0.70	11.4	0.94	0.94
tetracosane*	C24	0.80	0.67	0.67	4.27	1.25	14.4
pentacosane	C25	1.04	0.88	0.62	9.02	0.94	0.94
hexacosane	C26	0.65	0.48	0.84	7.28	0.94	9.72
heptacosane	C27	0.90	0.74	0.70	9.21	0.94	4.08
octacosane	C28	0.60	0.40	0.96	6.83	0.94	15.2
nonacosane	C29	1.47	1.18	0.74	9.67	0.94	2.51
triacontane*	C30	0.50	0.29	1.03	5.02	0.94	23.5
hentriacontane*	C31	1.69	1.12	1.10	3.52	0.94	13.3
dotriacontane*	C32	0.24	0.11	1.52	0.83	0.94	78.7
tritriacontane	C33	0.52	0.41	0.84	6.35	0.94	2.19
tettriacontane	C34	0.37	0.28	0.97	6.46	0.94	5.33
pentatriacontane	C35	0.29	0.23	0.85	5.85	0.94	5.33
hexatriacontane*	C36	0.15	0.10	1.06	4.38	0.94	22.6
heptatriacontane*	C37	0.11	0.06	1.19	3.43	0.94	39.8
octatriacontane*	C38	0.09	0.05	1.18	2.81	0.94	49.8
nonatriacontane*	C39	0.09	0.06	1.13	2.86	0.94	46.7
tetracontane*	C40	0.06	0.04	1.20	2.22	0.94	61.1
pentadecylcyclohexane*	cycC21	0.15	0.09	1.05	7.93	1.88	11.3
nonadecylcyclohexane*	cycC25	0.11	0.08	0.75	7.03	0.94	2.82
<i>PAHs</i>							
fluoranthene	Flu	0.14	0.10	0.81	5.90	0.94	5.02
pyrene	Pyr	0.08	0.06	0.91	6.37	0.94	3.45
benzo[ghi]fluoranthene	BghiF	0.06	0.04	0.89	9.41	0.94	0.63
cyclopenta[cd]pyrene	C-pyr	0.02	0.01	1.49	7.41	0.94	11.0
benz[a]anthracene	BaA	0.03	0.02	1.45	9.02	0.94	11.9
chrysene/triphenylene*	CT	0.14	0.11	0.80	13.2	0.94	0.00
benzo[b&k]fluoranthene	BbkF	0.16	0.11	1.08	8.45	0.94	0.63
benzo[j]fluoranthene*	BjF	0.00	0.00	2.29	0.69	0.94	86.5
benz[a&e]pyrene	BaeP	0.13	0.08	1.28	8.25	0.94	4.70
perylene*	Per	0.01	0.00	1.72	2.37	0.94	59.9
indeno[1,2,3-cd]pyrene	IP	0.05	0.04	0.84	9.45	0.94	5.33
benzo[ghi]perylene	BP	0.13	0.09	0.82	9.64	0.94	5.64
dibenz[ah]anthracene*	DahA	0.01	0.01	1.34	3.50	0.94	44.2
picene*	Pic	0.01	0.00	1.50	2.52	0.94	57.4
coronene	Cor	0.07	0.05	0.93	4.84	0.94	8.46
2-methylfluoranthene*	2M-Flu	0.13	0.09	1.03	5.93	0.94	3.76
methyl-202-PAH sum	M-202	0.36	0.22	1.12	6.08	0.94	2.19
retene	Ret	0.21	0.09	1.94	5.12	0.94	21.6
1-methylchrysene*	1M-Chr	0.02	0.01	1.29	3.15	0.94	52.4
methyl-228-PAH sum*	M-228	0.09	0.06	1.18	9.49	0.94	0.94

Table 5.S4 Continued

	Abbreviation.	Mean	Median	CV	S/N	Missing (%)	BDL (%)
<i>Oxy-PAHs</i>							
acenaphthenone*	Ace-O	0.06	0.03	2.41	0.57	1.25	65.8
fluorenone*	Flu-O	0.33	0.21	1.46	1.47	1.25	40.1
1H-phenalen-1-one*	Phe-O	0.40	0.21	1.16	7.00	1.25	7.52
xanthone*	Xan	0.22	0.19	0.67	8.24	0.94	0.31
1,8-naphthalic anhydride	Nap-DO	0.33	0.29	0.58	8.28	0.94	0.31
anthracene-9,10-dione	Ant-DO	0.43	0.35	0.68	8.22	0.94	0.00
benz[de]anthracene-7-one	BaA-O	0.06	0.04	1.08	11.7	0.94	0.31
<i>Steranes</i>							
20R-abb & 20S-aaa-cholestane	27-RS-C	0.13	0.10	0.76	11.6	0.94	1.25
20R & S-abb-methylcholestane	28-RS-M	0.09	0.07	0.97	12.6	0.94	0.63
20R & S-abb-ethylcholestane	29-RS-E	0.10	0.08	0.99	11.8	0.94	0.00
a-22,29,30-trisnorhopane	TS	0.08	0.06	0.74	12.9	0.94	1.25
ba-30-norhopane	ba-N	0.27	0.21	0.98	11.3	0.94	0.00
ab-hopane	ab-H	0.19	0.14	1.17	9.99	0.94	0.00
22S-ab-30-homohopane	31abS	0.08	0.06	1.23	10.3	0.94	0.94
22R-ab-30-homohopane	31abR	0.06	0.04	1.12	9.30	0.94	1.25
22S-ab-30-bishomohopane	32abS	0.05	0.04	1.08	9.22	0.94	1.25
22R-ab-30-bishomohopane	32abR	0.04	0.03	1.08	8.48	0.94	2.82
<i>Alkanoic acids</i>							
dodecanoic acid	C12:0	4.67	2.92	1.30	7.02	3.76	1.57
tridecanoic acid*	C13:0	0.59	0.43	1.46	2.49	3.76	48.6
tetradecanoic acid*	C14:0	5.50	3.74	0.95	6.01	2.19	0.63
pentadecanoic acid*	C15:0	1.28	0.94	0.91	5.73	2.19	2.82
hexadecanoic acid	C16:0	23.0	13.3	1.20	6.12	2.19	4.39
heptadecanoic acid*	C17:0	0.80	0.54	1.09	2.66	2.19	28.8
octadecanoic acid	C18:0	13.5	7.17	1.49	4.65	2.19	7.84
oleic acid*	C18:1	0.91	0.00	3.29	4.81	2.19	82.1
<i>Sterols and methoxyphenols</i>							
cholesterol*	Cho	0.25	0.10	1.60	1.91	0.94	69.3
stigmasterol*	Sti	0.27	0.17	1.54	2.33	0.94	36.1
acetovanillone	Acv	0.45	0.12	2.04	3.32	1.88	40.1
vanillin	Van	1.77	0.84	2.07	3.84	0.94	20.7
syringaldehyde	Syr	0.85	0.07	2.98	4.82	2.82	53.3
coniferaldehyde	Con	0.69	0.22	2.87	4.99	0.94	46.4
acetosyringone*	Ace	0.11	0.00	3.14	1.17	0.94	75.6

\* Species not included for source apportionment.

Table 5.S5 Statistics for PM<sub>2.5</sub> species quantified on sampling days with ambient temperature higher than 20°C (Hot) from January 27, 2003 – October 2, 2005.

	Abbreviation	Mean	Median	CV	S/N	Missing (%)	BDL (%)
<b>Bulk species</b>							
<b>(µg m<sup>-3</sup>)</b>							
Mass*		7.79	7.53	0.38	4.28	0.35	8.65
Nitrate*	Nitr	0.16	0.14	0.76	1.27	0.35	59.2
Sulfate	Sulf	1.22	1.11	0.46	14.0	0.35	1.04
Ammonium*	Ammo	0.51	0.48	0.54	10.3	39.8	0.35
EC		0.46	0.44	0.42	5.81	0.35	0.35
OC		3.45	3.35	0.31	9.62	0.35	0.00
<b>Organic molecular makers (ng m<sup>-3</sup>)</b>							
<i>Alkanes</i>							
docosane	C22	1.44	1.15	0.75	13.6	0.69	0.35
tricosane	C23	2.91	2.51	0.61	11.7	0.69	0.00
tetracosane*	C24	1.56	1.22	0.70	8.09	0.69	0.35
pentacosane	C25	2.11	1.70	0.57	9.86	0.69	0.00
hexacosane	C26	0.68	0.58	0.63	7.39	0.69	1.38
heptacosane	C27	1.03	0.92	0.51	12.1	0.69	0.00
octacosane	C28	0.46	0.35	0.86	6.42	0.69	3.46
nonacosane	C29	2.26	1.88	0.70	10.6	0.69	0.00
triacontane*	C30	0.36	0.24	1.13	4.45	0.69	5.19
hentriacontane*	C31	3.29	1.72	1.52	2.26	0.69	32.5
dotriacontane*	C32	0.14	0.05	1.66	0.23	0.69	84.8
tritriacontane	C33	0.59	0.48	0.78	6.88	0.69	0.35
tetratriacontane	C34	0.37	0.24	1.06	7.31	0.69	5.88
pentatriacontane	C35	0.26	0.23	0.78	6.81	0.69	5.54
hexatriacontane*	C36	0.11	0.09	1.09	4.09	0.69	21.5
heptatriacontane*	C37	0.08	0.05	1.38	2.69	0.69	44.6
octatriacontane*	C38	0.05	0.03	1.50	1.85	0.69	64.4
nonatriacontane*	C39	0.06	0.05	1.22	2.08	0.69	55.4
tetracontane*	C40	0.04	0.03	1.56	1.54	0.69	73.4
pentadecylcyclohexane*	cycC21	0.13	0.08	1.15	7.99	1.04	15.6
nonadecylcyclohexane*	cycC25	0.11	0.10	0.55	6.48	0.69	3.46
<i>PAHs</i>							
fluoranthene	Flu	0.12	0.09	0.92	6.27	0.69	1.38
pyrene	Pyr	0.07	0.05	1.06	6.28	0.69	1.38
benzo[ghi]fluoranthene	BghiF	0.03	0.02	0.89	7.01	0.69	2.08
cyclopenta[cd]pyrene*	C-pyr	0.01	0.00	2.63	4.67	0.69	28.7
benz[a]anthracene	BaA	0.02	0.01	1.69	6.09	0.69	22.8
chrysene/triphenylene*	CT	0.11	0.08	1.51	13.6	0.69	0.00
benzo[b&k]fluoranthene	BbkF	0.10	0.06	2.20	6.34	0.69	0.00
benzo[j]fluoranthene*	BjF	0.00	0.00	3.75	0.35	0.69	96.2
benz[a&e]pyrene	BaeP	0.06	0.04	1.87	5.63	0.69	7.96
perylene*	Per	0.01	0.00	3.94	1.33	0.69	68.5
indeno[1,2,3-cd]pyrene	IP	0.03	0.02	2.13	7.25	0.69	5.54
benzo[ghi]perylene	BP	0.07	0.05	1.44	8.34	0.69	4.15
dibenz[ah]anthracene*	DahA	0.01	0.00	4.69	1.43	0.69	80.3
picene*	Pic	0.01	0.00	10.5	1.67	0.69	79.9
coronene	Cor	0.04	0.03	1.95	4.90	0.69	9.34
2-methylfluoranthene*	2M-Flu	0.07	0.05	1.29	5.29	0.69	4.15
methyl-202-PAH sum	M-202	0.20	0.14	1.12	5.46	0.69	2.77
retene	Ret	0.06	0.03	1.44	3.12	0.69	41.2
1-methylchrysene*	1M-Chr	0.01	0.00	2.54	1.81	0.69	61.3
methyl-228-PAH sum*	M-228	0.06	0.04	1.17	8.00	0.69	0.69

Table 5.S5 Continued

	Abbreviation	Mean	Median	CV	S/N	Missing (%)	BDL (%)
<i>Oxy-PAHs</i>							
acenaphthenone*	Ace-O	NA <sup>a</sup>	NA	NA	NA	NA	NA
fluorenone*	Flu-O	0.12	0.05	3.24	0.75	1.38	60.9
1H-phenalen-1-one*	Phe-O	0.32	0.14	1.44	6.75	1.38	15.6
xanthone*	Xan	0.25	0.23	0.64	8.28	0.69	5.54
1,8-naphthalic anhydride*	Nap-DO	0.47	0.41	0.60	7.77	0.69	0.00
anthracene-9,10-dione	Ant-DO	0.60	0.47	0.70	7.87	0.69	0.00
benz[de]anthracene-7-one*	BaA-O	0.03	0.02	0.83	9.64	0.69	4.84
<i>Steranes</i>							
20R-abb & 20S-aaa-cholestane	27-RS-C	0.15	0.14	0.41	9.98	0.69	0.00
20R & S-abb-methylcholestane	28-RS-M	0.08	0.07	0.44	13.6	0.69	1.04
20R & S-abb-ethylcholestane	29-RS-E	0.08	0.08	0.44	9.20	0.69	0.00
a-22,29,30-trisnorhopane	TS	0.09	0.08	0.41	11.0	0.69	0.69
ba-30-norhopane	ba-N	0.24	0.22	0.42	10.1	0.69	0.00
ab-hopane	ab-H	0.14	0.14	0.43	8.87	0.69	0.00
22S-ab-30-homohopane	31abS	0.06	0.06	0.46	8.93	0.69	0.35
22R-ab-30-homohopane	31abR	0.05	0.04	0.50	7.73	0.69	2.08
22S-ab-30-bishomohopane	32abS	0.04	0.04	0.49	7.90	0.69	2.08
22R-ab-30-bishomohopane	32abR	0.04	0.03	0.53	7.36	0.69	3.81
<i>Alkanoic acids</i>							
dodecanoic acid*	C12:0	5.87	4.62	0.84	6.72	7.61	0.35
tridecanoic acid*	C13:0	0.64	0.47	1.20	2.73	7.61	38.8
tetradecanoic acid*	C14:0	9.86	7.95	0.67	8.17	2.42	0.00
pentadecanoic acid*	C15:0	2.11	1.60	0.73	7.24	2.42	2.08
hexadecanoic acid*	C16:0	28.66	19.48	0.82	5.10	2.42	6.92
heptadecanoic acid*	C17:0	1.38	1.02	0.81	2.78	2.42	2.77
octadecanoic acid*	C18:0	12.38	8.65	0.89	4.71	2.42	9.00
oleic acid*	C18:1	0.28	0.00	5.33	1.44	2.42	89.6
<i>Sterols and methoxyphenols</i>							
cholesterol*	Cho	0.18	0.09	1.91	1.28	0.69	80.3
stigmasterol*	Sti	0.17	0.10	2.58	0.76	0.69	57.4
acetovanillone*	Acv	0.03	0.00	7.51	0.18	0.69	94.1
vanillin*	Van	0.48	0.26	1.76	2.05	1.04	53.6
syringaldehyde*	Syr	0.23	0.00	4.81	2.08	2.77	74.4
coniferaldehyde*	Con	0.16	0.00	2.09	2.02	0.69	71.3
acetosyringone*	Ace	0.03	0.00	3.52	0.13	0.69	90.3

(a) Not available.

\* Species not included for source apportionment.

Table 5.S6 Median coefficient of variation (CV) of factor contributions for bootstrapped PMF solutions for each data set.

Factor	Data sets			
	Full	Cold	Warm	Hot
Inorganic ion or Sulfate	0.44	0.12	0.14	0.20
<i>n</i> -Alkane	0.22	0.18	0.18	0.28
EC/sterane	0.39	0.33	0.31	0.18
Light <i>n</i> -alkane/PAH	0.38	0.36	0.25	0.40
Medium alkane/alkanoic acid	1.35	0.49	0.51	/
PAH	0.27	0.24	0.35	0.74
Winter/methoxyphenol	0.90	0.48	1.05	/
Summer/odd <i>n</i> -alkane	0.54	/	/	/
Medium <i>n</i> -alkane/OC	/	/	/	0.35

CV = standard deviation/median factor contribution.

Table 6.S1 Statistics for the total (gas + particle phase) concentration of each SVOC estimated from January 27, 2003 – October 2, 2005 (Full data set).

	Abbreviation	Mean	Median	Mean fraction (%) <sup>a</sup>	S/N <sup>b</sup>	CV <sup>c</sup>	$\rho^{\circ}L$ (atm, 298K) <sup>d</sup>	$\Delta H_{\text{vap}}$ (kJ mol <sup>-1</sup> , 298K) <sup>e</sup>	$K_{\text{p,om}}$ (m <sup>3</sup> $\mu\text{g}^{-1}$ ) <sup>f</sup>
<b>Bulk species (<math>\mu\text{g m}^{-3}</math>)<sup>g</sup></b>									
Mass*		8.05	7.03		3.89	0.61			
Nitrate	Nitr	1.00	0.22		6.53	1.73			
Sulfate	Sulf	1.18	0.97		13.3	0.88			
Ammonium*	Ammo	0.56	0.43		14.2	1.04			
EC		0.51	0.43		6.53	0.69			
OC1*		1.28	1.20		4.58	0.43			
OC2		0.65	0.59		4.32	0.50			
OC3		0.43	0.41		2.39	0.43			
OC4*		0.03	0.01		1.00	2.30			
PC		0.38	0.28		2.46	0.94			
<b>Organic molecular makers (ng m<sup>-3</sup>)</b>									
<i>Alkanes</i>									
docosane	C22	32.8	15.1	16.1	1.46	2.48	3.2E-08	115	8.73E-02
tricosane	C23	23.7	7.40	28.6	1.60	2.38	1.2E-08	120	2.76E-01
tetracosane	C24	5.50	2.27	44.3	1.63	1.83	4.6E-09	124	8.72E-01
pentacosane	C25	3.73	2.02	61.3	1.36	2.21	1.7E-09	129	2.77E+00
hexacosane	C26	1.14	0.94	76.9	0.77	2.14	6.5E-10	133	8.83E+00
heptacosane	C27	1.27	1.09	88.3	0.67	2.48	2.5E-10	137	2.82E+01
octacosane	C28	0.79	0.54	94.8	0.93	2.28	9.3E-11	142	9.05E+01
nonacosane	C29	1.72	1.40	97.9	0.77	2.37	3.5E-11	146	2.91E+02
triacontane	C30	0.64	0.39	99.2	1.04	1.89	1.3E-11	151	9.39E+02
hentriacontane	C31	1.78	1.30	99.7	1.17	1.10	4.9E-12	155	3.04E+03
dotriacontane*	C32	0.51	0.23	99.9	1.61	0.70	1.9E-12	160	9.86E+03
tritriacontane	C33	0.62	0.47	100	0.88	1.95	7.0E-13	164	3.21E+04
tetraatriacontane	C34	0.45	0.31	100	1.01	2.01	2.6E-13	169	1.05E+05
pentatriacontane	C35	0.38	0.27	100	0.96	2.02	9.9E-14	173	3.42E+05
hexatriacontane	C36	0.20	0.12	100	1.19	1.75	3.7E-14	177	1.12E+06
heptatriacontane	C37	0.16	0.08	100	1.30	1.54	1.4E-14	182	3.69E+06
octatriacontane	C38	0.13	0.05	100	1.32	1.36	5.3E-15	186	1.22E+07
nonatriacontane*	C39	0.12	0.06	100	1.23	1.32	2.0E-15	191	4.02E+07
tetracontane*	C40	0.09	0.04	100	1.29	1.14	7.5E-16	195	1.33E+08
pentadecylcyclohexane	cycC21	8.07	3.62	7.30	1.67	2.01	8.2E-08	107	2.50E-02
nonadecylcyclohexane*	cycC25	0.26	0.21	61.3	0.77	1.86	1.6E-09	124	2.47E+00
<i>PAHs</i>									
fluoranthene	Flu	11.2	7.33	4.12	1.16	1.74	1.1E-07	92.2	1.14E-02
pyrene	Pyr	1.48	1.08	14.7	1.04	1.78	2.4E-08	98.9	6.40E-02
benzo[ghi]fluoranthene	BghiF	0.19	0.15	44.3	0.74	2.16	3.3E-09	104	5.64E-01
cyclopenta[cd]pyrene	C-pyr	0.06	0.03	48.5	1.38	2.10	2.5E-09	103	6.94E-01
benz[a]anthracene	BaA	0.12	0.08	44.6	1.16	2.12	3.4E-09	108	6.24E-01
chrysene/triphenylene*	CT	0.48	0.38	44.6	1.06	2.65	3.4E-09	108	6.24E-01
benzo[b&k]fluoranthene	BbkF	0.27	0.16	79.3	1.15	2.18	4.6E-10	113	5.54E+00
benzo[j]fluoranthene*	BjF	0.01	0.01	79.3	3.62	0.91	4.6E-10	113	5.54E+00
benz[a&e]pyrene	BaeP	0.20	0.09	93.7	1.35	2.25	1.0E-10	119	3.18E+01
perylene*	Per	0.02	0.01	93.7	1.66	1.28	1.0E-10	119	3.18E+01
indeno[1,2,3-cd]pyrene	IP	0.07	0.04	99.1	1.17	2.42	1.4E-11	124	2.85E+02
benzo[ghi]perylene	BP	0.18	0.11	99.8	1.14	2.60	3.2E-12	132	1.65E+03
dibenz[ah]anthracene*	DahA	0.02	0.01	99.0	1.47	1.76	1.5E-11	128	3.19E+02
picene*	Pic	0.02	0.01	99.0	3.70	1.45	1.5E-11	128	3.19E+02
coronene	Cor	0.11	0.06	100	1.27	1.53	9.7E-14	142	8.79E+04
methyl-202-PAH sum	M-202	8.70	6.67	9.90	0.93	1.70	4.0E-08	96.6	3.54E-02
retene*	Ret	1.98	1.17	20.9	1.40	1.54	1.6E-08	105	1.22E-01
methyl-228-PAH sum	M-228	0.19	0.13	63.0	0.99	2.26	1.3E-09	112	1.96E+00

Table 6.S1 Continued

	Abbreviation	Mean	Median	Mean fraction (%)	S/N	CV	$p^{\circ}L$ (atm) 298K	$\Delta H_{\text{vap}}$ (kJ mol <sup>-1</sup> ) 298K	$K_{p, \text{om}}$ (m <sup>3</sup> μg <sup>-1</sup> )
<i>Oxy-PAHs</i>									
acenaphthenone*	Ace-O	440	146	0.09	2.15	0.73	2.9E-06	71.4	2.17E-04
fluorenone*	Flu-O	467	226	0.29	2.23	0.88	1.1E-06	75.8	6.61E-04
1H-phenalen-1-one*	Phe-O	303	145	0.35	1.65	1.48	8.4E-07	74.7	8.19E-04
xanthone*	Xan	44.9	26.2	1.52	1.17	2.03	2.3E-07	81.8	3.70E-03
1,8-naphthalic anhydride	Nap-DO	64.7	41.7	1.61	1.11	2.00	1.9E-07	76.9	3.92E-03
anthracene-9,10-dione	Ant-DO	21.1	11.4	6.04	1.25	2.01	4.7E-08	79.9	1.73E-02
benz[de]anthracene-7-one	BaA-O	0.22	0.18	35.8	0.71	2.43	4.7E-09	96.2	3.02E-01
<i>Steranes</i>									
20R-abb & 20S-aaa-cholestane	27-RS-C	0.19	0.15	89.3	0.93	2.44	2.0E-10	121	1.77E+01
20R & S-abb-methylcholestane	28-RS-M	0.13	0.08	95.4	1.25	2.47	7.5E-11	125	5.59E+01
20R & S-abb-ethylcholestane	29-RS-E	0.14	0.09	98.2	1.13	2.62	2.8E-11	130	1.78E+02
a-22,29,30-trisnorhopane	TS	0.11	0.09	89.5	0.84	2.64	1.9E-10	117	1.58E+01
ba-30-norhopane	ba-N	0.38	0.25	98.3	1.16	2.50	2.7E-11	126	1.58E+02
ab-hopane	ab-H	0.27	0.16	99.3	1.30	2.42	1.0E-11	130	5.03E+02
22S-ab-30-homohopane	31abS	0.12	0.07	99.8	1.31	2.45	3.8E-12	134	1.61E+03
22R-ab-30-homohopane	31abR	0.09	0.05	99.8	1.30	2.39	3.8E-12	134	1.61E+03
22S-ab-30-bishomohopane	32abS	0.07	0.05	99.9	1.24	2.37	1.4E-12	139	5.14E+03
22R-ab-30-bishomohopane	32abR	0.06	0.04	99.9	1.16	2.31	1.4E-12	139	5.14E+03
<i>Alkanoic acids</i>									
dodecanoic acid	C12:0	710	336	2.81	1.63	1.70	1.8E-07	93.6	7.36E-03
tridecanoic acid*	C13:0	30.9	6.76	7.08	1.89	0.86	6.6E-08	98.1	2.29E-02
tetradecanoic acid*	C14:0	160	60.5	15.4	1.60	1.86	2.5E-08	103	7.12E-02
pentadecanoic acid*	C15:0	14.0	5.57	28.5	1.63	1.43	9.3E-09	107	2.23E-01
hexadecanoic acid	C16:0	90.6	44.6	45.3	1.44	1.60	3.5E-09	111	6.98E-01
heptadecanoic acid*	C17:0	2.18	1.35	63.3	1.33	0.71	1.3E-09	116	2.20E+00
octadecanoic acid	C18:0	17.6	11.5	79.2	1.13	1.55	5.0E-10	120	6.94E+00
oleic acid*	C18:1	2.59	0.18	82.3	2.67	1.91	3.9E-10	119	8.51E+00
<i>Sterols and methoxyphenols</i>									
cholesterol*	Cho	0.39	0.17	99.9	1.51	1.13	1.0E-12	136	6.52E+03
stigmasterol*	Sti	0.49	0.22	100	1.43	1.35	1.1E-13	144	8.20E+04
vanillin*	Van	271	163	1.58	1.43	1.28	2.2E-07	80.7	3.84E-03
acetovanillone*	Acv	67.3	32.1	1.45	1.42	0.83	2.0E-07	74.7	3.51E-03
coniferaldehyde*	Con	11.7	5.55	12.3	1.66	1.44	2.4E-08	88.4	4.50E-02
syringaldehyde*	Syr	5.55	2.64	26.1	2.13	1.43	8.0E-09	92.0	1.53E-01
acetosyringone*	Ace	1.88	0.46	25.6	1.79	0.78	7.2E-09	86.0	1.38E-01

(a) Mean fraction of particle-phase SVOC.

(b) Signal to noise ratio = mean concentration/mean uncertainty.

(c) Coefficient of variation = standard deviation/mean concentration.

(d) Pure compound vapor pressure at 298.15 K.

(e) Enthalpy of vaporization of the liquid.

(f) Average value.

(g) Obtained from filter measurement, not including gas phase.

\* Species not included for PMF analysis.

Table 6.S2 Statistics for the total (gas + particle phase) concentration of each SVOC estimated for sampling days with ambient temperature lower than 10 °C (Cold period).

