Development of PM$_{2.5}$ source impact spatial fields using a hybrid source apportionment air quality model

C. E. Ivey$^1$, H. A. Holmes$^2$, Y. T. Hu$^1$, J. A. Mulholland$^1$, and A. G. Russell$^1$

$^1$Georgia Institute of Technology, Atlanta, Georgia, USA
$^2$University of Nevada Reno, Reno, Nevada, USA

Received: 11 December 2014 – Accepted: 22 December 2014 – Published: 29 January 2015
Correspondence to: C. E. Ivey (sunni.ivey@gmail.com)
Published by Copernicus Publications on behalf of the European Geosciences Union.
Abstract

An integral part of air quality management is knowledge of the impact of pollutant sources on ambient concentrations of particulate matter (PM). There is also a growing desire to directly use source impact estimates in health studies; however, source impacts cannot be directly measured. Several limitations are inherent in most source apportionment methods, which has led to the development of a novel hybrid approach that is used to estimate source impacts by combining the capabilities of receptor modeling (RM) and chemical transport modeling (CTM). The hybrid CTM-RM method calculates adjustment factors to refine the CTM-estimated impact of sources at monitoring sites using pollutant species observations and the results of CTM sensitivity analyses, though it does not directly generate spatial source impact fields. The CTM used here is the Community Multi-Scale Air Quality (CMAQ) model, and the RM approach is based on the Chemical Mass Balance model. This work presents a method that utilizes kriging to spatially interpolate source-specific impact adjustment factors to generate revised CTM source impact fields from the CTM-RM method results, and is applied to January 2004 over the continental United States. The kriging step is evaluated using data withholding and by comparing results to data from alternative networks. Directly applied and spatially interpolated hybrid adjustment factors at withheld monitors had a correlation coefficient of 0.89, a linear regression slope of 0.83 ± 0.02, and an intercept of 0.14 ± 0.02. Refined source contributions reflect current knowledge of PM emissions (e.g., significant differences in biomass burning impact fields). Concentrations of 19 species and total PM$_{2.5}$ mass were reconstructed for withheld monitors using directly applied and spatially interpolated hybrid adjustment factors. The mean concentrations of total PM$_{2.5}$ for withheld monitors were 11.7(±8.3), 16.3(±11), 8.59(±4.7), and 9.20(±5.7) µg m$^{-3}$ for the observations, CTM, directly applied hybrid, and spatially interpolated hybrid predictions, respectively. Results demonstrate that the hybrid method along with a spatial extension can be used for large-scale, spatially resolved source apportionment studies where observational data are spatially and temporally
limited. Data withholding also provides an estimate of method uncertainty. Species concentrations were reconstructed using spatial hybrid results, and the error relative to observed concentrations was greatly reduced as compared to CTM-simulated concentrations.

1 Introduction

Variations in ambient pollutant species concentrations, including particulate matter (PM) and gases, are correlated with health outcomes such as lower birth weight (Darrow et al., 2011; Wang et al., 1997), higher occurrences of bradycardia and central apnea (Campen et al., 2001; Peel et al., 2011); decreased peak expiratory flows and increased respiratory symptoms in non-smoking asthmatics (Peters et al., 1997); and all-cause, lung cancer, and cardiopulmonary mortality (Pope et al., 2002). Additionally, nanotoxicological studies report that particle uptake by cells and entry into blood and lymph leads to oxidative stress in sensitive areas of the body such as lymph nodes, bone marrow, and the spleen (Oberdorster et al., 2005). Recently, in a study on the global burden of disease, of the 67 risk factors studied, exposure to ambient particulate matter pollution was the ninth highest risk factor leading to disability-adjusted life years (Lim et al., 2012). Many past epidemiological studies focused on associating PM mass (e.g., PM$_{2.5/10}$: PM with aerodynamic diameters less than 2.5 or 10 µm) with the health outcomes, as opposed to individual species or the sources of the PM due to limited data availability or difficulties in quantifying source impacts. Epidemiological studies are examining the associations between individual species and health outcomes using data from ground observation networks, such as the Chemical Speciation Network (CSN) and the Southeastern Aerosol Research and Characterization Network (SEARCH) (Dominici et al., 2010; Samet et al., 2000; Sarnat et al., 2008; Tolbert et al., 2007). It is of further interest to determine the degree to which individual sources are influencing health events and to link human exposure and subsequent adverse impacts to sources and multi-pollutant mixtures (Laden et al., 2000; Thurston
Attributing individual component concentrations and the overall mixture of observed air pollution to specific sources, and then linking those sources with adverse health impacts is challenging. Typically, receptor modeling (RM) is used to generate source apportionment (SA) results for epidemiological studies because longer time series are required (e.g., greater than two years) (Sarnat et al., 2008).

