Atmos. Chem. Phys. Discuss., 13, 26657–26698, 2013
www.atmos-chem-phys-discuss.net/13/26657/2013/
doi:10.5194/acpd-13-26657-2013

Atmospheric Chemistry and Physics Discussions

Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach

Y. Hu[1], S. Balachandran[1], J. E. Pachon[1,*], J. Baek[1,], C. Ivey[1], H. Holmes[1], M. T. Odman[1], J. A. Mulholland[1], and A. G. Russell[1]**

[1]School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, Georgia, USA
[*]now at: Program of Environmental Engineering, Universidad de La Salle, Bogota, Colombia
[**]now at: Center for Global and Regional Environmental Research, University of Iowa, Iowa City, Iowa, USA

Received: 31 August 2013 – Accepted: 18 September 2013 – Published: 15 October 2013

Correspondence to: Y. Hu (yh29@mail.gatech.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

A hybrid fine particulate matter ($PM_{2.5}$) source apportionment approach based on a receptor-model (RM) species balance and species specific source impacts from a chemical transport model (CTM) equipped with a sensitivity analysis tool is developed to provide physically- and chemically-consistent relationships between source emissions and receptor impacts. This hybrid approach enhances RM results by providing initial estimates of source impacts from a much larger number of sources than are typically used in RMs, and provides source-receptor relationships for secondary species. Further, the method addresses issues of source collinearities, and accounts for emissions uncertainties. Hybrid method results also provide information on the resulting source impact uncertainties. We apply this hybrid approach to conduct $PM_{2.5}$ source apportionment at Chemical Speciation Network (CSN) sites across the US. Ambient $PM_{2.5}$ concentrations at these receptor sites were apportioned to 33 separate sources. Hybrid method results led to large changes of impacts from CTM estimates for sources such as dust, woodstove, and other biomass burning sources, but limited changes to others. The refinements reduced the differences between CTM-simulated and observed concentrations of individual $PM_{2.5}$ species by over 98 % when using a weighted least squared error minimization. The rankings of source impacts changed from the initial estimates, revealing that CTM-only results should be evaluated with observations. Assessment with RM results at six US locations showed that the hybrid results differ somewhat from commonly resolved sources. The hybrid method also resolved sources that typical RM methods do not capture without extra measurement information on unique tracers. The method can be readily applied to large domains and long (such as multi-annual) time periods to provide source impact estimates for management- and health-related studies.

ACPD
13, 26657–26698, 2013

1 Introduction

Fine particulate matter ($PM_{2.5}$) with an aerodynamic diameter less than 2.5 µm is associated with adverse effects on human health (Dockery et al., 1993). From the perspective of linking health effects with air quality, and for assessing air quality management options, it is desirable to have the spatially and temporally resolved impacts of major emission sources. However, quantifying the impacts of all individual sources on the ambient concentration of fine particulate matter, better known as source apportionment (SA), is challenging. A fundamental issue with any SA method is that there is no way to directly measure source impacts, and, as such, it is difficult to assess the accuracy of source apportionment results. Tracer gases such as cyclic perfluoroalkanes and SF_6 can be utilized to help quantify source impacts (Martin et al., 2011). However, that is far from measuring multiple sources' impacts at the same time and is typically limited to special studies. Instead, source apportionment results are typically evaluated by comparing simulated concentrations of individual component and total mass of $PM_{2.5}$ with observations.

Receptor model (RM) approaches have long been used for $PM_{2.5}$ source apportionment (Chow et al., 1992; Cooper and Watson, 1980; Liu et al., 2006; Martello et al., 2008; Reff et al., 2007; Schauer et al., 1996; Swietlicki et al., 1996; Thurston et al., 2011; Viana et al., 2008b; Watson, 1984; Watson et al., 2008; Xie et al., 2013). These methods, such as the Chemical Mass Balance (CMB) (Watson et al., 1984) or Positive Matrix Factorization (PMF) (Pattero and Tapper, 1994), rely on using observed species concentrations of $PM_{2.5}$ at a receptor(s) and solve a set of species balance equations to estimate source impacts. RM methods typically do not use emissions estimates or explicitly account for the chemical and physical processes that governs pollutant transport and transformation after being emitted from a specific source. To address them, additional approaches had to be used (Blanchard et al., 2012; Chen et al., 2011; Lin and Milford, 1994; Roy et al., 2011; Watson et al., 2002; Wittig and Allen, 2008). In addition, receptor modeling typically accounts for a relatively small number of sources (on the

ACPD
13, 26657–26698, 2013

order of ten out of hundreds in the inventory), comprising about 80 % of the estimated emissions (Baek, 2009), leading to potential biases in the results. In RM methods, the common approach for assessing the accuracy of source apportionment results is to compare the calculated $PM_{2.5}$ composition concentrations and total mass to observations, and if they compare well, it is assumed that the results are reasonable. However, this type of evaluation does not use a set of observations that are totally independent from the ones used to obtain the source impacts, although non-fitting species comparison and other tests can be used to assist in the evaluation (USEPA, 2004). Further, similar estimated species concentrations, and hence similar performance, can result from very different combinations of source impacts. Results can also be quite sensitive to model inputs (e.g., source profiles for CMB), or the number of sources (or factors in PMF) used. Differences in source apportionment results for similar cases found between competing RM methods also suggest errors (Held et al., 2005; Laupsa et al., 2009; Lee et al., 2008; Lowenthal et al., 2010; Marmur et al., 2006; Rizzo and Scheff, 2007; Shi et al., 2009; Viana et al., 2008a; Watson et al., 2008). Several studies have tried to reconcile the results by refining source profiles and adding extra constraints (Lee and Russell, 2007; Marmur et al., 2007; Sheesley et al., 2007; Swietlicki et al., 1996; Watson et al., 2008). Extra species such as organic molecular markers and other unique tracers for certain sources have been utilized in RM modeling (Bullock et al., 2008; Lee et al., 2009; Schauer et al., 1996; Zheng et al., 2002) to improve the accuracy and identify additional sources, however measurements of those markers are not available on routine monitoring networks.

Source-oriented modeling (SM) approaches, such as chemical transport models (CTMs), follow the emission, transport, transformation and loss of chemical species in the atmosphere to simulate ambient concentrations and source impacts. CTMs can compensate for limitations in RM methods (Burr and Zhang, 2011a, b; Doraiswamy et al., 2007; Held et al., 2005; Henze et al., 2009; Kleeman et al., 2007; Lowenthal et al., 2010; Marmur et al., 2006; Russell, 2008; Schichtel et al., 2006; Wagstrom et al., 2008; Ying et al., 2008) because they describe processes affecting source-receptor

ACPD
13, 26657–26698, 2013

relationships from a first principles basis. For example, compared with RMs, CTMs directly account for secondary formation of $PM_{2.5}$ and nonlinearities in pollutant transformations and have the ability to quantify a more complete range of sources. Also, CTMs use knowledge of the specific location of emission sources in the region and their emission rates and can provide spatially resolved source impacts across the modeling domain. An important strength of using CTMs for source apportionment is that model evaluation relies on independent data. However, estimates of source strengths and characteristics (e.g., diurnal and day-to-day variations) are viewed as highly uncertain, meteorological inputs of CTMs contain errors, and there continue to be uncertainties in how various processes are described. Due to these uncertainties and the effort involved, the application of SM approaches for source apportionment is limited.

One way to take the advantages of SM approaches is to further improve SM source apportionment results by utilizing species concentration observations in a manner similar to RM approaches. Here, a hybrid SM-RM approach is developed and applied to obtain improved source impact estimates by integrating measurements with the CTM results, including uncertainty estimates of measurements and emissions. As developed, the method integrates the CMB method with CTM results at monitoring locations and measurement times, by adding additional information and constraints in a species balance approach similar to CMB. The improved source impact estimates at these sparse locations can potentially be utilized to obtain source impact fields using spatial and temporal interpolation that take advantage of the initial CTM estimates across the domain and over the time period of interest. In this study the hybrid approach is applied to a 36 km resolution CTM simulation over North America. Our focus is to demonstrate the hybrid method by closely examining SM-RM source apportionment results across all sites and with more detail at select locations.

ACPD
13, 26657–26698, 2013

2 Methods

2.1 CTM simulation and measurement data

Simulated three-dimensional concentration fields of trace chemical species are obtained using the Community Multiscale Air Quality model (CMAQ) (Byun and Schere, 2006) version 4.5 with strict mass conservation (Hu et al., 2006), the SAPRC-99 chemical mechanism (Carter, 2000) and the aerosol module described in Binkowski and Roselle (2003). The modeling domain (Fig. 1) covers the continental United States as well as portions of Canada and Mexico with 36 km × 36 km horizontal grids and 13 vertical layers of variable thickness extending from the surface to 70 hPa. The semi-normalized first-order sensitivity coefficients of pollutant concentrations to specific model inputs such as emission sources are computed using the decoupled direct method (DDM) (Dunker, 1981, 1984) applied to three dimensional air quality models (Cohan et al., 2005; Dunker et al., 2002; Yang et al., 1997) and extended to include the ability to follow $PM_{2.5}$ (called DDM-3D/PM hereafter) (Boylan et al., 2002, 2006; Koo et al., 2009; Napelenok et al., 2006). In DDM-3D/PM, the semi-normalized first-order sensitivity coefficients, $S^{(1)}_{i,j}(\bar{x}, t)$ (for simplicity, the notations for time (t) and space (x) dependences of variables are dropped below), are defined as the response of species i's concentration c_i to perturbations in a sensitivity parameter p_j (a model parameter or input such as an emission rate, initial condition, or boundary condition) by scaling the local sensitivities ($\partial c_i / \partial p_j$) by P_j(the unperturbed or "base case" value of the sensitivity parameter):

$$S^{(1)}_{i,j} = P_j \frac{\partial c_i}{\partial p_j} = P_j \frac{\partial c_i}{\partial \left(\varepsilon_j P_j\right)} = \frac{\partial c_i}{\partial \varepsilon_j} \quad (1)$$

where ε_j is a scaling variable, with a nominal value of 1 such that

$$p_j = \varepsilon_j P_j \quad (2)$$

ACPD
13, 26657–26698, 2013

The sensitivity parameter, P_j, can be defined as the emission rate of one of the emitted compounds, a group of emitted compounds, or the summation of all the emitted compounds from the same source.

Here, CMAQ uses meteorological fields generated by the Fifth-Generation PSU/NCAR Mesoscale Model (MM5) (Grell et al., 1994), which is run with 35 vertical levels using four dimensional data assimilation (FDDA) and the Pleim-Xiu land-surface model (Pleim and Xiu, 1995; Xiu and Pleim, 2001). Simulated meteorological fields were evaluated against surface hourly observations from the US and Canada (Table S1); performance was well within the typical range for regional air quality modeling (Emery et al., 2001; Hanna and Yang, 2001).

Emissions inputs used were developed from a 2004 inventory that was projected from the 2002 National Emissions Inventory (NEI2002, obtained from http://www.epa.gov/ttn/chief/emch/index.html#2002). Projection of the 2002 inventory to 2004 was conducted using growth factors obtained from the Economic Growth Analysis System (EGAS) Version 4.0, and control efficiency data obtained from EPA for the existing federal and local control strategies. In addition, US emissions from large NO_x and SO_2 point sources for 2004 were obtained from the continuous emissions monitoring (CEM) database http://ampd.epa.gov/ampd/). The inventory has emissions of seven criteria pollutants including $PM_{2.5}$. $PM_{2.5}$ emissions were split into major components (sulfate, nitrate, EC, OC and other) using source-specific speciation profiles from the SPECIATE program (Simon et al., 2010). The component historically called "unidentified" in the emissions modeling process, is called "other" here because this portion of $PM_{2.5}$ is derived from measurements that provide the composition of the emissions, and includes element species which can be used to track source specific impacts on primary $PM_{2.5}$. We group the emissions into 33 integrated source categories (a simple description of the source categories are in Table 1 and further detailed grouping information using source classification code (SCC) can be found in Table S2). The Sparse Matrix Operator Kernel for Emissions (SMOKE) model (CEP, 2003) is used to prepare gridded, CMAQ-ready emissions inputs.