	Abbreviation	Mean	Median	Mean fraction (%)	CV	S/N	$K_{p,om}$ (m <sup>3</sup> µg <sup>-1</sup> )
<b>Bulk species (µg m<sup>-3</sup>)</b>							
Mass*		9.57	7.64		0.70	4.13	
Nitrate	Nitr	2.23	1.41		1.04	12.2	
Sulfate	Sulf	1.30	0.79		1.13	13.5	
Ammonium*	Ammo	0.88	0.56		1.02	21.0	
EC		0.61	0.48		0.75	7.20	
OC1*		1.29	1.13		0.55	4.64	
OC2		0.56	0.48		0.54	4.61	
OC3		0.40	0.35		0.53	2.51	
OC4*		0.04	0.01		2.27	1.00	
PC		0.42	0.30		1.03	2.60	
<b>Organic molecular makers (ng m<sup>-3</sup>)</b>							
<i>Alkanes</i>							
docosane	C22	6.31	4.89	35.5	2.75	0.82	2.12E-01
tricosane	C23	3.11	2.54	57.8	2.45	0.69	6.76E-01
tetracosane	C24	1.56	1.43	77.7	1.53	0.54	2.16E+00
pentacosane	C25	1.59	1.39	90.4	2.35	0.60	6.92E+00
hexacosane	C26	1.22	1.07	96.4	2.50	0.66	2.22E+01
heptacosane	C27	1.37	1.19	98.7	2.60	0.69	7.15E+01
octacosane	C28	1.17	1.00	99.6	2.72	0.73	2.31E+02
nonacosane	C29	1.41	1.21	99.9	2.62	0.70	7.46E+02
triacontane	C30	0.98	0.89	100	2.25	0.78	2.42E+03
hentriacontane	C31	1.43	1.23	100	2.13	0.76	7.86E+03
dotriacontane*	C32	0.49	0.32	100	1.09	1.14	2.56E+04
tritriacontane	C33	0.73	0.55	100	2.04	0.90	8.36E+04
tetratriacontane	C34	0.58	0.43	100	2.12	0.91	2.73E+05
pentatriacontane	C35	0.55	0.42	100	2.22	0.85	8.97E+05
hexatriacontane	C36	0.32	0.22	100	2.01	0.96	2.95E+06
heptatriacontane	C37	0.27	0.18	100	1.85	0.99	9.72E+06
octatriacontane	C38	0.22	0.15	100	1.68	0.99	3.21E+07
nonatriacontane	C39	0.20	0.14	100	1.60	0.96	1.06E+08
tetracontane*	C40	0.15	0.10	100	1.40	1.00	3.52E+08
pentadecylcyclohexane	cycC21	1.93	1.30	16.7	1.92	1.06	5.93E-02
nonadecylcyclohexane*	cycC25	0.22	0.18	89.9	1.98	0.71	6.11E+00
<i>PAHs</i>							
fluoranthene	Flu	4.15	3.08	9.18	1.69	0.84	2.58E-02
pyrene	Pyr	0.81	0.70	31.1	1.84	0.72	1.48E-01
benzo[ghi]fluoranthene	BghiF	0.22	0.19	73.7	2.38	0.67	1.33E+00
cyclopenta[cd]pyrene	C-pyr	0.09	0.06	77.2	2.61	1.03	1.63E+00
benz[a]anthracene	BaA	0.15	0.11	74.9	2.58	0.89	1.48E+00
chrysene/triphenylene*	CT	0.42	0.33	74.9	2.86	0.74	1.48E+00
benzo[b&k]fluoranthene	BbkF	0.42	0.31	95.6	2.42	0.78	1.34E+01
benzo[j]fluoranthene*	BjF	0.02	0.01	95.6	1.44	3.53	1.34E+01
benz[a&e]pyrene	BaeP	0.36	0.24	99.1	2.47	0.93	7.79E+01
perylene*	Per	0.03	0.02	99.1	1.62	1.19	7.79E+01
indeno[1,2,3-cd]pyrene	IP	0.13	0.11	99.9	2.58	0.76	7.05E+02
benzo[ghi]perylene	BP	0.32	0.25	100	2.85	0.80	4.15E+03
dibenz[ah]anthracene*	DahA	0.04	0.03	99.9	2.25	0.92	7.97E+02
picene*	Pic	0.03	0.02	99.9	2.05	0.95	7.97E+02
coronene	Cor	0.18	0.13	100	1.50	0.91	2.24E+05
methyl-202-PAH sum	M-202	5.85	4.89	21.5	1.80	0.71	8.13E-02
retene*	Ret	2.44	1.74	43.1	1.81	1.22	2.87E-01
methyl-228-PAH sum	M-228	0.26	0.18	89.0	2.49	0.91	4.72E+00

Table 6.S2 Continued

	Abbreviation	Mean	Median	Mean fraction (%)	CV	S/N	$K_{p,om}$ ( $m^3 \mu g^{-1}$ )
<i>Oxy-PAHs</i>							
acenaphthenone*	Ace-O	463	176	0.20	1.18	2.88	4.42E-04
fluorenone*	Flu-O	455	162	0.61	1.16	3.25	1.38E-03
1H-phenalen-1-one*	Phe-O	204	89.8	0.75	1.11	1.82	1.70E-03
xanthone*	Xan	9.07	7.31	3.31	1.98	0.87	7.98E-03
1,8-naphthalic anhydride	Nap-DO	15.9	12.8	3.41	2.00	0.70	8.23E-03
anthracene-9,10-dione	Ant-DO	4.53	3.54	12.6	2.00	0.76	3.68E-02
benz[de]anthracene-7-one	BaA-O	0.23	0.19	62.9	2.77	0.63	6.92E-01
<i>Steranes</i>							
20R-abb & 20S-aaa-cholestane	27-RS-C	0.24	0.15	98.4	2.62	1.07	4.34E+01
20R & S-abb-methylcholestane	28-RS-M	0.20	0.11	99.5	2.42	1.18	1.39E+02
20R & S-abb-ethylcholestane	29-RS-E	0.22	0.14	99.8	2.92	1.02	4.45E+02
a-22,29,30-trisnorhopane	TS	0.13	0.09	98.3	2.96	0.97	3.85E+01
ba-30-norhopane	ba-N	0.59	0.37	99.8	2.68	1.07	3.93E+02
ab-hopane	ab-H	0.43	0.27	99.9	2.63	1.12	1.26E+03
22S-ab-30-homohopane	31abS	0.19	0.12	100	2.66	1.11	4.05E+03
22R-ab-30-homohopane	31abR	0.14	0.09	100	2.65	1.12	4.05E+03
22S-ab-30-bishomohopane	32abS	0.11	0.07	100	2.64	1.12	1.30E+04
22R-ab-30-bishomohopane	32abR	0.09	0.05	100	2.60	1.05	1.30E+04
<i>Alkanoic acids</i>							
dodecanoic acid	C12:0	108	63.8	6.35	1.97	1.32	1.67E-02
tridecanoic acid*	C13:0	4.06	1.42	15.8	0.77	1.90	5.27E-02
tetradecanoic acid*	C14:0	19.0	11.5	32.8	1.50	1.14	1.67E-01
pentadecanoic acid*	C15:0	2.33	1.66	55.4	1.38	0.89	5.28E-01
hexadecanoic acid	C16:0	34.0	25.1	76.2	1.96	0.93	1.68E+00
heptadecanoic acid*	C17:0	1.18	0.74	89.7	1.04	1.00	5.34E+00
octadecanoic acid	C18:0	15.5	10.5	96.1	1.74	0.96	1.70E+01
oleic acid*	C18:1	5.39	0.60	96.8	2.33	1.87	2.09E+01
<i>Sterols and methoxyphenols</i>							
cholesterol*	Cho	0.60	0.28	100	1.46	1.37	1.65E+04
stigmasterol*	Sti	0.77	0.49	100	2.27	1.13	2.10E+05
vanillin*	Van	328	224	3.41	1.51	1.13	8.23E-03
acetovanillone*	Acv	76.0	55.7	3.04	1.04	0.95	7.28E-03
coniferaldehyde*	Con	15.9	10.1	25.3	1.79	1.21	9.98E-02
syringaldehyde*	Syr	7.95	5.05	49.0	1.75	1.28	3.45E-01
acetosyringone*	Ace	1.98	1.19	47.1	1.80	1.20	3.03E-01

Table 6.S3 Statistics for the total (gas + particle phase) concentration of each SVOC estimated for sampling days with ambient temperature between 10 °C and 20 °C (Warm period).

	Abbreviation	Mean	Median	Mean fraction (%)	CV	S/N	$K_{p,om}$ (m <sup>3</sup> µg <sup>-1</sup> )
<b>Bulk species (µg m<sup>-3</sup>)</b>							
Mass*		6.54	6.18		0.44	3.26	
Nitrate	Nitr	0.37	0.17		1.58	2.56	
Sulfate	Sulf	1.01	0.90		0.67	12.4	
Ammonium*	Ammo	0.32	0.23		0.96	14.0	
EC		0.43	0.36		0.62	6.35	
OC1*		1.08	1.05		0.33	4.67	
OC2		0.57	0.54		0.46	4.76	
OC3		0.39	0.38		0.38	2.64	
OC4*		0.03	0.01		2.23	1.00	
PC		0.28	0.22		0.78	2.57	
<b>Organic molecular makers (ng m<sup>-3</sup>)</b>							
<i>Alkanes</i>							
docosane	C22	23.2	18.0	6.42	2.40	0.84	1.97E-02
tricosane	C23	12.2	8.53	15.9	2.33	0.91	5.62E-02
tetracosane	C24	2.56	2.15	33.7	1.41	0.80	1.60E-01
pentacosane	C25	1.86	1.63	57.1	2.07	0.66	4.56E-01
hexacosane	C26	0.81	0.61	77.8	1.94	0.86	1.30E+00
heptacosane	C27	0.98	0.84	90.3	2.25	0.70	3.71E+00
octacosane	C28	0.61	0.36	96.2	1.94	0.99	1.06E+01
nonacosane	C29	1.49	1.19	98.6	2.22	0.75	3.02E+01
triacontane	C30	0.49	0.27	99.5	1.59	1.07	8.61E+01
hentriacontane	C31	1.50	1.09	99.8	1.19	0.88	2.46E+02
dotriacontane*	C32	0.39	0.21	99.9	0.65	1.66	7.02E+02
tritriacontane	C33	0.51	0.41	100	1.84	0.84	2.00E+03
tetratriacontane	C34	0.37	0.28	100	1.84	0.98	5.73E+03
pentatriacontane	C35	0.28	0.23	100	1.76	0.86	1.64E+04
hexatriacontane	C36	0.14	0.09	100	1.48	1.09	4.68E+04
heptatriacontane	C37	0.11	0.05	100	1.27	1.22	1.34E+05
octatriacontane	C38	0.09	0.04	100	1.12	1.19	3.82E+05
nonatriacontane	C39	0.09	0.05	100	1.12	1.16	1.09E+06
tetracontane*	C40	0.06	0.04	100	0.96	1.16	3.13E+06
pentadecylcyclohexane	cycC21	7.11	4.84	2.36	1.99	1.00	6.85E-03
nonadecylcyclohexane*	cycC25	0.18	0.16	57.0	1.83	0.65	4.50E-01
<i>PAHs</i>							
fluoranthene	Flu	10.8	8.22	1.46	1.73	0.95	4.18E-03
pyrene	Pyr	1.32	1.05	6.70	1.76	0.79	2.05E-02
benzo[ghi]fluoranthene	BghiF	0.16	0.13	34.8	2.12	0.70	1.64E-01
cyclopenta[cd]pyrene	C-pyr	0.04	0.03	39.9	1.89	1.20	2.07E-01
benz[a]anthracene	BaA	0.09	0.05	34.9	2.05	1.17	1.66E-01
chrysene/triphenylene*	CT	0.39	0.33	34.9	2.50	0.67	1.66E-01
benzo[b&k]fluoranthene	BbkF	0.20	0.14	79.1	2.03	1.01	1.33E+00
benzo[j]fluoranthene*	BjF	0.01	0.01	79.1	0.60	1.26	1.33E+00
benz[a&e]pyrene	BaeP	0.13	0.08	94.6	2.05	1.26	6.52E+00
perylene*	Per	0.01	0.00	94.6	1.06	1.48	6.52E+00
indeno[1,2,3-cd]pyrene	IP	0.05	0.04	99.3	2.24	0.84	5.22E+01
benzo[ghi]perylene	BP	0.12	0.09	99.8	2.30	0.84	2.57E+02
dibenz[ah]anthracene*	DahA	0.01	0.01	99.3	1.35	1.28	5.29E+01
picene*	Pic	0.01	0.01	99.3	1.13	1.26	5.29E+01
coronene	Cor	0.07	0.05	100	1.59	0.95	1.01E+04
methyl-202-PAH sum	M-202	8.83	6.78	4.02	1.70	0.82	1.19E-02
retene*	Ret	1.80	0.91	10.6	1.54	1.52	3.43E-02
methyl-228-PAH sum	M-228	0.15	0.11	58.8	2.14	0.95	4.72E-01

Table 6.S3 Continued

	Abbreviation	Mean	Median	Mean fraction (%)	CV	S/N	$K_{p,om}$ (m <sup>3</sup> µg <sup>-1</sup> )
<i>Oxy-PAHs</i>							
acenaphthenone*	Ace-O	411	194	0.04	0.63	1.13	1.14E-04
fluorenone*	Flu-O	433	312	0.12	0.83	0.97	3.24E-04
1H-phenalen-1-one*	Phe-O	283	161	0.15	1.66	1.21	4.09E-04
xanthone*	Xan	43.6	34.1	0.58	2.02	0.82	1.64E-03
1,8-naphthalic anhydride	Nap-DO	54.2	46.9	0.67	2.02	0.57	1.88E-03
anthracene-9,10-dione	Ant-DO	18.2	13.8	2.73	2.01	0.85	7.90E-03
benz[de]anthracene-7-one	BaA-O	0.20	0.16	25.5	2.32	0.78	1.02E-01
<i>Steranes</i>							
20R-abb & 20S-aaa-cholestane	27-RS-C	0.14	0.11	90.4	2.40	0.72	3.50E+00
20R & S-abb-methylcholestane	28-RS-M	0.09	0.07	96.3	2.52	0.96	9.98E+00
20R & S-abb-ethylcholestane	29-RS-E	0.11	0.08	98.6	2.44	0.99	2.84E+01
a-22,29,30-trisnorhopane	TS	0.08	0.07	90.4	2.50	0.79	3.46E+00
ba-30-norhopane	ba-N	0.27	0.21	98.6	2.33	0.98	2.80E+01
ab-hopane	ab-H	0.19	0.14	99.5	2.20	1.17	7.99E+01
22S-ab-30-homohopane	31abS	0.08	0.06	99.8	2.26	1.23	2.28E+02
22R-ab-30-homohopane	31abR	0.06	0.04	99.8	2.17	1.13	2.28E+02
22S-ab-30-bishomohopane	32abS	0.05	0.04	99.9	2.16	1.08	6.50E+02
22R-ab-30-bishomohopane	32abR	0.04	0.03	99.9	2.10	1.09	6.50E+02
<i>Alkanoic acids</i>							
dodecanoic acid	C12:0	536	405	0.92	1.73	0.83	2.62E-03
tridecanoic acid*	C13:0	24.8	13.3	2.57	0.93	1.14	7.46E-03
tetradecanoic acid*	C14:0	96.4	66.4	6.89	1.68	1.14	2.12E-02
pentadecanoic acid*	C15:0	8.74	6.16	17.0	1.36	1.06	6.03E-02
hexadecanoic acid	C16:0	66.8	40.5	35.6	1.65	1.25	1.72E-01
heptadecanoic acid*	C17:0	1.34	0.94	59.3	0.73	1.09	4.89E-01
octadecanoic acid	C18:0	16.4	9.48	79.5	1.47	1.49	1.39E+00
oleic acid*	C18:1	1.14	0.16	82.9	1.42	2.88	1.75E+00
<i>Sterols and methoxyphenols</i>							
cholesterol*	Cho	0.30	0.13	100	0.96	1.27	8.77E+02
stigmasterol*	Sti	0.32	0.17	100	1.13	1.44	8.99E+03
vanillin*	Van	275	151	0.62	1.33	1.60	1.74E-03
acetovanillone*	Acv	70.8	30.3	0.62	0.96	1.56	1.75E-03
coniferaldehyde*	Con	10.5	3.68	5.86	1.47	2.29	1.77E-02
syringaldehyde*	Syr	4.43	0.99	16.2	1.37	2.65	5.63E-02
acetosyringone*	Ace	1.07	0.23	16.4	0.72	2.12	5.67E-02

Table 6.S4 Statistics for the total (gas + particle phase) concentration of each SVOC estimated for sampling days with ambient temperature above 20 °C (Hot period).

	Abbreviation	Mean	Median	Mean fraction (%)	CV	S/N	$K_{p,om}$ (m <sup>3</sup> µg <sup>-1</sup> )
<b>Bulk species (µg m<sup>-3</sup>)</b>							
Mass*		7.79	7.53		0.38	4.28	
Nitrate*	Nitr	0.16	0.14		0.76	1.27	
Sulfate	Sulf	1.22	1.11		0.46	14.0	
Ammonium*	Ammo	0.51	0.48		0.54	10.3	
EC		0.46	0.44		0.42	5.81	
OC1*		1.48	1.46		0.25	4.45	
OC2		0.86	0.84		0.38	3.85	
OC3		0.52	0.50		0.30	2.13	
OC4*		0.02	0.01		0.86	1.00	
PC		0.45	0.33		0.79	2.24	
<b>Organic molecular makers (ng m<sup>-3</sup>)</b>							
<i>Alkanes</i>							
docosane	C22	76.8	58.4	2.21	2.48	0.85	4.49E-03
tricosane	C23	62.5	47.8	5.67	2.39	0.81	1.21E-02
tetracosane*	C24	13.7	9.13	13.7	2.01	0.94	3.24E-02
pentacosane	C25	8.50	5.80	29.2	2.22	0.84	8.69E-02
hexacosane	C26	1.41	1.11	51.2	1.97	0.74	2.33E-01
heptacosane	C27	1.44	1.26	72.8	2.55	0.54	6.26E-01
octacosane	C28	0.52	0.40	87.3	1.85	0.85	1.68E+00
nonacosane	C29	2.38	1.99	94.7	2.30	0.70	4.52E+00
triacontane*	C30	0.36	0.24	97.9	1.48	1.14	1.21E+01
hentriacontane*	C31	2.53	1.60	99.2	0.78	1.27	3.26E+01
dotriacontane*	C32	0.67	0.21	99.7	0.54	1.76	8.77E+01
tritriacontane	C33	0.59	0.49	99.9	1.93	0.78	2.36E+02
tetratriacontane	C34	0.37	0.24	100	1.98	1.06	6.34E+02
pentatriacontane	C35	0.26	0.23	100	1.91	0.78	1.70E+03
hexatriacontane*	C36	0.11	0.09	100	1.43	1.08	4.59E+03
heptatriacontane*	C37	0.07	0.05	100	1.10	1.38	1.23E+04
octatriacontane*	C38	0.06	0.03	100	0.87	1.41	3.32E+04
nonatriacontane*	C39	0.06	0.04	100	0.89	1.06	8.94E+04
tetracontane*	C40	0.04	0.03	100	0.76	1.16	2.41E+05
pentadecylcyclohexane	cycC21	16.9	9.94	0.88	2.03	1.22	1.75E-03
nonadecylcyclohexane*	cycC25	0.41	0.36	30.1	1.80	0.62	9.06E-02
<i>PAHs</i>							
fluoranthene	Flu	20.4	15.6	0.65	1.76	0.82	1.29E-03
pyrene	Pyr	2.49	1.92	2.85	1.76	0.87	5.80E-03
benzo[ghi]fluoranthene	BghiF	0.17	0.13	17.7	1.89	0.81	4.35E-02
cyclopenta[cd]pyrene*	C-pyr	0.04	0.02	21.5	1.51	2.04	5.57E-02
benz[a]anthracene	BaA	0.12	0.07	17.0	1.70	1.45	4.17E-02
chrysene/triphenylene*	CT	0.65	0.49	17.0	2.61	1.22	4.17E-02
benzo[b&k]fluoranthene	BbkF	0.16	0.11	59.0	1.78	1.97	3.13E-01
benzo[j]fluoranthene*	BjF	0.01	0.01	59.0	0.55	1.20	3.13E-01
benz[a&e]pyrene	BaeP	0.07	0.05	85.9	1.68	1.80	1.40E+00
perylene*	Per	0.01	0.01	85.9	0.78	2.86	1.40E+00
indeno[1,2,3-cd]pyrene	IP	0.03	0.02	97.8	1.95	2.11	1.05E+01
benzo[ghi]perylene	BP	0.07	0.05	99.5	2.09	1.44	4.74E+01
dibenz[ah]anthracene*	DahA	0.01	0.00	97.7	0.88	3.33	1.01E+01
picene*	Pic	0.01	0.00	97.7	0.92	8.60	1.01E+01
coronene	Cor	0.04	0.03	100	1.57	1.95	1.60E+03
methyl-202-PAH sum	M-202	12.2	9.23	1.73	1.64	0.90	3.47E-03
retene*	Ret	1.59	0.81	4.28	1.21	1.52	8.91E-03
methyl-228-PAH sum*	M-228	0.16	0.13	34.8	2.00	0.97	1.12E-01

Table 6.S4 Continued

	Abbreviation	Mean	Median	Mean fraction (%)	CV	S/N	$K_{p,om}$ (m <sup>3</sup> µg <sup>-1</sup> )
<i>Oxy-PAHs</i>							
acenaphthenone*	Ace-O	442	91.0	0.02	0.55	1.62	4.65E-05
fluorenone*	Flu-O	518	220	0.06	0.73	1.60	1.24E-04
1H-phenalen-1-one*	Phe-O	451	198	0.08	1.68	1.58	1.59E-04
xanthone*	Xan	91.5	78.5	0.30	2.05	0.69	5.83E-04
1,8-naphthalic anhydride*	Nap-DO	138	114	0.36	1.99	0.63	7.13E-04
anthracene-9,10-dione	Ant-DO	45.4	36.9	1.45	2.00	0.74	2.88E-03
benz[de]anthracene-7-one*	BaA-O	0.22	0.18	13.0	2.18	0.74	3.00E-02
<i>Steranes</i>							
20R-abb & 20S-aaa-cholestane	27-RS-C	0.19	0.18	76.5	2.23	0.44	7.40E-01
20R & S-abb-methylcholestane	28-RS-M	0.08	0.08	89.4	2.56	0.43	1.99E+00
20R & S-abb-ethylcholestane	29-RS-E	0.09	0.08	95.7	2.14	0.44	5.33E+00
a-22,29,30-trisnorhopane	TS	0.11	0.10	77.4	2.37	0.43	7.72E-01
ba-30-norhopane	ba-N	0.25	0.23	95.9	2.22	0.42	5.56E+00
ab-hopane	ab-H	0.15	0.14	98.4	2.07	0.44	1.49E+01
22S-ab-30-homohopane	31abS	0.06	0.06	99.4	2.09	0.46	4.01E+01
22R-ab-30-homohopane	31abR	0.05	0.04	99.4	1.96	0.50	4.01E+01
22S-ab-30-bishomohopane	32abS	0.04	0.04	99.8	1.99	0.49	1.08E+02
22R-ab-30-bishomohopane	32abR	0.04	0.03	99.8	1.92	0.53	1.08E+02
<i>Alkanoic acids</i>							
dodecanoic acid*	C12:0	1662	1180	0.40	1.67	1.02	7.96E-04
tridecanoic acid*	C13:0	71.5	36.6	1.07	0.84	1.25	2.13E-03
tetradecanoic acid*	C14:0	407	301	2.80	1.94	0.84	5.72E-03
pentadecanoic acid*	C15:0	34.7	23.7	7.12	1.45	0.93	1.53E-02
hexadecanoic acid*	C16:0	188	120	16.8	1.51	0.98	4.11E-02
heptadecanoic acid*	C17:0	4.38	3.00	34.4	0.64	0.95	1.10E-01
octadecanoic acid*	C18:0	21.6	14.9	57.4	1.47	0.88	2.96E-01
oleic acid*	C18:1	0.64	0.24	63.1	0.86	2.91	3.79E-01
<i>Sterols and methoxyphenols</i>							
cholesterol*	Cho	0.24	0.13	99.8	0.78	1.29	1.50E+02
stigmasterol*	Sti	0.33	0.13	100	0.69	1.70	1.38E+03
vanillin*	Van	194	123	0.32	0.93	1.71	6.26E-04
acetovanillone*	Acv	52.7	8.50	0.35	0.53	1.94	6.82E-04
coniferaldehyde*	Con	7.81	3.24	2.83	0.94	1.45	5.75E-03
syringaldehyde*	Syr	3.76	0.68	8.08	0.99	3.53	1.75E-02
acetosyringone*	Ace	2.66	0.28	8.75	0.52	1.85	1.90E-02

Table 6.S5 Mean particle-phase fractions of selected SVOCs and data from other field studies.

Sample No.	This study (PM <sub>2.5</sub> )				Fraser et al. (1997,1998) (PM <sub>10</sub> )		Simcik et al. (1997, 1998) (TSP)		Tsapakis and Stephanou (2005) (TSP)		Mandalakis et al. (2002) (TSP)	
	Whole	Cold	Warm	Hot	Los Angeles, Summer	Chicago, Winter	Chicago, Summer	Lake Michigan, Summer	Heraklion (Greece), Annual	Athens (Greece), Summer	Athens (Greece), Summer	
Average Temp. (°C)	14	3.3	15	24	27	2.7	23	22	17	28		
Sample No.	970	364	318	288	32	1-4	1-18	2-15	16	4		
<i>Alkanes</i>												
docosane	16	36	6.4	2.2	2.8						17	
tricosane	29	58	16	5.7	10						26	
tetracosane	44	78	34	14	25						36	
pentacosane	61	90	57	29	32						56	
hexacosane	77	96	78	51	44						62	
heptacosane	88	99	90	73	72						81	
octacosane	98	100	96	87	94						66	
nonacosane	98	100	99	95	100						87	
triacontane	99	100	100	98	100						65	
hentriacontane	100	100	100	99	100						91	
dotriacontane	100	100	100	100	100						78	
tritriacontane	100	100	100	100	100						70	
tetatriacontane	100	100	100	100	100							
pentatriacontane	100	100	100	100	100							
hexatriacontane	100	100	100	100	100							
<i>PAHs</i>												
fluoranthene	4.1	9.2	1.5	0.65	0.71	55	17	9.7	3.7	6.4		
pyrene	15	31	6.7	2.8	1.0	54	20	14	4.7	9.0		
benzo[ghi]fluoranthene	44	74	35	18	18							
cyclopenta[cd]pyrene	49	77	40	22	35				32			
benz[a]anthracene	45	75	35	17	60	95	81	49	57	41		
chrysene/triphenylene	45	75	35	17	44	92	72	40	53	45		
benzo[ <i>b</i> , <i>k</i> ]fluoranthene	79	96	79	59	100	94	93	77	94	96		
benzo[ <i>j</i> ]fluoranthene	79	96	79	59	100	95	90	88	92	99		
benz[ <i>a</i> , <i>e</i> ]pyrene	94	99	95	86	100	95			91	100		
perylene	94	99	95	86	100				98	100		
indeno[1,2,3- <i>cd</i> ]pyrene	99	100	99	98	100			92	97	100		
benzo[ <i>ghi</i> ]perylene	100	100	100	100	100				100	100		
dibenz[ <i>ah</i> ]anthracene	99	100	99	98	100	96	95		100	100		
coronene	100	100	100	100	100				100	100		
retene	21	43	11	4.3					10			

Table 6.S5 Continued

	This study (PM <sub>2.5</sub> )				Fraser et al. (1997,1998) (PM <sub>1.6</sub> )	Mandalakis et al. (2002) (TSP)
	Whole	Cold	Warm	Hot	Los Angeles, Summer	Athens (Greece), Summer
Average Temp. (°C)	14	3.3	15	24	27	28
Sample No.	970	364	318	288	32	4
<i>Sterane and hopanes</i>						
20R-abb & 20S-aaa-cholestane	89	98	90	76	100	
a-22,29,30-trisnorhopane	89	98	90	77	100	40
ab-hopane	99	100	100	98	100	83
22S-ab-30-homohopane	100	100	100	100	100	93
22R-ab-30-homohopane	100	100	100	100	100	92
22S-ab-30-bishomohopane	100	100	100	100	100	100
22R-ab-30-bishomohopane	100	100	100	100	100	100

Table 6.S6 Correlation coefficients of factor contributions from full data set solution versus meteorological and trace gas measurements.

Factors	Temperature (°C)	Radiance (KW-hr m <sup>-2</sup> )	RH (%)	Ozone (ppm)	NO <sub>x</sub> (ppm)	CO (ppm)
Whole period						
Nitrate	<b>-0.60</b>	<b>-0.46</b>	<b>0.49</b>	<b>-0.51</b>	<b>0.45</b>	0.39
Sulfate	0.15	0.21	0.31	0.04*	-0.07*	-0.05*
<i>n</i> -Alkane	-0.28	-0.22	0.10*	-0.35	<b>0.45</b>	<b>0.43</b>
Sterane	-0.35	-0.32	0.17	<b>-0.47</b>	<b>0.64</b>	<b>0.58</b>
Light SVOCs	<b>0.73</b>	<b>0.50</b>	<b>-0.48</b>	0.34	-0.30	-0.20
PAH	-0.39	-0.34	0.05*	<b>-0.59</b>	<b>0.65</b>	<b>0.62</b>
Bulk carbon	0.39	0.33	-0.22	<b>0.46</b>	-0.10*	0.05*
Cold period						
Nitrate	-0.38	-0.10*	<b>0.40</b>	<b>-0.40</b>	0.31	0.29
Sulfate	-0.22	0.08*	<b>0.40</b>	-0.07*	-0.03*	-0.04*
<i>n</i> -Alkane	-0.19	-0.16	-0.11*	-0.33	<b>0.47</b>	<b>0.43</b>
Sterane	-0.09*	-0.28	-0.01*	<b>-0.50</b>	<b>0.57</b>	<b>0.54</b>
Light SVOCs	<b>0.61</b>	0.15	-0.34	0.24	-0.21	-0.15
PAH	0.03*	-0.06*	-0.30	<b>-0.46</b>	<b>0.57</b>	<b>0.60</b>
Bulk carbon	0.15	0.14	-0.16	-0.15	0.28	0.33
Warm period						
Nitrate	-0.20	-0.18	0.38	-0.23	0.23	0.27
Sulfate	0.16	0.16	<b>0.43</b>	0.12*	-0.16	-0.04*
<i>n</i> -Alkane	-0.05*	-0.06*	-0.03*	-0.05*	0.22	0.17
Sterane	-0.14*	-0.10*	-0.01*	-0.21	<b>0.60</b>	<b>0.55</b>
Light SVOCs	0.35	0.23	-0.33	0.05*	-0.08*	-0.06*
PAH	-0.25	-0.15*	-0.10*	<b>-0.43</b>	<b>0.62</b>	<b>0.61</b>
Bulk carbon	0.19	0.09*	0.28	0.08*	0.18	0.25
Hot period						
Nitrate	-0.02*	-0.04*	0.03*	0.03*	0.08*	0.06*
Sulfate	0.07*	0.11*	0.26	0.18	-0.09*	-0.03*
<i>n</i> -Alkane	-0.01*	0.08*	-0.03*	0.05*	0.31	0.16
Sterane	0.01*	0.03*	0.01*	-0.00*	0.36	0.34
Light SVOCs	0.16	-0.06*	<b>-0.45</b>	<b>-0.48</b>	0.01*	0.03*
PAH	-0.15*	-0.15*	-0.11*	<b>-0.46</b>	<b>0.44</b>	0.31
Bulk carbon	0.21	0.17	0.10*	0.36	0.21	0.39

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

Table 6.S7 Correlation coefficients of factor contributions from sub-data set solutions versus meteorological and trace gas measurements.

