

**AIR-QUALITY MODELING AND SOURCE-APPORTIONMENT OF
FINE PARTICULATE MATTER: IMPLICATIONS AND
APPLICATIONS IN TIME-SERIES HEALTH STUDIES**

A Dissertation
Presented to
The Academic Faculty

by

Amit Marmur

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
December 2006

**AIR-QUALITY MODELING AND SOURCE-APPORTIONMENT OF
FINE PARTICULATE MATTER: IMPLICATIONS AND
APPLICATIONS IN TIME-SERIES HEALTH STUDIES**

Approved by:

Dr. Armistead G. Russell, Advisor
School of Civil and Environmental
Engineering
Georgia Institute of Technology

Dr. Paige E. Tolbert
School of Public Health
Emory University

Dr. Michael E. Chang
School of Earth and Atmospheric Sciences
Georgia Institute of Technology

Dr. James A. Mulholland
School of Civil and Environmental
Engineering
Georgia Institute of Technology

Dr. Michael H. Bergin
School of Civil and Environmental
Engineering
Georgia Institute of Technology

Date Approved: September 19, 2006

To my wife, son, and family

ACKNOWLEDGEMENTS

I would like to express my gratitude to a number of people who have provided me with the support and encouragement needed for making this dissertation a reality. First, I would like to thank my advisor, Dr. Ted Russell, for his insightful guidance, support, and for many enriching scientific discussions throughout the years. I am also thankful for the guidance of Dr. Jim Mulholland, whose insight made me think and re-search. Thanks also to Dr. Paige Tolbert for the insight into the world of air-pollution epidemiology, and to Dr. Mike Bergin and Dr. Michael Chang for serving on my thesis committee.

Group-members and friends at Georgia-Tech have enriched both my educational and social experience, and thanks go to: Yongtao Hu (for his time and patience), Alper Unal, Sun-Kyoung Park, Di Tian, Kasemsan Manomaiphiboon, Dan Cohan, Jaemeen Baek, Jim Boylan, Maudood Khan, Sergey Napelenok, Sangil Lee, K.J. Liao, Efthimios Tagaris, Amir Hakami, Evan Cobb, Bo Yan, Burcak Kaynak, Santosh Chandru, Michelle Bergin, Kangwook Kim, Katie Wade, Talat Odman, Therese Rehkopf, Patricia Taboada, and Anup Shah. Collaborations with the Rollins School of Public Health at Emory University have both motivated and provided crucial aspects to this work, and I am especially thankful for the help of Jeremy Sarnat, Mitch Klein, Stafanie Sarnat, Jennifer Peel and Kristi Metzger. I would also like to thank my supervisors at the Georgia Environmental Protection Division for the much needed flexibility to complete this work.

Above all I would like to thank my wife, Efrat, for her love, patience, and support on this journey, and for fulfilling her goals and dreams as well during our stay in the USA; my son, Eli, for making me laugh; and my family for their encouragement.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	vii
LIST OF FIGURES	x
SUMMARY	xiii
<u>CHAPTER</u>	
1 INTRODUCTION	1
Structure and scope of thesis	5
References	8
2 TEMPORAL AND SPATIAL VARIABILITY IN MEASURED AND SIMULATED PM _{2.5} CONSTITUENTS IN ATLANTA, GA, AND IMPLICATIONS FOR TIME-SERIES HEALTH STUDIES	11
Abstract	11
2.1 Introduction	12
2.2 Methods	15
2.3 Results	18
2.4 Discussion	28
Acknowledgements	30
References	30
3 OPTIMIZATION BASED SOURCE APPORTIONMENT OF PM _{2.5} INCORPORATING GAS-TO-PARTICLE RATIOS	34
Abstract	34
3.1 Introduction	35
3.2 Model description	38

3.3 Model implementation	42
3.4 Results	48
Acknowledgements	61
References	62
4 OPTIMIZED VARIABLE SOURCE-PROFILE APPROACH FOR SOURCE APPORTIONMENT	65
Abstract	65
4.1 Background	66
4.2 Methods	68
4.3 Results	74
4.4 Conclusions	89
Acknowledgements	90
References	90
5 SOURCE APPORTIONMENT OF PM _{2.5} IN THE SOUTHEASTERN UNITED STATES USING RECEPTOR AND EMISSIONS-BASED MODELS: CONCEPTUAL DIFFERENCES AND IMPLICATIONS FOR TIME-SERIES HEALTH STUDIES	93
Abstract	93
5.1 Introduction	94
5.2 Methods	95
5.3 Results	106
5.4 Discussion	126
Acknowledgements	128
References	128
6 INTERMETHOD VARIABILITY IN ASSOCIATIONS BETWEEN SOURCE-APPORTIONED PM _{2.5} AND DAILY EMERGENCY-DEPARTMENT VISITS IN ATLANTA, GA	133
Abstract	133

6.1 Introduction	134
6.2 Methods	135
6.3 Results	145
6.4 Discussion	153
6.5 Summary	157
References	158
7 CONCLUSIONS AND FUTURE RESEARCH	160
Conclusions	160
Future research	164
Closing remarks	168
References	168
VITA	170

LIST OF TABLES

	Page
Table 2.1: Model performance statistics for 36km CMAQ modeling of the Atlanta aerosol.	19
Table 2.2: Comparison between model performance using a 36km and a 12km modeling grid (1/01-8/01).	19
Table 2.3: Model performance statistics for VISTAS 12km modeling of the Atlanta aerosol.	21
Table 2.4: Comparison between correlation coefficients based on the entire simulated dataset and shorter subparts of the dataset.	23
Table 3.1: Mean, standard-deviation, minimum, maximum of ambient levels of the species used for the source apportionment, JST site, Atlanta, GA.	44
Table 3.2: Particulate source-profiles used in the apportionment process.	47
Table 3.3: Gas-to-PM _{2.5} ratios used as constraints in the optimization process.	48
Table 3.4: Average and standard deviation of the source-contributions to PM _{2.5} levels measured at JST site, using CMB8-EV, CMB8-OWLS and LGO.	52
Table 3.5: Correlations (R) between source-contributions and ambient-levels of fitting species.	57
Table 3.6: Source inter-correlations (R) using CMB8-EV, CMB8-OWLS and LGO.	59
Table 3.7: Comparison between percent contributions to primary PM _{2.5} levels based on LGO, organic-marker CMB, and the five-county Atlanta metropolitan area emissions inventory.	61
Table 4.1: Chemicals from particles in different emissions sources.	69
Table 4.2: Lower and upper bounds for chemical fractions of total PM _{2.5} mass emitted in source profiles.	70
Table 4.3: Percent of cases in which the derived species fraction was at the lower limit / within the allowable range / at the upper limit.	81
Table 4.4: LGO derived PM _{2.5} source profiles based on 447 days in which all relevant data (ions, EC, OC, metals, CO, SO ₂ , NO _y) were available.	82
Table 4.5: Performance measures for the LDSP and MBSP solutions.	85

Table 4.6: Correlation matrix of source-contributions based on LDSP and MBSP solutions.	85
Table 4.7: Percent change in average source-attributions for a 50% increase in concentrations of several tracer species.	87
Table 4.8: Comparison between source apportionment results based on PMF, LSDP, and MBSP.	89
Table 5.1: Correlations between CMB-LGO and CMAQ based source contributions to PM _{2.5} at JST.	115
Table 5.2: Correlations between CMB-LGO and CMAQ based source contributions to PM _{2.5} at YK.	115
Table 5.3: Correlations between CMB-LGO and CMAQ based source contributions to PM _{2.5} at BHM.	116
Table 5.4: Correlations between CMB-LGO and CMAQ based source contributions to PM _{2.5} at CTR.	116
Table 5.5: Correlations between source contributions and ambient species concentrations measured at JST.	121
Table 5.6: Change in average species concentrations and average source-attributions corresponding to a decrease/increase of one standard-deviation of the log normalized species concentrations.	121
Table 5.7: Average and standard deviation of the ratio of ambient concentration to minimum detection limit.	122
Table 6.1: Average source contributions to PM _{2.5} in Atlanta (Jefferson St.) based on CMB-LGO and PMF for the period of November 1998 – August 2000.	140
Table 6.2: Correlation matrix of daily source contributions to PM _{2.5} in Atlanta (Jefferson St.) based on CMB-LGO and PMF for the period of November 1998 – August 2000.	140

LIST OF FIGURES

	Page
Figure 2.1: Location of SEARCH and ASACA PM _{2.5} monitoring stations in the Atlanta Metropolitan, overlaid by model grid cells corresponding to these sites (36 km and 12 km domains).	17
Figure 2.2: Comparison between daily modeled (36km grid) and measured (JST) sulfate concentrations, for the period of August 1999 – August 2001.	23
Figure 2.3: Comparison between daily modeled (36km grid) and measured (JST) EC concentrations, for the period of August 1999 – August 2001.	24
Figure 2.4: Comparison between CMAQ modeled SOA and unapportioned OM from a CMB analysis for 2002 at the JST site.	28
Figure 3.1: Source-contributions to PM _{2.5} levels at JST site, Atlanta, GA, using CMB8 EV solution, CMB8 OWLS solution and LGO.	51
Figure 4.1: Source profiles generated by LGO for gasoline-fueled vehicles, compared to a profile from the NFRAQS study.	78
Figure 4.2: Source profiles generated by LGO for diesel-fueled vehicles, compared to a profile from the NFRAQS study.	78
Figure 4.3: Source profiles generated by LGO for vegetative burning, compared to a vegetative burning profile from the BRAVO study.	79
Figure 4.4: Source profiles generated by LGO for soil-dust, compared to an Alabama soil dust profile.	79
Figure 4.5: Source profiles generated by LGO for coal-fired power plants, compared to a CFPP profile measured in Texas.	80
Figure 4.6: Average source contributions (8/1/98-8/31/00) to PM _{2.5} at the Atlanta Jefferson Street site, using LGO-derived source-profiles and measurement-based source-profiles.	83
Figure 5.1: Locations of SEARCH monitoring sites in Georgia (JST and YK) and Alabama (BHM and CTR), over plotted by the grids (36km) of the air quality model.	103
Figure 5.2: Average source contributions to primary PM _{2.5} over a two month period (July 2001 and January 2002) at four SEARCH sites in Georgia and Alabama.	107

Figure 5.3: Daily source-specific contributions to primary PM _{2.5} at JST from gasoline vehicles, diesel vehicles, soil dust, vegetative/wood burning and coal fired power plants for the periods July 2001 and January 2002 using CMAQ.	110
Figure 5.4: Daily source-specific contributions to primary PM _{2.5} at JST from gasoline vehicles, diesel vehicles, soil dust, vegetative/wood burning and coal fired power plants for July 2001 and January 2002 using CMB-LGO.	111
Figure 5.5: Daily fractional source contribution at JST using CMAQ.	112
Figure 5.6: Daily fractional source contribution at JST using CMB-LGO.	112
Figure 5.7: Average daily modeled emissions in the model cell where the JST site is located.	117
Figure 5.8: Location of JST and SDK monitoring sites with respect to major interstates in Atlanta.	124
Figure 5.9: Time series comparison between ambient concentrations at two Atlanta monitoring sites located 15.3 km apart (SDK and JST) for K, Si, Fe, Se, EC and Zn.	126
Figure 6.1: Daily PM _{2.5} source impacts of gasoline vehicles and coal-fired power plants for select periods.	137
Figure 6.2: Daily PM _{2.5} source impacts of vegetative-burning and “other” (unapportioned) OC based on CMB and CMB-LGO, along with the ratio of calculated-to-observed K.	138
Figure 6.3: CMB-LGO vs. PMF source contributions of gasoline vehicles, diesel vehicles, soil dust, wood/vegetative burning and CMB-LGO “other” OC/PMF gasoline for the period of November 1998 – August 2000.	141
Figure 6.4: Weekly and seasonal patterns in contributions of gasoline vehicles, diesel vehicles, wood burning and “other” OC to PM _{2.5} levels in Atlanta (Jefferson St.) based on CMB-LGO for the period of November 1998 – August 2000.	143
Figure 6.5: Weekly and seasonal patterns in contributions of gasoline vehicles, diesel vehicles and wood burning to PM _{2.5} levels in Atlanta (Jefferson St.) based on PMF for the period of November 1998 – August 2000.	143
Figure 6.6: Preliminary cardiovascular related risk-ratios for gasoline, diesel and total mobile-source related PM _{2.5} per change of one inter-quartile range.	149
Figure 6.7: Preliminary respiratory related risk-ratios for gasoline, diesel and total mobile-source related PM _{2.5} per change of one inter-quartile range.	149

Figure 6.8: Preliminary respiratory and cardiovascular related risk-ratios for wood burning related PM _{2.5} per change of one inter-quartile range.	150
Figure 6.9: Preliminary respiratory and cardiovascular related risk-ratios for soil PM _{2.5} per change of one inter-quartile range.	150
Figure 6.10: Preliminary respiratory and cardiovascular related risk-ratios for sulfate-related and total PM _{2.5} per change of one inter-quartile range.	151
Figure 6.11: Preliminary respiratory and cardiovascular related risk-ratios for nitrate-related and total PM _{2.5} per change of one inter-quartile range.	151
Figure 6.12: Preliminary respiratory and cardiovascular related risk-ratios for CMB-LGO “other”/unapportioned OC, total OC and total PM _{2.5} per change of one inter-quartile range.	152
Figure 6.13: Control outcome related risk-ratios for various CMB-LGO source categories and total PM _{2.5} per change of one inter-quartile range.	152
Figure 6.14: Scatter plot of preliminary risk-ratios for various respiratory and cardiovascular health-outcomes based on PMF and CMB-LGO source contributions.	154

SUMMARY

Particulate matter, especially that smaller than 2.5 microns in diameter ($PM_{2.5}$), has been associated with adverse health effects and mortality in studies covering more than 150 cities. Studies of the effects of air pollution on human health are typically conducted using ambient measurements to represent the air quality over cities or regions. However, the use of ambient data for such studies introduces several limitations such as spatial representativeness of the monitoring site, analytical uncertainties, and incompleteness and lack of continuity in data. The complex chemical composition of $PM_{2.5}$ and associated analytical uncertainties pose a further challenge when trying to investigate species-specific health effects. A complementary approach is to examine associations between health outcomes and sources contributing to ambient $PM_{2.5}$, which can provide regulators with important information to tighten controls on sources more prone to causing health effects.

Air-quality modeling tools may be useful in such investigations of the health effects of air-pollution and $PM_{2.5}$ specifically. Emissions-based three-dimensional air quality models may introduce several benefits when applied in epidemiologic studies, such as improved spatial representativeness and availability/continuity of data, as well as information on source impacts. Receptor-based models are a common tool for apportioning of ambient levels of pollutants among the major contributing sources, and can be useful in discerning the relative health impacts of different sources.

Results from a long term air quality simulation using EPA's Models-3 suite of models (MM5/SMOKE/CMAQ) were analyzed in terms of the model's ability to

simulate temporal and spatial variability in concentrations of both secondary and primary PM_{2.5} components in Atlanta, GA. Seasonal variations in sulfate and nitrate concentrations were well captured by the model, but the model's ability to capture shorter-term (e.g., daily) variations, typically of interest in time-series health studies of acute outcomes, was limited. Moreover, the spatial homogeneity in ambient concentrations of secondary PM_{2.5} constituents (such as sulfate and nitrate), suggests limited benefit in applying simulated concentration fields for these species in a time-series health study when ambient measurements are available. Concentrations of primary PM_{2.5} constituents (such as elemental carbon), on the other hand, have much greater spatial variability, and the short-term variability in these species is better captured by the air quality model.

A modified approach to receptor-based PM_{2.5} source apportionment was developed, using source indicative SO₂/PM_{2.5}, CO/PM_{2.5} and NO_x/PM_{2.5} ratios as constraints, in addition to the commonly used particulate-phase source profiles. Additional information from using gas-to-particle ratios assists in reducing collinearity between source profiles, a problem that often limits the source-identification capabilities and accuracy of traditional receptor models. The set of equations for the PM_{2.5} Chemical Mass Balance (CMB) receptor model were solved using a global-optimization program, Lipschitz Global Optimizer (LGO), subject to constraints on ambient gas-phase concentrations. Application of the CMB-LGO model to a 25 month dataset of daily PM_{2.5} measurements (total mass and composition) at the Atlanta Jefferson Street SEARCH site yielded source-contributions that seem more indicative of the named sources compared to particulate-phase source apportionment methods, based on correlations of the source

impacts and tracer species. Furthermore, collinearity between source-categories (e.g., soil-dust and primary PM_{2.5} from coal-burning; wood burning and “other” OC) was reduced based on higher source inter-correlations than in the “regular” CMB model.

A further expansion of the CMB-LGO approach for PM_{2.5} source apportionment was developed in which both the local source compositions and corresponding contributions were determined from ambient measurements and initial estimates of source compositions. Such an approach can serve as an alternative to using predetermined (measured) source profiles, as traditionally used in CMB applications, which are not always representative of the region and/or time period of interest. This technique was applied to a dataset of PM_{2.5} measurements at the former Atlanta supersite (Jefferson Street site), to apportion PM_{2.5} mass into nine source categories. Good agreement is found when these source impacts are compared with those derived based on measured source profiles as well as those derived using a factor analytical technique.

Two conceptually different approaches to source apportionment were compared: a receptor model and an emissions-based air-quality model. The receptor model captured more of the temporal variation in source impacts at a specific receptor site compared to the emissions-based model. Driven by data at a single site, receptor models may have some shortcomings with respect to spatial representativeness (unless a reduced study area is used or data from multiple sites are available), likely attenuating the observed association in a health study. Source apportionment results from emissions-based models, such as CMAQ, may be more spatially representative as they represent an average grid-cell value. Limitations in the ability to model daily fluctuations in emissions, however, lead to results being driven mainly by regional meteorological trends, likely

underestimating the true daily variations in local source impacts. These effects will likely introduce an attenuation of observed association in a health study and limit the model's usefulness in discerning the relative impacts of the sources on health outcomes.

Results from a preliminary source-specific $\text{PM}_{2.5}$ epidemiologic analysis were presented and analyzed for inter-method variability in risk-ratio estimates, based on source apportionment results from the Positive Matrix Factorization (PMF) and Chemical Mass Balance model incorporating the Lipschitz Global Optimizer (CMB-LGO) models, as well as the direct application of tracer species in the epidemiologic study. Despite methodological differences and uncertainties in the apportionment process, good agreement was observed between the CMB-LGO and PMF based risk ratios, suggesting the usefulness of applying apportionment methods in health studies. Preliminary epidemiologic analysis found mobile-source related $\text{PM}_{2.5}$ significantly associated with cardiovascular outcomes; wood burning $\text{PM}_{2.5}$ significantly associated with respiratory outcomes; soil dust significantly associated with asthma. "Other" OC was found significantly associated with various respiratory outcomes with high risk-ratios, though further analyses, such as based on longer datasets, are needed to support this finding. Analysis of temporal patterns in source impacts suggested that "other" OC is associated with secondary formation. Despite generally good agreement between risk-ratios estimates based on source-contributions and tracer-species, source-apportionment and sensitivity analyses are needed to determine the most suitable tracers for each source-category.

CHAPTER 1

INTRODUCTION

Particulate matter, especially particles smaller than 2.5 microns in diameter ($PM_{2.5}$), has been associated with adverse health effects and mortality in studies covering more than 150 cities (Dockery et al. 1993; Pope et al. 2002; Pope et al. 1995). Both acute and chronic exposures to $PM_{2.5}$ have been associated with increased mortality rates and hospital visits, as well as cardiopulmonary disease, heart attacks, decreased lung function, and asthma (Dockery et al. 1993; Ebel et al. 2000; Peters et al. 2001; Pope et al. 1995; Vedal 1997). Studies of the effects of air pollution on human health are typically conducted using ambient measurements to represent the air quality over cities or regions. However, the use of ambient data for such studies introduces several limitations: monitoring sites might not adequately represent air quality over the health study domain (a single central site is often used) or capture the spatial variability in concentration fields; measurement errors, especially for complex measurements of $PM_{2.5}$ species, might introduce noise to the epidemiologic analysis; incompleteness of data and lack of continuity in data might diminish the ability of the epidemiologic analysis to detect such health related associations; the need for large datasets often leads to conducting measurements over a multi-year period, but obtaining ambient measurements of $PM_{2.5}$ species is costly and time-consuming. In addition, since $PM_{2.5}$ is chemically complex (comprised of numerous primary and secondary components, including ionic and organic compounds and dozens of trace elements), recent epidemiologic studies also investigate whether specific components of $PM_{2.5}$ are more prone to cause specific health effects

(Hauck et al. 2004; Heal et al. 2005; Metzger et al. 2004; Metzger et al. 2004; Peel et al. 2005). However, the association between health outcomes and specific PM_{2.5} components raises several issues: it is not obvious that the major cause for the health outcome is actually measured (it is impractical to measure every single PM_{2.5} species) or is possibly measured inaccurately due to analytical issues; the actual health effects may be due to a combination of pollutants; many species are correlated which limits the ability to isolate species health impacts. A complementary approach is to examine associations between health outcomes and sources contributing to ambient PM_{2.5} (Laden et al. 2000; Manchester-Neesvig et al. 2003; Mar et al. 2000; Tsai et al. 2000). By means of source apportionment, source impacts on the receptor can be quantified, and their health impacts examined. A source-impact oriented approach could help target and regulate the sources that contribute most to adverse health effects. It could also allow for better multi-component epidemiologic modeling, as the number of major source-impact categories is typically far fewer than the number of PM components. Finally, this approach can help identify health effects of unmeasured species present in emissions from specific source categories. For example, preliminary studies have found an association between mortality and combustion-related PM_{2.5} (from motor vehicles, coal combustion and wood burning), but not soil-related PM_{2.5}, in both cohort (Laden et al. 2000) and time-series (Mar et al. 2000) studies.

Air-quality modeling tools may be useful in such investigations of the health effects of air-pollution and PM_{2.5} specifically. Emissions-based three-dimensional air quality models may introduce several benefits when applied in epidemiologic studies. First, an average value over a model cell of typical size (e.g., 36km x 36km or 12km x 12

km) may better represent the air quality over an applicable area as compared to a measurement at a single point within that area. Such a value can also assist in evaluating the representativeness and quality of measurements at different locations (stations) all residing within the range of the same model cell. Also, the ability to model episodes for which no measurements were performed may allow expanding the epidemiologic study to geographical areas for which no data were available, to past (historic) episodes, and to complete and extend existing datasets. Using air quality models is also less time-consuming and resource intensive as compared to the real-time required for measurements, often spanning several years. Air quality models can also provide information on source impacts, can simulate atmospheric concentrations for shorter time periods than many measurement techniques, and can simulate the levels of pollutants that are very difficult to measure (e.g., species found at very trace levels). Emissions-based, three dimensional photochemical air-quality models simulate the formation, transport and fate of atmospheric chemical constituents, both gaseous and in particulate form, by solving the conservation equation expressed as:

$$\frac{\partial c_i}{\partial t} + \nabla \cdot (U c_i) = \nabla \rho D_i \nabla \left(\frac{c_i}{\rho} \right) + R_i(c_1, c_2, \dots, c_n, T, t) + S_i(x, t) \quad i = 1, 2, \dots, n$$

where: c_i is the concentration of species i ; U is the wind velocity vector; D_i is the molecular diffusivity of species i ; R_i is the rate of concentration change of species i by chemical reaction; $S_i(x, t)$ is the source/sink of i at location x and time t ; ρ is the air density; n is the number of predicted species. The conservation equation describes the formation, transport and fate of air pollutants, including components for processing emissions, meteorology, topography, and atmospheric chemistry (Russell and Dennis 2000). Source apportionment can be performed using direct sensitivity methods such as

Direct Decoupled Method (DDM) (Dunker 1981; Hakami et al. 2003), inert tracer methods, or by multiple applications of the model with and without emissions from target sources (“brute force”). Here we applied the US-EPA’s Models-3 suite of models, including MM5 (Grell et al. 1999) as the meteorological model to simulate atmospheric physical dynamics; SMOKE (Houyoux et al. 2003) as the emissions processor to calculate spatial and temporal trends in emissions based on the annual emissions inventory; and the Community Multiscale Air Quality (CMAQ) model (Byun and Ching 1999) to simulate atmospheric reactivity, transport, and deposition of chemical contaminants.

Chemical Mass Balance (CMB) receptor models are a common tool for apportioning of ambient levels of pollutants among the major contributing sources. CMB combines the chemical and physical characteristics of particles or gases measured at sources and receptors to quantify the source contributions to the receptor. The quantification is based on the solution to a set of linear equations that express each receptor’s ambient chemical concentration as a linear sum of products of source-profile abundances and source contributions (US-EPA 1998; US-EPA 2001), as expressed by:

$$C_i = \sum_{j=1}^n f_{ij} S_j + e_i$$

where: C_i is the ambient concentration of chemical species i ($\mu\text{g}/\text{m}^3$); $f_{i,j}$ is the fraction of species i in emissions from source j ; S_j is the contribution (source-strength) of source j ($\mu\text{g}/\text{m}^3$); n is the total number of sources; e_i is the error term. The source profile abundances (f_{ij} , the mass fraction of a chemical in the emissions from each source type) and the receptor concentrations (C_i), along with uncertainty estimates, serve as input data

to the CMB model. The output consists of the contribution of each source category (S_j) to the measured concentration of different species at the receptor.

CMB models are based on the following assumptions (US-EPA 1998): ¹⁾ compositions of source emissions are constant over the period of ambient and source sampling; ²⁾ chemical species do not react with each other, i.e., they add linearly; ³⁾ all sources with a potential for significantly contributing to the receptor are included in the analysis; ⁴⁾ the source compositions are linearly independent of each other; ⁵⁾ the number of sources or source categories is less than or equal to the number of chemical species; ⁶⁾ measurement uncertainties are random, uncorrelated, and normally distributed.

In terms of PM_{2.5} source apportionment, the major difference between a receptor model and an emissions-based air quality model is the starting point. While a receptor model's starting point is the ambient measurement, from there going backwards to estimate source contributions, the starting point of the air-quality model is the processed emissions inventory, going forward by simulating the transport and transformation of pollutants and ultimate air quality impact.

Structure and scope of thesis

- **Chapter 2: Temporal and spatial variability in measured and simulated PM_{2.5} constituents in Atlanta, GA, and implications for time-series health studies.** Temporal and spatial patterns in both observed and simulated PM_{2.5} constituents are analyzed and an emissions-based air quality model (CMAQ) evaluated in terms of simulating short-term temporal (daily) variations in concentrations of both primary and secondary PM_{2.5} components, to assess the

potential of using simulated $\text{PM}_{2.5}$ concentrations in health studies, in lieu of measured data and/or to improve spatial representativeness compared to point measurements.

- **Chapter 3: Optimization based source apportionment of $\text{PM}_{2.5}$ Incorporating gas-to-particle ratios.** A modified approach to $\text{PM}_{2.5}$ source apportionment is developed, using source indicative $\text{SO}_2/\text{PM}_{2.5}$, $\text{CO}/\text{PM}_{2.5}$ and $\text{NO}_x/\text{PM}_{2.5}$ ratios as constraints, in addition to the commonly used particulate-phase source profiles. Additional information from using gas-to-particle ratios assists in reducing collinearity between source profiles, a problem that often limits the source-identification capabilities and accuracy of traditional receptor models.
- **Chapter 4: Optimized variable source-profile approach for source apportionment.** An expanded Chemical Mass Balance (CMB) approach for $\text{PM}_{2.5}$ source apportionment is developed in which both the local source compositions and corresponding contributions are determined from ambient measurements and initial estimates of source compositions using a global-optimization mechanism. Such an approach can serve as an alternative to using predetermined (measured) source profiles, as traditionally used in CMB applications and which are not always representative of the region and/or time period of interest.
- **Chapter 5: Source apportionment of $\text{PM}_{2.5}$ in the Southeastern United States using receptor and emissions-based models: conceptual differences and implications for time-series health studies.** Two conceptually different approaches to $\text{PM}_{2.5}$ source apportionment are compared: a receptor model and an

emissions-based air-quality model. Daily source impacts are calculated using CMB-LGO (Chemical Mass Balance model incorporating the Lipschitz Global Optimizer), an extended CMB receptor model, and EPA's Models-3 emissions-based air-quality modeling system (MM5-SMOKE-CMAQ). The temporal trends in source-impacts based on the two methods is analyzed and compared and implications to time-series health studies are discussed.

- **Chapter 6: Intermethod variability in associations between source-apportioned PM_{2.5} and daily emergency-department visits in Atlanta, GA.** Results from a preliminary source-specific PM_{2.5} epidemiologic analysis are presented and analyzed for intermethod variability in risk-ratio estimates based on source apportionment results from the Positive Matrix Factorization and CMB-LGO models, as well as the direct application of tracer species in the epidemiologic study. Advantages and disadvantages of applying source-apportionment results and tracer species in epidemiologic studies are discussed.
- **Chapter 7: Conclusions and future research.** Air-quality models, both emissions and receptor based, may be useful for epidemiologic studies of the health-effects of air-pollution. Emissions-based models such as CMAQ may help address some of the issues pertaining to spatial variability and representativeness of primary PM_{2.5} constituents. Receptor-based models can help in identifying whether specific sources of PM_{2.5} are more prone to causing certain types of health effects. An application of source-apportionment results in a preliminary epidemiologic analysis found associations between source-specific PM_{2.5} and

various health endpoints, and showed relatively good agreement in risk-ratio estimates across several source-apportionment methods.

References

- Byun DW, Ching JKS. 1999. Science Algorithms of the EPA-Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, Office of Research and Development, EPA/600/R-99/030.
- Dockery DW, Pope CA, Xu XP, Spengler JD, Ware JH, Fay ME. 1993. An Association Between Air-Pollution And Mortality In 6 United-States Cities. *New England Journal Of Medicine* 329(24):1753-1759.
- Dunker AM. 1981. Efficient Calculation Of Sensitivity Coefficients For Complex Atmospheric Models. *Atmospheric Environment* 15(7):1155-1161.
- Ebelt ST, Petkau AJ, Vedal S, Fisher TV, Brauer M. 2000. Exposure of chronic obstructive pulmonary disease patients to particulate matter: Relationships between personal and ambient air concentrations. *Journal of the Air & Waste Management Association* 50(7):1081-1094.
- Grell GA, Dudhia J, Stauffer DR. 1999. A description of the fifth-generation Penn State/NCAR mesoscale model (MM5).
- Hakami A, Odman MT, Russell AG. 2003. High-order, direct sensitivity analysis of multidimensional air quality models. *Environmental Science & Technology* 37(11):2442-2452.
- Hauck H, Berner A, Frischer T, Gomiscek B, Kundi M, Neuberger M, et al. 2004. AUPHEP-Austrian Project on Health Effects of Particulates - general overview. *Atmospheric Environment* 38(24):3905-3915.
- Heal MR, Hibbs LR, Agius RM, Beverland LJ. 2005. Total and water-soluble trace metal content of urban background PM₁₀, PM_{2.5} and black smoke in Edinburgh, UK. *Atmospheric Environment* 39(8):1417-1430.

- Houyoux M, Vukovich J, Brandmeyer JE, Seppanen C, Holland A. 2003. Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE): User Manual, Version2.0.
- Laden F, Neas LM, Dockery DW, Schwartz J. 2000. Association of fine particulate matter from different sources with daily mortality in six US cities. *Environmental Health Perspectives* 108(10):941-947.
- Manchester-Neesvig JB, Schauer JJ, Cass GR. 2003. The distribution of particle-phase organic compounds in the atmosphere and their use for source apportionment during the southern California children's health study. *Journal of the Air & Waste Management Association* 53(9):1065-1079.
- Mar TF, Norris GA, Koenig JQ, Larson TV. 2000. Associations between air pollution and mortality in Phoenix, 1995-1997. *Environmental Health Perspectives* 108(4):347-353.
- Metzger KB, Tolbert PE, Klein M, Peel JL, Flanders WD, Mulholland JA. 2004. Ambient air pollution and arrhythmic events in patients with implanted cardioverter defibrillators, 1993-2002. *Epidemiology* 15(4):S22-S23.
- Metzger KB, Tolbert PE, Klein M, Peel JL, Flanders WD, Todd K. 2004. Ambient air pollution and cardiovascular emergency department visits. *Epidemiology* 15(1):46-56.
- Peel JL, Tolbert PE, Klein M, Metzger KB, Flanders WD, Todd K. 2005. Ambient air pollution and respiratory emergency department visits. *Epidemiology* 16(2):164-174.
- Peters A, Dockery DW, Muller JE, Mittleman MA. 2001. Increased particulate air pollution and the triggering of myocardial infarction. *Circulation* 103(23):2810-2815.
- Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA-Journal of the American Medical Association* 287(9):1132-1141.
- Pope CA, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE. 1995. Particulate Air-Pollution as a Predictor of Mortality in a Prospective-Study of US

Adults. American Journal of Respiratory and Critical Care Medicine 151(3):669-674.

Russell A, Dennis R. 2000. NARSTO critical review of photochemical models and modeling. Atmospheric Environment 34(12-14):2283-2324.

Tsai FC, Apte MG, Daisey JM. 2000. An exploratory analysis of the relationship between mortality and the chemical composition of airborne particulate matter. Inhalation Toxicology 12:121-135.

US-EPA. 1998. CMB8 Application and Validation Protocol for PM_{2.5} and VOC; EPA-454/R-98-XXX; Office of Air Quality, Planning and Standards: Research Triangle Park, NC 27711.

US-EPA. 2001. CMB8 User's Manual, EPA-454/R-01-XXX, Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711.

Vedal S. 1997. Ambient particles and health: Lines that divide. Journal of the Air & Waste Management Association 47(5):551-581.

CHAPTER 2

TEMPORAL AND SPATIAL VARIABILITY IN MEASURED AND SIMULATED PM_{2.5} CONSTITUENTS IN ATLANTA, GA, AND IMPLICATIONS FOR TIME-SERIES HEALTH STUDIES

(A. Marmur, Y.Hu, J.A. Mulholland, P.E. Tolbert and A.G. Russell, prepared for the
Journal of the Air and Waste Management Association)

Abstract

Time-series health studies rely on the availability of long-term, accurate, spatially representative air quality data. This study examines whether an emissions-based air quality model (CMAQ) may be used, in lieu of measured data and/or to improve spatial representativeness compared to point measurements. Results from a long term air quality simulation are analyzed in terms of the model's ability to simulate temporal and spatial variability in concentrations of both secondary and primary PM_{2.5} components in Atlanta, GA, as part of an ongoing health study (ARIES). Seasonal variations in sulfate and nitrate concentrations were well captured by the model, but the model's ability to capture shorter-term (e.g., daily) variations, typically of interest in time-series health studies, was limited. Moreover, the spatial homogeneity in ambient concentrations of secondary PM_{2.5} constituents (such as sulfate and nitrate), suggests limited benefit in applying simulated concentration fields for these species in a time-series health study when ambient

measurements are available. Concentrations of primary PM_{2.5} constituents (such as elemental carbon), on the other hand, have much greater spatial variability, and short-term variability in these species is better captured by the air quality model. Thus, modeled concentrations of elemental carbon and organic carbon, which consists of primary and secondary components, may be more representative of a study area than point measurements. A comparison between estimates of secondary organic aerosol (SOA) concentrations from an air-quality model and from a Chemical Mass Balance analysis showed good agreement. As SOA formation chemistry continues to be studied and models updated accordingly, these may become useful tools for analyzing the health effects associated with SOA.

Key words: Air quality model, CMAQ, PM_{2.5}, epidemiologic study, health study, time series.

2.1 Introduction

In time-series studies of the effects of air pollution on human health, an association between a health endpoint of interest and short-term (e.g., daily) variability in species concentrations is sought. Typically, such studies are conducted using ambient measurements to represent the air quality over cities or regions. Specifically, many studies have been conducted to assess the health effects associated with fine particulate matter (aerodynamic diameter less than 2.5µm, PM_{2.5}) (Dockery et al. 1993; Pope et al. 2002; Pope et al. 1995), and more recently, health outcomes associated with species-specific PM_{2.5} are being investigated (Hauck et al. 2004; Heal et al. 2005; Metzger et al.

2004; Metzger et al. 2004; Peel et al. 2005). However, the use of ambient data for such studies introduces several limitations: monitoring sites might not adequately represent air quality over the health study domain (a single central site is often used) or capture the spatial variability in concentration fields; measurement errors, especially for complex measurements of PM_{2.5} species, might introduce noise to the epidemiologic analysis; incompleteness of data and lack of continuity in data might diminish the ability of the epidemiologic analysis to detect such health related associations; the need for large datasets often leads to conducting measurements over a multi-year period, but obtaining ambient measurements of PM_{2.5} species is costly and time-consuming. Use of emissions-based three-dimensional air quality models may introduce several benefits when applied in epidemiologic studies. First, an average value over a model cell of typical size (e.g., 36km x 36km or 12km x 12 km) may better represent the air quality over an applicable area as compared to a measurement at a single point within that area. Such a value can also assist in evaluating the representativeness and quality of measurements at different locations (stations) all residing within the range of the same model cell. Also, the ability to model episodes for which no measurements were performed may allow expanding the epidemiologic study to geographical areas for which no data were available, to past (historic) episodes, and to complete and extend existing datasets. Using air quality models is also less time-consuming and resource intensive as compared to the real-time required for measurements, often spanning several years. Air quality models can also provide information on source impacts, can simulate atmospheric concentrations for shorter time periods than many measurement techniques, and can simulate the levels of pollutants that are very difficult to measure (e.g., species found at very trace levels).

However, to be useful for time-series health studies, the air-quality model must be able to capture the temporal variability in pollutant concentrations. The focus on the temporal variability is different than in regulatory applications of air quality models, in which they are typically applied to compare the average response to changes in emissions for control strategy development.

The focus of this study is on application of an air-quality model to simulate PM_{2.5} components in Atlanta, GA, as part of an ongoing health study (ARIES: Aerosol Research Inhalation Epidemiologic Study). As part of ARIES, species specific PM_{2.5} health outcomes are being investigated (Metzger et al. 2004; Metzger et al. 2004; Peel et al. 2005), along with the temporal and spatial variability of PM_{2.5} species (Wade et al. 2004; Wade et al. 2006). The latter analysis indicated that the difference between measured ambient pollutant level and true ambient pollutant level, represented by the sum of instrument error and noise due to spatial variability, is greatest for primary PM_{2.5} constituents (EC and to a degree, OC), which can attenuate a health association with primary PM_{2.5} constituents. Currently, source-specific PM_{2.5} impacts are also being investigated, to assess whether specific sources of PM_{2.5} appear to be more related to specific health outcomes than others. As part of this analysis, source apportionment has been performed for the Atlanta aerosol using various methods, both receptor-based (such as Chemical Mass Balance) and emissions-based (such as Models-3 modeling system) (Kim et al. 2004; Marmur et al. 2006; Marmur et al. 2005). The current study evaluates air-quality model performance and addresses issues related specifically to application of such models in time-series health studies.

2.2 Methods

2.2.1 Air quality modeling

Fine particulate matter (PM_{2.5}) modeling was performed using components of the US-EPA's Models-3 modeling system, including the Penn-State/NCAR Meteorological Model (MM5) (Grell et al. 1999), the Carolina Environmental Program's (CEP) Sparse Matrix Operator Kernel Emissions (SMOKE) Modeling System version 1.5 (Houyoux et al. 2003), and CMAQ version 4.3, a three-dimensional (3-D) air quality model (Byun and Ching 1999). Speciated PM_{2.5} and gas phase pollutants were simulated for a three-year period, 1999-2001, using a grid of 36km by 36 km cells. The grid covered the entire eastern and central US, and was comprised of 78 by 66 cells laterally, and six vertical layers (Liu et al. 2006; Park et al. 2006). A sub-grid of 12km by 12km cells was placed over the northern part of Georgia, centered around Atlanta (14 by 14 cells laterally), and modeling using this grid was performed for the year 2001. Meteorological fields (e.g., temperature, relative humidity, three directional wind profiles etc.) were generated by MM5. Emissions from each grid cell were generated by SMOKE based on the 1999 National Emission Inventory, and subject to temporal trends (hour of day, day of week etc.) and meteorological parameters. Finally, pollutant concentrations, in the form of hourly averages, were calculated by CMAQ. The use of six vertical layers and might lead to underestimation of concentrations of PM_{2.5} components in cases where the actual mixing-height is lower than that of the model bottom layer. To address this issue, we also examined results from model simulations by the Visibility Improvement State and Tribal Association of the Southeast (VISTAS) (Morris et al. 2004), in which a finer vertical layer structure was used. Other differences include the use of the CBIV chemical

mechanism (instead of SAPRC99), and the inclusion of additional formation pathways for secondary organic aerosol (SOA) (Morris et al. 2006).

2.2.2 PM_{2.5} measurements

Total PM_{2.5} mass, major ions (SO_4^{2-} , NO_3^- , NH_4^+) and carbon fractions (elemental carbon, EC; organic carbon, OC) were measured daily at four locations in the Atlanta metropolitan area. Monitoring sites from which data were used are part of two different networks: SEARCH (Southeastern Aerosol Research and Characterization) network (Hansen et al. 2003), which includes the Jefferson Street (JST) site in Atlanta, and ASACA (Assessment of Spatial Aerosol Composition in Atlanta) network (Butler et al. 2003), which includes the South Dekalb (SD), Fort McPherson (FM) and Tucker (TU) sites (Figure 2-1). Measurements began in August 1998 at JST, and in August 1999 at the ASACA sites. A third site used was the South Dekalb Speciation Trends Network (STN) site, co-located at the SD ASACA site; however, data collection at the STN site started only on March 2001, on an every third day basis. For comparison with model simulations, the ambient data from August 1999-August 2001 were used (data from the ASACA network was not available for the latter part of 2001 and 2002). Results from the VISTAS modeling (available for 2002 only) were compared with the ambient data from JST and with the SD-STN data.

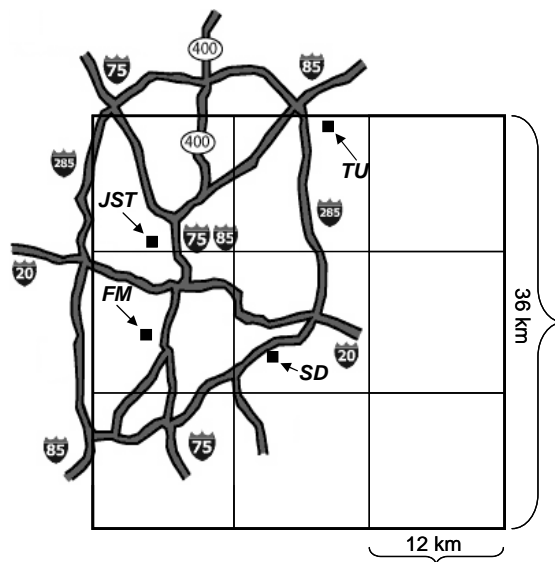


Figure 2-1. Location of SEARCH and ASACA PM_{2.5} monitoring stations in the Atlanta Metropolitan, overlaid by model grid cells corresponding to these sites (36 km and 12 km domains). Jefferson St. (JST) is a part of the SEARCH network; Fort McPherson (FM), South-Dekalb (SD), Tucker (TU) are a part of the ASACA network. In 2001, a PM_{2.5} STN site was co-located at the South-Dekalb site. All sites are co-located within one 36 km model grid-cell, but are located in different 12 km grid cells.

Total PM_{2.5} mass was measured by means of TEOM (Tapered Element Oscillating Microbalance). The JST site included also a gravimetric measurement of total PM_{2.5} mass. For the speciation of the PM_{2.5}, a manual, filter-based Particle Composition Monitor (PCM) was operated on a daily schedule. The PCM included three channels to collect 24 hour integrated samples for analysis of major ions, trace metals and organic and elemental carbon in the PM_{2.5} size range. Ion chromatography (IC) was used to quantify water soluble ionic species. Elemental and organic carbon collected on quartz filters were measured by Thermal Optical Transmittance (TOT) in the ASACA network, and by Thermal Optical Reflectance (TOR) in the SEARCH network (Butler et al. 2003; Hansen et al. 2003). Comparison of these two techniques indicates that while the total

carbon (TC) measurements are in good agreement, lower EC and higher OC values are obtained using TOT, compared to TOR (Chow et al. 2004).

2.3 Results

2.3.1 Air-quality model simulations

Concentrations of SO_4^{-2} , NO_3^- , NH_4^+ , EC, OC simulated using CMAQ for ARIES were compared with observations at four different sites in the Atlanta metropolitan area (JST, FM, SD, TU). All four sites used here are located within the same 36km CMAQ grid cell (Figure 2-1), allowing two issues to be addressed: evaluation of CMAQ performance (using a 36km grid resolution), and to suggest whether a single site exists which is more representative of the health study domain (the Atlanta metropolitan area) and hence more suitable for use in the epidemiologic study. Comparing model performance for major $\text{PM}_{2.5}$ components (Table 2-1) indicates that performance is better for sulfate than for nitrate and ammonium. Average sulfate concentrations are well simulated (same average for the JST site, slight underprediction for the other sites), and the correlation coefficient at JST is relatively high (0.72, 0.59-0.61 at the other sites). On the other hand, nitrate is substantially overpredicted by CMAQ, and the daily variations are not captured as well (as indicated by the relatively low correlations, in the range of 0.45-0.52). Similar findings have been reported elsewhere (Morris et al. 2005). EC and OC are underpredicted in this application of CMAQ, likely due to the relatively coarse vertical layer structure. This may cause artificial dispersion of primary $\text{PM}_{2.5}$ emitted locally at the ground level (such as EC and a fraction of OC), especially during strong, low-level inversion events. In contrast, sulfate levels are well predicted using the same

vertical layer structure, being a regional secondary pollutant. JST seems to be the most consistent with the model simulated concentrations when comparing model performance at the four sites. This is evident in the higher correlation and lower RMSE for sulfate, compared to FM, TU and SD sites, as well as in the higher correlations of EC and OC. This may be an indication of JST being more spatially representative than the other sites, or of better data quality (smaller measurement error) compared the other sites/datasets. To address the issue of grid resolution, results from a simulation using a 12km grid were used, but model performance did not change substantially (Table 2-2).