Several SA models have been developed to quantify source impacts on pollutant concentrations. Each model has its own unique characteristics and associated uncertainties (Balachandran et al., 2012; Seigneur et al., 2000). In an effort to improve the spatial and temporal resolution of SA data, chemical transport models (CTM) have been adapted to estimate emission impacts on pollutant concentrations. This work utilizes a hybrid CTM-RM method to provide spatial fields of source impacts for use in detailed health-related, spatiotemporal analyses (e.g., Sarnat et al., 2008).

The goal of this study is to create spatial fields of source impacts by spatially interpolating source impact adjustment factors (ratios, or $R$’s) and then applying those adjustments to CTM source impact fields. $R$’s are generated by a hybrid CTM-RM SA approach that integrates observational data and results from a CTM to calculate an emission-based adjustment of source impacts at receptor locations (Hu et al., 2014). Kriging is employed to generate spatial fields of $R$’s for 33 emissions sources. The spatial fields of adjustment factors are applied to original source impact fields to produce hybrid-adjusted source impact and species concentration fields for the continental US. The adjustments can also be interpolated in time to adjust source impact fields on days when speciated observations are not available. The performance of the spatial extension is evaluated by performing data withholding and by comparing results to observations from other monitoring networks. The hybrid CTM-RM method, along with the spatial extension, provides air quality data fields for health studies that require spatially-resolved exposure metrics. This approach can also be used to assist air quality planners in developing state implementation plans (SIPs) and assessing exceptional events, such as wildland fires.
2 Data and methods

2.1 Data

Observational data during January 2004 from 189 monitors in the Chemical Speciation Network (CSN) were used for model development and evaluation (Fig. 1). Data were obtained on one in every three or six days in January 2004 for a total of 9 days (e.g., 4, 7, 10...28 January), which led to varying sample sizes for each observation day. CSN monitor measurements include total PM$_{2.5}$, organic and elemental carbon, ions, and 35 metals. CSN monitors tend to be located in more densely populated areas such as urban and suburban areas, and data are more associated with high-population emissions sources such as mobile and cooking sources. Speciated PM$_{2.5}$ data are also available from the SEARCH (Hansen et al., 2003, 2006) and IMPROVE (Chow et al., 1993) networks, and those data were used for further model evaluation. The SEARCH network includes eight monitors in the southeastern US, configured as urban/rural pairs. IMPROVE monitors are mainly located in pristine locations such as national parks and wilderness areas. Thirty-eight IMPROVE monitors in the eastern US were used for model evaluation. Monitors in the eastern US were used due to their closer proximity with urban monitoring sites (e.g., less than 50 km), as opposed to western IMPROVE sites which are much more spatially sparse. Additionally, modeled processes have higher uncertainty for the western US due to complex terrain and meteorology, leading to added bias in the observation and model comparison (Baker et al., 2011).

2.2 CTM-RM hybrid method

This study utilizes a hybrid SA method that combines techniques of both CTMs and RMs to generate adjustment factors (symbolized by $R$) that improve source impact estimates. Hu et al. (2014) describe the hybrid approach in detail, but it is briefly summarized here. First, gridded concentrations and emissions sensitivities of PM$_{2.5}$...
species are generated using CMAQ-DDM (CMAQ v. 4.7 or 5.0.2) (Byun and Schere, 2006; Dunker, 1981, 1984; Napelenok et al., 2006) with traditional inputs for emissions (Sparse Matrix Operator Kernel Emissions, SMOKE, CEP, 2003), meteorology (Fifth-Generation PSU/NCAR Mesoscale Model, MM5, Grell et al., 1994), and terrain (Pleim and Xiu, 1995; Xiu and Pleim, 2001). CMAQ-DDM model sensitivities to emissions are designated as the original (base case) source impacts ($SA_{i,j}^{base}$) for species $i$ and source $j$. Next, the original source impacts, receptor observations, and uncertainties are used as inputs to the objective function (Eq. 1) of the hybrid SA model.