Factors	Temperature (°C)	Radiance (KW-hr m <sup>-2</sup> )	RH (%)	Ozone (ppm)	NO <sub>x</sub> (ppm)	CO (ppm)
Cold period						
Nitrate	<b>-0.43</b>	-0.09*	0.39	<b>-0.40</b>	0.32	0.29
Sulfate	-0.27	0.07*	<b>0.42</b>	-0.09*	-0.02*	-0.03*
<i>n</i> -Alkane	-0.22	-0.11*	-0.09*	-0.32	<b>0.43</b>	0.39
Sterane	-0.11*	-0.28	-0.00*	<b>-0.48</b>	<b>0.53</b>	<b>0.50</b>
Light SVOCs	<b>0.61</b>	0.17	<b>-0.40</b>	0.26	-0.22	-0.15
PAH	-0.10*	-0.13*	-0.24	<b>-0.44</b>	<b>0.49</b>	<b>0.52</b>
Bulk carbon	-0.04*	0.03*	-0.14*	<b>-0.51</b>	<b>0.76</b>	<b>0.76</b>
Warm period						
Nitrate	-0.06*	-0.03*	0.37	-0.23	0.24	0.29
Sulfate	0.19	0.21	<b>0.40</b>	0.16	-0.22	-0.10*
<i>n</i> -Alkane	-0.04*	-0.06*	-0.01*	-0.04*	0.22	0.16
Sterane	-0.12*	-0.03*	-0.01*	-0.20	<b>0.58</b>	<b>0.53</b>
Light SVOCs	0.33	0.19	-0.39	0.11*	-0.17	-0.15
PAH	-0.27	-0.21	-0.12*	<b>-0.43</b>	<b>0.61</b>	<b>0.61</b>
Bulk carbon	0.23	0.10*	0.26	0.03*	0.22	0.28
Hot period						
Sulfate	0.08*	0.13*	0.29	0.23	-0.11*	-0.07*
<i>n</i> -Alkane	0.02*	0.11*	-0.01*	0.16	0.15*	0.16
Sterane	-0.02*	0.05*	0.16	0.13*	0.20	0.24
Light SVOCs	0.07*	0.02*	-0.36	<b>-0.46</b>	0.01*	-0.06*
PAH	-0.19	-0.08*	-0.13*	<b>-0.44</b>	<b>0.47</b>	0.32
Bulk carbon	0.14*	0.05*	0.07*	0.16	0.31	<b>0.47</b>
Medium <i>n</i> -alkane	<b>0.59</b>	0.26	<b>-0.45</b>	0.15*	-0.20	-0.04*

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

Table 7.S1 Sampling dates and sampling matrices for gas- and particle-phase SVOCs.

Sampling dates	Sample ID	Quartz fiber filter (QFF) pack	PUF/XAD/PUF cartridge	
			<i>n</i> -alkanes and PAHs	2-methyltetrols and levoglucosan
08/22/12	S1	tQFF and bQFF <sup>c</sup>	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF <sup>d</sup>	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF <sup>d</sup>
08/28/12	S2	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF <sup>d</sup>	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
09/03/12	S3	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
09/09/12 <sup>a</sup>	S4	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-7 (5 g)/PUF <sup>d</sup>
09/15/12 <sup>a</sup>	S5	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-7 (5 g)/PUF
09/21/12 <sup>a</sup>	S6	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-7 (5 g)/PUF
09/27/12	S7	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
10/03/12	S8	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
10/09/12	S9	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
10/15/12	S10	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
10/21/12	S11	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
10/27/12	S12	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
11/02/12	S13	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
11/08/12 <sup>a</sup>	S14	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-7 (5 g)/PUF
11/14/12	S15	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
11/20/12	S16	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
11/26/12	S17	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
12/02/12	S18	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
12/08/12	S19	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
12/14/12	S20	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
01/13/13 <sup>a</sup>	S21	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-7 (5 g)/PUF
01/19/13	S22	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
01/25/13	S23	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
01/31/13	S24	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF
02/06/13	S25	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF

**Table 7.S1 Continued**

Sampling dates	Sample ID	Quartz fiber filter (QFF) pack	<i>n</i> -alkanes and PAHs	PUF/XAD/PUF cartridge	2-methylterols and levoglucosan
02/12/13	S26	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF	/
02/18/13 <sup>b</sup>	S27	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	/
02/24/13 <sup>b</sup>	S28	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	/
03/02/13 <sup>b</sup>	S29	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	/
03/08/13 <sup>b</sup>	S30	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	/
03/14/13	S31	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
03/19/13 <sup>a</sup>	S32	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF
04/04/13	S33	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
04/10/13	S34	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
04/16/13 <sup>a</sup>	S35	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF
04/22/13	S36	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
04/28/13	S37	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
05/04/13	S38	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
05/10/13	S39	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
05/16/13	S40	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
05/22/13	S41	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
05/28/13	S42	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
06/03/13	S43	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF <sup>e</sup>
06/20/13	S44	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
06/26/13	S45	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF <sup>e</sup>
07/02/13	S46	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
07/08/13 <sup>a</sup>	S47	tQFF and bQFF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF	two PUF(50 mm diam. × 20 mm length)/XAD-4 (5 g)/PUF
07/14/13	S48	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
07/20/13	S49	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF
07/25/13	S50	tQFF and bQFF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF	PUF (50 mm diam. × 40 mm length)/XAD-4 (5 g)/PUF

(a) Breakthrough samples were collected on these sampling days.

(b) PUF/XAD/PUF samples for 2-methylterols and levoglucosan were not collected.

(c) Top QFF (where PM<sub>2.5</sub> were loaded) and backup QFF (used to estimate positive sampling artifact of particulate organics).

(d) Samples used to test extraction and quantification methods, and the data obtained were not used for analysis.

(e) Samples extractions were failed, and not data were obtained.

Table 7.S2 Non-polar SVOC compounds and their referenced internal standards.

Compound Name (Empirical Formula)	M/Z <sup>a</sup>	Internal Standard (Empirical Formula) <sup>b</sup>
<i>n</i> -Alkanes		
dodecane (C <sub>12</sub> H <sub>26</sub> )	170	pentadecane (C <sub>15</sub> D <sub>34</sub> )
tridecane (C <sub>13</sub> H <sub>28</sub> )	184	pentadecane (C <sub>15</sub> D <sub>34</sub> )
tetradecane (C <sub>14</sub> H <sub>30</sub> )	198	pentadecane (C <sub>15</sub> D <sub>34</sub> )
pentadecane (C <sub>15</sub> H <sub>32</sub> )	212	pentadecane (C <sub>15</sub> D <sub>34</sub> )
hexadecane (C <sub>16</sub> H <sub>34</sub> )	226	pentadecane (C <sub>15</sub> D <sub>34</sub> )
heptadecane (C <sub>17</sub> H <sub>36</sub> )	240	pentadecane (C <sub>15</sub> D <sub>34</sub> )
octadecane (C <sub>18</sub> H <sub>38</sub> )	254	pentadecane (C <sub>15</sub> D <sub>34</sub> )
nonadecane (C <sub>19</sub> H <sub>40</sub> )	268	pentadecane (C <sub>15</sub> D <sub>34</sub> )
eicosane (C <sub>20</sub> H <sub>42</sub> )	282	eicosane (C <sub>20</sub> D <sub>42</sub> )
heneicosane (C <sub>21</sub> H <sub>44</sub> )	296	eicosane (C <sub>20</sub> D <sub>42</sub> )
docosane (C <sub>22</sub> H <sub>46</sub> )	310	eicosane (C <sub>20</sub> D <sub>42</sub> )
tricosane (C <sub>23</sub> H <sub>48</sub> )	324	eicosane (C <sub>20</sub> D <sub>42</sub> )
tetracosane (C <sub>24</sub> H <sub>50</sub> )	338	triacontane (C <sub>30</sub> D <sub>62</sub> )
pentacosane (C <sub>25</sub> H <sub>52</sub> )	352	triacontane (C <sub>30</sub> D <sub>62</sub> )
hexacosane (C <sub>26</sub> H <sub>54</sub> )	366	triacontane (C <sub>30</sub> D <sub>62</sub> )
heptacosane (C <sub>27</sub> H <sub>56</sub> )	380	triacontane (C <sub>30</sub> D <sub>62</sub> )
octacosane (C <sub>28</sub> H <sub>58</sub> )	394	triacontane (C <sub>30</sub> D <sub>62</sub> )
nonacosane (C <sub>29</sub> H <sub>60</sub> )	408	triacontane (C <sub>30</sub> D <sub>62</sub> )
triacontane (C <sub>30</sub> H <sub>62</sub> )	422	triacontane (C <sub>30</sub> D <sub>62</sub> )
hentriacontane (C <sub>31</sub> H <sub>64</sub> )	436	triacontane (C <sub>30</sub> D <sub>62</sub> )
dotriacontane (C <sub>32</sub> H <sub>66</sub> )	450	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
tritriacontane (C <sub>33</sub> H <sub>68</sub> )	464	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
tetratriacontane (C <sub>34</sub> H <sub>70</sub> )	478	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
pentatriacontane (C <sub>35</sub> H <sub>72</sub> )	492	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
hexatriacontane (C <sub>36</sub> H <sub>74</sub> )	506	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
heptatriacontane (C <sub>37</sub> H <sub>76</sub> )	520	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
octatriacontane (C <sub>38</sub> H <sub>78</sub> )	534	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
nonatriacontane (C <sub>39</sub> H <sub>80</sub> )	548	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )
tetracontane (C <sub>40</sub> H <sub>82</sub> )	562	hexatriacontane (C <sub>36</sub> D <sub>74</sub> )

Table 7.S2 Continued

Compound Name (Empirical Formula)	M/Z	Internal Standard (Empirical Formula)
PAHs		
naphthalene (C <sub>10</sub> H <sub>8</sub> )	128	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
2-methylnaphthalene (C <sub>11</sub> H <sub>10</sub> )	142	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
1-methylnaphthalene (C <sub>11</sub> H <sub>10</sub> )	142	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
acenaphthylene (C <sub>12</sub> H <sub>8</sub> )	152	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
acenaphthene (C <sub>12</sub> H <sub>10</sub> )	154	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
fluorene (C <sub>13</sub> H <sub>10</sub> )	166	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
2-methylfluorene (C <sub>14</sub> H <sub>12</sub> )	180	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
phenanthrene (C <sub>14</sub> H <sub>10</sub> )	178	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
anthracene (C <sub>14</sub> H <sub>10</sub> )	178	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
methyl-178-PAH sum (C <sub>15</sub> H <sub>12</sub> ) <sup>c</sup>	192	acenaphthene (C <sub>12</sub> D <sub>10</sub> )
fluoranthene (C <sub>16</sub> H <sub>10</sub> )	202	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
pyrene (C <sub>16</sub> H <sub>10</sub> )	202	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
methyl-202-PAH sum (C <sub>17</sub> H <sub>12</sub> ) <sup>c</sup>	216	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
benzo[ghi]fluoranthene (C <sub>18</sub> H <sub>10</sub> )	226	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
cyclopenta[cd]pyrene (C <sub>18</sub> H <sub>10</sub> )	226	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
benz[a]anthracene (C <sub>18</sub> H <sub>12</sub> )	228	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
chrysene/triphenylene (C <sub>18</sub> H <sub>12</sub> )	228	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
methyl-228-PAH sum (C <sub>19</sub> H <sub>14</sub> ) <sup>c</sup>	242	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
retene (C <sub>18</sub> H <sub>18</sub> )	234	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
benzo[b&k]fluoranthene (C <sub>20</sub> H <sub>12</sub> )	252	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
benzo[j]fluoranthene (C <sub>20</sub> H <sub>12</sub> ) <sup>c</sup>	252	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
benz[a&e]pyrene (C <sub>20</sub> H <sub>12</sub> )	252	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
perylene (C <sub>20</sub> H <sub>12</sub> )	252	benz[a]anthracene(C <sub>18</sub> D <sub>12</sub> )
indeno[1,2,3-cd]pyrene (C <sub>22</sub> H <sub>12</sub> )	276	coronene (C <sub>24</sub> D <sub>12</sub> )
benzo[ghi]perylene (C <sub>22</sub> H <sub>12</sub> )	276	coronene(C <sub>24</sub> D <sub>12</sub> )
dibenz[ah]anthracene (C <sub>22</sub> H <sub>14</sub> )	278	coronene (C <sub>24</sub> D <sub>12</sub> )
picene (C <sub>22</sub> H <sub>14</sub> )	278	coronene (C <sub>24</sub> D <sub>12</sub> )
coronene (C <sub>24</sub> H <sub>12</sub> )	300	coronene (C <sub>24</sub> D <sub>12</sub> )

(a) Mass to charge ratio (m/z);

(b) Referenced internal standard (IS) compounds isotopically labeled by deuterium (D) substitution.

(c) Additional compounds not included in the quantification standards. Calibration curves for the next closest compound based on molecular structure and weight were used for quantification of these compounds.

Table 7.S3 Recoveries of meso-erythritol and levoglucosan for all sampling matrices used in this work.

Recoveries (%)	QFF	XAD-7	PUF/XAD-7/PUF	PUF
No. of repetition	4	3	4	3
Meso-erythritol <sup>b</sup>	85.6 ± 1.9 <sup>c</sup>	100.7 ± 4.8 <sup>c</sup>	103.8 ± 11.0 <sup>c</sup>	90.0 ± 8.7 <sup>c</sup>
Levoglucosan	95.2 ± 0.8 <sup>c</sup>	90.5 ± 1.9 <sup>c</sup>	51.9 ± 4.4 <sup>c</sup>	70.7 ± 1.7 <sup>c</sup>
Recoveries (%)	XAD-4	PUF/XAD-4/PUF	PUF/XAD/PUF (BT) <sup>a</sup>	
No. of repetition	2	2	4	
Meso-erythritol <sup>b</sup>	86.0 ± 2.9 <sup>d</sup>	79.3 ± 7.4 <sup>d</sup>	85.1 ± 3.2 <sup>c</sup>	
Levoglucosan	77.8 ± 0.1 <sup>d</sup>	61.2 ± 12.1 <sup>d</sup>	63.3 ± 4.4 <sup>c</sup>	

(a) PUF/XAD/PUF sandwiches used for breakthrough experiments, and the four repetitions contains two PUF/XAD-7/PUF and two PUF/XAD-4/PUF sandwiches.

(b) Surrogate compound used to quantify 2-methyltetrols.

(c) Average ± standard deviation (SD).

(d) Average ± relative percent difference (RPD) between two measurements,  $RPD = \frac{|X1 - X2|}{(X1 + X2)/2} \times 100\%$ .

### Interpretation

The recoveries of meso-erythritol and levoglucosan were firstly investigated for QFF, XAD-7 resin (5 g), PUF (50 mm diam. × 40 mm length)/XAD-7 (5 g)/PUF sandwiches (for regular sampling) and PUF adsorbent (50 mm diam. × 40 mm length), and the number of repetitions was at least 3. The recoveries of the two species for QFF analysis were 85.6 ± 1.9% and 95.2 ± 0.8%, respectively. For those sampling matrices used for gaseous phase, the recoveries of levoglucosan decreased from 90.5 ± 1.9% to 51.9 ± 4.4% as the usage of PUF adsorbent increased, suggesting that the amount of PUF adsorbent used for sampling primarily affected the recoveries of levoglucosan. However, the differences across repeated observations were very small, as shown by the standard deviations of recoveries. When we replaced XAD-7 for XAD-4 in PUF/XAD/PUF sandwiches for gas-phase sampling, only 2 repetitions of recoveries were conducted for XAD-4 (5 g), PUF/XAD-4/PUF sandwiches (for regular sampling), and the relative percent differences between each two observations were smaller than 20%. Two PUF (50 mm diam. × 20 mm length)/XAD-7 (5g)/PUF and two PUF/XAD-4/PUF sandwiches were used to obtain the recoveries for breakthrough sample analysis. The recoveries of meso-erythritol and levoglucosan were comparable across the four observations (82.9 – 89.9%, meso-erythritol; 58.9 – 68.0%, levoglucosan), and thus were combined to calculate the average recoveries for breakthrough sample analysis.

Table 7.S4 Concentrations of polar SVOCs in different parts of selected PUF/XAD/PUF samples.

Compounds	S2 <sup>a</sup>	S3 <sup>a,b</sup>	S8 <sup>a</sup>	S26 <sup>a</sup>	S31 <sup>c</sup>	S37 <sup>c</sup>	S44 <sup>c</sup>	S46 <sup>c</sup>
tPUF								
2-Methylthreitol	4.07		/	/	/	/	1.56	1.02
2-Methylerythritol	7.51		/	/	/	/	4.31	4.33
Levoglucosan	2.18		1.10	2.18	2.39	4.90	44.1	3.44
Middle XAD resin								
2-Methylthreitol	/	/	/	/	/	/	/	/
2-Methylerythritol	/	/	/	/	/	/	/	/
Levoglucosan	/	/	/	/	/	/	/	/
bPUF								
2-Methylthreitol	/		/	/	/	/	/	/
2-Methylerythritol	/		/	/	/	/	/	/
Levoglucosan	/		/	/	/	/	/	/

(a) XAD-7 resin was used.

(b) The tPUF and bPUF of sample S3 were combined and analyzed together, and corrected by the recoveries of PUF/XAD-7/PUF sandwich in Table S3. The data for tPUF and bPUF could not be obtained.

(c) XAD-4 resin was used.

Table 7.S5 Statistics of *n*-alkanes and PAHs quantified from QFFs and PUF/XAD-4/PUF samples.

Chemical species (MW)	tQFF (N = 50)					bQFF (N = 50)					PUF/XAD-4/PUF (N = 48)				
	No. of Obs.	Median	Mean	Range	S/N <sup>a</sup>	No. of Obs.	Median	Mean	Range	S/N	No. of Obs.	Median	Mean	Range	S/N
<i>n</i> -Alkanes															
dodecane (170)	4	0.16	0.19	0.05 - 0.40	2.63	17	0.09	0.16	0.01 - 0.96	1.93	48	26.0	29.7	3.13 - 102	5.10
tridecane (184)	26	0.11	0.44	0.02 - 7.15	5.48	36	0.10	0.22	0.01 - 2.87	3.64	48	20.4	21.5	3.61 - 60.3	9.77
tetradecane (198)	7	0.09	0.11	0.01 - 0.22	2.48	38	0.12	0.15	0.00 - 0.92	2.29	48	26.1	27.7	5.16 - 59.0	13.8
pentadecane (212)	35	0.12	0.18	0.01 - 0.55	3.28	44	0.13	0.22	0.02 - 0.85	2.91	48	21.8	22.9	5.57 - 44.8	25.3
hexadecane (226)	48	0.41	0.48	0.02 - 1.44	5.31	47	0.24	0.40	0.01 - 1.77	3.79	48	17.1	18.2	4.55 - 37.5	17.1
heptadecane (240)	47	0.52	0.81	0.00 - 3.08	5.70	49	0.64	1.14	0.01 - 5.70	5.72	48	17.6	21.5	1.07 - 66.3	10.5
octadecane (254)	43	0.51	0.92	0.04 - 3.47	5.04	48	0.83	1.26	0.04 - 5.97	6.19	47	9.64	12.8	0.21 - 48.5	7.63
nonadecane (268)	49	0.68	1.14	0.00 - 6.29	5.72	50	0.85	1.30	0.05 - 5.11	7.08	46	3.42	6.00	0.05 - 24.0	6.84
eicosane (282)	49	0.65	0.82	0.01 - 2.87	12.6	48	0.52	0.65	0.02 - 2.24	10.5	40	1.38	2.28	0.02 - 8.64	5.86
heneicosane (296)	50	0.92	1.08	0.09 - 2.64	12.9	48	0.24	0.46	0.01 - 2.65	8.67	26	1.12	1.53	0.04 - 4.75	6.45
docosane (310)	50	1.00	1.06	0.13 - 2.46	7.94	40	0.11	0.34	0.01 - 1.87	5.60	27	0.31	0.57	0.00 - 2.54	2.93
tricosane (324)	50	1.53	1.88	0.36 - 6.44	7.81	34	0.13	0.38	0.00 - 1.92	5.38	26	0.19	0.43	0.02 - 1.68	2.25
tetracosane (338)	50	1.04	1.52	0.24 - 5.76	6.79	42	0.04	0.13	0.00 - 1.02	3.13	25	0.21	0.31	0.00 - 2.00	1.36
pentacosane (352)	50	1.13	1.54	0.31 - 5.67	7.47	42	0.05	0.10	0.00 - 0.59	2.34	27	0.17	0.25	0.01 - 2.17	1.33
hexacosane (366)	50	0.60	0.84	0.19 - 3.53	8.45	38	0.04	0.08	0.00 - 0.44	2.16	23	0.09	0.18	0.00 - 1.36	1.05
heptacosane (380)	50	0.82	1.09	0.26 - 3.65	10.9	32	0.06	0.10	0.00 - 0.64	2.49	17	0.14	0.23	0.00 - 1.31	1.55
octacosane (394)	50	0.44	0.62	0.14 - 3.02	11.1	29	0.06	0.10	0.00 - 0.35	2.35	11	0.18	0.26	0.01 - 0.99	1.79
nonacosane (408)	50	1.48	1.78	0.31 - 7.53	23.4	33	0.07	0.09	0.00 - 0.32	1.48	17	0.16	0.17	0.01 - 0.58	0.74
triacontane (422)	50	0.31	0.44	0.04 - 1.91	8.52	30	0.06	0.07	0.00 - 0.28	1.32	11	0.05	0.10	0.01 - 0.33	0.73
hentriacontane (436)	50	1.13	1.47	0.28 - 7.63	12.5	34	0.08	0.09	0.00 - 0.40	1.38	7	0.04	0.10	0.02 - 0.35	0.62
dotriacontane (450)	50	0.28	0.37	0.03 - 1.20	4.93	27	0.05	0.08	0.00 - 0.46	1.21	6	0.13	0.24	0.05 - 0.66	1.39
tritriacontane (464)	50	0.60	0.69	0.16 - 1.65	7.51	27	0.06	0.10	0.01 - 0.52	1.28	3	0.11	0.46	0.07 - 1.20	1.26
tetraatriacontane (478)	50	0.32	0.40	0.03 - 1.43	4.78	25	0.04	0.09	0.00 - 0.34	1.18	3	0.20	0.49	0.04 - 1.21	1.47
pentatriacontane (492)	50	0.31	0.37	0.05 - 1.16	4.54	29	0.04	0.07	0.00 - 0.35	1.10	3	0.20	0.49	0.04 - 1.21	1.47
hexatriacontane (506)	50	0.15	0.18	0.01 - 0.46	2.78	25	0.02	0.06	0.00 - 0.23	0.96	25	0.02	0.06	0.00 - 0.23	0.96
heptatriacontane (520)	50	0.19	0.32	0.00 - 1.86	4.96	25	0.02	0.05	0.00 - 0.18	0.71	25	0.02	0.05	0.00 - 0.18	0.71
octatriacontane (534)	45	0.11	0.13	0.02 - 0.36	1.98	18	0.04	0.06	0.01 - 0.19	0.96	18	0.04	0.06	0.01 - 0.19	0.96
nonatriacontane (548)	43	0.16	0.19	0.04 - 0.47	3.09	13	0.04	0.06	0.02 - 0.15	0.97	13	0.04	0.06	0.02 - 0.15	0.97
tetracontane (562)	33	0.10	0.12	0.05 - 0.31	1.77	9	0.05	0.06	0.03 - 0.12	0.86	9	0.05	0.06	0.03 - 0.12	0.86

**Table 7.S5 Continued**

Chemical species (MW)	tQFF				bQFF				PUF/XAD-4/PUF					
	No. of Obs.	Median	Mean	S/N	No. of Obs.	Median	Mean	Range	S/N	No. of Obs.	Median	Mean	Range	S/N
<i>PAHs</i>														
naphthalene (128)	15	0.03	0.16	0.00 - 0.53	1.14	0.04	0.47	0.00 - 3.79	2.50	48	58.6	70.3	10.8 - 186	4.40
2-methylnaphthalene (142)	31	0.03	0.07	0.00 - 0.47	2.84	0.03	0.18	0.00 - 2.42	4.46	48	26.4	29.1	5.58 - 68.6	9.70
1-methylnaphthalene (142)	4	0.07	0.64	0.00 - 2.47	8.65	0.04	0.44	0.00 - 3.85	4.71	48	10.9	12.3	1.73 - 44.0	7.81
acenaphthylene (152)	34	0.01	0.01	0.00 - 0.04	2.27	0.00	0.01	0.00 - 0.14	2.12	48	0.21	0.55	0.02 - 5.79	10.1
acenaphthene (154)	27	0.25	0.35	0.01 - 1.28	10.6	0.01	0.05	0.01 - 0.28	2.07	48	6.74	8.04	0.67 - 30.3	18.2
fluorine (166)	37	0.02	0.04	0.00 - 0.18	2.08	0.02	0.05	0.00 - 0.41	2.62	48	4.55	4.95	0.99 - 14.5	13.1
2-methylfluorene (180)	14	0.02	0.04	0.00 - 0.18	1.70	0.02	0.19	0.00 - 2.59	5.41	48	0.48	0.50	0.06 - 1.61	5.79
phenanthrene (178)	45	0.11	0.18	0.01 - 1.02	1.39	0.11	0.35	0.01 - 3.67	2.76	48	6.85	9.31	0.20 - 27.7	7.48
anthracene (178)	49	0.03	0.05	0.00 - 0.29	1.78	0.01	0.10	0.00 - 2.53	4.01	41	0.23	0.27	0.02 - 0.79	3.17
methyl-178-PAH (192)	48	0.25	0.42	0.03 - 1.84	2.24	0.15	0.54	0.02 - 9.87	2.93	46	2.25	2.71	0.07 - 7.79	3.89
fluoranthene (202)	50	0.10	0.19	0.01 - 0.95	5.26	0.05	0.09	0.00 - 0.32	4.99	47	0.63	0.95	0.03 - 3.28	4.49
pyrene (202)	50	0.07	0.12	0.01 - 0.90	6.00	0.01	0.02	0.00 - 0.20	3.37	44	0.15	0.28	0.01 - 1.05	3.14
methyl-202-PAH (216)	50	0.31	0.61	0.02 - 5.26	5.31	0.04	0.05	0.00 - 0.25	1.48	14	0.77	0.97	0.38 - 4.11	4.60
benzo[ghi]fluoranthene (226)	50	0.04	0.07	0.00 - 0.53	14.3	0.00	0.01	0.00 - 0.02	4.59	7	0.02	0.02	0.01 - 0.03	4.10
cyclopenta[cd]pyrene (226)	50	0.03	0.07	0.00 - 0.84	8.03	0.00	0.00	0.00 - 0.01	1.23	5	0.02	0.06	0.01 - 0.26	5.80
benz[a]anthracene (228)	49	0.02	0.06	0.00 - 0.45	12.0	0.00	0.00	0.00 - 0.00	1.04	10	0.01	0.06	0.00 - 0.55	7.88
chrysene/triphenylene (228)	50	0.10	0.15	0.02 - 0.82	16.5	0.01	0.01	0.00 - 0.04	4.09	10	0.03	0.05	0.01 - 0.23	6.89
retene (234)	49	0.38	1.59	0.01 - 18.8	6.13	0.05	0.11	0.01 - 0.84	4.95	12	0.21	0.32	0.03 - 1.51	0.97
methyl-228-PAH (242)	48	0.07	0.18	0.01 - 3.35	24.5									
benzo[b&k]fluoranthene (252)	50	0.10	0.17	0.02 - 0.94	12.1									
benz[a&e]pyrene (252)	50	0.09	0.17	0.01 - 1.05	9.61									
indeno[1,2,3-cd]pyrene (276)	50	0.04	0.05	0.01 - 0.29	12.3									
benzo[ghi]perylene (276)	50	0.07	0.11	0.02 - 0.61	11.0									
dibenz[ah]anthracene (278)	40	0.01	0.02	0.00 - 0.11	6.03									
picene (278)	36	0.01	0.02	0.00 - 0.06	6.89									
coronene (300)	50	0.02	0.04	0.00 - 0.30	6.99									

(a) Signal to noise ratio (mean concentration/mean uncertainty).

Table 8.S1 Properties of selected SVOCs for analysis.