Table 2-1. Model performance statistics for 36km CMAQ modeling of the Atlanta aerosol (for ARIES). For each site, the average daily concentration for the period of 8/99-8/01 is given (obs., $\mu\text{g}/\text{m}^3$), along with the correlation (R), and the root mean square error (RMSE, $\mu\text{g}/\text{m}^3$) compared with the daily CMAQ simulated concentrations ($\mu\text{g}/\text{m}^3$). Data for the SD site are from the ASACA network.

Species	CMAQ (avg. modeled conc.)	JST (obs./ R/ RMSE)	FM (obs. /R /RMSE)	TU (obs. /R /RMSE)	SD (ASACA) (obs. /R /RMSE)
SO ₄ ⁻²	5.2	5.2/ 0.72/ 2.7	4.7/ 0.60/ 3.1	4.6/ 0.61/ 3.1	4.6/ 0.59/ 3.2
NO ₃ ⁻	3.4	1.1/ 0.52/ 4.2	1.0/ 0.50/ 4.9	1.2/ 0.50/ 4.8	0.90/ 0.45/ 4.8
NH ₄ ⁺	2.8	2.7/ 0.42/ 1.6	1.7/ 0.44/ 2.0	1.7/ 0.42/ 2.0	1.6/ 0.44/ 2.1
EC	1.1	1.6/ 0.66/ 1.0	1.4/ 0.40/ 0.74	1.2/ 0.31/ 0.67	1.7/ 0.52/ 1.0
OC	2.6	4.2/ 0.58/ 2.5	4.5/ 0.43/ 3.3	4.6/ 0.45/ 3.2	4.7/ 0.47/ 3.3

Table 2-2. Comparison between model performance using a 36km and a 12km modeling grid (1/01-8/01), for ARIES. For each site, the correlation coefficients between observed and modeled values (both at 36km and 12 km), along with the RMSE ($\mu\text{g}/\text{m}^3$, calculated for both 36 km and 12 km modeling) are given. Data for the SD site are from the ASACA network.

Species	JST		FM		TU		SD (ASACA)	
	R (36/12)	RMSE (36/12)	R (36/12)	RMSE (36/12)	R (36/12)	RMSE (36/12)	R (36/12)	RMSE (36/12)
SO ₄ ⁻²	0.81/0.79	2.0/2.1	0.66/0.66	2.4/2.5	0.70/0.71	2.2/2.3	0.65/0.62	2.5/2.6
NO ₃ ⁻	0.60/0.57	3.1/3.2	0.72/0.71	3.3/3.2	0.70/0.69	3.1/3.2	0.61/0.61	3.4/3.3
NH ₄ ⁺	0.40/0.41	1.3/1.3	0.49/0.50	1.5/1.7	0.54/0.53	1.4/1.5	0.50/0.49	1.5/1.5
EC	0.58/0.59	0.9/0.8	0.31/0.34	0.9/0.8	0.30/0.30	0.7/0.7	0.50/0.47	1.0/1.0
OC	0.56/0.59	1.9/1.8	0.37/0.36	3.9/4.1	0.46/0.44	3.8/3.5	0.46/0.45	3.8/3.9

We also compared model simulations from the VISTAS 12 km modeling (Morris et al. 2004; Morris et al. 2005), in which a finer vertical layer structure was used (19 layers), with observations during 2002. In this modeling campaign, the CB-IV chemical mechanism and an updated SOA module (Morris et al. 2006) were used. Model performance (Table 2-3) is somewhat improved compared to the results presented previously (Tables 2-1 and 2-2), but the general trends remain the same, with sulfate performance being the highest, followed by EC/OC, and nitrate and ammonium performance being the poorest. However, these results do not exhibit an underprediction of carbonaceous PM_{2.5} components, due to a combination of more detailed vertical layering (EC and primary fraction of OC), and an enhanced SOA module which includes additional SOA formation and polymerization processes (Morris et al. 2006). Another difference is the improved model performance at the SD grid cell (data from the STN site were used), but similar patterns were also observed when the CMAQ modeling done as part of ARIES was evaluated using the available STN data for 2001 (instead of the ASACA data). Based on these findings from the model evaluation process, there is no evidence to support that the JST site is more/less spatially representative than the SD-STN site.

Table 2-3. Model performance statistics for VISTAS 12km modeling (Morris et al. 2004; Morris et al. 2005) of the Atlanta aerosol. For each site, the average daily concentration for the period of 1/02-12/02 is given (observed, $\mu\text{g}/\text{m}^3$), along with the correlation (R), and the root mean square error (RMSE, $\mu\text{g}/\text{m}^3$) compared with the daily CMAQ simulated concentrations (modeled, $\mu\text{g}/\text{m}^3$). Data for SD site are from the EPA STN.

Species	JST				SD (EPA STN)			
	Observed	Modeled	R	RMSE	Observed	Modeled	R	RMSE
SO_4^{-2}	4.3	5.2	0.68	2.8	4.6	4.8	0.82	1.7
NO_3^-	1.0	1.8	0.65	1.9	0.9	1.5	0.60	1.8
NH_4^+	2.1	2.1	0.53	1.0	1.3	1.9	0.62	0.9
EC	1.3	1.5	0.63	0.8	0.9	1.1	0.75	0.5
OC	4.2	5.6	0.63	2.6	5.4	5.2	0.70	2.0

2.3.2 Daily vs. seasonal variations

Correlation coefficients reported in Tables 2-1 - 2-3 represent temporal variations on both the daily scale (24 hour data are used for the comparison) and seasonal scale. As previously mentioned, to be applied in a time-series health study, it is important that the model capture the daily variability in concentrations of $\text{PM}_{2.5}$ components (seasonal variability is typically controlled for in the health study). Previous studies indicated that meteorological (MM5) and air-quality (CMAQ, applied to ozone) models capture inter-annual (seasonal) and synoptic-scale (lasting several days and longer) variability, while the magnitude of fluctuations on shorter time scales is underestimated (Hogrefe et al. 2004; Hogrefe et al. 2001). To address this issue specifically for $\text{PM}_{2.5}$ components, we computed correlations for running blocks of 30 days each during the period of 8/99-8/01 (731 blocks during a 762 day period). This allowed evaluating the variability captured by the model on shorter time scales, emphasizing daily and synoptic variability over seasonal/interannual variability. The average of these shorter time-scale correlation coefficients was compared to the correlations based on the entire dataset (Table 2-4).

Also provided are similar comparisons for observations at JST vs. SD-STN (3/01-12/02). For species with a strong seasonal pattern, such as sulfate and nitrate, the correlations between simulated and observed concentrations drop sharply when shorter time scales are used. For example, the correlation coefficient for sulfate based on the entire dataset is 0.72, and only 0.56 based on the average of 25 monthly coefficients. A similar pattern is observed for nitrate. This means that the model better simulates seasonal patterns (e.g., high sulfate levels in summer and low sulfate levels in winter) than daily fluctuations in concentrations (Figure 2-2). For species lacking a strong seasonal pattern, such as EC (Figure 2-3) and OC, no substantial difference between the two coefficients is observed. This implies that the previously reported correlation coefficients are likely driven by the model's ability to simulate daily fluctuations in concentrations of EC and OC. Also shown in Table 2-4 are correlation coefficients using the measurements at JST and SD-STN sites. These results address whether inter-site correlations are also driven by seasonal patterns for the major PM_{2.5} constituents. The long-term and short-term correlation coefficients between sites are similar, with the largest difference observed for nitrate. This result suggests that the daily variability is as spatially correlated as the seasonal variability. The amount of short-term temporal variability (as represented by the variance) captured by the model relative to that common to both sites (inter-site correlation) for the various species indicates that the model captures more of the short-term variability of primary components than that of secondary components. For example, the model captures 74% of the temporal variability (represented by the variance) in EC concentrations, compared to only 34% of the temporal variability in sulfate concentrations. Although different periods were used for model versus JST comparison

and the inter-site comparison due to availability of data, the results are informative on a comparative basis. Hence, despite similar short-term correlations between model results and JST measurements for EC and sulfate (0.64 and 0.56, respectively), we estimate that the model is capturing a substantially larger fraction of the spatially representative short-term variation in EC (75%) than in sulfate (34%).

Table 2-4 Comparison between correlation coefficients (R) (modeled vs. observed concentrations at JST, 8/99-8/01) based on the entire simulated dataset (daily and seasonal patterns represented) and shorter (30 day) subparts of the dataset (emphasis on daily variations). Also provided is a similar comparison for correlations between observations at JST and SD-STN sites (3/01-12/02)

Species	CMAQ (modeled) vs. JST (obs.), 8/99-8/01		JST (obs.) vs. SD-STN (obs.), 3/01-12/02		Short-term temporal variability ¹ captured by the model ² (%)
	Correlation based on entire dataset	Average of 30 day correlations	Correlation based on entire dataset	Average of 30 day correlations	
SO ₄ ⁻²	0.72	0.56	0.98	0.96	34% (e.g., 0.56 ² /0.96 ²)
NO ₃ ⁻	0.52	0.41	0.91	0.80	26%
NH ₄ ⁺	0.42	0.48	0.77	0.81	35%
EC	0.66	0.64	0.77	0.74	75%
OC	0.58	0.65	0.79	0.80	66%

¹ - represented by the variance

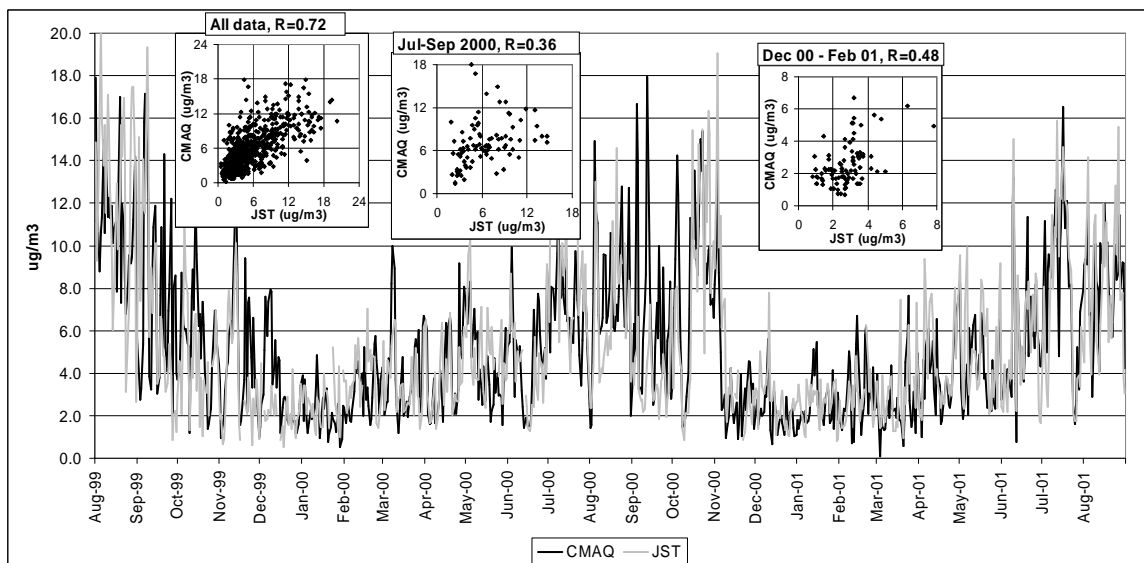


Figure 2-2. Comparison between daily modeled (36km grid) and measured (JST) sulfate concentrations, for the period of August 1999 – August 2001. Seasonal patterns are well captured by the model, but model performance is poor when shorter time scales are

considered. The correlation coefficient (R) for all data is 0.72; 0.36 for the period of July-September 2000; and 0.48 for the period of December 2000 - February 2001.

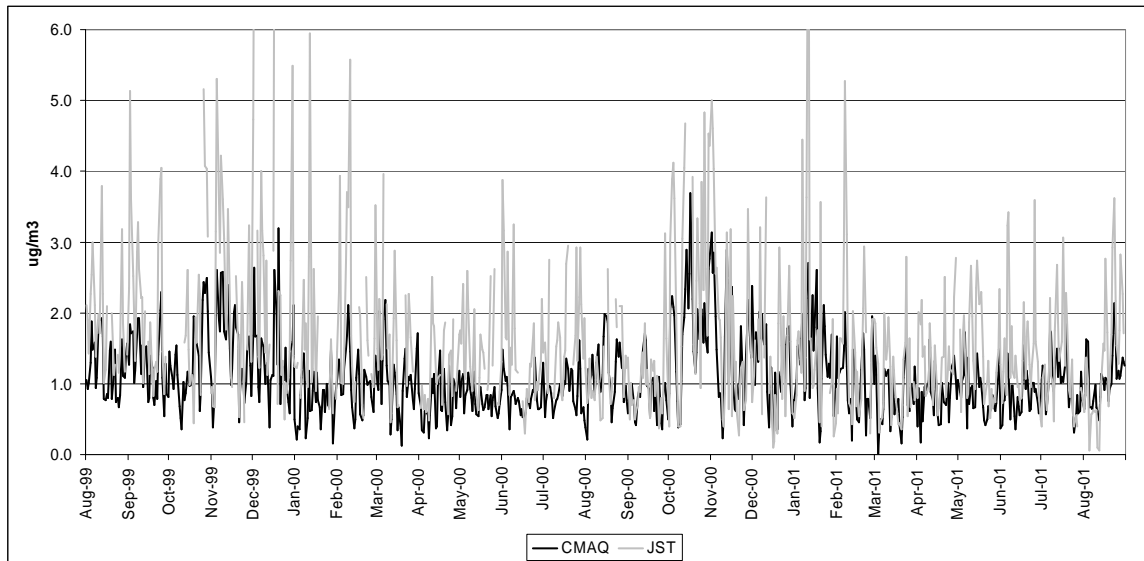


Figure 2-3. Comparison between daily modeled (36km grid) and measured (JST) EC concentrations, for the period of August 1999 – August 2001. Modeled EC is generally under predicted. With regards to temporal variations, in lieu of a strong seasonal pattern, no substantial difference in model performance is observed between shorter and longer time scales (as also indicated in Table 2-5).

2.3.3 Spatial representativeness of modeled grid-cell concentrations vs. point

measurements

Ambient measurements at a single location might not be representative of air-quality over a larger domain if the measurements are influenced by local sources. This is of special concern for primary $PM_{2.5}$ constituents, such as EC, as demonstrated by the data in Table 2-4. In contrast, air-quality models calculate volume average concentrations. These differences in spatial scales can contribute to the discrepancies between measured and simulated concentrations. Park et al. (2006) demonstrated that

model performance improves when it is compared with ambient data from several measurement sites (weighting for distance from the central site), compared to a single site, and that the improvement was the largest for primary $\text{PM}_{2.5}$ components. To address the issue of spatial representativeness within the Atlanta metropolitan area, we compared model performance with single site measurements and an average of sites, JST and SD-STN. The question then is whether a spatial average of data from these two sites is more correlated with the model than any single measurement. Given only two sites, and the fact that they both reside in the same 36km grid cell, we used a simple average of the two sites, for the period 3/01-12/01 (the overlapping period between the 36km modeling and the SD-STN measurements). For elemental carbon, the correlation between CMAQ modeled concentrations and measurements at JST was 0.69, the correlation between the model and the SD-STN data was 0.66, and the correlation with the average of these two sites was 0.72. These differences are subtle, likely indicating that neither of these sites is strongly influenced by local sources (the intersite correlation was 0.80), and that the model is only slightly more spatially representative than any of these single measurements. For sulfate, these correlations were 0.79, 0.80 and 0.80, respectively, with an intersite correlation of 0.98, indicating that the model is not more spatially representative than any single measurement.

2.3.3 Secondary organic aerosol

Of the secondary $\text{PM}_{2.5}$ components, secondary organic aerosol (SOA) is of special interest, as it comprises a large fraction of ambient $\text{PM}_{2.5}$, but cannot be measured directly. Thus, in order to study the health effects of SOA, indirect methods must be

applied. One such common method is the EC-tracer method (Turpin and Huntzicker 1995), in which ambient data are analyzed to identify an OC/EC ratio typical of emissions in the region, and secondary OC is then estimated as the difference between total OC and the product of EC and this primary OC/EC ratio. Another approach is to apply a Chemical Mass Balance (CMB) model, and to estimate secondary OC as the difference between total OC and OC apportioned to emission sources (Marmur et al. 2005; Zheng et al. 2002). This can be viewed as an “expanded” EC-tracer method, as the primary OC/EC ratio can vary based on the daily fluctuations in source impacts (e.g., biomass burning and diesel emissions have very different OC/EC ratios, and as their relative impact on the monitor changes, so will the primary OC/EC ratio). One drawback of this approach, however, is that this estimate is more accurately an estimate of unapportioned OC, which might not necessarily be entirely SOA (any sources of primary OC not included in the CMB analysis may be lumped into this category). Further, errors in the CMB source apportionment will impact the SOA allocation in an unknown fashion. A conceptually different approach would be to use an emissions-based air quality model, such as CMAQ, to generate concentration fields of SOA. Since the chemistry of SOA formation is not fully understood, not all processes leading to SOA formation are currently included in models such as CMAQ. In addition, it is not trivial to evaluate the performance of SOA modules, as a direct comparison with ambient data is not yet possible. Model performance for total OC is often evaluated, but it is then difficult to distinguish between performance for primary and secondary OC (e.g., overprediction of primary OC may compensate for an underprediction in SOA etc.). Morris et al. (2006) developed an enhanced SOA module for CMAQ, which includes the formation of SOA

from sesquiterpenes and isoprene, in addition to the monoterpenes oxidation processes accounted for in the current version of CMAQ. In addition, this module includes polymerization of SOA into non-volatile particles. Using this enhanced SOA module, Morris et al. (2006) report substantial improvements in model performance for OC. Even with the inclusion of these processes in the SOA module, much remains unknown regarding the chemistry leading to SOA formation. Recent evidence suggests a substantial contribution of anthropogenic volatile organic compounds (VOC) to SOA formation in the Atlanta area (Sullivan, 2006), in contrast to findings from CMAQ, in which the vast majority of SOA is formed from biogenic precursors (e.g., 90% of SOA is of biogenic origin, on an annual basis, based on modeling results presented here). Here we compare results from the VISTAS simulations (daily values for 2002) with results from a CMB analysis (Marmur et al. 2006; Marmur et al. 2005). A comparison of unapportioned organic matter (OM) (OC not apportioned to any of the sources of primary OC; multiplied by 1.4 to account for OM/OC ratio) and SOA from CMAQ shows similar trends (Figure 2-4) except for several outliers in the CMB analysis (extremely high unapportioned OM levels on specific winter days). Focusing on summertime (April-October), CMAQ SOA is slightly overestimated compared to the CMB unapportioned OM (4.3 vs. 3.6 $\mu\text{g}/\text{m}^3$, respectively), though an OM/OC ratio higher than 1.4 (used in the CMB analysis), as may be more suitable for SOA (Turpin and Lim 2001) would have made these more similar. The correlation coefficient (R) between these two methods is 0.70, slightly higher than the coefficient for total OC (Table 2-3). Comparing this correlation coefficient for all data between April-October 2002 with an average of the correlation coefficients during running 30 day blocks (to dampen out seasonal/long-term

effects) revealed no difference (both are equal to 0.70). Thus, there is substantial agreement in the daily variability of SOA levels estimated by these two methods, for the time period considered. This finding is somewhat surprising, given the limited understanding of SOA formation chemistry. However, as SOA formation chemistry continues to be studied and models updated accordingly, these may become useful tools in health studies.

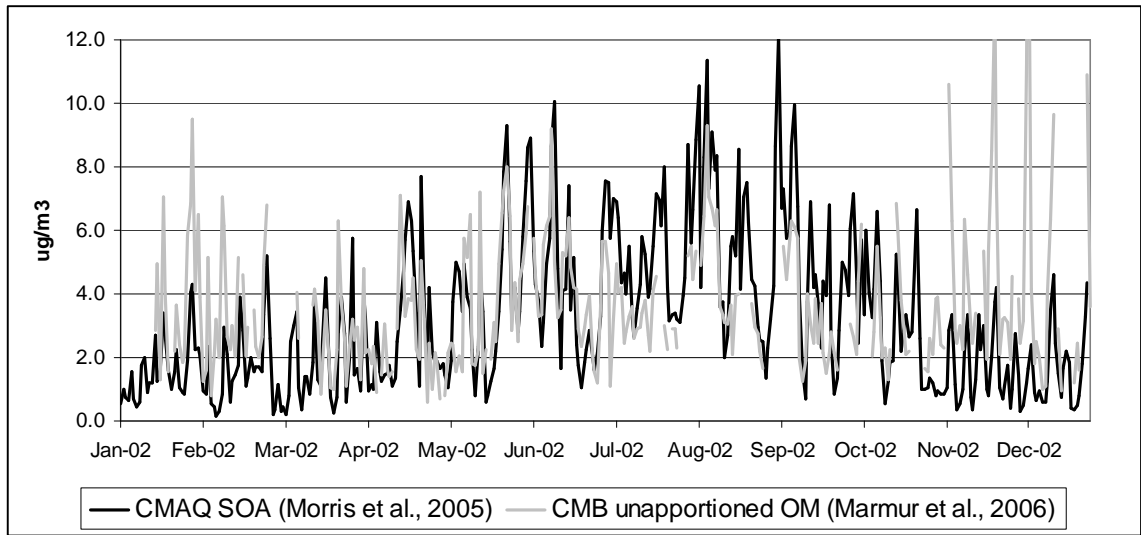


Figure 2-4. Comparison between CMAQ modeled SOA (Morris et al., 2005) and unapportioned OM from a CMB analysis (Marmur et al., 2006) for 2002 at the JST site

2.4 Discussion

An analysis of spatial and temporal variability in concentration fields (Wade et al. 2006) has indicated that the spatial variation in concentrations of secondary $PM_{2.5}$ components is lower than that of primary components. This can also be observed in the inter-site correlations presented here (JST and SD-STN, Table 2-4), being higher for sulfate (0.98) than for EC (0.77). This implies that the error introduced into the epidemiologic analysis due to spatial variability of ambient concentration fields would be

higher for primary than for secondary PM_{2.5} components, possibly attenuating observed associations with health-outcomes. In practice, these findings may also imply that the location of measurement is not of great significance when secondary PM_{2.5} components are of interest (e.g., measurements at any sampling site within the health study domain would likely be representative of the daily variability in sulfate concentrations), while the sampling location may be of greater significance when primary PM_{2.5} components are of interest (different sites might exhibit different daily variability in concentrations, possibly influencing the epidemiologic analysis). In terms of applying air-quality models in the epidemiologic analysis, these data suggest that such an application might not introduce much benefit for secondary PM_{2.5} components, if ambient data are available. A sensitivity analysis of the epidemiologic model would be needed in order to estimate the effect of dampened temporal variability of a given (simulated) pollutant on the derived health risk-ratios. Such information may be helpful in determining whether applying air-quality models in-lieu of ambient data would serve as a reasonable surrogate. Based on the results presented here (Table 2-4), the ability of CMAQ to model short term variations in concentrations of sulfate, nitrate and ammonium is limited, while for EC and OC, the model does capture more of the short-term variability in these concentrations. Regarding spatial representativeness, model performance has been shown to improve when model results are compared with a spatial average of ambient data, rather than one site, especially for primary PM_{2.5} components (slightly in this study, more substantially in (Park et al. 2006)). This implies that a volume average concentration, as from the output of an air-quality model, may be more spatially representative than any individual point measurement. These findings regarding the spatial and temporal variability suggest that

there may be some benefit in applying simulated concentrations of EC (and to a degree, OC) in time-series health studies, if the increased spatial representativeness outweighs the loss of some of the temporal variability.

Acknowledgements

This work was supported by subcontracts to Emory University under grants from the U.S. Environmental Protection Agency (R82921301-0) and the National Institute of Environmental Health Sciences (R01ES11199 and R01ES11294). We would also like to thank Georgia Power and Southern Company for continuing support and the individuals at ARA (Atmospheric Research and Analysis) for both providing access to data used in this analysis and ongoing discussions.

References

- Butler AJ, Andrew MS, Russell AG. 2003. Daily sampling of PM_{2.5} in Atlanta: results of the first year of the assessment of spatial aerosol composition in Atlanta study. *Journal of Geophysical Research-Atmospheres* 108(D1).
- Byun DW, Ching JKS. 1999. Science Algorithms of the EPA-Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, Office of Research and Development, EPA/600/R-99/030.
- Chow JC, Watson JG, Chen LWA, Arnott WP, Moosmuller H. 2004. Equivalence of elemental carbon by thermal/optical reflectance and transmittance with different temperature protocols. *Environmental Science & Technology* 38(16):4414-4422.
- Dockery DW, Pope CA, Xu XP, Spengler JD, Ware JH, Fay ME. 1993. An Association Between Air-Pollution And Mortality In 6 United-States Cities. *New England Journal of Medicine* 329(24):1753-1759.

- Grell GA, Dudhia J, Stauffer DR. 1999. A description of the fifth-generation Penn State/NCAR mesoscale model (MM5).
- Hansen DA, Edgerton ES, Hartsell BE, Jansen JJ, Kandasamy N, Hidy GM. 2003. The southeastern aerosol research and characterization study: Part 1-overview. *Journal of the Air & Waste Management Association* 53(12):1460-1471.
- Hauck H, Berner A, Frischer T, Gomiscek B, Kundi M, Neuberger M, et al. 2004. AUPHEP-Austrian Project on Health Effects of Particulates - general overview. *Atmospheric Environment* 38(24):3905-3915.
- Heal MR, Hibbs LR, Agius RM, Beverland LJ. 2005. Total and water-soluble trace metal content of urban background PM₁₀, PM_{2.5} and black smoke in Edinburgh, UK. *Atmospheric Environment* 39(8):1417-1430.
- Hogrefe C, Biswas J, Lynn B, Civerolo K, Ku JY, Rosenthal J, et al. 2004. Simulating regional-scale ozone climatology over the eastern United States: model evaluation results. *Atmospheric Environment* 38(17):2627-2638.
- Hogrefe C, Rao ST, Kasibhatla P, Hao W, Sistla G, Mathur R, et al. 2001. Evaluating the performance of regional-scale photochemical modeling systems: Part II - ozone predictions. *Atmospheric Environment* 35(24):4175-4188.
- Houyoux M, Vukovich J, Brandmeyer JE, Seppanen C, Holland A. 2003. Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE): User Manual, Version 2.0.
- Kim E, Hopke PK, Edgerton ES. 2004. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. *Atmospheric Environment* 38(20):3349-3362.
- Liu W, Wang Y, Marmur A, Russell AG, Edgerton E. 2006. Evaluation of CMAQ simulated atmospheric constituents with SEARCH observations in the factor projected space. *J Geophys Res* (submitted).
- Marmur A, Park SK, Mulholland JA, P.E. T, Russell AG. 2006. Source Apportionment of PM_{2.5} in the Southeastern United States Using Receptor and Emissions-based Models: Conceptual Differences and Implications for Time- Series Health Studies. *Atmospheric Environment* (in print).

- Marmur A, Unal A, Mulholland JA, Russell AG. 2005. Optimization Based Source Apportionment of PM_{2.5} Incorporating Gas-to-Particle Ratios. *Environmental Science and Technology* 39(9):3245-3254.
- Metzger KB, Tolbert PE, Klein M, Peel JL, Flanders WD, Mulholland JA. 2004. Ambient air pollution and arrhythmic events in patients with implanted cardioverter defibrillators, 1993-2002. *Epidemiology* 15(4):S22-S23.
- Metzger KB, Tolbert PE, Klein M, Peel JL, Flanders WD, Todd K. 2004. Ambient air pollution and cardiovascular emergency department visits. *Epidemiology* 15(1):46-56.
- Morris RE, Koo B, Guenther A, Yarwood G, McNally DE, Tesche TW, et al. 2006. Model Sensitivity Evaluation for Organic Carbon Using Two Multi-pollutant Air Quality Models that Simulate Regional Haze in the Southeastern United States. *Atmospheric Environment* (in press).
- Morris RE, Koo B, Tesche TW, Loomis C, Stella G, Tonnesen G, et al. 2004. Modeling Protocol for the VISTAS Phase II Regional Haze Modeling (http://pah.cert.ucr.edu/vistas2/reports/vistasII_Final_Protocol_05_10_2004.pdf, accessed 1/5/2006).
- Morris RE, McNally DE, Tesche TW, Tonnesen G, Boylan JW, Brewer P. 2005. Preliminary evaluation of the community multiscale air, quality model for 2002 over the southeastern United States. *Journal Of The Air & Waste Management Association* 55(11):1694-1708.
- Park SK, Cobb CE, Wade K, Mulholland JA, Hu Y, Russell AG. 2006. Uncertainty in Air Quality Model Evaluation for Particulate Matter due to Spatial Variation in Pollutant Concentrations. *Atmospheric Environment* (in press).
- Park SK, Marmur A, Kim SB, Tian D, McMurry P, Russell AG. 2006. Evaluation of fine particle number concentrations in CMAQ. *Aerosol Science And Technology* (in press).
- Peel JL, Tolbert PE, Klein M, Metzger KB, Flanders WD, Todd K. 2005. Ambient air pollution and respiratory emergency department visits. *Epidemiology* 16(2):164-174.

- Pope CA, Burnett RT, Thun MJ, Calle EE, Krewski D, Ito K. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA-Journal of the American Medical Association* 287(9):1132-1141.
- Pope CA, Thun MJ, Namboodiri MM, Dockery DW, Evans JS, Speizer FE. 1995. Particulate Air-Pollution as a Predictor of Mortality in a Prospective-Study of US Adults. *American Journal of Respiratory and Critical Care Medicine* 151(3):669-674.
- Sullivan AM. 2006. The ambient organic aerosol soluble in water: measurements, chemical characterization, and an investigation of sources. Ph.D. dissertation, School of Earth & Atmospheric Sciences, Georgia Institute of Technology
- Turpin BJ, Huntzicker JJ. 1995. Identification Of Secondary Organic Aerosol Episodes And Quantitation Of Primary And Secondary Organic Aerosol Concentrations During SCAQS. *Atmospheric Environment* 29(23):3527-3544.
- Turpin BJ, Lim HJ. 2001. Species contributions to PM_{2.5} mass concentrations: Revisiting common assumptions for estimating organic mass. *Aerosol Science And Technology* 35(1):602-610.
- Wade KS, Marmur A, Mulholland JA, Russell AG, Peel JL, Klein M, et al. 2004. Analyses of air quality data for an emergency department study, 1993-2002. *Epidemiology* 15(4):S61-S61.
- Wade KS, Mulholland JA, Marmur A, Russell AG, Hartsell B, Edgerton E, et al. 2006. Effects of instrument precision and spatial variability on the assessment of the temporal variation of ambient air pollution in Atlanta, Georgia. *Journal of the Air & Waste Management Association* (in press).
- Zheng M, Cass GR, Schauer JJ, Edgerton ES. 2002. Source apportionment of PM_{2.5} in the southeastern United States using solvent-extractable organic compounds as tracers. *Environmental Science & Technology* 36(11):2361-2371.

CHAPTER 3

OPTIMIZATION BASED SOURCE APPORTIONMENT OF PM_{2.5}

INCORPORATING GAS-TO-PARTICLE RATIOS

(A. Marmur, A. Unal, J.A. Mulholland, and A.G. Russell, *Environmental Science and Technology*, 39, 3245-3254, 2005)

Abstract

A modified approach to PM_{2.5} source apportionment is developed, using source indicative SO₂/PM_{2.5}, CO/PM_{2.5} and NO_x/PM_{2.5} ratios as constraints, in addition to the commonly used particulate-phase source profiles. Additional information from using gas-to-particle ratios assists in reducing collinearity between source profiles, a problem that often limits the source-identification capabilities and accuracy of traditional receptor models. This is especially true in the absence of speciated organic-carbon measurements. In the approach presented here, the solution is based on a global optimization mechanism, minimizing the weighted-error between apportioned and ambient levels of PM_{2.5} components, while introducing constraints on calculated source contributions that assure that the ambient gas-phase pollutants (SO₂, CO and NO_y) are reasonable. This technique was applied to a 25 month dataset of daily PM_{2.5} measurements (total mass and composition) at the Atlanta Jefferson Street SEARCH site. Results indicate that this technique was able to split the contributions of mobile sources (gasoline and diesel vehicles) more accurately than particulate-phase source apportionment methods.

Furthermore, this technique was able to better quantify the direct contribution (primary PM_{2.5}) of coal-fired power-plants to ambient PM_{2.5} levels.

Keywords: CMB, LGO, optimization, source-apportionment, PM_{2.5}, two-phase-source-profile.

3.1 Introduction

Chemical Mass Balance (CMB) receptor models are a common tool for apportioning of ambient levels of pollutants (mainly particulate matter) among the major contributing sources. CMB combines the chemical and physical characteristics of particles or gases measured at sources and receptors to quantify the source contributions to the receptor. The quantification is based on the solution to a set of linear equations that express each receptor's ambient chemical concentration as a linear sum of products of source-profile abundances and source contributions (1,2), as expressed by:

$$C_i = \sum_{j=1}^n f_{ij} S_j + e_i \quad (3-1)$$

where:

C_i = ambient concentration of chemical species i ($\mu\text{g}/\text{m}^3$);

$f_{i,j}$ = fraction of species i in emissions from source j ;

S_j = contribution (source-strength) of source j ($\mu\text{g}/\text{m}^3$);

n = total number of sources;

e_i = error term;

The source profile abundances (f_{ij} , the mass fraction of a chemical in the emissions from each source type) and the receptor concentrations (C_i), along with uncertainty estimates, serve as input data to the CMB model. The output consists of the contribution of each source category (S_j) to the measured concentration of different species at the receptor.

In CMB8 (2), the effective variance (EV) weighing for least squares calculations is applied, to find the best solution to the set of equations given by Equation 1. The effective weighing method takes into account both the uncertainties in the ambient measurements and the uncertainties in the source-profile compositions. In practice, CMB8 performs a series of matrix operations to minimize χ^2 , given as (3):

$$\chi^2 = \sum_{i=1}^m \frac{\left(C_i - \sum_{j=1}^n f_{ij} S_j \right)^2}{\sigma_{C_i}^2 + \sum_{j=1}^n \sigma_{f_{ij}}^2 S_j^2} \quad (3-2)$$

where:

σ_{C_i} = one standard deviation precision of the C_i measurement;

$\sigma_{f_{ij}}$ = one standard deviation of the f_{ij} measurement;

m = total number of species;

If the $\sigma_{f_{ij}}$ are set to zero, the solution reduces to the ordinary weighted least square (OWLS) solution (3), taking only the uncertainties in the ambient measurements into account.

CMB models are based on the following assumptions (2):

1. Compositions of source emissions are constant over the period of ambient and source sampling.
2. Chemical species do not react with each other, i.e., they add linearly.
3. All sources with a potential for significantly contributing to the receptor are included in the analysis.
4. The source compositions are linearly independent of each other.
5. The number of sources or source categories is less than or equal to the number of chemical species.
6. Measurement uncertainties are random, uncorrelated, and normally distributed.

Of these, one of the major assumptions limiting the ability of CMB models to identify and quantitatively provide impacts of the major sources is the linear independence of source profiles, when those profiles are based solely on traditional species. For apportionment of PM_{2.5} (particulate matter with a diameter less than 2.5µm), source profiles including major ions (SO₄⁻², NO₃⁻, NH₄⁺, Cl⁻), elemental and organic carbon fractions (EC, OC) and trace metals are typically used. Some source categories share relatively similar profiles (e.g. diesel and gasoline vehicles), limiting the ability of CMB to accurately and consistently apportion the PM mass between those sources, particularly in the presence of other sources of OC and EC. To address this issue, recent source-apportionment studies make use of speciated organic compounds (“organic markers”) to apportion OC (4-6), a major component in emissions from mobile sources, vegetative burning and meat-charbroiling. However, some sources share organic markers (e.g.,

hopanes and steranes in both gasoline and diesel vehicles), making it difficult to accurately and consistently apportion the OC mass between those sources. In addition, speciated ambient OC data are not yet commonly available.

3.2 Model description

3.2.1 Incorporating gas-to-particle ratios in PM_{2.5} source-apportionment

Here we apply an extended CMB approach for PM_{2.5} source-apportionment which incorporates source-indicative SO₂/PM_{2.5}, CO/PM_{2.5} and NO_x/PM_{2.5} ratios, in addition to the commonly used PM_{2.5} source profiles. Such ratios, along with ambient gas phase data, can further assist in identifying sources, as sources that may have fairly similar PM_{2.5} emissions, may have significantly different gaseous emissions. Such gas-to-particle ratios may vary during transport from source to receptor, due to different deposition rates and reactivity. However, the atmospheric lifetimes of SO₂ (about a week), CO (1-4 months), and NO_x (1-7 days) (7) are long enough to assume that no major change in the gas-to-particle ratio will occur within an urban to regional airshed, given that the typical lifetime of a fine particle is in the order of days to weeks, as well (8). Even so, variations in the gas-to-particle ratios, along with uncertainties in the initial estimate used, need to be considered.

A few studies have shown the increased resolution in source apportionment of two-phase receptor models (9-13), though this is not a common practice in the source apportionment literature. Applying a two-phase receptor model for PM₁₀ and non-methane-hydro-carbons (NMHC) has shown to significantly reduce the collinearity problem (9). A study dealing with decay-adjusted receptor modeling (13) has shown

small improvements in the agreement between CMB-predicted and observed concentrations of individual VOCs, but did not significantly change the estimated emissions contributions. These studies made use of two phase source profiles in which the profile included the fractional composition of both PM and gas phase data (speciated VOC, NO_x, SO₂ and CO) in a single profile, and χ^2 was minimized based on all these species. However, when Equation 1 is solved in this manner, several issues arise. First, since these gas phase species are reactive, the numerator in χ^2 cannot be simply expected to approach zero. In addition, the uncertainty in the measured ambient concentration is likely lower for major gas phase species, compared to speciated PM_{2.5} components. Hence these major gaseous species are likely to drive the minimization of χ^2 (assuming uncertainties in the source profile compositions are comparable), despite the fact that for many sources of PM_{2.5}, the fraction of PM_{2.5} emissions is much smaller than that of gas phase emissions. For example, data from the national emission inventory for the USA (14) indicate that only about 0.6% of the total mass emissions from coal power-plants are PM_{2.5}, the remaining and major part being gases (SO₂ and NO_x).

To avoid inaccuracies evolving from the use of two-phase source profiles and the straight forward minimization of χ^2 including gaseous species as fitting species, we suggest using ratios of SO₂/PM_{2.5}, CO/PM_{2.5} and NO_x/PM_{2.5} in emissions from the various sources to bound acceptable solutions to the source apportionment problem (Equation 1), without directly including these data in the process of minimizing χ^2 . That is, this information is used as a constraint, but not directly in the source profiles used by CMB. This information adds three constraint equations to the apportionment process,

based on the same principles as in equation 1. The ambient SO₂ levels can then be expressed as:

$$[SO_2] = \sum_{j=1}^n \left(\frac{SO_2}{PM_{2.5}} \right)_j S_j \quad (3-3)$$

where:

$[SO_2]$ = ambient SO₂ concentration (µg/m³);

$\left(\frac{SO_2}{PM_{2.5}} \right)_j$ = SO₂/PM_{2.5} ratio in emissions from source j (mass/mass);

S_j = contribution (source-strength) of source j (µg/m³) to the PM_{2.5} loading;

n = total number of sources;

Similar equations can be expressed also for ambient CO and NO_x. Due to uncertainties in the initial estimate of the gas-to-particle ratio at the source, and to account for possible changes to these ratios during transport, we suggest using these equations (Equation 3) to bound acceptable solutions to the PM_{2.5} source-apportionment problem (Equation 1), but not as part of the error minimization process. In practice, we suggest that such an acceptable solution is one that predicts the ambient SO₂, CO and NO_x concentrations within a factor of three (under or over prediction) of the observed value (sensitivity to this factor is addressed shortly). Hence, the goal is to find an optimum solution based on the particulate-phase data, which adheres to somewhat more flexible constraints on the gaseous side.

3.2.2 Use of global optimization models for source-apportionment

To solve the PM_{2.5} source-apportionment problem (Equation 1), subject to gas-phase constraints, we use a global optimization program. A large variety of quantitative decision problems in the applied sciences, engineering and economics can be described by constrained optimization models. In these models, the best decision is sought that satisfies all stated feasibility constraints and maximizes (or minimizes) the value of a given objective function. The general mathematical form of these models is summarized as (15,16):

$$\max f(x); a \leq x \leq b; g(x) \leq 0 \quad (3-4)$$

where:

x = a real n -vector (to describe feasible decisions)

a, b = finite, component-wise vector bounds imposed on x

$f(x)$ = a continuous function (to describe the model objective)

$g(x)$ = a continuous vector function (to describe the model constraints; the inequality is interpreted component-wise).

The objective of global optimization is to find the best solution of nonlinear decision models, in the possible presence of multiple locally optimal solutions. Here, LGO (Lipschitz(-Continuous) Global Optimizer) is used (15,16). LGO integrates a suite of robust and efficient global and local scope solvers. These include: global adaptive partition and search (branch-and-bound); adaptive global random search; local (convex)

unconstrained optimization; and local (convex) constrained optimization. The LGO implementation of these methods does not require derivative information. Their operations are based exclusively on the computation of the objective and constraint function values, at algorithmically selected search points.

Here, LGO was applied to identify and quantify the sources contributing to ambient levels of particulate matter. In practice, LGO was applied to solve the set of equations represented by Equation 1 (22 equations for 4 ions, 2 carbon fractions, and 16 trace metals), by setting χ^2 (Equation 2) as the objective function to be minimized. The solution was set subject to the constraint that the total apportioned levels of SO₂, CO and NO_x (as calculated by Equation 3) lie within a factor of three of the observed ambient levels.

3.3 Model implementation

3.3.1 Test Case: SEARCH 25 month dataset, Jefferson St., Atlanta, Georgia

To evaluate this modified approach for source-apportionment, we used the SEARCH (Southeastern Aerosol Research and Characterization) 25 month (8/98-8/00) dataset for Jefferson St. (JST) site in Atlanta, GA (17,18), which included data on total PM_{2.5} mass (gravimetric measure) and its components. The JST site is located 4 km northwest of downtown Atlanta, in an industrial and commercial area. The main objectives of SEARCH include the understanding of composition and sources of PM in the southeast (17,18). SEARCH data is being used for the Aerosol Research Inhalation Epidemiological Study (ARIES) air-quality health study in Atlanta, GA (17), and one motivation of this work is to assess the possibility of using source information derived

from receptor modeling in epidemiologic studies.. For the speciation of $PM_{2.5}$, a manual, filter-based, Particle Composition Monitor (PCM) was operated daily. The PCM included three channels to collect 24 hour integrated samples for analysis of major ions, trace metals, organic and elemental carbon in $PM_{2.5}$ size range (17). Ion Chromatography (IC) was used to quantify water soluble ionic species. Elemental and organic carbon collected on quartz filters were measured by Thermal Optical Reflectance (TOR). Trace metals were measured by x-ray fluorescence (XRF). Ambient values of daily SO_2 , CO and NO_y were reported as well. Mean values and standard deviations measured at the JST site, for the species and time period (8/98-8/00) used in this analysis, are given in Table 1. Note that NO_y was used rather than NO_x , to account for the amount of NO and NO_2 oxidized to other nitrogen forms, such as HNO_3 and Peroxy Acetyl Nitrate (PAN). The average NO_x/NO_y mass ratio was 0.89, indicating “fresh” local emissions (compared to 0.63 at the rural Yorkville site, 55 km west northwest of Atlanta). The concentration values were used for the measured data, and the summation of the analytical uncertainty and 1/3 of the detection limit value was used as the overall uncertainty assigned to each measured value (18). Values below the detection limit were replaced by half of the detection limit values, and their overall uncertainties were set at 5/6 of the detection limit values (18). Missing values were replaced by the geometric mean of the measured values, and their accompanying uncertainties were set at 4 times this geometric mean value (18).

Table 3-1. Mean, standard-deviation, minimum, maximum of ambient levels of the species used for the source apportionment, JST site, Atlanta, GA

Species	Mean ($\mu\text{g}/\text{m}^3$)	StDev ($\mu\text{g}/\text{m}^3$)	Min ($\mu\text{g}/\text{m}^3$)	Max ($\mu\text{g}/\text{m}^3$)
PM _{2.5}	19.1	8.9	1.9	54.6
SO ₄ ⁻²	5.41	3.65	0.53	20.8
NO ₃ ⁻	1.12	0.87	0.00	7.49
Cl ⁻	0.11	0.08	0.02	0.83
NH ₄ ⁺	2.79	1.60	0.30	10.3
EC	1.98	1.36	0.17	11.9
OC	4.46	2.21	0.66	18.4
Al	1.61E-02	4.52E-02	6.16E-03	9.00E-01
As	1.42E-03	1.35E-03	5.05E-04	1.51E-02
Ba	1.81E-02	8.01E-03	1.45E-02	5.69E-02
Br	4.04E-03	7.97E-03	2.60E-04	2.07E-01
Ca	5.37E-02	4.48E-02	4.04E-03	5.02E-01
Cu	3.70E-03	4.57E-03	6.15E-04	4.19E-02
Fe	8.92E-02	7.45E-02	5.34E-03	1.05E+00
K	6.51E-02	5.86E-02	6.37E-03	8.27E-01
Mn	1.91E-03	1.54E-03	4.00E-04	1.31E-02
Pb	6.40E-03	7.49E-03	1.17E-03	7.83E-02
Sb	3.34E-03	4.40E-03	2.13E-03	1.07E-01
Se	1.32E-03	1.26E-03	3.50E-04	1.01E-02
Si	1.12E-01	1.15E-01	1.05E-02	1.83E+00
Sn	4.32E-03	1.92E-03	3.53E-03	1.72E-02
Ti	4.78E-03	4.38E-03	2.14E-03	5.46E-02
Zn	1.63E-02	1.61E-02	4.23E-04	2.11E-01
SO ₂	16.6	12.3	1.4	98.1
CO	560	423	180	4020
NO _y	108	68.2	12.4	590

The major source categories used to in the source apportionment included light duty gasoline vehicles (LDGV), heavy duty diesel vehicles (HDDV), fugitive soil dust (SDUST), vegetative burning (BURN), coal fired power plants (CFPP) and cement kilns (CEM). To address the formation of secondary pollutants, we also included theoretical profiles based on the molecular weight fractions, for ammonium-sulfate (AMSULF), ammonium-bisulfate (AMBSLF), ammonium-nitrate (AMNIT) and secondary/other OC (OTHEROC). The secondary/other OC category will include any OC not apportioned to one of the primary source categories above. Of special note are emissions from meat charbroiling, dominated almost solely by OC emissions (19), with no unique inorganic marker, and characterized by low SO₂/PM_{2.5} and CO/PM_{2.5} ratios (19,20). This makes it

very difficult to distinguish between emissions from meat charbroiling and secondary OC formation. For this reason, meat charbroiling emissions were not apportioned directly, but were rather lumped into the secondary/other OC category.