$$\chi^2 = \sum_{i=1}^{N} \left[ \left( \frac{c_{i}^{obs} - c_{i}^{sim}}{\sigma_{i,obs}^2 + \sigma_{i,SP}^2} \right) - \sum_{j=1}^{J} SA_{i,j}^{base} (R_j - 1) \right]^2 + \Gamma \sum_{j=1}^{J} \frac{\ln(R_j)^2}{\sigma_{\ln(R_j)}^2}$$  (1)

where the adjustment factors $R_j$ are optimized by minimizing the objective function, $\chi^2$. The terms $c_{i}^{obs}$ and $c_{i}^{sim}$ represent the observed and CMAQ-simulated concentrations, respectively; $\Gamma$ weights the amount of change in source impact. Uncertainties in observation measurement ($\sigma_{i,obs}$), source profiles ($\sigma_{i,SP}$), and source strength ($\sigma_{\ln(R_j)}$) are also included in the model. The objective function is minimized by using a nonlinear optimization approach known as sequential quadratic programming (Fletcher, 1987; Gill et al., 1981). The function is modeled using a ridge regression structure, as demonstrated by the second term, and uses an effective variance approach to balance model outputs (Watson et al., 1984).

Source profiles are derived from the information provided by Reff et al. (2009). Uncertainties in the first term of the objective function serve as effective variances of the numerator and are specified for each species $i$. Finally, $R_j$ are applied to $SA_{i,j}^{base}$ to adjust original source impact estimates (Eq. 2) and reconstruct simulated concentrations.
\( c_{i}^{\text{adj}} \) at receptors to more closely reflect observations (Eq. 3).

\[
SA_{i,j}^{\text{adj}} = R_{j}SA_{i,j}^{\text{base}}
\]

(2)

\[
c_{i}^{\text{adj}} = c_{i}^{\text{sim}} + \sum_{j=1}^{J} SA_{i,j}^{\text{base}}(R_{j} - 1)
\]

(3)

Given that many of the source profiles are similar between categories such that colinearities are present, the variation of the \( R_{j} \)'s are constrained to \( 0.1 \leq R_{j} \leq 10 \).

The hybrid CTM-RM method produces results that more closely reflect observations than the original CTM results, which are often biased (Hu et al., 2014). It accounts for more known source categories than traditional RM approaches (e.g., 33 vs. 6), and it links sources and observations both temporally and spatially. Additionally, the hybrid CTM-RM method generates estimates of the uncertainty in source impact predictions and identifies potential errors in source strength and composition. One limitation of the hybrid CTM-RM method is that results are only available at receptor locations when observations are available, limiting its spatial and temporal scope. In this paper, the spatial hybrid method is presented and evaluated, and it extends the benefits of the hybrid CTM-RM method through spatial interpolation.

### 2.3 Development of spatiotemporal fields

Spatial and temporal source impact fields can be developed by combining the hybrid CTM-RM method and geostatistical techniques (see Fig. S1 in the Supplement for a flow diagram of the methods). Hybrid-generated \( R_{j} \) were spatially interpolated for each observation day using kriging to generate spatial fields of source impact adjustment factors. Matlab\textsuperscript{©} (v. 7.14.0.739) was used to perform all geostatistical and optimization calculations.

Daily-averaged spatial fields of CMAQ-DDM source impacts are adjusted by grid-by-grid multiplication of the original fields by the corresponding adjustment factor field,
resulting in spatial fields of hybrid-adjusted source impacts that are available every third day, as are observations. In later work, source impact fields for intervening periods are developed by interpolation of the $R_j$’s. Temporally interpolating $R_j$ and then applying those adjustments to simulated source impact fields is preferred over simply interpolating the 1-in-3 day hybrid-adjusted source impact fields because temporally interpolating adjusted source impacts would smooth the fields and the day-specific spatial and temporal variability in the emissions and meteorology captured by the CTM would be lost.

### 2.4 Method evaluation

Performance of the spatial extension was evaluated using a data withholding approach for which 10% of CSN monitors with associated hybrid data were randomly removed from the data set for each observation day ($N = 75$ total removed points). The remaining 90% were used to fit the variogram models used for kriging to produce spatial fields of $R_j$. Spatially interpolated $R_j$ values were extracted for grids containing the withheld monitors. The hybrid optimization is directly applied to withheld receptors to assess the performance of the kriging model. Concentrations are reconstructed using Eq. (3) and the spatially interpolated adjustment factors. Finally, the original CMAQ-DDM, directly applied hybrid (CTM-RM), and spatial hybrid (SH) concentrations are compared to observed concentrations at withheld receptors. Linear regression was used to assess correlations between observations and modeled concentrations. Results were also evaluated at SEARCH and IMPROVE locations, where CMAQ-DDM and hybrid concentrations were compared to observations. Note that the application of the hybrid method, as conducted here, did not include SEARCH and IMPROVE data, and CTM-RM/SH results are independent of those observation data. Also note that 41 species, including total PM, were used for spatial field construction, but only results for 20 species are presented for comparison of CSN results and 15 species for SEARCH and IMPROVE results, as measurements for some trace metals are seldom above measurement detection limit.
3 Results

