Response to Referee #1

Interactive comment on "Development of $PM_{2.5}$ source impact spatial fields using a hybrid source apportionment air quality model" by C. E. Ivey et al.

Anonymous Referee #1

Received and published: 25 March 2015

General comments The paper introduces and discusses 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. The method is then applied to January 2004 over the continental United States. The paper addresses a relevant issue concerning the growing need to produce detailed and sound estimations of the contribution of the different emission sources to the PM concentration. To this aim the paper introduces a novel approach, partially based on a previous work, combining features of both source and receptor oriented modelling techniques. For this reason this work certainly fits the scope of GMD. The paper is well written, with concise and clear statements, however there a few general and specific questions that should be addressed before publications. General and specific questions are detailed in the following.

We would first like to thank the referee for the constructive review of our manuscript. As noted, the aim of this paper is to develop a modeling method that addresses the need for creating spatially and temporally complete source impact fields, particularly to serve as exposure surrogates for health studies. We believe this work will positively benefit those conducting health studies as well as policy assessments regarding ambient $PM_{2.5}$ concentrations.

First of all there are some inconsistencies in figures and table citations; particularly the order of figures and tables does not exactly reflect the corresponding order of citation in the text. Moreover there are some figures and tables included in the paper but never cited or commented.

We thank the reviewer for catching those and they have been fixed. We renumbered the figures and tables according to their appearance in the text. We also eliminated figures (previously Figs. S8-10) in the supplementary information that were not mentioned in the text.

The abstract is concise and complete, but I suggest to introduce also a few quantitative evaluation of the improvement in model performance (e.g some measure of error and its corresponding reduction).

We added the following performance metrics in the abstract: % error relative to the observed quantity for trace metals (e.g., V, Mg, P) and correlation coefficients for ions (e.g., NO_3^{-2}, SO_4^{-2}, NH_4^+).

"Correlations improved for concentrations of major ions, including nitrate (CMAQ-DDM: 0.404, SH: 0.449), ammonium (CMAQ-DDM: 0.454, SH: 0.492), and sulfate (CMAQ-DDM: 0.706, SH: 0.730). Errors in simulated concentrations of trace metals were reduced considerably: 295% (CMAQ-DDM) to 139% (SH) for vanadium; and 1340% (CMAQ-DDM) to 326% (SH) for manganese. Errors in simulated concentrations of very trace components are expected to remain given the uncertainties in source profiles."

I would also suggest to introduce one or two figures describing the average spatial field of R's (likewise figure 5 does for concentrations).

Average Rj values were presented for withheld CSN observation locations in the supplementary information (Table S1). We added an additional figure to display the 9-day averaged Rj spatial fields for four sources (dust, on-road diesel, on-road gasoline, woodstoves) similar to Figure 2 (Fig. S1).

There are also a few general questions that should be addressed in the discussion, namely: 1. The objective function described in (1) introduces a set of "adjustment" factors that allows filling the gap between observed and CMAQ concentrations modulating the influence of each source to the total concentration. But we know that the discrepancy between modelled and measured concentrations can rely not only on the emission strengths but also on other inputs (e.g. meteorology) as well as model formulation. Supposing for example that the overestimation of dust is related to an overestimation of wind, either that the overestimation of biomass burning is yielded by a too efficient SOA formation, how can the CTM-RM hybrid method improve the model performance for the right reason?

The referee brings up an important point that we clarify in the revised manuscript (conclusion). The purpose of the CTM-RM hybrid method is not to create a new chemical transport or atmospheric physics model, but to use statistical methods and measurements of PM species concentrations to improve source impact estimates. This improvement is two fold, 1) using the observed species concentrations to improve the predicted species concentrations and 2) incorporating emissions and atmospheric transformation processes from the model improves the secondary PM impact estimates beyond what is achievable using only a receptor model. We are not adjusting the source strengths, but adjusting the source impacts, and this distinction is important. We agree with the reviewer that other processes (e.g., wind, rain, or model parameter errors) can also impact the simulated impact of a source at a specific receptor, though errors in emissions (particularly highly variable ones such as dust) on a specific day can be large. Thus, the adjustment, as implemented here, is not on the emissions, but on the source-specific impacts. That is why we interpolate the receptor-specific adjustment spatially as opposed to using the adjustment to estimate how much the source strength should be adjusted and re-running the model.

The CTM-RM adjustment takes into account several sources of uncertainty: measurement uncertainty, modeled concentration uncertainty, as well as uncertainties in source impact estimates. These are represented as $\sigma_{i,obs}^2$, $\sigma_{i,SP}^2$ (now named $\sigma_{i,CTM}^2$, see question 3 below), and $\sigma_{\ln(R_j)^2}$, respectively. Additionally, using measured concentrations as the main point of reference for adjusting source impacts directly addresses the uncertainty of modeled processes (e.g.. wind-blown dust, SOA yield) by increasing or decreasing source impacts in light of an under- or overestimation in modeled concentrations. The 2nd term in the objective function prevents a non-real adjustment by further weighting the term with R_j uncertainty.

Since biomass burning and SOA formation have different compositions and source impact profiles, distinguishing between the two discrepancies can be determined by examining source impact adjustments. Further, collinearity is not an issue with this method because each source impact profile is different.

2. The paper introduces the concept of CTM-RM hybrid approach, also specifying that the RM model is CMB. But where does the RM actually contribute to the hybrid modelling approach? It seems that the objective function takes into account the source profiles, but not the results of a concentration apportionment. Probably more details are needed.

The RM portion of the hybrid approach is based on minimizing the weighted difference between observed concentrations and simulated concentrations, while also accounting for uncertainties in source strengths, source profiles and observations, in essence an extended chemical mass balance (CMB) approach. In section 2.2 (pg. 650, line 17), we note that the CTM-RM "uses an effective variance approach to balance model outputs." The effective variance approach is also utilized by versions of the CMB approach (Watson et al., 1984). As suggested by the referee, we extended this section to add additional details and make this clearer in the description of the CTM-RM hybrid method in section 2.2.

"The effective variance approach is also utilized by versions of CMB (Watson et al., 1984), and the optimization method used here is, in essence, an extended CMB approach. Uncertainties in the first term of the objective function serve as effective variances of the numerator and are specified for each species *i*."

3. Related to the previous point, there is also another issue concerning the proper definition of "source profiles". Do they represent an "emission speciation profile", thus describing the source fingerprint at the emission point? Either do they represent the source fingerprint at the receptor? In case they correspond to the first definition, have they been somehow compared to the emission speciation profiles adopted by SMOKE/CMAQ to define the 33 emission categories? In the second case (source profile at the receptor) I suppose they cannot be considered totally invariant, because the fingerprint of a source can change according to the travel time (e.g. different deposition rates for primary compounds; chemical transformation for secondary compounds). Please, briefly discuss this issue, if the authors consider that is relevant for the optimization process.