Source profiles used for LDGV and HDDV were based on measurements as part of the Northern Front Range Air Quality Study (NFRAQS) (21). The profiles used for vegetative burning, power plants, and cement kilns were based on measurements done as part of the Big Bend Regional Aerosol Visibility and Observational (BRAVO) study (19). The soil dust profile used was from more regionally-representative measurements in Alabama (22). A summary of the source-profiles used in this study is given in Table 2. The LDGV profile is characterized by high carbon content and a high OC/EC ratio (2.3). The HDDV profile is also characterized by high carbon content, but there the OC/EC ratio is much lower (0.27). The LDGV had a higher abundance of trace metals, compared to the HDDV profile. However, the relative amounts of EC and OC in emissions from both gasoline and diesel vehicles is highly variable, and there is significant overlap in the range of values between the two mobile source types (23). Therefore, trying to distinguish gasoline and diesel contributions separately on the basis of just EC and OC mass fractions is suspect (23). This further indicates the need for additional markers to accurately separate the emissions from each of these sources. The BURN profile is characterized by high carbon content and a high OC/EC ratio (4.1), but also by a high K content (0.057), which can serve as a marker for vegetative burning. Crustal elements, Al, Ca, Fe and Si, along with OC, are abundant in the SDUST profile. The CFPP is characterized by high fractions of SO_4^{-2} , OC, Al, Ca and Si, and by a relatively high Se content, compared to other sources. Selenium can therefore serve as a marker for coal-

fired power plants. Emissions from cement kilns are characterized by high fractions of SO_4^{-2} , NO_3^- , OC, Al, Ca, Fe, K, and Si. This shows that differentiating emissions from power plants, cement kilns and fugitive soil dust might be subject to collinearity.

Adding information on gaseous emissions, in the form of gas-to-particle ratios, can further assist in identifying sources. Gas-to-particle ratios for mobile-sources (LDGV, HDDV), based on the 1999 National Emissions Inventory (14), show very different patterns, LDGV being characterized by a significantly higher $\text{CO}/\text{PM}_{2.5}$ ratio than HDDV (Table 3). Uncertainties in these ratios were not available, but are likely not large enough to mask the major differences between gaseous emissions from gasoline and diesel vehicles. Ratios for vegetative burning, coal-fired power plants, and cement-kilns were determined based on data from the BRAVO study (19) and the emission inventory for the State of Georgia (24). CO and NO_x ratios for cement kilns were modified to describe the kiln, rather than the entire plant emissions, as given by the inventory (which includes high particulate matter emissions from all grinding operations). The modification was based on the SO_2 ratio for kilns (19), compared to the SO_2 ratio obtained from the inventory. The high $\text{SO}_2/\text{PM}_{2.5}$ ratios in power plants and cement plants (Table 3) can assist in separating these emissions from fugitive soil dust (no gaseous emissions). The higher NO_x emissions from cement kilns, along with the differences in PM emissions, can assist in separating cement kilns and coal-fired power plant emissions. The relatively low $\text{CO}/\text{PM}_{2.5}$ ratios in vegetative burning emissions can serve as an additional marker to assist in separating this source from LDGV emissions (along with potassium).

Table 3-2. Particulate source-profiles used in the apportionment process (fraction of total PM_{2.5} emissions and standard deviations over multiple measurements)

Species	LDGV ¹	HDDV ¹	SDUST ²	BURN ³	CFPP ³
SO ₄ ²⁻	0.0133±0.0056	0.0046±0.0048	0.0010±0.0004	0.0239±0.0227	0.2874±0.2256
NO ₃ ⁻	0.0000±0.0052	0.0020±0.0014	0.0010±0.0004	0.0024±0.0018	0.0069±0.0109
Cl ⁻	0.0000±0.0100	0.0011±0.0003	0.0007±0.0005	0.0761±0.0730	0.0089±0.0157
NH ₄ ⁺	0.0000±0.0100	0.0000±0.0100	0.0000±0.0000	0.0165±0.0253	0.0179±0.0213
EC	0.2355±0.0277	0.7351±0.1014	0.0060±0.0040	0.1575±0.1545	0.0138±0.0222
OC	0.5486±0.0642	0.1981±0.0774	0.0440±0.0170	0.6441±0.1645	0.2718±0.2577
Al	0.0019±0.0024	0.0000±0.0100	0.0950±0.0010	0.0011±0.0010	0.0530±0.0326
As	0.0000±0.0006	0.0000±0.0001	0.0000±0.0000	0.0002±0.0007	0.0000±0.0006
Ba	0.0000±0.0100	0.0000±0.0100	0.0000±0.0000	0.0000±0.0003	0.0107±0.0101
Br	0.0000±0.0003	0.0000±0.0000	0.0000±0.0000	0.0008±0.0009	0.0003±0.0006
Ca	0.0118±0.0016	0.0006±0.0005	0.0180±0.0040	0.0040±0.0050	0.1655±0.1053
Cu	0.0004±0.0006	0.0000±0.0001	0.0003±0.0003	0.0000±0.0000	0.0009±0.0007
Fe	0.0120±0.0016	0.0002±0.0001	0.0530±0.0060	0.0007±0.0008	0.0361±0.0202
K	0.0001±0.0015	0.0001±0.0002	0.0092±0.0033	0.0573±0.0563	0.0052±0.0026
Mn	0.0001±0.0008	0.0000±0.0001	0.0016±0.0007	0.0000±0.0000	0.0012±0.0011
Pb	0.0006±0.0008	0.0000±0.0001	0.0001±0.0000	0.0000±0.0000	0.0006±0.0009
Sb	0.0000±0.0100	0.0000±0.0100	0.0000±0.0000	0.0000±0.0001	0.0001±0.0005
Se	0.0000±0.0003	0.0000±0.0001	0.0000±0.0000	0.0000±0.0000	0.0058±0.0083
Si	0.0121±0.0193	0.0000±0.0100	0.2660±0.0140	0.0030±0.0032	0.1069±0.0681
Sn	0.0000±0.0100	0.0000±0.0100	0.0000±0.0000	0.0000±0.0001	0.0001±0.0004
Ti	0.0001±0.0067	0.0000±0.0011	0.0100±0.0010	0.0001±0.0001	0.0085±0.0052
Zn	0.0091±0.0010	0.0006±0.0003	0.0001±0.0000	0.0003±0.0002	0.0031±0.0033
Species	CEM ³	AMSULF ⁴	AMBSLF ⁴	AMNITR ⁴	OTHEROC ⁴
SO ₄ ²⁻	0.3138±0.0837	0.727±0.036	0.835±0.042	0.000±0.000	0.00±0.00
NO ₃ ⁻	0.0891±0.0734	0.000±0.000	0.000±0.000	0.775±0.039	0.00±0.00
Cl ⁻	0.0712±0.1255	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
NH ₄ ⁺	0.0236±0.0187	0.273±0.014	0.156±0.008	0.225±0.011	0.00±0.00
EC	0.0296±0.0250	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
OC	0.1278±0.0603	0.000±0.000	0.000±0.000	0.000±0.000	1.00±0.00
Al	0.0106±0.0035	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
As	0.0000±0.0002	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Ba	0.0004±0.0012	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Br	0.0011±0.0013	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Ca	0.1748±0.0526	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Cu	0.0002±0.0001	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Fe	0.0134±0.0052	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
K	0.1159±0.0618	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Mn	0.0010±0.0004	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Pb	0.0006±0.0008	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Sb	0.0000±0.0003	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Se	0.0001±0.0000	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Si	0.0426±0.0219	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Sn	0.0001±0.0002	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Ti	0.0015±0.0007	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00
Zn	0.0041±0.0059	0.000±0.000	0.000±0.000	0.000±0.000	0.00±0.00

1- from the NFRAQS study (21)

2- from Cooper (22)

3- from Chow et al. (19)

4- based on molecular-weight fractions

Table 3-3. Gas-to-PM_{2.5} ratios used as constraints in the optimization process (mass/mass)

Source	SO ₂ /PM _{2.5}	CO/PM _{2.5}	NO _x /PM _{2.5}
LDGV	4.0 ¹	800 ¹	83.7 ¹
HDDV	0.71 ¹	13.4 ¹	21.9 ¹
BURN	0.013 ± 0.0004 ²	10.1 ± 1.1 ²	0.24 ± 0.06 ²
CFPP	128 ± 29.4 ³	2.1 ± 0.7 ³	41.0 ± 14.5 ³
CEM	316 ± 210 ^{3,4}	5.3 ± 6.5 ^{3,4}	270 ± 344 ^{3,4}

1- based on emission inventory data, no variability provided

2- based on emission inventory data; standard deviations based on county level, therefore low

3- based on emission inventory data; standard deviations based on plant-level, therefore higher

4- based on source-profile measurements (19)

3.4 Results

Source apportionment was performed on the SEARCH 25 month dataset using three different techniques. First, CMB8 (1,2) was used, applying effective variance weighting for least squares (EV) calculations (3) for PM_{2.5} components only (i.e., gaseous species were not used as fitting species). Then, the uncertainties in the source profiles were set equal to zero, and CMB8 was run again, using the ordinary weighted least square (OWLS) solution (3) (once more, without using gaseous species in the weighing procedure). Finally, the Lipschitz(-Continuous) Global Optimizer (LGO) (15,16) was applied to perform the OWLS solution, forcing constraints on the calculated levels of SO₂, CO and NO_x. An LGO derived OWLS solution without forcing gas-phase constraints was similar to the CMB OWLS solution. Applying an EV solution to LGO and forcing gas-phase constraints turned out to be too irregular, due to the rigidity of the EV weighing function. The measures used to evaluate each individual solution achieved were the chi-square (Equation 2), the correlation coefficient, the fraction of total PM_{2.5} mass apportioned and the calculated-to-observed ratios for the individual ratios. However, the chi-square values from EV are not comparable with the ones achieved by

OWLS, since the denominator in its formula is different. Therefore, as a convenient uniform measure of the quality of the fit, we also calculated daily values for the normalized-mean-square-error (NMSE), given as:

$$NMSE = \frac{\sum_{i=1}^m (C_i - \sum_{j=1}^n f_{i,j} S_j)^2}{\sum_{i=1}^m (C_i \sum_{j=1}^n f_{i,j} S_j)} \quad (3.4)$$

The NMSE has a range of $0 \leq NMSE \leq \infty$, 0 meaning perfect agreement in value between modeled and ambient values. A NMSE value of 0.5 represents a factor of two, on the average, between the two sets of data.

The average source-contributions, based on the entire 25 month dataset (average of 762 daily values) and using these three techniques, indicate that a major part of the ambient $PM_{2.5}$ is of secondary origin (Figure 3-1; Table 3-4. The apportionment of the primary pollutants differed among the three techniques used. The CMB8 EV solution apportioned $3.4 \mu\text{g}/\text{m}^3$ to mobile sources, with a diesel-to-gasoline ratio of 0.97. A slightly lower contribution was apportioned to mobile sources using the OWLS solution ($3.3 \mu\text{g}/\text{m}^3$), with a similar diesel-to-gasoline ratio. The LGO based mobile source contribution was slightly lower ($3.2 \mu\text{g}/\text{m}^3$), with a higher diesel-to-gasoline ratio (1.53). This lower gasoline-vehicle contribution is also evident in the lower calculated-to-observed ratio for CO based on the LGO solution, compared to the EV and OWLS solutions (all over predicted). Another notable difference between the three solutions was the amount of $PM_{2.5}$ attributed to power-plants. CMB8-EV estimated that contribution at

0.29 $\mu\text{g}/\text{m}^3$, CMB-OWLS at 0.62 $\mu\text{g}/\text{m}^3$, while LGO estimated the contribution at 0.15 $\mu\text{g}/\text{m}^3$. These differences are also evident in the calculated-to-observed ratios for SO_2 , which are significantly overpredicted in the EV and OWLS solutions. The amount attributed to vegetative burning was fairly similar in the EV and OWLS solutions (1.9 and 2.0 $\mu\text{g}/\text{m}^3$ respectively), significantly higher than in the LGO solution (1.1 $\mu\text{g}/\text{m}^3$). Potassium, a marker for vegetative burning, is overpredicted in the EV and OWLS solutions (calculated-to-observed ratios of 2.2), and better predicted in the LGO solution (ratio of 1.2). Differences were also noticed in the fugitive soil dust contributions. The amount attributed to the “Other OC” category was lower in the two CMB applications compared to the LGO solution, most likely due to over estimation of the OC contribution from gasoline vehicles and vegetative burning.

Interesting to note that all three solutions are characterized by high correlation coefficients for the fit obtained (0.97-0.99), good mass closure (91-93 %), and calculated-to-observed ratios nearing one for the major $\text{PM}_{2.5}$ components. In the EV solution, the average chi-square value, and most individual values, lied within the acceptable range (<4) (1,2). The chi-square values based on the OWLS and LGO solutions are not comparable to that of the EV solution. The chi-square based on the LGO solution is significantly higher than in the OWLS solution, and reflects the “penalty” of bounding acceptable solutions based on the gas phase species. However, the correlation coefficient is higher, and the overall and trace-metal based NMSE values are lower for the LGO solution compared to the OWLS solution.

To address the sensitivity of the solution obtained to the factor used as a constraint for the gas phase species, we also conducted the same analysis using a factor of

two (instead of three). Results obtained were nearly identical, with mass contributions differing by less than 7% for most sources. The major difference observed was for the average LDGV contribution, 0.2 $\mu\text{g}/\text{m}^3$ (18%). The source cross correlations between these two sets of solution were higher than 0.92 for all sources but cement kilns ($R=0.62$, but mass contribution being extremely low).

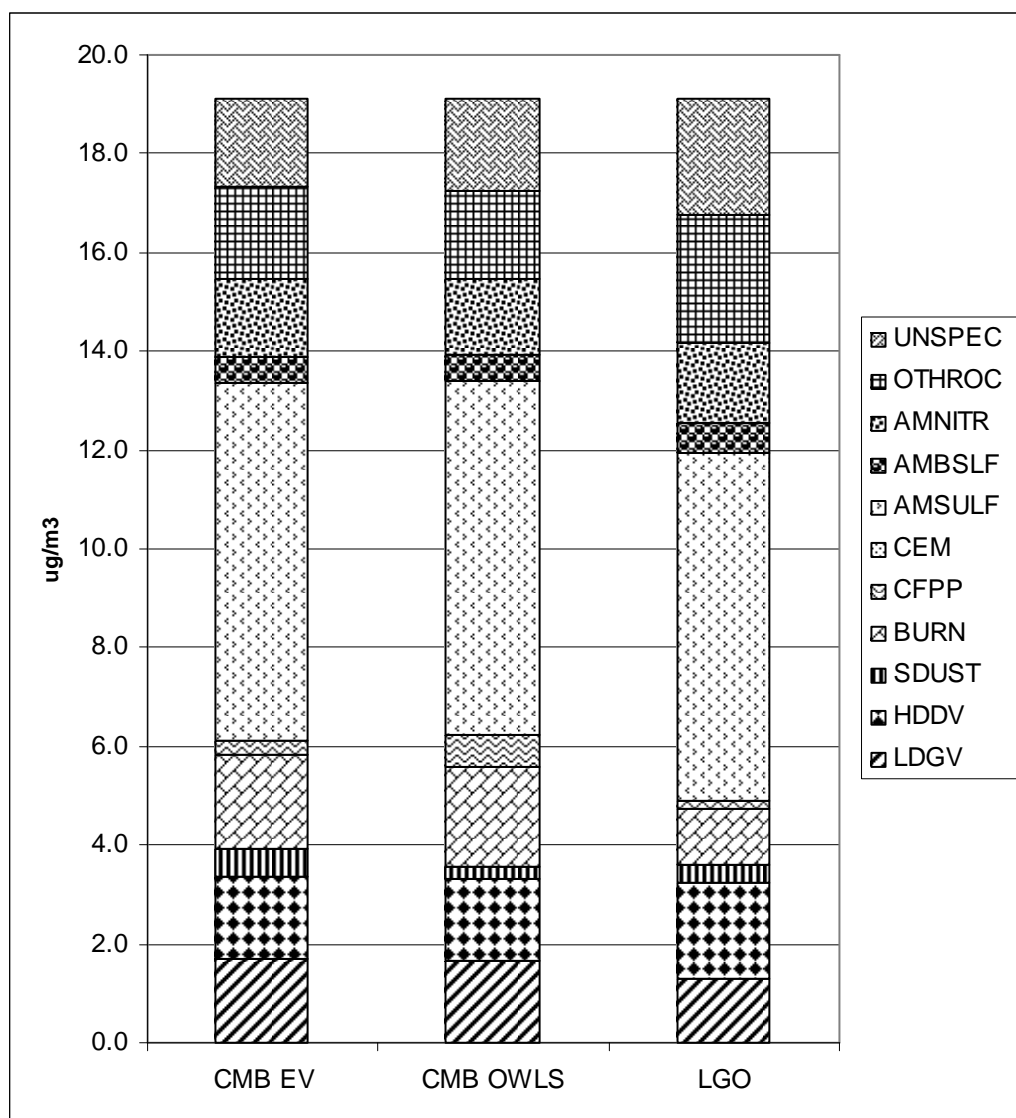


Figure 3-1. Source-contributions to PM_{2.5} levels at JST site, Atlanta, GA, using CMB8 EV solution, CMB8 OWLS solution and LGO

Table 3-4. Average and standard deviation of the source-contributions to PM_{2.5} levels measured at JST site, Atlanta, GA, using CMB8-EV, CMB8-OWLS and LGO. Also reported are the correlation (R), NMSE, % total mass, chi-square and calculated-to-observed ratios

	CMB8-EV	CMB8-OWLS	LGO
	Mean (StDev)	Mean (StDev)	Mean (StDev)
R ¹	0.9734 (0.0298)	0.9661 (0.0357)	0.9879 (0.0324)
NMSE PM _{2.5} ¹	0.161 (0.362)	0.0327 (0.131)	0.026 (0.096)
NMSE metals ²	0.801 (0.966)	0.714 (0.901)	0.249 (0.346)
% total mass ³	93.4 (18.2)	93.1 (18.7)	90.5 (17.4)
Chi-square ⁴	3.16 (3.47)	4.475 (6.447)	20.3 (16.8)
LDGV (µg/m ³)	1.72 (1.61)	1.68 (1.55)	1.28 (0.90)
HDDV (µg/m ³)	1.66 (1.53)	1.62 (1.52)	1.96 (1.63)
SDUST (µg/m ³)	0.55 (0.61)	0.28 (0.45)	0.39 (0.48)
BURN (µg/m ³)	1.90 (1.29)	2.01 (1.50)	1.13 (0.69)
CFPP (µg/m ³)	0.29 (0.48)	0.62 (0.74)	0.15 (0.12)
CEM (µg/m ³)	0.006 (0.04)	0.012 (0.08)	0.004 (0.02)
AMSULF (µg/m ³)	7.23 (5.20)	7.19 (5.17)	7.03 (5.12)
AMBSLF (µg/m ³)	0.54 (1.30)	0.50 (1.28)	0.64 (1.46)
AMNITR (µg/m ³)	1.57 (1.25)	1.55 (1.25)	1.60 (1.34)
OTHEROC (µg/m ³)	1.86 (1.55)	1.76 (1.50)	2.59 (1.64)
SO ₄ ⁻² ratio ⁵	1.16 (0.47)	1.10 (0.12)	1.07 (0.07)
NO ₃ ⁻ ratio ⁵	1.25 (1.02)	1.14 (0.75)	1.18 (0.87)
Cl ⁻ ratio ⁵	1.81 (1.49)	1.85 (1.36)	1.06 (0.63)
NH ₄ ⁺ ratio ⁵	0.93 (0.38)	0.89 (0.14)	0.88 (0.15)
EC ratio ⁵	1.09 (0.78)	0.97 (0.14)	0.98 (0.13)
OC ratio ⁵	1.06 (0.51)	1.01 (0.19)	1.00 (0.03)
Al ratio ^{5,6}	7.82 (6.37)	6.64 (5.11)	4.67 (2.81)
As ratio ⁵	0.53 (0.47)	0.55 (0.51)	0.32 (0.25)
Ba ratio ⁵	0.18 (0.28)	0.38 (0.45)	0.10 (0.08)
Br ratio ⁵	0.68 (0.92)	0.68 (0.84)	0.39 (0.38)
Ca ratio ⁵	1.87 (1.48)	2.92 (2.60)	1.15 (0.34)
Cu ratio ⁵	0.59 (0.57)	0.60 (0.52)	0.42 (0.39)
Fe ratio ⁵	0.84 (0.66)	0.73 (0.38)	0.55 (0.17)
K ratio ⁵	2.23 (1.72)	2.20 (1.55)	1.19 (0.49)
Mn ratio ⁵	1.06 (1.08)	0.94 (0.88)	0.69 (0.58)
Pb ratio ⁵	0.37 (0.38)	0.36 (0.36)	0.27 (0.23)
Sb ratio ⁵	0.02 (0.02)	0.03 (0.03)	0.01 (0.01)
Se ratio ⁵	1.99 (4.05)	4.66 (7.72)	1.11 (1.20)
Si ratio ⁵	2.18 (1.54)	1.64 (0.77)	1.28 (0.12)
Sn ratio ⁵	0.02 (0.02)	0.02 (0.02)	0.01 (0.00)
Ti ratio ⁵	2.12 (1.59)	2.01 (1.57)	1.27 (0.70)
Zn ratio ⁵	1.42 (1.95)	1.22 (0.73)	1.01 (0.35)
SO ₂ ratio ^{5,6}	4.37 (8.73)	8.24 (23.4)	1.99 (0.97)
CO ratio ^{5,6}	3.18 (3.43)	3.07 (3.56)	2.06 (0.83)
NO _y ratio ^{5,6}	2.11 (1.83)	2.07 (1.51)	1.58 (0.66)

1- calculated based on all PM_{2.5} components

2- calculated based on trace metals only

3- % of apportioned mass to total PM_{2.5}

4- chi-square is not comparable for the EV case and the two OWLS cases, as the denominator in its formula is different

5- ratio of apportioned mass to ambient level (ideally would approach 1 for all species)

6- not used as a fitting species

To analyze the driving forces in the apportionment process, we calculated the correlations between the daily contributions of the various sources and the daily ambient levels of the different species. These correlations (R values, Table 5) indicate which are the species most highly correlated with each source category, therefore driving the apportionment. This is done on the entire dataset, as opposed to the transpose of the normalized modified pseudo-inverse matrix (MPIN) (2), which indicates the degree of influence each species concentration has on the contribution, on a case by case basis. Note that the correlations used here are not normalized, hence the species with the highest correlations are considered the ones most influential, even if the actual correlation is somewhat low. The following is stated based on these correlations:

LDGV: The LDGV contribution based on the CMB-EV solution is correlated mainly with Zn, OC and EC. A low correlation with CO and NO_y is observed in the EV solution. The OWLS solution showed a fairly similar pattern, with slightly higher correlations with CO and NO_y, and a fairly high correlation with Pb. However, the LGO solution was highly correlated with CO and NO_y, along with much of the same PM species as the EV and OWLS solutions.

HDDV: The HDDV contribution generated by all three source-apportionment techniques used here was most correlated with EC, which is the major component of diesel emissions. Stronger correlations with EC were observed in the OWLS and LGO solutions. Correlation with NO_y was the highest in the LGO solution. Such a correlation is expected as NO_x emissions from diesel vehicles, on a per-mile basis, are higher than from gasoline vehicles (25,26).

BURN: The vegetative burning contribution from the EV and OWLS solution was correlated with chlorine, potassium, EC, OC and bromine. The LGO solution was correlated with the same species but bromine. The correlation with potassium was much higher in the LGO solution (0.62), compared to the EV and OWLS solutions (0.37 and 0.43 respectively).

SDUST: Soil-dust is characterized by a high abundance of crustal elements, such as Al, Ca, Fe, Si and Ti. Results from the all three solutions are correlated with these elements. However, the EV solution is most correlated with Fe, and to a degree with Si, Ti and Mn, while the OWLS and LGO solutions are correlated mainly with Si, Al, Ti and Fe. As expected, low correlations with the gaseous species were found. However, the EV solution seems to have picked a contribution associated with motor vehicles (possibly resuspended paved road dust), as it is somewhat correlated with EC and CO, as opposed to the OWLS and LGO solutions.

CFPP: The EV-generated power-plant contribution is mostly correlated with Cu, SO_4^{-2} , Ca, and to a degree with Se and Fe. The OWLS solution is mostly correlated with Cu, Fe and Mn. Both these solutions show no correlation with SO_2 , and the OWLS solution shows a correlation with CO and NO_y , indicative of mobile sources. The LGO solution, however, is mostly correlated with Ca, Se and SO_2 . Se is a unique marker for coal-fired power plant emissions (2). The LGO solution, being correlated with both Se and SO_2 , is likely truly indicative of power plants.

CEM: It is difficult to evaluate the driving species for the cement kiln contribution, as it is very low. Non-zero contributions were generated in only 33, 36 and 65 cases (out of 762 cases), using EV, OWLS and LGO respectively. Based on this limited data, the EV

solution was mainly correlated with Ca; the OWLS solution with Br; and the LGO solution with Ca and the NO_x.

OTHEROC: The other OC category includes any OC not apportioned to one of the previous categories. If most of the primary OC was accounted for, this category would include mainly secondary OC. Since EC and OC often share the same sources (25), a high correlation of the OTHEROC category with either EC or OC would indicate a primary OC contribution. A good reference point is the correlation between OC and EC in the ambient data, which is 0.82. The correlations between EC and the contributions to the OTHEROC category are lower: 0.44, 0.60 and 0.65 for the EV, OWLS and LGO solutions, respectively. This indicates a secondary component in the OTHEROC category. These values, along with the correlations with OC and the magnitude of the contribution, suggest that LGO solution includes more primary OC than the EV and OWLS solutions. This is likely due to an over estimation of the mobile-source contribution by both EV and OWLS, leaving less OC to be apportioned to the OTHEROC category. This does mean, however, that there is an unexplained source of OC in the LGO solution. One likely source would be meat charbroiling, which as previously mentioned, emits almost solely OC, and is characterized by low gas-to-particle ratios. For this reason, it is difficult to distinguish meat charbroiling from secondary OC formation using either CMB (without organic markers) or this application of LGO. An organic marker, such as cholesterol, is needed to identify and quantify meat-charbroiling emissions.

AMSULF, AMBSLF, AMNITR: These secondary “sources” were all correlated with their major components, and low correlations with gaseous pollutants were observed, for all three cases.

To further illustrate the differences between the EV, OWLS and LGO solutions, we also calculated source inter-correlations using these three solutions (Table 6). The diagonal terms in these matrices indicate that the contributions of ammonium-sulfate, ammonium-bisulfate, and ammonium-nitrate are fairly similar in all three cases. However, major differences are observed for the primary source categories. The EV and OWLS gasoline vehicle contribution is significantly different than the LGO LDGV contribution, as shown by the low correlations. The differences in the HDDV contributions are more subtle. Another major difference is observed in the CFPP contribution: The OWLS CFPP contribution is correlated more with the LDGV contribution from LGO than the corresponding CFPP contributions, likely due to collinearity.

Table 3-5. Correlations (R) between source-contributions and ambient-levels of fitting species

	LDGV			HDDV			BURN		
	CMB EV	CMBOWLS	LGO	CMB EV	CMBOWLS	LGO	CMB EV	CMBOWLS	LGO
SO ₄ ⁻²	0.03	0.01	0.04	0.22	0.22	0.21	0.10	0.07	0.14
NO ₃ ⁻	0.22	0.31	0.30	0.14	0.18	0.19	0.15	0.23	0.33
Cl ⁻	0.16	0.17	0.09	0.05	0.04	0.11	0.35	0.45	0.52
NH ₄ ⁺	0.08	0.05	0.04	0.27	0.25	0.26	0.22	0.23	0.18
EC	0.42	0.53	0.39	0.69	0.94	0.96	0.34	0.43	0.41
OC	0.45	0.55	0.45	0.53	0.68	0.74	0.35	0.49	0.49
Al	0.00	0.04	0.07	0.05	0.09	0.08	0.03	0.04	0.13
As	0.26	0.32	0.27	0.21	0.23	0.29	0.30	0.42	0.29
Ba	-0.01	0.01	0.03	0.09	0.13	0.10	0.04	0.05	0.10
Br	0.13	0.15	0.13	0.35	0.34	0.17	0.38	0.41	0.14
Ca	0.24	0.30	0.39	0.25	0.34	0.32	0.17	0.19	0.19
Cu	0.38	0.50	0.36	0.25	0.30	0.38	0.22	0.38	0.29
Fe	0.36	0.46	0.40	0.45	0.56	0.60	0.24	0.34	0.33
K	0.23	0.27	0.22	0.24	0.30	0.35	0.37	0.43	0.62
Mn	0.41	0.51	0.36	0.41	0.53	0.58	0.24	0.34	0.32
Pb	0.42	0.65	0.34	0.19	0.22	0.33	0.18	0.32	0.33
Sb	0.06	0.06	0.06	0.06	0.11	0.07	0.25	0.26	0.00
Se	0.13	0.17	0.13	0.34	0.38	0.41	0.23	0.27	0.17
Si	0.12	0.17	0.20	0.22	0.27	0.27	0.12	0.16	0.22
Sn	0.16	0.21	0.16	0.13	0.14	0.14	0.08	0.13	0.11
Ti	0.14	0.18	0.18	0.27	0.35	0.37	0.19	0.25	0.29
Zn	0.58	0.86	0.46	0.31	0.38	0.42	0.20	0.37	0.33
SO ₂	0.18	0.23	0.25	0.21	0.31	0.33	0.17	0.20	0.20
CO	0.31	0.36	0.74	0.23	0.37	0.32	0.15	0.25	0.24
NO _y	0.31	0.40	0.66	0.28	0.42	0.45	0.19	0.29	0.25
	SDUST			CFPP			CEM		
	CMB EV	CMBOWLS	LGO	CMB EV	CMBOWLS	LGO	CMB EV	CMBOWLS	LGO
SO ₄ ⁻²	0.20	0.17	0.24	0.43	0.28	0.28	-0.11	0.15	0.02
NO ₃ ⁻	0.01	-0.09	-0.09	0.07	0.17	0.14	0.29	0.37	0.04
Cl ⁻	0.02	-0.01	-0.01	0.01	0.10	0.08	-0.08	0.15	-0.18
NH ₄ ⁺	0.29	0.21	0.28	0.43	0.33	0.24	-0.12	0.10	-0.06
EC	0.40	0.21	0.25	0.34	0.54	0.42	-0.15	0.06	0.13
OC	0.33	0.13	0.19	0.27	0.48	0.38	-0.13	0.17	0.17
Al	0.29	0.91	0.87	0.03	0.13	0.09	0.09	-0.09	-0.13
As	0.16	0.05	0.05	0.07	0.23	0.18	0.24	0.46	0.23
Ba	0.11	0.21	0.22	0.23	0.25	0.12	-0.04	-0.08	0.12
Br	0.24	0.04	0.01	0.10	0.16	0.12	0.01	0.58	0.11
Ca	0.33	0.62	0.57	0.43	0.46	0.56	0.44	0.37	0.56
Cu	0.37	0.31	0.24	0.46	0.74	0.21	0.24	0.14	-0.03
Fe	0.57	0.76	0.77	0.36	0.62	0.40	0.18	0.28	0.05
K	0.30	0.42	0.29	0.17	0.34	0.17	-0.15	0.05	-0.05
Mn	0.46	0.53	0.53	0.33	0.61	0.33	0.18	0.23	-0.05
Pb	0.24	0.07	0.06	0.13	0.49	0.11	-0.07	0.18	0.16
Sb	0.05	0.00	-0.01	0.08	0.13	0.08	-0.28	0.04	0.11
Se	0.15	0.12	0.15	0.36	0.29	0.50	0.12	-0.01	0.24
Si	0.48	0.95	0.99	0.22	0.34	0.28	0.22	0.07	0.02
Sn	0.08	0.06	0.12	0.02	0.21	0.05	-0.07	0.12	0.15
Ti	0.44	0.83	0.84	0.27	0.43	0.27	0.05	0.06	-0.13
Zn	0.29	0.08	0.12	0.14	0.50	0.27	0.29	0.24	0.25
SO ₂	0.09	-0.03	-0.03	0.11	0.21	0.45	0.18	-0.12	0.25
CO	0.23	0.11	0.14	0.17	0.43	0.19	0.11	0.06	0.18
NO _y	0.20	0.04	0.07	0.16	0.41	0.31	0.14	0.11	0.40

Table 3-5 (cont.). Correlations (R) between source-contributions and ambient-levels of fitting species

	AMSULF			AMBSULF			AMNITR		
	CMB EV	CMBOWLS	LGO	CMB EV	CMBOWLS	LGO	CMB EV	CMBOWLS	LGO
SO ₄ ⁻²	0.85	0.94	0.93	0.46	0.52	0.54	0.04	0.05	0.06
NO ₃ ⁻	-0.06	-0.01	-0.01	-0.01	0.04	0.05	0.68	0.87	0.82
Cl ⁻	0.04	0.07	0.07	-0.05	0.03	0.05	0.22	0.30	0.26
NH ₄ ⁺	0.85	0.95	0.94	0.29	0.36	0.40	0.08	0.11	0.11
EC	0.16	0.22	0.22	-0.07	0.03	0.00	0.17	0.22	0.21
OC	0.22	0.28	0.28	0.01	0.09	0.06	0.20	0.23	0.23
Al	0.00	0.03	0.03	0.09	0.15	0.09	-0.04	-0.03	-0.04
As	-0.02	0.00	0.00	-0.06	-0.01	-0.02	0.10	0.14	0.14
Ba	0.14	0.18	0.18	0.09	0.06	0.01	0.01	0.02	0.02
Br	-0.02	0.01	0.01	0.00	0.07	0.09	0.08	0.09	0.09
Ca	0.24	0.30	0.30	0.07	0.16	0.17	0.04	0.06	0.05
Cu	0.07	0.10	0.10	-0.02	0.06	-0.02	0.08	0.12	0.11
Fe	0.20	0.27	0.27	0.04	0.14	0.08	0.07	0.12	0.10
K	0.12	0.15	0.15	0.04	0.09	0.04	0.09	0.10	0.11
Mn	0.17	0.23	0.23	0.01	0.07	0.02	0.13	0.18	0.16
Pb	-0.05	-0.03	-0.02	-0.05	0.05	-0.04	0.14	0.22	0.20
Sb	-0.06	-0.05	-0.04	-0.07	0.02	-0.01	-0.01	0.01	0.01
Se	0.32	0.37	0.37	0.14	0.18	0.23	0.12	0.12	0.12
Si	0.23	0.28	0.28	0.22	0.25	0.23	-0.04	-0.04	-0.05
Sn	0.01	0.03	0.03	0.07	0.09	0.03	0.10	0.08	0.07
Ti	0.25	0.31	0.30	0.11	0.28	0.22	-0.03	-0.01	-0.01
Zn	0.01	0.04	0.05	-0.07	-0.03	-0.06	0.16	0.19	0.17
SO ₂	-0.12	-0.08	-0.08	-0.08	0.00	-0.02	0.16	0.18	0.18
CO	-0.03	0.00	0.00	-0.05	-0.02	-0.01	0.13	0.15	0.16
NO _y	-0.10	-0.07	-0.07	-0.09	-0.05	-0.06	0.18	0.22	0.21

	OTHER OC		
	CMB EV	CMBOWLS	LGO
SO ₄ ⁻²	0.22	0.23	0.29
NO ₃ ⁻	0.08	0.08	0.14
Cl ⁻	0.05	-0.01	0.08
NH ₄ ⁺	0.25	0.23	0.29
EC	0.44	0.60	0.65
OC	0.63	0.81	0.92
Al	-0.03	0.00	-0.04
As	0.12	0.15	0.27
Ba	0.06	0.12	0.10
Br	0.29	0.30	0.14
Ca	0.15	0.20	0.21
Cu	0.30	0.28	0.32
Fe	0.32	0.36	0.37
K	0.33	0.45	0.36
Mn	0.24	0.23	0.34
Pb	0.25	0.25	0.35
Sb	0.03	0.07	0.07
Se	0.09	0.12	0.24
Si	0.15	0.21	0.16
Sn	0.09	0.11	0.12
Ti	0.17	0.23	0.21
Zn	0.27	0.24	0.39
SO ₂	0.08	0.12	0.24
CO	0.24	0.27	0.24
NO _y	0.19	0.26	0.29

Table 3-6. Source inter-correlations (R) using CMB8-EV, CMB8-OWLS and LGO

		EV									
		LDGV	HDDV	SDUST	BURN	CFPP	CEM	AMSULF	AMBSLF	AMNITR	OTHROC
OWLS	LDGV	0.64	0.30	0.31	0.22	0.14	0.23	-0.03	-0.05	0.19	0.25
	HDDV	0.26	0.69	0.33	0.21	0.34	-0.20	0.16	-0.10	0.12	0.39
	SDUST	0.00	0.13	0.42	0.08	0.15	-0.10	0.16	0.25	-0.08	0.07
	BURN	0.37	0.26	0.28	0.69	0.17	-0.08	0.07	-0.06	0.15	0.21
	CFPP	0.41	0.40	0.41	0.28	0.65	0.15	0.25	-0.09	0.13	0.24
	CEM	0.29	0.24	0.32	0.21	0.00	0.81	0.11	0.25	0.31	0.61
	AMSULF	0.04	0.25	0.26	0.15	0.46	-0.14	0.90	0.19	-0.01	0.22
	AMBSLF	-0.03	0.00	0.16	-0.05	0.02	0.13	0.22	0.91	0.01	0.10
	AMNITR	0.18	0.12	0.01	0.11	0.06	0.18	-0.04	0.01	0.77	0.10
	OTHROC	0.16	0.44	0.16	0.15	0.17	-0.29	0.17	0.05	0.08	0.68
		LGO									
		LDGV	HDDV	SDUST	BURN	CFPP	CEM	AMSULF	AMBSLF	AMNITR	OTHROC
EV	LDGV	0.44	0.36	0.07	0.30	0.16	0.15	0.04	-0.02	0.16	0.32
	HDDV	0.17	0.74	0.19	0.28	0.31	0.19	0.25	0.00	0.11	0.45
	SDUST	0.23	0.38	0.47	0.23	0.16	-0.03	0.26	0.08	0.00	0.23
	BURN	0.16	0.32	0.10	0.46	0.16	0.07	0.14	0.04	0.11	0.24
	CFPP	0.14	0.33	0.20	0.13	0.37	0.31	0.46	0.02	0.06	0.21
	CEM	0.18	-0.23	0.20	-0.06	0.25	1.00	-0.14	0.22	0.16	-0.23
	AMSULF	-0.01	0.16	0.23	0.07	0.21	-0.05	0.89	0.27	-0.03	0.22
	AMBSLF	-0.03	-0.09	0.22	0.00	0.02	0.24	0.18	0.86	0.00	0.04
	AMNITR	0.20	0.12	-0.07	0.24	0.06	-0.05	-0.01	0.05	0.76	0.13
	OTHROC	0.18	0.43	0.13	0.26	0.17	0.35	0.21	0.09	0.09	0.66
		LGO									
		LDGV	HDDV	SDUST	BURN	CFPP	CEM	AMSULF	AMBSLF	AMNITR	OTHROC
OWLS	LDGV	0.54	0.45	0.11	0.33	0.25	0.20	0.01	-0.03	0.24	0.40
	HDDV	0.24	0.93	0.23	0.29	0.40	0.13	0.23	0.01	0.17	0.58
	SDUST	0.08	0.18	0.96	0.16	0.16	-0.02	0.20	0.19	-0.08	0.06
	BURN	0.26	0.40	0.13	0.59	0.16	0.12	0.12	0.02	0.16	0.34
	CFPP	0.43	0.51	0.31	0.31	0.37	0.00	0.32	-0.08	0.17	0.33
	CEM	0.06	0.11	0.09	0.29	0.13	0.30	0.11	0.45	0.35	0.08
	AMSULF	0.02	0.23	0.28	0.13	0.26	-0.02	0.99	0.27	0.00	0.28
	AMBSLF	-0.02	0.02	0.24	0.09	0.13	0.51	0.20	0.95	0.05	0.14
	AMNITR	0.25	0.19	-0.07	0.27	0.10	0.03	-0.02	0.09	0.97	0.16
	OTHROC	0.18	0.58	0.18	0.24	0.21	0.31	0.20	0.08	0.12	0.88

Results from the LGO solution (based on inorganic markers and inorganic gases) were also compared with results from an organic-markers source apportionment study (4) and the five county Atlanta metropolitan area emissions inventory (24) (Table 7). In the Zheng et al. study (4), average monthly contributions to PM_{2.5} were calculated for the months of April, July, August, October of 1999, and January 2002, for the JST site, and are averaged here. Source categories included were: diesel exhaust; gasoline exhaust; vegetative detritus; meat cooking; road dust; wood combustion; and natural gas combustion. Both the LGO solution and the Zheng et al. (4) results indicate the dominance of contributions from mobile sources to primary PM_{2.5} levels, but the magnitude is somewhat different (66% and 58%, respectively). The split between

gasoline and diesel vehicles also was different: a diesel-to-gasoline ratio of 1.5 using LGO and 6.7 using organic tracers. For comparison, the diesel-to-gasoline ratio in the emissions inventory for the five-county Atlanta metropolitan area is 3.0 (22). Other differences were the somewhat lower vegetative burning contribution using LGO compared to organic markers (23% and 30%), the higher LGO soil dust contribution (7.9% compared to 2.5%), and the meat-charbroiling contribution identified by the organic markers study (6.4%). The LGO solution generated a higher “secondary/other OC” contribution compared to the organic marker study (not presented in Table 7), which may include meat cooking emissions (characterized almost solely by OC emissions). In contrast to the receptor model results, the emissions inventory is dominated by area sources other than soil dust and wood combustion (44% of total $PM_{2.5}$ emissions). Wood combustion and road dust are the next two major sources in the inventory (15% each), followed by diesel and gasoline engines (11.5% and 3.9% respectively). The road dust emissions seem to be over estimated, as shown by the receptor model results and measured levels of crustal species. Incorporating preliminary data on emissions from meat charbroiling (27) into the inventory suggests these emissions contribute 5% of total $PM_{2.5}$ emissions (not formally reported in the inventory, 24). Given the dominance of “other” area sources (waste disposal treatment, recovery and incineration; industrial oil and gas production; agriculture production; other sources) in the emissions inventory, it is difficult to compare the inventory to the source apportionment results directly. It seems that the “other” area source category is over estimated, as four different sets of source apportionment results presented here (CMB EV, CMB OWLS, LGO and the Zheng et al. study from 2002) indicate to the dominance of contributions from diesel and gasoline

engines to primary PM_{2.5} levels (58-66% of PM_{2.5} emissions, compared to 15% in the inventory). The contributions of coal-fired power plants and cement kilns, as indicated by the inventory, are relatively small, similar to findings from the LGO solution.

Table 3-7. Comparison between percent contributions to primary PM_{2.5} levels based on LGO, organic-marker CMB (4), and the five-county Atlanta metropolitan area emissions inventory (24). The Zheng et al. results (4) were averaged to represent a yearly pattern by weighting July and August results as “summer”, October as “fall”, January as “winter” and April as “spring”. Cooking emissions reported here are based on a preliminary estimate (27), and are not reported in the inventory (24). Inventoried emissions are given with and without “other” sources for more direct comparison.

Source category (% contribution)	LGO	CMB using organic tracers (4)	Five county Atlanta metro emissions inventory (with / without “other” sources)
Gasoline engines	26.0	7.5	3.9 / 7.6
Diesel engines	39.9	50.3	11.5 / 22
Fugitive soil dust	7.9	2.5	15.1 / 29 ⁵
Vegetative burning/ Wood combustion	23.0	29.7	15.1 / 29
Coal-fired power plants	3.1	-	0.8 / 1.5
Cement kilns	0.1	-	0.3 / 0.4
Meat charbroiling	-	6.4	5.0 / 9.7 ⁴
Vegetative detritus	-	2.7	-
Natural gas combustion	-	1.0	-
Other area sources ¹	-	-	43.9
Other point sources ²	-	-	2.7
Other non-road sources ³	-	-	1.8

1- other than soil dust and wood combustion

2- other than coal-fired power plants and cement kilns

3- other than gasoline and diesel engines

4- based on a preliminary estimate by Baek et al. (26)

5- A large fraction of the fugitive dust emissions are expected to be removed locally

Acknowledgements

This work was supported by grants to Emory University from the U.S. Environmental Protection Agency (R82921301-0) and the National Institute of Environmental Health Sciences (R01ES11199 and R01ES11294) and to Georgia Tech (EPA grants R831076 and R830960) We would also like to thank Georgia Power and Southern Company for continuing support and the individuals at ARA (Atmospheric Research and Analysis) for both providing access to data used in this analysis and ongoing discussions.