Species	Formula	MW (g mol <sup>-1</sup> )	$p_L^{\circ}$ (atm) <sup>a</sup>	$\Delta H_{\text{vap}}$ (kJ mol <sup>-1</sup> ) <sup>a</sup>	$K_{p,OM}^t$ <sup>b</sup>
docosane <sup>c</sup>	C <sub>22</sub> H <sub>46</sub>	310	$3.25 \times 10^{-8}$	115	0.029
tricosane <sup>c</sup>	C <sub>23</sub> H <sub>48</sub>	324	$1.22 \times 10^{-8}$	120	0.085
fluoranthene <sup>c</sup>	C <sub>16</sub> H <sub>10</sub>	202	$1.07 \times 10^{-7}$	92.2	0.0058
pyrene <sup>c</sup>	C <sub>16</sub> H <sub>10</sub>	202	$2.43 \times 10^{-8}$	98.9	0.029
acenaphthenone <sup>c</sup>	C <sub>12</sub> H <sub>8</sub> O	168	$2.88 \times 10^{-6}$	71.4	0.00015
fluorenone <sup>c</sup>	C <sub>13</sub> H <sub>8</sub> O	180	$1.08 \times 10^{-6}$	75.8	0.00043
2-methyltetrols <sup>d</sup>	C <sub>5</sub> H <sub>12</sub> O <sub>4</sub>	136	$1.47 \times 10^{-9}$	38.4	0.17
levoglucosan <sup>e</sup>	C <sub>6</sub> H <sub>10</sub> O <sub>5</sub>	162	$2.38 \times 10^{-10}$	84.0	2.51

(a) Values were obtained for 298.15 K.

(b) Median values.

(c) Values of  $p_L^{\circ}$  and  $\Delta H_{\text{vap}}$  were estimated using the group contribution methods (GCMs) SIMPOL (Pankow and Asher, 2008).

(d) Values of  $p_L^{\circ}$  and  $\Delta H_{\text{vap}}$  were obtained from Couvidat and Seigneur (2011).

(e) Values of  $p_L^{\circ}$  and  $\Delta H_{\text{vap}}$  were obtained from Booth et al. (2011).

Table 8.S2 Statistics for bulk carbon ( $\mu\text{g m}^{-3}$ ) and all selected SVOCs ( $\text{ng m}^{-3}$ ) concentrations in QFFs and PUF/XAD/PUF samples.

	No. of obs.	Median	Mean	Range	S/N <sup>a</sup>	Breakthrough, %
tQFF						
EC <sup>b</sup>	48	0.28	0.40	0.06 – 1.83	3.2	/
OC <sup>c</sup>	48	3.06	3.56	0.94 – 10.9	3.6	/
docosane	50	1.00	1.06	0.13 – 2.46	7.9	/
tricosane	50	1.53	1.88	0.36 – 6.44	7.8	/
fluoranthene	50	0.10	0.19	0.01 – 0.95	5.3	/
pyrene	50	0.07	0.12	0.01 – 0.90	6.0	/
acenaphthenone <sup>d</sup>	47	0.13	0.18	0.02 – 0.92	6.3	/
fluorenone <sup>d</sup>	50	0.36	0.42	0.02 – 1.67	5.8	/
2-methyltetrols	38	1.22	3.80	0.11 – 27.5	/	/
levoglucosan	48	27.1	57.6	2.48 – 434	/	/
bQFF						
OC <sup>e</sup>	49	0.45	0.48	0.18 – 1.13	3.1	/
docosane	40	0.11	0.34	0.01 - 1.87	5.6	/
tricosane	34	0.13	0.38	0.00 - 1.92	5.4	/
fluoranthene	47	0.05	0.09	0.00 - 0.32	5.0	/
pyrene	41	0.01	0.02	0.00 - 0.20	3.4	/
acenaphthenone <sup>d</sup>	43	0.06	0.15	0.01 - 2.37	7.8	/
fluorenone <sup>d</sup>	49	0.35	0.63	0.01 - 9.57	8.2	/
2-methyltetrols	12	1.02	2.88	0.05 - 17.0	/	/
levoglucosan	3	0.33	1.88	0.09 - 5.22	/	/
PUF/XAD/PUF						
docosane	27	0.31	0.57	0.00 - 2.54	2.9	0 <sup>f</sup> (3 <sup>g</sup> )
tricosane	26	0.19	0.43	0.02 - 1.68	2.3	0 <sup>f</sup> (3 <sup>g</sup> )
fluoranthene	47	0.63	0.95	0.03 - 3.28	4.4	0 <sup>f</sup> (6 <sup>g</sup> )
pyrene	44	0.15	0.28	0.01 - 1.05	3.1	0 <sup>f</sup> (3 <sup>g</sup> )
acenaphthenone <sup>d</sup>	37	0.32	0.37	0.03 - 1.10	6.4	1.46 <sup>f</sup> (7 <sup>g</sup> )
fluorenone <sup>d</sup>	44	0.33	0.65	0.00 - 2.20	5.2	1.41 <sup>f</sup> (8 <sup>g</sup> )
2-methyltetrols	14	5.61	6.62	1.54 - 15.4	/	0 <sup>f</sup> (3 <sup>g</sup> )
levoglucosan	39	6.71	9.31	0.60 - 49.1	/	11.0 <sup>f</sup> (5 <sup>g</sup> )

(a) Signal to noise ratio (mean concentration/mean uncertainty).

(b) Not observed on bQFF.

(c) Sum of OC fractions (OC1 + OC2 + OC3 + OC4 + PC).

(d) Quantified as fluoranthene, and acenaphthene-D10 was the internal standard.

(e) Only include OC1 fraction that observed on bQFF.

(f) Average breakthrough.

(g) Number of observations.

Table 8.S3 Statistics of  $K_{p,OM}^m$  values with different approaches for artifact corrections.

Species	C1 (Positive artifact)			C2 (Negative artifact)			C3 (Positive offsets negative)		
	No. of obs.	median	mean	No. of obs.	median	mean	No. of obs.	median	mean
dococane	38	0.85	4.06	27	0.61	9.33	43	1.21	5.17
tricosane	36	1.23	7.98	26	1.62	4.02	36	1.54	11.9
fluoranthene	43	0.02	0.12	47	0.04	0.26	48	0.03	0.18
pyrene	43	0.11	1.16	44	0.11	0.44	45	0.11	1.55
acenaphthenone	24	0.12	0.46	34	0.10	0.38	44	0.07	0.47
fluorenone	23	0.27	1.23	44	0.57	3.57	48	0.13	0.90
	C'1 (No TMF absorption)			C'2 (TMF absorption similar as QFF)					
	No. of obs.	median	mean	No. of obs.	median	mean			
2-methyltetrol	14	0.20	0.23	14	0.17	0.20			
levoglucosan	39	0.97	2.59	39	0.93	2.54			

## 11.2 SUPPLEMENTAL FIGURES

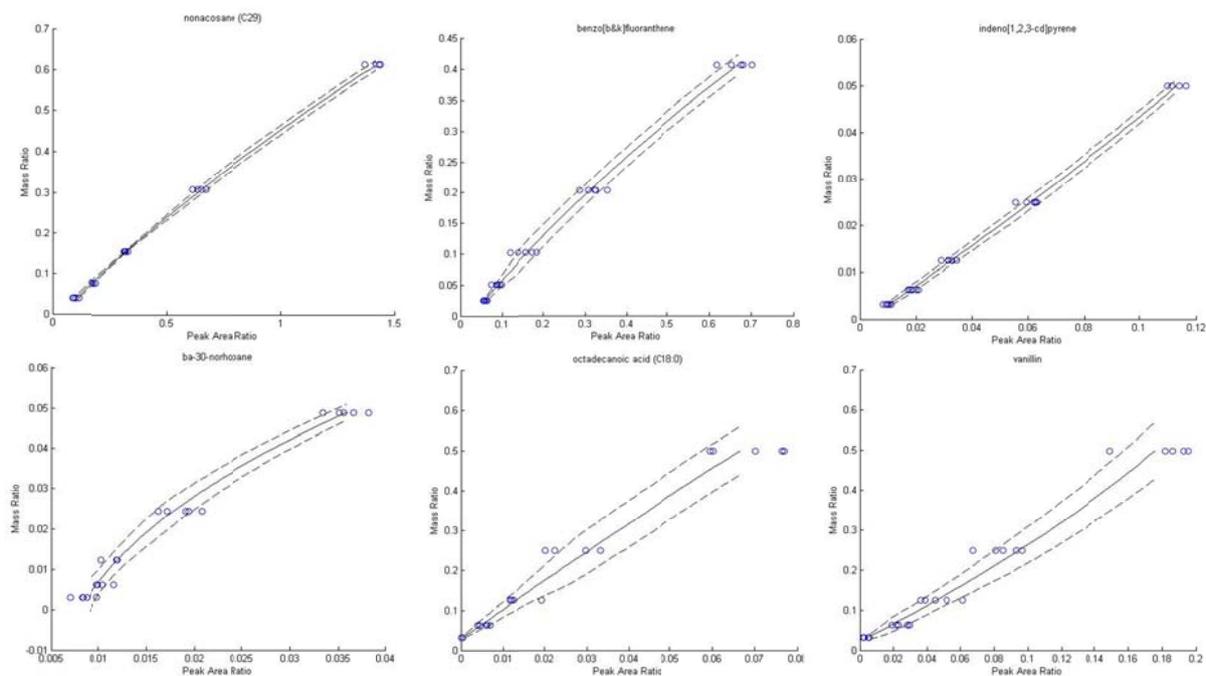
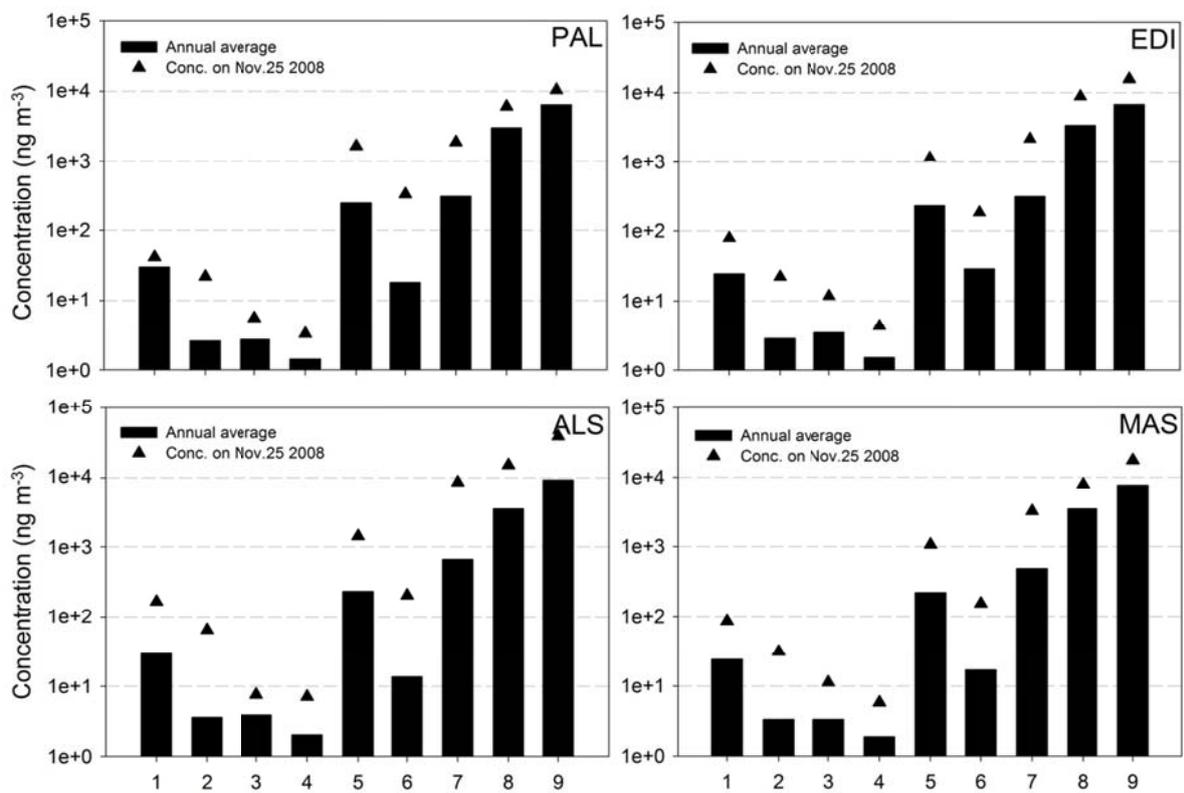


Figure 2.S1 Calibration curves for selected species from one of the three quantification batches.

The six plots above are calibration curves used for quantification, which determines the final mass amount of each molecular maker by converting peak area ratios to mass ratios. The dashed lines represent uncertainties associated with mass ratios that correspond to the peak area ratios of each compound in samples. For measurements below the range covered by the quantification standard dilutions, the calibration curve was extrapolated linearly through zero, and the absolute uncertainty was extrapolated down. While for the measurements above the range of quantification standard dilutions, the calibration curve was extrapolated upward by following the slope at the top end of the quadratic calibration curve, and the relative uncertainty was extrapolated up (Dutton et al., 2009a).



1. *n*-alkanes; 2. PAHs; 3. Oxy-PAHs; 4. Steranes; 5. *n*-alkanoic acids; 6. Sterols and Methoxyphenols; 7. EC; 8. OC; 9. PM<sub>2.5</sub> mass

Figure 2.S2 Concentrations of different groups of PM<sub>2.5</sub> species on November 25<sup>th</sup> 2008 (black triangles) vs. annual averages (bar plots).

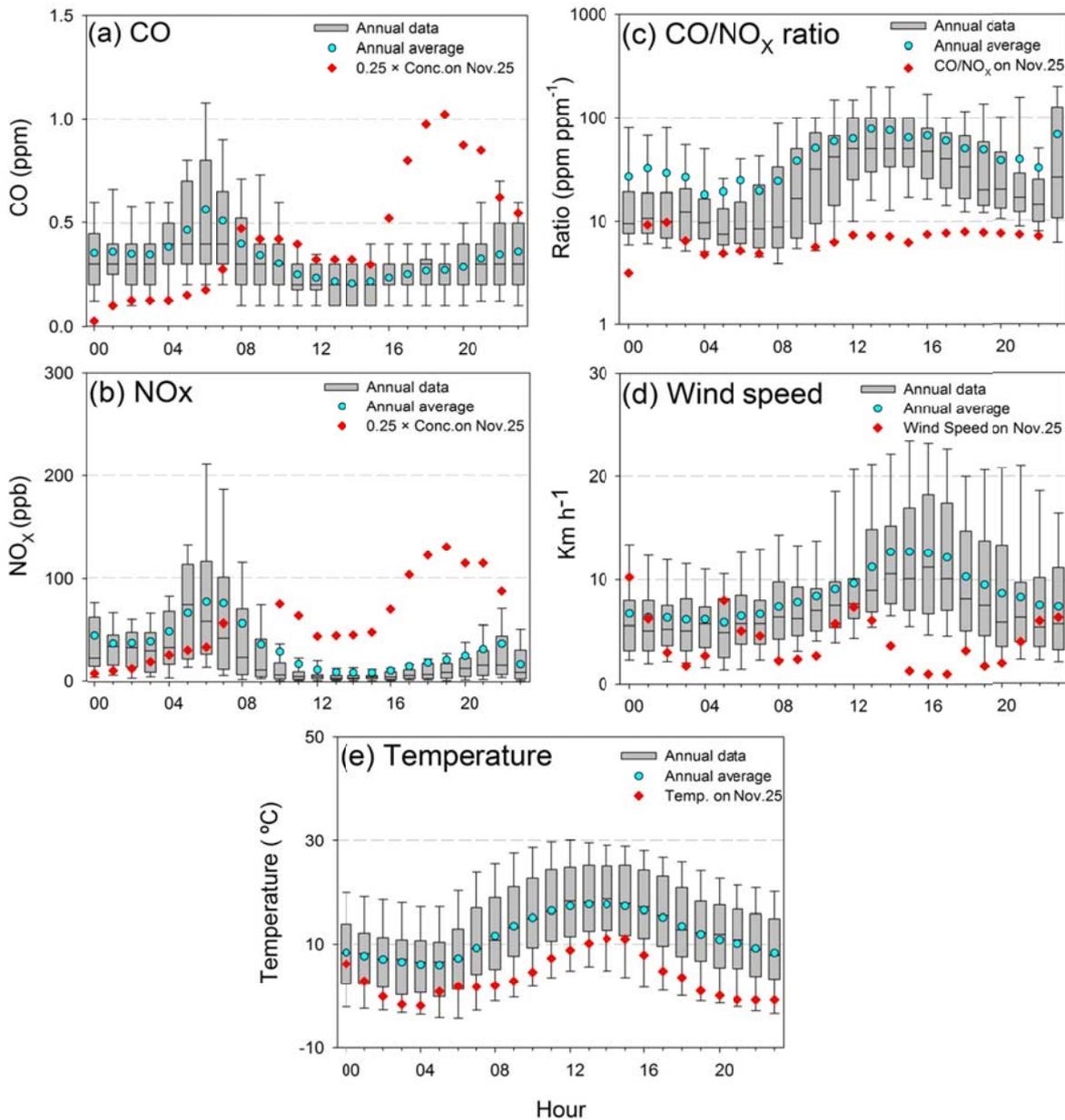


Figure 2.S3 Hourly observations for (a) CO, (b) NO<sub>x</sub>, (c) CO/NO<sub>x</sub> ratios, (d) wind speed, (e) ambient temperature on November 25, 2008 (red diamond) vs. annual distribution data (box plots). The boxes depict the median (black line), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and mean (light blue circle).

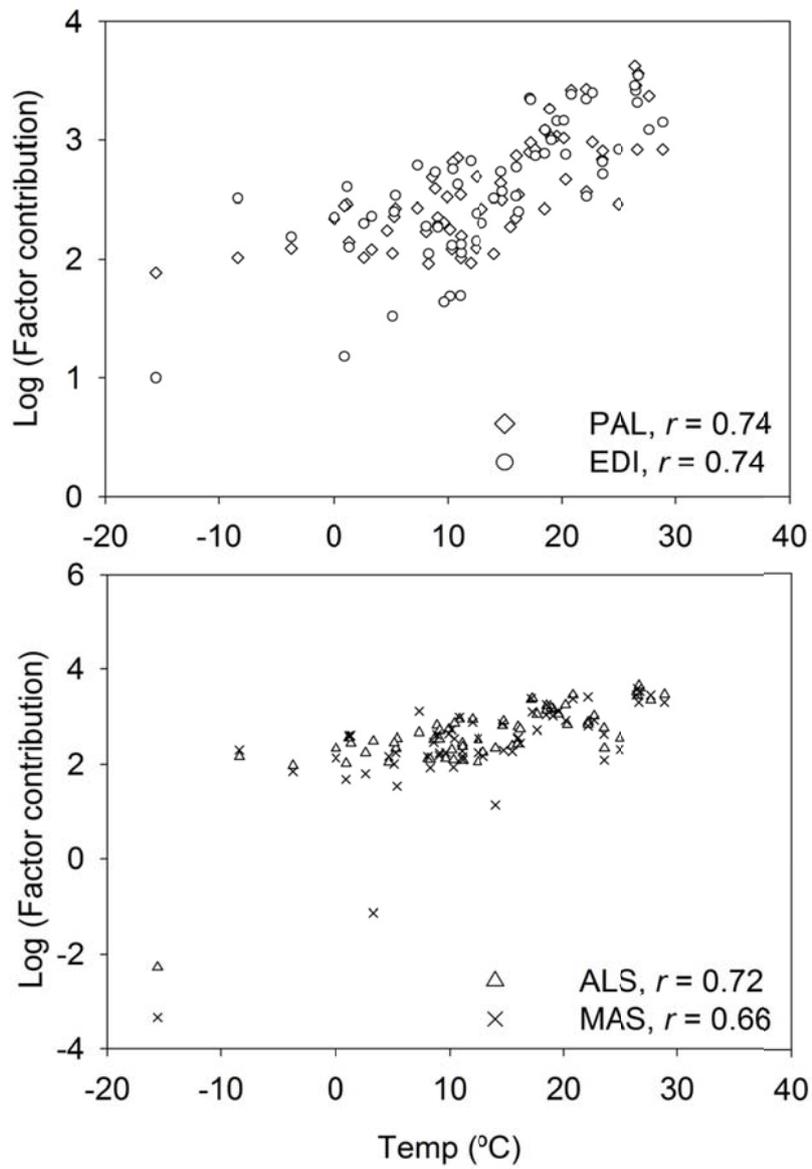


Figure 3.S1 Correlations of median daily contribution of Factor 2 versus ambient temperature at the four sites.

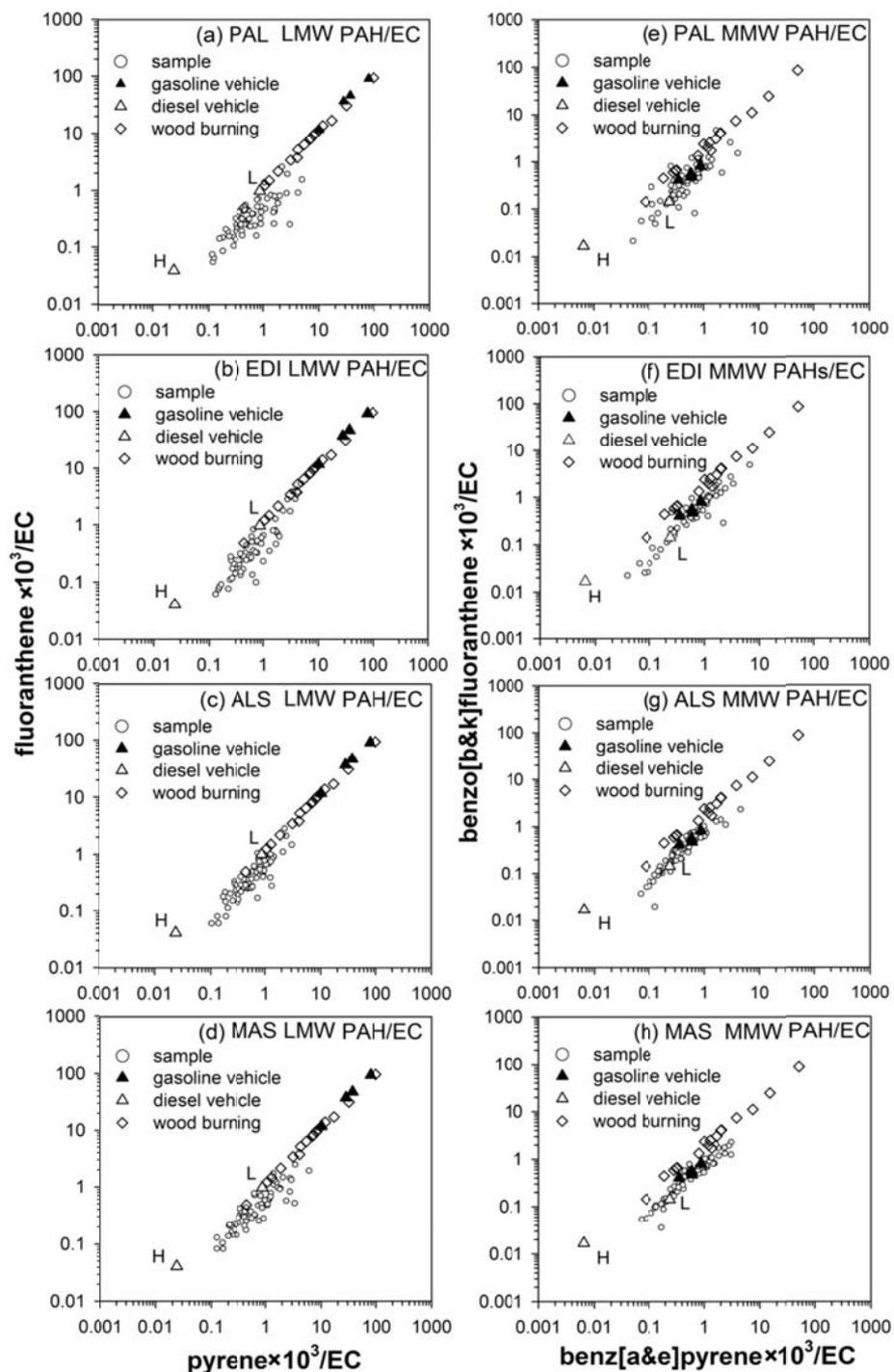


Figure 3.S2 Ratio-ratio plots for two low molecular weight (LMW) PAHs (a-d) and two medium MW PAHs (e-h) normalized by EC at each of the four sites. Black and blank triangles represent gasoline and diesel (L light duty, H heavy duty) source profiles respectively. Blank diamonds represent wood burning source profiles. Source profile data are obtained from Zielinska et al. (1998).

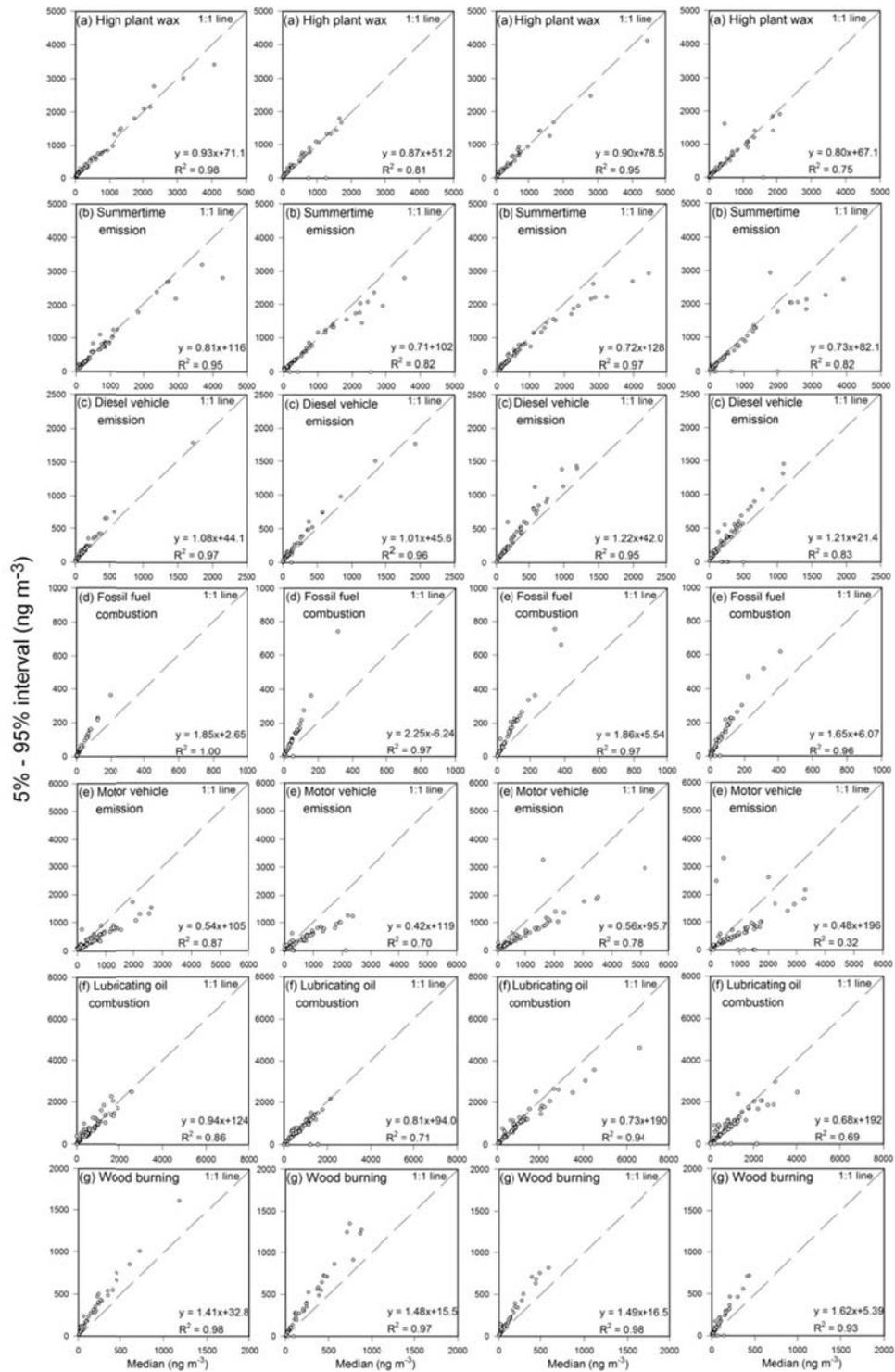


Figure 3.S3 Comparison of median factor contributions versus uncertainties represented by 5<sup>th</sup> – 95<sup>th</sup> percentile range of factor contribution from bootstrap solutions for each factor, separated by site.