This is an important point covered in Hu et al. (2014) and is related to the prior response (point 1). In this manuscript, "source profiles" describe the source fingerprint at the receptor, i.e., that the source profile can be altered, e.g., by the formation of secondary species. However, for many of the species, there is no secondary formation. We assume that within the accumulation mode (which contains most of the fine PM mass in CMAQ) that the composition of the primary portion of the PM2.5 from any source is the same, but secondary species can be formed, and primary components lost to deposition, altering the source profile at the receptor. The specific steps taken in applying source profiles to CMAQ-generated data, and those steps are described as follows. In a publication by Adam Reff et al. (2009), source profiles for 84 source categories were presented, which were aggregated from roughly 300 $PM_{2.5}$ SPECIATE v4.0 profiles and contain estimates of trace metal contributions. The 84 PM2.5 profiles were further aggregated into 33 categories, consistent with the sources of interest in this study. Then the trace metals contributions in the 33 profiles were used to speciate the "unidentified" portion of PM2.5 (species name: A25) as output by CMAQ (v4.7.1). The contributions of 35 trace metal species were normalized to one and then used to split the unidentified PM2.5, and results for these species are used. At the receptor, both the primary and secondary PM2.5 contribution at the receptor is used to determine the new, receptor-oriented, source profile. This same approach was used to generate trace metal species concentrations in the preceding publication by Hu et al., and more details about this method may be found there (Hu et al., 2014, *ACP*).

We do agree that source fingerprints may change from source to receptor due to physical and chemical changes. But CMAQ, in principle, captures those changes for emissions from each source from their emitting point to the receptor point and represents those changes in the calculated source impact values. We included a more detailed discussion of the source profiles in section 2.2 in the revision.

"Source impact profiles are derived from the information provided by Reff et al (2009). In this manuscript, "source impact profiles" are different than "source profiles" in that they describe the source fingerprint at the receptor. In other words, the source profile can be altered, for example by the formation of secondary species. However, for many of the species, there is no secondary formation. It is assumed that within the accumulation mode, which contains most of the fine PM mass in CMAQ, the composition of the primary portion of the PM2.5 from any source is the same, but secondary species can be formed, altering the source profile at the receptor. The specific steps taken in applying source profiles to CMAQ-generated data, and those steps are described as follows. Source profiles for 84 source categories were presented in Reff et al. (2009), which were aggregated from roughly 300 PM2.5 SPECIATE v4.0 profiles and contain estimates of trace metal contributions. The 84 PM2.5 profiles were further aggregated into 33 categories, consistent with the sources of interest in this study. Then the contributions in the 33 profiles were used to speciate the "other" (sometimes called unidentified) portion of PM2.5 (species name: A25) as output by CMAQ. The contributions of the 35 trace species were then used to split the "other" PM2.5 in to individual species, and results for these species, along with the other primary and secondary species are used. At the receptor, both the primary and secondary PM2.5 contribution at the receptor are used to determine the new, receptor-oriented, source profiles. This same approach was used to generate receptor-oriented profiles in the preceding publication by Hu et al. (2014)."

4. The authors correctly point out that the "SAs" terms cannot strictly be considered as Source Contribution Estimates (SCEs), while they should be seen as sensitivity terms. This discrepancy can be relevant for sources like livestock that strongly contributes to ammonia emissions but not to NOX. As a consequence, in terms of SCEs they contribute just to ammonia, but in terms of sensitivity they influence both ammonia and nitrate. Therefore, in case of livestock, considering the sensitivity analysis as a source apportionment may imply an overestimation of the contribution of this source category. Do the authors consider that this aspect would help to explain the increased ranking of livestock after the adjustment? Do they consider this increase reliable?

We would first like to clarify the meaning of SAs in this manuscript. SAs denote the sensitivity of ambient particulate matter concentrations to emissions. In agreement with the referee, SAs are not total source contribution estimates (SCEs). However, we do not agree that our methods lead necessarily to an overestimation of the contribution. We do a mass balance on each individual speies as part of the method, and given that we have the inventoried sources (we use a comprehensive emissions inventory), so the result should not be biased high for any specific source. Our methods are novel in that, although some sources may not emit a certain pollutant, there still be some interactions with emissions from other sources, and they capture those interactions. For example, in the case of agricultural emissions, although little NOx is directly emitted, the influence on nitrate concentrations is precisely what we hope to estimate, as quantifying inter-source interactions (traditionally not quantified in source apportionment methods) are important in determining the primary and secondary impacts of sources on air quality. Our hybrid source- and receptor-oriented approach takes this into account, and can help elucidate impacts from source interactions. In this case, source impacts are the concentrations of pollutants that arise as a result of direct emissions and secondary interactions (both formation and destruction processes), and there is no truly unique impact of a source on secondary species when source interactions are involved. We added this explanation in the discussion section in the revised manuscript.

"The spatial hybrid method is also novel in that, although some sources may not emit a certain pollutant, there still may be some interactions with emissions from other sources leading to those species being part of the source impact. For example, in the case of agricultural fertilizer emissions, although NOx is not directly emitted, the influence on nitrate concentrations is calculated, taking account of inter-source interactions (traditionally not quantified in receptor-oriented source apportionment methods) that are important in determining the primary and secondary impacts of sources on air quality. This hybrid source- and receptor-oriented approach takes this into account and can determine impacts from complex source interactions. However, this also shows that the formation of secondary species is often dependent upon multiple sources, and the impact of one source is dependent upon other sources, leading to ambiguity in source attribution. The approach here uses the sensitivities at current conditions."

5. The key aspect of novelty of the paper concerns the development of gridded CTM-RM source apportionment results, based on findings at the receptor sites. Did the authors investigate the issues related to the spatial representativeness of the measurement sites?

First, we thank reviewer for noting spatial novel aspects of the manuscript. The spatial representativeness of our hybrid results were tested/evaluated by performing cross-validation through data withholding and by comparing to an independent dataset. Data withholding was used to evaluate the initial model development for spatial interpolation, where 10% of the monitors were randomly removed from the spatial dataset and the remaining 90% of points were kriged. The Rj value at the monitor and the corresponding kriged value from the spatial grid were compared. Evaluation metrics are presented in the supplementary information (Table S2 and Figure S3) and indicate that the observation data available for this study can adequately serve as inputs for spatial interpolation.

Following the model evaluation using data withholding an independent dataset was used to evaluate the model. Gridded hybrid concentrations were evaluated using observation data from the IMPROVE monitoring network. Data from IMPROVE was not used for model development. Results of this evaluation are found in the supplementary information (Tables S7 and S8). Correlation coefficients comparing observed and gridded concentrations are highest for more abundantly available species concentrations (higher frequency of measurement being above detection limit).

We addressed this comment in section 2.4 in the revised manuscript:

"This study uses available speciated CSN data over the entire U.S., thereby providing a very spatially heterogeneous dataset that is representative of key emissions and meteorology in each region. The lack of rural data available may present uncertainties in the spatial representativeness of R_j values outside of urban regions."

Specific comments P650 – (1). Though the objective function is properly referred (Hu et al., 2014), I suggest to add a few details to make it more readable. - Are all daily quantities ? - Add a few details about: Uncertainties in observation measurement, source profiles and source strength - Are Rj expressed as function of receptor and time?