References

1. US-EPA. *CMB8 User's Manual*. EPA-454/R-01-XXX, Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711, 2001.
2. US-EPA. *CMB8 Application and Validation Protocol for PM_{2.5} and VOC*. EPA-454/R-98-XXX, Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711, 1998.
3. Watson, J.G.; Cooper, J.A.; Huntzicker, J.J. The effective variance weighting for least squares calculations applied to the mass balance receptor model. *Atmos. Environ.*, **1984**, 18, 1347-1355.
4. Zheng, M.; Cass, G.R.; Schauer, J.J.; Edgerton, E.S. Source apportionment of PM_{2.5} in the southeastern United States using solvent-extractable organic compounds as tracers. *Environ. Sci. Technol.*, **2002**, 36, 2361-2371.
5. Schauer, J.J.; Kleeman, M.J.; Cass, G.R.; Simoneit, A.T. Measurement of emissions from air pollution sources. 3. C₁-C₂₉ organic compounds from fireplace combustion of wood. *Environ. Sci. Technol.*, **2001**, 35, 1716-1728.
6. Schauer, J.J.; Kleeman, M.J.; Cass, G.R.; and Simoneit, A.T. Measurement of emissions from air pollution sources. 2. C₁ through C₃₀ organic compounds from medium duty diesel trucks *Environ. Sci. Technol.*, **1999**, 33, 1578-1587.
7. Seinfeld, J.H.; Pandis, S.N. *Atmospheric Chemistry and Physics*; John Wiley & Sons, Inc. 1998.
8. Wilson, W.E.; Suh, H.H. Fine particles and coarse particles: Concentration relationships relevant to epidemiologic studies. *J. Air Waste Manage. Assoc.*, **1997**, 47:1238-1249.
9. Lin, J.; Scheff, P.A.; Wadden, R.A. *Development of a Two-Phase Receptor Model for NMHC and PM₁₀ Air Pollution Sources in Chicago*. Presented at the 86th annual meeting & exhibition of the Air & Waste Management Association, Denver, Colorado, 1993.
10. Wadden, R.A.; Scheff, P.A.; Lin, J.; Lee, H.; Keil, C.; Graf-Teterycz, J.; Keehan, K.; Kenski, D.; Milz, S.; Holsen, T.M.; Khalili, N. *Two Phase Receptor*

Modeling. Presented at the 84th annual meeting & Exhibition of the Air & Waste Management Association, Vancouver, British Columbia, 1991.

11. McKee, G.A.; Wadden, R.A.; and Scheff, P.A. *Development of a Two-Phase Chemical Mass Balance Receptor Model*. Presented at the 83rd annual meeting & Exhibition of the Air & Waste Management Association, Pittsburgh, Pennsylvania, 1990.
12. Scheff, P.A., Wadden, R.A, and Allen, R.J. Development and Validation of a Chemical Element Mass Balance for Chicago. *Environ. Sci. Technol.*, **1984**, 18, 923-931.
13. Lin, C.; Milford, J.B. Decay-adjusted chemical mass balance receptor modeling for volatile organic compounds. *Atmos. Environ.*, **1994**, 28, 3261-3276.
14. US-EPA; 1999 National Emission Inventory (NEI): Air Pollutant Emission Trend (www.epa.gov/ttn/chief/trends/, accessed March 12, 2004).
15. Pinter, J.D. *Global Optimization in Action*, Kluwer Academic Publishers, The Netherlands, 1996.
16. Pintér, J.D. LGO - A Program System for Continuous and Lipschitz Global Optimization. In *Developments in Global Optimization*; Bomze, I.M., Csendes, T., Horst, R., Pardalos, P.M., Eds. Kluwer Academic Publishers, Dordrecht, Boston, London, 1997.
17. Hansen, D.A.; Edgerton, E.S.; Hartsell, B.E.; Jansen, J.J.; Kandasamy, N.; Hidy, G.M.; Blanchard, C.L. The Southeastern Aerosol Research and Characterization Study: part 1—Overview. *Journal of the Air and Waste Management Association*, **2003**, 53, 1460-1471.
18. Kim, E.; Hopke, P.K.; Edgerton, E.S. Source identification of Atlanta aerosol by positive matrix factorization. *J. Air Waste Manage. Assoc.*, **2003**, 53, 731-739.
19. Chow, J.C.; Watson, J.G.; Kuhns, H.; Etyemezian, V.; Lowenthal, D.H.; Crow, D.; Kohl, S.D.; Engelbrecht, J.P.; and Green, M.C. source profiles for industrial, mobile, and area sources in the Big Bend Regional Aerosol Visibility and Observational study. *Chemosphere*, **2004**, 54, 185-208.

20. McDonald, J.D.; Zielinska, B.; Fujita, E.M.; Sagebiel, J.C.; Chow, J.C.; and Watson, J.G. Emissions from charbroiling and grilling of chicken and beef. *J. Air Waste Manage. Assoc.*, **2003**, 53, 185-194.
21. Zielinska, B.; McDonald, J.D.; Hayes, T.; Chow, J.C.; Fujita, E.M.; Watson, J.G. *Northern Front Range Air Quality Study Final Report. Volume B: Source Measurements*. (<http://www.nfraqs.colostate.edu/nfraqs/index2.html>, accessed April 25, 2004)
22. Cooper, J.A. *Determination of Source Contributions to Fine and Coarse Suspended Particulate Levels in Petersville, Alabama*. Report to Tennessee Valley Authority by NEA, Inc., 1981.
23. Gillies, J.A. and Gertler, A.W. Comparison and evaluation of chemically speciated mobile source PM_{2.5} particulate matter profiles *J. Air Waste Manage. Assoc.*, **2000**, 50:1459-1480.
24. Unal, A.; Tian, D.; Hu, Y.; and Russell, A. *2000 Emissions Inventory for Fall Line Air Quality Study (FAQS)*. Prepared for Georgia Department of Natural Resources, Environmental Protection Division, April 2003.
25. Health Effects Institute (HEI). *Emissions from Diesel and Gasoline Engines Measured in Highway Tunnels*. Research Report # 107, 2002.
26. Marmur, A. and Mamane, Y. Comparison and evaluation of several mobile-source and line-source models in Israel *Trans. Res. D.*, **2003**, 8, 249–265.
27. Baek, J. and Russell, A.G. 2004. Private communication.

CHAPTER 4

OPTIMIZED VARIABLE SOURCE-PROFILE APPROACH FOR SOURCE APPORTIONMENT

(A. Marmur, J.A. Mulholland, and A.G. Russell, *Atmospheric Environment*, in press)

Abstract

An expanded Chemical Mass Balance (CMB) approach for PM_{2.5} source apportionment is presented in which both the local source compositions and corresponding contributions are determined from ambient measurements and initial estimates of source compositions using a global-optimization mechanism. Such an approach can serve as an alternative to using predetermined (measured) source profiles, as traditionally used in CMB applications, which are not always representative of the region and/or time period of interest. Constraints based on ranges of typical source profiles are used to ensure that the compositions identified are representative of sources and are less ambiguous than the factors/sources identified by typical Factor Analysis (FA) techniques. Gas-phase data (SO₂, CO and NO_y) are also used, as these data can assist in identifying sources. Impacts of identified sources are then quantified by minimizing the weighted-error between apportioned and measured levels of the fitting species. This technique was applied to a dataset of PM_{2.5} measurements at the former Atlanta Supersite (Jefferson Street site), to apportion PM_{2.5} mass into nine source categories. Good agreement is found when these source impacts are compared with those

derived based on measured source profiles as well as those derived using a current FA technique, Positive Matrix Factorization. The proposed method can be used to assess the representativeness of measured source-profiles and to help identify those profiles that may be in significant error, as well as to quantify uncertainties in source-impact estimates, due in part to uncertainties in source compositions.

Keywords: CMB-LGO, optimization, source-apportionment, PM_{2.5}, PMF, health-study.

4.1 Background

Chemical Mass Balance (CMB) receptor models are a common tool for apportioning ambient levels of pollutants (mainly particulate matter) among the major contributing sources. CMB combines the chemical and physical characteristics of particles or gases measured at sources and receptors to quantify the source contributions to the receptor. Quantification is based on the solution to a set of linear equations that express each receptor's ambient chemical concentration as a linear sum of products of source-profile fractions and source contributions (US-EPA, 2004a; US-EPA, 2004b), as expressed by:

$$C_i = \sum_{j=1}^n f_{ij} S_j + e_i \quad (4-1)$$

where:

C_i = ambient concentration of chemical species i ($\mu\text{g}/\text{m}^3$);

$f_{i,j}$ = fraction of species i in emissions from source j ;

S_j = contribution (source-strength) of source j ($\mu\text{g}/\text{m}^3$);

n = total number of sources;

e_i = error term.

Source profile fractions (f_{ij}) and the receptor concentrations (C_i), along with uncertainty estimates, serve as input data to the CMB model. Results consist of the contribution of each source category (S_j) to the measured concentration of different species at the receptor. A frequent source of uncertainty in the implementation of CMB is the choice of source profiles used as input. There is a wide variety of source profiles in the literature, but these are not always representative of the region and/or time of interest. Some examples of this are as follows: soil (dust) composition often varies geographically; emission composition from biomass burning is dependent on the type of vegetation or wood burned (e.g. agriculture burning, soft or hard wood residential combustion); emissions from coal-fired power plants may vary depending on the types of coal used; mobile source emissions can vary from region to region and temporally due to different fuels, fleet composition, or driving conditions. To date, the most common approach to addressing this variability has been to select profiles that are most representative of the region and time period of interest from those that are available. In many cases, however, specific profiles are not available. Moreover, a profile derived from any one source at one time may not be representative due to variability in time and space. Due to these reasons, Factor Analysis (FA) techniques have been developed (Hopke, 1988; Paatero and Tapper, 1994) and are often applied to characterize and quantify the sources contributing to ambient particulate matter levels (Kim et al., 2003; Kim et al., 2004a; Kim et al., 2004b; Maykut et al., 2003). FA models do not require the

use of predetermined source profiles, but results are often difficult to interpret as factors do not necessarily represent specific sources (Seinfeld and Pandis, 1998). The underlying assumption in all FA models is that the chemical composition of ambient particulate samples includes information about the fingerprints of the sources affecting the receptor, and that this information can be used to derive the source compositions. The procedure for characterizing these sources (or factors) is based on correlations between ambient levels of the different species, a high correlation indicating that the species share a common source (Seinfeld and Pandis, 1998). One of the more commonly used FA methods in recent years is Positive Matrix Factorization (PMF) (Paatero and Tapper, 1994). In PMF, factors are constrained to have non-negative $f_{i,j}$'s, and no sample can have a negative source contribution. Application of PMF requires that error estimates for the data be chosen judiciously so that the estimates reflect the quality and reliability of each of the data points. A critical step in PMF analysis is the determination of the number of factors (Paatero, 2004).

4.2 Methods

4.2.1 CMB model expansion to include variable source compositions

This study combines concepts from FA and CMB applications to calculate source contributions to ambient $\text{PM}_{2.5}$ without relying solely on emissions composition studies or on interpretation of factors obtained by FA as sources. The technique is based on solving the same set of equations used in CMB modeling (Equation 4-1), but instead of using predetermined source profiles, ranges for different fractions in source-indicative profiles are used as input. The model then optimizes the fractions of different species

within each profile by minimizing residual mass, subject to several constraints. Lower and upper bounds for the fractions of species in the various source profiles are set based on knowledge of typical compositions of various sources. Instead of deriving the contributing factors by FA, and then identifying (interpreting) them as sources based on knowledge of typical composition, this information is used beforehand to constrain the model while searching for the best combination of sources to describe the ambient levels of PM_{2.5}. The choice of source categories to include likewise is made beforehand, in contrast to FA.

As a basis for setting the constraints for the fractions of various species in the source profiles, suggested values are taken from in the validation protocol for CMB8.2 (US-EPA, 2004b) (Table 4-1).

Table 4-1. Chemicals from particles in different emissions sources (US-EPA, 1998)

Source Type	Dominant Particle Size	Chemical fractions			
		<0.001	0.001 - 0.01	0.01 - 0.1	>0.1
Motor vehicles	Fine	Cr, Ni, Y	NH ₄ ⁺ , Si, Cl, Al, Si, P, Ca, Mn, Fe, Zn, Br, Pb	Cl ⁻ , NO ₃ ⁻ , SO ₄ ⁻² , NH ₄ ⁺ , S	OC, EC
Vegetative burning	Fine	Ca, Mn, Fe, Zn, Br, Rb, Pb	NO ₃ ⁻ , SO ₄ ⁻² , NH ₄ ⁺ , Na ⁺ , S	Cl ⁻ , K ⁺ , Cl, K	OC, EC
Coal-fired boiler	Fine	Cl, Cr, Mn, Ga, As, Se, Br, Rb, Zr	NH ₄ ⁺ , P, K, Ti, V, Ni, Zn, Sr, Ba, Pb	SO ₄ ⁻² , OC, EC, Al, S, Ca, Fe	Si
Road dust	Course	NO ₃ ⁻ , NH ₄ ⁺ , P, Zn, Sr, Ba	SO ₄ ⁻² , Na ⁺ , K ⁺ , P, S, Cl, Mn, Ba, Ti	EC, OC, Al, K, Ca, Fe	OC, Si

These bounds on the abundance of species were slightly modified (Table 4-2) and several additional constraints were added to better characterize the different sources.

Table 4-2: Lower and upper bounds for chemical fractions of total PM_{2.5} mass emitted in source profiles

Species	Gasoline vehicles (GV)		Diesel vehicles (DV)		Road dust (DUST)		Vegetative burning (BURN)		Coal power plants (CFPP)	
	lower	upper	lower	upper	lower	upper	lower	upper	lower	upper
EC	0.05	1	0.4	1	0	0.01	0.01	0.3	0.01	0.1
OC	0.3	1	0.1	1	0.01	0.1	0.3	1	0.01	0.3
SO ₄ ⁻²	0.01	0.1	0.01	0.1	10 ⁻³	0.01	10 ⁻³	0.01	0.01	0.3
NO ₃ ⁻	0.01	0.1	0.01	0.1	0	10 ⁻³	10 ⁻³	0.01	0	0
Cl ⁻	0.01	0.1	0.01	0.1	10 ⁻³	0.01	0.01	0.1	0	0.01
NH ₄ ⁺	10 ⁻³	0.01	10 ⁻³	0.01	0	10 ⁻³	10 ⁻³	0.01	10 ⁻³	0.01
Al	10 ⁻³	0.05	10 ⁻⁴	10 ⁻³	0.01	0.2	0	0	0.01	0.1
As	0	0	0	0	0	0	0	0	0	10 ⁻³
Ba	0	0	0	0	10 ⁻³	0.01	0	0	10 ⁻³	0.01
Br	10 ⁻⁴	0.01	10 ⁻⁵	10 ⁻³	0	0	0	10 ⁻³	0	10 ⁻³
Ca	10 ⁻³	0.05	10 ⁻⁴	10 ⁻³	0.01	0.2	0	10 ⁻³	0.01	0.2
Fe	10 ⁻³	0.05	10 ⁻⁴	10 ⁻³	0.01	0.2	0	10 ⁻³	0.01	0.1
K	0	10 ⁻³	0	10 ⁻⁴	0.01	0.1	0.01	0.1	10 ⁻³	0.01
Mn	10 ⁻⁴	0.01	10 ⁻⁴	10 ⁻³	10 ⁻³	0.05	0	10 ⁻³	0	0.005
Pb	0	10 ⁻³	0	10 ⁻³	0	0	0	10 ⁻³	0	0.01
Se	0	0	0	0	0	0	0	0	10 ⁻⁴	0.01
Si	10 ⁻³	0.05	10 ⁻⁴	0.01	0.1	1	0	0	0.05	0.2
Ti	0	0	0	0	10 ⁻³	0.05	0	0	10 ⁻³	0.01
Zn	10 ⁻³	0.02	10 ⁻⁴	10 ⁻³	0	10 ⁻³	0	10 ⁻³	10 ⁻³	0.01
Additional constraints	Sum≤1; OC/EC≥1; TC≥0.5; OM/OC≥1.4		Sum≤1; OC/EC≤1; TC≥0.5; OM/OC≥1.4		Sum≤1; Sum metal oxides≤1		Sum≤1; OC/EC≥3; TC≥0.5; OM/OC≥1.4		Sum≤1; Sum metal oxides≤1;	

Emissions from light-duty gasoline vehicles (LDGVs) usually contain more OC than EC (Gillies and Gertler, 2000), so a constraint of OC/EC≥1 was used for LDGVs, and an opposite constraint was used for heavy duty diesel vehicles (HDDVs). However, the relative amount of EC and OC components in PM emissions from both gasoline and diesel vehicles is highly variable (Gillies and Gertler, 2000), and there is significant overlap in the range of values between the two mobile source types. Therefore, trying to distinguish gasoline and diesel contributions separately on the basis of just EC and OC mass fractions is suspect (Gillies and Gertler, 2000). For this reason, we also incorporated information on typical CO/PM_{2.5}, NO_x/PM_{2.5} and SO₂/PM_{2.5} ratios in the emissions from these, as well as other sources (Marmur et al., 2005). Higher bounds for trace metals are set for gasoline vehicles, compared to diesel vehicles (HEI, 2002; Manchester-Neesvig et

al., 2003). For vegetative burning, a constraint of $OC/EC \geq 3$ was used, as this source is characterized by high OC to EC ratios (US-EPA, 2004b). A relatively large fraction (0.01-0.1) of potassium in biomass burning emissions (US-EPA, 2004b) is also used. For all sources, the sum of fractions over all species was constrained to be less than or equal to unity. In the case of soil dust and power-plants, oxidized forms of the metals are assumed (such as Al_2O_3 , SiO_2 etc.). Organic material (OM) fractions in the primary emissions were bounded by a minimum contribution of 1.4 times the fraction of OC in the profile. These constraints are summarized by the following equations:

$$f_{i,j \text{ lower}} \leq f_{i,j} \leq f_{i,j \text{ upper}} \quad (4-2)$$

$$f_{OC,j}/f_{EC,j} \geq R_{OC/EC} \quad (4-3)$$

$$f_{OC,j} + f_{EC,j} \geq R_{TC} \quad (4-4)$$

$$1.4f_{OC,j} + f_{i(\text{excluding } OC),j} \leq 1.0 \quad (4-5)$$

(for all sources but soil-dust and power-plants)

$$1.89f_{Al,j} + 1.40f_{Ca,j} + 1.43f_{Fe,j} + 1.20f_{K,j} + 2.14f_{Si,j} + 1.67f_{Ti,j} + 1.4f_{OC,j} + f_{other,j} \leq 1.0 \quad (4-6)$$

(for soil-dust and power-plants)

where:

$f_{i,j \text{ lower}}, f_{i,j \text{ upper}}$ = lower and upper bound on fraction of species i in source j (Table 4-2);

$R_{OC/EC}$ = bound on OC/EC ratio (≥ 1 for gasoline vehicles; ≥ 3 for vegetative burning; ≤ 1 for diesel vehicles; Table 4-2);

R_{TC} = bound on TC (EC+OC) fraction (≥ 0.5 for gasoline and diesel vehicles, vegetative burning; Table 4-2);

1.89, 1.40, 1.43, 1.20, 2.14, 1.67 = ratios of molecular weights of metal-oxide/metal for Al_2O_3 , CaO, Fe_2O_3 , K_2O , SiO_2 and TiO_2 , respectively.

To address the formation of secondary pollutants, four pure component profiles were used for ammonium-sulfate (AMSULF; 73% SO_4^{-2} , 27% NH_4^+), ammonium-bisulfate (AMBSLF; 84% SO_4^{-2} , 16% NH_4^+), ammonium-nitrate (AMNITR; 78% NO_3^- , 22% NH_4^+) and other/secondary OC (OTHROC; 100% OC), based on the molecular weights of the components (Marmur et al., 2005).

For each sample, Equation 1 was solved by minimizing χ^2 :

$$\chi^2 = \sum_{i=1}^m \frac{\left(C_i - \sum_{j=1}^n f_{ij} S_j \right)^2}{\sigma_{C_i}^2} \quad (4-7)$$

where:

σ_{C_i} is the uncertainty of the C_i measurement. This is solved subject to the constraints on the $\text{PM}_{2.5}$ source compositions (Equations 4-2 - 4-6 and Table 4-2), as well as the requirement to reasonably reconstruct ambient gas-phase (SO_2 , CO, and NO_y) concentrations:

$$\frac{1}{b} [\text{GS}] \leq \sum_{j=1}^n \left(\frac{\text{GS}}{\text{PM}_{2.5}} \right)_j S_j \leq b [\text{GS}] \quad (4-8)$$

where:

$[\text{GS}]$ = ambient concentration of gaseous-species (CO, SO_2 , NO_y ; $\mu\text{g}/\text{m}^3$);

$\left(\frac{\text{GS}}{\text{PM}_{2.5}} \right)_j$ = GS/ $\text{PM}_{2.5}$ mass ratio in emissions from source j (see Marmur et al., 2005 for

values used);

S_j = contribution (source-strength) of source j ($\mu\text{g}/\text{m}^3$) to the $\text{PM}_{2.5}$ loading;

n = total number of sources;

b = bound for gas-species mass reconstruction (typically $b=3$, to account for uncertainties in initial GS/PM_{2.5} ratios and changes in these ratios during transport from source to receptor; (Marmur et al., 2005))

This latter requirement has been shown to reduce collinearity between source-compositions and to achieve more plausible source-apportionment results (Marmur et al., 2005).

A global optimization program, Lipschitz global optimizer (LGO) (Pinter, 1996; Pintér, 1997), was utilized to find the optimal solution (by minimizing χ^2), subject to the above mentioned constraints. In LGO, the best solution is sought that satisfies all stated feasibility constraints and maximizes (or minimizes) the value of a given objective function (Pinter, 1996; Pintér, 1997). The objective of global optimization is to find the best solution of nonlinear decision models, in the possible presence of multiple locally optimal solutions. LGO integrates a suite of robust and efficient global and local scope solvers. These include: global adaptive partition and search (branch-and-bound); adaptive global random search; local (convex) unconstrained optimization; and local (convex) constrained optimization. The LGO implementation of these methods does not require derivative information. Their operations are based exclusively on the computation of the objective and constraint function values, at algorithmically selected search points.

4.2.2 SEARCH 25 month dataset, Jefferson St., Atlanta, Georgia

Evaluation of this expanded CMB approach involved using the SEARCH (Southeastern Aerosol Research and Characterization) 25 month (8/98-8/00) dataset for the Jefferson St. (JST) monitoring site in Atlanta, GA (Hansen et al., 2006; Hansen et al., 2003; Kim et al., 2003; Marmur et al., 2005), which includes data on total PM_{2.5} mass (gravimetric measure) and its components (major ions by Ion Chromatography; trace metals by x-ray fluorescence; organic and elemental carbon by Thermal Optical Reflectance), as well as ambient concentrations of SO₂, CO and NO_y. Summation of the analytical uncertainty and 1/3 of the detection limit value was used as the overall uncertainty assigned to each measured value. Values below the detection limit were replaced by half of the detection limit values, and their overall uncertainties were set at 5/6 of the detection limit values. Missing values were replaced by the geometric mean of the measured values, and their accompanying uncertainties were set at 4 times this geometric mean value (Marmur et al., 2005). Five variable source-profiles (GV, DV, DUST, BURN, CFPP; Table 4-2) and four constant ones (AMSULF, AMBSLF, AMNITR, OTHROC) were included in the analysis.

4.3 Results

4.3.1 Derived source-profiles

First, source profile compositions were determined (for five sources: GV, DV, DUST, BURN, CFPP; for the species in Table 4-2) using, initially, a subset of the data. In this way, a separate set of data could be used in the evaluation. Out of the total of 762 days, we identified 447 days in which all of the data (all ions, carbon fractions, metals, CO, SO₂, NO_y) were available. From those 447 days, we chose all the January, April,

July and October samples (133 days) to represent the four seasons. LGO then was applied to find the optimal solution (Equation 4-1) based on the ordinary weighted least-squares (OWLS) approach (Equation 4-7) (Christensen and Gunst, 2004; Friedlan, 1973) while adjusting the source profile fractions within the allowable bounds (Table 4-2, Equations 4-2 – 4-6) and subject to the gas-phase constraints (Equation 4-8). This becomes an optimization problem with more than one hundred decision variables (fractions and contributions), requiring several minutes of computational workload (on a Pentium 4.0 PC) and several tens of thousands of iterations per sample to reach a global minimum point. The computational workload for a solution using predetermined source-profiles is much smaller, reaching convergence within several seconds and several hundreds of iterations per sample. Source profiles obtained are analyzed for how often bounds (lower or upper) on individual species are met and for compositional variability between samples. The process was evaluated by repeating the analysis using all the February, May, August (excluding 1998), and November samples (149 days). No significant seasonal variability in source composition was observed. Therefore, the comparison will focus on average source compositions for each of the two test cases.

Average source-profile compositions for the two scenarios (two subsets of the data: Case 1 based on 133 samples; Case 2 based on 149 samples) show little difference (Figures 4-1 – 4-5). When compared to several source profiles from the literature (Chow et al., 2004; Cooper, 1981; Zielinska et al., 1998) differences arise, but the species driving source apportionment modeling (Marmur et al., 2006) are the same. Major differences are observed for primary sulfate, nitrate and ammonium content in various sources because LGO assigns most of that mass into the secondary sulfate and nitrate

categories. These estimates of primary sulfate and nitrate content are therefore highly uncertain; however, these species do not serve as markers for any of the sources of primary emissions. Hence, the effect of this uncertainty on the overall source-apportionment process is likely a bias in predicted impacts rather than uncertainty in the prediction of day-to-day variation in source impacts.

LGO generated gasoline-vehicle profiles (GV1 and GV2 in Figure 4-1) are comparable to the profile from NFRAQS (Zielinska et al., 1998), being characterized by an OC/EC ratio of 1.6 on average (a constraint of ≥ 1.0 was used), compared to 2.3 in the NFRAQS profile. The total carbon content (TC) is 0.67 (a constraint of ≥ 0.5 was used), compared to 0.78, in the NFRAQS profile. Content of Zn, a good marker for gasoline vehicles (Marmur et al., 2006; Marmur et al., 2005), is similar (average of 0.008 vs. 0.009). The content of other trace metals (Al, Ca, Fe, Si) is similar.

Diesel vehicle profiles generated by LGO (DV1 and DV2) are characterized by an OC/EC ratio of 0.36 (a constraint of ≤ 1.0 was used) compared to 0.27 in the NFRAQS profile (Zielinska et al., 1998). The TC content is 0.77 on average (a constraint of ≥ 0.5 was used), compared to 0.93. Metal content is similar among all diesel vehicle profiles.

Vegetative burning profiles from LGO were fairly similar to the BURN profile from BRAVO (Chow et al., 2004), with an OC/EC ratio of 4.7 on average, compared to 4.1 from BRAVO. Potassium content in the LGO profiles is 0.063 on average, compared to 0.056 in the BRAVO profile. The LGO profiles' chlorine content is roughly half of that in the measured profile (0.037 compared to 0.076).

The LGO-derived soil dust profiles are similar to the Alabama soil-dust profile from Cooper et al. (Cooper, 1981) with respect to Si, Ti, Mn and Fe content, but Al, Ca

and K content differed significantly. The high Al content in the Alabama profile (Cooper, 1981) seems to be an overestimate for Atlanta aerosol (Marmur et al., 2005).

Some differences arise when comparing the LGO derived CFPP profile to the one from BRAVO (Chow et al., 2004), which are based on measurements in Texas. SO_4^{-2} , EC, OC, Al, Ca, Fe and Si are the most abundant species in both sets of profiles, but differences in their content is evident, especially in SO_4^{-2} and OC content, though neither is an important tracer for primary CFPP $\text{PM}_{2.5}$. As previously mentioned, LGO assigns most of the SO_4^{-2} to the secondary “ammonium-sulfate” category, likely underestimating sulfate content in primary emissions. OC is apportioned to carbon-rich source-categories such as GV, DV and BURN based on constraints on both OC content and OC/EC ratios, and to the “other OC” category (secondary and un-apportioned organic carbon). For relatively carbon-lean source-categories, such as DUST and CFPP, for which knowledge on typical OC/EC ratios is limited, LGO tends to suggest a lower fraction of OC. The content of selenium, a unique tracer for CFPP, is very similar, 0.0061 in the LGO profiles, compared to 0.0058 in the BRAVO profile.

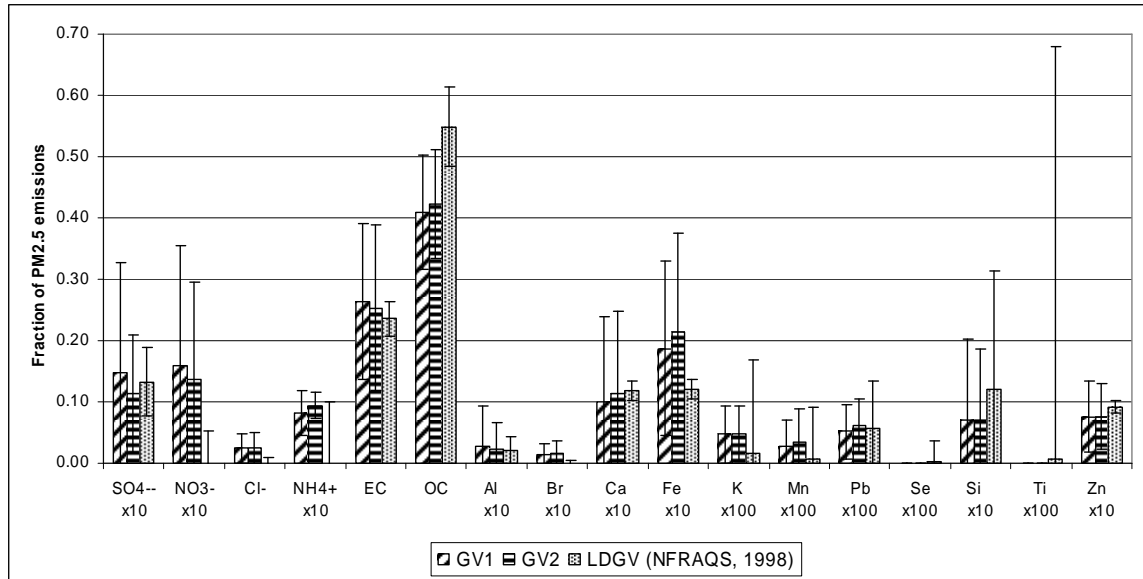


Figure 4-1. Source profiles generated by LGO for gasoline-fueled vehicles (GV1 based on 133 cases, GV2 based on 149 cases), compared to a profile from the NFRAQS study (Zielinska et al., 1998), previously used to apportion $PM_{2.5}$ in Atlanta (Marmur et al., 2005). Bars represent \pm one standard-deviation of the LGO estimated (over 133 and 149 cases) or measured fractions.

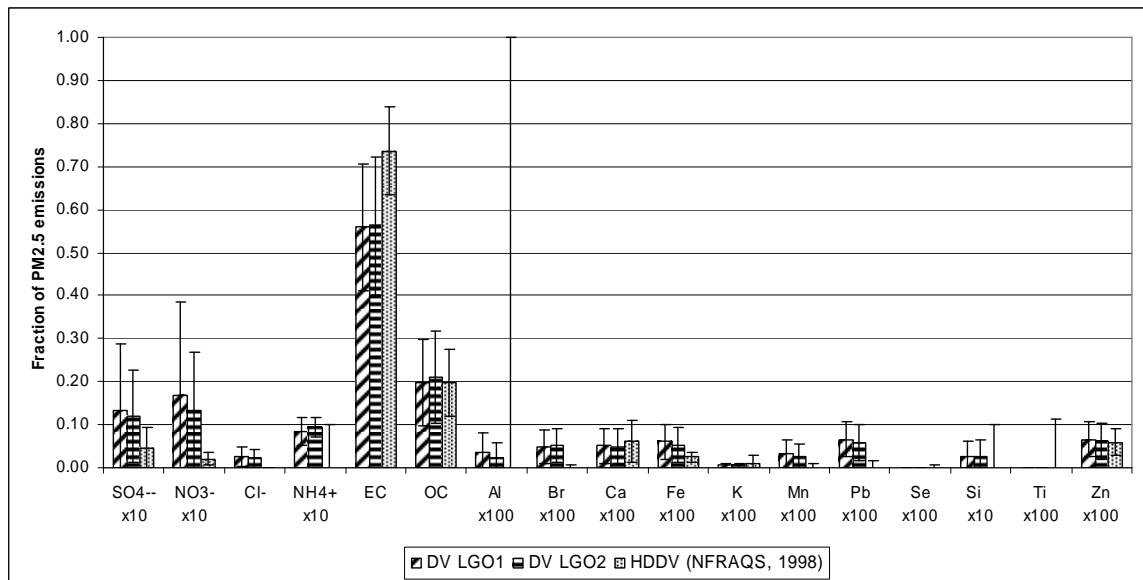


Figure 4-2. Source profiles generated by LGO for diesel-fueled vehicles (DV1 based on 133 cases, DV2 based on 149 cases), compared to a profile from the NFRAQS study (Zielinska et al., 1998), previously used to apportion $PM_{2.5}$ in Atlanta (Marmur et al., 2005). Bars represent \pm one standard-deviation of the LGO estimated (over 133 and 149 cases) or measured fractions.

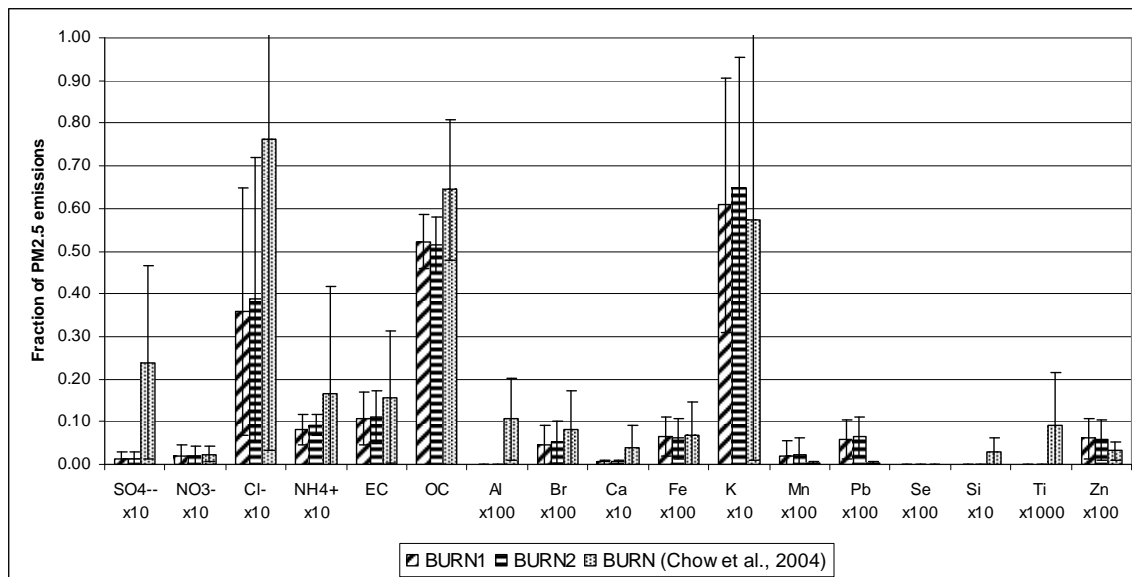


Figure 4-3. Source profiles generated by LGO for vegetative burning (BURN1 based on 133 cases, BURN2 based on 149 cases), compared to a vegetative burning profile from the BRAVO study (Chow et al., 2004), previously used to apportion PM_{2.5} in Atlanta (Marmur et al., 2005). Bars represent \pm one standard-deviation of the LGO estimated (over 133 and 149 cases) or measured fractions.

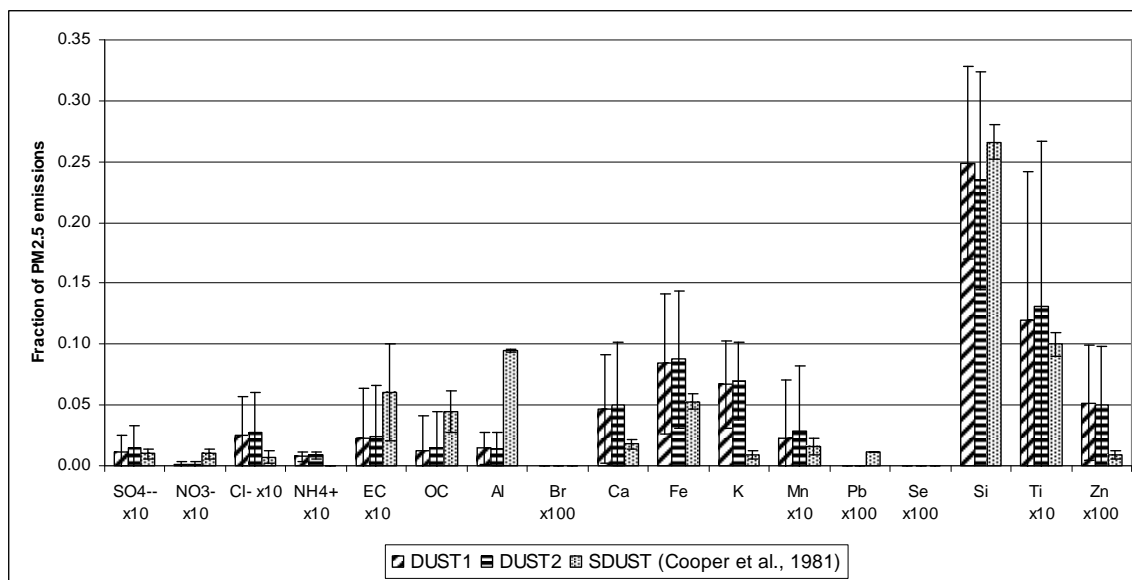


Figure 4-4. Source profiles generated by LGO for soil-dust (DUST1 based on 133 cases, DUST2 based on 149 cases), compared to an Alabama soil dust profile (Cooper, 1981), previously used to apportion PM_{2.5} in Atlanta (Marmur et al., 2005). Bars represent \pm one standard-deviation of the LGO estimated (over 133 and 149 cases) or measured fractions.

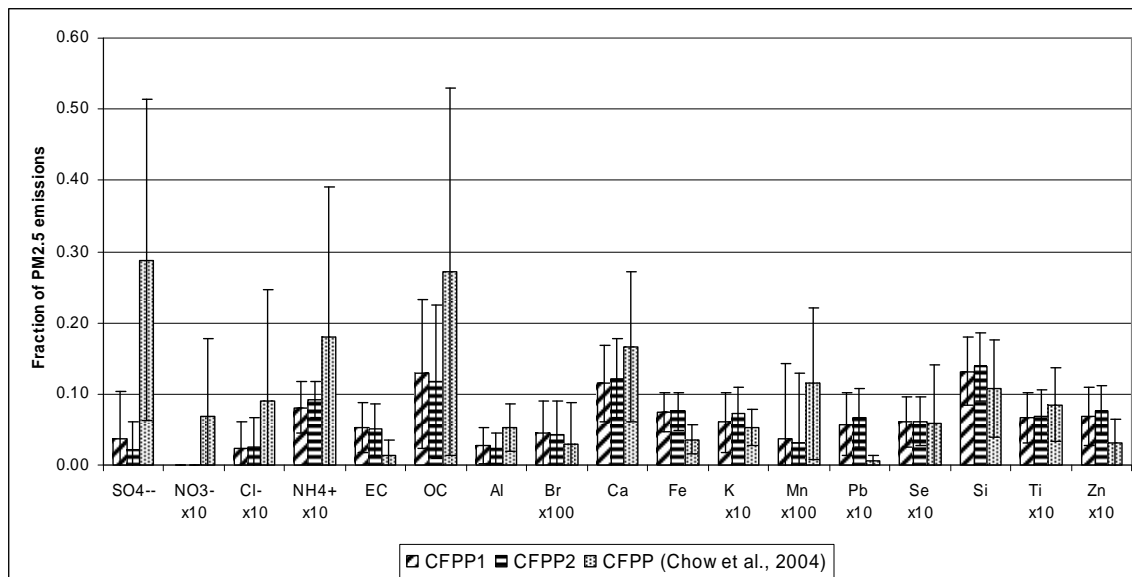


Figure 4-5. Source profiles generated by LGO for coal-fired power plants (CFPP1 based on 133 cases, CFPP2 based on 149 cases), compared to a CFPP profile measured in Texas (Chow et al., 2004), previously used to apportion PM_{2.5} in Atlanta (Marmur et al., 2005). Bars represent \pm one standard-deviation of the LGO estimated (over 133 and 149 cases) or measured fractions.

The role of the constraints/bounds used to derive the source compositions were analyzed using the entire dataset of derived source-profiles (447 cases), in terms of percent of cases in which either bound (lower/upper) were found limiting (Table 4-3). These data indicate that the constraints are most often limiting for species that are not unique tracers or key driving species of a given category. Examples are sulfate, nitrate and ammonium content in all sources of primary PM_{2.5}, OC and non-crustal elements in soil dust, and various metals such as Al, K, Mn and Si in both types of mobile sources. However, a key success of the source-profile derivation process is LGO's ability to estimate the fraction of unique/key species well within the allowable range in most cases. Examples are EC, OC, TC, OC/EC ratio and Zn for gasoline vehicles; EC, OC, TC and OC/EC ratio for diesel vehicles; EC, OC, TC, OC/EC ratio and K for wood burning; Ca, Fe, Si, and Ti for dust; and Ca and Se for coal-fired power plants.

Table 4-3: Percent of cases (447 total) in which the derived species fraction was at the lower limit / within the allowable range / at the upper limit.

Species	LDGV (low/within/upper)	HDDV (low/within/upper)	BURN (low/within/upper)	DUST (low/within/upper)	CFPP (low/within/upper)
EC	20 / 80 / 0	23 / 77 / 0	18 / 82 / 0	63 / 17 / 20	25 / 52 / 23
OC	0 / 100 / 0	35 / 65 / 0	0 / 100 / 0	66 / 27 / 7	19 / 79 / 3
OC/EC	15 / 85 / 0	0 / 100 / 0	3 / 97 / 0	-	-
TC	1 / 99 / 0	4 / 96 / 0	1 / 99 / 0	-	-
SO ₄ ⁻²	92 / 7 / 1	90 / 8 / 2	89 / 9 / 2	84 / 14 / 2	77 / 22 / 1
NO ₃ ⁻	88 / 9 / 3	82 / 13 / 4	83 / 10 / 7	78 / 15 / 7	-
Cl	45 / 52 / 2	39 / 59 / 2	41 / 50 / 9	66 / 20 / 14	59 / 25 / 16
NH ₄	11 / 7 / 82	11 / 6 / 83	11 / 11 / 78	13 / 18 / 69	11 / 13 / 76
Al	64 / 35 / 1	51 / 31 / 19	-	80 / 20 / 0	53 / 45 / 3
As	-	-	-	-	1 / 11 / 88
Ba	-	-	-	7 / 15 / 78	2 / 12 / 87
Br	18 / 81 / 1	30 / 48 / 23	34 / 34 / 32	-	37 / 33 / 29
Ca	30 / 66 / 3	36 / 39 / 25	38 / 27 / 35	33 / 64 / 3	3 / 87 / 10
Fe	14 / 78 / 8	32 / 40 / 28	26 / 32 / 42	14 / 81 / 6	6 / 63 / 32
K	35 / 33 / 32	38 / 23 / 39	6 / 75 / 19	14 / 57 / 28	30 / 25 / 46
Mn	54 / 45 / 0	46 / 48 / 6	55 / 35 / 9	66 / 34 / 0	66 / 31 / 3
Pb	20 / 39 / 41	15 / 48 / 37	23 / 28 / 49	-	19 / 34 / 47
Se	-	-	-	-	2 / 71 / 26
Si	44 / 53 / 3	41 / 49 / 10	-	11 / 89 / 0	8 / 81 / 11
Ti	-	-	-	23 / 74 / 3	15 / 41 / 44
Zn	4 / 89 / 7	22 / 42 / 36	25 / 30 / 45	32 / 27 / 41	19 / 25 / 56

4.3.2 Source apportionment based on the derived PM_{2.5} source profiles

Using LGO-derived source profiles (LDSP) based on all available samples (447 cases for the period of 8/1/1998-8/31/2000, Table 4-4) to apportion daily PM_{2.5} levels measured at the Jefferson Street site in Atlanta, typically led to similar results as when measurement-based source profiles (MBSP) were used (Marmur et al., 2005) (Figure 4-6; Tables 4-5 – 4-7), though with a couple major differences.

PM_{2.5} attributed to wood burning was 0.66 µg/m³, on average, using LDSP versus 1.1 µg/m³ using MBSP. This is driven, in part, by the higher potassium fraction in the LGO derived DUST profile, compared to the measurement-based DUST profile. Other

differences include diesel PM_{2.5} (2.3 µg/m³ using LDSP, 1.9 µg/m³ using MBSP) and “other OC” (3.1 and 2.5 µg/m³, respectively).