ACPD
13, 26657–26698, 2013

We apply the above modeling system to simulate $PM_{2.5}$ and gaseous concentrations for the month of January 2004, with 1–3 January as ramp-up days. The simulations of major $PM_{2.5}$ and gaseous species were compared against measurements from multiple monitoring networks (Table S3) with performance statistics well within the normal range of current state-of-the-art CTM's (Boylan and Russell, 2006; Tesche et al., 2006). We also computed DDM-3D/PM first-order sensitivity coefficients for each source except SEASALT, as well as boundary and initial conditions (the sensitivity parameters are defined as the summation of all emitted compounds). The sensitivity coefficients of boundary and initial conditions were found minimal and therefore ignored in our calculations. For SEASALT we directly used the simulated concentrations of Na^+ and Cl^- from sea salt emissions in the model, as sensitivities of Na^+ or Cl^- to SEASALT emissions. Sensitivities of other species (including other elements, ions and total mass of $PM_{2.5}$) to SEASALT emissions were derived by applying the composition profile (Table S4) for each species relative to the Na^+ sensitivities. For the other 32 sources, element (metals and minerals) sensitivity coefficients that are not explicitly simulated by CMAQ are derived by applying composition profiles (Table S4) for those elements relative to the modeled, source specific, "other" $PM_{2.5}$ sensitivities, respectively. The source composition profiles of all the 33 categories are assembled from the 86 profiles examined in Reff et al. (2009) by emissions-weighted averaging of corresponding member profiles (determined by SCC groupings). 24 h average simulated $PM_{2.5}$ species, including derived elements' concentrations, are paired with 24 h average measurements at Chemical Speciation Network (CSN) sites (Fig. 1) for further use in optimization. We use 35 elements in $PM_{2.5}$ that are measured at each CSN site along with major components and total mass (Table S5). Detection limit and measurement uncertainty were used to screen for measurements that are invalid or below the detection limit (DL). Values below DL were set to one half of the detection limit and the uncertainty was set to 2/3 of the DL (Marmur et al., 2006). Organic and elemental carbon measurements were artifact corrected and converted from thermal optical transmittance (TOT) values to thermal optical reflectance (TOR) equivalents using the method (Malm et al., 2011) recommended

ACPD
13, 26657–26698, 2013

by US EPA (http://www.epa.gov/ttn/naaqs/standards/pm/data/20120614Frank.pdf; see Note S1).

2.2 CTM source apportionment

Source impacts (and initial and boundary condition impacts) can be estimated using a Taylor series approach (Cohan et al., 2005):

$$SA_{i,j}^{\mathrm{CTM}} = \sum_{k=1}^{K} \left[P_{j,k} \frac{\partial c_i}{\partial p_{j,k}} + \frac{P_{j,k}^2}{2} \frac{\partial^2 c_i}{\partial p_{j,k}^2} + \sum_{l=1;l\neq k}^{L} \frac{P_{j,k} P_{j,l}}{2} \frac{\partial^2 c_i}{\partial p_{j,k} \partial p_{j,l}} \right] + HOT \quad (3)$$

where $SA_{i,j}^{\mathrm{CTM}}$ is the CTM simulated impact (source apportionment result) of source $j (j = 1, \ldots J^{\mathrm{CTM}}$, with J^{CTM} being the total number of sources that are included in the CTM simulation, treating initial and boundary conditions as "sources") on $PM_{2.5}$ species i ($i = 1, \ldots N$ with N being the total number of such species) at the receptor; $P_{j,k}$ is either the emission rate of compound k ($k = 1, \ldots, K$) (k can be different than i, accounting for species transformations) from source j, i.e. $E_{j,k}$, or the initial or boundary concentration of compound k; l and L are the same as k and K; $p_{j,k}(p_{j,l})$ is the sensitivity parameter for $P_{j,k}(P_{j,l})$, and HOT stands for high order terms. The total impact of source j on the $PM_{2.5}$ concentration using CTM method (SR_j^{CTM}) is found by summing its impact on each species concentration:

$$SR_j^{\mathrm{CTM}} = \sum_{i=1}^{N} SA_{i,j}^{\mathrm{CTM}} \quad (4)$$

Here, only the first order terms will be used assuming higher order terms are typically small (Hakami et al., 2004), particularly for primary $PM_{2.5}$ species that do not undergo

ACPD
13, 26657–26698, 2013

significant transformation, so

$$SA_{i,j}^{\mathrm{CTM}} \approx \sum_{k=1}^{K} P_{j,k} \frac{\partial c_i}{\partial p_{j,k}} = \sum_{k=1}^{K} S_{i,j,k}^{(1)} = S_{i,j}^{(1)} \tag{5}$$

where $S_{i,j,k}^{(1)}$ is the semi-normalized first-order sensitivity of species i's concentration to emission rate (or initial and boundary conditions) of compound k from source j, while $S_{i,j}^{(1)}$ is the similar first-order sensitivity to the emissions of all compounds from source j, here as computed by CMAQ with DDM-3D/PM. Again, the notations for time and space dependencies are dropped for simplicity.

This result can be compared with the CMB method, which is based on apportioning each species proportional to the relative amount of that species in the $PM_{2.5}$ emissions from a source:

$$SA_{i,j}^{\mathrm{CMB}} = \frac{E_{j,i}}{E_j} SR_j^{\mathrm{CMB}} = f_{i,j} SR_j^{\mathrm{CMB}} \tag{6}$$

Where $f_{i,j} = \frac{E_{j,i}}{E_j}$ represents the original source profile used by CMB, i.e. the emission fraction of species $i(E_{j,i})$ of the total $PM_{2.5}$ (E_j) emitted from source j ($j = 1, \ldots J^{\mathrm{CMB}}$ with J^{CMB} being the total number of emission sources that the CMB approach considers, source j here can be different than the sources CTM includes) and SR_j^{CMB} is the CMB-calculated impact of source j on total $PM_{2.5}$ concentration. One can extend the definition of $f_{i,j}$ for CTMs using Eq. (5) that includes the source impacts on condensed secondary pollutants in the analysis. Hence, an effective $f_{i,j}^*$ is found as:

$$f_{i,j}^* = \frac{SA_{i,j}^{\mathrm{CTM}}}{SR_j^{\mathrm{CTM}}} = \frac{S_{i,j}^{(1)}}{\sum_{i=1}^{N} S_{i,j}^{(1)}} = \frac{\sum_{k=1}^{K} S_{i,j,k}^{(1)}}{\sum_{i=1}^{N} \sum_{k=1}^{K} S_{i,j,k}^{(1)}} \tag{7}$$

ACPD
13, 26657–26698, 2013

Equation (7) reveals that when there are no emissions of $PM_{2.5}$ component i from source j, $f^*_{i,j}$ can still be non-zero, as the source could still contribute to secondary production of $PM_{2.5}$.

2.3 CTM-CMB hybrid source apportionment approach

At monitoring locations, on days with sufficient $PM_{2.5}$ composition measurements available, the following species balance equations can be built for a CMB solution:

$$c_i^{obs} = \sum_{j=1}^{J^{CMB}} f_{i,j} SR_j^{CMB} + e_i^{CMB} \quad (8)$$

where c_i^{obs} is the measured concentration for the ith $PM_{2.5}$ species, and e_i^{CMB} is the concentration prediction error to be minimized. CMB solves the species balance equations to calculate a set of SR_j^{CMB} using fixed source profiles $f_{i,j}$ (with uncertainties) that minimizes the weighted squared error in the simulated concentrations (Watson, 1984).

Likewise, similar species balance equations can be built at the same receptors using the initial source apportionments from CMAQ DDM-3D/PM results as follows:

$$c_i^{obs} = \sum_{j=1}^{J^{CTM}} SA_{i,j}^{CTM} + e_i^{CTM} = \sum_{j=1}^{J^{CTM}} \sum_{k=1}^{K} S_{i,j,k}^{(1)} + e_i^{CTM} = \sum_{j=1}^{J^{CTM}} S_{i,j}^{(1)} + e_i^{CTM} \quad (9)$$

The extension to using CTM results is shown in the second through the fourth equalities where e_i^{CTM} is the prediction error of CTM for the ith $PM_{2.5}$ species. This equation is applied at specific receptor locations and times.

ACPD
13, 26657–26698, 2013

Utilizing Eq. (9) we can evaluate the initial source apportionment results for a measurement at a receptor by calculating the square prediction error as:

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{\left(c_i^{\text{obs}} - \sum_{j=1}^{J^{\text{CTM}}} SA_{i,j}^{\text{CTM}} \right)^2}{\sigma^2_{c_i^{\text{obs}}}} \right] \tag{10}$$

where $\sigma_{C_i^{\text{obs}}}$ is the uncertainty in the measured concentration of species i obtained from the CSN measurement uncertainty.

Equation (10) also sheds light on an opportunity to further minimize the CTM's prediction error in a least square solution that mimics the CMB method. This leads to a new method of conducting source apportionment in a SM-RM hybrid approach. One way to achieve this is to calculate a new set of SR_j^{CTM} using the extended $f^*_{i,j}$ that minimizes the weighted squared error in the simulated concentrations as follows:

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{\left(c_i^{\text{obs}} - \sum_{j=1}^{J^{\text{CTM}}} f^*_{i,j} SR_j^{\text{CTM}} \right)^2}{\sigma^2_{c_i^{\text{obs}}}} \right] \tag{11}$$

While this approach is similar to CMB, it accounts for secondary contributions and other atmospheric processing using the extended $f^*_{i,j}$. If Eq. (11), alone, were used to develop revised source impacts, it would not fully take into account the information provided by the CTM about the estimated size and location of various emission sources and their probable impact on pollutant concentrations at a receptor, i.e., the initial source impact

ACPD
13, 26657–26698, 2013

estimates $SR_j^{CTM} = \sum_{i=1}^{N} SA_{i,j}^{CTM}$. As formulated in Eq. (11), this information is only used in the calculation of $f_{i,j}^*$, but the magnitudes of the source impacts are lost. Further, collinearity and uniqueness issues, such as different sources sharing similar source profiles, would still impact the solution of the system of equations.

Instead of the above approach, the CMB concept is extended to directly use the initial estimates of $SA_{i,j}^{CTM}$ as well as the initial simulated concentrations c_i^{init} from the CTM to refine the estimated source impacts. Defining R_j as a scale factor applied to the initial estimate of impact of source j (or initial or boundary conditions), $SA_{i,j}^{refnd}$, the refined CTM-simulated impact of source j on species i is obtained as:

$$SA_{i,j}^{refnd} = R_j SA_{i,j}^{init} \tag{12}$$

Here $SA_{i,j}^{init}$ is the initial source impact ($SA_{i,j}^{init}$ is the same as previous $SA_{i,j}^{CTM}$ and is used from now on to distinguish from $SA_{i,j}^{refnd}$). As such, refinements to source impacts can be found in a similar fashion to traditional CMB approaches by solving for R_j to minimize χ^2, where:

$$\chi^2 = \sum_{i=1}^{N} \frac{\left(c_i^{obs} - c_i^{init} - \sum_{j=1}^{J^{CTM}} (R_j - 1) SA_{i,j}^{init} \right)^2}{\sigma_{c_i^{obs}}^2} \tag{13}$$

However, without further constraints R_j can be physically unrealistic and would not account for the knowledge provided by the CTM about the source impacts or the uncertainties in emission estimates. Here, additional constraints and a term that penalizes moving away from the initial source impact estimates are added to find an optimized

ACPD
13, 26657–26698, 2013

R_j:

$$\chi^2 = \sum_{i=1}^{N} \left[\frac{\left(c_i^{\text{obs}} - c_i^{\text{init}} - \sum_{j=1}^{J^{\text{CTM}}} (R_j - 1) SA_{i,j}^{\text{init}} \right)^2}{\sigma_{c_i^{\text{obs}}}^2 + \sigma_{SR_i^{\text{CTM}}}^2} \right] + \Gamma \sum_{j=1}^{J^{\text{CTM}}} \frac{(\ln R_j)^2}{\sigma_{\ln R_j}^2} \quad (14)$$

where $\sigma_{SR_i^{\text{CTM}}}$ is the a priori uncertainty in CTM-derived total sources' impact on the ith species, which is added to give weight for initial source impact estimates for different species and represents model errors. One can estimate $\sigma_{SR_i^{\text{CTM}}}$ as proportional to observed concentration $\sigma_{SR_i^{\text{CTM}}} = \delta_i * c_i^{\text{obs}}$, with δ_i as normalized model errors. The second term of the equation accounts for uncertainties in the CTM-derived individual source impacts due to emissions error. $\sigma_{\ln R_j}$ is the a priori uncertainty of the natural log of source j's scale factor. The logarithmic form is used as it has the same value on a relative basis. This naturally constrains R_j to be positive. Γ is introduced to balance the two terms in Eq. (14).