In revision, we added more details about the objective function, its inputs, and the spatiotemporal characteristics of the R outputs (Section 2.2). R values are specific to one site and one day, as the method is applied at monitors when speciated $PM_{2.5}$ data is available on observation days.

"The initial R_j values are specific to one site and one day, as the method is applied at monitors when speciated PM2.5 data is available on observation days, and are then kriged and interpolated. The terms c_i^{obs} and c_i^{sim} represent the observed and CMAQ-simulated concentrations, respectively; Γ weights the amount of change in source impact. Uncertainties in observation measurement ($\sigma_{i,obs}$), modeled concentrations ($\sigma_{i,CTM}$), and source strength ($\sigma_{\ln(R_j)}$) are also included in the model. Specifically, $\sigma_{i,obs}$ is reported with measurements for each day from the CSN network; $\sigma_{i,CTM}$ is modeled error, which is proportional to observed concentrations and remains constant for all sites and days; and $\sigma_{\ln(R_j)}$ is uncertainty in source contribution expressed as the log of the factor of uncertainty, which also remains constant for each site and day."

P650 R19 – As already mentioned authors should briefly discuss the definition of "source profiles" (general question #3)

Please see our response to general question #3. We elaborated further on the source profiles in section 2.2 of the revised manuscript.

P652 R12 – if there are 189 CSN stations with 9 days, why N= 75 instead of about 170?

On page 652, line 12, N = 75 refers to the number of withheld CSN observations, which were used for model cross-validation. These 75 observations (space-time pairs) were randomly selected by removing 10% of the available observations with speciated $PM_{2.5}$ data on each observation day. Speciation is conducted every three or six days. We clarified this in Section 2.4 of the revision:

Response to Referee #1

"Performance of the spatial extension was evaluated using a data withholding approach. To evaluate the method, we removed 10% of the available observation (75 sets of observations at the monitors with speciated PM2.5 data) and re-ran the spatial hybrid model. This led to a total of 75 observation sets being used in the model evaluation. All references to "withheld CSN data" refer to these 75 sets of withheld data. "

P652 R24 – Authors say that 41 species were used for CTM-RM/SH optimization, though some of them were seldom above the detection limit. Can these species introduce too much uncertainty in the optimization phase?

The referee raises an important note regarding the introduction of added uncertainty due to some species concentrations measured at CSN sites being below detection limit. We also considered the possibility of added uncertainty due to detection limit issues. We tested the optimization with the absence of species with limited availability, and we found no significant differences in model performance. Further, knowing that the observations are below the detection limit is useful information. The uncertainty value used in weighting the observations reduces the impact of those observations, but the information is still captured. We also added a summary of this explanation in Section 2.4 of the revised manuscript:

"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. The possibility of added uncertainty in the optimization step due to detection limit issues was considered. The optimization was tested with the absence of species with limited availability, and no significant differences in model performance were found. The use of the measurement uncertainty in the objective function minimizes the role of those measurements on days when they are below the detection limit, but still accounts for the levels being low. Using all available measurements in the optimization model is the preferred approach."

P653 R5-8 – The maps show several "hot spots" with strong spatial gradients. Could this effect be related to the spatial representativeness of measurement sites? (see also General question #5). In some cases the observed value does not correspond to the surrounding gridded value. Are they withheld data?

The referee points out that in Figure 2 (cited as Figure 3 in text) the Rjs at the measurement sites do not correspond to the surrounding kriged results. We found that the spatial grid is shifted slightly to the east, hence this issue is not related to spatial representativeness. The observations align well with the kriged fields when re-plotted. Figure 2, now Figure 3 in the revision, was replotted.

P653 R9 – How many data are considered in table S2?

In Table S2, 75 data points, corresponding to 10% of monitors with available speciated PM2.5 data from each observation day. All references to metrics at "withheld CSN data" refer to these 75 monitoring locations throughout the manuscript. We made designation clearer in the method evaluation section (2.4), as well as in the table headers. Please see the response above.

P653 R10 – Figure S3 shows that almost all Rs are < 1.0 suggesting that CMAQ-DDM estimations are always overestimated at all sites, for all sources. Any comment?

We observed that most R values are less than 1.0, which indeed indicates that the hybrid-adjustment is reducing the initial CMAQ-DDM estimated source impact. However, for some sites and days, R values are greater than one (see Figure S2). We replotted Figure S2 with the y-axis on a log scale in order to better see the values greater than 1.0. Further, in Figure S3, the cumulative distribution plots exceed 1.0 (x-axis) for dust, lawn waste burning, prescribed burning, and woodstoves. These sources are highly variable day-to-day, and underestimations are possible in cases where the original emissions missed an actual burn or dust event. We addressed this comment in section 3.1 of the revision.

"The cumulative distribution plots exceed 1.0 (x-axis) for dust, lawn waste burning, prescribed burning, and woodstoves. These sources are highly variable day-to-day, and underestimations are possible in cases where the original emissions missed an actual burn or dust event."

P654 R9-13 – Are the overestimations concerning Fig S6 and S7 expressed in terms of "factors", likewise fig. S5? They seem too high.

The referee addresses that the factors presented in reference to the differences between the gridded spatial fields of CMAQ-DDM and spatial hybrid concentrations. The factors indeed do not appear to be reflected in the spatial field plots (Figs. S5-S7). However, the factors were calculated based on the average ratio of CMAQ-DDM to spatial hybrid grids over the entire domain. For instance, on Jan. 4, on average, the CMAQ-DDM grid values were a factor of 3 times higher than spatial hybrid grid values for biomass burning (Fig. S5). Over a large portion of the domain (boundaries, central US), impacts are near zero, and ratios may be influenced by numerical noise. We should note that there was a small error in the plots (Rjs spatial hybrid impacts fields were oriented in reverse), which produced plots that would lead the referee to believe that our estimates were too high. We have replotted Figures S5-7, and the new plots reflect the factors explained above.

P654 R22-25 – See General question #4

We addressed this comment in general question #4.

P655 (4) – What do "i" and "N" account for?

In Equation 4, *i* represents monitors and *N* represents the total number of monitors withheld for evaluation. We will clarify this in the next revision. Error of the modeled concentration of each species is calculated as the average of the errors over all withheld observations. The notation of Equation 4 will also be modified for clarity (see below).

$$Error = \frac{1}{N}\sum_{i=1}^{N}\frac{|obs_i - sim_i|}{obs_i} \quad (4)$$

P656 R5-10 – Authors discuss categories showing high absolute values of RMSE. Maybe some comments could be added also for categories showing a relevant RMSE with respect to the corresponding average and median (e.g. dust)

We included the discussion of sources with similar mean and median values and correspondingly low RMSEs, such as livestock, Mexican combustion, and nonroad natural gas combustion (Table S1 in the revision).

"Sources such as diesel, liquid petroleum gas, non-road natural gas, and Mexican combustion all had very low RMSEs, mean R values near 1, and median R values near 1. This indicates that there is little to no adjustment to these source impacts and that kriging captures the R values calculated by the CTM-RM application."