Table 4-4: LGO derived PM_{2.5} source profiles based on 447 days in which all relevant data (ions, EC, OC, metals, CO, SO₂, NO_y) were available (GV- gasoline vehicles; DV- diesel vehicles; BURN- vegetative burning; DUST- soil dust; CFPP- coal fired power plants)

Species	GV	DV	DUST	BURN	CFPP
SO ₄ ²⁻	0.0129 ± 0.0138	0.0128 ± 0.0139	0.0013 ± 0.0015	0.0013 ± 0.0016	0.0307 ± 0.0563
NO ₃ ⁻	0.0144 ± 0.0174	0.0161 ± 0.0205	0.0001 ± 0.0003	0.0018 ± 0.0025	0.0000 ± 0.0000
Cl ⁻	0.0240 ± 0.0236	0.0238 ± 0.0222	0.0026 ± 0.0033	0.0374 ± 0.0320	0.0023 ± 0.0039
NH ₄	0.0088 ± 0.0031	0.0088 ± 0.0030	0.0008 ± 0.0004	0.0086 ± 0.0032	0.0087 ± 0.0031
EC	0.2575 ± 0.1323	0.5654 ± 0.1570	0.0024 ± 0.0041	0.1093 ± 0.0609	0.0522 ± 0.0357
OC	0.4176 ± 0.0914	0.2063 ± 0.1059	0.0150 ± 0.0301	0.5225 ± 0.0626	0.1280 ± 0.1036
Al	0.0032 ± 0.0073	0.0003 ± 0.0004	0.0150 ± 0.0154	0.0000 ± 0.0000	0.0253 ± 0.0238
As	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0010 ± 0.0002
Ba	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0089 ± 0.0028	0.0000 ± 0.0000	0.0095 ± 0.0019
Br	0.0014 ± 0.0018	0.0005 ± 0.0004	0.0000 ± 0.0000	0.0005 ± 0.0005	0.0004 ± 0.0005
Ca	0.0109 ± 0.0133	0.0005 ± 0.0004	0.0467 ± 0.0469	0.0005 ± 0.0005	0.1157 ± 0.0552
Fe	0.0210 ± 0.0157	0.0005 ± 0.0004	0.0867 ± 0.0582	0.0006 ± 0.0005	0.0745 ± 0.0280
K	0.0005 ± 0.0005	0.0001 ± 0.0000	0.0668 ± 0.0342	0.0628 ± 0.0304	0.0066 ± 0.0041
Mn	0.0003 ± 0.0007	0.0003 ± 0.0003	0.0025 ± 0.0046	0.0002 ± 0.0004	0.0004 ± 0.0011
Pb	0.0006 ± 0.0004	0.0006 ± 0.0004	0.0000 ± 0.0000	0.0006 ± 0.0005	0.0063 ± 0.0043
Se	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0061 ± 0.0035
Si	0.0076 ± 0.0130	0.0026 ± 0.0037	0.2419 ± 0.0897	0.0000 ± 0.0000	0.1341 ± 0.0475
Ti	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0120 ± 0.0125	0.0000 ± 0.0000	0.0069 ± 0.0036
Zn	0.0074 ± 0.0054	0.0006 ± 0.0004	0.0005 ± 0.0005	0.0006 ± 0.0005	0.0075 ± 0.0037

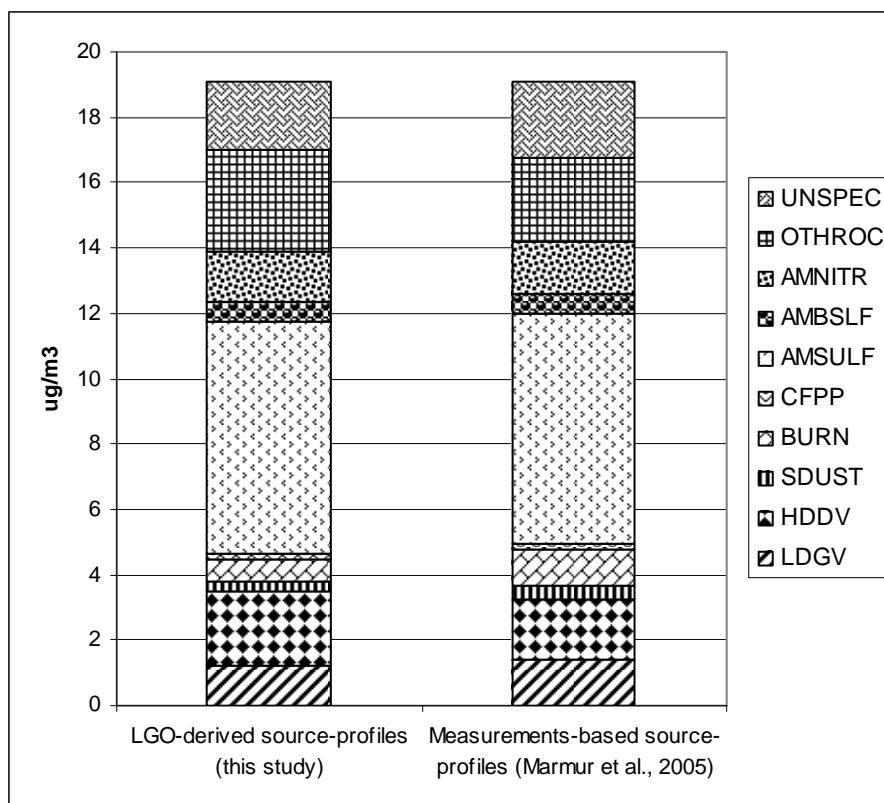


Figure 4-6. Average source contributions (8/1/98-8/31/00) to PM_{2.5} at the Atlanta Jefferson Street site, using LGO-derived source-profiles (this study) and measurement-based source-profiles (Marmur et al., 2005) (GV- gasoline vehicles; DV- diesel vehicles; BURN- vegetative burning; DUST- soil dust; CFPP- coal fired power plants; AMSULF- ammonium sulfate; AMBSLF- ammonium-bisulfate; AMNTR- ammonium-nitrate; OTHROC – Other OC; UNSPEC – unspecified)

Comparing the quality of fit achieved in the two cases (Table 4-5), finds a significantly lower chi-square value (error function being minimized) (Marmur et al., 2005) using LDSP (12.6) compared to MBSP (20.3). This is driven by several trace species, such as Br, Ca, Fe, K, Pb and Si, for which their ambient concentrations were better reconstructed using LDSP (Al was not used as a fitting species in the MBSP solution). However, EC, Cl⁻ and Zn were better fit using the MBSP. The improved fit for potassium using LDSP can partially explain the lower mass contribution of BURN using LDSP, compared to MBSP. The improved fit for Si, Fe and Al using LDSP may indicate

that the DUST profile derived by LGO is more representative of Atlanta soil dust, compared to the Alabama soil profile used (Cooper, 1981). Aluminum was excluded as a fitting species in the MBSP study (Marmur et al., 2005) because including it did not improve the fit significantly (calculated/observed ratio for Al was 4.0 when included, 4.7 when excluded), but the chi-square increased (22.4 vs. 20.3), indicating that the DUST impact was driven by another species (Si) (Marmur et al., 2006), and that the Al/Si ratio in the soil profile is too high compared to ambient measurements in Atlanta.

To assess the difference in daily variability in source impacts based on LDSP and MBSP, we also computed correlations between the various source-contribution estimates (Table 4-6). Of the five source categories for which profiles have been derived, the source inter-correlations are high for DUST (0.97), GV (0.93), and CFPP (0.89), slightly lower for BURN (0.83), and relatively low for DV (0.68). DV and BURN were previously mentioned for differences in their average source contributions based on the two methods (Figure 4-6). The correlations for all the secondary PM_{2.5} categories are high (0.95-1.00).

Table 4-5. Performance measures for the LDSP and MBSP solutions: Chi-square (error function), correlation (R) between ambient and reconstructed PM_{2.5}, percent of total mass explained, and calculated-to-observed ratios (ideally would approach 1). Bolded values are superior compared to the other solution.

	LDSP (this study) Mean (StDev)	MBSP (Marmur et al., 2005) Mean (StDev)
Chi-square (error function)	12.4 (12.0)	20.3 (16.8)
R	0.9836 (0.0349)	0.9879 (0.0324)
% total mass	92.3 (18.7)	90.5 (17.4)
SO ₄ ⁻² ratio	1.06 (0.07)	1.07 (0.07)
NO ₃ ⁻ ratio	1.16 (0.79)	1.18 (0.87)
Cl ⁻ ratio	1.31 (0.77)	1.06 (0.63)
NH ₄ ⁺ ratio	0.88 (0.15)	0.88 (0.15)
EC ratio	0.94 (0.37)	0.98 (0.13)
OC ratio	1.00 (0.03)	1.00 (0.03)
Al ratio ¹	1.50 (0.86)	4.67 (2.81)
As ratio	0.19 (0.18)	0.32 (0.25)
Ba ratio	0.26 (0.20)	0.10 (0.08)
Br ratio	1.26 (1.36)	0.39 (0.38)
Ca ratio	1.08 (0.34)	1.15 (0.34)
Fe ratio	0.85 (0.19)	0.55 (0.17)
K ratio	1.05 (0.46)	1.19 (0.49)
Mn ratio	1.57 (1.23)	0.69 (0.58)
Pb ratio	1.12 (0.95)	0.27 (0.23)
Se ratio	1.25 (1.27)	1.11 (1.20)
Si ratio	1.07 (0.19)	1.28 (0.12)
Ti ratio	1.16 (0.70)	1.27 (0.70)
Zn ratio	0.92 (0.37)	1.01 (0.35)
SO ₂ ratio	2.13 (0.89)	1.99 (0.97)
CO ratio	1.75 (0.84)	2.06 (0.83)
NO _y ratio	1.63 (0.68)	1.58 (0.66)

1 – Al was not included as a fitting species in the MBSP solution (Marmur et al., 2005). To allow for a full comparison, we also reran the MBSP analysis with Al as fitting species. Al fit was slightly improved (calculated/observed ratio of 4.0 vs. 4.7), though still very much overestimated, while the overall fit (as expressed by chi-square) worsened (22.4 vs. 20.3).

Table 4-6. Correlation matrix (R) of source-contributions based on LDSP and MBSP solutions

		LDSP (this study)								
		GV	DV	DUST	BURN	CFPP	AMSULF	AMBSLF	AMNITR	OTHROC
MBSP (Marmur et al., 2005)	GV	0.93	0.43	0.12	0.33	0.12	0.07	-0.02	0.30	0.29
	DV	0.22	0.68	0.19	0.62	0.22	0.26	-0.03	0.21	0.41
	DUST	0.17	0.33	0.97	0.17	0.22	0.33	-0.08	-0.08	0.17
	BURN	0.10	0.02	-0.15	0.83	-0.08	-0.04	0.05	0.22	0.12
	CFPP	0.28	0.39	0.15	0.16	0.89	0.21	0.06	0.08	0.31
	AMSULF	0.03	0.23	0.29	0.13	0.24	1.00	-0.08	-0.03	0.30
	AMBSLF	0.03	-0.04	-0.07	0.05	0.07	-0.12	0.95	0.17	0.01
	AMNITR	0.26	0.15	-0.09	0.25	0.02	-0.05	0.18	0.98	0.11
	OTHROC	0.42	0.70	0.20	0.28	0.30	0.29	0.03	0.14	0.97

Further analyzing differences in the DV and BURN source contributions predicted by the two methods, MBSP DV source impact is highly correlated with EC (0.96) whereas the LDSP DV source impact has a lower correlation with EC (0.72). This indicates that EC is more of a driving force in the MBSP solution than in the LDSP solution. For BURN, the correlations with K were more similar, 0.62 based on the MBSP solution, 0.67 based on LDSP. Effects of fluctuations in tracer concentrations on source contributions are investigated further by a sensitivity analysis, in which the ambient concentrations of one $PM_{2.5}$ component at a time were increased by 50% and the resulting effects on the source-attributions (using the fix derived source profiles) were analyzed. These results are compared to a similar analysis performed on the MBSP solution (Marmur et al., 2006) (Table 4-7). DV source contributions are driven mainly by EC in both solutions (62% and 70% increase in DV contribution based on LDSP and MBSP due to a 50% increase in EC concentrations), but the LDSP is less sensitive to EC and more sensitive to Si compared to the MBSP solution. In addition, EC has a bigger effect on the spilt between gasoline and diesel vehicles in the LDSP solution, scavenging more mass from the GV category compared to the MBSP case. Similarly, mass is scavenged from the BURN category using the LDSP, not so using MBSP. The BURN impact based on the LDSP solution is more sensitive to K.

Table 4-7. Percent change in average source-attributions for a 50% increase in concentrations of several tracer species (increased one at a time)

Species	% change in source-attribution for a 50% increase in the corresponding species-concentration (LDSP / MBSP)					
	GV	DV	DUST	BURN	CFPP	OTHROC
EC	-17 / -6.8	62 / 70	2.8 / 1.2	-6.7 / 0	0.5 / -1.2	-5.1 / -5.6
OC	0 / 0.6	0 / 0	0 / 0	-0.1 / 0.6	0 / 0	73 / 82
Al	0.6 / 0	-0.1 / 0.3	0.5 / 0.7	-0.4 / 1	0.8 / 1.4	-0.1 / 1.7
Br	4.8 / 0	3.1 / -0.5	-1.2 / -0.5	4.1 / 2.8	-2.7 / 0	-1.5 / -0.9
Ca	1.5 / 1.3	-0.5 / -1.3	1.7 / -2	-2.7 / 1.3	8.1 / 29	-0.1 / -0.7
Fe	6.7 / 0	-4.9 / 0	14.3 / 0	-13.9 / 0	-0.9 / 0	0.1 / 0
K	-4.7 / 0	-5.6 / -8.5	1.3 / -1.7	110 / 40	0.7 / -3.4	-7.9 / -12
Mn	0.4 / 0	0.6 / 0	4.2 / 3	-1.1 / 0	-1.4 / 0	-0.2 / 0
Pb	0 / 1.0	0.1 / -0.3	-0.7 / 0	1.2 / 0.3	0.7 / -0.3	-0.3 / -0.3
Se	-1.5 / -0.9	0.5 / 0	-1.2 / -1.4	0.8 / 0	3.1 / 8.6	0.2 / 0
Si	-5.0 / -1.0	5.3 / 0.7	27 / 37	-21 / -1.4	-1.8 / -5.5	1.5 / 0.3
Ti	-1.5 / -0.9	0.7 / 0	6.6 / 7.3	-2.4 / 0	0 / 0.9	0.6 / 0
Zn	38 / 13	-6.9 / -4	-6.1 / -2.3	6 / 0.6	-6.2 / -2.8	-2.9 / -2.8
SO ₂	-2.5 / -0.8	0.4 / 0.5	-7 / -4.6	4.8 / 1.3	48 / 32	-0.2 / -0.3
CO	7.0 / 5.6	-1.8 / -2.1	0 / -0.7	0 / 0	-1.5 / -2.8	-0.3 / -1.4
NO _y	3.2 / 3.4	0.9 / 1.7	-2.6 / -0.6	-0.3 / -0.6	-2.9 / -1.1	-0.4 / -0.6

The selection of bounds is a critical step in the analysis, and this choice can have an effect on the solutions obtained. However, setting the bounds based on well based knowledge of typical source compositions (such as in Tables 4-1 – 4-2) reduces the possibility of noise or randomness in the source-attributions. To assess the effect of bound selection on the source-apportionment results, we repeated the analysis, this time relaxing both the lower (dividing) and upper (multiplying) species fraction bounds by a factor of two. This had little effect on the temporal patterns in source contributions, with source inter-correlations of 0.93, 0.88, 0.99, 0.90, and 0.97 for GV, DV, DUST, BURN, and CFPP, respectively, for the sensitivity and baseline cases. Inter-correlations for the four secondary PM_{2.5} categories were near perfect (0.98-1.00). Average mass attributions changed by 10% or less for all source categories except GV (23% reduction compared to base-case) and DV (31% reduction). To assess whether solutions obtained are unique, we also repeated the analysis this time changing the optimization starting point. Changing the initial estimate of the mass apportioned to each category (while keeping source compositions fixed) had no effect on the final results obtained, but changing the initial

estimate of the source compositions (from the midpoint of the allowed range to the extreme) did change the results slightly, with source inter-correlations of 0.95 and above for all sources except for GV (0.88).

4.3.3 Comparison with FA results

Both sets of source apportionment results (LDSP and MBSP) provide reasonable estimates of the impacts of various sources on ambient $PM_{2.5}$ levels, as reflected by the sensitivity analysis results and correlations with ambient tracer concentrations (Marmur et al., 2006; Marmur et al., 2005). For most sources, there is significant agreement between the two, indicating that the approach presented here is capable of deriving source profiles from the receptor data and producing source impacts without the use of predetermined (measured) source profiles specific to the regions or exceptionally recent. There are, however, some differences in both total mass apportioned to some of the categories (e.g., BURN and HDDV) and the in the magnitude of the driving force for some of the species. While there is no standard by which to compare the accuracies of the LDSP and MBSP results, a comparison of these results with FA results provides an indication of consistency across methods. We compare the LDSP and MBSP results to those from a PMF study (Kim et al., 2004a) for the JST site for the period of 11/98-8/00 (Table 4-8). Only overlapping source categories are compared (gasoline, diesel, soil, wood). There is more agreement between the results based on PMF and MBSP than PMF and LDSP for diesel vehicles, wood burning and soil dust. This is expressed by both the average mass apportioned to the various categories and by the correlations between the various source-apportionment methods (Table 4-8). Impacts of gasoline vehicles and the split between diesel and gasoline vehicles (diesel/gasoline ratio of 1.6, 1.7 and 1.2 based

on PMF, LDSP and MBSP results, respectively) are more alike in PMF and LDSP results. The overall greater agreement between PMF and the MBSP results is surprising given that in theory LDSP and PMF are more alike (both derive source compositions using the ambient data). However, the differences between method inter-correlations (Table 4-8) are minor, and are a reflection of the small differences between the LDSP and MBSP results. These results might also suggest that the differences between CMB and FA results are more due to what the sources and factors, respectively, represent than due to the accuracy of the source profiles used in CMB.

Table 4-8. Comparison between source apportionment results (total mass and correlations) based on PMF (Kim et al., 2004a), LSDP (this study), and MBSP (Marmur et al., 2005)

Source category	LDSP/PMF mass ratio	MBSP/PMF mass ratio	Correlation (R): PMF, LDSP	Correlation (R): PMF, MDSP
Gasoline	1.15	1.31	0.52	0.50
Diesel	1.23	0.98	0.72	0.78
Soil	0.57	0.72	0.93	0.97
Wood	0.62	1.00	0.71	0.78

4.4 Conclusions

Results from an expanded CMB approach deriving source-compositions based on ambient data were compared with CMB results based on measured source-profiles. For most sources, there is significant agreement between the two methods. Despite overall lower residual mass obtained by the expanded approach presented here, there is no standard by which to compare the accuracy of these two methods, especially in how well they capture the temporal trends in source impacts. As such, the approach presented here can be viewed as one method to assess the representativeness of measured source-profiles

and to help identify those profiles that may be in significant error. It can also be used to quantify uncertainties in source-impact estimates, which are in part due to uncertainties in source compositions.

Acknowledgements

This work was supported by grants to Emory University from the U.S. Environmental Protection Agency (R82921301-0) and the National Institute of Environmental Health Sciences (R01ES11199 and R01ES11294) and to Georgia Tech (EPA grants R831076 and R830960). We also thank ARA (Atmospheric Research and Analysis) for access to data.

References

- Chow, J.C., Watson, J.G., Kuhns, H., Etyemezian, V., Lowenthal, D.H., Crow, D., Kohl, S.D., Engelbrecht, J.P., Green, M.C. 2004. Source profiles for industrial, mobile, and area sources in the Big Bend Regional Aerosol Visibility and Observational study. *Chemosphere*, 54(2): 185-208.
- Christensen, W.F. and Gunst, R.F., 2004. Measurement error models in chemical mass balance analysis of air quality data. *Atmospheric Environment*, 38(5): 733-744.
- Cooper, J.A., 1981. Determination of Source Contributions to Fine and Coarse Suspended Particulate Levels in Petersville, Alabama. Report to Tennessee Valley Authority, NEA Inc.
- Friedlan, S., 1973. Chemical Element Balances And Identification Of Air-Pollution Sources. *Environmental Science & Technology*, 7(3): 235-240.
- Gillies, J.A. and Gertler, A.W., 2000. Comparison and evaluation of chemically speciated mobile source PM_{2.5} particulate matter profiles. *Journal of the Air & Waste Management Association*, 50(8): 1459-1480.

- Hansen D.A., Edgerton E.S., Hartsell B.E., Jansen J.J., Hidy G.M. 2006. Air Quality Measurements for the Aerosol Research and Inhalation Epidemiology Study. Journal of the Air and Waste Management Association (in press).
- Hansen D.A., Edgerton E.S., Hartsell B.E., Jansen J.J., Kandasamy N., Hidy G.M. 2003. The southeastern aerosol research and characterization study: Part 1-overview. Journal of the Air & Waste Management Association 53(12):1460-1471.
- HEI, 2002. Emissions from Diesel and Gasoline Engines Measured in Highway Tunnels. 107.
- Hopke, P.K., 1988. Target Transformation Factor-Analysis As An Aerosol Mass Apportionment Method - A Review And Sensitivity Study. Atmospheric Environment, 22(9): 1777-1792.
- Kim, E., Hopke, P.K. and Edgerton, E.S., 2003. Source identification of Atlanta aerosol by positive matrix factorization. Journal of the Air & Waste Management Association, 53(6): 731-739.
- Kim, E., Hopke, P.K. and Edgerton, E.S., 2004a. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. Atmospheric Environment, 38(20): 3349-3362.
- Kim, E., Hopke, P.K., Larson, T.V., Maykut, N.N. and Lewtas, J., 2004b. Factor analysis of Seattle fine particles. Aerosol Science and Technology, 38(7): 724-738.
- Manchester-Neesvig, J.B., Schauer, J.J. and Cass, G.R., 2003. The distribution of particle-phase organic compounds in the atmosphere and their use for source apportionment during the southern California children's health study. Journal of the Air & Waste Management Association, 53(9): 1065-1079.
- Marmur, A., Park, S.K., Mulholland, J.A., Tolbert, P.E. and Russell, A.G., 2006. Source apportionment of PM_{2.5} 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.
- Marmur, A., Unal, A., Mulholland, J.A. and Russell, A.G., 2005. Optimization Based Source Apportionment of PM_{2.5} Incorporating Gas-to-Particle Ratios. Environmental Science and Technology, 39(9): 3245-3254.

- Maykut, N.N., Lewtas, J., Kim, E. and Larson, T.V., 2003. Source apportionment of $PM_{2.5}$ at an urban IMPROVE site in Seattle, Washington. *Environmental Science & Technology*, 37(22): 5135-5142.
- Paatero, P., 2004. User's Guide for Positive Matrix Factorization programs PMF2 and PMF3, Part 1: tutorials.
- Paatero, P. and Tapper, U., 1994. Positive Matrix Factorization - A Nonnegative Factor Model With Optimal Utilization Of Error-Estimates of Data Values. *Environmetrics*, 5(2): 111-126.
- Pinter, J.D., 1996. *Global Optimization in Action*. Kluwer Academic Publishers, The Netherlands.
- Pintér, J.D., 1997. LGO - A Program System for Continuous and Lipschitz Global Optimization. In *Developments in Global Optimization*. Kluwer Academic Publishers, Dordrecht, Boston, London.
- Seinfeld, J.H. and Pandis, S.N., 1998. *Atmospheric Chemistry and Physics*. John Wiley & Sons, Inc.
- US-EPA, 2004a. EPA-CMB8.2 Users Manual (EPA-452/R-04-011), Office of Air Quality Planning & Standards, Research Triangle Park, NC 27711. Available at <http://www.epa.gov/scram001/models/receptor/>, accessed February 5, 2003.
- US-EPA, 2004b. Protocol for Applying and Validating the CMB Model for $PM_{2.5}$ and VOC (EPA-451/R-04-001), Office of Air Quality Planning & Standards, Research Triangle Park, NC 27711.
- Zielinska, B., McDonald, J.D., Hayes, T., Chow, J.C., Fujita, E.M., Watson, J.G., 1998. Northern Front Range Air Quality Study Final Report. Volume B: Source Measurements. Available at <http://www.nfraqs.colostate.edu/nfraqs/index2.html>, accessed April 25, 2004.

CHAPTER 5

SOURCE APPORTIONMENT OF PM_{2.5} IN THE SOUTHEASTERN UNITED STATES USING RECEPTOR AND EMISSIONS-BASED MODELS: CONCEPTUAL DIFFERENCES AND IMPLICATIONS FOR TIME-SERIES HEALTH STUDIES

(A. Marmur, S.K. Park, J.A. Mulholland, P.E. Tolbert and A.G. Russell, *Atmospheric Environment*, 40, 2533-2551, 2006)

Abstract

Elevated levels of fine particulate matter (PM_{2.5}) have been associated with adverse effects on human health, but whether specific components of PM_{2.5} are responsible for specific health effects is still under investigation. A complementary approach to examining species-specific associations is to assess associations between health outcomes and sources contributing to PM_{2.5}. This approach could help target and regulate the sources that contribute most to adverse health effects. Various techniques have been developed to quantify source impacts on air quality, allowing examination of their health impacts. We compare two conceptually different approaches to source apportionment: a receptor model and an emissions-based air-quality model. Daily source impacts for July 2001 and January 2002 at four sites in the southeastern US were calculated using CMB-LGO (Chemical Mass Balance model incorporating the Lipschitz Global Optimizer), an extended CMB receptor model, and EPA's Models-3 emissions-based air-quality modeling system (MM5-SMOKE-CMAQ). The receptor model

captured more of the temporal variation in source impacts at a specific receptor site compared to the emissions-based model. Driven by data at a single site, receptor models may have some significant shortcomings with respect to spatial representativeness (unless a reduced study area is used or data from multiple sites are available). Source apportionment results from emissions-based models, such as CMAQ, may be more spatially representative as they represent an average grid-cell value. Limitations in the ability to model daily fluctuations in emissions, however, lead to results being driven mainly by regional meteorological trends, likely underestimating the true daily variations in local source impacts. Using results from either approach in a health study would likely introduce an attenuation of the observed association, due to limited spatial representativeness in receptor modeling results and to limited temporal representativeness in emissions-based models results.

Keywords: CMAQ, CMB, LGO, source-apportionment, health-effects, PM_{2.5}

5.1 Introduction

Particulate matter, especially particles smaller than 2.5 microns in diameter (PM_{2.5}), has been associated with adverse health effects and mortality in studies covering more than 150 cities (Dockery et al., 1993; Pope et al., 2002; Pope et al., 1995). Both acute and chronic exposures to PM_{2.5} have been associated with increased mortality rates and hospital visits, as well as cardiopulmonary disease, heart attacks, decreased lung function, and asthma (Dockery et al., 1993; Ebelt et al., 2000; Peters et al., 2001; Pope et al., 1995; Vedal, 1997). Particulate matter is chemically complex, being comprised of

numerous primary and secondary components, including ionic and organic compounds and dozens of trace elements. It is still unknown which specific components of $PM_{2.5}$ are more prone to cause specific health effects, although recent studies have started to address this issue (e.g., Metzger et al. 2004a; Metzger et al. 2004b; Peel et al. 2005; Hauck et al. 2004; Heal et al. 2005). However, the association between health outcomes and specific $PM_{2.5}$ components raises several issues: it is not obvious that the major cause for the health outcome is actually measured (it is impractical to measure every single $PM_{2.5}$ species) or is possibly measured inaccurately due to analytical issues; the actual health effects may be due to a combination of pollutants; many species are correlated which limits the ability to isolate species health impacts. A complementary approach is to examine associations between health outcomes and sources contributing to ambient $PM_{2.5}$ (Laden et al., 2000; Manchester-Neesvig et al., 2003; Mar et al., 2000; Tsai et al., 2000). By means of source apportionment (SA), source impacts on the receptor can be quantified, and their health impacts examined. A source-impact oriented approach could help target and regulate the sources that contribute most to adverse health effects. It could also allow for better multi-component epidemiologic modeling, as the number of major source-impact categories is typically far fewer than the number of PM components. Finally, this approach can help identify health effects of unmeasured species present in emissions from specific source categories. For example, preliminary studies have found an association between mortality and combustion-related $PM_{2.5}$ (from motor vehicles, coal combustion and wood burning), but not soil-related $PM_{2.5}$, in both cohort (Laden et al., 2000) and time-series (Mar et al., 2000) studies. Ito et al. (2004) mention that source-

oriented evaluation of PM health effects need to take into consideration the uncertainty associated with spatial representativeness of the species measured at a single monitor.

Here we compare two conceptually different approaches to PM_{2.5} source apportionment, receptor-based modeling (represented in this study by the Chemical Mass Balance model, or CMB) and emissions-based air-quality models (represented in this study by EPA's Models-3 suite of models). We address issues associated with using these techniques for time-series health studies, with special emphasis on the degree to which these approaches provide source impact estimates that are both spatially and temporally representative. So far, the focus in the source-apportionment and health literature has been on the use of various factor analytical (FA) techniques (such as PCA or PMF), to associate health outcomes with factors associated with sources of PM_{2.5} (Laden et al., 2000; Manchester-Neesvig et al., 2003; Mar et al., 2000; Tsai et al., 2000). Here we address a slightly different approach, using the CMB receptor model, which, especially in its extended form presented here (CMB-LGO), is more explicit in terms of identifying the sources in question. The main difference between CMB and factor analytical approaches is that CMB uses emission composition data to derive source impacts, whereas FA techniques derive the source compositions from trends in the ambient data during the process of estimating source impacts. In an FA application, the investigator assigns names to the obtained factors (i.e., identifies the factors as sources) based on their chemical composition; the factors are unique for each dataset/site analyzed. For example, Ito et al. (2004) illustrate that uncertainties in "naming" the factors in FA applications might cause source impacts to appear less spatially representative than the true case because the derived factor for a given source category will be different at each site. Both

approaches have advantages and disadvantages; however in terms of assessing the spatial and temporal representativeness of receptor and emissions-based models for use in health studies, FA techniques share many of the same characteristics and issues as presented here for CMB (both being based on measured ambient data).

5.2 Methods

We used both receptor-based (Chemical Mass Balance, or CMB) and emissions-based (EPA's Models-3) air-quality modeling approaches to conduct source apportionment of PM_{2.5} in Atlanta, GA and other sites in the southeastern US. The CMB receptor-based model (US-EPA, 2001) makes use of speciated ambient PM_{2.5} measurements (major ions, carbon fractions, trace elements) and typical compositions of emissions from various source categories to quantify the source contribution to measured concentrations at the receptor. It is based on the following mass balance equation, which is solved for S_j (a vector of source contributions) (US-EPA, 2001):

$$C_i = \sum_{j=1}^n f_{ij} S_j + e_i \quad (5-1)$$

where:

C_i = ambient concentration of chemical species i (mg m⁻³ in PM_{2.5});

$f_{i,j}$ = fraction of species i in emissions from source j ;

S_j = contribution (source-strength) of source j (mg m⁻³ in PM_{2.5});

n = total number of sources;

e_i = error term;

Here we applied both an extended version of CMB, referred to as CMB-LGO (Marmur et al., 2005) and the standard CMB. CMB-LGO (Lipschitz Global Optimizer) uses measured ambient gaseous concentrations (SO_2 , CO, and NO_y) to bound acceptable solutions to the mass balance equation above. For each solution obtained, ambient SO_2 , CO and NO_y concentrations are calculated based on the obtained $\text{PM}_{2.5}$ source contribution and typical gas-to-particle ratios at the source (e.g., $\text{SO}_2/\text{PM}_{2.5}$ from the various sources). The calculated value is then compared to the ambient measurement and, in case of significant differences (e.g., under/over prediction of more than a factor 3), the model then searches for the next best fit to the mass balance equation that would also adhere to the gas-phase constraints (Marmur et al., 2005). This approach reduces collinearity between sources, which is one of the major limitations to source identification using CMB (i.e., the inability to distinguish between sources with similar $\text{PM}_{2.5}$ emissions compositions). The reduced collinearity is because sources that share fairly similar $\text{PM}_{2.5}$ composition may have very different gaseous emissions. For example, gasoline and diesel engines have fairly similar $\text{PM}_{2.5}$ compositions (rich in organic and elemental carbon), but differ significantly with regards to CO and NO_x emissions. Likewise, collinearity caused by crustal elements found in both soil dust and coal-fired power plant emissions is significantly reduced with the introduction of the SO_2 constraint (SO_2 is abundant in power plant emissions, but not present in resuspended soil dust). However, use of CMB-LGO has some limitations, some of which might be important in time-series health studies. First, collinearity is not eliminated completely, and part of the daily fluctuations in the amount of mass apportioned between several sources might be due to collinearity, hence introducing an error to the time-series

epidemiologic analysis. There are also uncertainties associated with the source profiles used. The composition might not necessarily represent typical local source compositions since locally accurate source profiles are not always available. The composition may vary temporally (e.g., the effect of driving mode on the composition of PM_{2.5} from mobile sources) whereas constant values are used. Finally, the model relies on local (receptor) ambient measurements which might be significantly affected by local sources within 1-2 km and, therefore, might not represent the health study area, and might also contain measurement errors.

For these reasons, we also evaluated the use of emissions-based air-quality models for epidemiologic analyses. Such models have been used for gas-phase simulations and source-impact analysis of ozone for decades, and are typically three-dimensional representations of the atmosphere. More recently, they are being applied to PM_{2.5} source apportionment as well, apportioning mass the either sources or regions (Boylan et al., 2002; Held et al., 2005; Odman et al., 2004). The three-dimensional air quality model simulates the source impacts by solving the conservation equation expressed as:

$$\frac{\partial c_i}{\partial t} + \nabla \cdot (U c_i) = \nabla \rho D_i \nabla \left(\frac{c_i}{\rho} \right) + R_i(c_1, c_2, \dots, c_n, T, t) + S_i(x, t) \quad i = 1, 2, \dots, n \quad (5-2)$$

where:

c_i = concentration of species i ;

U = wind velocity vector;

D_i = molecular diffusivity of species i ;

R_i = rate of concentration change of species i by chemical reaction;

$S_i(x,t)$ = source/sink of i at location x and time t ;

ρ = air density;

n = number of predicted species;

The conservation equation describes the formation, transport and fate of air pollutants, including components for processing emissions, meteorology, topography, and atmospheric chemistry (Russell and Dennis, 2000). Source apportionment can be performed using direct sensitivity methods such as Direct Decoupled Method (DDM) (Dunker, 1981; Hakami et al., 2003), inert tracer methods, or by multiple applications of the model with and without emissions from target sources (“brute force”). Here we applied brute-force to the US-EPA’s Models-3 suite of models, including MM5 (Grell et al., 1999) as the meteorological model to simulate atmospheric physical dynamics; SMOKE (Houyoux et al., 2003) as the emissions processor to calculate spatial and temporal trends in emissions based on the annual emissions inventory; and the Community Multiscale Air Quality (CMAQ) model (Byun and Ching, 1999) to simulate atmospheric reactivity, transport, and deposition of chemical contaminants.

In terms of $PM_{2.5}$ source apportionment, the major difference between a receptor model and an emissions-based air quality model is the starting point. While a receptor model’s starting point is the ambient measurement, from there going backwards to estimate source contributions, the starting point of the air-quality model is the processed emissions inventory, going forward by simulating the transport and transformation of pollutants and ultimate air quality impact. The emissions used are typically processed from annual, county level emissions, using statistical daily/weekly/seasonal temporal

trends (such as peak morning rush hour and lower weekend emissions from mobile sources) and spatial information. Starting from an estimate of expected emissions at a location and time, atmospheric processes taking place during transport from source to receptor are simulated. CMAQ results provide more regionally representative values than those provided by an analysis based on a local measurement. However, results from air quality models include uncertainties arising from each step of the process (meteorological modeling, emission estimates, air quality modeling). Studies have shown that the emission inventory is one of the more uncertain, but particularly important, inputs to the air quality modeling process (NRC, 1991; Seinfeld, 1988). This is especially important in the current application where daily variability in source impacts is sought.

5.2.1 Model application

Source apportionment using CMAQ was performed on a daily basis for the months of July 2001 and January 2002. For actually applying either of these methods to a time series health study, a much larger dataset is desired (typically, several years of data). However, focusing on the reduced time periods allows more detailed examination of the issues involved in using results from either approach. Also, obtaining several years of CMAQ-based source apportionment results using the brute-force method requires significant computational resources. As tracer and other direct source-apportionment methods become available in CMAQ, it will be possible to expand this type of analysis by tracking source impacts efficiently, without having to re-apply to model for each source category examined. A longer CMB-LGO analysis has been completed and analyzed (Marmur et al., 2005).

The Models-3 air-quality modeling system was applied over much of the US using a 36 km grid, and a finer 12 km grid was used over northern Georgia (Figure 5-1). More detailed model information is presented elsewhere (Park et al., 2006a; Park et al., 2006b; Park et al., 2006c). We focused on two urban (Atlanta, GA and Birmingham, AL) and two rural (Yorkville, GA and Centerville, AL) Southeastern Aerosol Research and Characterization (SEARCH) sites (Figure 1), which include data on total $PM_{2.5}$ mass (gravimetric measure) and its components (Hansen et al., 2003; Kim et al., 2003). The main objectives of SEARCH include the understanding of composition and sources of PM in the southeast (Hansen et al., 2003; Kim et al., 2003). The SEARCH data are also being used for health studies in Atlanta in which associations of $PM_{2.5}$ with respiratory illnesses and cardiovascular disease have been observed (Metzger et al., 2004; Peel et al., 2005). In the southeastern US, a major part of the total $PM_{2.5}$ is secondary (Hansen et al., 2003; Kim et al., 2003; Marmur et al., 2005), i.e. formed in the atmosphere from precursor gases. The major secondary $PM_{2.5}$ components are sulfate from the oxidation of SO_2 and secondary OC particles formed from VOC (volatile organic compounds) emissions. The emphasis in this paper is on sources of primary $PM_{2.5}$, i.e. particulate matter emitted directly from emissions sources, due to the fact that receptor models are limited in their ability to link secondary compounds to emission sources (Burnett et al., 1998). We focused on primary $PM_{2.5}$ emissions from five source categories: gasoline vehicles, diesel vehicles, soil dust, vegetative/wood burning and coal-fired power plants (in the CMB-LGO analysis these were noted as LDGV, HDDV, SDUST, BURN and CFPP, respectively). These categories were previously identified as affecting the Atlanta airshed (Kim et al., 2004; Marmur et al., 2005; Zheng et al., 2002). Other categories, such

as meat cooking, were not quantified due to lack of tracers and collinearity issues (Marmur et al., 2005), and were likely apportioned to the unexplained fraction of OC (which also includes secondary organic aerosol). Regarding the major secondary PM_{2.5} compounds, relating ambient sulfate concentrations to power-plant emissions is fairly straight forward since these account for the vast majority of SO₂ emissions (US-EPA, 1999). The sources and chemistry of secondary organic aerosol (SOA) are still being investigated; biogenic compounds, such as monoterpene emissions from vegetation, are believed to be a major source of SOA (Carreras-Sospedra et al., 2005; Lim and Turpin, 2002). As the chemistry of SOA becomes better understood, air-quality models such as CMAQ may be able to provide a useful tool to assess the health outcomes associated with elevated levels of SOA.

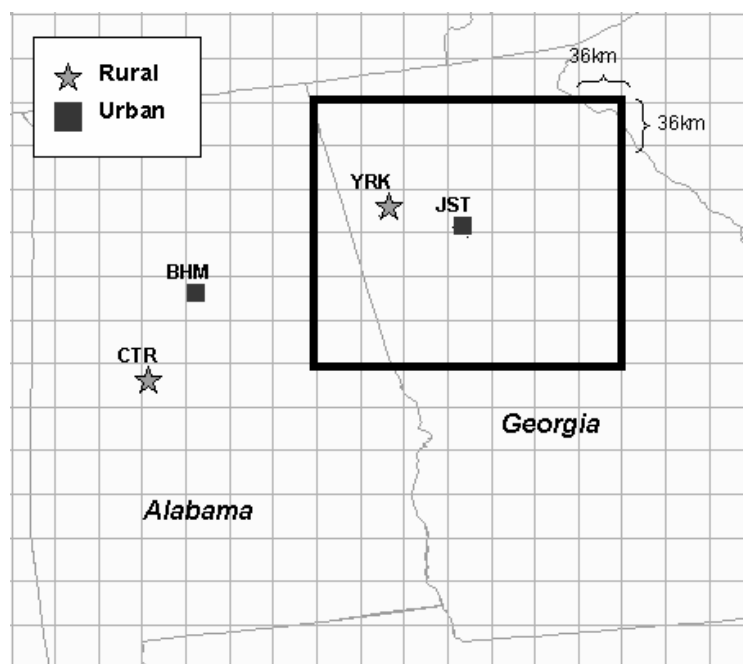


Figure 5-1. Locations of SEARCH monitoring sites in Georgia (JST and YK) and Alabama (BHM and CTR), over plotted by the grids (36km) of the air quality model (only a part of the 36 km domain is shown here). The rectangle that contains the YRK and JST sites is the sub-domain of the air quality model with a grid size of 12km (12km grids are not shown here).

For each one of the primary source categories investigated, emissions from SMOKE were tracked separately, and the air-quality model was applied six times: a base-case run including all sources, and five additional runs withholding one source category at a time (domain-wide exclusion) (Park et al., 2006b; Park et al., 2006c). A “source-category” in this case is defined as a summation of emissions from numerous source classification codes (SCC) from the emissions inventory. Specifically, the “gasoline” SMOKE category is a summation of emissions from on-road and off-road gasoline-engine powered vehicles; the “diesel” category is a summation of emissions from on-road and off-road diesel-engine powered vehicles; “dust” emissions are a summation of emissions from both paved and unpaved roads; “wood-burning” emissions are a summation of emissions from industrial and residential wood burning, prescribed burning and wildfires; “power-plant” emissions are a summation of emissions from all coal-fired boilers. Emissions from natural-gas combustion and meat charbroiling were tracked separately as well (Park et al., 2006b; Park et al., 2006c), but are not presented here, as they were not resolved by the receptor model (Marmur et al., 2005). Withholding emissions of primary PM_{2.5} particles may shift secondary formation products to other particles, changing their size and deposition velocities and change their rate of coagulation with particles from other sources. However, a comparison of the results obtained by the brute force method and a tracer method (where no sources are being withheld) showed very minor differences in the amount of mass apportioned to each source category and near perfect agreement in modeled daily variability of source impacts (Baek et al., 2005). CMB-LGO (Marmur et al., 2005) was also applied for the same time periods, using source profiles representing these five source categories (Chow

et al., 2004b; Cooper, 1981; Zielinska et al., 1998). We compared the two models used here in terms of their ability to estimate long-term (monthly) and short-term (daily) source impacts, the degree of correlation between the various source categories, and the factors driving each model. In addition, we also addressed the issue of spatial variability in source impacts using concentrations of major PM_{2.5} tracers from an additional monitoring site in the Atlanta metropolitan area (South-Dekalb site).

5.3 Results

Source apportionment results were analyzed for average source contributions, daily variability, and factors/species driving the apportionment process. Source impacts at four sites were studied using CMB-LGO and CMAQ: Jefferson Street (JST), an urban site in Atlanta, GA; Yorkville (YK), a rural site in northwestern GA; Birmingham (BHM), urban site in Birmingham, AL; and Centerville (CTR), a rural site in AL (Figure 5-1). We report the calculated source contributions using CMB-LGO and CMAQ, as well as those using regular CMB (without incorporating the gaseous data as in CMB-LGO) for reference.

When analyzing these results for use in a time-series health study, it is important to consider two aspects: temporal (daily) variation and spatial representativeness. Inaccuracies in either introduce errors in the epidemiologic analysis. The temporal (daily, in this case) variation in source impacts is the major factor driving a time-series health study, as short-term health effects are sought (other trends in the data, such as seasonal effects, are controlled for). For other purposes, such as air quality management and control strategy development, longer term results (e.g., seasonal) can be used. Obtaining

such results involves less uncertainty; as demonstrated below (Figure 5-2), the two approaches agree reasonably well on a monthly average basis, but differences in the daily source impacts are evident (Figures 5-3 - 5-6). Spatial representativeness of air quality data is an important issue when such data are used in health and exposure studies. This is especially true in this type of application because the spatial heterogeneity of source impacts is likely higher than that of individual species. Though there are not many data available to fully investigate this issue, consider the extreme example of two different but constant sources, both emitting significant amounts of the same pollutant (e.g., OC from wood burning and meat charbroiling), each located next to a monitoring site. Both sites will show an OC impact and likely follow similar meteorological trends, resulting in high correlations between OC levels at the two sites. However, the dominant source contributing to the OC is completely different. This need for increased temporal and spatial accuracy places greater burdens on the SA approach than might be demanded for air quality management that focuses on reducing annual averaged levels.