The objective function expressed as Eq. (14) can be minimized by using various optimization algorithms available for nonlinear optimization problems with constraints. We have tested a few such algorithms, including the algorithm of Sequential Least-Square Quadratic Programming (SLSQP) (Kraft, 1988, 1994) and L-BFGS, a limited-memory quasi-Newton optimization function (Liu and Nocedal, 1989; Nocedal, 1980). With both the SLSQP and L-BFGS methods one can set lower and upper limits on R_j for each individual source. We chose L-BFGS for our demonstration case study. As R_j is optimized, the refined estimates of individual source impacts by species at a specific location are then given by Eq. (12). The level of remaining error in the refined concentration predictions can be found using Eq. (13).

ACPD
13, 26657–26698, 2013

2.4 Application and case study

The hybrid method was applied for January of 2004 to calculate $PM_{2.5}$ source impact scale factors at 164 CSN monitors for which we had valid speciated $PM_{2.5}$ data. By using the valid measurements at each of these CSN sites for each valid day, the initial source impacts were evaluated through Eq. (14) to obtain impact scale factors and refined source impacts estimates. The L-BFGS algorithm was used with box constraints that limited R_j to be between 0.1 and 10.0 (different sets of limits have been tested, up to the range of between 0.02 and 50.0). Two steps were used to apply L-BFGS to find the final optimized R_j. First, an initial choice for Γ was set as $\Gamma = \frac{N}{J^{CTM}} = \frac{41}{33} = 1.24$ to equally weigh the two terms in the objective function and obtain the initial optimal R_j. Then, the initial optimal R_j were used to create a new Γ as the value of the first term of the objective function divided by J^{CTM}. The new Γ (typically about 20) is then applied to obtain the final optimized R_j. Here $\sigma_{\ln R_j}$ are determined by considering the daily emission estimates uncertainties for each source (Table S2) derived from literature (Hanna et al., 1998, 2001, 2005). In general, regulated sectors such as industrial, on-road and non-road sources have lower uncertainties, non-regulated sectors such as residential related sources, dust and biomass-burning have much higher uncertainties. Because the refinements are applied daily, the uncertainties used account for the day-to-day variability in source strengths. For example, prescribed burning events can be quite variable in time. For traffic, day-specific emissions patterns are used, so the source strength's variability is smaller. Sources for which direct emissions monitoring is available are assigned the lowest uncertainty. To determine $\sigma_{SR_i^{CTM}}$, δ_i (Table S6) are chosen as the typical normalized prediction errors of $PM_{2.5}$ species as found in regional applications of state-of-the-art CTM models (Appel et al., 2008; Boylan and Russell, 2006; Tesche et al., 2006). Results were found not very sensitive to the range of values of $\sigma_{\ln R_j}$ and $\sigma_{SR_i^{CTM}}$ tested.

We chose six CSN sites, each representing a major US metropolitan area, for close examination of the method and further analysis. These six sites are located in the At-

ACPD
13, 26657–26698, 2013

lanta, Chicago, Detroit, Los Angeles, New York, and Pittsburgh areas, representing urban/suburban locations across the country. Additional information for these six sites can be found in Tables S7 (basic site information) and Table S8 (emissions estimates surrounding each site). For comparison, we also conducted CMB modeling at the Atlanta site using the same measurement dataset and collected source apportionment results from literature at the other five sites.

3 Results

3.1 Impact scale factors and refined concentration predictions

The hybrid method was applied to obtain R_j and to further refine the initial source impact estimates. An R_j of less than 1.0 means that the refined impact is reduced from the original (suggesting that the emissions are biased high or that the CTM is leading to a high bias in the source-receptor relationship) while larger than 1.0 means that the impact is increased from the initial simulation. The R_j values obtained for the 33 sources have means ranging from 0.15 to 1.0 with sources of higher uncertainties having larger standard deviations (Table 2). In general, sources that are commonly considered as having high uncertainties were found to have R_j values deviating the most from 1.0, while those sources considered less uncertain were found to have R_j values near 1.0. This is expected, in part because of the second term in the weighting function. The scale factors are also found to be quite consistent (i.e. in same directions), in general, for the same source between locations and between days at the same location (Table S9). Most significantly, R_j's cumulative distribution functions are found to be distinctive between sources (Fig. S1). This is true even between biomass-burning sources although most of them have a similar composition in emissions (Fig. S1a). Dust, lawn waste burning (LWASTEBURN) and woodstove impacts (and other biomass-burning sources as well, although to a lesser extent) are found to be biased high (R_j values typically ~ 0.1). This is consistent with findings of prior studies (Baek, 2009; Chow

ACPD
13, 26657–26698, 2013

et al., 2007; Tian et al., 2009) that emission rates for these sources were overestimated. Also, prescribed burning impacts are found to be biased low (R_j values being close to 10.0) a small portion of the time due to its high day-to-day variations. Typically, prescribed burn emissions are distributed uniformly over time in the inventories while in realty burns occur on days with favorable burning conditions. For most other sources (Fig. S1b–d), impact scale factors are typically closer to 1.0, with most of the R_j values between 0.8 and 1.1, with the exceptions of metal processing, cooking processes, fuel oil and natural gas combustion, on-road gasoline vehicle and "others" sources. These six sources have more diverse R_j values among locations and/or between days.

An indication of the magnitude of the refinements can be found by comparing the initial and refined individual species concentrations to the observations and can be quantified using the weighted least square error (i.e. χ^2 as expressed in Eq. 13). The simulated concentrations are found to be improved substantially compared to the initial simulation after refining source-impact estimates for major individual components and for most of the elements (Fig. 2 and Table 3). Note that several elements with very low ambient concentrations (e.g. near the measurement uncertainty) were found to have slightly deteriorated agreement with observations (Table 3). However, results show that the refined $\chi^2_{\text{c, refnd}}$ (Eq. 13 with obtained R_j), an overall index for remaining error, were reduced by over 98 % on average (Fig. 3).

3.2 Initial and refined source impacts

Significant day-to-day variations are found in the initial source impact estimates (e.g. Table S10, as renormalized by total source impact), being more pronounced for some sources, such as power plants (i.e. coal combustion) and industrial sources. For example, in Atlanta, power plants (coal combustion) can contribute over 30 % on one day but only about 5 % on other days (primarily as secondary sulfates). In Chicago, metal processing contributes 20 % on some days but less than 10 % on other days. On-road gasoline impact can also vary significantly day to day, such as in Detroit, it varies from ~ 18 % to ~ 3 %. Biomass-burning sources such as prescribed burns and agriculture

ACPD
13, 26657–26698, 2013

burns, contribute significantly on some days in Atlanta, but have virtually zero impact on other days.

Refined source impacts changed significantly from the initial estimates for sources with high uncertainties, such as woodstove and dust, as well as other biomass-burning sources, but changed much less or little for other sources (compare left and right columns in Tables 4 and S10). Woodstove and dust were top ranked at all six sites from the initial estimates; however, refinement significantly lowered those sources' impacts (Table 5). The differing adjustments between sources resulted in the rankings of top contributors changing. This indicates that estimates from SM-only methods might result in misleading source apportionment outcome due to the errors in emission estimates on the specific day, as well as meteorological field and model parameter errors. For example, Marmur et al. (2006) found that the CMAQ-calculated impact of soil dust at Jefferson Street, Atlanta, GA (and other locations) was high when compared with two CMB estimates. This shows that it is necessary to evaluate SM source apportionment results using measurements.

The hybrid method can separate sources with similar composition, e.g., woodstove and prescribed burns, especially noting the different changes of these two sources between their initial and refined impacts in Table S10a and d, as well as on-road and non-road diesel vehicles. This is because it starts from integrating estimated emissions from the inventory with source specific spatial and temporal resolution, instead of starting from only the source composition like RMs do. In addition, with the hybrid method, secondary pollutants are apportioned to specific sources while in RMs they are aggregated together. For example, after the hybrid method refinement livestock impacts advance in rank among top contributors in Midwestern cities: Chicago, Detroit and Pittsburgh (Table 5), mostly through the secondary formation of ammonium and the associated nitrate from NH_3 emissions. Also, the two most common major contributors across the cities become coal combustion (except Los Angeles; Table 5), mainly due to the sulfate formation from SO_2 emissions, and on-road gasoline vehicles, partially due to nitrate and organic matter formation from NO_x and VOC emissions.

ACPD
13, 26657–26698, 2013

3.3 Comparison of refined source impacts with results from RM methods

In order to compare with other source apportionment studies, we first reduced the number of sources from 33 to 13 by aggregating source impacts (Table 6). The 13 aggregated sources are chosen to cover the range of various sources in different locations as identified in prior studies. Sources having similar composition, e.g. various gasoline and diesel vehicular sources, were merged accordingly. "AllOthers" included sources typically not resolved in traditional SA studies, e.g. livestock, biogenic and solvents as well as minor combustion and industrial sources. "AllOthers" (due to its large secondary contribution), and gasoline and diesel vehicles are top ranked in all six cities (Table 6). To make hybrid results directly comparable to that of RM methods, we further separated the primary and secondary contributions in the aggregated source impacts and merged the secondary portions correspondingly into ammonium sulfates, ammonium nitrate, and secondary organic carbon (Details are discussed in Note S2). We compare the regrouped hybrid results to results of RM methods conducted at the same location by this or prior studies (Coutant et al., 2003; Gildemeister et al., 2007; Maranche, 2006; Pham et al., 2008; Rizzo and Scheff, 2007) in Table 7. All the RM results were based on CSN measurements, though time periods for other RM results may differ (details of RM model applications are found in Note S3).

The hybrid approach resolved extra sources (with the total impacts of extra sources ranging between 20–30 % at the six sites) that are typically missing from RM results (Table 7). This is consistent with ~ 20 % of the emissions that Baek (2009) found were not captured in most RM source apportionment applications. For example, CMB-LGO did not capture the aircraft source impact at the Atlanta site (Balachandran et al., 2012) as the profile is uncertain and similar to diesel combustion. However, measurement (Herndon et al., 2008; Lee et al., 2011) and modeling (Unal et al., 2005) studies have both suggested that the commercial aircraft engine emissions from the Atlanta Airport had significant impacts on local air quality including $PM_{2.5}$ concentrations. Natural gas combustion and cooking process are the two sources usually not resolved by

ACPD
13, 26657–26698, 2013

RM methods using CSN data because their identification needs extra measurement information. For instance, CMB with particle-phase organic compounds as tracers using measurements collected at the Jefferson street site has identified that natural gas combustion had a 1.1 % impact on $PM_{2.5}$ in Atlanta (Zheng et al., 2002). Subramanian et al. (2007) used CMB with molecular markers and found that the impact of cooking processes range from 1–5 % on $PM_{2.5}$ concentrations in Pittsburgh. Compared to the hybrid results, the coal combustion primary impact estimates from RM methods are either missing or too low. This is because the trace element markers for coal combustion, Se and Sr, were not detected consistently in CSN samples due to low signal to noise ratios (Chen et al., 2010).