Response to Referee #1

P657 R1 and R4 – What does N represent?

In these references, N represents the number of monitors used for evaluation. We changed the notation to "N = ## monitors."

P657 R24 – Could it be useful adding also some information about emissions of the main precursors (NOX, NH3, SO2 and VOC)?

We added domain totals of the emissions of precursors and discussed their role in the formation of secondary $PM_{2.5}$ in Section 3.1.

"Coal combustion, which includes the secondary formation of sulfate, remains in the top three sources for average hybrid PM2.5 source contributions at withheld observation locations, as its emissions uncertainties are low due to the availability of continuous emission monitoring data. SO_2 emissions are large (Jan. 2004 domain totals: 72924.7 metric tons per day), as are NOx (74619.7 metric tons per day) (Table S9). During the study period, coal combustion had the highest contribution to SO2 emissions (35080.3 metric tons/day) and the second highest contribution to NOx emissions (14250.1 metric tons per day) behind mobile sources. The source impacts found here account for the transformation of these gaseous emissions from coal combustion."

P658 R25-P659 R14 – The concept of source profile and its role should be better clarified (see also general question #3)

Please see response to general question #3. We clarified the role of source profiles in the revision.

P659 R14-16 – Authors state that just through changes in emissions they can improve the model results and performance. But, how can they deal with discrepancies not directly related to emissions? See also general question #1

Please see the response to general question #1.

P670-671 figure 5 – Authors may also include a pair of total $PM_{2.5}$ gridded fields, also overlapping observed data. This would give an idea of the actual improvement before and after the implementation of the correction factors.

We added plots of CMAQ-DDM and spatial hybrid total $PM_{2.5}$ fields to Fig. 5, as well as a discussion of the performance in estimating $PM_{2.5}$. We plotted overlapped observations for one observation day in January, which gives a better idea of improvement after implementing the adjustment factors (Fig 5).

Tables S3-S5 – Some error metrics (e.g. RMSE) could be added for each species to quantify the changes in model performance between CMAQ-DDM and the hybrid approaches

We added columns for RMSE in Tables S3, S5, and S6.

Tables S6-S8 – They should be commented because some results are not very clear. For example Beta coefficient for PM2.5 in table S6 decreases from 0.43 to 0.27 and 0.24. I would expect an increase of beta coefficient toward 1.0, in case of improved model performance.

The referee raises an interesting point. Indeed some performance indicators for some species indicate poorer correlation, such as the beta values for calcium for CMAQ-DDM (beta = 1.22) and spatial hybrid (beta = 0.16) comparison (Table S6). However, all metrics presented must be taken into account and evaluated holistically. The alpha values for calcium indicate an improvement in performance, as the spatial hybrid value (alpha = 0.044) is closer to 0.0 than the CMAQ-DDM value (alpha = 0.13). Further, mean concentrations at withheld observation locations also indicate better performance of the spatial hybrid model, where mean calcium concentrations were 0.0407 (observed), 0.182 (CMAQ-DDM), and 0.0501 (spatial hybrid) (Table S3). According to the mean concentrations, the spatial hybrid method performs best. Throughout the analysis, CMAQ-DDM estimates of trace metal concentrations were orders of magnitude too high, while spatial hybrid results were closer to observations. While some individual metrics indicate better performance of the base CMAQ-DDM, overall performance of the spatial hybrid method is most favorable. An important point is that the species where performance is less good are typically those species that have a smaller role in determining source impacts, e.g., they are very trace species and/or have high uncertainties (relative to their observed concentrations) in the measurements or source profiles. We added this discussion to Section 3.1.

Technical corrections P654 R8 – Figure 2?

This part of the manuscript is in reference to Figs. S5-7.

P657 R24 – Why Tables S1 is placed before Table S2?

As addressed earlier, tables and figures will be renumbered according to their appearance in the text.

P668 Figure 3 – Is it cited in the text?

Tables and figures will be renumbered according to their appearance in the text.

Tables S4 and S5 – HYB should be SH (spatial Hybrid)?

Yes, HYB should be SH, as this was an oversight. Throughout the course of the production of this manuscript, we changed the notation of the results to spatial hybrid.

Tables S6-S8 – they are cited but not commented

We added relevant comments about these tables in the next revision.

Figures S4 and S8 – they seem not cited

These figures were removed in the revision.

Response to Referee #2

Interactive comment on "Development of $PM_{2.5}$ source impact spatial fields using a hybrid source apportionment air quality model" by C. E. Ivey et al.

Anonymous Referee #2
Received and published: 15 May 2015

This manuscript seems to be prepared with care. However, it is not clear the substantial contribution to modelling science compared with the previous paper of Hu et al. (2014). Hu et al. (2014) developed a hybrid CTM-RM method to calculate adjustment factors to refine the CTM-estimated source impacts at monitoring sites. This paper simply spatially interpolates the impact adjustment factors using the commercial MATLAB software. If this is the "new method" developed in this paper, it has no enough novelty for publication in GMD, a premier international journal. If the authors can revise the manuscript to highlight what new scientific findings they made by spatially interpolating the adjustment factors (rather than emphasizing on the method itself), it is still possible for this manuscript to be published.

We would first like to thank the referee for the constructive review of our manuscript and his/her note about the care with which the manuscript is prepared. The integration of the Hu et al. (2014), Chemical Transport Model-Chemical Mass Balance Method with kriging and interpolation in time allows both increasing our understanding of the spatiotemporal characteristics of source impacts in the United States, as well as developing accurate source impact fields for health studies. What we have done is linked together multiple modeling approaches, e.g., CMAQ-DDM3D responses are integrated into a chemical mass balance method, which is then linked with spatial kriging and temporal interpolation, thus developing a larger modeling system approach. We use commercially available packages to do the integration because they are very efficient and this also allows other modelers who are interested in using the modeling approach descried here to more readily develop variants (e.g., alternative chemical transport models, optimization techniques, etc.).

In regards to the new scientific findings, in the initially submitted version of the current paper we downplayed the scientific findings that the spatial extension brings, as we emphasized the method, which is the typical focus of GMD. In the revised paper we expand the discussion of its application and develop increased scientific findings (e.g., biomass burning impacts, dust impacts, mobile source impacts). We will also extend the discussion of the use of the method as an approach to provide spatiotemporal fields of source impacts on air quality for use in acute health studies, the main purpose for the development of this modeling approach. We are currently using this approach to develop 8 years of source impact fields for use in epidemiologic health studies. When using the initial CTM-RM method, source impact estimates are available for 1-2 locations in a metro region. After applying the current method to get spatially resolved source impact estimates, we are able to see the spatial variability in 33 source categories at more than 20 grid locations in a metropolitan statistical area (MSA). This allows the spatial variability in sources to be detected through the modeling domain. The spatial extension is necessary for this application, and it is very unique. In the revision, we emphasize the scientific findings that were elucidated in our results, along with the application in health studies.

In addition, there are also some other major comments that should be addressed before publication, as described below. Specific comments:

Response to Referee #2

(1) Section 1. Introduction: The authors only used 2-3 sentences to describe the previous studies and their shortcomings. I think this part should be significantly expanded.