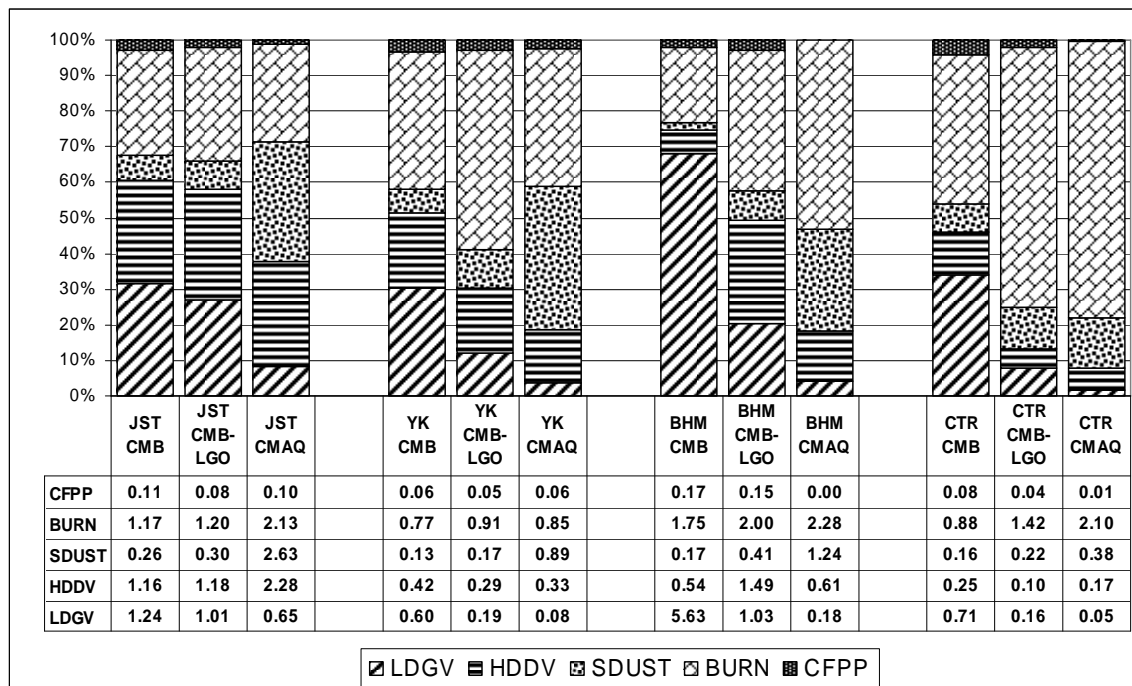


Figure 5-2. Average source contributions ($\mu\text{g m}^{-3}$) to primary $\text{PM}_{2.5}$ over a two month period (July 2001 and January 2002) of gasoline vehicles (LDGV), diesel vehicles (HDDV), soil dust (SDUST), vegetative/wood burning (BURN) and coal fired power plants (CFPP) at four SEARCH sites in Georgia and Alabama

5.3.1 Average source impacts

Analyzing the average source contributions at the four sites examined (Figure 5-2), biomass burning appears to be a major source of primary $\text{PM}_{2.5}$ in the region, with contributions ranging between 27%-77% (higher fractions in the rural sites). Average wood burning contributions obtained using the three techniques were fairly similar at YRK and BHM, while CMAQ values were significantly higher at CTR and JST. Calculated average source contributions of primary $\text{PM}_{2.5}$ from coal-fired power plants are small, less than 4% for all sites using all techniques, and the average values obtained using the various techniques are in good agreement. Calculated impacts from diesel vehicles were quite different at JST (CMAQ value higher) and BHM (CMB-LGO value

higher). The average contribution ranged between 5%-31% of the primary $\text{PM}_{2.5}$, with higher fractions at the urban sites. Major differences were observed for the gasoline vehicle primary $\text{PM}_{2.5}$ contributions, with consistently higher values obtained by CMB compared to CMB-LGO. This is likely caused by some collinearity between the gasoline vehicle source profile (OC fraction of 0.55) and the secondary/other OC profile (OC fraction of 1.0), included to account for secondary OC (and any OC not apportioned to one of the primary sources included in the analysis). Without bounding acceptable solutions based on the gaseous species, the optimal solution obtained by CMB contains a high gasoline vehicle contribution, likely including some secondary OC formation. Such a high gasoline vehicle contribution would suggest much higher ambient CO concentrations should be present than are measured. When the acceptable solutions are bounded by CO, as done in CMB-LGO, lower gasoline vehicle contributions and higher secondary/other OC contributions are obtained, $3.1 \mu\text{g m}^{-3}$ on average, compared to $1.2 \mu\text{g m}^{-3}$ using regular CMB. The solutions obtained by CMB indicate that even at the rural sites (YK and CTR), the contributions from mobile sources comprise approximately 50% of the primary $\text{PM}_{2.5}$, and that the gasoline vehicle contribution at BHM is nearly 70% of the primary $\text{PM}_{2.5}$, with an extremely high gasoline to diesel ratio of 10.4. On the other hand, solutions obtained by CMB-LGO and CMAQ indicate that the gasoline-to-diesel ratio is smaller than 1 and that wood burning is the major source of primary $\text{PM}_{2.5}$ at the rural sites. These results demonstrate the collinearity problem often encountered with regular CMB source apportionment. Significant differences in the soil-dust contribution are observed when CMAQ estimates are compared to the CMB and CMB-LGO results. CMAQ soil-dust impacts seem to be extremely overpredicted, up to ten times compared

to the receptor based results (the ambient data shows much lower levels of crustal elements than is modeled using CMAQ). This is a well known issue in CMAQ, where resuspended dust is assumed to be uniformly vertically mixed in the bottom layer of the model grid, while in practice, much of the resuspended dust is removed locally by impaction to surfaces (vehicles, leaves etc.). Often a 75% removal factor (DRI, 2000) is applied for soil dust (as in this case); however, soil dust impacts are still over-predicted.

5.3.2 Conceptual differences between CMB and CMAQ

To illustrate the conceptual differences between source apportionment from a receptor model and a 3-D air quality model, we will focus on the results from CMB-LGO and CMAQ for the Atlanta urban site (JST). Different source impacts using CMAQ follow a similar day-to-day trend (Figure 5-3), driven mainly by meteorology (mixing height and ventilation). This trend is also similar to the temporal trend in modeled EC levels (also shown in Figure 5-3), which further indicates the strong effect of meteorology on the temporal variation. CMB-LGO results, on the other hand, exhibit less correlated source category trends (Figure 5-4). In the case of CMB-LGO, the trends in the diesel source-category and measured EC levels (also shown in Figure 5-4) are similar, but these trends are different than the trends of other source categories. These differences between the daily trends in CMAQ and CMB-LGO are more clearly shown when the fraction of each source category's contribution to the total is plotted as a time series (Figures 5-5 and 5-6). CMAQ source apportionment results show little variation in relative source impact while CMB-LGO results show substantial variation.

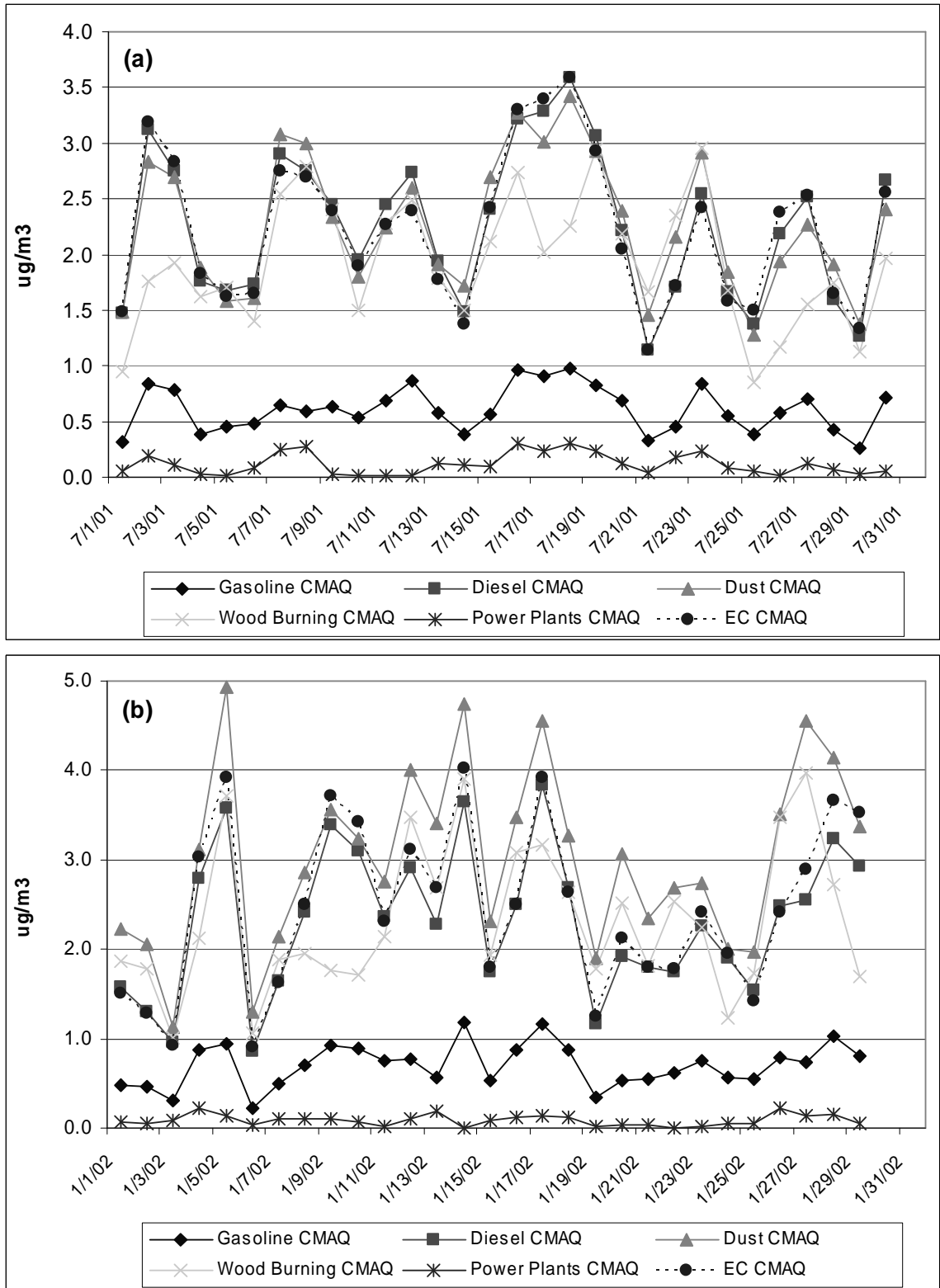


Figure 5-3. Daily source-specific contributions ($\mu\text{g m}^{-3}$) to primary PM_{2.5} at JST from gasoline vehicles, diesel vehicles, soil dust, vegetative/wood burning and coal fired power plants for the periods July 2001 (a) and January 2002 (b) using CMAQ. Also plotted are modeled EC levels at JST.

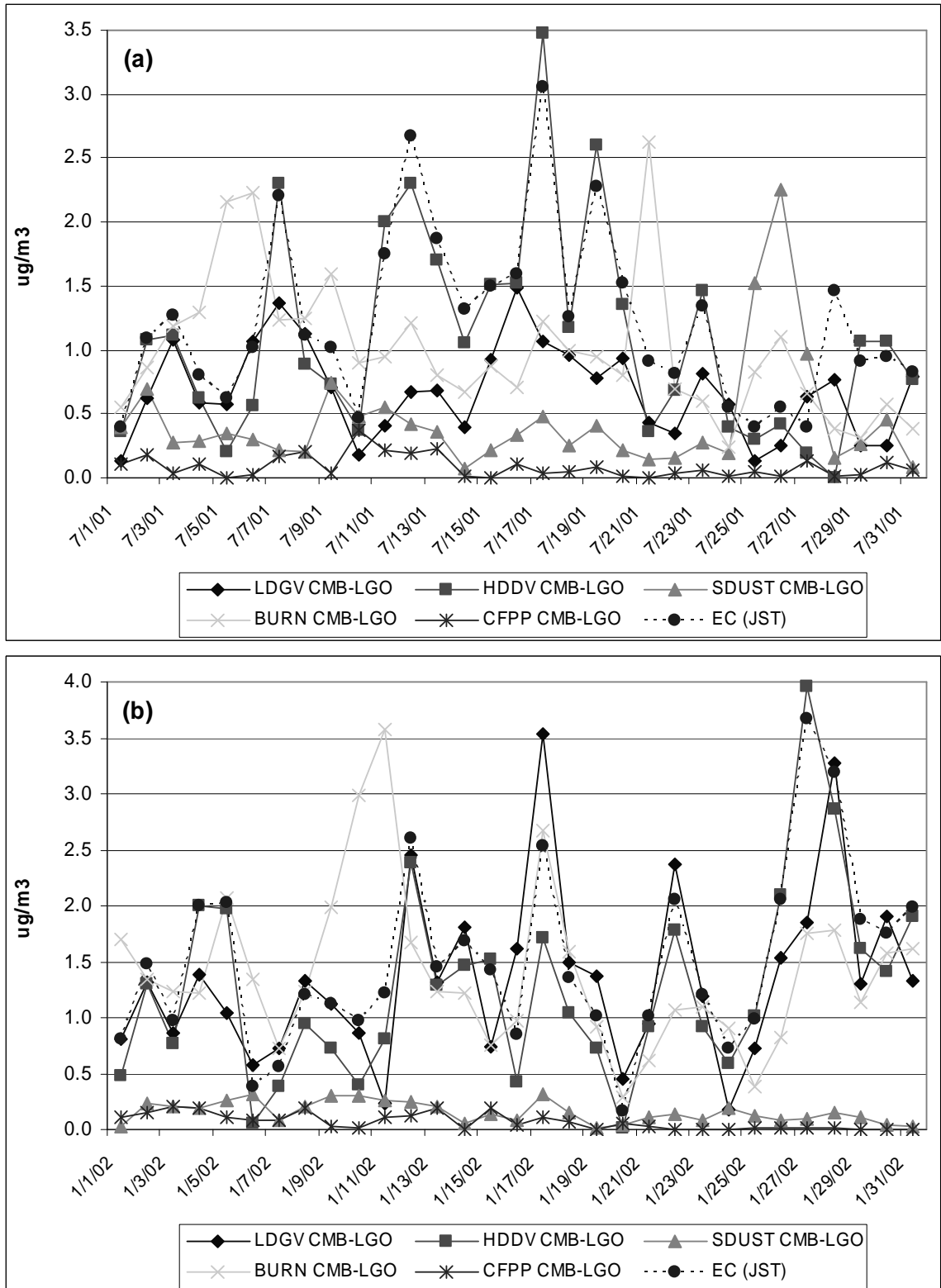


Figure 5-4. Daily source-specific contributions ($\mu\text{g m}^{-3}$) to primary PM_{2.5} at JST from gasoline vehicles (LDGV), diesel vehicles (HDDV), soil dust (SDUST), vegetative/wood burning (BURN) and coal fired power plants (CFPP) for the periods July 2001 (a) and January 2002 (b) using CMB-LGO. Also plotted are measured EC levels at JST.

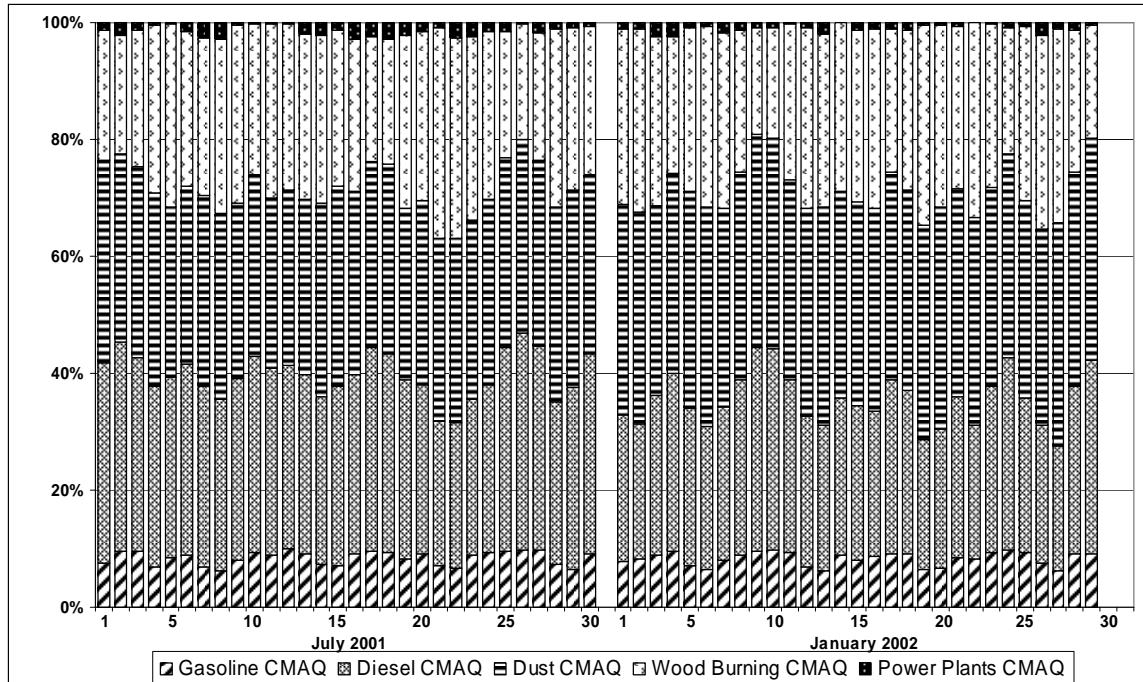


Figure 5-5. Daily fractional source contribution (relative to overall contribution from the five categories examined) at JST using CMAQ

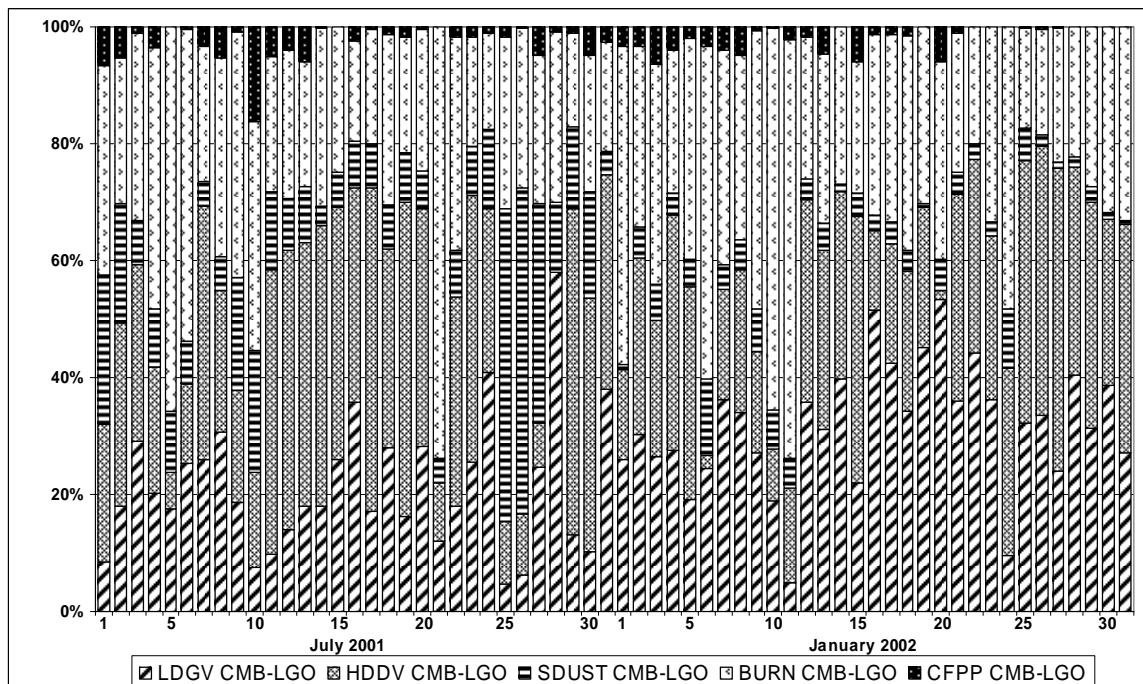


Figure 5-6. Daily fractional source contribution (relative to overall contribution from the five categories examined) at JST using CMB-LGO

Another useful way of interpreting these results is by means of a correlation matrix (R values in Table 5-1). Relatively high correlations between the CMAQ source categories and low correlations between the CMB-LGO categories are observed. As an example, in terms of the health study, the high correlation (0.94) between the gasoline and diesel categories using CMAQ might limit the ability to distinguish between the health outcomes of these two categories (even if these were a true reflection of the source impacts), while the results from CMB-LGO (significantly lower correlation, $R=0.54$) will allow such differentiation (even if these variations were mainly a reflection of collinearity). These correlations (Table 5-1) also indicate the level of agreement between the two techniques. The sources of the discrepancy include errors of both CMAQ and CMB results. However, these have different magnitudes with respect to different sources. Correlations between the corresponding source impacts using the two techniques are poor for soil dust (-0.24), wood burning (0.18), and power plants (0.16), and more reasonable for the mobile sources (0.58 for gasoline vehicles, 0.52 for diesel vehicles, and 0.59 for the sum of these two categories).

One of the major sources of possible error in results from CMAQ comes from the uncertainty in emissions estimates. Emissions used as input to CMAQ have little daily variation (Figure 5-7). Soil dust, wood burning, and power plant $PM_{2.5}$ daily emissions are assumed to be constant (except for power plants on the 4th of July and New Year's Day holidays). Mobile source emissions are given a weekly trend, with emissions modified on the 4th of July and New Year's Day holidays. In reality, soil dust emissions would depend on wind speed, humidity and recent rain fall, all having a significant effect on the amount of soil resuspended into the air (Hien et al., 2002). Stronger winds will

lead to increased resuspension of dust (though this also leads to increased mixing of the atmosphere, lowering concentrations of pollutants). On the other hand, relatively small amounts of dust will be resuspended following rain events. Such effects are not captured by SMOKE, which can explain the low correlation with the CMB-LGO results (-0.24). Also, wood burning is a source with varying activity, including prescribed and agricultural burning, residential wood combustion, and industrial use of wood bark as fuel, that is not captured by SMOKE (unless specific information about wild fires and prescribed burning is incorporated). The constant emission rate used might explain the low correlation with the receptor model results (0.18). The differences between the variations in power plant contributions may partially be due to a relatively coarse plume characterization using a 12 km grid (and a 36km grid for the Alabama sites), though uncertainties in the Se measurement and source profile fraction may contribute to the discrepancy as well. The temporal pattern of mobile sources emissions have been studied in detail (Sawyer et al., 2000). Therefore, actual mobile source emissions are probably more like the typical trends used in the model than the emissions from other sources. So it is not surprising that the correlations for the mobile sources are significantly higher (0.59 for the sum of the mobile sources). Further, mobile source emissions are more ubiquitous, so a specific local source will have a smaller effect. Results from the other three sites studied, BHM, CTR and YK, show similar trends in differences between CMAQ and CMB-LGO results, as shown by correlation matrices (Tables 5-2 – 5-4).

Table 5-1. Correlations (R) between CMB-LGO and CMAQ based source contributions to PM_{2.5} at JST

	LDGV (CMB- LGO)	HDDV (CMB- LGO)	SDUST (CMB- LGO)	BURN (CMB- LGO)	CFPP (CMB- LGO)	gasoline (CMAQ)	diesel (CMAQ)	dust (CMAQ)	Wood- burning (CMAQ)	power- plant (CMAQ)
LDGV (CMB-LGO)	1.00									
HDDV (CMB-LGO)	0.53	1.00								
SDUST (CMB-LGO)	-0.32	-0.16	1.00							
BURN (CMB-LGO)	0.31	0.10	-0.04	1.00						
CFPP (CMB-LGO)	-0.13	0.06	0.10	-0.06	1.00					
gasoline (CMAQ)	0.58	0.52	-0.05	0.28	-0.01	1.00				
diesel (CMAQ)	0.49	0.52	0.03	0.27	0.05	0.94	1.00			
dust (CMAQ)	0.70	0.59	-0.24	0.30	-0.06	0.85	0.85	1.00		
wood- burning (CMAQ)	0.62	0.61	-0.33	0.18	-0.02	0.67	0.63	0.86	1.00	
power- plant (CMAQ)	0.25	0.42	-0.10	-0.10	0.16	0.40	0.49	0.39	0.38	1.00

Table 5-2. Correlations (R) between CMB-LGO and CMAQ based source contributions to PM_{2.5} at YK

	LDGV (CMB- LGO)	HDDV (CMB- LGO)	SDUST (CMB- LGO)	BURN (CMB- LGO)	CFPP (CMB- LGO)	gasoline (CMAQ)	diesel (CMAQ)	dust (CMAQ)	Wood- burning (CMAQ)	power- plant (CMAQ)
LDGV (CMB-LGO)	1.00									
HDDV (CMB-LGO)	0.14	1.00								
SDUST (CMB-LGO)	-0.35	-0.18	1.00							
BURN (CMB-LGO)	0.10	-0.09	0.02	1.00						
CFPP (CMB-LGO)	0.33	0.11	-0.26	-0.07	1.00					
gasoline (CMAQ)	0.47	0.64	-0.29	0.15	0.11	1.00				
diesel (CMAQ)	0.42	0.57	-0.29	0.24	0.15	0.97	1.00			
dust (CMAQ)	0.57	0.41	-0.31	0.29	0.15	0.83	0.88	1.00		
wood- burning (CMAQ)	0.33	0.29	-0.26	0.26	0.03	0.66	0.72	0.77	1.00	
power- plant (CMAQ)	0.13	0.45	-0.17	0.10	0.00	0.70	0.70	0.56	0.28	1.00

Table 5-3. Correlations (R) between CMB-LGO and CMAQ based source contributions to PM_{2.5} at BHM

	LDGV (CMB- LGO)	HDDV (CMB- LGO)	SDUST (CMB- LGO)	BURN (CMB- LGO)	CFPP (CMB- LGO)	gasoline (CMAQ)	diesel (CMAQ)	dust (CMAQ)	Wood- burning (CMAQ)	power- plant (CMAQ)
LDGV (CMB-LGO)	1.00									
HDDV (CMB-LGO)	0.65	1.00								
SDUST (CMB-LGO)	-0.03	0.08	1.00							
BURN (CMB-LGO)	0.20	0.16	0.11	1.00						
CFPP (CMB-LGO)	0.58	0.71	0.00	0.12	1.00					
gasoline (CMAQ)	0.42	0.52	-0.07	0.31	0.37	1.00				
diesel (CMAQ)	0.39	0.48	-0.05	0.29	0.35	0.97	1.00			
dust (CMAQ)	0.50	0.50	-0.21	0.22	0.41	0.86	0.88	1.00		
wood- burning (CMAQ)	0.38	0.47	-0.02	0.28	0.39	0.76	0.80	0.79	1.00	
power- plant (CMAQ)	-	-	-	-	-	-	-	-	-	-

Table 5-4. Correlations (R) between CMB-LGO and CMAQ based source contributions to PM_{2.5} at CTR

	LDGV (CMB- LGO)	HDDV (CMB- LGO)	SDUST (CMB- LGO)	BURN (CMB- LGO)	CFPP (CMB- LGO)	gasoline (CMAQ)	diesel (CMAQ)	dust (CMAQ)	Wood- burning (CMAQ)	power- plant (CMAQ)
LDGV (CMB-LGO)	1.00									
HDDV (CMB-LGO)	0.36	1.00								
SDUST (CMB-LGO)	-0.27	-0.19	1.00							
BURN (CMB-LGO)	-0.12	-0.16	-0.02	1.00						
CFPP (CMB-LGO)	0.72	0.34	-0.20	-0.12	1.00					
gasoline (CMAQ)	0.36	0.30	-0.23	0.25	0.47	1.00				
diesel (CMAQ)	0.35	0.34	-0.23	0.28	0.47	0.94	1.00			
dust (CMAQ)	0.57	0.44	-0.36	0.19	0.56	0.81	0.85	1.00		
wood- burning (CMAQ)	0.20	0.39	-0.22	0.29	0.20	0.65	0.64	0.74	1.00	
power- plant (CMAQ)	0.18	-0.07	-0.09	0.01	0.28	0.48	0.46	0.25	0.10	1.00

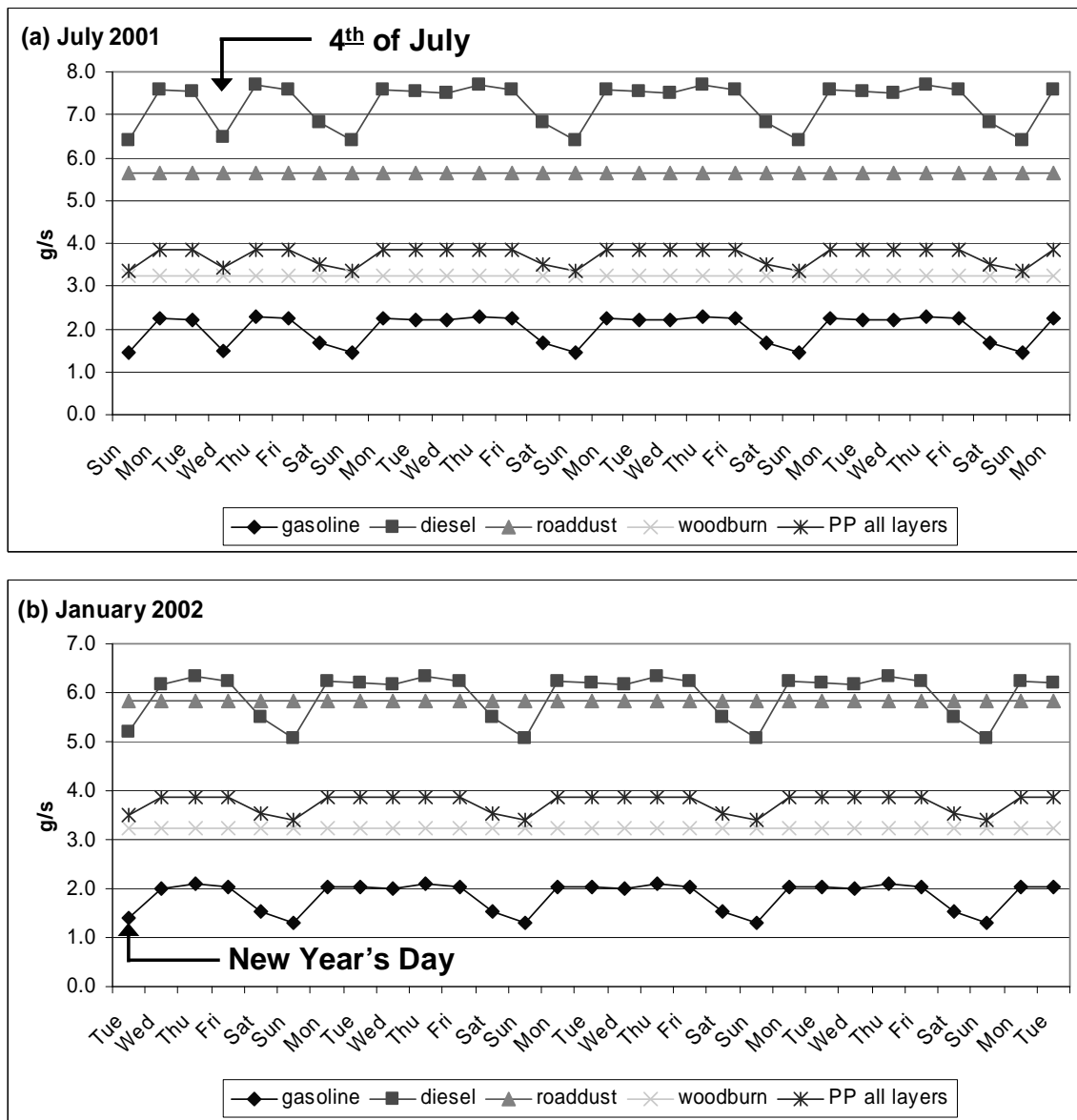


Figure 5-7. Average daily modeled emissions in the model cell where the JST site is located

When comparing the results it is important to recognize some of the limitations and issues arising from the use of CMB-LGO, contributing to the discrepancy between the two sets of results. First, as previously mentioned, there is some collinearity between the source categories, introducing more fluctuations in source impacts than is likely true. In addition, CMB-LGO is limited by the availability of ambient data. One such example is the zero contribution of diesel vehicles on July 28, 2001 estimated using CMB-LGO (see Figure 4), which is unlikely. This “error” is the result of there being no EC measurement on that day, so EC concentration was estimated as the monthly average. The high uncertainty associated with that value resulted in a low EC weighting in the error function; thus, EC levels were very poorly reconstructed for that day (17% of the filled-in value) and nearly no mass was apportioned to the HDDV category.

There are also issues pertaining to the use of temporally constant source profiles in CMB applications. For instance, $PM_{2.5}$ composition from mobile sources depends on driving mode (Shah et al., 2004), and the composition of $PM_{2.5}$ from wood burning depends on the type of wood and the burning practices (Chow et al., 2004b). Also, any receptor-based analysis is driven by point measurements, with very local influences. Finally, measurement error of $PM_{2.5}$ components introduces noise to the apportionment process.

5.3.3 Temporal variation in source impacts

To assess the degree to which results from CMB-LGO and CMAQ track the temporal variation in expected source tracer species, we calculated the correlations between these modeled daily source impacts and the daily ambient species concentrations at JST (Table 5-5). CMB-LGO source impacts are correlated with the expected tracers:

EC, OC, Zn, CO, and NO_y for gasoline and diesel vehicles; crustal elements (e.g. Si and Al) for the soil dust contribution; potassium (K) for wood burning; selenium (Se) and SO₂ for power plants. CMAQ source apportionment is not based on the ambient data, but a similar test gives additional insight to the factors driving the CMAQ source apportionment results. All of the source impacts are most correlated with EC. EC is the most abundant component of PM_{2.5} that is entirely primary, and, therefore, is in part an indicator of atmospheric stability. (Other major PM_{2.5} components, such as sulfate or OC, are either nearly entirely secondary in nature, or contain a large portion that is secondary.) Variations in CMAQ-based source impacts are mainly due to variations in meteorology, and more specifically to the results of meteorological modeling; hence, these impacts are mostly correlated with measured EC (R values ranging between 0.36-0.64). (Hogrefe et al., 2001) showed that meteorological models, and hence air quality simulations based on those results, do not capture fine scale temporal and spatial variations.

We also performed a sensitivity analysis for the CMB-LGO solution, changing the ambient concentrations of one PM_{2.5} component by one (\pm) standard-deviation of the log-normalized values per analysis. Results from this analysis (Table 5-6) indicate that the gasoline vehicle contribution is most sensitive to Zn, CO and NO_y concentrations (Zn is present in the lubricating oil of both gasoline and diesel vehicles; however, its fraction in emissions from gasoline vehicles is higher.) The diesel vehicle category is most sensitive to EC. Soil-dust is most sensitive to Si concentrations (Al was not included in the model error function (Marmur et al., 2005) and, therefore, was not included in the sensitivity analysis). Vegetative burning is most sensitive to K levels. The power-plant

contribution is most sensitive to SO₂ and Ca. The sensitivity of the secondary/other OC category to OC levels, and the lack of such sensitivity in the other categories, indicates that OC is not a driver of the source apportionment of primary PM_{2.5} (OC is first apportioned to the sources of primary PM_{2.5}, and only thereafter any OC unaccounted for is apportioned to the other/secondary OC category; hence, any change in OC levels would affect the other/secondary OC category first).

It is important to analyze the measurement accuracy of the driving species, especially in the case of trace level metals, to ensure that they are significantly above the detection limit. Average ratios of ambient concentrations to the minimum detection limits for various trace metals (Table 5-7) indicate that Zn concentrations are typically well above the detection limit, adding confidence in the gasoline-vehicle source attribution; so are the concentrations of Si, K and Ca, which were identified as key species in the attribution to the soil-dust, vegetative burning and power-plant categories, respectively. One interesting result of the sensitivity analysis was that the CFPP contribution was not very sensitive to Se levels, even though Se is a unique marker for coal combustion. However, the data in Table 7 indicate that the accuracy of the Se measurement is low compared to Ca, and in some cases Se concentrations may be near or lower than the detection limit. This explains why lowering the Se concentrations did not cause a reduction in the CFPP contribution (as the concentration approaches the detection limit, the weight in the error function is reduced), and why increasing Se concentration did cause an increase in this contribution (as Se levels increase above the detection limit, so does the weight in the error function).

Table 5-5. Correlations (R) between source contributions and ambient species concentrations measured at JST

	CMB-LGO					CMAQ				
	LDGV	HDDV	SDUST	BURN	CFPP	Gasoline	Diesel	Dust	Wood Burning	Power Plants
EC	0.67	0.94	-0.22	0.25	0.00	0.56	0.55	0.64	0.64	0.36
OC	0.69	0.76	-0.25	0.34	0.13	0.48	0.45	0.58	0.57	0.30
Al	-0.23	-0.12	0.96	-0.02	0.06	-0.03	0.03	-0.18	-0.25	-0.07
As	0.21	0.23	-0.01	0.15	0.34	0.15	0.14	0.19	0.15	-0.06
Br	0.59	0.40	-0.24	0.46	-0.02	0.41	0.40	0.46	0.43	0.29
Ca	-0.08	0.17	0.63	0.19	0.53	0.18	0.28	0.00	-0.05	0.00
Fe	0.37	0.42	0.60	0.29	0.18	0.42	0.44	0.30	0.16	0.15
K	0.46	0.24	0.09	0.71	-0.01	0.25	0.24	0.29	0.23	-0.04
Mn	0.19	0.29	0.43	0.14	0.24	0.42	0.41	0.20	0.15	0.24
Pb	-0.04	0.03	0.03	-0.03	0.51	-0.02	-0.02	-0.07	-0.06	-0.07
Se	-0.08	0.15	0.13	0.08	0.58	0.07	0.09	-0.05	-0.03	0.01
Si	-0.19	-0.10	0.98	0.03	0.19	0.02	0.10	-0.14	-0.26	-0.06
Ti	-0.12	0.00	0.90	0.09	0.15	0.07	0.12	-0.10	-0.23	-0.06
Zn	0.82	0.44	-0.32	0.41	-0.11	0.54	0.42	0.58	0.53	0.17
SO ₂	0.41	0.24	-0.15	0.34	0.58	0.32	0.28	0.40	0.31	0.20
CO	0.81	0.67	-0.24	0.34	-0.06	0.58	0.53	0.77	0.67	0.31
NO _y	0.90	0.63	-0.23	0.33	-0.09	0.57	0.49	0.73	0.64	0.17

Table 5-6. Change in average species concentrations and average source-attributions corresponding to a decrease/increase of one standard-deviation of the log normalized species concentrations Correlations (R) between source contributions and ambient species concentrations measured

	Average conc. relative to base case (- σ_{\log} / + σ_{\log})	Average source-attribution relative to base case (- σ_{\log} / + σ_{\log})					
		LDGV	HDDV	SDUST	BURN	CFPP	OtherOC
EC	0.55/ 1.81	1.04/ 0.89	0.34/ 2.13	0.99/ 1.02	0.98/ 1.00	1.03/ 0.98	1.06/ 0.91
OC	0.57/ 1.77	0.97/ 1.01	1.02/ 1.00	1.00/ 1.00	0.95/ 1.01	1.01/ 1.00	0.30/ 2.27
As	0.45/2.22	1.00/1.00	1.00/0.99	1.00/1.00	0.98/1.04	1.00/1.00	1.01/0.99
Br	0.48/2.07	1.00/1.00	1.01/0.99	1.01/0.99	0.94/1.06	1.00/1.00	1.02/0.98
Ca	0.57/ 1.76	0.98/ 1.02	1.01/ 0.98	1.01/ 0.97	0.98/ 1.02	0.72/ 1.44	1.01/ 0.99
Cu	0.31/3.20	0.99/1.03	1.00/0.99	1.00/1.00	1.00/1.01	1.00/1.00	1.00/0.99
Fe	0.60/ 1.67	0.98/ 1.00	1.01/ 1.00	0.91/ 1.00	1.01/ 1.00	0.99/ 1.00	1.00/ 1.00
K	0.63/ 1.59	1.01/ 1.00	1.06/ 0.90	1.01/ 0.98	0.69/ 1.47	1.03/ 0.96	1.09/ 0.86
Mn	0.46/2.18	1.00/1.00	1.00/1.00	0.95/1.07	1.00/1.00	1.00/1.00	1.00/1.00
Pb	0.34/2.93	0.99/1.04	1.00/0.99	1.00/1.00	1.00/1.01	1.00/0.99	1.00/0.99
Se	0.48/ 2.10	1.00/ 0.98	1.00/ 1.00	1.00/ 0.97	1.03/ 1.00	1.03/ 1.19	1.00/ 1.00
Si	0.41/ 2.46	1.01/ 0.97	0.99/ 1.02	0.40/ 2.07	1.02/ 0.96	1.04/ 0.84	0.99/ 1.01
Ti	0.64/1.55	1.01/0.99	1.00/1.00	0.94/1.08	1.00/1.00	0.99/1.01	1.00/1.00
Zn	0.53/ 1.88	0.62/ 1.23	1.12/ 0.93	1.05/ 0.96	1.00/ 1.01	1.22/ 0.95	1.07/ 0.95
SO ₂	0.34/ 2.95	0.89/ 0.97	0.99/ 1.02	1.10/ 0.82	1.66/ 1.05	0.37/ 2.23	1.00/ 0.99
CO	0.58/ 1.72	0.72/ 1.08	1.14/ 0.97	1.03/ 0.99	0.99/ 1.00	1.19/ 0.96	1.05/ 0.98
NO _y	0.53/ 1.88	0.71/ 1.06	0.93/ 1.03	1.05/ 0.99	1.09/ 0.99	1.06/ 0.98	1.04/ 0.99

Table 5-7. Average and standard deviation of the ratio of ambient concentration (Conc.) to minimum detection limit (MDL). The higher the ratio, the more accurate the measurement.

Species	(Conc./MDL) _{avg}
As	1.82±2.01
Ba	0.58±0.25
Br	5.55±3.26
Ca	4.15±2.59
Fe	161±89.1
K	8.68±4.58
Mn	2.01±1.43
Pb	4.02±11.63
Se	2.11±1.86
Si	9.91±11.5
Ti	0.77±0.56
Zn	14.5±8.81

In terms of the number of species influencing each category, if each category was driven by only one species, one might consider doing the health analysis using ambient concentrations of that species, without the apportionment into categories. However, most of these key species are not unique indicators of a single source category; they are present in emissions from several categories (e.g., Si and Ca in both soil dust and power plant emissions) and do not represent one specific source category. Some of the categories are driven by more than one species, such as Zn, CO and NO_y for LDGV and SO₂ and Ca for CFPP, and others have “secondary” driving species, such as Fe and Mn for SDUST and Br for BURN. Finally, source-apportionment results can be used to reveal which are the species most associated with various source categories, if one were to interpret an association with a species as an indication of a source-related health outcome.

5.3.4 Spatial representativeness

A major issue regarding the use of receptor-based source apportionment results in health studies is the spatial representativeness of the site. For example, Ito et al. (2004)

report that the temporal correlation of source impacts across three monitors in New-York city, 2-6 miles apart, varied significantly for sources of primary particulate matter. To address this issue for the Atlanta study, we also examined speciated $PM_{2.5}$ data from the EPA-STN (Speciation Trends Network) monitor (US-EPA) at South-Dekalb (SDK), located 15.3 km south east of the JST site. The SDK site is located near the I-285 interstate (“perimeter”) that encircles much of the Atlanta metropolitan area, while the JST site is located more towards the center of Atlanta, approximately 2 km east of the I-75/85 interstate (Figure 5-8). As the STN and SEARCH networks differ with respect to the carbon analyses method, Thermal Optical Transmittance (TOT) and Thermal Optical Reflectance (TOR), respectively (Chow et al., 2004a), it is not possible to conduct a CMB analysis of the SDK data using the same source profiles used for the JST case (these were based on TOR carbon measurements). Also, SO_2 and CO were not monitored at the SDK site, so CMB-LGO cannot be used. Instead, we compared levels and variations in major tracers for the various categories at the two sites. Data from March 2001 through December 2002 were used (samples at SDK were collected every third day, so 220 samples were available for comparison). We looked at K as an indicator for wood burning, Si and Fe as indicators for soil dust, and Se for coal fired power plants (even though the CFPP contributions were more sensitive to Ca concentrations, Ca is by no means a unique tracer for CFPP, and is often correlated with elements such as Si and Fe, indicative of soil dust). There is no unique $PM_{2.5}$ marker to separate gasoline and diesel vehicles contributions. To evaluate mobile sources as a whole, EC seems to be the most suitable (OC is partially secondary), but the comparison between EC at JST and SDK will include some noise due to the two different techniques used (TOR and TOT). EC is

also emitted from wood burning and other processes, so it cannot be view entirely as mobile-source related. We also examine Zn data, as it was correlated with gasoline vehicles impacts at the other sites examined.

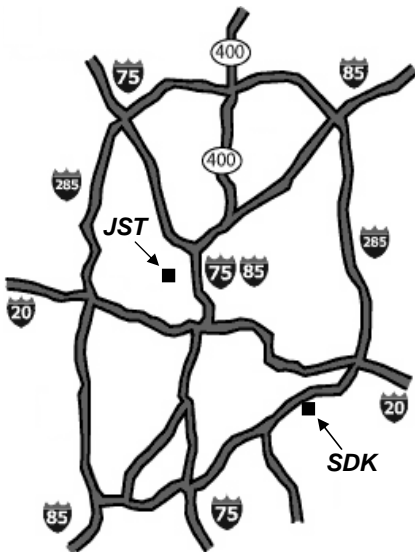


Figure 5-8. Location of JST and SDK monitoring sites with respect to major interstates in Atlanta

Potassium levels at JST and SDK are highly correlated (Figure 5-9). This likely indicates a spatially homogeneous source (residential/industrial wood combustion) or distant plume sources (prescribed agricultural burning) hitting the two monitors similarly. Crustal elements (Si, Fe) are also highly correlated, indicating regional/global dust events, and/or soil moisture resuspension effects, assuming that rain events occur similarly at the two sites, and that soil moisture and wind speed have a strong influence on the resuspension of local dust. However, when high Si events are excluded, the correlation is lower, indicating local effects (though Fe correlations still remain high). Se is poorly correlated between the two sites, likely representing the directionality of

impacts from power plant plumes. This is also demonstrated in (Wade et al., 2004), where SO_2 concentrations (as a power plant marker) were the least spatially homogeneous of the primary gaseous pollutants. EC correlations are surprisingly high, considering the differences in measurement (TOR vs. TOT) and location in proximity to major highways. One possible explanation for the high correlation can be the role of atmospheric stability in daily variations of EC (and other pollutants) concentrations. Zn correlations are also relatively high, but drop when outliers are excluded. These findings are in overall agreement with data from (Wade et al., 2004) in which CO (as a mobile source tracer) measurements at three sites in Atlanta, 11.5-16.8 km apart, were correlated at levels of 0.65-0.76. Such results provide information about spatial variability of source impacts and site representativeness in the Atlanta area. Wood burning and soil dust contributions found at either site seem to be relatively spatially representative of the Atlanta urban area, as indicated by the correlations of potassium, silicon and iron. Power plant impacts seem to be local, based on the low inter-site selenium correlations. It is difficult to draw conclusions regarding the spatial representativeness of mobile source impacts due to the lack of a unique marker and CO data. However, weighing both the EC and zinc correlations, it seems that mobile sources impacts are “intermediately” representative, i.e., likely more spatially representative than power plant impacts, but less than wood burning and soil dust impacts. Note that these results are based on a preliminary analysis of two sites in Atlanta, and may represent a local phenomenon.