3.1 Hybrid adjustment factors

After application of the hybrid CTM-RM method at each of the monitors and spatial extension over the continental US, it was found that while many of the source impacts were adjusted relatively little (i.e., $R \approx 1.0$), dust- and biomass burning-related impacts were often biased high in the original CMAQ-DDM simulation and therefore considerably reduced. Spatial fields of hybrid adjustment factors are presented for dust, on-road diesel and gasoline combustion, and woodstove sources (Fig. 3); mean and median values for $R_j$’s for all sources are presented in the Supplement (Table S2) as well as source-specific probability distribution functions (Fig. S3). Typically, $R_j$’s were less than one for dust and woodstove impacts, indicating a high bias in those source impacts in the base CMAQ-DDM simulations. Spatial field values for on-road diesel and gasoline combustion $R_j$’s are generally near one over most of the US, though $R_j$’s for those sources tend be below one in the southeastern region of the US. The distribution of all $R_j$ values was approximately lognormal, and an analysis was performed to determine whether log-transformation of $R_j$ prior to the kriging step was necessary to reduce bias in source impact and concentration estimates. From the analysis it was determined that lognormal transformation of $R_j$ values was not necessary, as little difference was observed in reconstructed concentrations and source impact fields as a result of the transformation.

3.2 Refined spatial fields

Base CMAQ-DDM spatial fields were refined by applying kriged fields of hybrid adjustment factors (Fig. 2). Sources with high occurrences ($\sim>50\%$) of adjustment factors less than 1 include biomass burning, metals processing, and natural gas combustion; refined spatial fields are presented in the Supplement (Figs. S5–S7). Biomass burning includes impacts from agricultural burning, lawn waste burning, open fires, prescribed
burning, wildfires, woodfuel burning, and woodstoves. The SH method significantly decreases impacts from biomass burning on 4 and 22 January for the eastern US and for portions of the west coast (Fig. S5), largely driven by the observed potassium and OC levels being lower than simulated levels. On average, CMAQ-DDM simulated levels were a factor of 3.1 ± 1.1 times higher than SH values on 4 January, and a factor of 5.2 ± 1.0 times higher on 22 January. Metal processing impacts were reduced for areas highly impacted by smelting and metal works industries including the Ohio River Valley and Mid-Atlantic regions. On average, the base simulated values were 21 ± 21% higher than SH values on 4 January, and 25 ± 21% higher on 22 January for metal processing impacts. Natural gas combustion source impacts (area and point sources only) were reduced for the southeastern US, the Ohio River Valley Region, the Gulf Coast states, and parts of California and Texas. On average, simulated levels were 35 ± 14% higher than SH values on 4 January, and 72 ± 28% higher on 22 January for natural gas combustion impacts.

### 3.3 Refined source impacts

Average source contributions to PM$_{2.5}$ at withheld CSN monitors were ranked from largest to smallest for the base CMAQ-DDM (without any adjustment), directly applied hybrid (CTM-RM, available at the monitors), and interpolated spatial hybrid (SH) results (Table 1). The top three sources were woodstoves, dust, and livestock emissions for base CMAQ-DDM simulations, the latter capturing the influence of ammonia emissions on the formation of nitrate. The livestock category includes impacts from other agricultural/farming activities. For CTM-RM and SH results, woodstove (10th for both) and dust (13th for CTM-RM, 14th for SH) were ranked much lower. Livestock emissions were ranked 1st for both the CTM-RM and SH hybrid applications. Source ranking for open fires was reduced from 10th (CMAQ-DDM) to 20th for both the CTM-RM and SH applications. The fuel oil source impact ranking increased from 12th for the base CMAQ-DDM simulation to 6th and 5th for CTM-RM and SH results, respectively.
3.4 Refined concentration estimates