We extended this section. In particular, we make additional references to previous receptor- and source-oriented apportionment studies and discuss their strengths, limitations and uncertainties. In the initial version, we attempted to make the manuscript as concise as we thought reasonable.

(2) Section 2. Data and method: The configurations of the CMAQDDM modeling system should be described, e.g., the modeling domain, geographical projection, physical and chemical mechanisms, initial and boundary conditions, and emission inventory, etc.

We now provide the additional modeling configuration in the revised version.

(3) Section 2.4 Model evaluation: The authors should explain the objectives of each evaluation method.

We included an explanation of the objectives of each evaluation method Section 2.4:

"The hybrid optimization is directly applied to withheld observations to assess the performance of the kriging model. Concentrations are reconstructed using Eq. 3 and the spatially interpolated adjustment factors. Then the original CMAQ-DDM, directly applied hybrid (CTM-RM), and spatial hybrid (SH) concentrations are compared to observed concentrations at withheld observation locations to evaluate the performance of each method in simulating concentrations. Linear regression was used to assess correlations between observations and modeled concentrations for each method.

In order to evaluate prediction performance in remote locations and in locations independent of CSN, CMAQ-DDM and hybrid concentrations were compared to observations at SEARCH and IMPROVE locations."

(4) Section 3. Results; Section. 4 Discussion: These two sections should be reorganized. Firstly, the spatial extension method should be evaluated before any discussion of the modeling results, so Section 3.5 and 3.6 should be moved ahead of Section 3.1-3.4. Secondly, the discussion section is long and complex, with model performance, modeling results, advantage/shortcomings mixed up; the majority of this section is actually "results" rather than "discussions". I would suggest the author to merge the "results" and "discussion" sections and move most of the "discussions" to the corresponding sub-sections of "results".

We reorganized the results and discussion section so that scientific findings information is located in the in a separate section of the paper. We moved the spatial evaluation before the results section.

(5) Section 3.1, P653 L10-14: The authors should explain why the adjustment factors for specific sources are less than one, near one, or more than one.

The value of the adjustment factor Rj is specified depending on the initial CMAQ-DDM simulation of the source impact. If, after optimization, the Rj value is less than one, then the initial CMAQ-DDM estimate will be reduced to be more consistent with observations. In turn, if the Rj value is greater than one, then the initial CMAQ-DDM estimate will be increased to be more consistent with observations. An Rj value of 1 indicates that no adjustment to the CMAQ-DDM is necessary to improve consistency with observations. We included this explanation in the revision:

"In general, for an R value less than one, the initial CMAQ-DDM estimate was reduced to be more consistent with observations. In turn, for an R value greater than one, the initial CMAQ-DDM estimate was increased to be more consistent with observations. An R value of 1 indicates that no adjustment to the CMAQ-DDM is necessary to improve consistency with observations."

(6) Section 3.1, P653 L14-17: How much is the difference between these two methods?

The difference in the two approaches is the log-transformation of the Rj values before kriging. In one approach, we log-transform the Rj values at the monitors before kriging, and then the kriged values are unlogged before use in reconstruction. In the second approach, we do not log-transform before kriging. In this manuscript we have chosen the second approach, as the additional log-transformation step produces similar results as the second approach.

(7) Section 3.2: This section has the same problem with Section 3.1. Lots of modeling results are shown, but their reasons/implications are seldom illustrated.

We initially placed the reasons/implications in the discussion section. We address the referee's concerns by rearranging the information presented in the results and discussion, as stated in our response to specific comment (4).

(8) Section 3.3: This manuscript focus on the development of source impact spatial fields. However, only the source impacts at the withheld CSN monitors are illustrated in this section.

Section 3.3 is meant to compare kriged results to direct application of the hybrid method. Other sections address performance of the spatial extension, and we feel that section 3.3 brings added perspective of the benefits of the hybrid method. We have reorganized the results section and have expanded the discussion of refined spatial fields in Section 3.2.

(9) A conclusion section should be added. The limitations of this method, except for that described in the last paragraph of the manuscript, should also be summarized. For example, the accuracy of this method still needs to be further improved by optimizing source profiles etc., and this method cannot capture the nonlinearity in the source-receptor relationships.

We included a discussion of CTM and RM limitations in the introduction:

"Several receptor-oriented SA models have been developed to quantify emission source impacts on pollutant concentrations. Each model has its own unique characteristics and associated uncertainties (Balachandran et al. 2012; Seigneur et al. 2000). Schauer and Cass (2000) used organic tracers for source apportionment using the Chemical Mass Balance (CMB) method at two urban sites and one background site in central California (Watson et al., 1984). Their implementation addressed the improper accounting of VOCs from motor vehicle exhaust and wood combustion. Watson, Chow, and Fujita (2001) reviewed several studies that used CMB for source apportionment, and reported that uncertainties in source contributions of VOCs led to uncertainties in impacts from important sources such as off-road vehicles, solvent use, diesel and gasoline exhaust, meat cooking, and biomass burning. The authors also describe several limitations of CMB, including reliance on existing observations and overlooking profiles that change between source and receptor due to factors such as dilution, aerosol aging, and deposition. Maykut et al. (2003) used Positive Matrix Factorization (PMF) for source apportionment at an urban

Seattle, Washington (USA) site with selected trace elements to distinguish combustion sources (Pattero and Tapper, 1994). Temperature-resolved organic and elemental carbon fractions were also used in Unmix to distinguish diesel and other mobile sources but did not lead to significantly different results (Henry 2005). There was also difficulty in distinguishing small sodium-rich industrial sources due to the similarity to the aged marine aerosol source profile.

In an effort to improve the spatial and temporal resolution of SA data and improve source distinction, chemical transport models (CTM) have been adapted to estimate emission impacts on pollutant concentrations. Marmur et al. (2006) conducted a comparison of source-oriented and receptor-oriented modeling results for a winter and summer month in the Southeastern U.S. The brute force method was used in the Community Multiscale Air Quality (CMAQ) model to calculate impacts from mobile sources, biomass burning, coal-fired power plants, and dust. The authors determined that meteorological effects had a strong impact on the temporal variation of CMAQ source impacts, where receptor model results exhibited more day-to-day variability. Koo et al. (2009) used the decoupled direct method (DDM) in the Comprehensive Air Quality Model with extensions (CAMx) to determine the sensitivity of particle sulfate concentration to changes in emissions of SO2 and NOX from point sources; NOX, VOC, and NH3, from area sources, and all emissions from on-road mobile sources (Byun and Schere, 2006; Dunker, 1981, 1984; Napelenok et al., 2006). DDM first order-sensitivities under-estimated the impacts on sulfate concentration when all emissions are removed due to nonlinearities, as compared to brute force method results. Zhang et al. (2012) addressed this issue by calculating second order sensitivities of inorganic aerosols using DDM, which better captured nonlinear responses to changes in emissions up to 50%."