Hybrid results estimated total vehicle impacts (ranging between 14–22 %) were comparable to the RM results found at the same urban/suburban locations, with an exception in Chicago (Table S11). In Chicago, Rizzo and Scheff (2007) also conducted PMF modeling using the same composite data and their PMF results differ from CMB results, e.g. for biomass burning (5 % vs. 11 %) and vehicle (23 % vs. 31 %) source impacts. The PMF results were closer to the hybrid findings. At three of the four sites that the RM methods separated vehicle impacts between diesel and gasoline, the hybrid results do not agree with the RM methods on the diesel/gasoline split (Table S11): the hybrid method found higher impacts of diesel vs. gasoline (by a factor of 1.97–2.62), while the RMs found the opposite (0.28–0.49). The ratios of diesel/gasoline emissions surrounding the sites are in the range of 1.67–3.58 (Table S11). Subramanian at al. (2006) also found that diesel impacts in Pittsburgh tend to dominate by utilizing molecular markers. The split between diesel and gasoline vehicular impacts at Minnesota CSN sites from CMB solutions has been found to be inaccurate (Chen et al., 2011) when only regular measurements were used. Chow et al. (2007) suggested the difficulty of CMB to make an accurate gasoline/diesel split without organic marker compounds.

Hybrid results tend to find lower secondary contributions than the RM methods, except in Chicago and Pittsburgh (Table S12). While the hybrid and RMs agree well on the ammonium sulfates at all six sites (16–37 % vs. 20–38 %; Table 7), the hybrid method

ACPD
13, 26657–26698, 2013

estimated lower secondary organic carbon (4.8 % vs. 11.7 %) in Atlanta, and they differ the most on secondary nitrate impacts (3–27 % vs. 20–44 %; Table 7). The difficulties in simulating particulate nitrate have been noted previously (Chang et al., 2011).

4 Discussion

The hybrid source apportionment method developed and applied here has been demonstrated to be a novel way to improve SM-only CTM results by utilizing independent measurements. It also has advantages over RM methods. First, some limitations of RM methods are addressed (depending upon RM method): (1) the assumption that emissions are inert, with no chemical reactions, (2) not all source categories are considered, (3) potential collinearities between source compositions and, (4) inconsistent or unrealistic results because receptor models do not include information on the strength and location of source emissions, and (5) not accounting for physical process such as complex meteorology. Second, the refinement and evaluation of the source impact estimates use measurement data that are independent from those used to develop the initial source impact estimates. Additionally, the hybrid method can be applied to obtain spatial fields of source impacts providing hourly spatial fields.

A number of potential uncertainties from the CTM modeling can lead to uncertainties in the estimated impacts from the hybrid approach. The assumption for deriving concentrations and sensitivities for the elements that are not explicitly simulated in the CTM model might not hold always. The missing pathways of secondary organic aerosol formation and inaccurate representation of nitrate formation in the CTM model can lead to underestimation of secondary source impacts. Errors in the meteorology may result in errors in the source fingerprints ($f^*_{i,j}$). Errors in the initial emissions inventory, particularly in the spatial and/or temporal variability and in the composition of the emissions, also introduce potential errors, particularly when using the model to temporally interpolate the impact adjustments, i.e., to provide 1 h impact fields after using

ACPD
13, 26657–26698, 2013

the 24 h, speciated PM measurements. Thus, it is best to consider using results of this approach applied to 24 h averaged fields.

On the other hand, evaluating the hybrid model results on a species basis can help identify errors in the original source profiles. Additionally, including measurements from multiple sites in a region and/or spatially dense satellite retrievals in the process of adjusting emissions can further help stabilize R_j. This will provide more accurate refinements and address the possibility of the measurements taken at a single point being overly influenced by local sources. In this direction, the hybrid source results can be more accurate representations of the pollutant levels spatially because they integrate estimates of the spatial distribution of emissions and the local chemical and physical atmospheric processes.

Supplementary material related to this article is available online at http://www.atmos-chem-phys-discuss.net/13/26657/2013/acpd-13-26657-2013-supplement.pdf.

Acknowledgements. This publication was made possible by funding from the US EPA under grants RD834799 and RD833866, NASA under grant NNX11AI55G and Georgia Power. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the supporting agencies. Further, The US Government does not endorse the purchase of any commercial products or services mentioned in the publication.

References

Appel, K. W., Bhave, P. V., Gilliland, A. B., Sarwar, G., and Roselle, S. J.: Evaluation of the community multiscale air quality (CMAQ) model version 4.5: sensitivities impacting model performance; Part II – particulate matter, Atmos. Environ., 42, 6057–6066, 2008.

ACPD
13, 26657–26698, 2013

Baek, J.: Improving aerosol simulations: assessing and improving emissions and secondary organic aerosol formation in air quality modeling, Ph.D. Dissertation, Georgia Institute of Tecnology, Atlanta, GA, 140 pp., 2009.
Balachandran, S., Pachon, J. E., Hu, Y., Lee, D., Mulholland, J. A., and Russell, A. G.: Ensemble-trained source apportionment of fine particulate matter and method uncertainty analysis, Atmos. Environ., 61, 387–394, 2012.
Binkowski, F. S. and Roselle, S. J.: Models-3 Community Multi-scale Air Quality (CMAQ) model aerosol component: 1. Model description, J. Geophys. Res., 108, 4183, doi:10.1029/2001JD001409, 2003.
Blanchard, C. L., Tanenbaum, S., and Hidy, G. M.: Source contributions to atmospheric gases and particulate matter in the Southeastern united states, Environ. Sci. Technol., 46, 5479–5488, 2012.
Boylan, J. W. and Russell, A. G.: PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, Atmos. Environ., 40, 4946–4959, 2006.
Boylan, J. W., Odman, M. T., Wilkinson, J. G., Russell, A. G., Doty, K. G., Norris, W. B., and McNider, R. T.: Development of a comprehensive, multiscale "one-atmosphere" modeling system: application to the Southern Appalachian Mountains, Atmos. Environ., 36, 3721–3734, 2002.
Boylan, J. W., Odman, M. T., Wilkinson, J. G., and Russell, A. G.: Integrated assessment modeling of atmospheric pollutants in the Southern Appalachian Mountains: Part II – $PM_{2.5}$ and visibility, J. Air Waste Manage., 56, 12–22, 2006.
Bullock, K. R., Duvall, R. M., Norris, G. A., McDow, S. R., and Hays, M. D.: Evaluation of the CMB and PMF models using organic molecular markers in fine particulate matter collected during the Pittsburgh Air Quality Study, Atmos. Environ., 42, 6897–6904, 2008.
Burr, M. J. and Zhang, Y.: Source apportionment of fine particulate matter over the Eastern US Part II: Source apportionment simulations using CAMx/PSAT and comparisons with CMAQ source sensitivity simulations, Atmos. Pollut. Res., 2, 318–336, doi:10.5094/APR.2011.5037, 2011a.
Burr, M. J. and Zhang, Y.: Source apportionment of fine particulate matter over the Eastern US Part I: Source sensitivity simulations using CMAQ with the Brute Force method, Atmos. Pollut. Res., 2, 300–317, doi:10.5094/APR.2011.5036, 2011b.

ACPD
13, 26657–26698, 2013

Byun, D. and Schere, K. L.: Review of the governing equations, computational algorithms, and other components of the Models-3 Community Multiscale Air Quality (CMAQ) modeling system, Appl. Mech. Rev., 59, 51–77, 2006.

Carter, W. P. L.: Documentation of the SAPRC-99 Chemical Mechanism for VOC Reactivity Assessment, Contract No. 92–329 and 95–308, California Air Resources Board, 2000.

CEP: Sparse Matrix Operator Kernel Emissions Modeling System (SMOKE) User Manual, Carolina Environmental Program – The University of North Carolina at Chapel Hill, Chapel Hill, NC, 2003.

Chang, W. L., Bhave, P. V., Brown, S. S., Riemer, N., Stutz, J., and Dabdub, D.: Heterogeneous atmospheric chemistry, ambient measurements, and model calculations of N_2O_5: a review, Aerosol Sci. Tech., 45, 665–695, 2011.

Chen, L. W. A., Watson, J. G., Chow, J. C., DuBois, D. W., and Herschberger, L.: Chemical mass balance source apportionment for combined $PM_{2.5}$ measurements from US non-urban and urban long-term networks, Atmos. Environ., 44(38), 4908–4918, 2010.

Chen, L. W. A., Watson, J. G., Chow, J. C., DuBois, D. W., and Herschberger, L.: $PM_{2.5}$ source apportionment: reconciling receptor models for US nonurban and urban long-term networks, J. Air Waste Manage., 61, 1204–1217, 2011.

Chow, J. C., Watson, J. G., Lowenthal, D. H., Solomon, P. A., Magliano, K. L., Ziman, S. D., and Richards, L. W.: PM_{10} source apportionment in California San Joaquin Valley, Atmos. Environ. A, 26, 3335–3354, 1992.

Chow, J. C., Watson, J. G., Lowenthal, D. H., Chen, L. W. A., Zielinska, B., Mazzoleni, L. R., and Magliano, K. L.: Evaluation of organic markers for chemical mass balance source apportionment at the Fresno Supersite, Atmos. Chem. Phys., 7, 1741–1754, doi:10.5194/acp-7-1741-2007, 2007.

Cohan, D. S., Hakami, A., Hu, Y., and Russell, A. G.: Nonlinear response of ozone to emissions: source apportionment and sensitivity analysis, Environ. Sci. Technol., 39, 6739–6748, 2005.

Cooper, J. A. and Watson, J. G.: Reseptor oriented methods of air particulate source apportionment, J. Air Pollut. Control Assoc., 30, 1116–1125, 1980.

Coutant, B. W., Holloman, C. H., Swinton, K. E., and Hafner, H. R.: Eight-site source apportionment of $PM_{2.5}$ speciation trend data, EPA Contract No. 68-D-02-061, Work Assignment 1-05, 2003.

ACPD
13, 26657–26698, 2013

Dockery, D. W., Pope, C. A., Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., and Speizer, F. E.: An association between air pollution and mortality in six US cities, The New Engl. J. Med., 329, 24, 1753–1759, 1993.
Doraiswamy, P., Davis, W. T., Miller, T. L., and Fu, J. S.: Source apportionment of fine particles in Tennessee using a source-oriented model, J. Air Waste Manage., 57, 407–419, 2007.
Dunker, A. M.: Efficient calculation of sensitivity coefficients for complex atmospheric models, Atmos. Environ., 15, 1155–1161, 1981.
Dunker, A. M.: The decoupled direct method for calculating sensitivity coefficients in chemical-kinetics, J. Chem. Phys., 81, 2385–2393, 1984.
Dunker, A. M., Yarwood, G., Ortmann, J. P., and Wilson, G. M.: Comparison of source apportionment and source sensitivity of ozone in a three-dimensional air quality model, Environ. Sci. Technol., 36, 2953–2964, 2002.
Emery, C., Tai, E., and Yarwood, G.: Enhanced Meteorological Modeling and Performance Evaluation for two Texas Ozone Episodes, Prepared for the Texas Natural Resource Conservation Commissions, ENVIRON International Corporation, Novato, CA, 2001.
Gildemeister, A. E., Hopke, P. K., and Kim, E.: Sources of fine urban particulate matter in Detroit, MI, Chemosphere, 69, 1064–1074, 2007.
Grell, G., Dudhia, J., and Stauffer, D.: A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5), NCAR/TN 398+STR, NCAR Technical Note: NCAR/TN-398+STR, 1994.
Hakami, A., Odman, M. T., and Russell, A. G.: Nonlinearity in atmospheric response: a direct sensitivity analysis approach, J. Geophys. Res.-Atmos., 109(D15), 2004.
Hanna, S., Chang, J., and Fernau, M.: Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables, Atmos. Environ., 32, 3619–3628, 1998.
Hanna, S. R. and Yang, R.: Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths, J. Appl. Meteorol., 40, 1095–1104, 2001.
Hanna, S. R., Lu, Z., Frey, H. C., Wheeler, N., Vukovich, J., Arumachalam, S., and Fernau, M. E.: Uncertainties in predicted ozone concentration due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain, Atmos. Environ., 35, 891–903, 2001.
Hanna, S. R., Russell, A. G., Wilkinson, J., Vukovich, J., and Hansen, D. A.: Monte Carlo estimation of uncertainties in BEIS3 emission outputs and their effects on uncertainties in chemical