In the method evaluation, 10% of the CSN monitors were withheld, and source impacts were calculated at the monitor location using the SH method for comparison to the observed species concentrations. The mean concentrations of total PM$_{2.5}$ for withheld monitors were 11.7(±8.3), 16.3(±11), 8.59(±4.7), and 9.2(±5.7) µg m$^{-3}$ for the observations, CMAQ-DDM, CTM-RM, and SH estimations, respectively. Levels of crustal metals (Al, Ca, Fe, and Si), K, and Cl were biased very high in the base CMAQ-DDM simulation, oftentimes an order of magnitude greater than observations. SH concentrations of metals were closer to the CSN observations. Error in simulated (sim) concentrations is calculated using Eq. (4):

$$\text{Error} = \frac{1}{N} \sum_{i=1}^{N} \frac{|\text{obs}_i - \text{sim}_i|}{\text{obs}_i}$$

(4)

For example, the error was 295 and 139% for CMAQ-DDM vs. observations and SH vs. observations, respectively for vanadium; and 1340 and 326% for CMAQ-DDM vs. observations and SH vs. observations, respectively for manganese. Mean observed and modeled concentrations for total PM mass, five major species, and other metals can be found in the Supplement (Table S3).

3.5 Spatial extension evaluation

CTM-RM and SH species concentrations and adjustment factors at withheld monitors were compared to evaluate the spatial interpolation that was performed using kriging. For each observation day (9 days), 10% of available monitors were randomly withheld, resulting in a total of 2475 data pairs (75 withheld sites × 33 source categories). Five outlying data pairs (< 0.5%) were removed from this regression ($R_{\text{Hyb}}>2$). The remaining CTM-RM and SH factors had a Pearson correlation coefficient of 0.89, a linear
regression slope of $0.83 \pm 0.02$, and an intercept of $0.14 \pm 0.02$ (Fig. 4). Root mean square errors (RMSE) were calculated for the adjustment factors by source (Eq. 5):

$$\text{RMSE}_j = \sqrt{\frac{\sum_{i=1}^{N} (R_{iCTM-RM}^j - R_{iSH}^j)^2}{N}}, \quad j = 1 \ldots J \text{ sources, } N = 75 \text{ sites}$$

RMSEs for all sources were less than 0.4, with the exception of RMSEs for lawn waste burning, prescribed burning, and woodstoves (Table S2). This is expected given the uncertainty in the burn emissions. Mean and median $R_j$’s for each source are also calculated, and values are within 20% for most source (Table S2). The overall mean $R_j$ at withheld monitors for all sources for CTM-RM and SH adjustment factors was 0.84 and 0.83 respectively, indicating a high bias in CMAQ-DDM overall, as expected from the base model performance evaluation ($\text{PM}_{2.5}$ was biased approximately 40% high).

Probability distributions were examined for CTM-RM and SH adjustment factors for each source (Fig. S3). Cumulative distributions of adjustment factors were highly correlated for each source. The spatial interpolation captured CTM-RM trends for sources dominated by adjustment factors near 0.1, such as dust, lawn waste burning, prescribed burning, and woodstoves, though did not capture all of the extremely low adjustments (e.g., meat cooking in some locations). Sources that found little adjustment ($R_j = 1$) include aircraft, diesel combustion (stationary sources), fuel oil burning, Mexican combustion, non-road liquid petroleum gasoline combustion, and seasalt, and were well captured by the spatial extension, as demonstrated by nearly identical PDFs.

### 3.6 SEARCH and IMPROVE comparison

The spatial extension of the hybrid method was further evaluated by comparing simulated concentrations to independent data from the SEARCH and IMPROVE networks (Tables S4–S5 and S7–S8). The mean concentrations over observation days were compared, as well as regression statistics for observations vs. modeled results. For
the SEARCH network \((N = 8)\), 15 species were compared to observations. Error in mean concentrations for crustal elements was significantly decreased (CMAQ-DDM and SH): Al, 2203 to 540 \%; Si, 1228 to 271 \%; K, 365 to 61 \%; Ca, 402 to 61 \%; Fe, 260 to 3 \%; Cu, 231 to 38 \%; and Se, 63 to 25 \%. For the IMPROVE network \((N = 38)\), errors in mean concentrations for crustal elements were also significantly decreased: Al, 704 to 24 \%; Si, 371 to 24 \%; K, 599 to 48 \%; Ca, 361 to 36 \%; Fe, 334 to 18 \%; Cu, 186 to 57 \%; and Se, 22 to 11 \%. The large remaining errors stem from the source profiles leading to some elements being biased consistently high, others low. Further work to optimize source profiles can reduce residual errors.