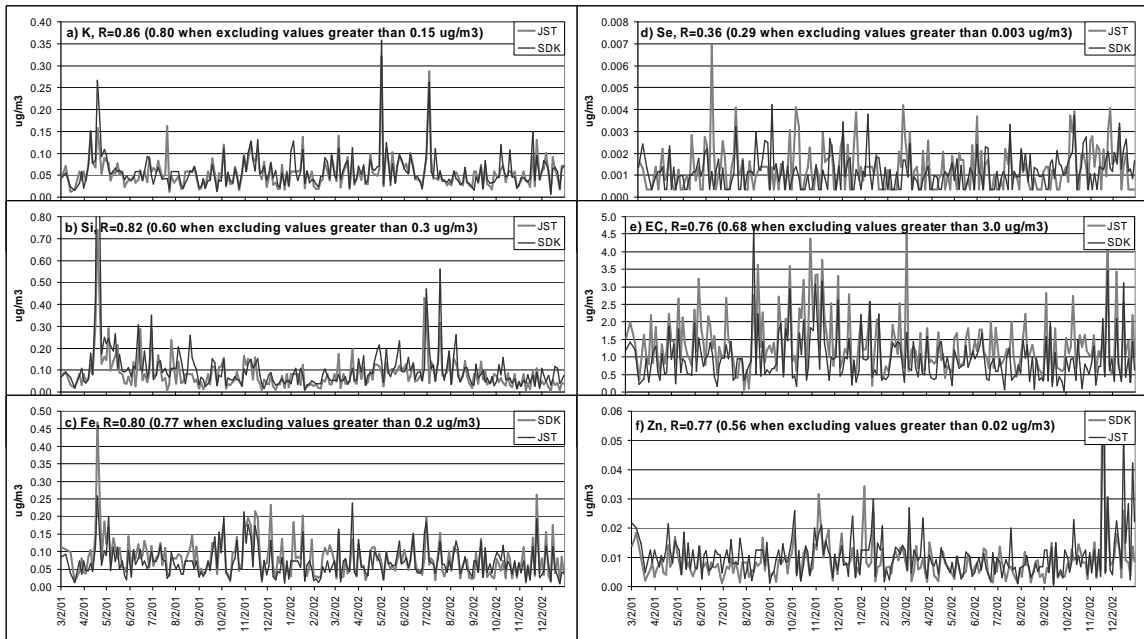


Figure 5-9. Time series comparison between ambient concentrations at two Atlanta monitoring sites located 15.3 km apart (SDK and JST) for K (a), Si (b), Fe (c), Se (d), EC (e) and Zn (f)

5.4 Discussion

Associating health outcomes with sources, rather than pollutants, may have several advantages relating to both the epidemiologic modeling process and the regulatory process. For such analyses, source impacts that capture both the temporal and spatial variability need to be generated. Receptor models, such as CMB-LGO, capture more of the temporal variation in source impacts at a specific receptor site, compared to emissions-based models, though this variation might be overestimated due to collinearity between sources. Being driven by data at a single site, receptor models may have some significant shortcomings with respect to spatial representativeness and exposure issues. Source apportionment results from emission-based models, such as CMAQ, may be more

spatially representative, as they represent an average grid-cell value. However, limitations in the ability to model fine-scale meteorological fluctuations and daily fluctuations in emissions lead to results being driven mainly by regional meteorological trends (atmospheric stability), likely underestimating the true daily variations in source impacts.

The impact of a lack of spatial representativeness of estimated source impacts, anticipated in receptor modeling output, would likely introduce a bias to the null in epidemiologic models (i.e., an attenuation of the observed association). The degree of spatial representativeness varies by source, and results for those sources with poor spatial representativeness (such as power plants) will have a greater degree of bias to the null. By characterizing the degree of spatial representativeness, investigators can take measures to handle this issue, such as reducing the study area included in analyses for the less representative sources, or possibly using data from several monitoring sites, if available. The impact of relatively limited capture of true day-to-day variation in the source impacts, anticipated to be more of an issue with emissions-based models than receptor models (though these may over-estimate the temporal variation), is also likely to be a bias to the null in the epidemiologic models. In the case of emissions-based models, the fact that the day-to-day variation in the source apportionment estimates is in large part a result of the meteorological conditions, and also influenced to a lesser extent by fixed day-of-week and seasonal patterns in the estimated emissions input, could lead to model instability, intractable confounding by meteorological conditions and temporal factors such as day-of-week and season, as well as limited usefulness in discerning the relative impacts of the sources on health outcomes.

Acknowledgements

This work was supported by grants to Emory University from the U.S. Environmental Protection Agency (R82921301-0) and the National Institute of Environmental Health Sciences (R01ES11199 and R01ES11294) and to Georgia Tech (EPA grants R831076 and R830960). We would also like to thank Georgia Power and Southern Company for continuing support of Dr. Russell's work in this area. We are grateful to the Electric Power Research Institute and ARA (Atmospheric Research and Analysis) for both providing access to data used in this analysis and ongoing discussions.

References

- Baek, J., Park, S.K. and Russell, A.G., 2005. Source apportionment of fine organic aerosol using CMAQ Tracer, Proceedings of the Models-3 User's workshop, Research Triangle Park, North Carolina.
- Boylan, J.W., Odman, M.T., Wilkinson, J.G., Russell, A.G., Doty K., Norris W., McNider, R., 2002. Development of a comprehensive, multiscale "one-atmosphere" modeling system: application to the Southern Appalachian Mountains. *Atmospheric Environment*, 36(23): 3721-3734.
- Burnett, R.T., Cakmak, S., Raizenne, M.E., Stieb, D., Vincent, R., Krewski, D., Brook, J.R., Philips, O., Ozkaynak, H., 1998. The association between ambient carbon monoxide levels and daily mortality in Toronto Canada. *Journal of the Air & Waste Management Association*, 48(8): 689-700.
- Byun, D.W. and Ching, J.K.S., 1999. Science Algorithms of the EPA-Models-3 Community Multiscale Air Quality (CMAQ) Modeling System, EPA Report, EPA/600/R-99/030, NERL, Research Triangle Park, NC.
- Carreras-Sospedra, M., Griffin, R.J. and Dabdub, D., 2005. Calculations of incremental secondary organic aerosol reactivity. *Environmental Science & Technology*, 39(6): 1724-1730.

- Chow, J.C., Watson, J.G., Chen, L.W.A., Arnott, W.P. and Moosmuller, H., 2004a. Equivalence of elemental carbon by thermal/optical reflectance and transmittance with different temperature protocols. *Environmental Science & Technology*, 38(16): 4414-4422.
- Chow, J.C., Watson, J.G., Kuhns, H., Etyemezian, V., Lowenthal, D.H., Crow, D., Kohl, S.D., Engelbrecht, J.P., Green, M.C. 2004b. Source profiles for industrial, mobile, and area sources in the Big Bend Regional Aerosol Visibility and Observational study. *Chemosphere*, 54(2): 185-208.
- Cooper, J.A., 1981. Determination of Source Contributions to Fine and Coarse Suspended Particulate Levels in Petersville, Alabama. Report to Tennessee Valley Authority, NEA Inc.
- Dockery, D.W. et al., 1993. An Association Between Air-Pollution And Mortality In 6 United-States Cities. *New England Journal Of Medicine*, 329(24): 1753-1759.
- DRI, 2000. Reconciling Urban Fugitive Dust Emissions Inventory and Ambient Source Contribution Estimates: Summary of Current Knowledge and Needed Research, Desert Research Institute Document No. 6110.4F, Reno, NV 89512.
- Dunker, A.M., 1981. Efficient Calculation Of Sensitivity Coefficients For Complex Atmospheric Models. *Atmospheric Environment*, 15(7): 1155-1161.
- Ebelt, S.T., Petkau, A.J., Vedal, S., Fisher, T.V. and Brauer, M., 2000. Exposure of chronic obstructive pulmonary disease patients to particulate matter: Relationships between personal and ambient air concentrations. *Journal of the Air & Waste Management Association*, 50(7): 1081-1094.
- Grell, G.A., Dudhia, J. and Stauffer, D.R., 1999. A description of the fifth-generation Penn State/NCAR mesoscale model (MM5).
- Hakami, A., Odman, M.T. and Russell, A.G., 2003. High-order, direct sensitivity analysis of multidimensional air quality models. *Environmental Science & Technology*, 37(11): 2442-2452.
- Hansen D.A., Edgerton E.S., Hartsell B.E., Jansen J.J., Kandasamy N., Hidy G.M. 2003. The southeastern aerosol research and characterization study: Part 1-overview. *Journal of the Air & Waste Management Association* 53(12):1460-1471.
- Held, T., Ying, Q., Kleeman, M.J., Schauer, J.J. and Fraser, M.P., 2005. A comparison of the UCD/CIT air quality model and the CMB source-receptor model for primary airborne particulate matter. *Atmospheric Environment*, 39: 2281-2297.

- Hien, P.D., Bac, V.T., Tham, H.C., Nhan, D.D. and Vinh, L.D., 2002. Influence of meteorological conditions on PM_{2.5} and PM_{2.5-10} concentrations during the monsoon season in Hanoi, Vietnam. *Atmospheric Environment*, 36(21): 3473-3484.
- Hogrefe, C. et al., 2001. Evaluating the performance of regional-scale photochemical modeling systems: Part II - ozone predictions. *Atmospheric Environment*, 35(24): 4175-4188.
- Houyoux, M., Vukovich, J., Brandmeyer, J.E., Seppanen, C. and Holland, A., 2003. Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE): User Manual, Version 2.0.
- Ito, K., Xue, N. and Thurston, G., 2004. Spatial variation of PM_{2.5} chemical species and source-apportioned mass concentrations in New York City. *Atmospheric Environment*, 38(31): 5269-5282.
- Kim, E., Hopke, P.K. and Edgerton, E.S., 2003. Source identification of Atlanta aerosol by positive matrix factorization. *Journal of the Air & Waste Management Association*, 53(6): 731-739.
- Kim, E., Hopke, P.K. and Edgerton, E.S., 2004. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. *Atmospheric Environment*, 38(20): 3349-3362.
- Laden, F., Neas, L.M., Dockery, D.W. and Schwartz, J., 2000. Association of fine particulate matter from different sources with daily mortality in six US cities. *Environmental Health Perspectives*, 108(10): 941-947.
- Lim, H.J. and Turpin, B.J., 2002. Origins of primary and secondary organic aerosol in Atlanta: Results of time-resolved measurements during the Atlanta supersite experiment. *Environmental Science & Technology*, 36(21): 4489-4496.
- Manchester-Neesvig, J.B., Schauer, J.J. and Cass, G.R., 2003. The distribution of particle-phase organic compounds in the atmosphere and their use for source apportionment during the southern California children's health study. *Journal of the Air & Waste Management Association*, 53(9): 1065-1079.
- Mar, T.F., Norris, G.A., Koenig, J.Q. and Larson, T.V., 2000. Associations between air pollution and mortality in Phoenix, 1995-1997. *Environmental Health Perspectives*, 108(4): 347-353.

- Marmur, A., Unal, A., Mulholland, J.A. and Russell, A.G., 2005. Optimization Based Source Apportionment of PM_{2.5} Incorporating Gas-to-Particle Ratios. *Environmental Science and Technology*, 39: 3245-3254.
- Metzger, K.B., Tolbert P.E., Klein M., Peel J.L., Flanders, W.D., Mulholland, K., 2004. Ambient air pollution and arrhythmic events in patients with implanted cardioverter defibrillators, 1993-2002. *Epidemiology*, 15(4): S22-S23.
- NRC, 1991. Rethinking the Ozone Problem in Urban and Regional Air Pollution. National Academy Press, Washington, DC.
- Odman, M.T., Russell, A.G. and Boylan, J.W., 2004. Estimates of PM_{2.5} levels in the southeastern United States for the year 2010: What else can be done? *Fuel Processing Technology*, 85(6-7): 631-639.
- Park, S.K., Cobb, C.E., Wade, K., Mulholland, J.A., Hu, Y., Russell, A.G., 2006a. Uncertainty in Air Quality Model Evaluation for Particulate Matter due to Spatial Variation in Pollutant Concentrations. *Atmospheric Environment* (in press).
- Park, S.K., Ke, L., Yan, B., Russell, A.G. and Zheng, M., 2006b. Source apportionment of PM_{2.5} using a three-dimensional air quality model and a receptor model, *Proceedings of an AAAR international specialty conference - Particulate Matter Supersites Program and Related Studies*, Atlanta, Georgia.
- Park, S.K., 2006c. Source Apportionment of PM_{2.5}: Comparison between Receptor and Air Quality Models. *Environmental Science & Technology* (in press).
- Peel, J.L., Tolbert, P.E., Klein, M., Metzger, K.B., Flanders, W.D., Todd, K., 2005. Ambient air pollution and respiratory emergency department visits. *Epidemiology*, 16(2): 164-174.
- Peters, A., Dockery, D.W., Muller, J.E. and Mittleman, M.A., 2001. Increased particulate air pollution and the triggering of myocardial infarction. *Circulation*, 103(23): 2810-2815.
- Pope, C.A., Burnett R.T., Thun, M.J., Calle, E.E., Krewski, D., Ito, K. 2002. Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. *JAMA-Journal of the American Medical Association* 287(9):1132-1141.
- Pope, C.A., Thun, M.J., Namboodiri, M.M., Dockery, D.W., Evans, J.S., Speizer, F.E. 1995. Particulate Air-Pollution as a Predictor of Mortality in a Prospective-Study of US Adults. *American Journal of Respiratory and Critical Care Medicine* 151(3):669-674.

- Russell, A. and Dennis, R., 2000. NARSTO critical review of photochemical models and modeling. *Atmospheric Environment*, 34(12-14): 2283-2324.
- Sawyer, R.F., Harley, R.A., Cadle, S.H., Norbeck, J.M., Slott, R., Bravofet, H.A., 2000. Mobile sources critical review: 1998 NARSTO assessment. *Atmospheric Environment*, 34(12-14): 2161-2181.
- Seinfeld, J.H., 1988. Ozone Air Quality Models: a critical review. *Journal of the Air Pollution Control Association*, 38: 616-645.
- Shah, S.D., Cocker, D.R., Miller, J.W. and Norbeck, J.M., 2004. Emission rates of particulate matter and elemental and organic carbon from in-use diesel engines. *Environmental Science & Technology*, 38(9): 2544-2550.
- Tsai, F.C., Apte, M.G. and Daisey, J.M., 2000. An exploratory analysis of the relationship between mortality and the chemical composition of airborne particulate matter. *Inhalation Toxicology*, 12: 121-135.
- US-EPA, PM_{2.5} Monitoring Information: Speciation Information.
- US-EPA, 1999. 1999 National Emission Inventory (NEI): Air Pollutant Emission Trend (www.epa.gov/ttn/chief/trends/, accessed March 12, 2004).
- US-EPA, 2001. CMB8 User's Manual, EPA-454/R-01-XXX, Office of Air Quality, Planning and Standards, Research Triangle Park, NC 27711.
- Vedal, S., 1997. Ambient particles and health: Lines that divide. *Journal of the Air & Waste Management Association*, 47(5): 551-581.
- Wade, K.S., Marmur A., Mulholland J.A., Russell A.G., Peel J.L., Klein M. 2004. Analyses of air quality data for an emergency department study, 1993-2002. *Epidemiology* 15(4):S61-S61. et al., 2004. Analyses of air quality data for an emergency department study, 1993-2002. *Epidemiology*, 15(4): S61-S61.
- Zheng, M., Cass, G.R., Schauer, J.J. and Edgerton, E.S., 2002. Source apportionment of PM_{2.5} in the southeastern United States using solvent-extractable organic compounds as tracers. *Environmental Science & Technology*, 36(11): 2361-2371.
- Zielinska, B., McDonald, J.D., Hayes, T., Chow, J.C., Fujita, E.M., Watson, J.G., 1998. Northern Front Range Air Quality Study Final Report. Volume B: Source Measurements. Available at <http://www.nfraqs.colostate.edu/nfraqs/index2.html>, accessed April 25, 2004.

CHAPTER 6

INTERMETHOD VARIABILITY IN ASSOCIATIONS BETWEEN SOURCE-APPORTIONED PM_{2.5} AND DAILY EMERGENCY- DEPARTMENT VISITS IN ATLANTA, GA

Abstract

Recent particulate-matter (PM) health effects research has focused on identifying the specific components of aerosol that pose the greatest health risks, but few epidemiologic studies have included source apportionment data in their examinations of PM health effects. This paper presents and analyzes results from source-specific epidemiologic analyses in Atlanta, Georgia, using data from several source apportionment methods (Positive Matrix Factorization [PMF] and the first application of a Chemical Mass Balance [CMB] model in a health-study). Atlanta is a unique location for conducting this type of health effects study given the existence of an extensive time-series of daily chemically-resolved aerosol measurements, detailed gaseous pollutant monitoring and corresponding hospital records. Despite methodological differences and uncertainties in the apportionment process, good agreement was observed between the CMB-LGO (Lipschitz Global Optimizer) and PMF based risk ratios, indicating to the usefulness of applying apportionment methods in health studies. Preliminary epidemiologic analysis found mobile-source related PM_{2.5} significantly associated with cardiovascular outcomes; wood burning PM_{2.5} significantly associated with respiratory outcomes; soil dust significantly associated with asthma. “Other” OC was found

significantly associated with various respiratory outcomes, with high risk-ratios (RR), though further analyses, such as based on longer datasets, are needed to support this finding. Temporal patterns in source impacts suggests that “other” OC is associated with secondary formation. Despite generally good agreement between RR estimates based on source contributions and tracer species, source apportionment and sensitivity analyses are needed to determine the most suitable tracers for each source-category. Source apportionment methods (CMB-LGO in this case) can also serve as a useful tool for estimating SOA concentrations (especially, in lieu of a SOA tracer) and associated health effects.

6.1 Introduction

Recent particulate matter (PM) health effects research has focused on identifying the specific components of aerosol that pose the greatest health risks (Hauck et al., 2004; Heal et al., 2005; Mar et al., 2000; Metzger et al., 2004a; Metzger et al., 2004b; Peel et al., 2005). Few epidemiologic studies, however, have included source apportionment data in their examinations of PM health effects (Laden et al., 2000; Mar et al., 2000; Thurston et al., 2005). Associating health outcomes with source-specific PM has several advantages relating to the epidemiologic modeling process, such as better treatment of multi-component interactions, and can provide important information to regulators to tighten controls on sources more prone to causing health outcomes. There are, however, uncertainties regarding optimal methods for conducting PM source apportionment, as well as a lack of suitable air quality and health effects data for analysis. A recent study examined the association between mortality and source-resolved PM measurements using

several principal components and factor analysis approaches (Thurston et al., 2005), and found that variability among the source-specific mortality risks was small when compared to overall model uncertainty, suggesting that these apportionment methods may be useful in discerning source-specific health effects. These findings were based on a relatively limited sample size and, thus, were not able to robustly identify specific source categories (e.g., mobile sources). Questions also remain concerning the generalizability of these findings to other locations with different aerosol compositions and whether analyses using other source apportionment methods, notably chemical mass balance (CMB) and source-based modeling, will show the same pattern of agreement.

In this chapter, results from source-specific epidemiologic analyses in Atlanta, Georgia, using data from several source apportionment methods are presented and compared. Atlanta is a unique location for conducting this type of health effects study given the existence of an extensive time-series of daily chemically-resolved aerosol measurements and corresponding hospital records, which have been characterized in several atmospheric dynamics (Hansen et al., 2006; Hansen et al., 2003), source apportionment (Kim et al., 2004; Marmur et al., 2006; Marmur et al., 2005) and epidemiologic analyses (Metzger et al., 2004a; Metzger et al., 2004b; Peel et al., 2005).

6.2 Methods

6.2.1 Source apportionment of the Atlanta aerosol

Several source-apportionment studies have been conducted on speciated PM_{2.5} (PM smaller than 2.5 µm in size) data collected at the SEARCH Jefferson Street (Hansen et al., 2006; Hansen et al., 2003), which also served as the former Atlanta PM Supersite

(Solomon, 2003). Source-apportionment methods applied include both the traditional Chemical Mass Balance (CMB) model and an expanded CMB model (CMB-LGO, incorporating the Lipschitz Global Optimizer) (Marmur et al., 2005), Positive Matrix Factorization (PMF) (Kim et al., 2004) and Models-3/CMAQ (Community Multiscale Air Quality model), an emissions-based air-quality modeling system (Marmur et al., 2006). Findings from these studies assisted in the selection of the most suitable source-apportionment results to be applied in the epidemiologic study.

A comparison of results based on CMB and CMB-LGO (Marmur et al., 2005) indicated that collinearity in source compositions might lead to “misplacement” of emissions between source-categories and to increased daily variability in source impacts when using the traditional CMB model. Such issues are dampened out significantly when the CMB-LGO approach is applied, as shown for gasoline-vehicles (GV) and primary PM_{2.5} coal-fired power plant (CFPP) contributions (Figure 6-1). By applying the CMB-LGO approach, more plausible results were also obtained for soil-dust, which is anti-correlated with the CFPP contributions (both rich in crustal material) in CMB ($R=-0.23$ for the period of 8/1998-8/2000, as it is zeroed-out whenever CFPP contribution peaks), and not correlated with CFPP in CMB-LGO ($R=0.07$; expected, due to the episodic nature of these two categories, expressed by dust storms and plume fumigation events). More plausible results were also obtained for vegetative-burning (BURN) and “other” (unapportioned) OC. “Other” OC includes secondary organic aerosol (SOA) as well as any primary OC of sources not accounted for (if present). It is therefore unlikely to observe no “other” OC/SOA in Atlanta during summertime, as estimated by CMB on specific August days (Figure 6-2). This is likely the result of overprediction of the BURN

impacts (rich in OC) by CMB, as also observed by the over-allocation of K on those days (Figure 6-2). These trends are also observed in the correlations (R) between the BURN and “other” OC categories (-0.14 in CMB; 0.30 in CMB-LGO). These increased source-intercorrelations are an indication of reduced collinearity in the CMB-LGO solution compared to “regular” CMB.

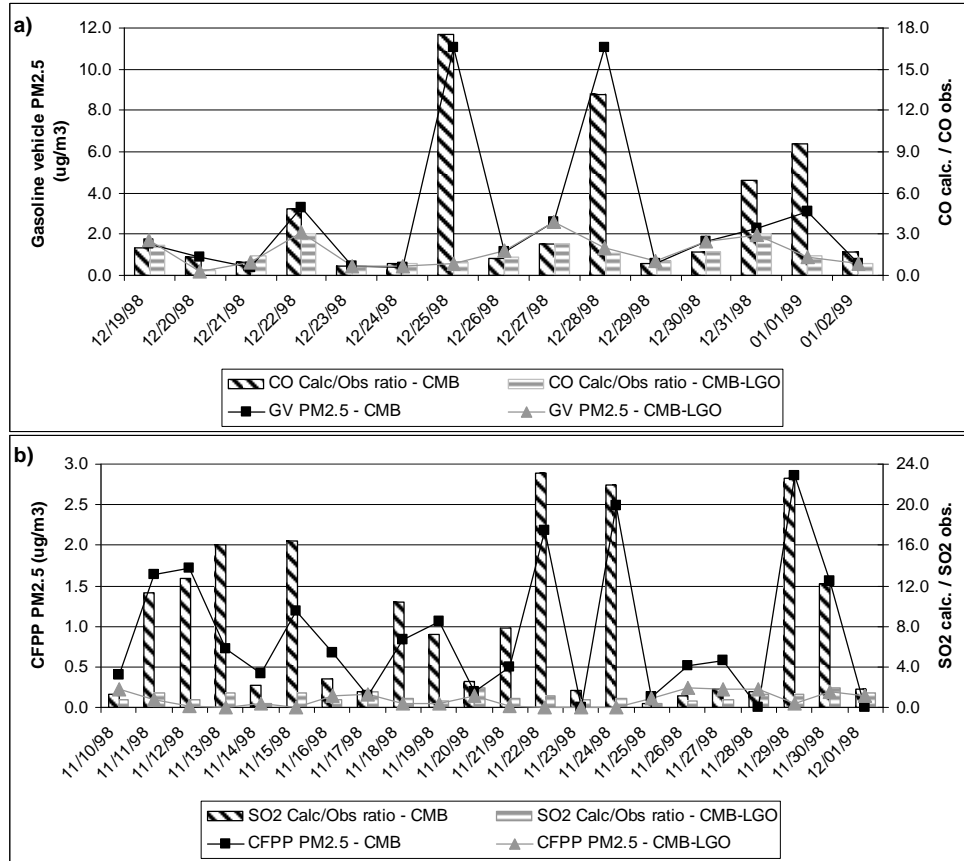


Figure 6-1. Daily PM_{2.5} source impacts of gasoline vehicles (GV; a) and coal-fired power plants (CFPP; b) for select periods. CMB results exhibit more variability in daily source impacts than results based on CMB-LGO, with extreme contributions accompanied by extreme over-predictions of CO and SO₂ concentrations.

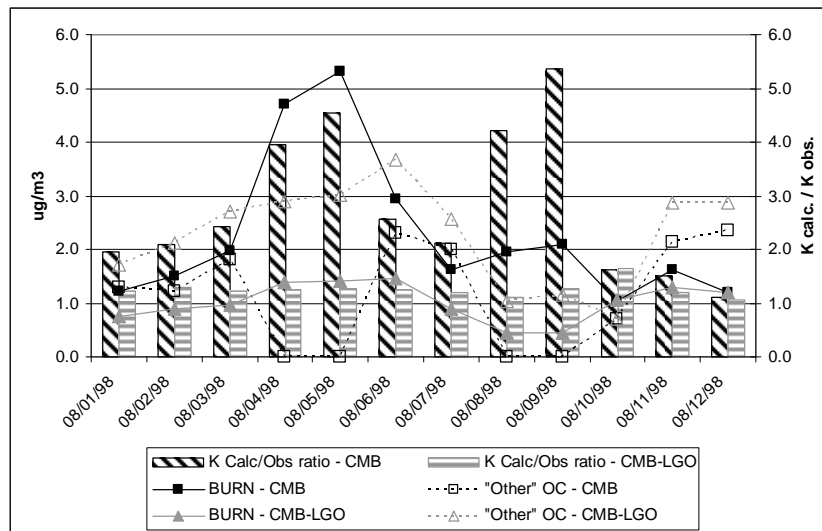


Figure 6-2. Daily $PM_{2.5}$ source impacts of vegetative-burning (BURN) and “other” (unapportioned) OC based on CMB and CMB-LGO, along with the ratio of calculated-to-observed K.

Source apportionment of the Atlanta aerosol has also been performed using the PMF model (Kim et al., 2004). Comparing these results to those based on CMB-LGO (Marmur et al., 2005) for the period of November 1998 - August 2000, average source contributions are in good agreement (Table 6-1). Some differences arise regarding which specific categories are identified such as coal-fired power plants, identified in CMB-LGO only, and cement, railroad, bus/highway and metals processing factors, identified in PMF only. One of the major differences between PMF and CMB-LGO is the treatment of secondary organic aerosol (SOA). In PMF, no specific factor was identified as SOA; rather, SOA mass is distributed between several categories, primarily the secondary sulfate 2 factor as well as other factors (Kim et al., 2004). In CMB-LGO, SOA mass is apportioned to the “other” (unapportioned) organic material category, which accounts for nearly 20% of the $PM_{2.5}$ mass. Comparing the correlations between daily source impacts based on these two methods (Table 6-2), good agreement is observed for the diesel, soil dust, wood burning, ammonium-sulfate/secondary-sulf-1 and ammonium-

nitrate/secondary-nitrate source-categories/factors. Relatively poor agreement is observed for the gasoline vehicle contributions ($R=0.50$). In fact, gasoline vehicle $PM_{2.5}$ based on PMF is relatively highly correlated with CMB-LGO “other” OC ($R=0.79$). Scatter plots of CMB-LGO versus PMF source contributions (Figure 6-3) show numerous cases when the PMF based diesel contribution is zero, as well as a few similar cases for both the PMF and CMB-LGO gasoline contributions. Such results are unlikely to be the case in the center of a large metropolitan city such as Atlanta. Also observed are several outliers in the CMB-LGO wood/vegetative burning contributions. These represent samples collected on the days surrounding July 4th of 1999 and 2000, and contained high concentrations of potassium due to fireworks. As no “fireworks” source was included in the analysis, CMB-LGO apportions much of the potassium into the vegetative-burning category, resulting in high $PM_{2.5}$ contributions due to the associated OC in the source profile for this category. In the PMF analysis, potassium content was adjusted to exclude the contribution of fireworks, resulting in negligible wood burning contributions. Also evident (Figure 6-3) is the higher correlation between CMB-LGO “other” OC and PMF-gasoline, compared to the inter-method gasoline contributions.

Table 6-1. Average source contributions to PM_{2.5} in Atlanta (Jefferson St.) based on CMB-LGO and PMF for the period of November 1998 – August 2000

Source category	% of PM _{2.5} , CMB-LGO	% of PM _{2.5} , PMF
Gasoline vehicles	7.6	5.8
Diesel vehicles	9.1	9.3
Soil dust	1.9	2.7
Vegetative/wood burning	5.8	5.8
Coal-fired power plants (CMB-LGO)	0.8	-
Cement (PMF)	-	1.8
Ammonium-sulfate (CMB-LGO) / Secondary sulfate 1 (PMF)	38.5	46.3
Ammonium-bisulfate (CMB-LGO) / Secondary sulfate 2 (PMF)	2.6	5.5
Ammonium-nitrate (CMB-LGO) / Secondary nitrate 1 (PMF)	8.6	7.7
Railroad (PMF)	-	2.3
Bus/Highway (PMF)	-	1.6
Metal processing (PMF)	-	2.9
"Other" organic material (CMB-LGO)	19.8	-
Unspecified mass	5.2	8.2

Table 6-2. Correlation matrix (R) of daily source contributions to PM_{2.5} in Atlanta (Jefferson St.) based on CMB-LGO (columns) and PMF (rows) for the period of November 1998 – August 2000

	Gasoline	Diesel	Soil	Burn.	CFPP	AmSulf	AmBslf	AmNit	OtherOC
Gasoline	0.50	0.62	0.12	0.54	0.05	0.03	-0.02	0.21	0.79
Diesel	0.64	0.78	0.15	0.50	0.22	0.14	0.02	0.17	0.68
Soil	0.19	0.15	0.97	0.22	0.19	0.08	-0.05	-0.14	0.10
Wood	0.31	0.25	0.03	0.78	-0.04	0.05	0.05	0.16	0.39
Cement	0.43	0.19	0.18	0.10	0.51	0.20	-0.06	0.00	0.05
Sec.Sulf1	0.02	0.17	0.22	0.16	0.24	0.95	0.21	-0.11	0.30
Sec.Sulf2	-0.05	0.11	0.08	0.02	0.17	0.42	-0.01	-0.07	0.34
Sec. Nitr	0.26	0.09	-0.27	0.17	-0.01	-0.15	0.22	0.82	0.01
Railroad	-0.12	0.16	0.04	-0.38	0.02	0.09	-0.13	-0.22	0.00
Bus/Hwy	0.44	0.33	0.13	0.26	0.06	0.05	-0.01	0.08	0.20
Metal prc	0.67	0.31	0.10	0.27	0.13	0.06	-0.01	0.21	0.31

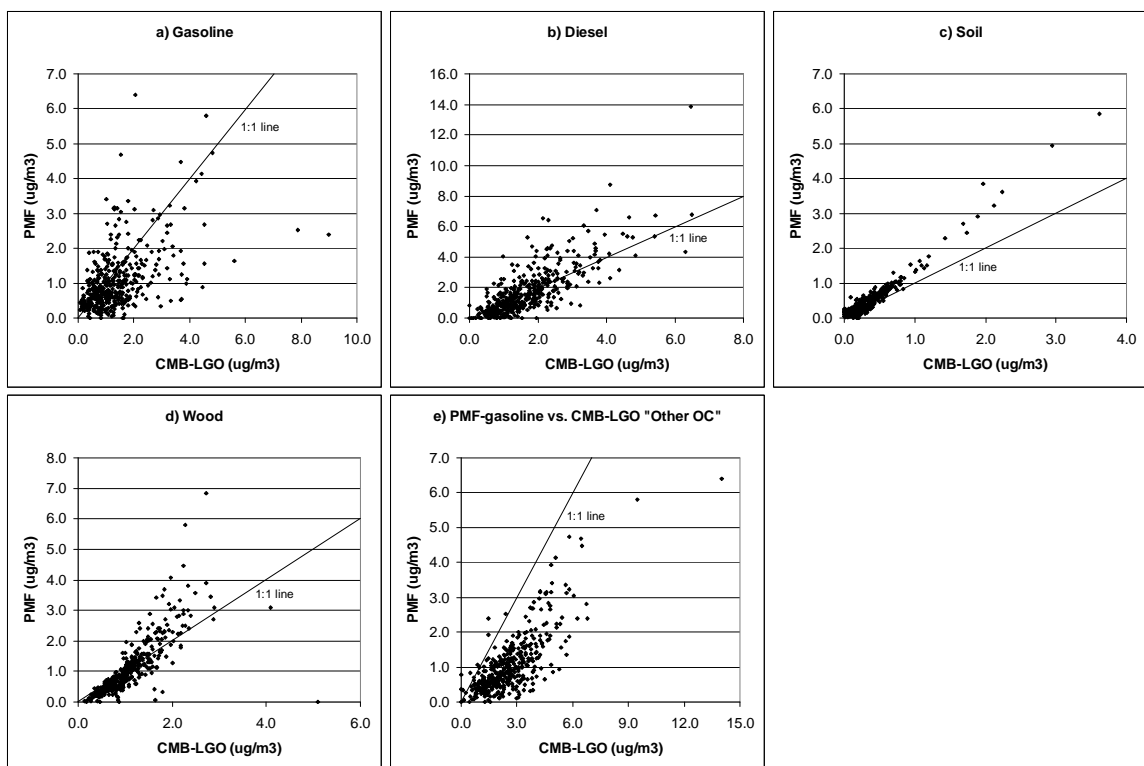


Figure 6-3. CMB-LGO vs. PMF source contributions of gasoline vehicles (a; $R=0.50$), diesel vehicles (b; $R=0.78$), soil dust (c; $R=0.97$), wood/vegetative burning (d; $R=0.78$, $R=0.87$ when July 4th fireworks related samples are excluded) and CMB-LGO “other” OC/PMF gasoline (e; $R=0.79$) for the period of November 1998 – August 2000

To better understand what the “other” (unapportioned) OC category represents, we analyzed the weekly and seasonal patterns in “other” OC contributions, as well as in other carbon-rich categories (gasoline, diesel, wood; Figure 6-4). “Other” OC contributions exhibit a fairly flat weekly pattern, in contrast to the strong weekday/weekend pattern exhibited for mobile sources. In addition, “other” OC contributions peak in summertime (quarter 3 in Figure 6-4b), while the other categories peak in winter, likely due to reduced atmospheric mixing and increased emissions (wood combustion and cold-start emissions from mobile sources). These trends support the assumption that “other” OC is indicative of SOA concentrations. Similarly, a sensitivity analysis for CMB-LGO (Marmur et al., 2006) indicated that “other” OC is strongly

sensitive to OC levels, and weakly sensitive to potassium and EC, which strongly influence the wood burning and diesel contributions respectively. An equivalent analysis for gasoline, diesel and wood burning contributions based on PMF indicated similar seasonal patterns, but the weekly patterns were not as typical (Figure 6-5).

In addition to the receptor-based source apportionment methods (CMB, PMF), an emissions-based approach to source-apportionment was also considered. However, analyzing source apportionment results from an emissions-based air quality model (Models-3/CMAQ) and comparing those to ambient levels of tracer species and to receptor based source-apportionment results (Marmur et al., 2006) found that the air-quality model is driven in large by regional meteorological trends and likely underestimates the temporal variability in source impacts, and may therefore be less useful in discerning the relative impacts of sources on health outcomes in time-series studies of acute outcomes.

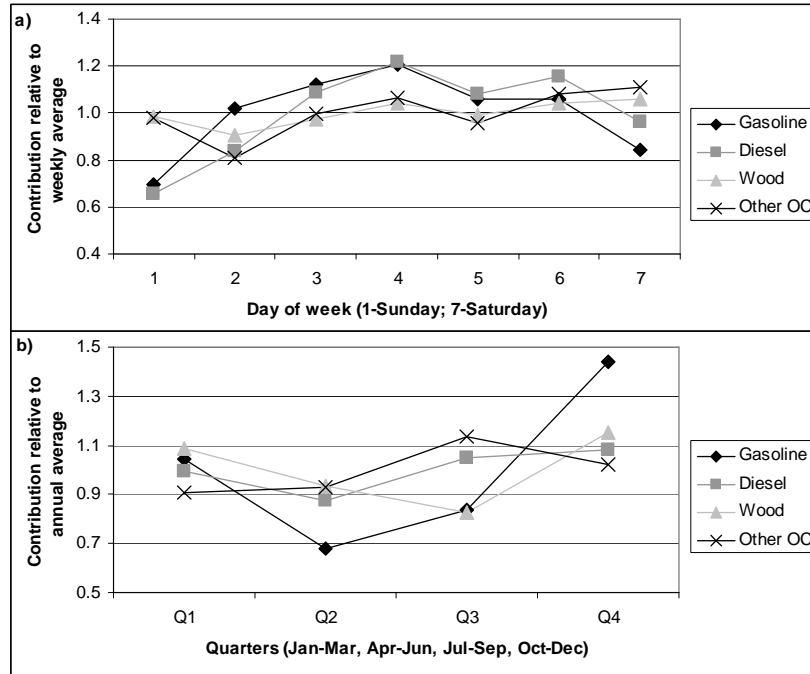


Figure 6-4. Weekly (a) and seasonal (b) patterns in contributions of gasoline vehicles, diesel vehicles, wood burning and “other” OC to PM_{2.5} levels in Atlanta (Jefferson St.) based on CMB-LGO for the period of November 1998 – August 2000

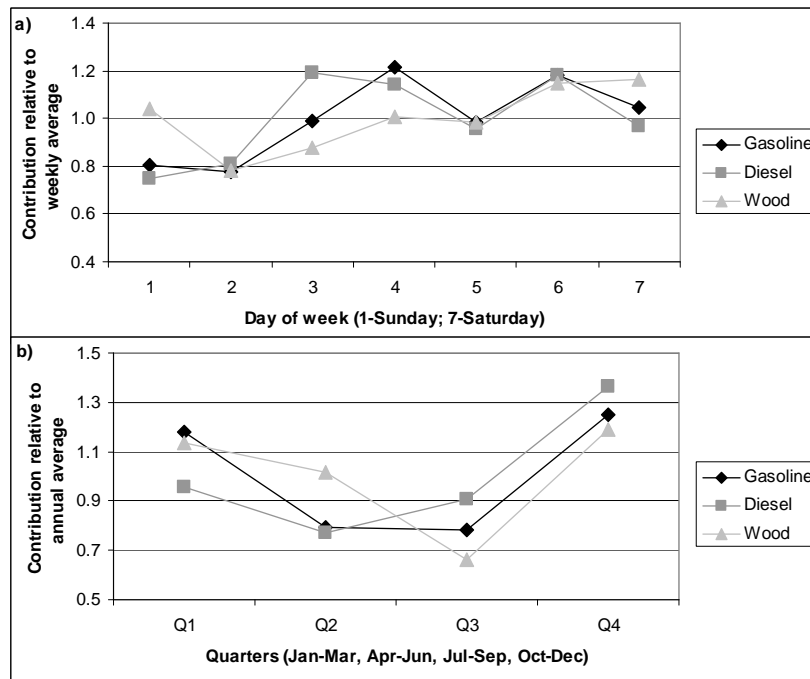


Figure 6-5. Weekly (a) and seasonal (b) patterns in contributions of gasoline vehicles, diesel vehicles and wood burning to PM_{2.5} levels in Atlanta (Jefferson St.) based on PMF for the period of November 1998 – August 2000

6.2.2 Epidemiologic time-series analysis of source-apportioned PM_{2.5} and daily emergency-department visits

Based on the findings regarding the temporal trends in source impacts from the various methods presented (CMB, CMB-LGO, PMF, Models-3/CMAQ), epidemiologic analysis was conducted by collaborators at the Rollins School of Public-Health at Emory University, using source-apportionment results from CMB-LGO and PMF, as well as ambient levels of source-indicative tracer species (Fe, Zn and EC as possible tracers for mobile sources, Si and coarse-PM for soil-dust; K for wood burning; Se for power plants; Ca for cement) (Sarnat et al., 2006). Relative risks (RRs) associated with 3-day moving averages of PM_{2.5} levels were estimated using Poisson generalized linear models (Sarnat et al., 2006), comparable to those used in previous analyses of Atlanta data (Metzger et al., 2004b; Peel et al., 2005). Respiratory outcomes included were asthma, chronic obstructive pulmonary disease (COPD), pneumonia, upper respiratory infection (URI) and the sum of these outcomes (“all respiratory”). Cardiovascular outcomes included were cardiac arrest (CA), congestive heart failure (CHF), dysrhythmia (DYS), ischemic heart disease (IHD), myocardial infarction (MI), peripheral vascular and cerebrovascular disease (PERI), the sum of these categories (CVD) and the sum of all circulatory diseases (CIRC) between codes 390-459 in the International Classification of Disease (ICD) codes (i.e., the CVD chapter in the ICD codes). The epidemiologic analysis was conducted for the period of 11/19/1998-8/31/2000, with 503 days included (149 of the 652 days had incomplete data).

6.3 Results

Source-specific risk-ratios (RR) were computed for gasoline vehicles, diesel vehicles, a combined mobile-source impact, wood/vegetative burning, soil dust, ammonium-sulfate (secondary sulfate 1,2) and ammonium-nitrate (secondary nitrate) based on both CMB-LGO and PMF, for various respiratory and cardiovascular health outcomes. Also computed were RR for coal-fired power plants and other/unapportioned OC (CMB-LGO only), and cement, railroad, bus/highway and metals processing factors (PMF only). RR for EC, OC, Zn, Fe, Si, PM_{2.5}, PM_{coarse}, K, SO₄⁻², NO₃⁻, Se, CO and SO₂ were computed as well, to allow a comparison to the source-specific RR. Since the emphasis here is on source-specific health-outcomes, risk-ratios and associations are sorted and presented by the various PM source categories (e.g., soil PM_{2.5} was found to be significantly associated with asthma). However, a causal relationship, if such exists, would be in the opposite direction (e.g., short-term increase in incidence of asthma could be partially explained by elevated levels of soil PM_{2.5}).

Diesel and total mobile-source (sum of gasoline and diesel) related PM_{2.5} was found to be significantly associated with the sum of all CVD (Figure 6-6), with an RR of 1.03 per change in one inter-quartile range (IQR). RR for specific outcomes such as CA, DYS, IHD, and MI were usually found to be insignificant, possibly due to the increased uncertainties from the smaller sample pools. Relatively strong associations were found between both gasoline and diesel PM_{2.5} and CHF, with RR in the range of 1.08-1.14 (Figure 6-6). PMF derived gasoline PM_{2.5} was found significantly associated with the sum of all respiratory diseases (RR=1.02), and both PMF-gasoline and PMF-diesel

related $PM_{2.5}$ were found significantly associated with URI (Figure 6-7). Significant associations with pneumonia, asthma and COPD were not found. In general, RR for respiratory outcomes based on the PMF results were higher than those based on CMB-LGO results. When comparing mobile source related RR to those based on possible tracers such as Zn, Fe and EC, as well as total $PM_{2.5}$ levels, better agreement was found with Fe and EC, than Zn in CVD RR (Figure 6-6). No significant difference was found between RR based on total $PM_{2.5}$ and that based on mobile source related $PM_{2.5}$ for both cardiovascular and respiratory outcomes, with the exception of URI (Figure 6-7). CO was generally more significantly associated with respiratory health outcomes than cardiovascular outcomes.

Wood/vegetative burning related $PM_{2.5}$ was found to be significantly associated with several respiratory outcomes, such as COPD, URI and pneumonia. RR for total respiratory ED visits were in the range of 1.03-1.05 (Figure 6-8). Significant associations with cardiovascular outcomes were not found (Figure 6-8). An opposite trend was observed for OC and total $PM_{2.5}$. Associations with potassium followed the same trends as the apportioned wood-burning $PM_{2.5}$, but potassium-based RR were typically lower than those based on CMB-LGO and PMF. RR based on CMB-LGO results were typically higher than those based on PMF.

Soil related $PM_{2.5}$ was found to be significantly associated with asthma, in contrast to total $PM_{2.5}$, and CMB-LGO derived soil $PM_{2.5}$ was also significantly associated with the sum of all respiratory outcomes (Figure 6-9). No significant association was found with any of the cardiovascular ED visits (Figure 6-9). Associations with Si in the $PM_{2.5}$ size range were fairly similar, while no significant associations, both

respiratory and cardiovascular, were found with coarse PM ($PM_{2.5-10}$). RR based on CMB-LGO results were typically higher than those based on PMF.

No significant associations, for both respiratory and cardiovascular, were found with coal-burning related $PM_{2.5}$ (CMB-LGO) and the cement, bus/highway, railroad and metals-processing factors (PMF), except for CHF with both cement and bus/highway. However, primary $PM_{2.5}$ contributions from point sources, such as coal-fired power plants, have been shown to vary spatially (Marmur et al., 2006), which could potentially attenuate a true association, if such exists.

Associations between both cardiovascular and respiratory ED visits and sulfate related $PM_{2.5}$ were insignificant (Figure 6-10), in contrast to those with total $PM_{2.5}$. This included both the CMB-LGO AmSulf contributions, and the PMF secondary sulfate 1 and 2 factors, and is consistent with previous findings from the Atlanta epidemiologic study (Metzger et al., 2004b; Peel et al., 2005). Similar patterns were also observed for nitrate-related $PM_{2.5}$ from both CMB-LGO and PMF, but ambient nitrate was found significantly associated with asthma/wheeze and the sum of all respiratory outcomes (Figure 6-11).