ACPD
13, 26657–26698, 2013

transport model predictions, J. Geophys. Res., 110, D01302, doi:10.1029/02004JD004986, 2005.
Held, T., Ying, Q., Kleeman, M. J., Schauer, J. J., and Fraser, M. P.: A comparison of the UCD/CIT air quality model and the CMB source-receptor model for primary airborne particulate matter, Atmos. Environ., 39, 2281–2297, 2005.
Henze, D. K., Seinfeld, J. H., and Shindell, D. T.: Inverse modeling and mapping US air quality influences of inorganic $PM_{2.5}$ precursor emissions using the adjoint of GEOS-Chem, Atmos. Chem. Phys., 9, 5877–5903, doi:10.5194/acp-9-5877-2009, 2009.
Herndon, S. C., Jayne, J. T., Lobo, P., Onasch, T. B., Fleming, G., Hagen, D. E., Whitefield, P. D., and Miake-Lye, R. C.: Commercial aircraft engine emissions characterization of in-use aircraft at Hartsfield-Jackson Atlanta International Airport, Environ. Sci. Technol., 42, 1877–1883, 2008.
Hu, Y., Odman, M. T., and Russell, A. G.: Mass conservation in the Community Multiscale Air Quality model, Atmos. Environ., 40, 1199–1204, 2006.
Kleeman, M. J., Ying, Q., Lu, J., Mysliwiec, M. J., Griffin, R. J., Chen, J. J., and Clegg, S.: Source apportionment of secondary organic aerosol during a severe photochemical smog episode, Atmos. Environ., 41, 576–591, 2007.
Koo, B., Wilson, G. M., Morris, R. E., Dunker, A. M., and Yarwood, G.: Comparison of source apportionment and sensitivity analysis in a particulate matter air quality model, Environ. Sci. Technol., 43, 6669–6675, 2009.
Kraft, D.: A software package for sequential quadratic programming, Technical Report DFVLR-FB 88-28, Institut für Dynamik der Flugsysteme, Oberpfaffenhofen, July 1988.
Kraft, D.: Algorithm 733: TOMP – Fortran modules for optimal control calculations, ACM Trans. Math. Softw., 20, 262–281, 1994.
Laupsa, H., Denby, B., Larssen, S., and Schaug, J.: Source apportionment of particulate matter ($PM_{2.5}$) in an urban area using dispersion, receptor and inverse modelling, Atmos. Environ., 43, 4733–4744, 2009.
Lee, B. H., Wood, E. C., Miake-Lye, R. C., Herndon, S. C., Munger, J. W., and Wofsy, S. C.: Reactive chemistry in aircraft exhaust implications for air quality, Transp. Res. Record, 2206, 19–23, 2011.
Lee, D., Balachandran, S., Pachon, J., Shankaran, R., Lee, S., Mulholland, J. A., and Russell, A. G.: Ensemble-trained $PM_{2.5}$ source apportionment approach for health studies, Environ. Sci. Technol., 43, 7023–7031, 2009.

ACPD
13, 26657–26698, 2013

Lee, S. and Russell, A. G.: Estimating uncertainties and uncertainty contributors of CMB $PM_{2.5}$ source apportionment results, Atmos. Environ., 41, 9616–9624, 2007.
Lee, S., Liu, W., Wang, Y. H., Russell, A. G., and Edgerton, E. S.: Source apportionment of $PM_{2.5}$: comparing PMF and CMB results for four ambient monitoriniz sites in the southeastern United States, Atmos. Environ., 42, 4126–4137, 2008.
Lin, C. and Milford, J. B.: Decay-adjusted chemical mass-balance receptor modeling for volatile organic-compounds, Atmos. Environ., 28, 3261–3276, 1994.
Liu, D. C. and Nocedal, J.: On the limited memory method for large scale optimization, Math. Progr. B, 45, 503–528, 1989.
Liu, W., Wang, Y. H., Russell, A., and Edgerton, E. S.: Enhanced source identification of southeast aerosols using temperature-resolved carbon fractions and gas phase components, Atmos. Environ., 40, S445–S466, 2006.
Lowenthal, D. H., Watson, J. G., Koracin, D., Chen, L. W. A., Dubois, D., Vellore, R., Kumar, N., Knipping, E. M., Wheeler, N., Craig, K., and Reid, S.: Evaluation of regional-scale receptor modeling, J. Air Waste Manage., 60, 26–42, 2010.
Malm, W. C., Schichtel, B. A., and Pitchford, M. L.: Uncertainties in $PM_{2.5}$ gravimetric and speciation measurements and what we can learn from them, J. Air Waste Manage., 61, 1131–1149, 2011.
Maranche, J.: Allegheny County $PM_{2.5}$ Source Apportionment Results using the Positive Matrix Factorization Model, Allegheny County Health Department, Pittsburgh, PA, available at: http://www.achd.net/airqual/pubs/pdf/pmf0106.pdf, 2006.
Marmur, A., Park, S. K., Mulholland, J. A., Tolbert, P. E., and Russell, A. G.: 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, Atmos. Environ., 40, 2533–2551, 2006.
Marmur, A., Mulholland, J. A., and Russell, A. G.: Optimized variable source-profile approach for source apportionment, Atmos. Environ., 41, 493–505, 2007.
Martello, D. V., Pekney, N. J., Anderson, R. R., Davidson, C. I., Hopke, P. K., Kim, E., Christensen, W. F., Mangelson, N. F., and Eatough, D. J.: Apportionment of ambient primary and secondary fine particulate matter at the Pittsburgh National Energy Laboratory particulate matter characterization site using positive matrix factorization and a potential source contributions function analysis, J. Air Waste Manage., 58, 357–368, 2008.

ACPD
13, 26657–26698, 2013

Martin, D., Petersson, K. F., and Shallcross, D. E.: The use of cyclic perfluoroalkanes and SF_6 in atmospheric dispersion experiments, Q. J. Roy. Meteor. Soc., 137, 2047–2063, 2011.
Napelenok, S. L., Cohan, D., Hu, Y., and Russell, A. G.: Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM), Atmos. Environ., 40, 6112–6121, 2006.
Nocedal, J.: Updating quasi-Newton matrices with limited storage, Math. Comput., 35, 773–782, 1980.
Pattero, P. and Tapper, U.: Positive matrix factorization – a nonnegative factor model with optimal utilization of error-estimates of data values, Environmetrics, 5, 111–126, 1994.
Pham, T.-L., Kim, B. M., Katzenstein, A., and Goss, T.: $PM_{2.5}$ Source Apportionment for the South Coast Air Basin Using Chemical Mass Balance Receptor Model, MATES III Final Report, available at: http://www.aqmd.gov/prdas/matesIII/Final/Appendices/g-MATESIIIAppendixVIIFinal92008.pdf, 2008.
Pleim, J. E. and Xiu, A.: Development and testing of a surface flux and planetary boundary layer model with explicit soil moisture parameterization for application in mesoscale models, J. Appl. Meteorol., 34, 16–32, 1995.
Reff, A., Eberly, S. I., and Bhave, P. V.: Receptor modeling of ambient particulate matter data using positive matrix factorization: review of existing methods, J. Air Waste Manage., 57, 146–154, 2007.
Reff, A., Bhave, P. V., Simon, H., Pace, T. G., Pouliot, G. A., Mobley, J. D., and Houyoux, M.: Emissions inventory of $PM_{2.5}$ trace elements across the United States, Environ. Sci. Technol., 43, 5790–5796, 2009.
Rizzo, M. J. and Scheff, P. A.: Fine particulate source apportionment using data from the USEPA speciation trends network in Chicago, Illinois: comparison of two source apportionment models, Atmos. Environ., 41, 6276–6288, 2007.
Roy, A. A., Wagstrom, K. M., Adams, P. J., Pandis, S. N., and Robinson, A. L.: Quantification of the effects of molecular marker oxidation on source apportionment estimates for motor vehicles, Atmos. Environ., 45, 3132–3140, 2011.
Russell, A. G.: EPA supersites program-related emissions-based particulate matter modeling: initial applications and advances, J. Air Waste Manage., 58, 289–302, 2008.
Schauer, J. J., Rogge, W. F., Hildemann, L. M., Mazurek, M. A., and Cass, G. R.: Source apportionment of airborne particulate matter using organic compounds as tracers, Atmos. Environ., 30, 3837–3855, 1996.

ACPD
13, 26657–26698, 2013

Schichtel, B. A., Malm, W. C., Gebhart, K. A., Barna, M. G., and Knipping, E. M.: A hybrid source apportionment model integrating measured data and air quality model results, J. Geophys. Res.-Atmos., 111, D7, D07301, doi:10.1029/2005JD006238, 2006.

Sheesley, R. J., Schauer, J. J., Zheng, M., and Wang, B.: Sensitivity of molecular marker-based CMB models to biomass burning source profiles, Atmos. Environ., 41, 9050–9063, 2007.

Shi, G. L., Li, X., Feng, Y. C., Wang, Y. Q., Wu, J. H., Li, J., and Zhu, T.: Combined source apportionment, using positive matrix factorization-chemical mass balance and principal component analysis/multiple linear regression-chemical mass balance models, Atmos. Environ., 43, 2929–2937, 2009.

Simon, H., Beck, L., Bhave, P. V., Divita, F., Hsu, Y., Luecken, D., Mobley, J. D., Pouliot, G. A., Reff, A., Sarwar, G., and Strum, M.: The development and uses of EPA's SPECIATE database, Atmos. Pollut. Res., 1, 196–206, 2010.

Subramanian, R., Donahue, N. M., Bernardo-Bricker, A., Rogge, W. F., and Robinson, A. L.: Contribution of motor vehicle emissions to organic carbon and fine particle mass in Pittsburgh, Pennsylvania: effects of varying source profiles and seasonal trends in ambient marker concentrations, Atmos. Environ., 40, 8002–8019, 2006.

Subramanian, R., Donahue, N. M., Bernardo-Bricker, A., Rogge, W. F., and Robinson, A. L.: Insights into the primary-secondary and regional-local contributions to organic aerosol and $PM_{2.5}$ mass in Pittsburgh, Pennsylvania, Atmos. Environ., 41, 7414–7433, 2007.

Swietlicki, E., Puri, S., Hansson, H. C., and Edner, H.: Urban air pollution source apportionment using a combination of aerosol and gas monitoring techniques, Atmos. Environ., 30, 2795–2809, 1996.

Tesche, T. W., Morris, R., Tonnesen, G., McNally, D., Boylan, J., and Brewer, P.: CMAQ/CMAx annual 2002 performance evaluation over the eastern US, Atmos. Environ., 40, 4906–4919, 2006.

Thurston, G. D., Ito, K., and Lall, R.: A source apportionment of US fine particulate matter air pollution, Atmos. Environ., 45, 3924–3936, 2011.

Tian, D., Hu, Y., Wang, Y., Boylan, J., Zheng, M., and Russell, A. G.: Assessment of Biomass Burning Emissions and Their Impacts on Urban and Regional $PM_{2.5}$: A Georgia Case Study, Environ. Sci. Technol., 43, 299–305, doi:10.1021/es801827s, 2009.

Unal, A., Hu, Y. T., Chang, M. E., Odman, M. T., and Russell, A. G.: Airport related emissions and impacts on air quality: application to the Atlanta International Airport, Atmos. Environ., 39, 5787–5798, 2005.

ACPD
13, 26657–26698, 2013

USEPA: Protocol for Applying and Validating the CMB Model for $PM_{2.5}$ and VOC, Report Number EPA-451/R-04-001, US Environmental Protection Agency, Research Triangle Park, NC, 2004.

Viana, M., Pandolfi, M., Minguillon, M. C., Querol, X., Alastuey, A., Monfort, E., and Celades, I.: Inter-comparison of receptor models for PM source apportionment: case study in an industrial area, Atmos. Environ., 42, 3820–3832, 2008a.