### S4-1 Factor 1 High plant wax

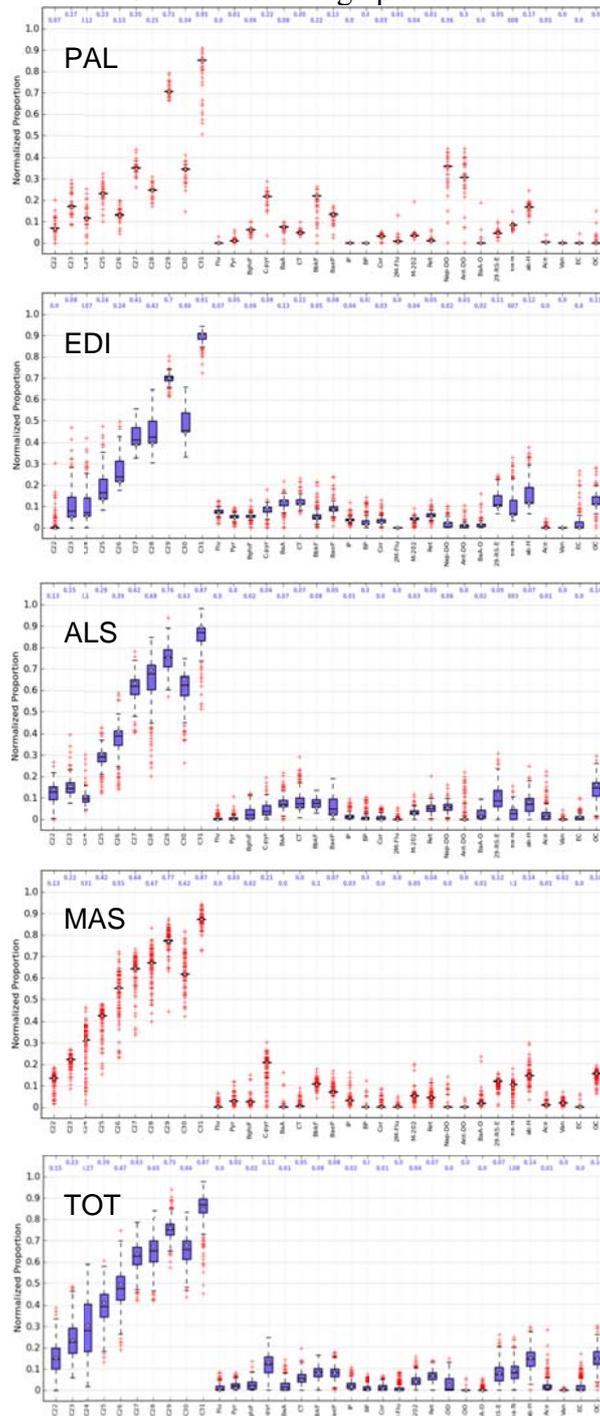


Figure 3.S4 Comparison of normalized factor profiles from PMF bootstrap solutions using site specific data sets versus those obtained from pooled data set (TOT). The boxes depict the median (dark line), inner quartile range (blue box), lower quartile –  $1.5 \times$  IQR and upper quartile +  $1.5 \times$  IQR (whisker), base case (blank star) and outliers (red plus). Values of the medians are labeled on the top of each plot.

### S4-2 Factor 2 Summertime emission

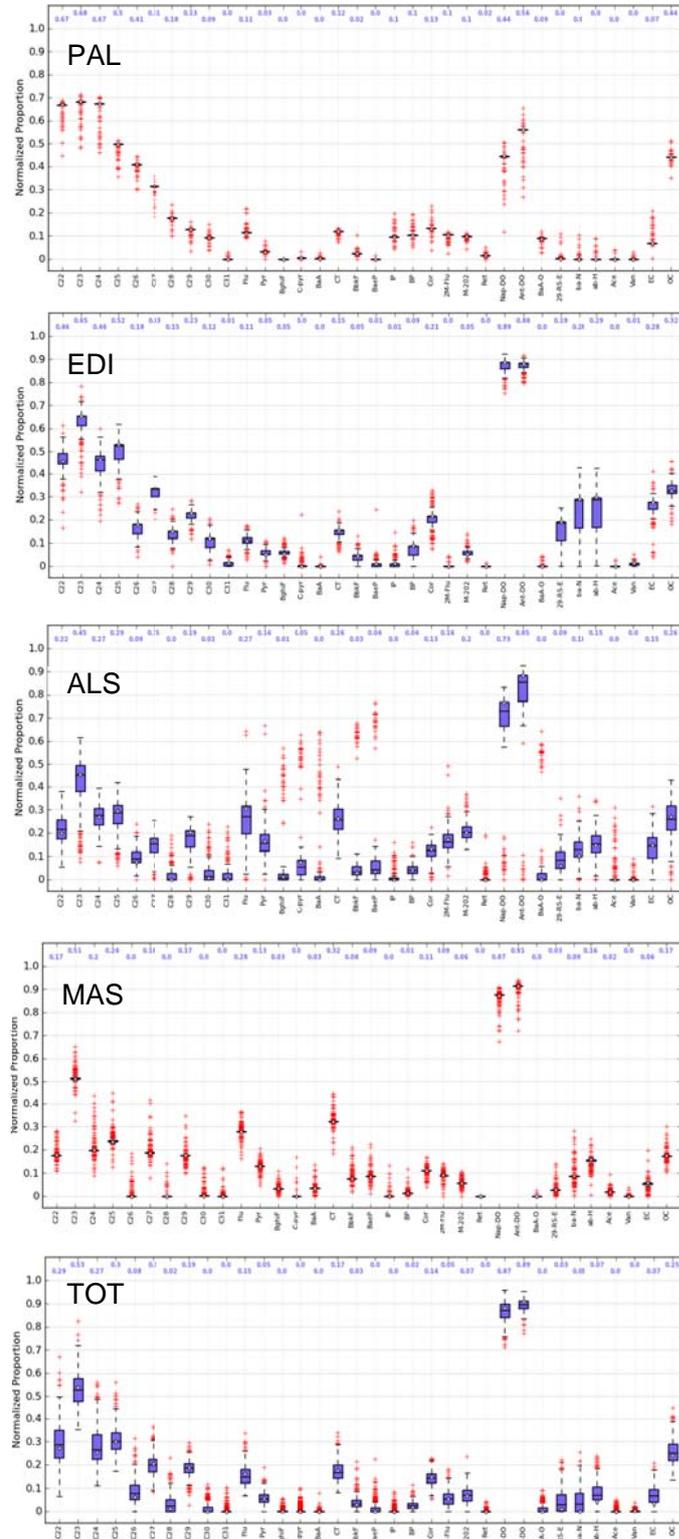
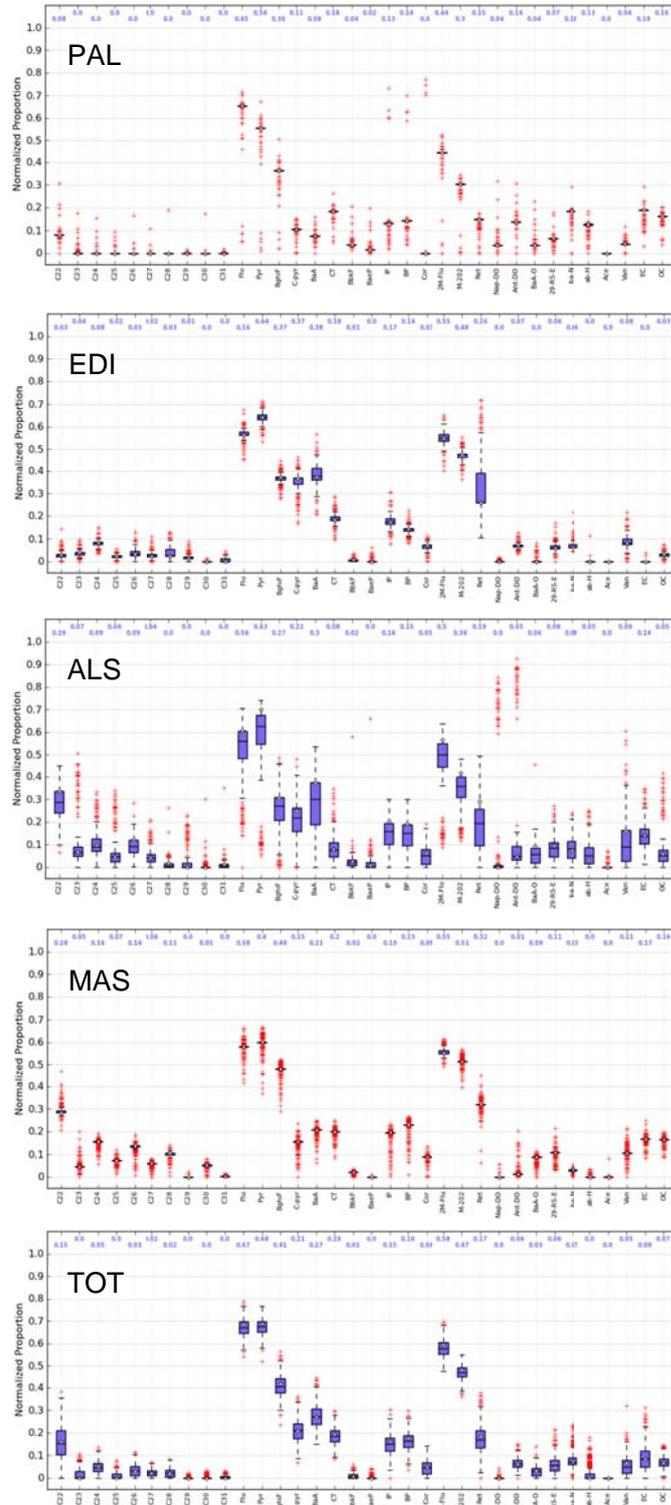


Figure 3.S4 Continued.

### S4-3 Factor 3 Diesel vehicle emission





### S4-5 Factor 5 Motor vehicle emission

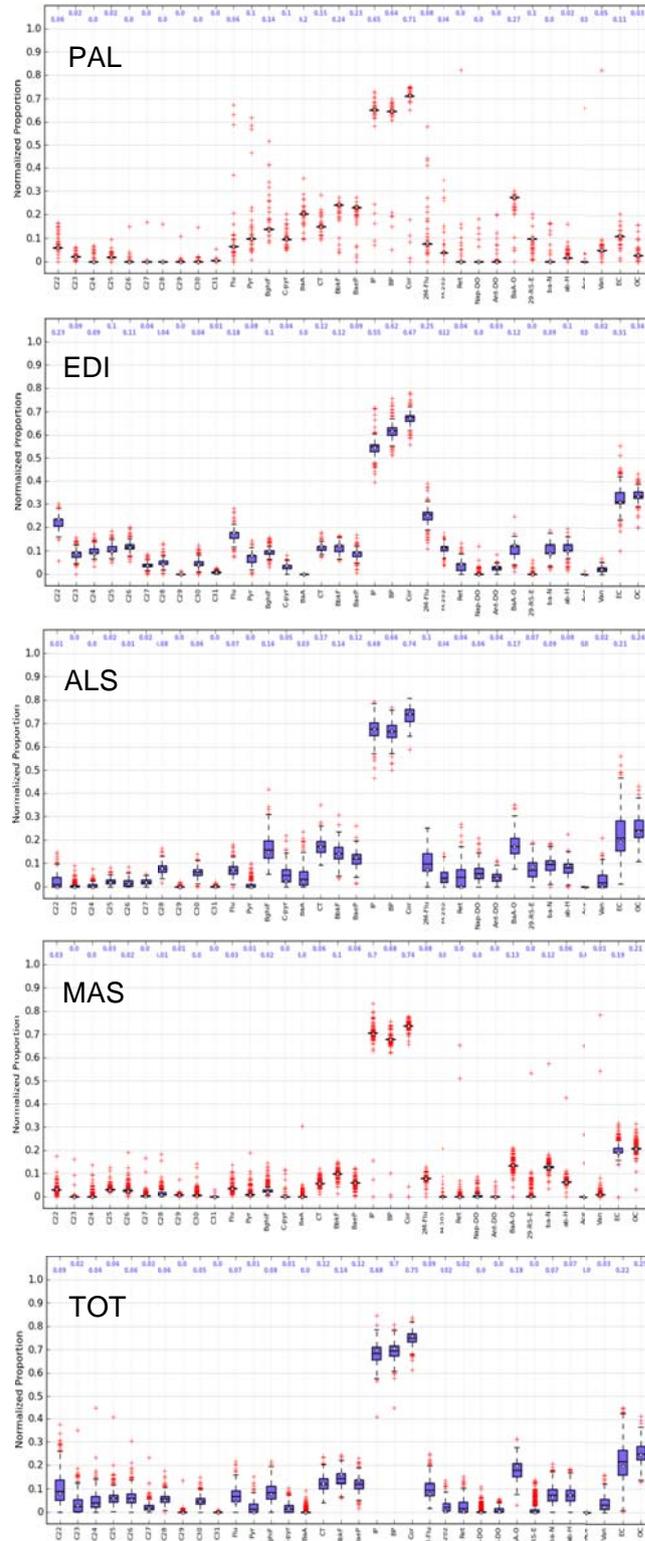


Figure 3.S4 Continued.

### S4-6 Factor 6 Lubricating oil combustion

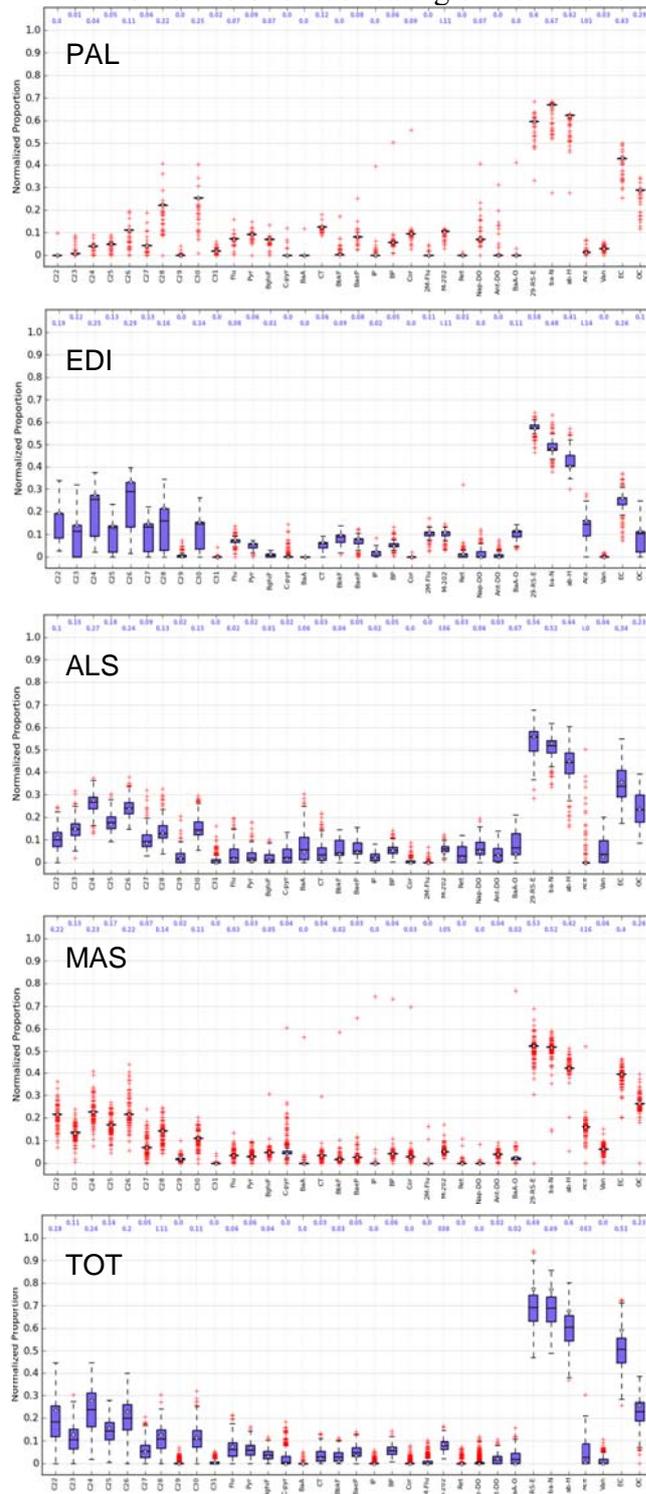


Figure 3.S4 Continued.



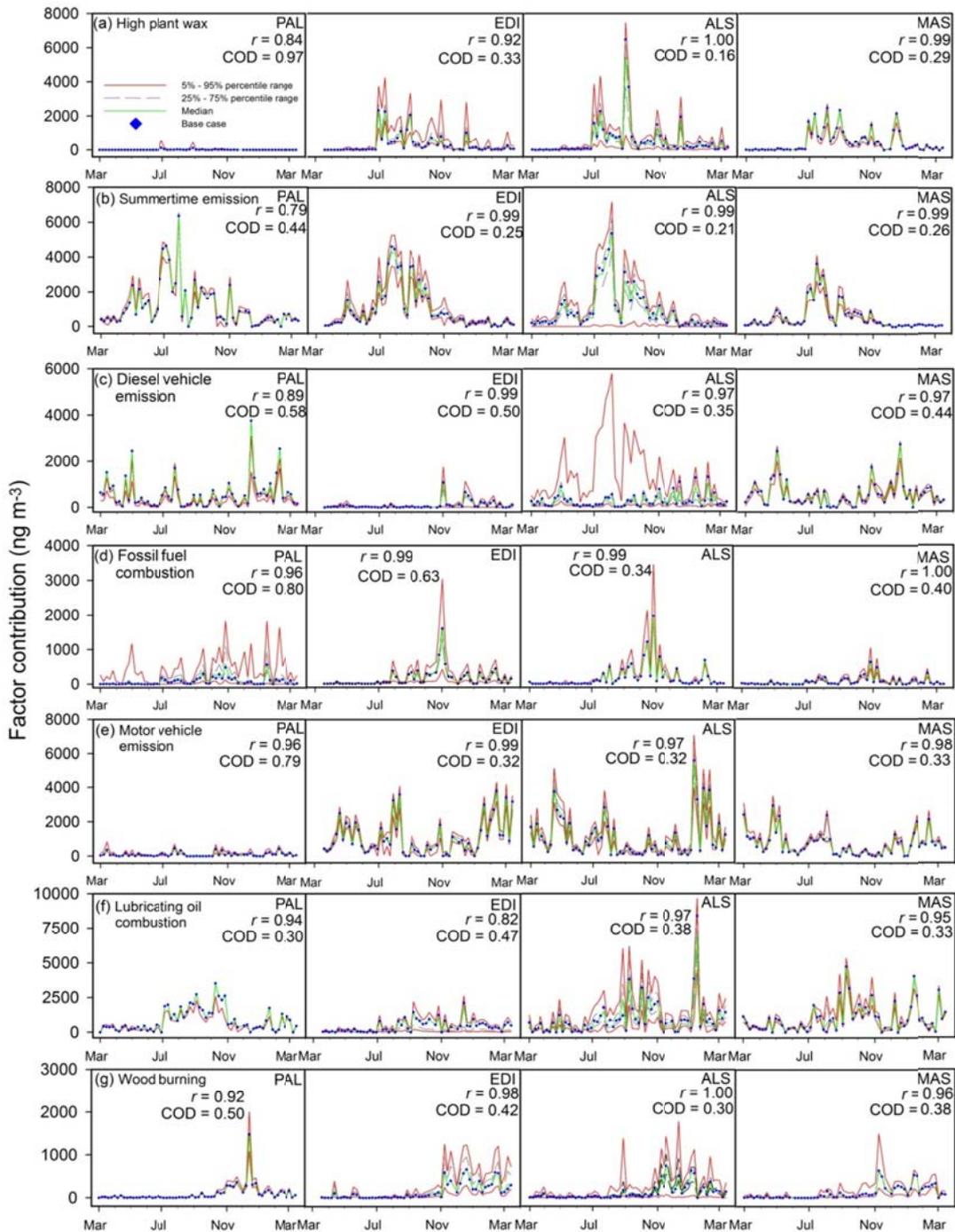


Figure 3.S5 Factor contribution time series from PMF bootstrap solutions for site specific data. Each plot shows the time series of daily factor contribution for base case (blue diamond), median (green line), and two bands based on the empirical quantiles of the bootstrap solutions (25<sup>th</sup> – 75<sup>th</sup> percentiles, gray dashed line; 5<sup>th</sup> – 95<sup>th</sup> percentiles, red solid line). Correlation coefficient ( $r$ ) and coefficient of divergence (COD) listed show the relationship between median factor contributions derived from site specific data sets and those from pooled data set.

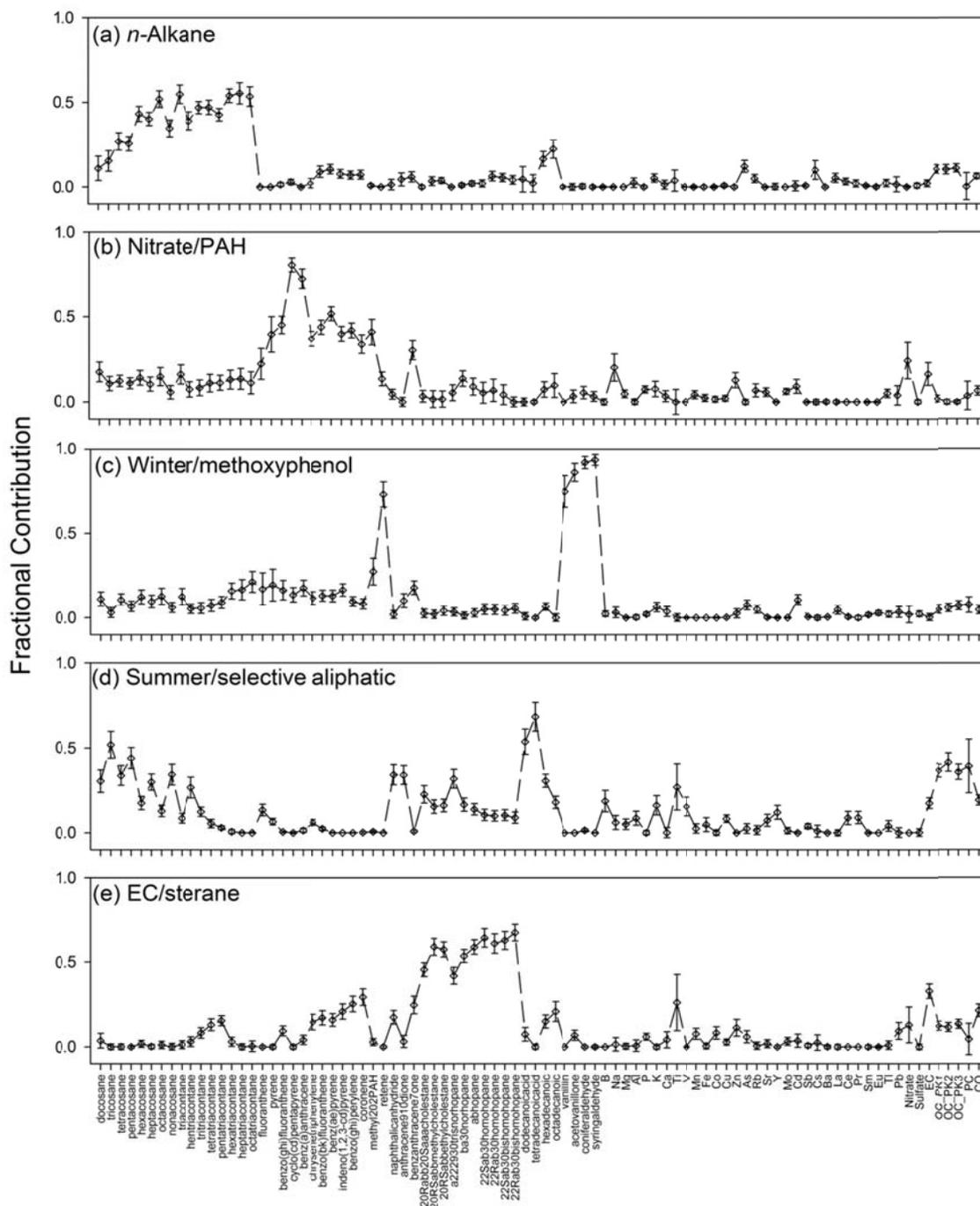


Figure 4.S1 Median normalized factor profiles derived from PMF bootstrap solutions using Composite data set including all species. The whiskers represent standard deviation.



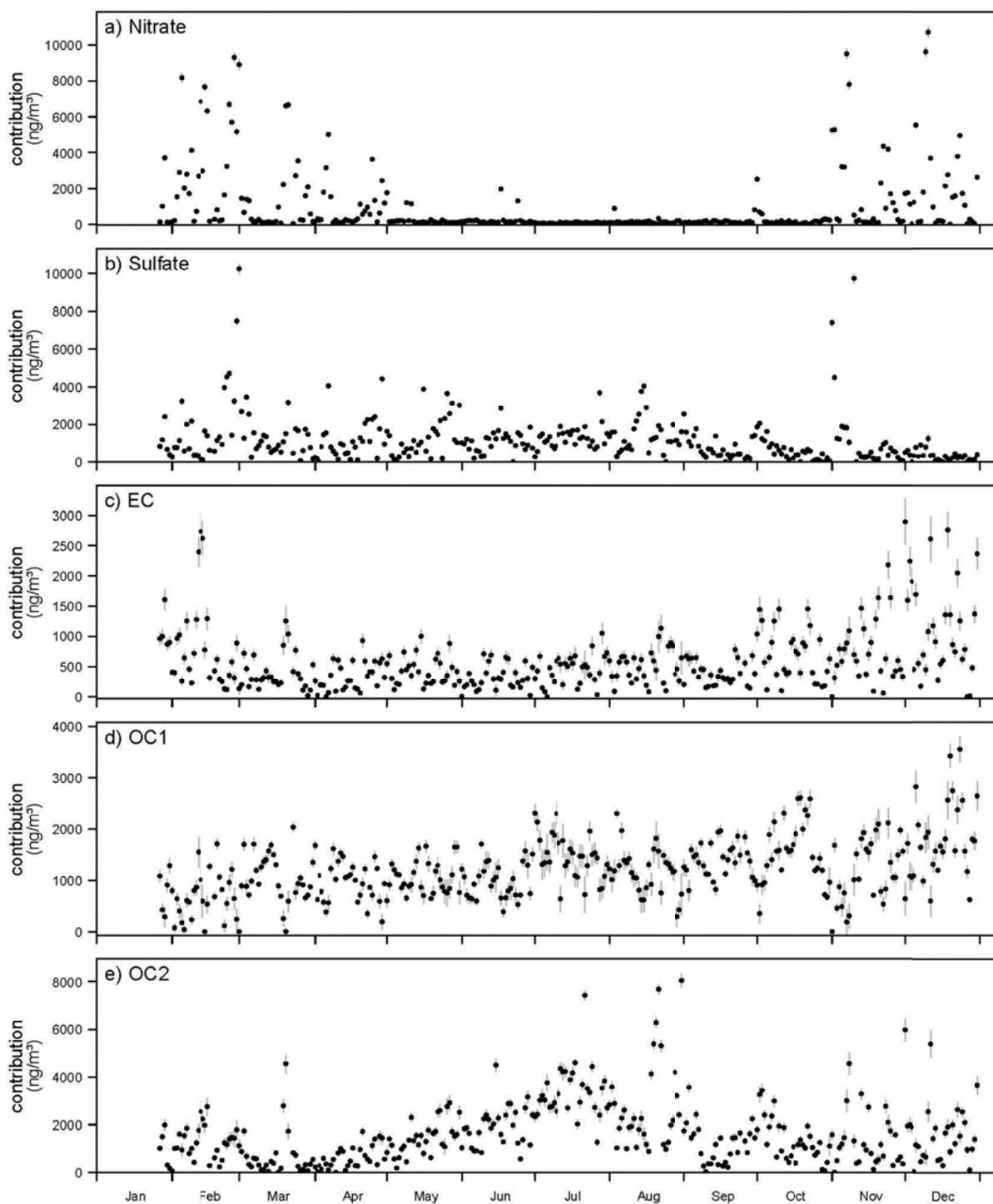


Figure 4.S2 Median PMF factor contributions for the Bulk data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD).

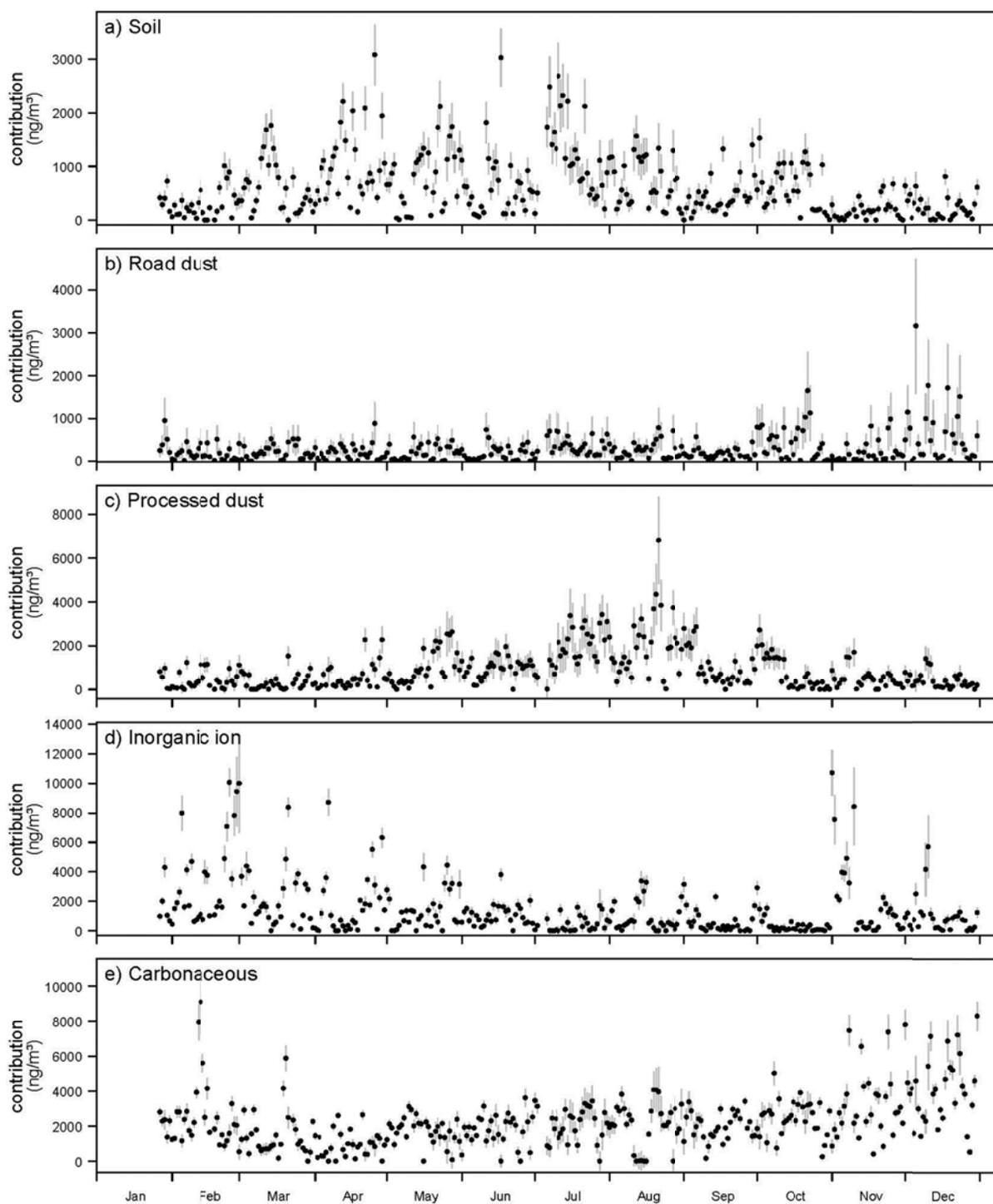


Figure 4.S3 Median PMF factor contributions for the WSE+Bulk data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD).