We also included a summary of limitations of the spatial hybrid method in the discussion:

"Spatial hybrid inputs, methods, and results have inherent uncertainties and challenges that are associated with implementation. Input uncertainties include measurement error and challenges are posed with temporal availability and spatial representativeness of concentrations. Emissions inputs for each source are available at different temporal and spatial scales. For instance point source emissions are available at hourly intervals in some cases, while dust emissions are highly variable, both spatially and temporally. Area source emissions are estimated at weekly or monthly intervals and averaged source fingerprints for the primary components of the PM2.5 emissions are used, which removes the consideration of locally-varying source composition. Physical processes in CMAQ-DDM are uncertain as modeling atmospheric behavior is a complex undertaking. Also, first-order sensitivity approaches may not capture all nonlinearities in source-receptor relationships. SH results are also subject to potential systematic bias from the optimization and kriging steps, though our evaluation suggests those biases are minimal"

(10) The figures and tables are not consistent with the citations in the main text. For example, in P653 L8, "Fig. 3" should be "Fig. 2". In addition, some figures and tables included in this manuscript are never cited or described in the main text, and some are cited but not described.

Thanks for catching this. We corrected the table and figure labels and removed those that are not referenced in the manuscript. This comment was also addressed in the review from referee #1.

Technical corrections:

P649, L3-4: "Chemical Speciation Network (CSN)" should be "CSN"

Response to Referee #2

Yes, CSN was already defined in the introduction. We corrected this in the revision.

P655, L22: How do you determine outlying data pairs?

Outlying data pairs are determined by examining the distribution of the directly calculated R values (mean = 0.837, stdev = 0.478) and the kriged R values (mean = 0.828, stdev = 0.294) at the withheld observation locations. Data pairs were removed if either value was more than six standard deviations from the mean R value. The removed data points (5 points out of 2475) were well outside of this range; hence we concluding that removing them would not alter the correlation. We added this explanation in Section 3.1.

"Outlying data pairs are determined by examining the distribution of the directly calculated R_j values (mean = 0.84, stdev = 0.48) and the kriged R_j values (mean = 0.83, stdev = 0.30) at the withheld observation locations. Data pairs were removed if either value was more than six standard deviations from the mean R_j value. The removed data points (5 points out of 2475) were well outside of this range."

P657, L1 and L4: What is the meaning of N? I deduce that it represents the number of observation sites but it should be explained.

Yes, N represents the number of monitoring locations from each network that were used for evaluation. We changed this to (N = 8 monitoring sites) and (N = 38 monitoring sites) to enumerate the number of monitors from each network that were used for evaluation.

Table S4 and S5: I guess "HYB" should be "SH" because only the latter is described in the main text.

Yes, HYB should be SH. We changed this in the revision.

Development of $PM_{2.5}$ ~~Source Impact Spatial Fields Using~~source impact spatial fields using a ~~Hybrid Source Apportionment Air Quality Model~~hybrid source apportionment air quality model

Formatted: Subscript

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}

Correspondence to: C. E. Ivey (sunni.ivey@gmail.com)

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~~ motivating 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~~for January 2004 over the continental United States. The kriging step is evaluated using data withholding and by comparing results to data from alternative networks. Data withholding also provides an estimate of method uncertainty. Directly applied (HYB) and spatially interpolated (spatial hybrid, SH) hybrid adjustment factors at withheld ~~monitors~~observation sites had a correlation coefficient of 0.~~885~~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~~observation sites using ~~directly applied~~HYB and ~~spatially interpolated hybrid~~ SH adjustment factors. The mean concentrations of total $PM_{2.5}$ ~~for~~at withheld ~~monitors~~observation sites were 11.7 (± 8.3), 16.3 (± 11), 8.59 ~~±~~(± 4.7~~,~~), and 9.2 (± 5.7) μg m^{-3} for the observations, CTM, ~~directly applied hybrid~~HYB, and ~~spatially interpolated hybrid~~SH predictions, respectively. Correlations improved for concentrations of major ions, including nitrate (CMAQ-DDM: 0.404, SH: 0.449), ammonium (CMAQ-DDM: 0.454, SH: 0.492), and sulfate (CMAQ-DDM: 0.706, SH: 0.730). Errors in simulated concentrations of trace metals were reduced considerably: 295% (CMAQ-DDM) to 139% (SH) for vanadium; and 1340% (CMAQ-DDM) to 326% (SH) for manganese. Errors in simulated concentrations of very trace components are expected to remain given the uncertainties in source profiles. Species concentrations were reconstructed using spatial hybrid results, and the error relative to observed concentrations was greatly reduced as compared to CTM-simulated concentrations. 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,

Formatted: Subscript

Formatted: Subscript

Formatted: Indent: First line: 0.3"

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 μm 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~~; Samet et al. 2000; Dominici et al. 2010~~). 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 et al. 2005). 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 receptor-oriented SA models have been developed to quantify emission source impacts on pollutant concentrations. Each model has its own unique characteristics and associated uncertainties (Balachandran et al. 2012; Seigneur et al. 2000). Schauer and Cass (2000) used organic tracers for source apportionment using the Chemical Mass Balance (CMB) method at two urban sites and one background site in central California (Watson et al., 1984). Their implementation addressed the improper accounting of VOCs from motor vehicle exhaust and wood combustion. Watson, Chow, and Fujita (2001) reviewed several studies that used CMB for source apportionment, and reported that uncertainties in source contributions of VOCs led to uncertainties in impacts from important sources such as off-road vehicles, solvent use, diesel and gasoline exhaust, meat cooking, and biomass burning. The authors also describe several limitations of CMB, including reliance on existing observations and overlooking profiles that change between source and receptor due to factors such as dilution, aerosol aging, and deposition. Maykut et al. (2003) used Positive Matrix Factorization (PMF) for source apportionment at an urban Seattle, Washington (USA) site with selected trace elements to distinguish combustion sources (Pattero and Tapper, 1994). Temperature-resolved organic and elemental carbon fractions were also used in Unmix to

Formatted: Subscript

distinguish diesel and other mobile sources but did not lead to significantly different results (Henry 2005). There was also difficulty in distinguishing small sodium-rich industrial sources due to the similarity to the aged marine aerosol source profile.

In an effort to improve the spatial and temporal resolution of SA data and improve source distinction, chemical transport models (CTM) have been adapted to estimate emission impacts on pollutant concentrations. Marmur et al. (2006) conducted a comparison of source-oriented and receptor-oriented modeling results for a winter and summer month in the Southeastern U.S. The brute force method was used in the Community Multiscale Air Quality (CMAQ) model to calculate impacts from mobile sources, biomass burning, coal-fired power plants, and dust. The authors determined that meteorological effects had a strong impact on the temporal variation of CMAQ source impacts, where receptor model results exhibited more day-to-day variability. Koo et al. (2009) used the decoupled direct method (DDM) in the Comprehensive Air Quality Model with extensions (CAMx) to determine the sensitivity of particle sulfate concentration to changes in emissions of SO_2 and NO_X from point sources; NO_X, VOC, and NH_3, from area sources, and all emissions from on-road mobile sources (Byun and Schere, 2006; Dunker, 1981, 1984; Napelenok et al., 2006). DDM first order-sensitivities under-estimated the impacts on sulfate concentration when all emissions are removed due to nonlinearities, as compared to brute force method results. Zhang et al. (2012) addressed this issue by calculating second order sensitivities of inorganic aerosols using DDM, which better captured nonlinear responses to changes in emissions up to 50%.