Viana, M., Kuhlbusch, T. A. J., Querol, X., Alastuey, A., Harrison, R. M., Hopke, P. K., Winiwarter, W., Vallius, A., Szidat, S., Prevot, A. S. H., Hueglin, C., Bloemen, H., Wahlin, P., Vecchi, R., Miranda, A. I., Kasper-Giebl, A., Maenhaut, W., and Hitzenberger, R.: Source apportionment of particulate matter in Europe: a review of methods and results, J. Aerosol Sci., 39, 827–849, 2008b.

Wagstrom, K. M., Pandis, S. N., Yarwood, G., Wilson, G. M., and Morris, R. E.: Development and application of a computationally efficient particulate matter apportionment algorithm in a three-dimensional chemical transport model, Atmos. Environ., 42, 5650–5659, 2008.

Watson, J. G.: Overview of receptor model principles, J. Air Pollut. Control Assoc., 34, 619–623, 1984.

Watson, J. G., Cooper, J. A., and Huntzicker, J. J.: The effective variance weighting for least-squares calculations applied to the mass balance receptor model, Atmos. Environ., 18, 1347–1355, 1984.

Watson, J. G., Chow, J. C., Lowenthal, D. H., Robinson, N. F., Cahill, C. F., and Blumenthal, D. L.: Simulating changes in source profiles from coal-fired power stations: use in chemical mass balance of $PM_{2.5}$ in the Mount Zirkel Wilderness, Energ. Fuels, 16, 311–324, 2002.

Watson, J. G., Chen, L. W. A., Chow, J. C., Doraiswamy, P., and Lowenthal, D. H.: Source apportionment: findings from the US supersites program, J. Air Waste Manage., 58, 265–288, 2008.

Wittig, A. E. and Allen, D. T.: Improvement of the Chemical Mass Balance model for apportioning – sources of non-methane hydrocarbons using composite aged source profiles, Atmos. Environ., 42, 1319–1337, 2008.

Xie, M. J., Piedrahita, R., Dutton, S. J., Milford, J. B., Hemann, J. G., Peel, J. L., Miller, S. L., Kim, S. Y., Vedal, S., Sheppard, L., and Hannigan, M. P.: Positive matrix factorization of a 32-month series of daily $PM_{2.5}$ speciation data with incorporation of temperature stratification, Atmos. Environ., 65, 11–20, 2013.

ACPD
13, 26657–26698, 2013

Xiu, A. and Pleim, J. E.: Development and application of a land-surface model. I: Application in a mesoscale meteorology model, J. Appl. Meteorol., 40, 192–209, 2001.

Yang, Y.-J., Wilkinson, J., and Russell, A. G.: Fast, direct sensitivity analysis of multidimensional photochemical models, Environ. Sci. Technol., 31, 2859–2868, 1997.

Ying, Q., Lu, J., Kaduwela, A., and Kleeman, M.: Modeling air quality during the California Regional $PM_{10}/PM_{2.5}$ Air Quality Study (CPRAQS) using the UCD/CIT Source Oriented Air Quality Model – Part II. Regional source apportionment of primary airborne particulate matter, Atmos. Environ., 42, 8967–8978, 2008.

Zheng, M., Cass, G. R., Schauer, J. J., and Edgerton, E. S.: Source apportionment of $PM_{2.5}$ in the southeastern United States using solvent-extractable organic compounds as tracers, Environ. Sci. Technol., 36, 2361–2371, 2002.

ACPD
13, 26657–26698, 2013

Table 1. Emissions source categories used in the hybrid method application.

Top tier sectors	Combustion	On-road	Non-road	Biomass-burning	Industrial and others
33 source categories	COALCMB, DIESELCMB, FUELOILCMB, LPGCMB, NAGASCMB, OTHERCMB, MEXCMB_M	ORDIESEL, ORGASOLINE	AIRCRAFT, NRDIESEL, NRFUELOIL, NRGASOLINE, NRLPG, NRNAGAS, NROTHERS, RAILROAD	AGRIBURN WILDFIRE, OPENFIRE, PRESCRBURN, LWASTEBURN, WOODFUEL, WOODSTOVE	BIOGENIC, DUST, LIVESTOCK, METALPRODUCT, MEATCOOKING, MINERALPRODUCT, SEASALT, SOLVENT, OTHERS

Note: COALCMB – coal combustion, DIESELCMB – diesel combustion, FUELOILCMB – fuel oil combustion, LPGCMB – liquid petroleum gas combustion, NAGASCMB – natural gas combustion, OTHERCMB – other fuel combustion, MEXCMB_M – Mexican combustion mix fuel, ORDIESEL – on-road diesel vehicles, ORGASOLINE – on-road gasoline vehicles, AIRCRAFT – aircraft operation in airports, NRDIESEL – non-road diesel, NRFUELOIL – non-road fuel oil, NRGASOLINE – non-road gasoline, NRLPG – non-road liquid petroleum gas, NRNAGAS – non-road natural gas, NROTHERS – non-road other fuel, RAILROAD – railroad, AGRIBURN – agricultural burn, WILDFIRE – wildfire, OPENFIRE – open fire, PRESCRBURN – prescribed fire, LWASTEBURN – lawn waste burning, WOODFUEL – wood fuel boiler combustion, WOODSTOV – woodstove and fireplace, BIOGENIC – biogenic, DUST – fugitive dust, LIVESTOCK – livestock mostly ammonia, METALPRODUCT – metal products, MEATCOOKING – meat cooking, frying, charcoal broil, MINERALPRODUCT – mineral products, SEASALT – sea salts, SOLVENT – solvents, OTHERS – others not in previous categories. See Table S2 for source classification codes grouped in each category.

ACPD
13, 26657–26698, 2013

Table 2. Calculated source impact scale factors (R_j) across 164 CSN sites, Jan 2004: mean and standard deviation.

Source	Mean	Stdev
AGRIBURN	0.702	0.334
AIRCRAFT	0.998	0.015
BIOGENIC	0.997	0.052
COALCMB	0.953	0.056
DIESELCM	1.000	0.001
DUST	0.150	0.269
FUELOILC	0.879	0.186
LIVESTOCK	0.989	0.043
LPGCMB	0.999	0.006
LWASTEBU	0.193	0.541
MEATALPR	0.738	0.224
MEATCOOK	0.817	0.305
MEXCMB_M	0.999	0.007
MINERALP	0.879	0.106
NAGASCMB	0.522	0.227
NRDIESEL	0.987	0.056
NRFUELOI	0.994	0.018
NRGASOL	0.988	0.054
NRLPG	1.000	0.003
NRNAGAS	1.000	0.001
NROTHERS	1.000	0.001
OPENFIRE	0.552	0.421
ORDIESEL	0.968	0.059
ORGASOL	0.862	0.172
OTHERCMB	0.910	0.130
OTHERS	0.521	0.222
PRESCRBU	0.961	1.122
RAILROAD	0.998	0.013
SEASALT	0.991	0.025
SOLVENT	0.895	0.163
WILDFIRE	0.836	0.256
WOODFUEL	0.904	0.184
WOODSTOVE	0.208	0.582

ACPD
13, 26657–26698, 2013

Table 3. Initial and refined modeled concentrations vs. observed across 164 CSN sites, Jan 2004: average and standard deviation.

Species	Observed		Initial		Refined	
	Avg ($\mu g\,m^{-3}$)	Stdev ($\mu g\,m^{-3}$)	Avg ($\mu g\,m^{-3}$)	Stdev ($\mu g\,m^{-3}$)	Avg ($\mu g\,m^{-3}$)	Stdev ($\mu g\,m^{-3}$)
PM_{25}	11.31	7.19	17.89	11.88	8.80	4.74
OC25	2.12	1.98	3.85	3.72	1.29	1.00
EC25	0.81	0.75	1.07	1.13	0.62	0.59
NO325	2.61	3.05	2.07	2.11	1.87	1.89
NH425	1.27	1.11	1.50	0.98	1.21	0.75
SO425	2.03	1.28	2.84	1.82	2.30	1.48
Na	7.6E–02	9.0E–02	1.2E–01	1.1E–01	3.6E–02	2.4E–02
Mg	1.4E–02	1.7E–02	2.9E–02	2.3E–02	1.2E–02	8.6E–03
Al	1.7E–02	1.9E–02	1.9E–01	1.6E–01	4.4E–02	2.8E–02
Si	8.5E–02	7.4E–02	5.6E–01	4.8E–01	1.2E–01	7.6E–02
P	4.5E–03	1.6E–03	6.9E–03	5.3E–03	2.3E–03	1.4E–03
Cl	5.3E–02	1.7E–01	4.5E–01	4.8E–01	8.7E–02	6.8E–02
K	6.8E–02	7.1E–02	5.3E–01	6.5E–01	8.1E–02	7.7E–02
Ca	4.4E–02	5.9E–02	1.9E–01	1.5E–01	5.0E–02	3.1E–02
Ti	4.1E–03	3.7E–03	2.1E–02	1.8E–02	7.1E–03	4.8E–03
V	2.4E–03	2.8E–03	1.5E–03	1.4E–03	5.2E–04	3.5E–04
Cr	2.3E–03	5.7E–03	3.4E–03	4.6E–03	1.3E–03	1.0E–03
Mn	3.6E–03	3.7E–02	5.6E–03	5.5E–03	1.5E–03	1.0E–03
Fe	6.4E–02	9.7E–02	1.6E–01	1.4E–01	4.2E–02	2.7E–02
Co	7.9E–04	4.2E–04	1.5E–04	1.3E–04	3.8E–05	2.6E–05
Ni	2.0E–03	4.8E–03	3.9E–03	5.6E–03	1.6E–03	1.2E–03
Cu	3.2E–03	4.9E–03	3.6E–03	4.8E–03	1.8E–03	1.3E–03
Zn	1.6E–02	2.9E–02	1.2E–02	1.3E–02	3.4E–03	2.5E–03
Ga	1.7E–03	9.0E–04	2.1E–05	2.2E–05	1.3E–05	9.1E–06
As	1.7E–03	1.9E–03	2.6E–04	3.6E–04	8.6E–05	7.4E–05
Se	1.8E–03	2.2E–03	1.9E–03	1.6E–03	1.3E–03	1.0E–03
Br	3.3E–03	4.2E–03	2.9E–03	2.7E–03	8.5E–04	5.6E–04
Rb	9.3E–04	5.1E–04	1.3E–03	1.1E–03	3.3E–04	2.3E–04
Sr	1.7E–03	1.2E–03	1.5E–03	1.1E–03	6.7E–04	5.0E–04
Zr	1.9E–03	1.3E–03	4.5E–04	3.7E–04	1.2E–04	7.3E–05
Mo	4.0E–03	1.7E–03	8.4E–04	1.2E–03	4.6E–04	3.5E–04
Ag	5.3E–03	3.4E–03	5.2E–04	8.1E–04	1.6E–04	1.3E–04
Cd	7.2E–03	6.3E–03	4.2E–03	1.7E–02	1.0E–03	2.2E–03
In	7.3E–03	4.2E–03	2.2E–04	2.2E–04	6.2E–05	4.4E–05
Sn	1.0E–02	5.0E–03	1.3E–03	1.2E–03	6.5E–04	4.9E–04
Sb	1.4E–02	1.1E–02	6.9E–04	8.6E–04	3.2E–04	2.3E–04
Ba	1.5E–02	1.8E–02	9.0E–03	6.7E–03	4.9E–03	3.4E–03
La	1.5E–02	1.9E–02	1.8E–03	1.8E–03	8.0E–04	5.7E–04
Ce	1.9E–02	2.4E–02	2.1E–04	3.9E–04	8.4E–05	7.8E–05
Hg	2.0E–03	9.3E–04	1.3E–05	1.0E–05	6.5E–06	4.4E–06
Pb	4.8E–03	6.1E–03	1.9E–03	2.0E–03	6.4E–04	4.6E–04

ACPD
13, 26657–26698, 2013

Table 4. January 2004 average initial and refined absolute ($\mu g\,m^{-3}$) and percentage (%) source impacts on $PM_{2.5}$ at the six sites.