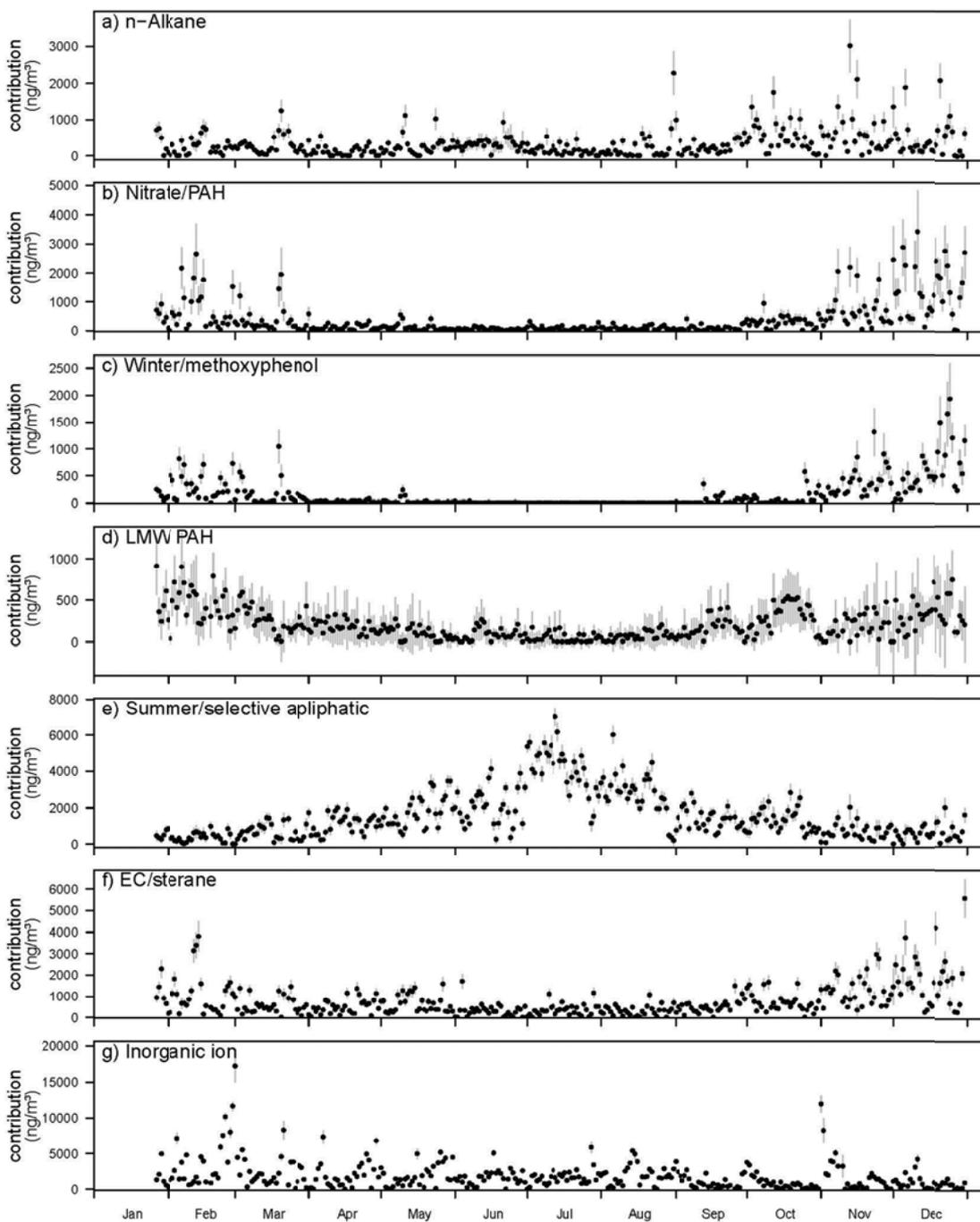


Figure 4.S4 Median PMF factor contributions for the OMM+Bulk data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD).

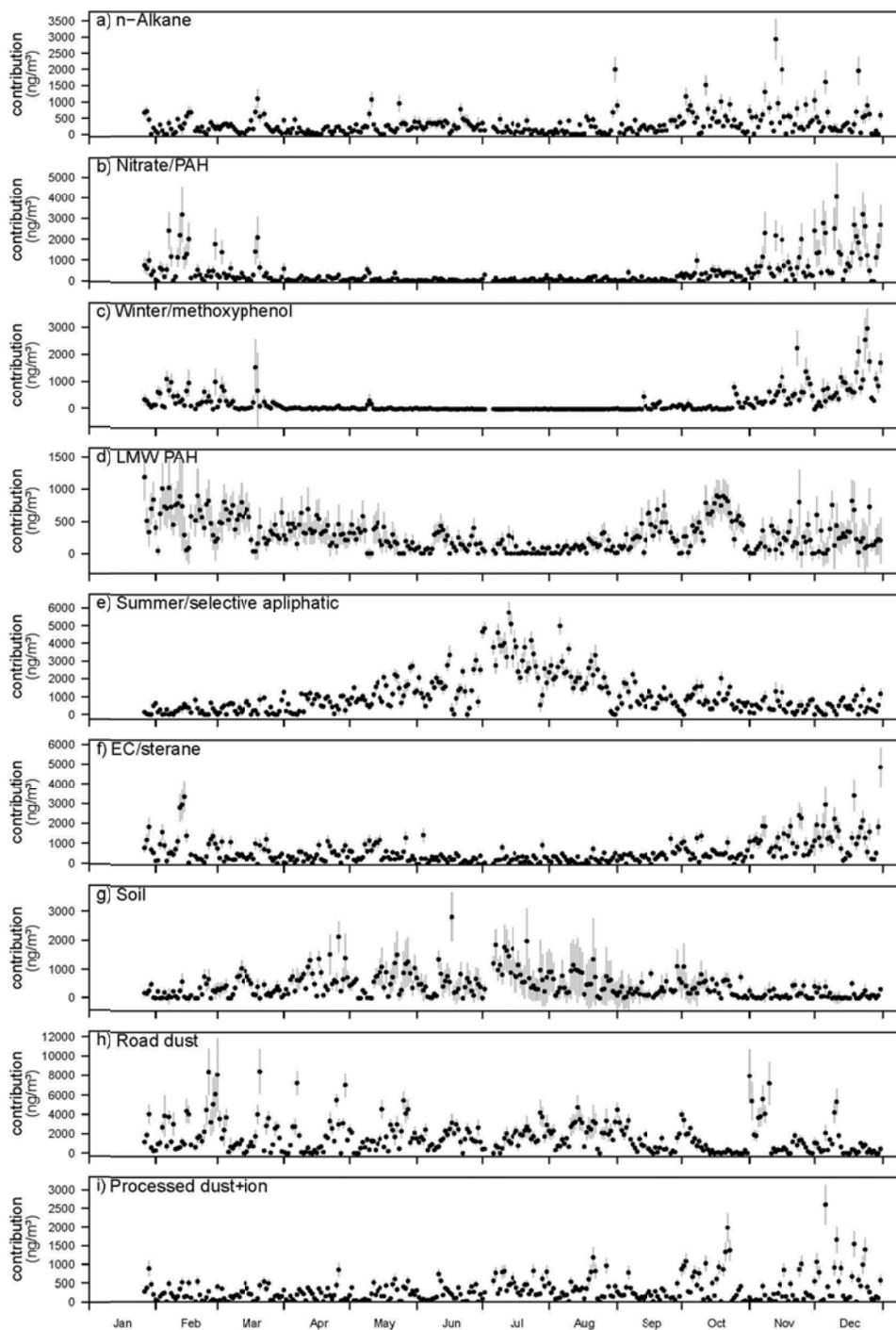


Figure 4.S5 Median PMF factor contributions for the Composite data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD).

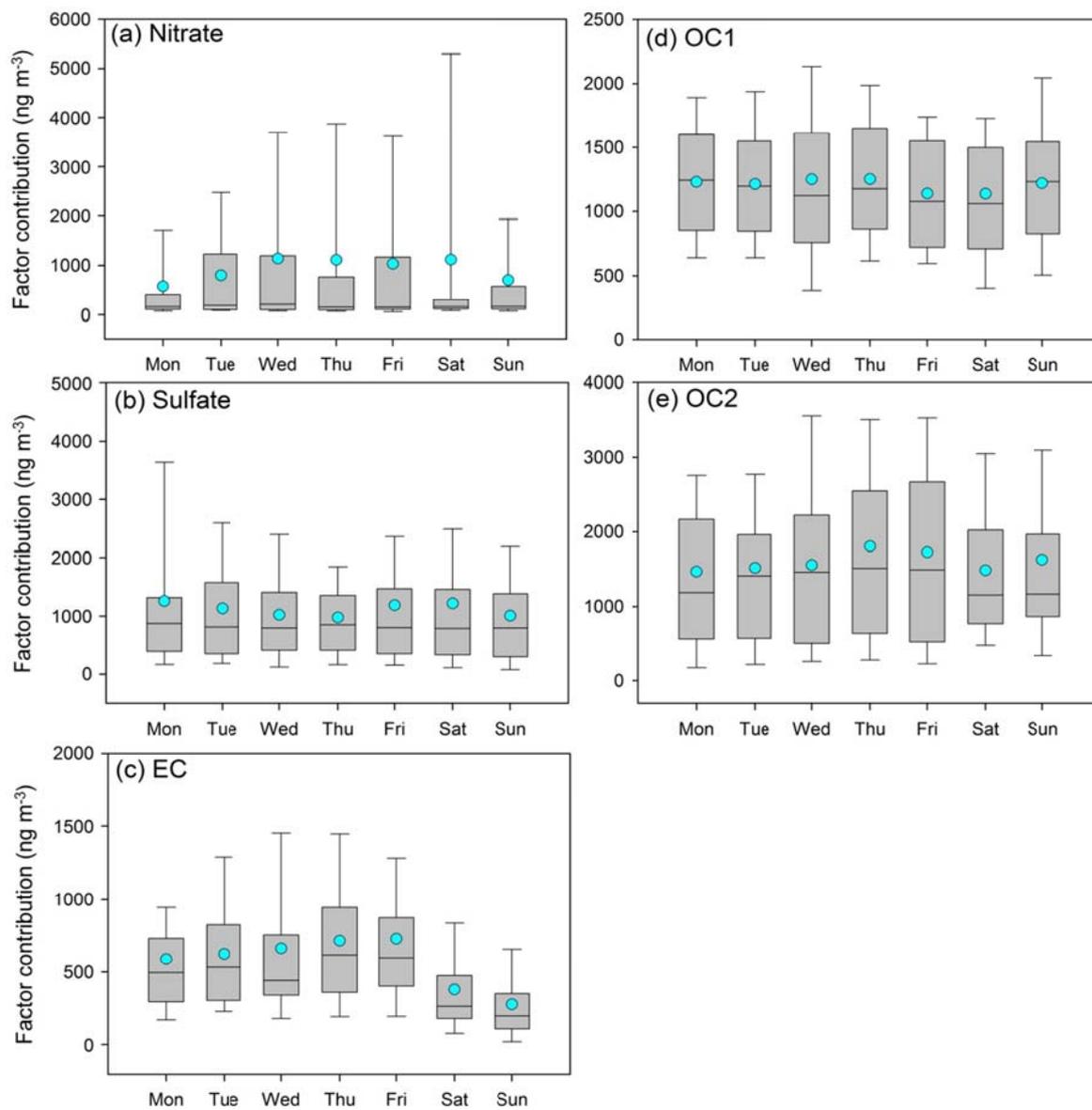


Figure 4.S6 Day of the week box plots for factor contributions from the Bulk data set. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle).

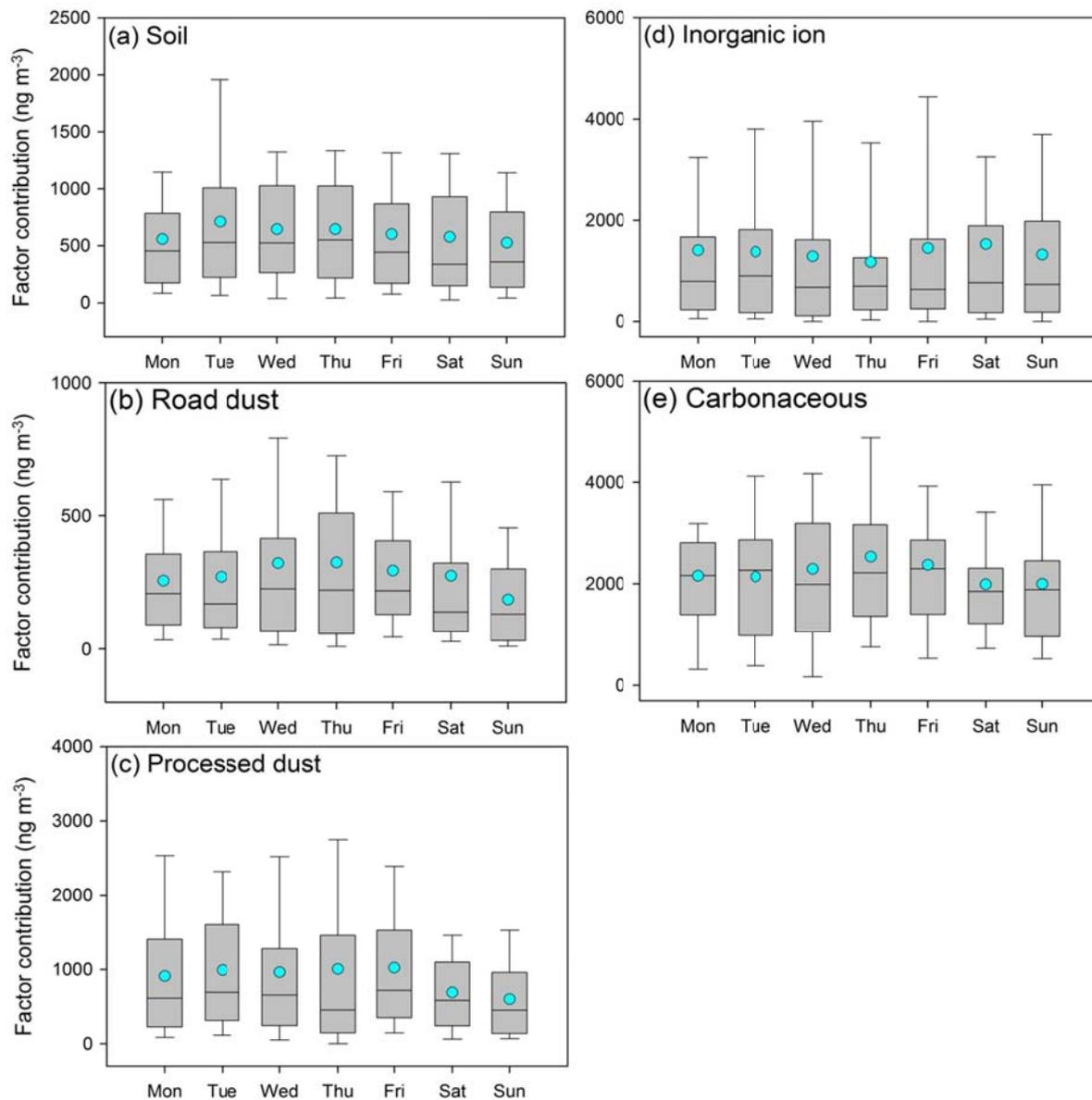


Figure 4.S7 Day of the week box plots for factor contributions from the WSE+Bulk data set. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle).

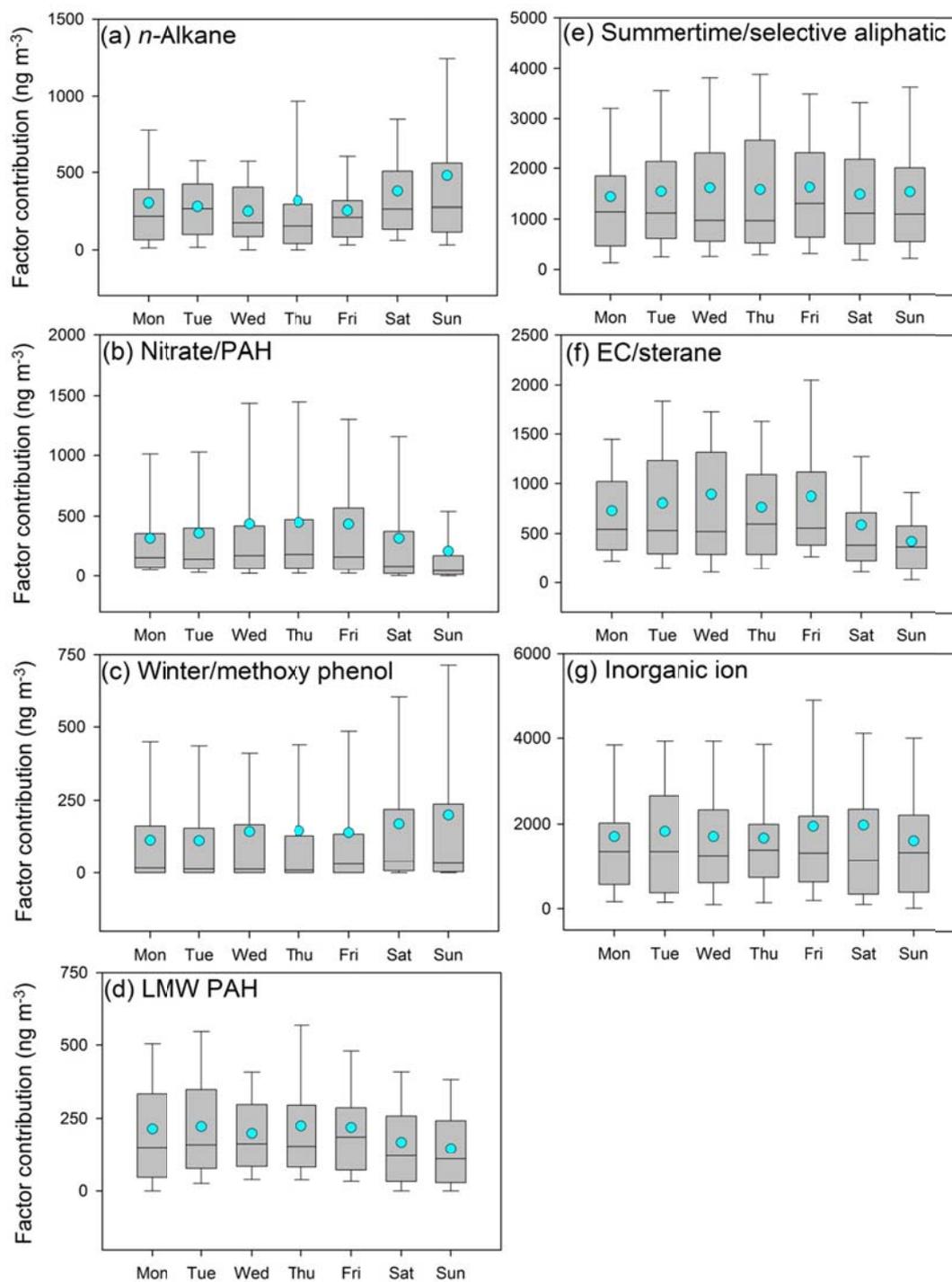


Figure 4.S8 Day of the week box plots for factor contributions from the OMM+Bulk data set. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle).

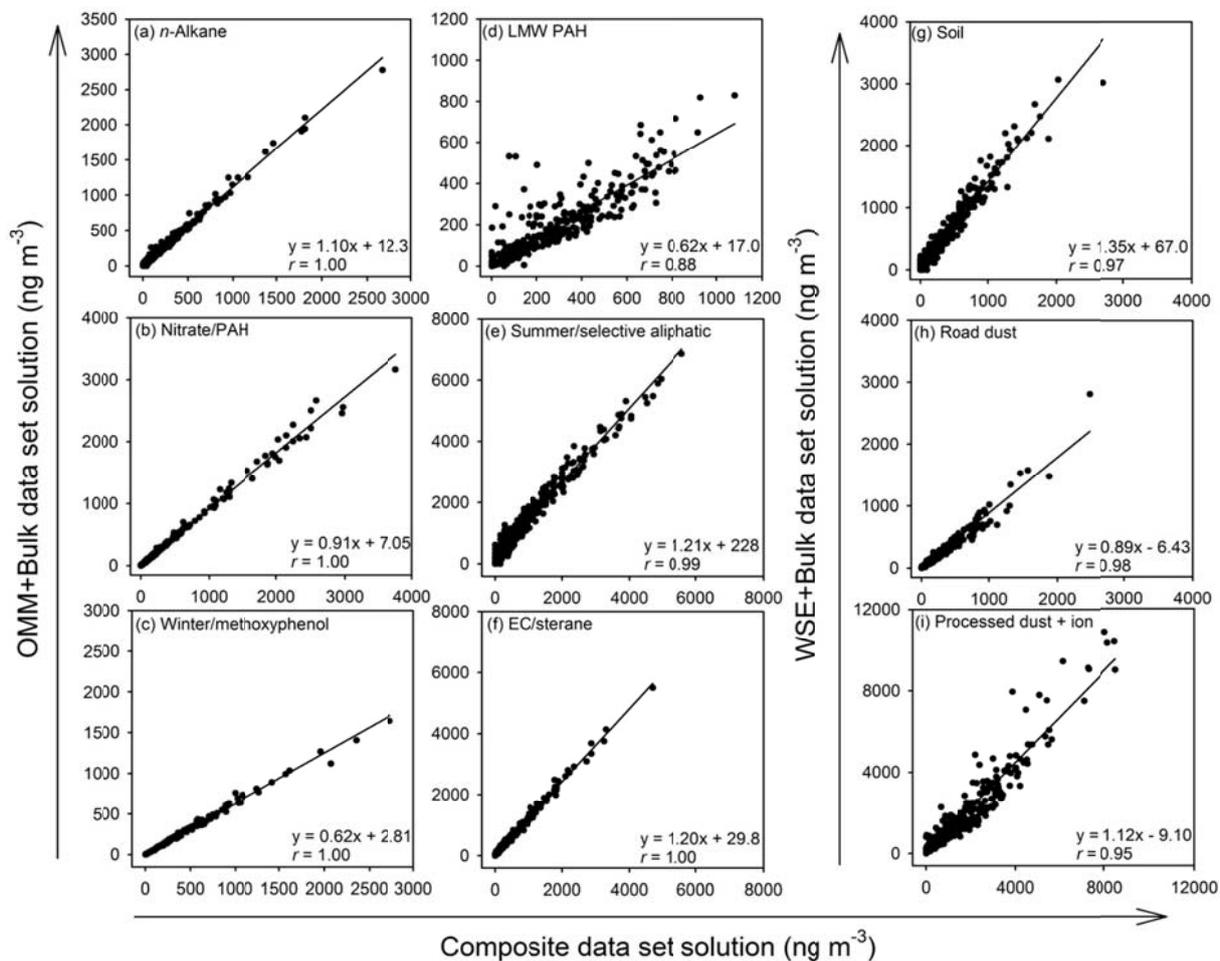


Figure 4.S9 Comparisons of median factor contribution time series between factors resolved by the Composite data set and the OMM+Bulk and WSE+Bulk data sets.

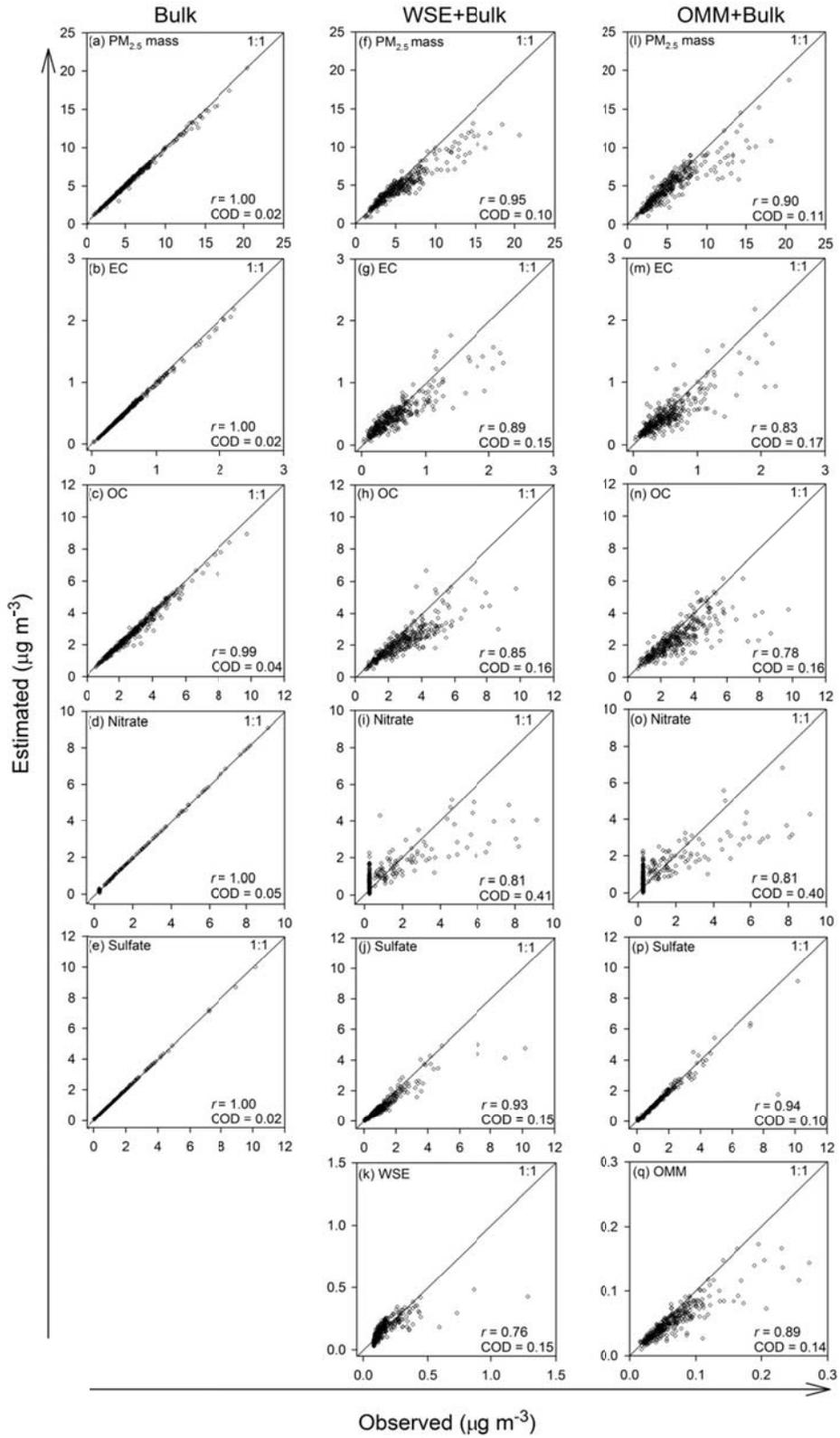


Figure 4.S10 Comparisons of PMF estimated vs. observed time series of PM<sub>2.5</sub> components.