4 Discussion

CTM-RM and SH adjustment factors are well-correlated for each source category for withheld monitors. Kriging captures the original distributions of adjustment factors for each source as simulated by the CTM-RM analysis. Spatial interpolation of hybrid adjustment factors is well-suited for providing spatial fields of refined source impacts for multiple distinct source categories. Kriged adjustment factors also provide refined source contribution estimates that are more consistent with observations than CMAQ-DDM.

The order of source contributions at withheld monitors for the CTM-RM and SH applications compared well when ranked by magnitude in descending order, though often differed greatly from the base CMAQ-DDM application. The difference in rankings between CTM-RM and SH contributions was, at most, 2 positions. The top three sources of primary PM\(_{2.5}\) for January 2004, based on source emissions, were dust, woodstoves, and coal combustion, estimated at 1275, 5301, and 3407 metric tons per day, respectively (Table S1). However, uncertainties associated with dust and woodstove emissions are much higher than most of the other sources, a factor of 10 and 5 respectively (Hanna et al., 1998, 2001; Hu et al., 2014). This uncertainty is driven,
in part, by source variability. This large uncertainty and potential bias is reflected in the large shift in rankings for dust and woodstove source contributions to PM$_{2.5}$. Other biomass burning sources such as lawn waste burning and wildfires have similarly large emissions uncertainties, and likely large temporal variabilities, and their rankings were also significantly decreased. Coal combustion, which includes secondary formation of sulfate, remains in the top three sources for average hybrid PM$_{2.5}$ source contributions at withheld monitors, as its emissions uncertainties are fairly low due to the availability of continuous emission monitoring data.

Secondary formation processes increase the impact of coal combustion, biogenic and livestock emissions relative to their initial primary PM contribution. Coal combustion was ranked 9th, 4th, and 3rd for CMAQ-DDM, CTM-RM, and SH hybrid contributions, respectively. January 2004 primary PM emissions estimates for biogenic and livestock were ranked 33 and 31, respectively. However, CMAQ-DDM, CTM-RM, and SH hybrid contributions ranked both sources significantly higher (biogenic rankings: 14, 11, and 9, respectively; livestock rankings: 3, 1, and 1, respectively). Although primary PM$_{2.5}$ emissions from these sources are not large, secondary processes lead to high source contributions. Biogenic sources emit large quantities of volatile organic compounds which go on to form secondary organic aerosols. Livestock emissions, (i.e., gaseous ammonia), react with sulfate, nitrate, and other acids to form ammonium salts. The hybrid CTM-RM method captures and refines impacts from sources that contribute precursors of PM$_{2.5}$.

Refined biomass burning and dust source impacts led to better agreement between simulated and observed concentrations of crustal (Al, Ca, Fe, Si) and biomass burning-derived elements (Cl, K). Original CMAQ-DDM simulations were biased very high for these species compared to observed concentrations. This is due to the apparently high bias in source profile estimates for biomass burning sources, which don’t take into account long-range transport and deposition of biomass burning-related PM. Results suggest that due to atmospheric transformation processes, the source profiles are in error for some species, similar to the findings in Balachandran et al. (2013). Ob-
servational data for some elemental species (Mg, P, Se, V) were highly influenced by measurement limitations (i.e., at or below MDL) and showed the poorest correlation with simulated observations. Additionally, conversion of carbon species between analytical methods, from total optical transmittance to total optical reflectance equivalents, introduced potential bias into concentration comparisons. However, other studies have shown that conversions may overcorrect observations of carbon species (Balachandran et al., 2013).

In summary, the SH method uses observations and modeled concentrations of species to adjust impacts on a source by source basis to provide spatially and temporally detailed source impact fields. Figure 5 shows spatial fields of source impacts for soil/crustal material, coal combustion, and other sources, which demonstrates the spatial and temporal completeness of the data that is provided by the SH method. The SH method also captures the impacts of secondary aerosol formation from precursor emission sources. Hybrid adjustment factors can be used to estimate the amount of change in emissions necessary for modeled results to better reflect observations, as emissions are roughly proportional to source impacts for primary sources. Kriging is an effective spatial interpolation method for spatially extending the CTM-RM model and generating spatial fields of adjustment factors. Adjusted CMAQ-DDM spatial fields of source impacts capture prior knowledge of emissions impacts, meteorology, and chemistry. Kriging does not introduce significant error, as the adjusted spatial fields maintain the spatial and temporal variability of the original source impact fields, and this application led to simulated PM$_{2.5}$ mass concentrations being closer to observations. Applying the hybrid model and spatial extension to original CMAQ-DDM source impact estimates also improves simulated estimates of crustal and trace metals.