Strong significant associations were found between “Other”/unapportioned OC from CMB-LGO and several respiratory outcomes (asthma/wheeze, URI, pneumonia and sum of all respiratory ED visits; RR between 1.04-1.12) (Figure 6-12), in contrast to the associations with total OC and with $PM_{2.5}$ and these respiratory outcomes, though further analyses, such as based on longer datasets, are needed to support this finding. Associations between cardiovascular ED visits and “other” OC were also significant, but similar to those with total OC and with $PM_{2.5}$ (Figure 6-12).

No significant associations were found between any of these PM_{2.5} source-categories and two control outcomes (appendicitis, finger wounds) (Figure 6-13).

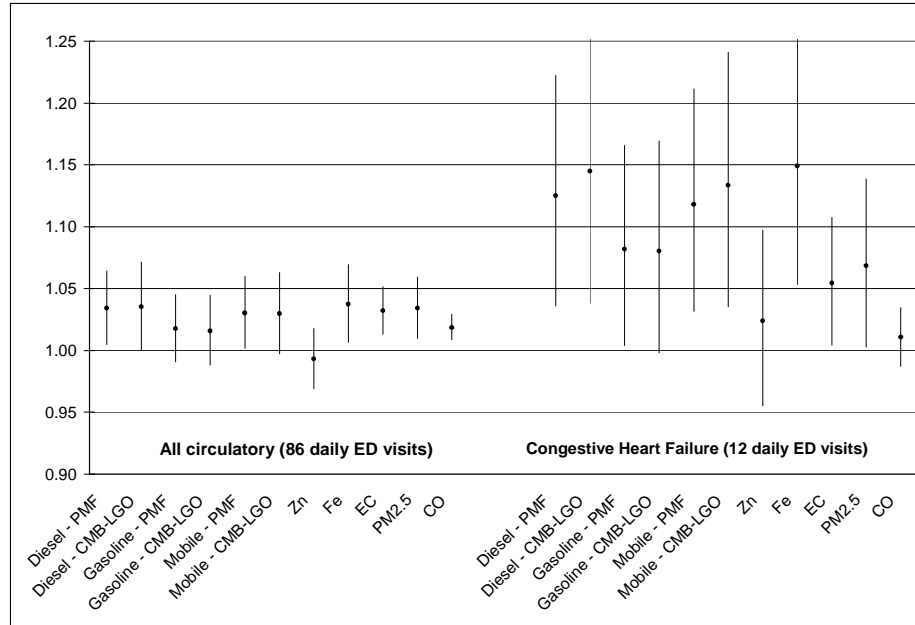


Figure 6-6. Preliminary cardiovascular related risk-ratios (95% confidence interval [CI]) for gasoline, diesel and total mobile-source related $PM_{2.5}$ per change of one inter-quartile range (IQR). Also shown are the RR for Zn, Fe, EC, total $PM_{2.5}$ and CO. Associations in which the error bar does not encompass $RR=1.00$ are considered significant.

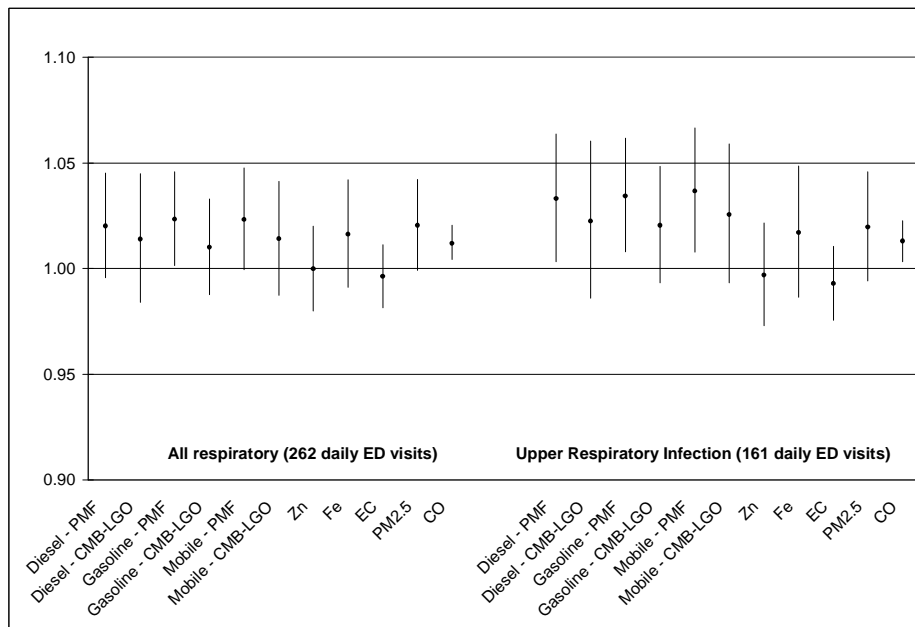


Figure 6-7. Preliminary respiratory related risk-ratios (95% CI) for gasoline, diesel and total mobile-source related $PM_{2.5}$ per change of one IQR. Also shown are the RR for Zn, Fe, EC, total $PM_{2.5}$ and CO. Associations in which the error bar does not encompass $RR=1.00$ are considered significant.

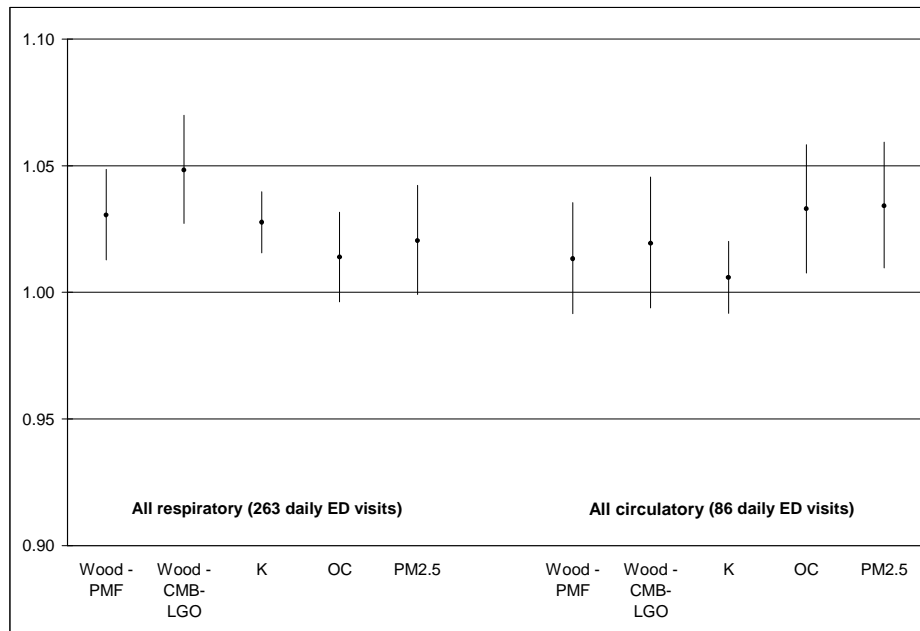


Figure 6-8. Preliminary respiratory and cardiovascular related risk-ratios (95% CI) for wood/vegetative burning related $PM_{2.5}$ per change of one IQR. Also shown are the RR for potassium (K), OC and total $PM_{2.5}$. Associations in which the error bar does not encompass $RR=1.00$ are considered significant.

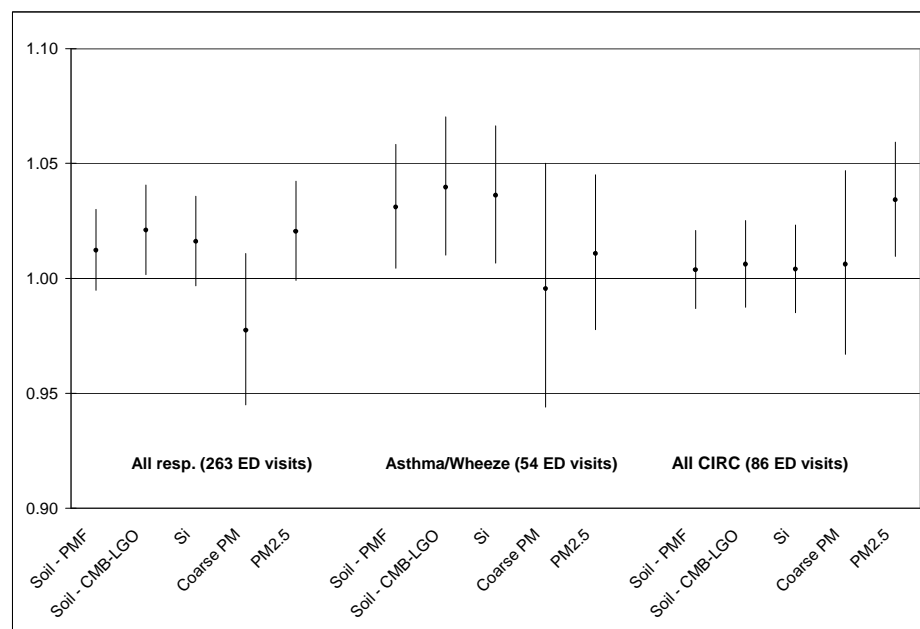


Figure 6-9. Preliminary respiratory and cardiovascular related risk-ratios (95% CI) for soil $PM_{2.5}$ per change of one IQR. Also shown are the equivalent RR for Si, coarse PM and $PM_{2.5}$. Associations in which the error bar does not encompass $RR=1.00$ are considered significant.

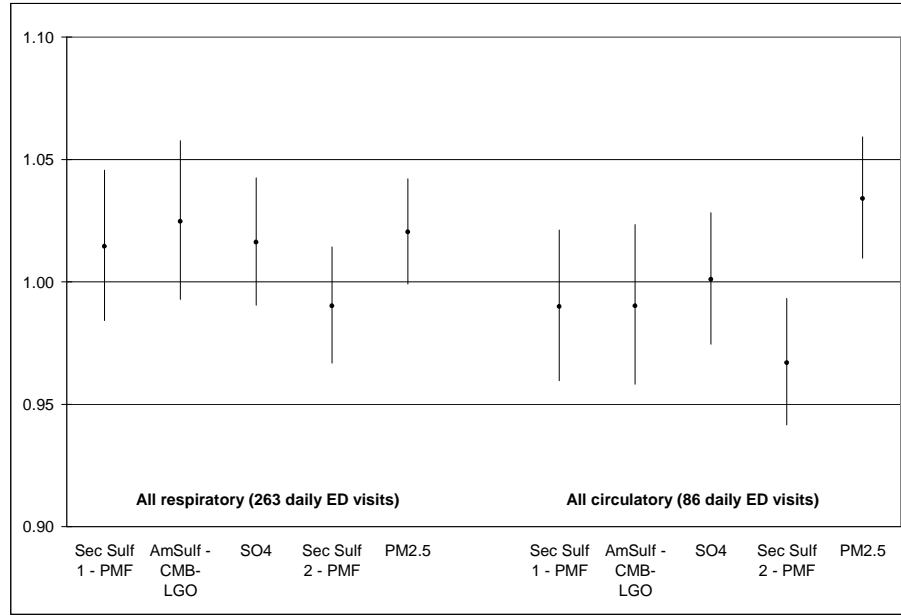


Figure 6-10. Preliminary respiratory and cardiovascular related risk-ratios (95% CI) for sulfate-related and total $PM_{2.5}$ per change of one IQR. Associations in which the error bar does not encompass $RR=1.00$ are considered significant.

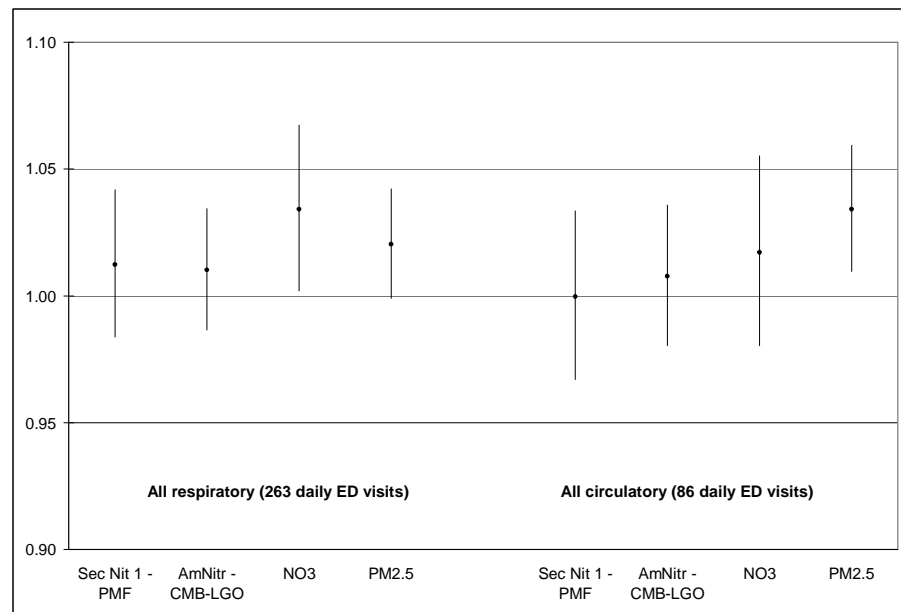


Figure 6-11. Preliminary respiratory and cardiovascular related risk-ratios (95% CI) for nitrate-related and total $PM_{2.5}$ per change of one IQR. Associations in which the error bar does not encompass $RR=1.00$ are considered significant.

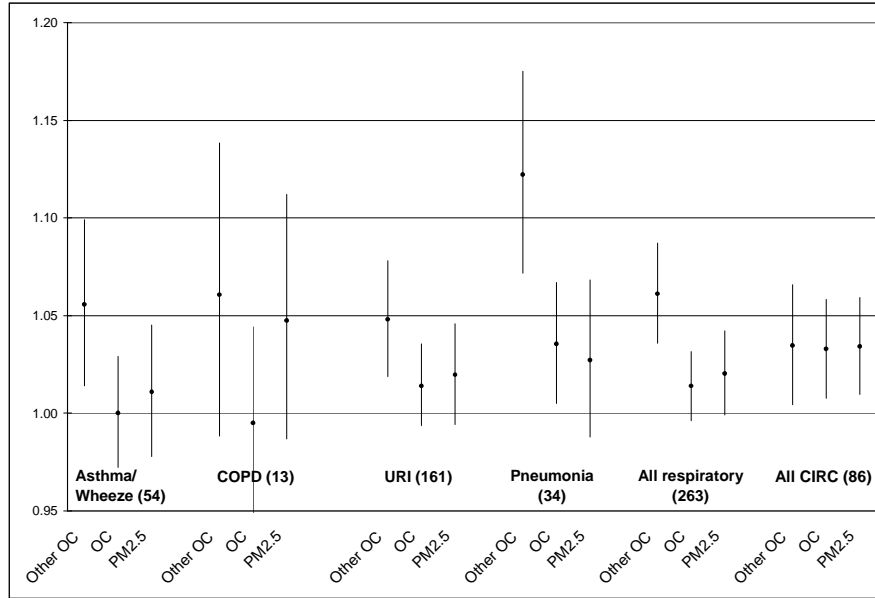


Figure 6-12. Preliminary respiratory (asthma, COPD, URI, pneumonia, all respiratory) and cardiovascular related risk-ratios (95% CI) for CMB-LGO “other”/unapportioned OC, total OC and total PM_{2.5} per change of one IQR. Associations in which the error bar does not encompass RR=1.00 are considered significant.

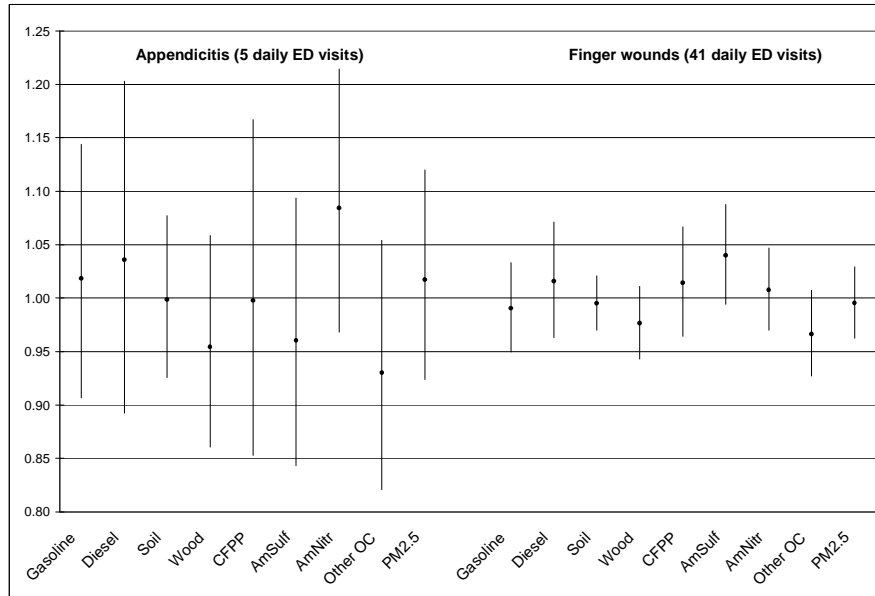


Figure 6-13. Control outcome (appendicitis, finger wounds) related risk-ratios (95% CI) for various CMB-LGO source categories and total PM_{2.5} per change of one IQR. None of the associations are significant.

6.4 Discussion

Several significant associations were found between health-outcomes and source-specific $PM_{2.5}$ in this preliminary analysis. Mobile-source related $PM_{2.5}$ was found significantly associated with cardiovascular outcomes, though the RR were fairly similar to those based on total $PM_{2.5}$ and tracers such as EC and Fe. Associations with Zn, a species which has been shown to strongly affect gasoline-vehicles contributions in CMB-LGO (Marmur et al., 2006), were largely insignificant. Wood burning $PM_{2.5}$ was found significantly associated with respiratory outcomes, with higher RR compared to K, OC (insignificant association) and total $PM_{2.5}$ (barely significant association). Soil dust and Si were found significantly associated with asthma, while the association between total $PM_{2.5}$ and asthma was not significant. “Other”/unapportioned OC was found significantly associated with various respiratory outcomes, with fairly high RR (1.04-1.12), while the equivalent associations with total OC and with $PM_{2.5}$ were insignificant, though further analyses, such as based on longer datasets, are needed to support this finding.

Comparing the RR estimates based on CMB-LGO and PMF (Figure 6-14), a fairly high correlation is observed across all source categories ($R=0.84$), except for the nitrate related RR (surprising considering the good inter-method correlation between the nitrate source contributions; see Table 6-2). No single method consistently provides more significant associations or higher RR. This is an indication that collinearity-generated “noise” in CMB-LGO was minimal, likely not substantially higher than in PMF, in which orthogonality between factors is maximized. The overall high correlation between RR based on CMB-LGO and PMF may be interpreted as an indication that despite methodological differences and uncertainties, source-apportionment methods may be

useful in discerning source-specific health effects. Similar findings have been reported by (Thurston et al., 2005), however their analysis was limited to factor analytical techniques (PCA, PMF, UNMIX). To our knowledge, this is the first time that a CMB model (CMB-LGO, in this case) has been applied in a health study, and the findings presented here suggest that despite larger methodological differences between PMF and CMB-LGO compared to the methods used in (Thurston et al., 2005), the conclusion regarding the usefulness of applying source-apportionment methods in health studies holds.

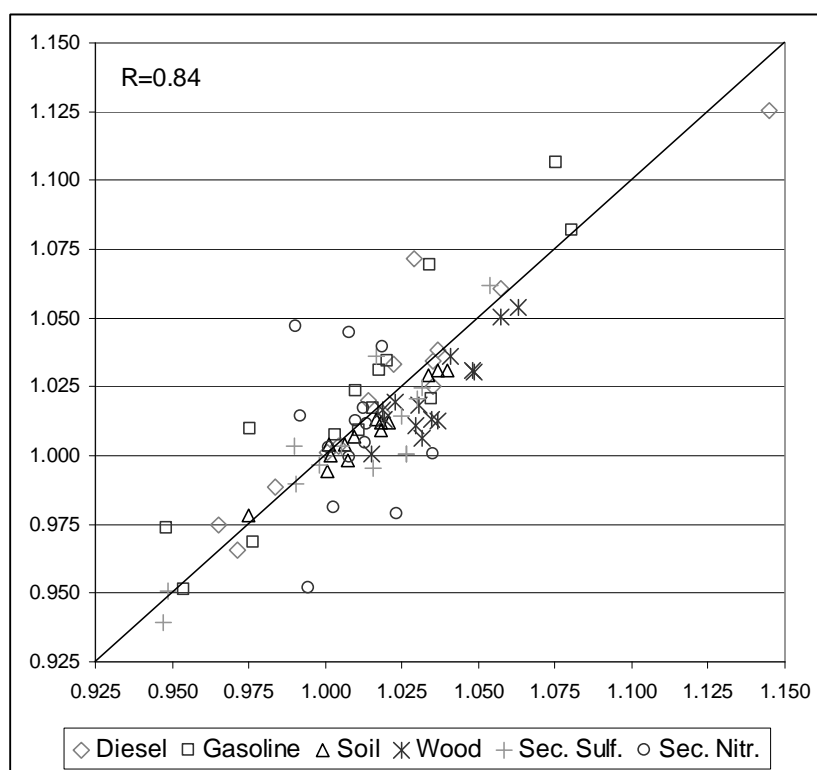


Figure 6-14. Scatter plot of preliminary risk-ratios for various respiratory and cardiovascular health-outcomes based on PMF and CMB-LGO source contributions.

An alternative approach to applying source-apportionment methods is to use concentrations of tracer species directly in the health analysis. For example, associations

and risk-ratios for Fe and EC were found fairly similar to those based on mobile-source contributions (Figure 6-6, 6-7), and similarities were also observed between K and wood burning PM_{2.5} (Figure 6-8) and Si and soil PM_{2.5} (Figure 6-9). However, the RR based on Zn, which has been shown to strongly affect gasoline-vehicles contributions in CMB-LGO (Marmur et al., 2006), exhibited a different pattern compared to that of the mobile-sources based RR. This serves as an example that an expected tracer might not always yield similar associations and RR as its suspected/assumed source. This is likely the result of that tracer not being a unique tracer for a specific source-category (i.e., it is being emitted from various source categories). An analysis of the source-apportionment results based on (Marmur et al., 2005) revealed several days in which (high) Zn concentrations could not be explained by contributions of gasoline-vehicles, likely indicating the presence of another source of Zn on those days (though on average, gasoline vehicles contributed 82% of the ambient Zn concentrations). If each source category emitted a tracer unique to that category, then there would be no benefit in applying source-apportionment estimates in the epidemiologic analysis, as the temporal variation in the source-estimates would be identical to that of each category's unique tracer. Given that potassium (K) is present in soil dust as well as in wood burning PM_{2.5} and that Fe is emitted from many sources other than mobile sources, one cannot conclude *a priori* that these can be used as surrogates for wood-burning and mobile sources in a health analysis. The validity of such a determination would depend on the dominant sources and source mixture in each air-shed/region studied. In the case of Atlanta based on CMB-LGO results, 92% of K and 89% of Si are from wood-burning and soil dust, respectively. This can explain the good agreement between the species and source-impact

based RR for wood burning and soil PM_{2.5}. On the other hand, mobile sources contribute only 23% of total ambient Fe (41% of apportioned Fe; only 55% of Fe is accounted for by the various source impacts), hence the good agreement between the Fe and mobile-sources based RR is surprising. In addition, Fe did not affect the mobile-source contributions in CMB-LGO, based on a sensitivity analysis (Marmur et al., 2006). In summary, despite generally good agreement between RR estimates based on source-impacts and tracer species, one cannot assume *a priori* what tracer species can be used as surrogates (unless these have been proven to be unique per source-category). Instead, it is recommended to perform a source-apportionment and sensitivity analysis, to identify the key species driving the apportionment process, and apply both tracers and source impacts in the health study.

Another interesting finding of the health study was the strong association between “other” OC and various respiratory outcomes (Figure 6-12). “Other” OC includes both secondary organic aerosol (SOA) and any primary OC not accounted for by the other source categories. While it is impossible to estimate what portion of “other” OC is indeed SOA, weekly and seasonal patterns (Figure 6-4), as well as the results from a sensitivity analysis (Marmur et al., 2006) and the fact that the sources of primary OC included in this analysis comprise most of the OC in the emissions inventory (Park et al., 2006) are all indicative of SOA. If so, this serves as a good example of the usefulness of applying source-estimates in a health studies, as tracers for SOA are still being investigated and measurement techniques developed and refined. Examples of potential tracers are oxidation products of terpenes (pinonic acid, pinic acid and nopinone) and isoprene (methyl vinyl ketone and methacrolein), as tracers of biogenic SOA. However, the

sources of SOA and the relative contributions of anthropogenic vs. biogenic SOA are still being investigated and debated (Sullivan, 2006). These issues are further complicated by the fact that speciation studies often identify only 10-20% of the organic compounds present, because the organic fraction covers a wide range of chemical and thermodynamic properties (Maria, 2003). One could consider applying the EC-tracer approach (Turpin and Huntzicker, 1995) to estimate SOA concentrations, but would then need to assume a constant mixture of primary sources of OC and EC (constant primary OC to EC ratio), an assumption not needed/made in CMB-LGO.

6.5 Summary

Source-contributions based on CMB-LGO and PMF have been applied in an emergency-department health study. Despite methodological differences and uncertainties in the apportionment process, good agreement was observed between the CMB-LGO and PMF based risk ratios, supporting the usefulness of applying apportionment methods in health studies. Preliminary epidemiologic analysis found mobile-sources related $PM_{2.5}$ significantly associated with cardiovascular outcomes; wood burning $PM_{2.5}$ significantly associated with respiratory outcomes; soil dust significantly associated with asthma and “other” OC (indicative of SOA) was found significantly associated with various respiratory outcomes, with high risk-ratios, though further analyses, such as based on longer datasets, are needed to support this finding. Despite generally good agreement between RR estimates based on source-contributions and tracer-species, source-apportionment and sensitivity analyses are needed to determine which are the most suitable tracers for each source-category. Source-apportionment

methods (CMB-LGO in this case) can also serve as a useful tool for estimating SOA concentrations (especially, in lieu of a SOA tracer) and associated health effects.

References

- Hansen, D.A., Edgerton E.S., Hartsell B.E., Jansen, J.J., Hidy, G.M. 2006. Air Quality Measurements for the Aerosol Research and Inhalation Epidemiology Study. Journal of the Air and Waste Management Association (in press).
- Hansen D.A., Edgerton E.S., Hartsell B.E., Jansen J.J., Kandasamy N., Hidy G.M. 2003. The southeastern aerosol research and characterization study: Part 1-overview. Journal of the Air & Waste Management Association 53(12):1460-1471.
- Hauck, H., Berner, A., Frischer, T., Gomiscek, B., Kundi, M., Neuberger, M., 2004. AUPHEP-Austrian Project on Health Effects of Particulates - general overview. Atmospheric Environment 38(24):3905-3915, 2004.
- Heal, M.R., Hibbs, L.R., Agius, R.M. and Beverland, L.J., 2005. Total and water-soluble trace metal content of urban background PM₁₀, PM_{2.5} and black smoke in Edinburgh, UK. Atmospheric Environment, 39(8): 1417-1430.
- Kim, E., Hopke, P.K. and Edgerton, E.S., 2004. Improving source identification of Atlanta aerosol using temperature resolved carbon fractions in positive matrix factorization. Atmospheric Environment, 38(20): 3349-3362.
- Laden, F., Neas, L.M., Dockery, D.W. and Schwartz, J., 2000. Association of fine particulate matter from different sources with daily mortality in six US cities. Environmental Health Perspectives, 108(10): 941-947.
- Mar, T.F., Norris, G.A., Koenig, J.Q. and Larson, T.V., 2000. Associations between air pollution and mortality in Phoenix, 1995-1997. Environmental Health Perspectives, 108(4): 347-353.
- Maria, S.F., 2003. Source signatures of carbon monoxide and organic functional groups in Asian Pacific Regional Aerosol Characterization Experiment submicron aerosol types. Journal of Geophysical Research-Atmospheres, 108(D23).
- Marmur, A., Park, S.K., Mulholland, J.A., Tolbert, P.E. and Russell, A.G., 2006. Source apportionment of PM_{2.5} 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.

- Marmur, A., Unal, A., Mulholland, J.A. and Russell, A.G., 2005. Optimization-based source apportionment of PM_{2.5} incorporating gas-to-particle ratios. *Environmental Science & Technology*, 39(9): 3245-3254.
- Metzger, K.B., Tolbert, P.E., Klein, M., Peel, J.L., Flanders, W.D., Mulholland, J.A., 2004a. Ambient air pollution and arrhythmic events in patients with implanted cardioverter defibrillators, 1993-2002. *Epidemiology*, 15(4): S22-S23.
- Metzger, K.B., Tolbert, P.E., Klein, M., Peel, J.L., Flanders, W.D., Todd, K., 2004b. Ambient air pollution and cardiovascular emergency department visits. *Epidemiology*, 15(1): 46-56.
- Park, S.K. et al., 2006. Source Apportionment of PM_{2.5}: Comparison between Receptor and Air Quality Models. *Environmental Science & Technology* (in press).
- Peel, J.L., Tolbert, P.E., Klein, M., Metzger, K.B., Flanders, W.D., Todd, K., 2005. Ambient air pollution and respiratory emergency department visits. *Epidemiology*, 16(2): 164-174.
- Sarnat, J.A., Marmur, A., Kim, E., Mulholland, J.A., Hopke, P.K., Russell, A.G., Tolbert, P.E., 2006. Examining the Cardiovascular Health Effects of Atlanta Aerosol using Three Source Apportionment Techniques, 2006 International Aerosol Conference, St. Paul, Minnesota USA.
- Solomon, P.A., 2003. Overview of the 1999 Atlanta Supersite Project. *Journal of Geophysical Research-Atmospheres*, 108(D7):8428.
- Sullivan, A.M., 2006. The ambient organic aerosol soluble in water: measurements, chemical characterization, and an investigation of sources. Ph.D. dissertation, School of Earth & Atmospheric Sciences, Georgia Institute of Technology
- Thurston, G.D., Ito, K., Mar, T., Christensen, W.F., Eatough, D.J., Henry, R.C., Kim, E., Laden, F., Lall, R., Larson, T.V., Liu, H., Neas, L., Pinto, J., Stölzel, M., Suh, H., Hopke, P.K., 2005. Workgroup report: Workshop on source apportionment of particulate matter health effects - Intercomparison of results and implications. *Environmental Health Perspectives*, 113(12): 1768-1774.
- Turpin, B.J. and Huntzicker, J.J., 1995. Identification of Secondary Organic Aerosol Episodes and Quantification of Primary and Secondary Organic Aerosol Concentrations During SCAQS. *Atmospheric Environment*, 29(23): 3527-3544.

CHAPTER 7

CONCLUSIONS AND FUTURE RESEARCH

Conclusions

Particulate matter, especially particles smaller than 2.5 microns in diameter ($PM_{2.5}$), has been associated with adverse health effects and mortality in studies covering more than 150 cities. Studies of the effects of air pollution on human health are typically conducted using ambient measurements to represent the air quality over cities or regions. However, the use of ambient data for such studies introduces several limitations such as spatial representativeness of the monitoring site, analytical uncertainties, and incompleteness and lack of continuity in data. The complex chemical composition of $PM_{2.5}$ and associated analytical uncertainties pose a further challenge when trying to investigate species specific health effects. A complementary approach is to examine associations between health outcomes and sources contributing to ambient $PM_{2.5}$, which can provide regulators with important information to tighten controls on sources more prone to causing health effects.

Temporal and spatial variability in measured and simulated $PM_{2.5}$ constituents in Atlanta, GA, and implications for time-series health studies

Results from a long term air quality simulation were analyzed in terms of the model's ability to simulate temporal and spatial variability in concentrations of both

secondary and primary PM_{2.5} components in Atlanta, GA. Seasonal variations in sulfate and nitrate concentrations were well captured by the model, but the model's ability to capture shorter-term (e.g., daily) variations, typically of interest in time-series health studies of acute outcomes, was limited. Moreover, the spatial homogeneity in ambient concentrations of secondary PM_{2.5} constituents (such as sulfate and nitrate), suggests limited benefit in applying simulated concentration fields for these species in a time-series health study when ambient measurements are available. Concentrations of primary PM_{2.5} constituents (such as elemental carbon), on the other hand, have much greater spatial variability, and short-term variability in these species is better captured by the air quality model.

Optimization-based source apportionment of PM_{2.5} incorporating gas-to-particle ratios

A modified approach to PM_{2.5} source apportionment was developed, using source indicative SO₂/PM_{2.5}, CO/PM_{2.5} and NO_x/PM_{2.5} ratios as constraints, in addition to the commonly used particulate-phase source profiles. Additional information from using gas-to-particle ratios assists in reducing collinearity between source profiles, a problem that often limits the source-identification capabilities and accuracy of traditional receptor models. The set of equations for the PM_{2.5} Chemical Mass Balance (CMB) receptor model were solved using a global-optimization program, Lipschitz Global Optimizer (LGO), subject to constraints on ambient gas-phase concentrations. Application of the CMB-LGO model to a 25 month dataset of daily PM_{2.5} measurements (total mass and composition) at the Atlanta Jefferson Street SEARCH site yielded source-contributions

that seem more indicative of the named sources compared to particulate-phase source apportionment methods, based on correlations of the source impacts and tracer species. Furthermore, collinearity between source-categories (e.g., soil-dust and primary $PM_{2.5}$ from coal-burning; wood burning and “other” OC) was reduced in CMB-LGO, based on higher source inter-correlations than in the “regular” CMB model.

Optimized variable source-profile approach for source apportionment

A further expansion of the CMB-LGO approach for $PM_{2.5}$ source apportionment was developed in which both the local source compositions and corresponding contributions were determined from ambient measurements and initial estimates of source compositions. Such an approach can serve as an alternative to using predetermined (measured) source profiles, as traditionally used in CMB applications, which are not always representative of the region and/or time period of interest. This technique was applied to a dataset of $PM_{2.5}$ measurements at the former Atlanta supersite (Jefferson Street site), to apportion $PM_{2.5}$ mass into nine source categories. Good agreement is found when these source impacts are compared with those derived based on measured source profiles as well as those derived using a factor analytical technique.

Source apportionment of $PM_{2.5}$ in the Southeastern United States using receptor and emissions-based models: conceptual differences and implications for time-series health studies

Two conceptually different approaches to source apportionment were compared: a receptor model and an emissions-based air-quality model. The receptor model captured

more of the temporal variation in source impacts at a specific receptor site compared to the emissions-based model. Driven by data at a single site, receptor models may have some shortcomings with respect to spatial representativeness (unless a reduced study area is used or data from multiple sites are available), likely attenuating the observed association in a health study. Source apportionment results from emissions-based models, such as CMAQ, may be more spatially representative as they represent an average grid-cell value. Limitations in the ability to model daily fluctuations in emissions, however, lead to results being driven mainly by regional meteorological trends, likely underestimating the true daily variations in local source impacts. These effects will likely introduce an attenuation of observed association in a health study and limit the model's usefulness in discerning the relative impacts of the sources on health outcomes.

Intermethod variability in associations between source-apportioned PM_{2.5} and daily emergency-department visits in Atlanta, GA

Results from a preliminary source-specific PM_{2.5} epidemiologic analysis were presented and analyzed for inter-method variability in risk-ratio estimates based on source apportionment results from the Positive Matrix Factorization (PMF) and CMB-LGO models, as well as the direct application of tracer species in the epidemiologic study. Despite methodological differences and uncertainties in the apportionment process, good agreement was observed between the CMB-LGO and PMF based risk ratios, indicating the usefulness of applying apportionment methods in health studies. Preliminary epidemiologic analysis found mobile-source related PM_{2.5} significantly associated with cardiovascular outcomes; wood burning PM_{2.5} significantly associated

with respiratory outcomes; soil dust significantly associated with asthma. “Other” OC was found significantly associated with various respiratory outcomes, with high risk-ratios (RR), though further analyses, such as based on longer datasets, are needed to support this finding. Temporal patterns in source impacts suggest that “other” OC is associated with secondary formation. Despite generally good agreement between RR estimates based on source-contributions and tracer-species, source-apportionment and sensitivity analyses are needed to determine the most suitable tracers for each source-category.

Future Research

Application of CMB-LGO other regions/periods

The advantages of incorporating gas-phase data in PM_{2.5} source apportionment, as done in CMB-LGO, were demonstrated throughout this thesis. However, such information is typically not included in source-apportionment studies. Moreover, ambient measurements of CO and SO₂ are limited in number, as very few areas throughout the USA are in non-attainment of the National Ambient Air Quality Standards (NAAQS) for these pollutants. However, these pollutants are useful as tracers, even if their significance from a direct regulatory standpoint has declined in recent decades. The findings presented in this dissertation should promote the use of gaseous data in source-apportionment and encourage states and other regulatory agencies to monitor CO and SO₂ as part of their observational network for understanding the sources of PM_{2.5} in their region. Applying CMB-LGO for other periods and regions will allow expanding and conducting more

epidemiologic studies to broaden our understanding of the associations between PM_{2.5} sources and health.

Application of CMB-LGO to organic tracer measurements

The use of organic tracers for PM_{2.5} source-apportionment is an emerging-field, and organic-tracer based source-apportionment methods are becoming a more common practice for apportioning OC mass between the contributing sources based on detailed organic source “fingerprints” (Lee et al. 2005; Schauer and Cass 2000; Schauer et al. 2001; Schauer et al. 2002; Schauer et al. 2002; Zheng et al. 2002; Zheng et al. 2006). The ability to detect dozens of organic species enhances the source-identification capabilities compared to inorganic-based source-apportionment. However, as with inorganic-based source-apportionment studies, organic-tracer based CMB analyses are often limited by collinearity issues and uncertainties in representativeness of source-profiles, as well as detection-level issues in ambient measurements. CMB-LGO can be applied to organic-tracer measurements in combination with gaseous measurements, to reduce collinearity issues in the same manner as demonstrated for inorganic measurements. For example, mobile-source markers such as hopanes and steranes are present in both gasoline and diesel vehicles, and splitting the contributions of these two categories is a challenge even with organic-tracers. As with inorganic tracers, CO and NO_x data can be helpful in reducing collinearity between these two sources. In addition, CMB-LGO can be applied to derive estimates of source-compositions and help in identifying suitable literature based source-profiles to be applied.

Enhanced emissions-based source-apportionment

Comparison of emission and receptor based source-apportionment results found relatively poor agreement in the temporal variation in source-contributions. This is in part due to little temporal variability in emission estimates in the emissions-based approach and to the relatively coarse grid-resolution used (12 km) compared to a point measurement. The purpose of this study is to enhance the source-apportionment capabilities of the Models-3 emissions-based modeling system. Developing and applying a neighborhood-scale (~1 km grid resolution) meteorological model (Dandou et al. 2005; Otte et al. 2004) may enable simulating local flow patterns from sources to the receptor. To improve the estimates of mobile-source contributions, data from “Commute Atlanta” (Guensler 2006; Li et al. 2004; Ogle et al. 2005), which includes the collection and analysis of second-by-second vehicle speed, position, and engine operating data from 470 vehicles in representative Atlanta households, would be incorporated. Information gathered by “Commute Atlanta” provides near real-time information on traffic conditions and will allow to compute more accurate estimates of time-dependant mobile-source emissions, in contrast to the “typical emissions” approach currently used in Models-3 (Mobile-6/SMOKE specifically), which is based on statistical analyses of traffic counts and the generation of “typical” temporal patterns in emissions, without adjusting PM emissions for driving speeds. Actual fire emissions would also be incorporated, based on information on the occurrence of major fire events, instead of using typical emissions, and may improve the source contribution estimates for the biomass-burning category. Soil-dust emissions would be estimated as a function of surface wind-speed and soil-

moisture, not accounted for in the current version of SMOKE, to improve the estimates of soil-dust contributions to PM_{2.5} levels.

The effect of dampened temporal variability on risk-ratio estimates

Evaluating spatial and temporal patterns in measured and simulated concentrations of secondary PM_{2.5} components such as sulfate, nitrate and ammonium indicated that when ambient data are available, there does not seem to be a benefit in using simulated concentrations of these species in a health study. This conclusion was based on the ability of the air-quality model (CMAQ) to capture the short-term variations in species concentrations and on the spatial heterogeneity in ambient concentrations of secondary PM_{2.5} species. One direction for future research would be to examine the effect of reduced temporal variability in species concentrations on robustness of findings from an epidemiologic study. Adjustments would be made to the temporal variability in ambient data to dampen out the variations around the average in different levels of adjustment. Then, the sensitivity of the risk-ratio estimates to the temporal variability would be analyzed. The expected effect of dampened temporal variability is an attenuation of the risk-ratio estimates. However, at some level of adjustment, the risk-ratios are expected to become insignificant. By applying different levels of dampening to temporal variability in ambient data, a “cutoff” value may be identified for when risk-ratios become insignificant (depending on pollutant and outcome). Such a “cutoff” value can be useful in determining whether the use of simulated concentration fields, exhibiting dampened temporal variability compared to the variations in ambient data, are expected to yield robust estimates of significance of risk-ratios.

The effect of analysis-period on risk-ratio estimates

To evaluate the robustness of risk-ratio estimates, a sensitivity analysis of risk-ratios to the analysis-period will be performed. Using the “Jackknife” statistical process, subsets of data would be systematically dropped out one at a time and the resulting variation in the risk-ratio estimated will be assessed. This analysis will enable evaluating whether the risk-ratios obtained were driven by a fairly small number of days within dataset of source-apportioned $PM_{2.5}$ and emergency-department visits, or are representative of recurring patterns in the data.

Closing Remarks

While much work remains, important insights have been derived from this thesis regarding source contributions to $PM_{2.5}$ levels in the southeastern U.S.; performance of various air-quality modeling techniques, both existing and newly developed, in simulating short-term variations in concentrations and source-contributions to $PM_{2.5}$; and the implementation of some of these methods in time-series health studies of the associations between $PM_{2.5}$ and various health endpoints. The findings of this thesis along with the above recommendations open numerous avenues for potentially fruitful future investigation.

References

Dandou A, Tombrou M, Akylas E, Soulakellis N, Bossioli E. 2005. Development and evaluation of an urban parameterization scheme in the Penn State/NCAR Mesoscale Model (MM5). *Journal of Geophysical Research-Atmospheres* 110(D10).

- Guensler RL. 2006. Faculty web page
(www.ce.gatech.edu/fac_staff/research_bio.php?active_id=rg49), accessed July 21, 2006.
- Lee S, Baumann K, Schauer JJ, Sheesley RJ, Naeher LP, Meinardi S, et al. 2005. Gaseous and particulate emissions from prescribed burning in Georgia. *Environmental Science & Technology* 39(23):9049-9056.
- Li H, Guensler R, Ogle J, Wang J. 2004. Using global positioning system data to understand day-to-day dynamics of morning commute behavior, *Transportation Planning and Analysis*, 78-84.
- Ogle J, Guensler R, Elango V. 2005. Georgia's commute Atlanta value pricing program: 28 Recruitment methods and travel diary response rates, *Transportation Planning And Analysis*, 28-37.
- Otte TL, Lacser A, Dupont S, Ching JKS. 2004. Implementation of an urban canopy parameterization in a mesoscale meteorological model. *Journal of Applied Meteorology* 43(11):1648-1665.
- Schauer JJ, Cass GR. 2000. Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds as tracers. *Environmental Science & Technology* 34(9):1821-1832.
- Schauer JJ, Kleeman MJ, Cass GR, Simoneit BRT. 2001. Measurement of emissions from air pollution sources. 3. C-1-C-29 organic compounds from fireplace combustion of wood. *Environmental Science & Technology* 35(9):1716-1728.
- Schauer JJ. 2002. Measurement of emissions from air pollution sources. 4. C-1-C-27 organic compounds from cooking with seed oils. *Environmental Science & Technology* 36(4):567-575.
- Schauer JJ. 2002. Measurement of emissions from air pollution sources. 5. C-1-C-32 organic compounds from gasoline-powered motor vehicles. *Environmental Science & Technology* 36(6):1169-1180.
- Zheng M, Cass GR, Schauer JJ, Edgerton ES. 2002. Source apportionment of PM_{2.5} in the southeastern United States using solvent-extractable organic compounds as tracers. *Environmental Science & Technology* 36(11):2361-2371.
- Zheng M, Ke L, Edgerton ES, Schauer JJ, Dong MY, Russell AG. 2006. Spatial distribution of carbonaceous aerosol in the southeastern United States using molecular markers and carbon isotope data. *Journal of Geophysical Research-Atmospheres* 111(D10).

VITA

AMIT MARMUR

Amit Marmur was born and raised in Haifa, Israel. He obtained his B.Sc. in Chemical Engineering and M.Sc. in Environmental Engineering from the Technion – Israel Institute of Technology, where his interest in the environment and air-quality in particular was sparked. In between his academic degrees, Amit worked as a consultant in various environmental related analyses. In the summer of 2001 Amit moved to Atlanta, Georgia to pursue a Ph.D. degree in Environmental Engineering at the Georgia Institute of Technology, and completed his degree in September 2006. He also obtained a second M.Sc. degree during his time at Georgia-Tech. For his graduate research, Amit received the *2004 Georgia Chapter Air and Waste Management Association Student Award* and the *2005/06 Jean-Lou Chameau Research Excellence Award*. Amit is the author of several scientific papers and numerous conference presentations on the topics of particulate matter source-apportionment, air-quality modeling and health/policy implications. Since 2004 Amit has been working for the Georgia Department of Natural Resources, Environmental Protection Division on strategic regional air-resources management. Amit is the proud husband of Efrat and proud father of Eli.