Category	Atlanta				Chicago				Detroit				Los Angeles				New York				Pittsburgh			
	Init.		Refnd.		Init.		Refnd.		Init.		Refnd.		Init.		Refnd.		Init.		Refnd.		Init.		Refnd.	
AGRIBURN	0.1	0.5	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0
AIRCRAFT	2.5	11.4	2.2	18.7	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.3	0.0	0.1	0.0	0.3	0.0	0.1	0.0	0.2	0.0	0.1	0.0	0.3
BIOGENIC	0.9	4.3	0.9	7.8	0.1	0.3	0.1	0.9	0.1	0.6	0.1	1.4	1.5	3.7	1.5	9.6	0.3	1.4	0.3	2.7	0.2	1.1	0.2	1.9
COALCMB	2.5	11.8	2.3	19.8	0.8	3.7	0.8	8.6	0.9	4.4	0.8	9.6	0.2	0.5	0.2	1.2	3.0	12.9	2.5	20.8	1.9	12.7	1.6	19.6
DIESELCMB	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1
DUST	2.8	13.2	0.3	2.4	2.2	9.5	0.2	2.4	2.0	10.2	0.2	2.3	3.9	9.8	0.5	3.3	2.0	8.4	0.2	1.8	1.8	12.3	0.2	2.2
FUELOILCMB	0.9	4.1	0.7	6.4	0.7	2.9	0.5	5.0	0.5	2.7	0.4	4.7	2.8	6.9	1.3	8.7	2.6	10.9	1.8	14.9	0.4	2.9	0.4	4.3
LIVESTOCK	0.8	3.5	0.7	6.2	1.5	6.6	1.4	15.9	1.1	5.5	1.1	12.5	0.5	1.3	0.5	3.3	0.8	3.3	0.8	6.3	1.3	9.1	1.3	16.0
LPGCMB	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.3	0.0	0.1	0.0	0.3	0.6	1.6	0.6	3.9	0.0	0.1	0.0	0.2	0.0	0.1	0.0	0.2
LWASTEBURN	0.9	4.0	0.1	0.7	0.6	2.7	0.1	1.0	1.0	5.4	0.1	1.4	2.8	6.9	0.3	1.8	0.4	1.6	0.1	0.4	1.2	8.2	0.1	1.4
METALPRODUCT	0.2	0.7	0.1	0.8	3.5	15.2	0.5	5.6	0.5	2.5	0.3	3.6	0.0	0.1	0.0	0.3	0.3	1.5	0.3	2.5	0.7	5.1	0.4	4.6
MEATCOOKING	0.0	0.1	0.0	0.1	0.7	3.2	0.4	4.2	0.6	3.3	0.3	3.3	5.5	13.6	1.5	9.7	1.7	7.4	0.8	6.3	0.4	2.6	0.2	2.8
MEXCMB_M	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MINERALPRODUCT	0.2	1.1	0.2	1.7	0.3	1.4	0.2	2.6	0.2	0.8	0.1	1.7	0.8	1.9	0.5	3.3	0.2	0.7	0.1	1.1	0.3	1.8	0.2	2.3
NAGASCMB	0.8	3.7	0.3	2.2	3.1	13.6	0.8	8.7	2.2	11.3	0.8	9.0	3.5	8.7	0.9	5.8	1.4	6.0	0.6	5.1	0.7	4.8	0.3	3.7
NRDIESEL	0.5	2.4	0.5	4.2	0.6	2.6	0.6	6.3	0.7	3.5	0.6	6.9	1.3	3.1	1.2	7.9	0.9	3.9	0.9	7.3	0.6	3.8	0.5	6.2
NRFUELOIL	0.0	0.1	0.0	0.1	0.1	0.3	0.1	0.7	0.1	0.3	0.1	0.7	0.0	0.1	0.0	0.3	0.0	0.1	0.0	0.2	0.1	0.6	0.1	1.1
NRGASOLINE	0.2	1.1	0.2	2.0	0.4	1.6	0.4	3.9	0.5	2.4	0.4	5.2	0.9	2.3	0.9	5.9	0.5	2.1	0.5	3.9	0.2	1.5	0.2	2.5
NRLPG	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.2	0.0	0.1	0.0	0.2	0.0	0.1	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.1
NRNAGAS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
NROTHERS	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	–.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
OPENFIRE	0.2	0.9	0.0	0.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	1.5	6.4	0.2	1.4	0.0	0.3	0.0	0.2
ORDIESEL	0.6	2.9	0.6	4.8	0.3	1.4	0.3	3.4	0.7	3.7	0.6	7.1	0.6	1.5	0.6	3.6	0.6	2.4	0.5	4.4	0.3	1.9	0.3	3.2
ORGASOLINE	2.2	10.3	1.4	11.9	1.7	7.3	1.1	12.2	1.5	7.8	1.2	13.7	2.2	5.6	1.5	9.5	1.3	5.7	1.0	8.0	1.3	9.1	1.2	14.0
OTHERCMB	0.1	0.6	0.1	0.7	1.2	5.3	0.7	7.4	0.2	0.9	0.2	1.8	0.1	0.3	0.1	0.6	0.1	0.5	0.1	0.9	0.1	0.7	0.1	1.1
OTHERS	0.5	2.5	0.2	1.9	2.1	9.3	0.3	3.4	0.8	4.3	0.4	4.5	3.2	8.1	1.3	8.1	1.2	5.3	0.4	3.7	1.1	7.2	0.5	6.5
PRESCRBURN	0.5	2.4	0.1	0.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.3	0.2	1.0	0.2	0.9	0.1	0.7	0.0	0.1	0.0	0.1
RAILROAD	0.1	0.3	0.1	0.6	0.1	0.7	0.1	1.6	0.1	0.4	0.1	0.8	0.3	0.7	0.3	1.9	0.0	0.1	0.0	0.2	0.1	0.5	0.1	0.8
SEASALT	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
SOLVENT	0.2	1.1	0.2	1.8	0.5	2.0	0.2	2.3	0.2	1.0	0.1	1.7	0.6	1.4	0.5	3.3	0.4	1.6	0.3	2.5	0.1	0.9	0.1	1.4
WILDFIRE	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.1	0.0	0.0	0.0	0.3	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1
WOODFUEL	0.1	0.3	0.1	0.4	0.0	0.1	0.0	0.2	0.0	0.2	0.0	0.5	0.0	0.1	0.0	0.1	0.1	0.4	0.1	0.6	0.0	0.1	0.0	0.2
WOODSTOVE	3.6	16.8	0.4	3.1	2.3	10.1	0.3	2.7	5.4	28.1	0.5	6.4	8.6	21.3	0.9	5.5	3.7	16.0	0.4	3.6	1.8	12.4	0.3	3.0
Total Impacts	21.6	**100.0**	11.8	**100.0**	22.8	**100.0**	9.1	**100.0**	19.3	**100.0**	8.5	**100.0**	40.3	**100.0**	15.6	**100.0**	23.5	**100.0**	12.1	**100.0**	14.7	**100.0**	8.4	**100.0**
Observed Concentration			12.1				8.7				10.				22.3				11.7				8.7	

ACPD
13, 26657–26698, 2013

Table 5. Initial vs. refined calculated largest five contributing sources (January 2004).

Site		1st	2nd	3rd	4th	5th
Atlanta	Initial	woodstove	dust	coal combustion	aircraft	on-road gasoline
	Refined	coal combustion	aircraft	on-road gasoline	biogenic	fuel oil combustion
Chicago	Initial	metal products	natural gas combustion	woodstove	dust	others
	Refined	livestock	on-road gasoline	natural gas combustion	coal combustion	other fuel combustion
Detroit	Initial	woodstove	natural gas combustion	dust	on-road gasoline	livestock
	Refined	on-road gasoline	livestock	coal combustion	natural gas combustion	on-road diesel
Los Angeles	Initial	woodstove	meat cooking	dust	natural gas combustion	others
	Refined	meat cooking	biogenic	on-road gasoline	fuel oil combustion	others
New York	Initial	woodstove	coal combustion	fuel oil combustion	dust	meat cooking
	Refined	coal combustion	fuel oil combustion	on-road gasoline	non-road diesel	meat cooking
Pittsburgh	Initial	coal combustion	woodstove	dust	on-road gasoline	livestock
	Refined	coal combustion	livestock	on-road gasoline	others	non-road diesel

ACPD
13, 26657–26698, 2013

Table 6. January 2004 average initial and refined source impacts on $PM_{2.5}$ at the six sites: regrouped to 13 sources.

Category (each includes both primary and secondary impacts)	Atlanta				Chicago				Detroit				Los Angeles				New York				Pittsburgh			
	Init.		Refnd.		Init.		Refnd.		Init.		Refnd.		Init.		Refnd.		Init.		Refnd.		Init.		Refnd.	
	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank	%	Rank
LDGV[a]	11.3	6	13.9	4	8.9	6	16.2	2	10.2	4	19.0	2	7.8	6	15.5	2	7.9	6	11.9	5	10.6	5	16.6	3
HDDV[a]	5.6	7	9.7	5	4.7	7	11.4	3	7.6	6	15.1	3	5.4	8	13.5	3	6.5	8	12.0	4	6.3	6	10.4	4
DUST	13.2	2	2.4	8	9.5	5	2.4	11	10.2	5	2.3	10	9.8	4	3.3	9	8.4	5	1.8	10	12.3	4	2.2	11
BURN[a]	24.9	1	5.5	7	13.0	4	4.2	9	33.8	1	8.6	6	28.7	1	9.0	5	25.4	1	6.8	6	21.1	1	5.1	5
COALCMB	11.8	4	19.8	1	3.7	8	8.7	4	4.4	7	9.7	4	0.5	10	1.2	10	12.9	2	20.9	1	12.7	3	19.8	2
MEATCOOKING	0.1	12	0.1	12	3.2	9	4.2	8	3.3	8	3.3	9	13.6	3	9.8	4	7.4	7	6.3	7	2.6	10	2.8	9
SEASALT	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13	0.0	13
METALPRODUCT	0.7	11	0.8	11	15.2	2	5.7	6	2.5	10	3.7	8	0.1	11	0.3	12	1.5	10	2.5	9	5.1	7	4.6	6
MINERALPRODUCT	1.1	10	1.7	10	1.4	11	2.6	10	0.8	11	1.7	11	1.9	9	3.3	8	0.7	11	1.1	11	1.8	11	2.4	10
NATURALGAS[a]	3.7	9	2.2	9	13.6	3	8.7	5	11.3	3	9.1	5	8.7	5	5.9	7	6.1	9	5.1	8	4.8	8	3.8	8
FUELOIL[a]	4.2	8	6.4	6	3.1	10	5.1	7	3.1	9	4.7	7	7.1	7	8.7	6	11.0	4	15.0	3	3.6	9	4.4	7
AIRCRAFT	11.4	5	18.7	2	0.0	12	0.1	12	0.1	12	0.3	12	0.1	12	0.3	11	0.1	12	0.2	12	0.1	12	0.3	13
AllOthers[a]	12.0	3	18.6	3	23.7	1	30.7	1	12.6	2	22.6	1	16.4	2	29.2	1	12.2	3	16.4	2	19.2	2	27.6	1

[a] Regrouped sources: LDGV – light-duty gasoline vehicles, merged from NRGASOLINE and ORGASOLINE; HDDV – heavy-duty diesel vehicles, merged from NRDIESEL, ORDIESEL, RAILROAD and DIESELCMB; BURN – vegetative burning, merged from AGRIBURN, LWASTEBURN, OPENFIRE, PRESCRBU, WILDFIRE, WOODFUEL and WOODSTOVE; NATURALGAS – merged from NAGASCMB and NRNAGAS; FUELOIL – merged from FUELOILC and NRFUELOI; AllOthers – merged from the leftover hybrid sources: BIOGENIC, LIVESTOCK, LPGCMB, MEXCMB_M, NRLPG, NROTHERS, OTHERCMB, OTHERS and SOLVENT.

ACPD
13, 26657–26698, 2013

Table 7a. Refined source impacts results regrouped to 13 primary sources and compared to results from using RM methods: Atlanta, Chicago and Detroit.