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).~~

2008). Spatially-resolved source impacts and concentrations are key inputs for residential or county level exposure studies that investigate the impact of air pollution on regional health outcomes (Bell, 2006). The CTM-RM method combines the strengths of both source apportionment techniques in an effort to reduce uncertainty in source impact estimates. The goal of this study is to create spatial fields of source impacts by spatially interpolating source impact adjustment factors (ratios, or ~~Rs~~*R's*) and then applying those adjustments to CTM source impact fields. ~~Rs~~*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 ~~Rs~~*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 U.S. 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 CSN 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., Jan. 4th, 7th, 10^{th} ~~…~~ … 28th), which led to varying sample sizes for each observation day. The number of available monitors with speciated $PM_{2.5}$ data on observation days ranged from approximately 40 to 150 and each site had 5 to 9 observations over the period examined. 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; Hansen et al~~.~~., 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 U.S., 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 U.S. were used for model evaluation. ~~Monitors~~IMPROVE monitors in the eastern U.S. 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 U.S. due to

Formatted: Superscript

Formatted: Subscript

Formatted: Subscript

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 (v. 4.5). 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. ~~CMAQ v. 4.7 or 5.0.2) (Dunker, 1981, 1984;Byun and Schere, 2006;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~~ CMAQ-DDM was run 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 contains the continental U.S., southern Canada, and northern Mexico, with 36-km grid resolution, Lambert Conformal Conic geographic projection, and 13 vertical layers of variable thickness extending from the surface to 70 hPa. Meteorological inputs were generated using the Fifth-Generation PSU/NCAR Mesoscale Model (MM5) with 35 vertical layers, implemented with the Pleim-Xiu land surface model ~~(Pleim and Xiu, 1995;Xiu and Pleim, 2001)~~(Grell et al., 1994, Pleim and Xiu, 1995; Xiu and Pleim, 2001)~~.~~. Emissions inputs were processed using the Sparse Matrix Operator Kernel Emissions (SMOKE) model. Emisisons data originated from a 2004 inventory that was projected from the 2002 National Emissions Inventory (NEI2002). Please refer to the preceding publication by Hu et. al (2014) for additional details about the emissions inventory ~~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~~.

Next, the original source impacts, receptor observations, and uncertainties are used as inputs to the objective function (Eq. ~~$X^2 = \sum_{i=1}^{N}\left[\frac{[(c_i^{obs}-c_i^{sim})-\sum_{j=1}^{J} SA_{i,j}^{base}(R_j-1))]^2}{\sigma_{i,obs}^2+\sigma_{i,SP}^2}\right] + \Gamma\sum_{j=1}^{J}\frac{ln(R_j)^2}{\sigma_{\ln(R_j)}^2}$~~ 1) of the hybrid SA model,

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

where the adjustment factors R_j are optimized by minimizing the objective function, ~~χ^2.~~ χ^2. The initial R_j values are specific to one site and one day, as the method is applied at monitors when speciated $PM_{2.5}$ data is available on observation days, and are then kriged and interpolated. The terms c_i^{obs} and c_i^{sim} represent the observed and CMAQ-simulated concentrations, respectively; ~~Γ~~Γ 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.~~modeled concentrations ($\sigma_{i,CTM}$), and source strength ($\sigma_{\ln(R_j)}$) are also included in the model. Specifically, $\sigma_{i,obs}$ is reported with measurements for each day from the CSN network; $\sigma_{i,CTM}$ is error in modeled concentrationserror, which is proportional to observed concentrations and remains constant for all sites and days; and $\sigma_{\ln(R_j)}$ is uncertainty in source contribution expressed as the log of the factor of uncertainty, which also remains constant for each site and day.

The uncertainties weight the adjustment of modeled source impacts, in that components with larger uncertainties are weighted less. The objective function is minimized by using a nonlinear optimization approach known as sequential quadratic programming (Fletcher, 1987). The function is modeled using a ridge regression structure, as demonstrated by the second term, and uses an effective variance approach ~~(Watson et al., 1984) to balance model outputs.~~

~~Source profiles are derived from the information provided by Reff et al (2009).~~to balance model outputs. The effective variance approach is also utilized by versions of CMB, and the optimization method used here is, in essence, an extended CMB approach (Watson et al., 1984). 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^{adj}) at receptors to more closely reflect observations (Eq. 3).

$$SA_{i,j}^{adj} = R_j SA_{i,j}^{base} \tag{2}$$

$$c_i^{adj} = c_i^{sim} + \sum_{j=1}^{J} SA_{i,j}^{base}(R_j - 1) \tag{3}$$

Given that many of the source impact profiles are similar between categories such that colinearities are present, the variation of the R_j~~'s~~ values are constrained to $0.1 \le R_j \le 10$.

Source impact profiles are derived from the information provided by Reff et al (2009). In this manuscript, "source impact profiles" are different than "source profiles" in that they describe the source fingerprint at the receptor. In other words, the source profile can be altered, for example by the formation of secondary species. However, for many of the species, there is no secondary formation. It is assumed that within the accumulation mode, which contains most of the fine PM mass in CMAQ, the composition of the primary portion of the $PM_{2.5}$ from any source is the same, but secondary species can be formed, altering the source profile at the receptor. The specific steps taken in applying source profiles to CMAQ-generated data, and those steps are described as follows. Source profiles for 84 source categories were presented in Reff et al. (2009), which were aggregated from roughly 300 $PM_{2.5}$ SPECIATE v4.0 profiles and contain estimates of trace metal contributions. The 84 $PM_{2.5}$ profiles were further aggregated into 33 categories, consistent with the sources of interest in this study. Then the contributions in the 33 profiles were used to speciate the "other" (sometimes called unidentified) portion of $PM_{2.5}$ (species name: A25) as output by CMAQ. The contributions of the 35 trace species were then used to split the "other" $PM_{2.5}$ in to individual species, and results for these species, along with the other primary and secondary species are used. At the receptor, both the primary and secondary $PM_{2.5}$ contribution at the receptor are used to determine the new, receptor-oriented, source impact profiles. This same approach was used to generate receptor-oriented profiles in the preceding publication by Hu et al. (2014).

The hybrid method produces results that more closely reflect observations than the original CTM results, which are often biased (Hu et al., ~~2014b~~2014). It accounts for more known source categories than traditional RM approaches (e.g~~.~~., 33 versus 6), and it links sources and observations both temporally and spatially. Additionally, the hybrid method generates estimates of the uncertainty in source impact predictions and identifies potential

Formatted: Indent: First line: 0.4"

errors in source strength and composition. One limitation of the hybrid method is that results are only available at receptor locations when observations are available, limiting its spatial and temporal scope. In this paper, ~~the~~a 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 supplementary information for a flow diagram of the methods).~~. Hybrid-generated R_j ~~–~~values were spatially interpolated for each observation day using kriging to generate spatial fields of source impact adjustment factors. Matlab © (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~~Source impact fields for intervening periods are developed by interpolation of the ~~R_j's.~~R_j spatial fields. Temporally interpolating R_j values 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.