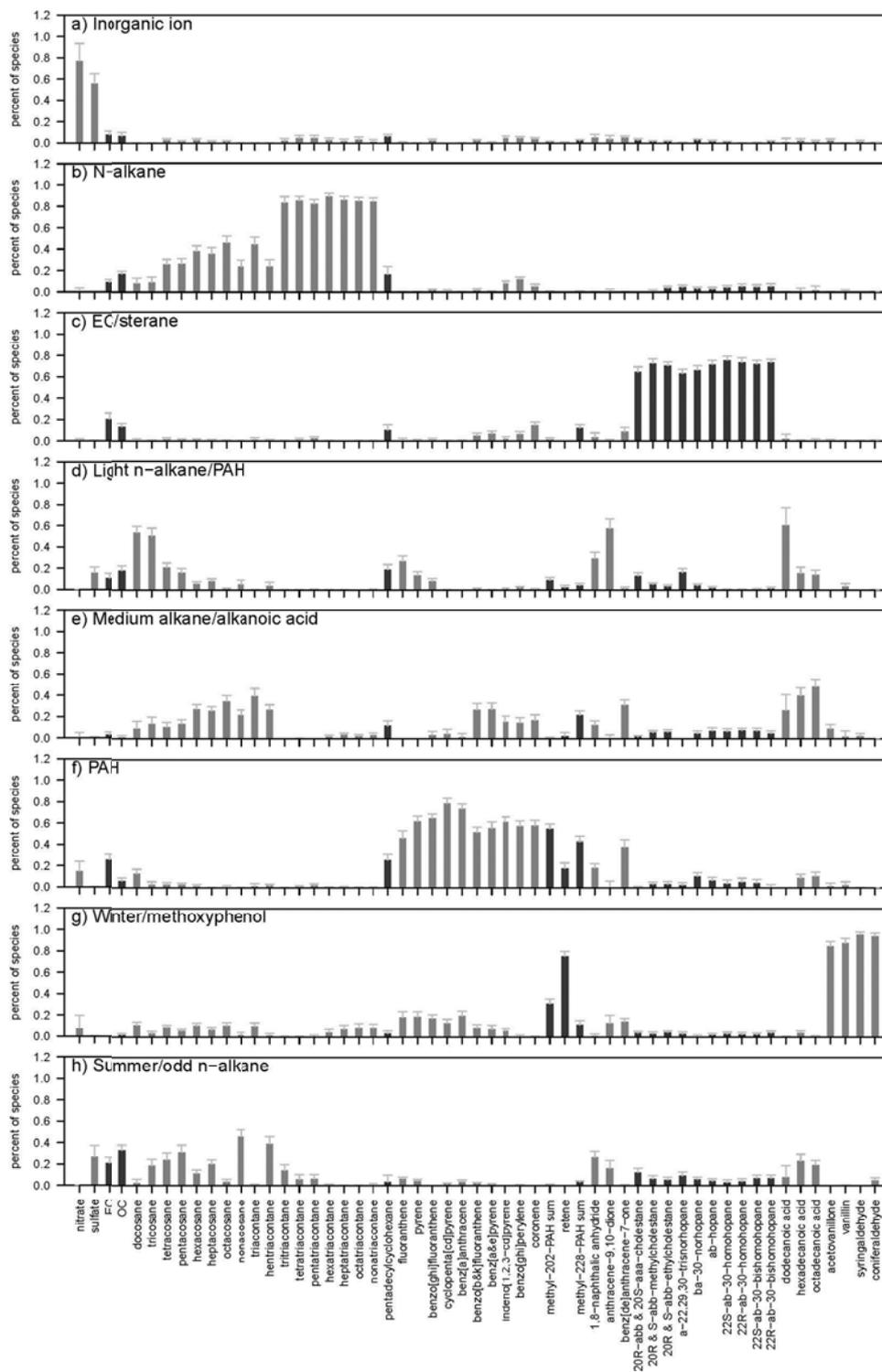


Figure 5.S1 Median PMF factor profiles for the full data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations.

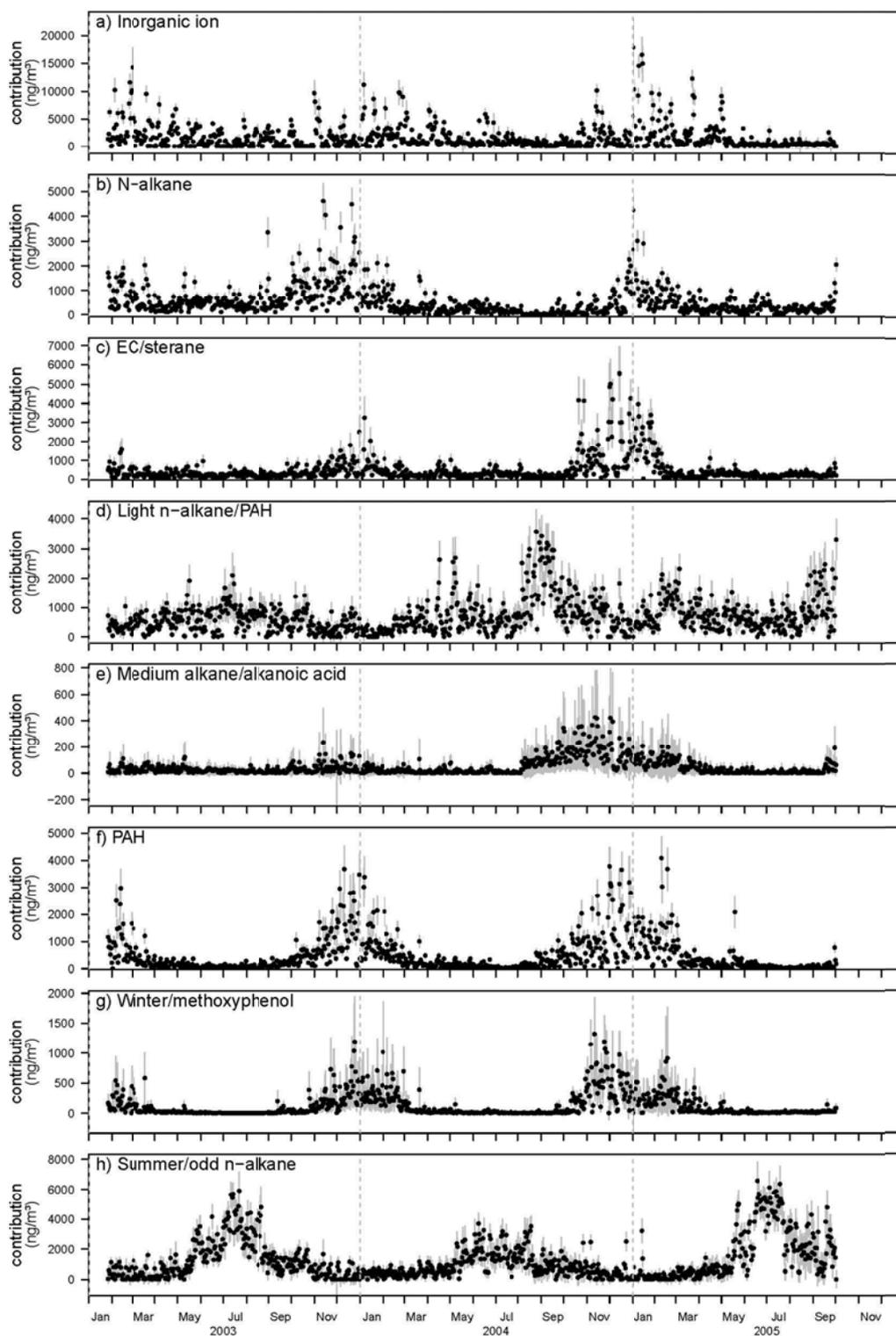


Figure 5.S2 Median PMF factor contributions for the full data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD).

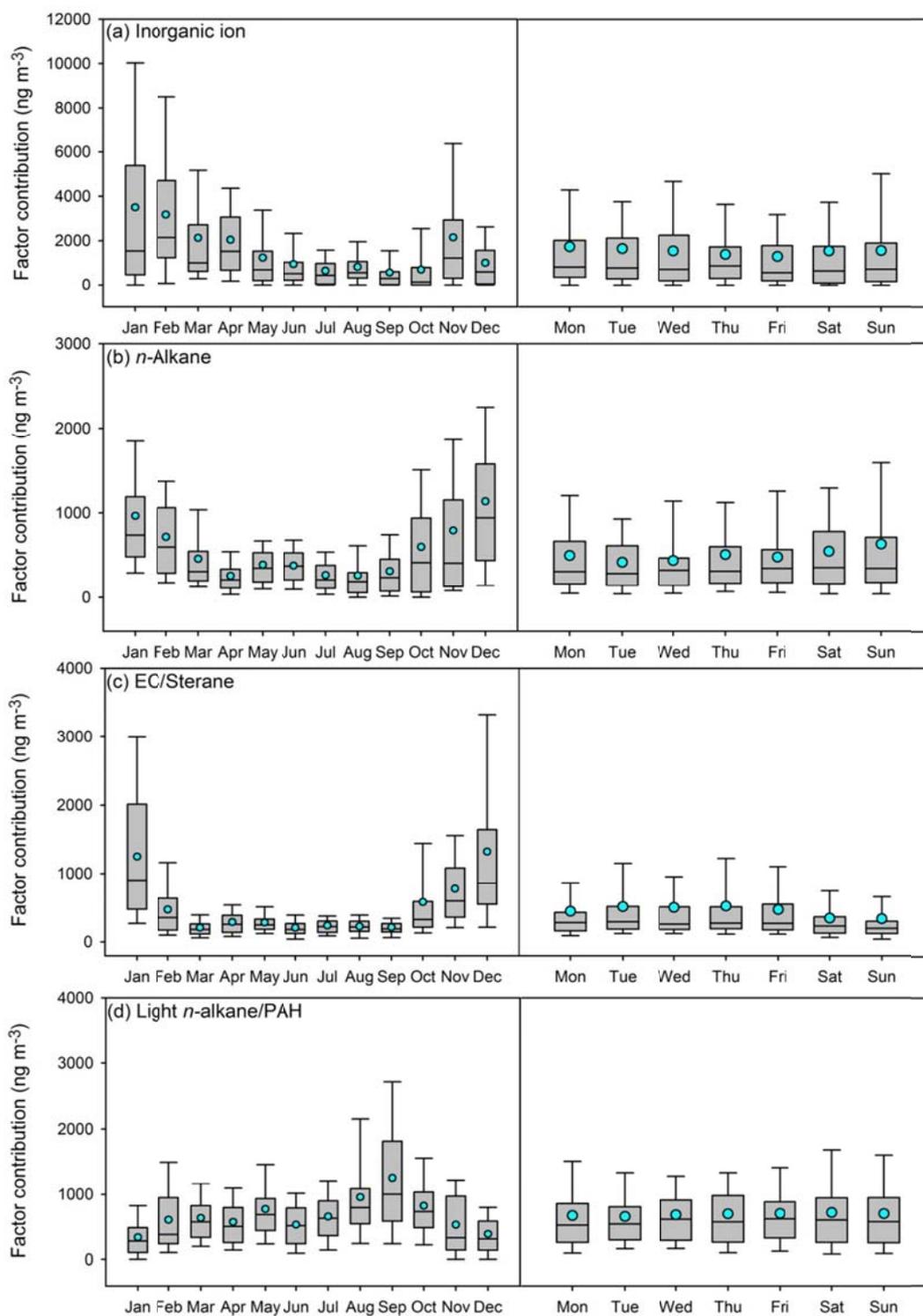


Figure 5.S3 Day of the week and seasonal box plots for PMF factor contributions of the full data set solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle).

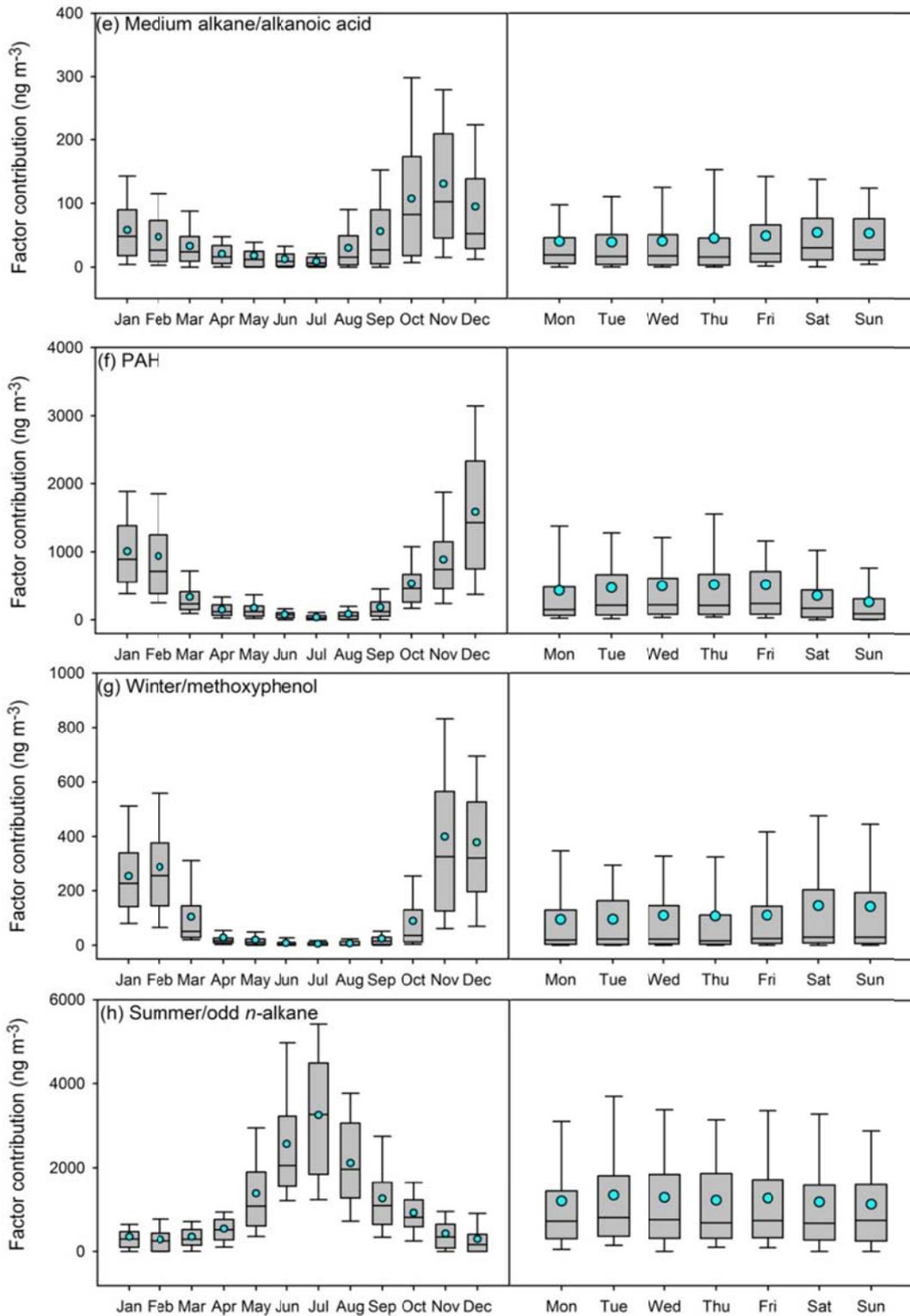


Table 5.S3 Continued

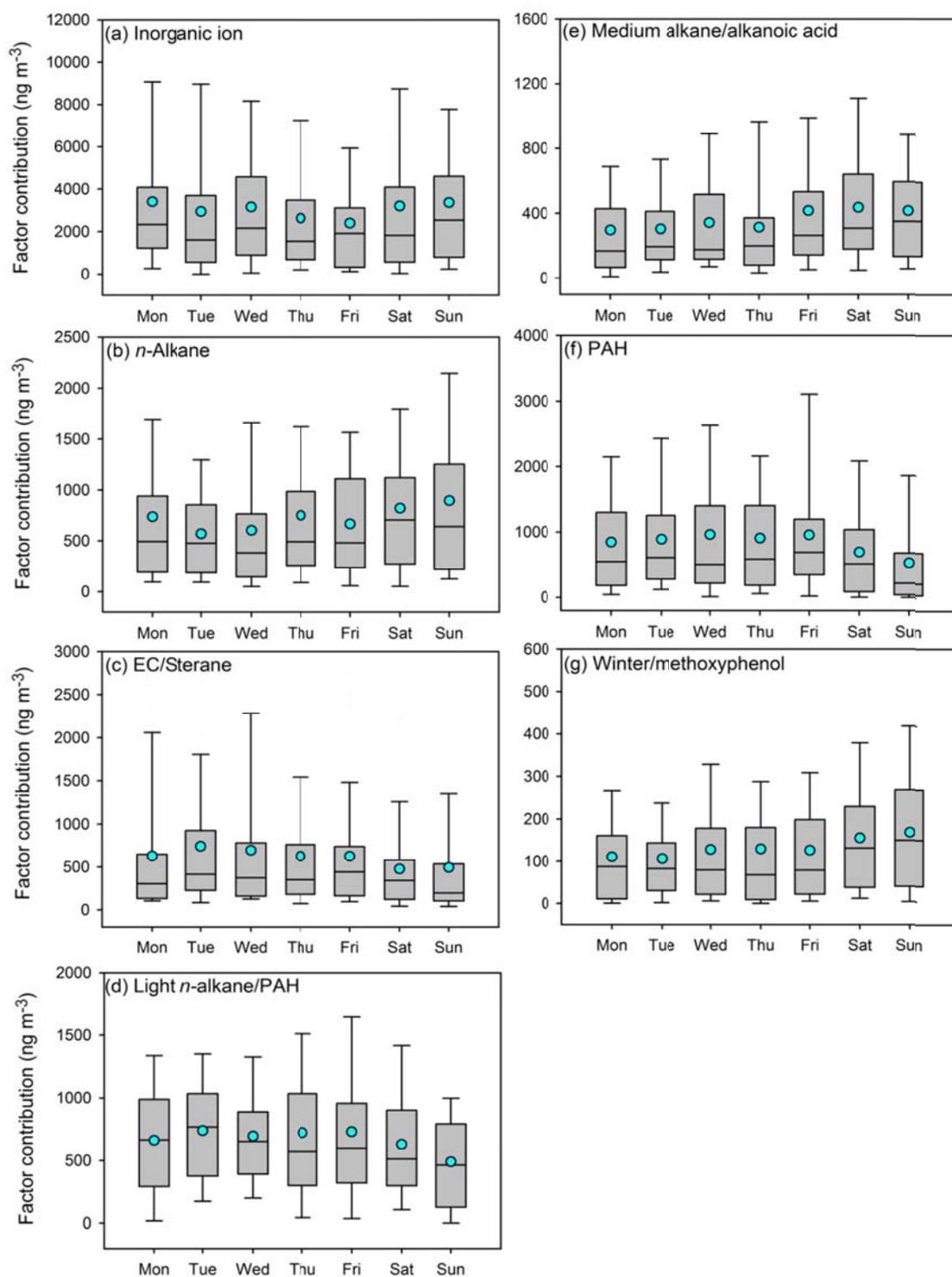


Figure 5.S4 Day of the week box plots for PMF factor contributions of the cold period solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle).

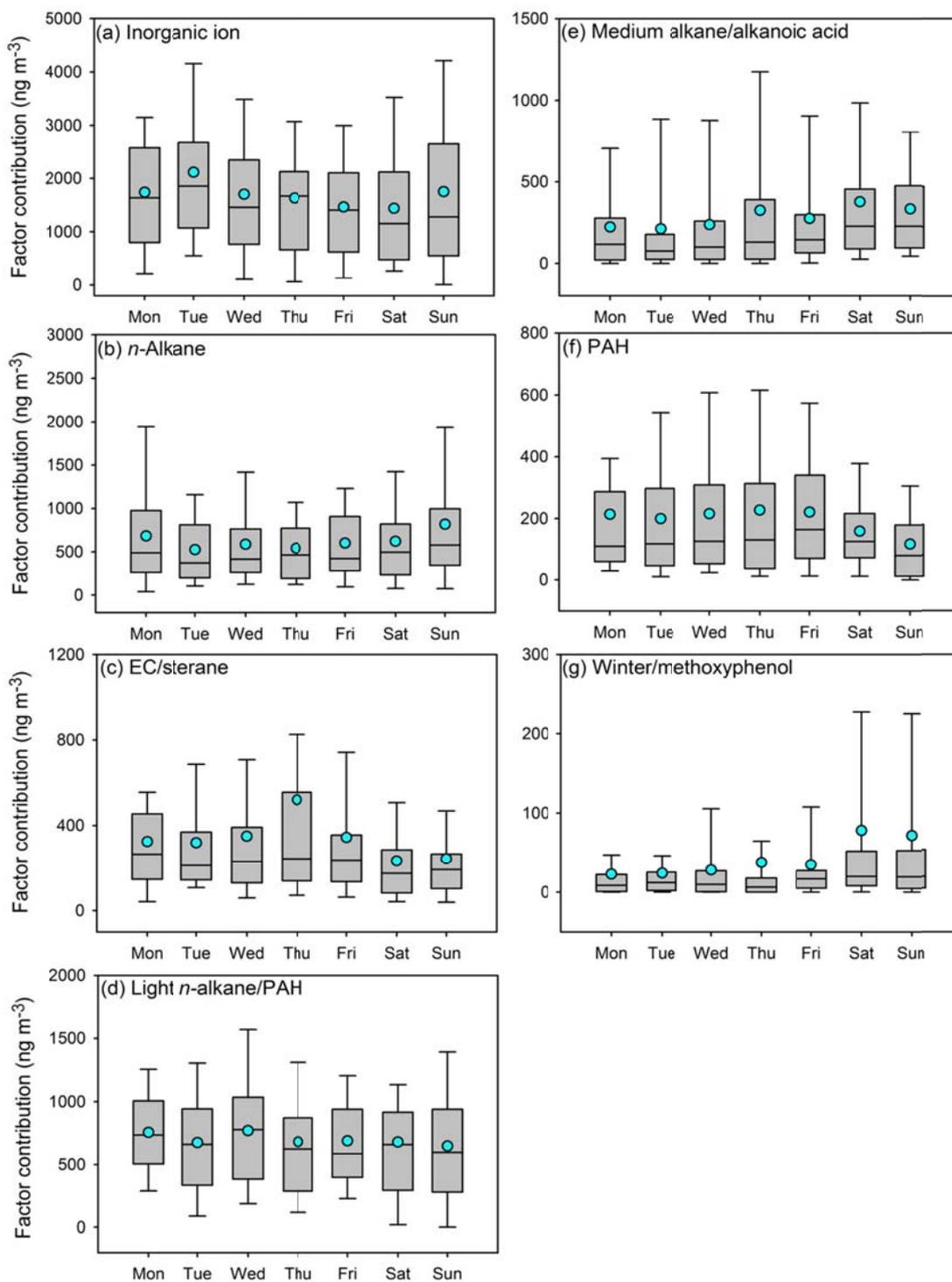


Figure 5.S5 Day of the week box plots for PMF factor contributions of the warm period solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle).

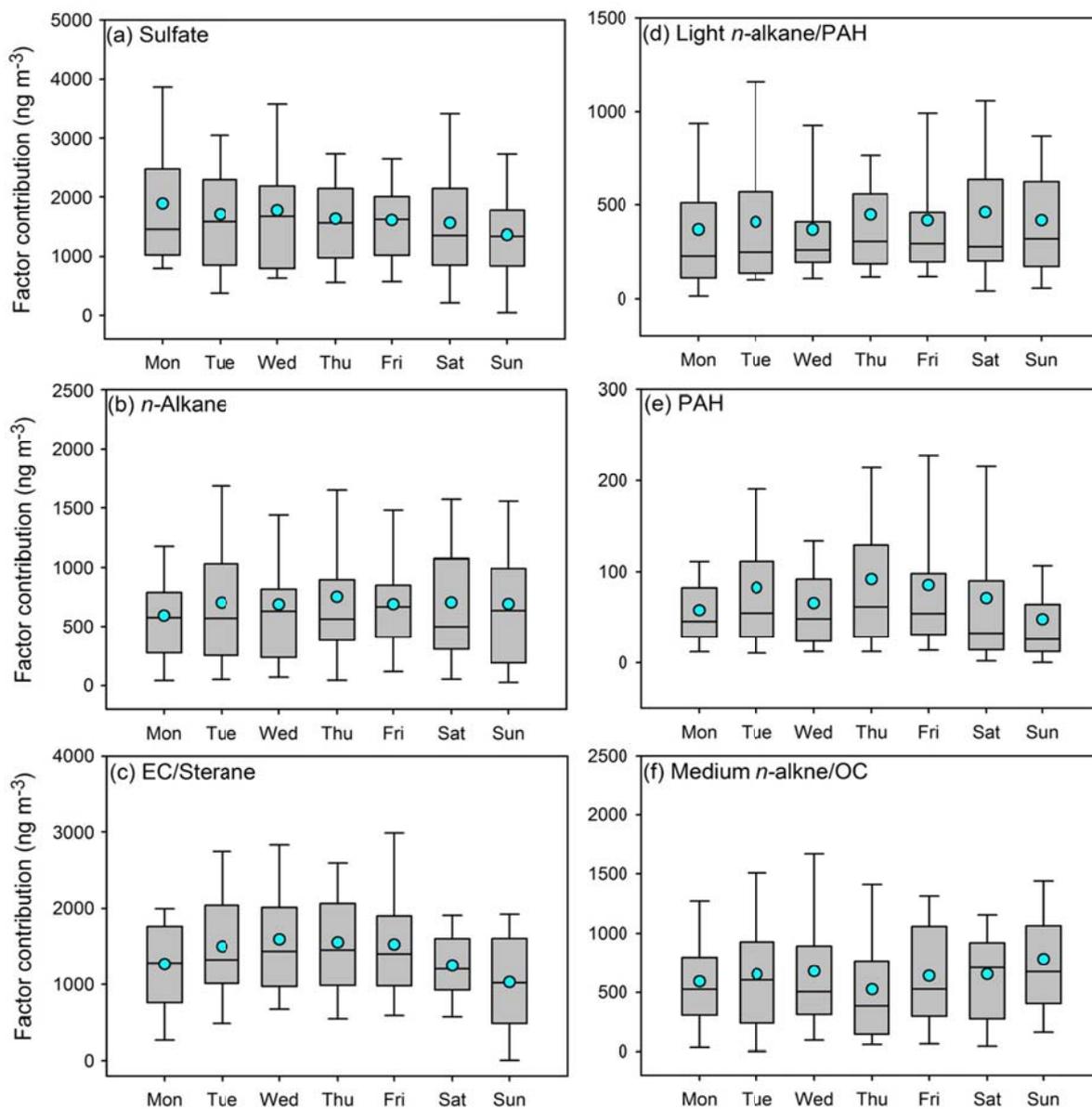


Figure 5.S6 Day of the week box plots for PMF factor contributions of the hot period solution. The boxes depict the median (dark line in the box), inner quartile range (gray box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (cyan circle).

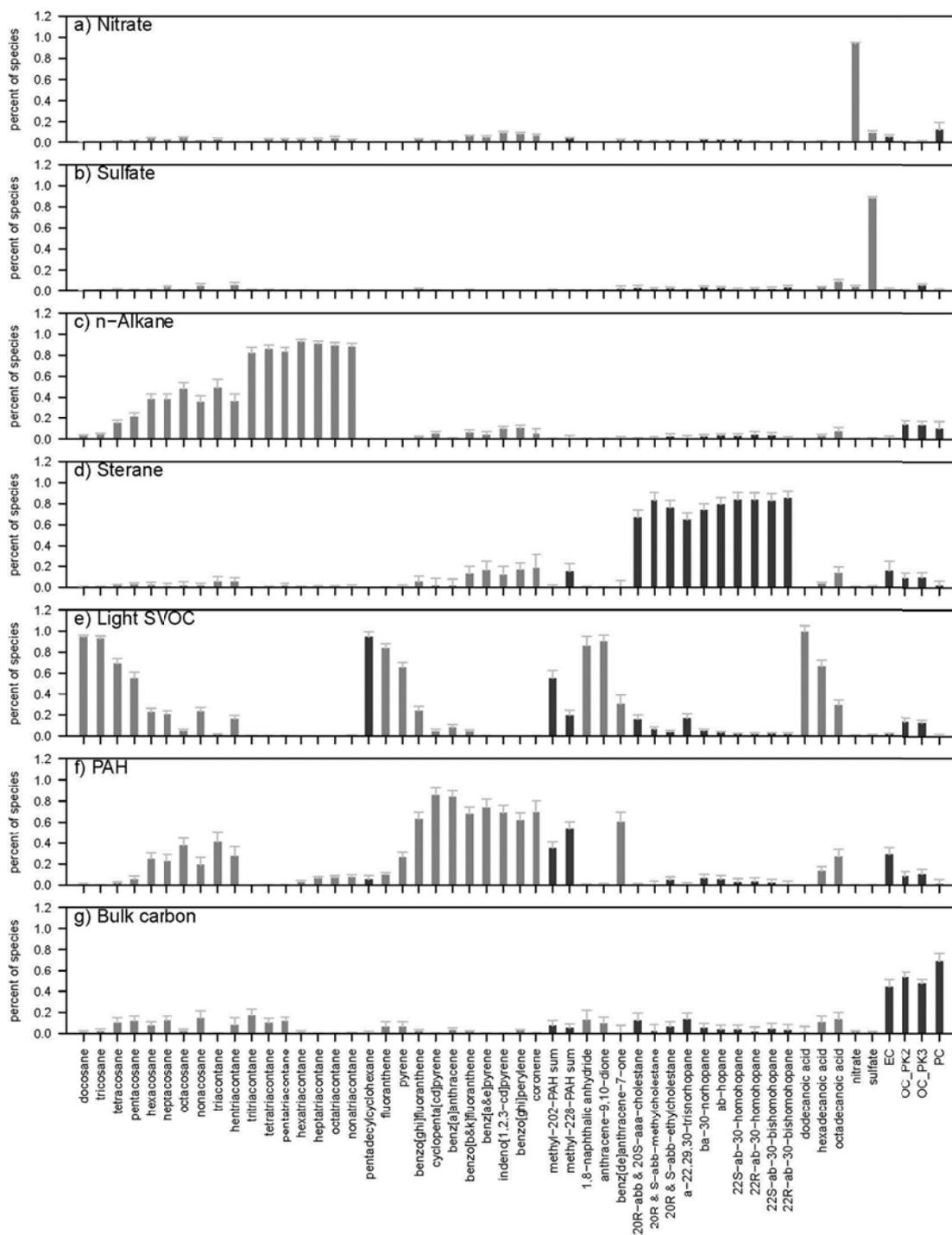


Figure 6.S1 Median PMF factor profiles for the full data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations.