The hybrid model is an effective approach for reducing the error in simulated source impacts through statistical optimization, instead of rerunning CMAQ-DDM which is more computationally expensive. Moreover, the methods presented generate daily, spatially complete fields that can be utilized by atmospheric scientists, air quality managers, and epidemiologists in health-related analyses. In future studies, the model will
be extended temporally to generate daily, adjusted spatial fields for the continental US for multiple years and to develop improved source profiles for emissions characterization.

The Supplement related to this article is available online at doi:10.5194/gmdd-8-645-2015-supplement.

Acknowledgements. This publication was made possible in part by USEPA STAR grants R833626, R833866, R834799 and RD83479901, STAR Fellowship FP-91761401-0, and by NASA under grant NNX11AI55G. Its contents are solely the responsibility of the grantee and do not necessarily represent the official views of the US government. Further, US government does not endorse the purchase of any commercial products or services mentioned in the publication. We also acknowledge the Southern Company and the Alfred P. Sloan Foundation for their support and thank Eric Edgerton of ARA, Inc. for access to the SEARCH data.

References


Development of PM$_{2.5}$ source impact spatial fields

C. E. Ivey et al.


Table 1. Source category abbreviations with average CMAQ-DDM, CTM-RM, and SH (spatial hybrid) source contributions to PM$_{2.5}$ concentrations for withheld CSN monitors ($N = 75$ monitors) for January 2004. Note: All averages and SDs are expressed in $\mu g \cdot m^{-3}$. Average total mass over withheld monitors for observations, CMAQ-DDM, CTM-RM, and SH was 11.7(±8.3), 16.3(±11), 8.59(±4.7), and 9.2(±5.7) $\mu g \cdot m^{-3}$. NR = Nonroad, CM = Combustion.