Metro area (site ID)	Atlanta (130890002)				Chicago (170310076)				Detroit (261630001)			
Study method	Hybrid		CMB-LGO (This study)		Hybrid		CMB (Rizzo and Scheff, 2007)		Hybrid		PMF (Gildemeister et al., 2007)	
Period of measurements	Jan 2004		Jan 2004		Jan 2004		2001–2003		Jan 2004		Dec 2000–Apr 2005	
Source (primary and secondary impacts separated)	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%
LDGV[a]	0.45	3.9	1.39	13.7	0.38	4.2	4.8	31	0.60	7.1	2.53	15.0
HDDV[a]	1.20	10.2	0.59	5.9	0.83	9.2			1.19	14.0	0.67	4.2
DUST[a]	0.28	2.4	0.18	1.8	0.22	2.4	0.39	2	0.19	2.3	1.29	8.0
BURN[a]	0.60	5.1	1.06	10.4	0.35	3.9	1.71	11	0.68	8.1	0.51	3.2
COALCMB[a]	0.64	5.4	0.01	0.1	0.31	3.4	0.19	1	0.25	3.0		
MEATCOOKING[a]	0.01	0.1			0.38	4.2			0.27	3.2		
SEASALT[a]	2.0E–3	1.7E–2			1.8E-4	1.9E–3	0.21[b]	1	1.9E–4	2.3E–3	0.57[b]	4.0
METALPRODUCT[a]	0.06	0.5			0.41	4.5	0.31	2	0.20	2.4	0.51	3.2
MINERALPRODUCT[a]	0.12	1.0			0.19	2.1			0.10	1.2		
NATURALGAS[a]	0.19	1.6			0.52	5.8			0.58	6.8		
FUELOIL[a]	0.09	0.7			0.15	1.6			0.10	1.2		
AIRCRAFT[a]	2.11	17.9			0.01	0.1			0.02	0.2		
AllOthers[a]	0.56	4.7			0.70	7.8			0.45	5.3		
AMSULFT	4.06	34.5	3.45	34.1	2.07	22.8	4.79	31	2.46	29.1	4.99	31.1
AMNITR	0.84	7.2	2.25	22.3	2.47	27.3	3.18	20	1.26	14.8	4.49	28.0
OTHROC	0.56	4.8	1.19	11.7	0.07	0.8			0.10	1.2		
Total Impacts	11.78	100.0	10.11	100.0	9.06	100.0	15.58	99	8.46	100.0	15.56	96.7
Modeled Concentration ($\mu g\,m^{-3}$)	14.69		10.11		10.27		15.58		10.15		15.56	
Observed Concentrations ($\mu g\,m^{-3}$)	12.07		12.07		8.68				9.95			

ACPD
13, 26657–26698, 2013

Table 7b. Refined source impacts results regrouped to 13 sources and compared to results from using RM methods: Los Angeles, New York and Pittsburgh.

Metro area	Los Angeles (060658001)				New York (360050083)				Pittsburgh (420030008)			
Study method	Hybrid		CMB (Pham et al., 2008)		Hybrid		PMF (Coutant et al., 2003)		Hybrid		PMF (Maranche, 2006)	
Period of measurements	Jan 2004		Apr 2004–Mar 2005		Jan 2004		3 Sep 2000–29 Jan 2002		Jan 2004		Jul 2003–Aug 2005	
Source (primary and secondary impacts separated)	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%	$\mu g\,m^{-3}$	%
LDGV[a]	0.93	6.0	0.85	3.7	0.72	6.0	2.5	15.5	0.37	4.5	1.37	9.5
HDDV[a]	1.87	12.0	2.54	11.1	1.56	13.0			0.82	9.9	0.68	4.7
DUST[a]	0.52	3.3	0.78	3.4	0.21	1.8	1.0	6.0	0.18	2.2	1.18	8.2
BURN[a]	1.27	8.2	0.38	1.6	0.75	6.2			0.39	4.8	2.4[d]	16.7
COALCMB[a]	0.03	0.2			1.32	10.9			0.87	10.5		
MEATCOOKING[a]	1.51	9.7	1.44	6.3	0.75	6.2			0.23	2.8		
SEASALT[a]	3.2E–3	2.0E–2	1.38	6.0	2.0E–3	1.7E–2	0.3	1.9	2.0E–4	2.5E–3		
METALPRODUCT[a]	0.02	0.1			0.09	0.8			0.27	3.3		
MINERALPRODUCT[a]	0.39	2.5	0.71	3.1	0.09	0.8			0.14	1.7		
NATURALGAS[a]	0.68	4.4			0.49	4.1			0.23	2.7		
FUELOIL[a]	0.80	5.1	0.27	1.2	0.62	5.2	1.2	7.6	0.07	0.8	0.45[e]	3.1
AIRCRAFT[a]	0.01	0.1			0.02	0.2			0.02	0.3		
AllOthers[a]	1.93	12.4			0.71	5.9	1.8[c]	11.3	0.38	4.6		
AMSULFT	2.47	15.9	4.51	19.7	4.24	35.2	5.3	32.9	3.10	37.4	5.49	38.2
AMNITR	2.32	14.9	10.08	44.0	0.32	2.6	4.1	25.4	1.10	13.3	2.81	19.5
OTHROC	0.79	5.1			0.16	1.3			0.12	1.4		
Total Impacts	15.56	100.0	22.93	100.0	12.05	100.0	16.1	100.0	8.31	100.0	14.4	100.0
Modeled Concentrations ($\mu g\,m^{-3}$)	16.49		22.93		13.62		16.1		8.64		14.4	
Observed Concentrations ($\mu g\,m^{-3}$)	22.35		23.54		11.70				8.71			

[a] Primary impacts only, secondary portion of the impacts are removed from these sources and merged into the secondary sources: AMSULFT – ammonium sulfate plus ammonium bisulfate, AMNITR – ammonium nitrate, and OTHROC – secondary organic carbon. [b] Road salts. [c] Industrial. [d] Burning and cooking. [e] Incinerator.

ACPD
13, 26657–26698, 2013

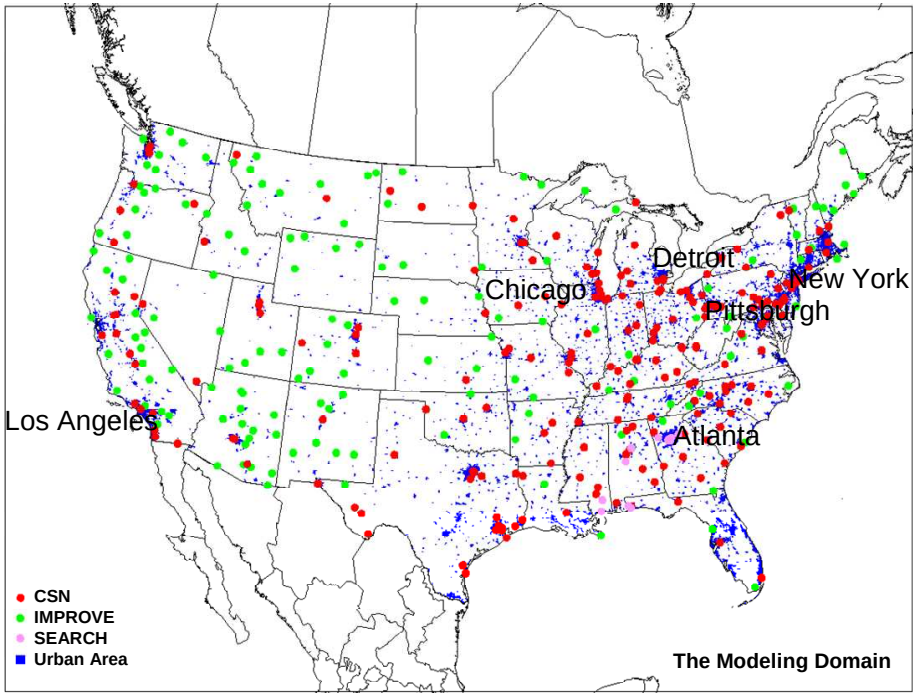

Fig. 1. Modeling domain and monitoring sites.

ACPD
13, 26657–26698, 2013

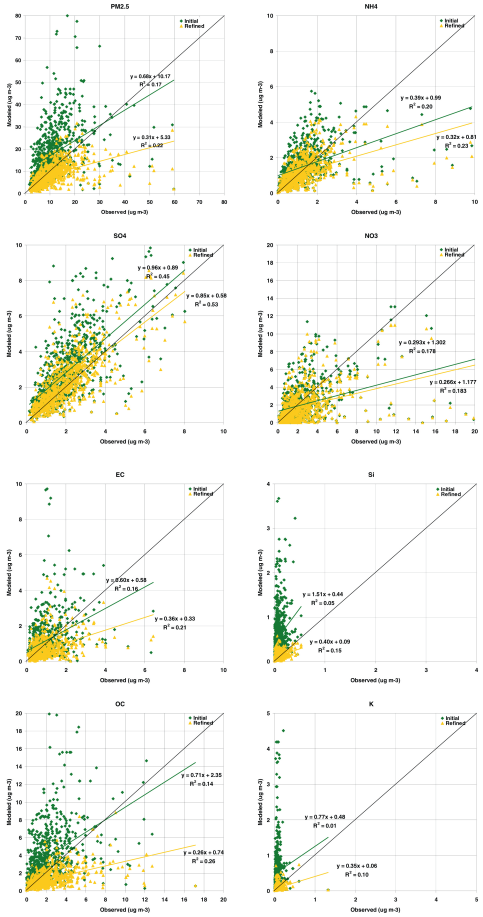

Fig. 2. Scatter plots of initial and refined concentration predictions against observations for $PM_{2.5}$ total mass and select components and elements.

ACPD
13, 26657–26698, 2013

A hybrid approach to source apportionment

Y. Hu et al.

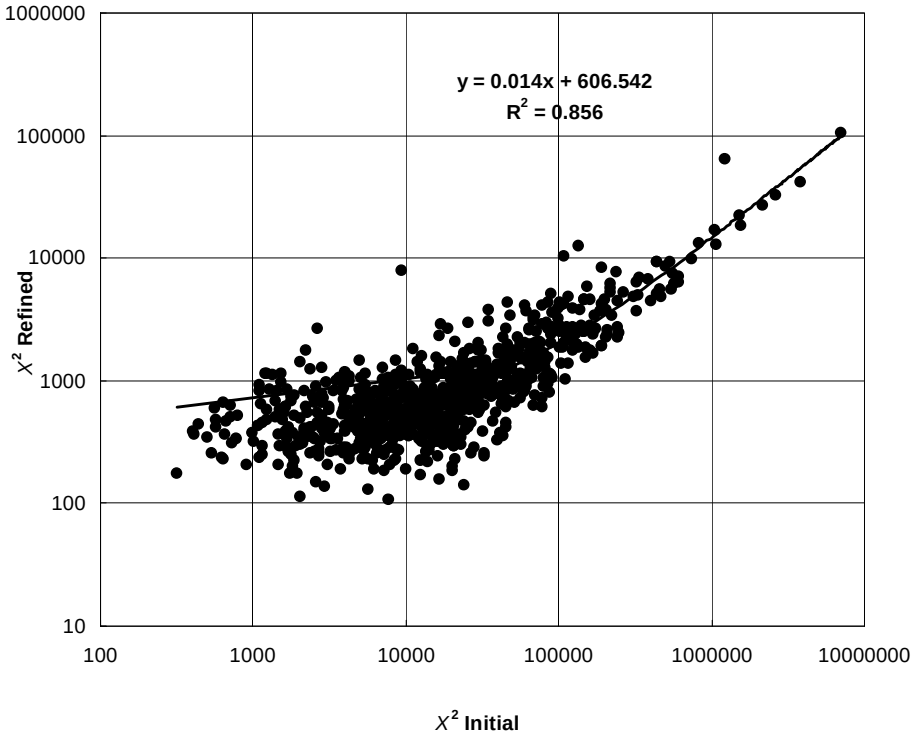

Fig. 3. Refined $\chi^2_{c,\,refnd}$ vs. initial $\chi^2_{c,\,init}$ (in logarithmic scales) for each measurement day during January 2004 at 164 CSN sites.

ACPD
13, 26657–26698, 2013