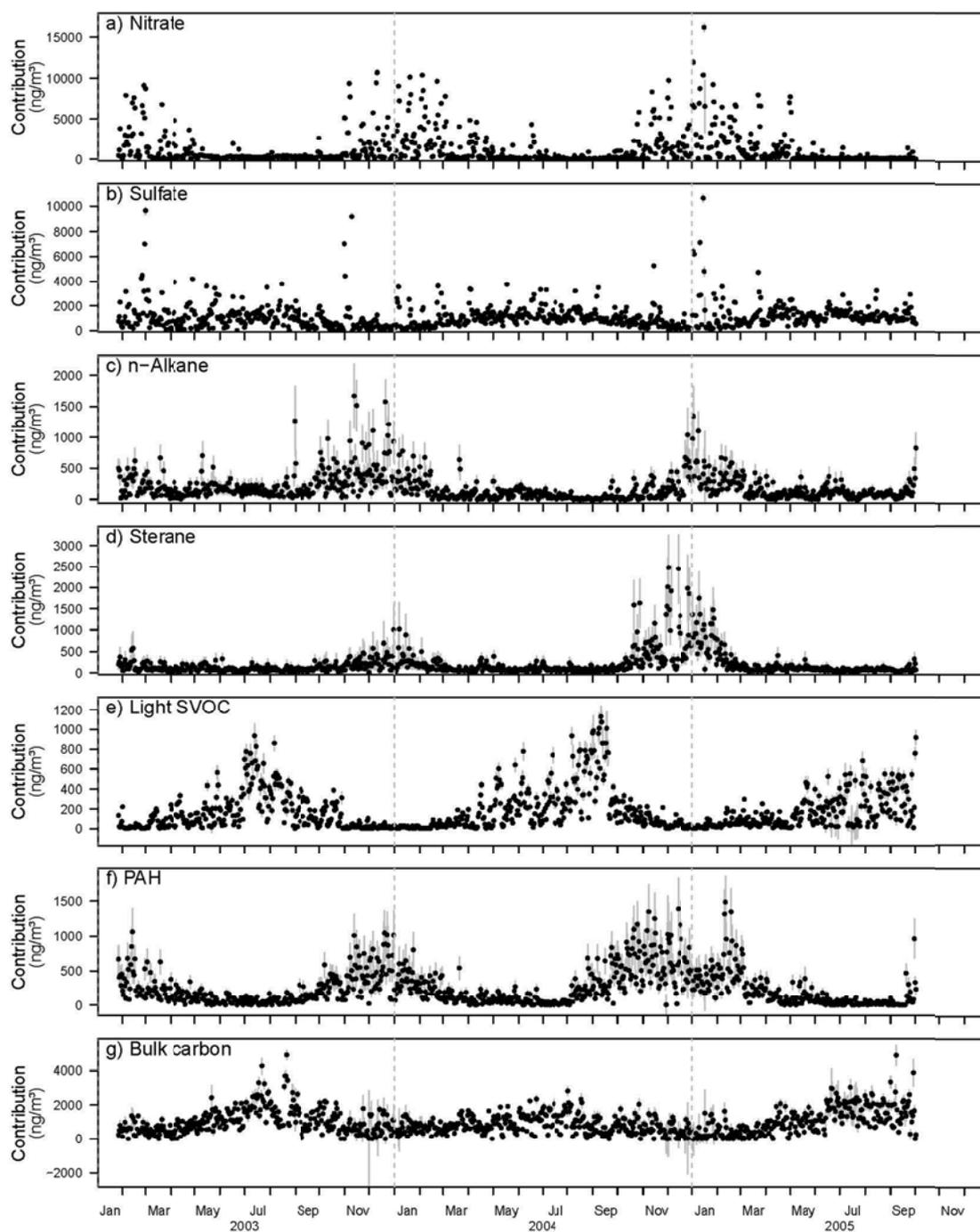


Figure 6.S2 Median PMF factor contributions for the full data set solution (black points). The gray bars represent the variability in the estimation of factor contribution from bootstrapped PMF solutions ( $\pm 1$  SD).

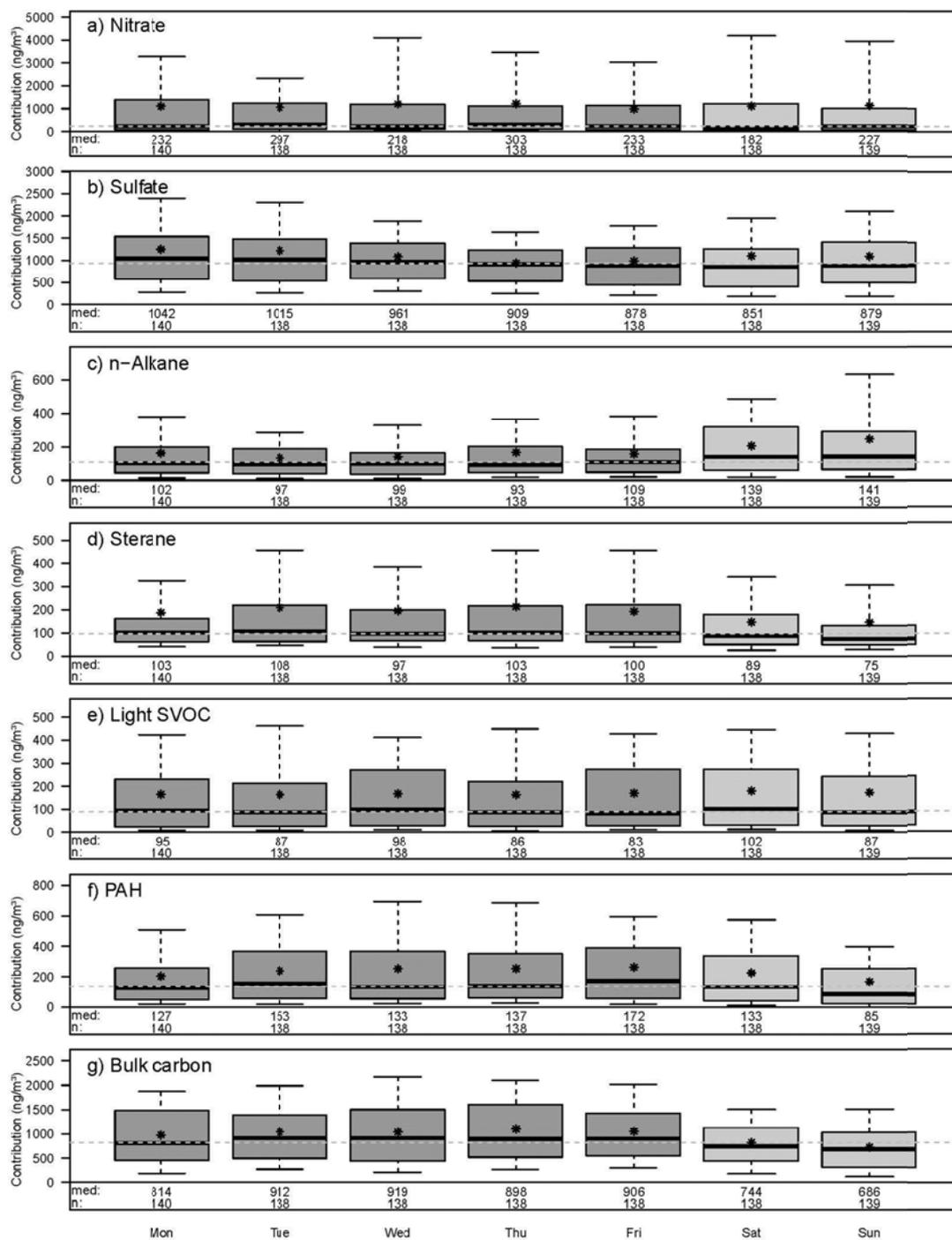


Figure 6.S3 Day of the week box plots for the factor contributions of full data set solution. The boxes depict the median (dark line), inner quartile range (shaded box), 10<sup>th</sup> and 90<sup>th</sup> percentiles (whiskers) and the mean (asterisk). The dashed line across the plot is the overall median. The individual median values and the number of points contained within each box are listed below the box.

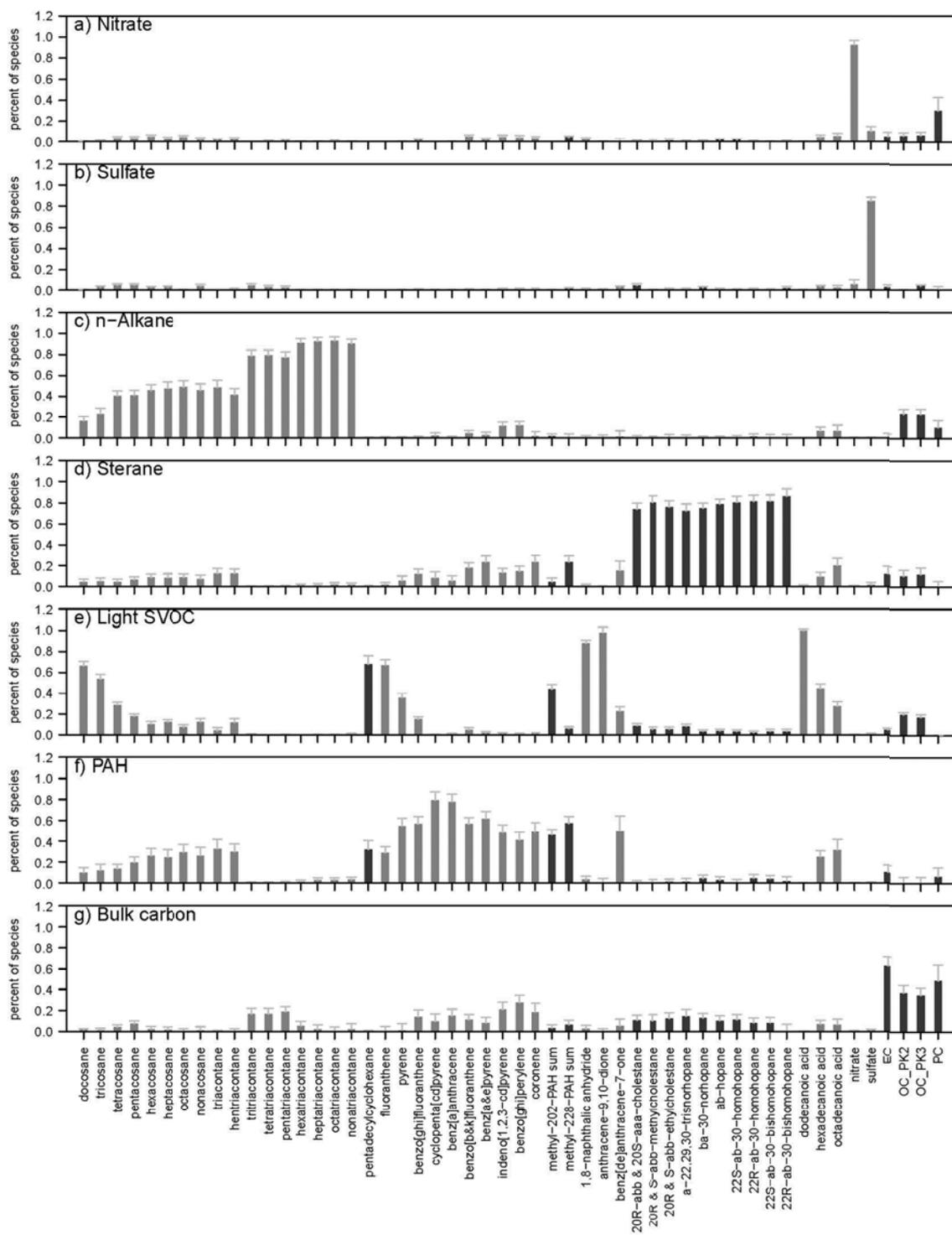


Figure 6.S4 Median PMF factor profiles for the cold period sub-data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations.

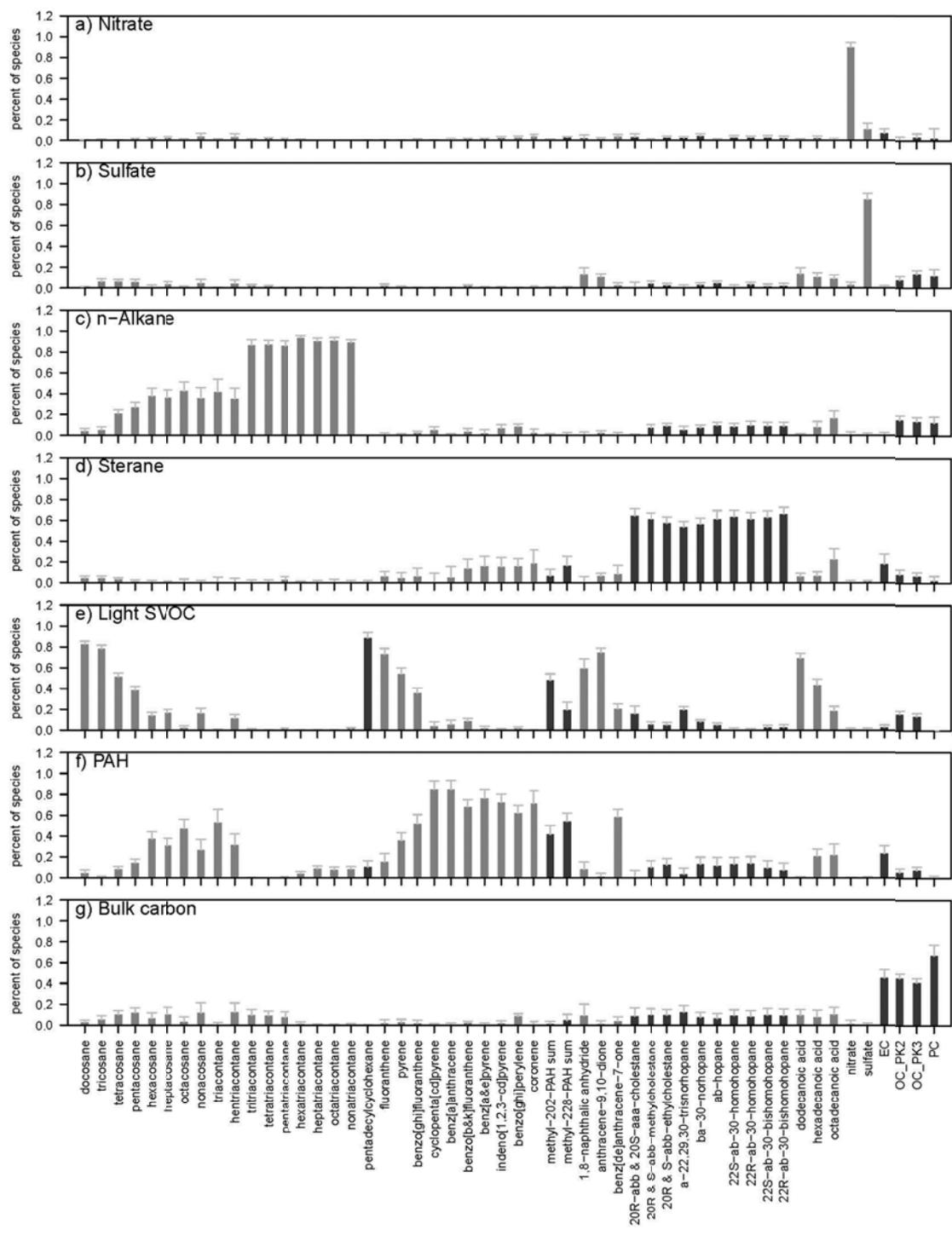


Figure 6.S5 Median PMF factor profiles for the warm period sub-data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations.

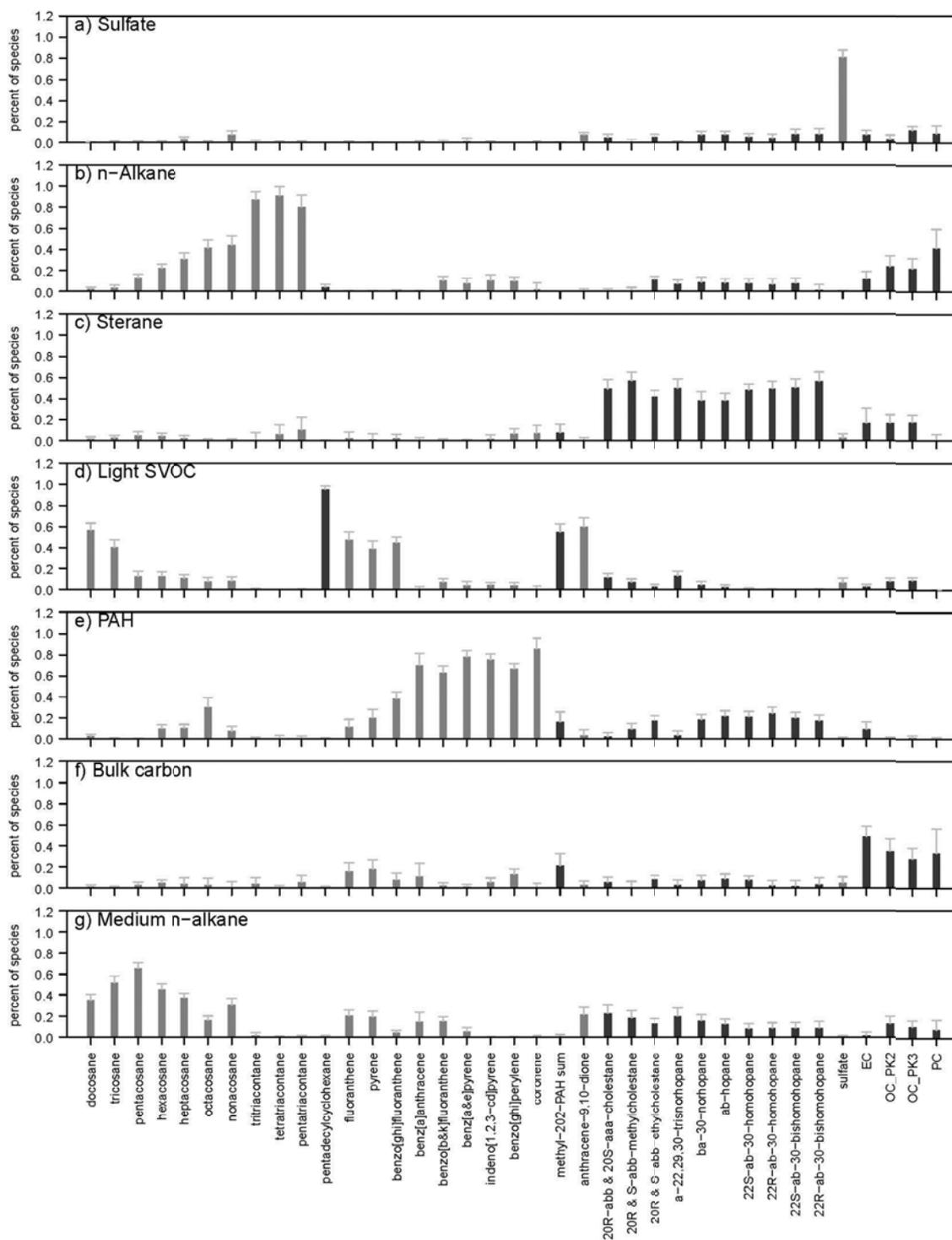


Figure 6.S6 Median PMF factor profiles for the hot period sub-data set solution (shaded bars). The whiskers represent the variability in factor profile derived from bootstrapped PMF solutions (+ 1 SD). The bars are shaded to show compound class separations.

## 11.3 SUPPLEMENTAL DISCUSSIONS

### 11.3.1 Extreme spike during Thanksgiving holiday (for CHAPTER 2)

The compositional data of samples collected on November 25<sup>th</sup> 2008, two days before Thanksgiving Day, were compared with annual averages in Figure 2.S2, from which we can observe a pronounced increase in all groups of species. The OC/EC ratios over the four sampling sites on that day were 1.6 - 4.0, 2 - 3 times lower than their corresponding annual averages, suggesting a sharp increase of contribution from primary emissions.

In order to find a plausible explanation for this spike, hourly CO and NO<sub>x</sub> mixing ratios as well as wind speed and temperature measured at the Welby site were obtained from the Colorado Department of Public Health and Environment (CDPHE). The Welby station is located 1 mile northwest of ALS. On Nov. 25<sup>th</sup> 2008, mixing ratios of CO and NO<sub>x</sub> are much higher than on other sampling days, with peaks occurring in the late morning and early evening (Figure 2.S3a, b). CO/NO<sub>x</sub> ratios lie between 3.1 and 9.8 (average 6.8) with little variation after 4:00 am (Figure 2.S3c). This ratio is much lower than the annual average and closer to the value for traffic emissions (8.8 - 9.4 ppm ppm<sup>-1</sup>) (Kirchstetter et al., 1999) than that for smoke from burning various prescribed biomass fuels (> 17 ppm ppm<sup>-1</sup>) (Andreae and Merlet, 2001) and wildfire (> 34 ppm ppm<sup>-1</sup>) (Ward et al., 1992), suggesting a stronger motor vehicle contribution on that day. In addition, wind speeds on Nov.25<sup>th</sup> 2008 were lower than annual averages during most time of the day (Figure 2.S3d). Hourly average wind speeds were negatively correlated with concentrations of CO ( $r = 0.42$ ) and NO<sub>x</sub> ( $r = 0.45$ ). Ambient temperature exhibited similar hourly variations as on other sampling days with lower values (Figure 2.S3e). As a result, the unusual extreme spike during the Thanksgiving holiday was likely initiated by heavy traffic activity in the late morning, emitting a large amount of pollution that accumulated in a stagnant

and cold atmosphere, and then was enhanced by more vehicle emissions and residential wood combustions in the evening. A body of evidence developed in the past shows increased daily mortality and morbidity associated with exposures to fine particulate air pollution (Dockery, 2001; de Hartog et al., 2008), thus a sudden acute increase in PM<sub>2.5</sub> mass and the related components as observed in this study might pose a threat to public health in a short term.

### 11.3.2 Impact of quantification uncertainty on correlation coefficient ( $r$ ) and COD (for CHAPTER 2)

Simulated data were used to evaluate the effect of quantification uncertainty on the values of  $r$  and COD observed CHAPTER 2. As mentioned in section 2.2.3 of CHAPTER 2, the uncertainty is composed of those in instrumental analysis, blank correction and sample air volume, and expressed as S/N ratios in Tables 2.S2 and 2.S3. For this analysis, pairs of bivariate lognormal random variables ( $X_f^*$ ,  $X_h^*$ ) were generated ( $N = 10^6$ ) with specified correlation coefficients ( $\rho^*$ ), simulating the *true* concentrations of a given species in side-by-side samples or at two sampling sites. To most effectively mimic our study,  $X_f^*$  and  $X_h^*$  were generated such that the means and variances of  $\text{Ln}(X_f^*)$  and  $\text{Ln}(X_h^*)$  were estimated from the time series concentrations observed for the DASH study. Simulated *measured* concentrations including quantification uncertainty were then calculated as  $x_i = x_i^* + u_i \times m_i$ , where  $x_i^*$  is an element of the true concentration vector  $X_f^*$  or  $X_h^*$ ;  $u_i$  is the quantification uncertainty, which is estimated by the ratio of  $x_i^*$  to S/N (signal to noise ratio, Table 2.S2 and 2.S3); and  $m_i$  is a random number from a normal distribution with a mean of 0 and standard deviation of 1. Values of the correlation coefficient ( $\rho$ ) and COD are then calculated from the new pair of bivariate variables ( $X_f$ ,  $X_h$ ). This process is repeated with a range of values of  $\rho^*$  until the correlation coefficient ( $\rho$ ) between  $X_f$  and  $X_h$  matches the observed value of  $r$  within  $\pm 0.5\%$ . The percentage differences

between  $\rho^*$  and  $\rho$ , and between the CODs of  $(X_f^*, X_h^*)$  and  $(X_f, X_h)$ , were used to estimate the sensitivity of these statistics to quantification uncertainty.

Table 2.S4 shows the results for 16 species with different quantification uncertainties. For the side-by-side samples, the effect on the statistics is generally greatest for those species with the lowest signal to noise ratios (e.g., tridecanoic acid, heptadecanoic acid), supporting the hypothesis that quantification uncertainties contribute significantly to the high side-by-side divergence observed for some species. On a relative basis, quantification uncertainty also contributes significantly to the divergence in side-by-side measurements of bulk EC and OC, although their absolute coefficients of divergence are low (COD = 0.14 and 0.10; Table 2.S3). However, the estimated quantification uncertainties do not account for all of the observed divergence in the side-by-side measurements, suggesting there are unaccounted for differences between the samplers or the air parcels being sampled. The influence of quantification uncertainty on comparisons across locations is illustrated in Table 2.S4 with results for PAL and EDI. For the multi-site comparisons, species with higher quantification uncertainty (lower S/N ratio) show higher influence of quantification uncertainty on  $r$  and COD. The COD value for EC is also highly impacted, on a relative basis. Thus large uncertainties in species quantification can have an impact on the assessment of spatial variability using  $r$  and COD. The simulation approach demonstrated here may be useful for understanding the impact of quantification uncertainties on cross-site comparisons in other studies where side-by-side measurements are not available for comparison.

### 11.3.3 Ratio-ratio plot analysis (for CHAPTER 3)

The ratio-ratio plot analysis is a technique for investigating the relative influence of similar sources on organic molecular markers (Robinson et al., 2006a; Robinson et al., 2006d, b, c), as a

complement to traditional quantitative source apportionment analysis. The essence of this method is to construct plots of ratios comparing two species concentrations, each normalized by a third species. If all ambient data cluster to a point on the ratio-ratio plot, then one dominant source class with consistent emission ratios for all three species could be inferred. If the ambient data in the plots organize on a line, this suggests two sources with varying emission rates for the three species. If the ambient data are constrained within a region and appear scattered, the three species are likely emitted by three or more source classes with unique emission rates. In this study, ratio-ratio plots were used to compare pairs of dominant organic molecular markers in one PMF factor that likely arose from multiple sources, assisting in the interpretation of the factor profiles.

#### 11.3.4 Comparing two correlation coefficients ( $r_1$ & $r_2$ ; for CHAPTER 8)

Assume  $r_1$  the correlation between X and Y in one population, while  $r_2$  is the correlation between X' and Y' in another population. The  $n_1$  and  $n_2$  are the numbers of sample pairs for  $r_1$  and  $r_2$ .

H<sub>0</sub>:  $r_1 = r_2$

First, Fisher's transformation is applied to both correlation coefficients:

$$R_1 = \frac{1}{2} \ln \left( \frac{1+r_1}{1-r_1} \right)$$

$$R_2 = \frac{1}{2} \ln \left( \frac{1+r_2}{1-r_2} \right)$$

If the null hypothesis of no difference between  $r_1$  and  $r_2$  is true, the variable ( $R_1 - R_2$ ) should satisfy a normal distribution with mean equal to 0 and the variance equal to:

$$\sigma^2 = \frac{1}{n_1-3} + \frac{1}{n_2-3}$$

Second, test statistics was computed:

$$z = \frac{R1 - R2}{\sigma}$$

Then obtain level of significance ( $p$  value) for the computed  $z$ :

$$p = 2\Phi\left(-\left|\frac{R1 - R2}{\sigma}\right|\right)$$

### 11.3.5 Compare the slope of linear regression to 1.0 (for CHAPTER 8)

A linear regression equation can be written as

$$y = b + mx$$

$y$ : dependent variable;

$x$ : independent variable;

$b$ : intercept;

$m$ : slope of regression line.

Expression for  $m$  and the corresponding standard errors ( $s.e.(m)$ ) could be written as

$$m = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

$$s.e.(m) = \sqrt{\frac{\frac{1}{n-2} \sum (y_i - \bar{y})^2 - b^2 \sum (x_i - \bar{x})^2}{\sum (x_i - \bar{x})^2}}$$

The statistic

$$t_0 = \frac{|m - 1|}{s.e.(m)}$$

follows a  $t$ -distribution with  $n-2$  degrees of freedom. We would reject the null assumption:  $m$

=1 if

$$t_0 > t_{\alpha/2, n-2}$$

where  $\alpha$  is the selected significance level. Here we use 0.05.