<table>
<thead>
<tr>
<th>Source Categories</th>
<th>Abbreviation</th>
<th>CMAQ-DDM</th>
<th>CTM-RM</th>
<th>SH Hybrid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural Burning</td>
<td>AGRIBURN</td>
<td>0.0040</td>
<td>0.003</td>
<td>25</td>
</tr>
<tr>
<td>Aircraft Emissions</td>
<td>AIRCRAFT</td>
<td>0.0038</td>
<td>0.013</td>
<td>26</td>
</tr>
<tr>
<td>Biogenic Emissions</td>
<td>BIOGENIC</td>
<td>0.074</td>
<td>0.22</td>
<td>14</td>
</tr>
<tr>
<td>Coal CM</td>
<td>COALCM</td>
<td>0.16</td>
<td>0.39</td>
<td>9</td>
</tr>
<tr>
<td>Diesel CM</td>
<td>DIESELCM</td>
<td>0.00060</td>
<td>0.0017</td>
<td>30</td>
</tr>
<tr>
<td>Dust</td>
<td>DUST</td>
<td>0.36</td>
<td>0.095</td>
<td>2</td>
</tr>
<tr>
<td>Fuel Oil CM</td>
<td>FUELOILC</td>
<td>0.14</td>
<td>0.54</td>
<td>12</td>
</tr>
<tr>
<td>Livestock Emissions</td>
<td>LIVEST2</td>
<td>0.31</td>
<td>0.89</td>
<td>3</td>
</tr>
<tr>
<td>Liquid Petroleum Gas CM</td>
<td>LPGCM</td>
<td>0.0043</td>
<td>0.013</td>
<td>24</td>
</tr>
<tr>
<td>Lawn Waste Burning</td>
<td>LWASTEBU</td>
<td>0.10</td>
<td>0.032</td>
<td>13</td>
</tr>
<tr>
<td>Metal Processing</td>
<td>METALPR</td>
<td>0.18</td>
<td>0.16</td>
<td>7</td>
</tr>
<tr>
<td>Meat Cooking</td>
<td>MEATCOOK</td>
<td>0.034</td>
<td>0.089</td>
<td>19</td>
</tr>
<tr>
<td>Mexican CM</td>
<td>MEXCM_B_M</td>
<td>0.00070</td>
<td>0.0028</td>
<td>29</td>
</tr>
<tr>
<td>Mineral Processing</td>
<td>MINERALP</td>
<td>0.030</td>
<td>0.062</td>
<td>21</td>
</tr>
<tr>
<td>Natural Gas CM</td>
<td>NAGASCMB</td>
<td>0.17</td>
<td>0.21</td>
<td>8</td>
</tr>
<tr>
<td>NR Diesel CM</td>
<td>NRDIESEL</td>
<td>0.14</td>
<td>0.48</td>
<td>11</td>
</tr>
<tr>
<td>NR Fuel Oil CM</td>
<td>NRFUELOI</td>
<td>0.010</td>
<td>0.036</td>
<td>23</td>
</tr>
<tr>
<td>NR Gasoline CM</td>
<td>NRGASOL</td>
<td>0.063</td>
<td>0.22</td>
<td>16</td>
</tr>
<tr>
<td>NR Liquid Petroleum Gas CM</td>
<td>NRLPG</td>
<td>0.0014</td>
<td>0.0056</td>
<td>28</td>
</tr>
<tr>
<td>NR Natural Gas CM</td>
<td>NRNAGAS</td>
<td>0.0005</td>
<td>0.0014</td>
<td>31</td>
</tr>
<tr>
<td>Other NR Sources</td>
<td>NROTHERS</td>
<td>0.0005</td>
<td>0.0012</td>
<td>32</td>
</tr>
<tr>
<td>Onroad Diesel CM</td>
<td>ORDIESEL</td>
<td>0.070</td>
<td>0.17</td>
<td>15</td>
</tr>
<tr>
<td>Onroad Gasoline CM</td>
<td>ORGASOL</td>
<td>0.27</td>
<td>0.60</td>
<td>4</td>
</tr>
<tr>
<td>Other CM Sources</td>
<td>OTHERCMB</td>
<td>0.040</td>
<td>0.072</td>
<td>18</td>
</tr>
<tr>
<td>Other PM Sources</td>
<td>OTHERS2</td>
<td>0.18</td>
<td>0.22</td>
<td>6</td>
</tr>
<tr>
<td>Prescribed Burning</td>
<td>PRESCRBU</td>
<td>0.032</td>
<td>0.054</td>
<td>20</td>
</tr>
<tr>
<td>Railroad Emissions</td>
<td>RAILROAD</td>
<td>0.013</td>
<td>0.046</td>
<td>22</td>
</tr>
<tr>
<td>Seashall</td>
<td>SEASALT</td>
<td>0.0001</td>
<td>0.0005</td>
<td>33</td>
</tr>
<tr>
<td>Solvent Emissions</td>
<td>SOLVENT</td>
<td>0.051</td>
<td>0.094</td>
<td>17</td>
</tr>
<tr>
<td>Wildfires</td>
<td>WILDFIRE</td>
<td>0.0018</td>
<td>0.0034</td>
<td>27</td>
</tr>
<tr>
<td>Woodfuel Burning</td>
<td>WOODFUEL</td>
<td>0.22</td>
<td>0.28</td>
<td>5</td>
</tr>
<tr>
<td>Woodstoves</td>
<td>WOODSTOV</td>
<td>0.62</td>
<td>0.44</td>
<td>1</td>
</tr>
</tbody>
</table>
Figure 1. Modeling domain (dotted, red line) and CSN, SEARCH, and IMPROVE monitors used for model development, application, and evaluation.
Figure 2. Spatial fields of kriged adjustment factors ($R_{j}^{SH}$) for dust, on-road diesel combustion, on-road gasoline combustion, and woodstove sources for 4 January 2004. Adjustment factors at CSN monitors (denoted by circles) were generated using hybrid (CTM-RM) source apportionment. Note that each figure has a different scale.
Figure 3. Hybrid-kriging adjustment of the dust impacts on PM$_{2.5}$ on 22 January 2004. (a) Original CMAQ-DDM simulation of dust source impacts. (b) Spatial field of hybrid adjustment factors for dust ($R_{ij}^{SH}$). (c) Adjusted spatial field of dust source impacts.
Figure 4. CTM-RM vs. Spatial Hybrid adjustment factors for withheld CSN monitors. Regression statistics: intercept, $\alpha = 0.14 \pm 0.02$; slope, $\beta = 0.84 \pm 0.02$; and correlation coefficient, $r = 0.89$. 
Figure 5.
Figure 5. Average CMAQ-DDM and spatial hybrid source impacts on PM$_{2.5}$ for observation days in January 2004 for seven source categories. Impact of (a, b) soil/crustal material, (c, d) traffic-related sources, (e, f) coal combustion, (g, h) sea salt aerosol, (i, j) metals-related sources, (k, l) fuel oil combustion, and (m, n) biomass burning.