Formatted: Indent: First line: 0.4"

2.4 Method Evaluation

Performance of the spatial extension was evaluated using a data withholding approach~~ for which~~, as well as by comparison with data from the SEARCH and IMPROVE networks. For data withholding, we removed 10% of ~~CSN~~ the available observations (75 sets of observations at the monitors with ~~associated~~ speciated $PM_{2.5}$ data) and re-ran the spatial hybrid ~~data were randomly removed from the data set for each~~model. This led to a total of 75 observation ~~day (N = 75 total removed points).~~sets being used in the model evaluation. All references to "withheld CSN data" refer to these 75 sets of withheld data. The remaining 90% of the available observations were used to fit the variogram models, which were used ~~for~~in kriging to produce spatial fields of R_j~~. R_j~~ values. Concentrations are reconstructed

Formatted: Font color: Auto, German (Germany)

Formatted: Font color: Auto, German (Germany)

Formatted: Font color: Auto, German (Germany)

Formatted: Font color: Auto, German (Germany)

Formatted: Font color: Auto, German (Germany)

Formatted: Font color: Red, German (Germany) using Eq. ~~were extracted for grids containing the withheld monitors. The~~3 with the spatially interpolated adjustment factors. Additionally, hybrid CTM-RM optimization is directly applied to withheld ~~receptors~~observation sites to assess the performance of the kriging model. Then~~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.~~measurements at withheld observation locations to evaluate the performance of each method in simulating concentrations. Linear regression was used to assess correlations between observations and modeled concentrations~~. Results were also evaluated at SEARCH and IMPROVE locations, where~~ for each method.

In order to evaluate prediction performance in remote locations and in locations independent of CSN, CMAQ-DDM and ~~hybrid~~SH concentrations were compared to observations~~.~~ at SEARCH and IMPROVE locations. Note that the application of the CTM-RM hybrid method, as conducted here, did not include SEARCH and IMPROVE data, and CTM-RM/SH results are independent of those observation data. The SEARCH and IMPROVE comparisons also address the issue of spatial representativeness of using only CSN data to produce spatial fields. This study uses available speciated CSN data over the entire U.S., thereby providing a very spatially heterogeneous dataset that is representative of key emissions and meteorology in each U.S. region. The lack of rural data available may present uncertainties in the spatial representativeness of R_j values outside of urban regions.

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. The possibility of added uncertainty in the optimization step due to detection limit issues was considered. Optimization was tested with the absence of species with limited availability, and no significant differences in model performance were found. The use of the measurement uncertainty in the objective function minimizes the role of those measurements on days when they are below the detection limit, but still accounts for the concentration levels being low. Using all available measurements in the optimization model is the preferred approach.

Formatted: Indent: First line: 0.3"

Formatted: Font color: Auto, German (Germany)

3 Results

Formatted: German (Germany)

~~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 (See supplemental information). 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 supplemental information (SI Table S2) as well as source-specific probability distribution functions (SI 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 U.S. 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~~1.1 ~~Refined Spatial Fields~~

~~Base CMAQ-DDM spatial fields were refined by applying kriged fields of hybrid adjustment factors (Fig. 2). Sources with high occurrences (~ >50%) of adjustment factors less than 1 include biomass burning, metals processing, and natural gas combustion; refined spatial fields are presented in the supplemental information (SI Figs. S5-7). 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 Jan. 4th and 22nd for the eastern U.S. and for portions of the west coast (See supplemental information), 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 Jan. 4th, and a factor of 5.2 ± 1.0 times~~

~~higher on Jan. 22nd. 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 Jan. 4th, and 25 ± 21% higher on Jan. 22nd for metal processing impacts. Natural gas combustion source impacts (area and point sources only) were reduced for the southeastern U.S., the Ohio River Valley Region, the Gulf States, and parts of California and Texas. On average, simulated levels were 35 ± 14% higher than SH values on Jan. 4th, and 72 ± 28% higher on Jan. 22nd for natural gas combustion impacts.~~

~~3.31.1 Refined Source Impacts~~

~~Average source contributions to PM2.5 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.~~

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

~~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 species concentrations found from the observed data (see Supplemental Information). 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, Si, Ca, and Fe), 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. 8:~~

$$\cancel{Error(\%) = \sum_{i=1}^{I} \frac{|obs_i - sim_i|}{obs_i}} \qquad \cancel{(8)}$$

~~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 supplemental information.~~

~~3.5~~3.1 Spatial Extension Evaluation

CTM-RM and SH ~~species concentrations and~~ adjustment factors at withheld ~~monitors~~observation locations were compared using regression to evaluate the spatial interpolation that was performed using kriging. For each observation day (9 days), 10% of available ~~monitors~~observations were randomly withheld, resulting in a total of 2,475 R_j data ~~pairs~~points (75 ~~withheld sites~~observations locations x 33 source categories). Five outlying data pairs (< 0.5%) were removed from this regression ~~($R_{Hyb} > 2$).~~. Outlying data pairs are determined by examining the distribution of the directly calculated R_j values (mean = 0.84, stdev = 0.48) and the kriged R_j values (mean = 0.83, stdev = 0.30) at the withheld observation locations. Data pairs were removed if either value was more than six standard deviations from the mean R_j value. The removed data points (5 points out of 2475) were well outside of this range. The remaining CTM-RM and SH factors had a Pearson correlation coefficient of 0.~~885~~89, a linear regression slope of 0.83 ± 0.02, and an intercept of 0.14 ± 0.02 (Fig. ~~4).~~ 2).

Root mean square errors (RMSE) were calculated for the adjustment factors by source (Eq. ~~7~~5):

$$RMSE_j = \cancel{\sqrt{\frac{\sum_{i=1}^{N}\left(R_j^{CTM-RM} - R_j^{SH}\right)^2}{n}}} \sqrt{\frac{\sum_{i=1}^{N}\left(R_j^{CTM-RM} - R_j^{SH}\right)^2}{N}}, \quad j = 1..J \text{ sources}, \ N = 75\ sites$$

(~~7~~5)

RMSEs for all sources were less than 0.4, with the exception of RMSEs for lawn waste burning, prescribed burning, and woodstoves (Table ~~S2~~S1). This is expected given the

Formatted: Indent: First line: 0.4"

uncertainty in the burn emissions~~.~~ (Table S2). Sources such as diesel, liquid petroleum gas, non-road natural gas, and Mexican combustion all had very low RMSEs, mean R_j values near 1, and median R_j values near 1. This indicates that there is little to no adjustment to these source impacts and that kriging captures the R_j values calculated by the CTM-RM application. Mean and median ~~R_j's for each source are also calculated, and~~ R_j values are within 20% for most ~~source~~sources (Table ~~S2~~S1). The overall mean R_j value at withheld ~~monitors~~observation locations 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 ($PM_{2.5}$ was biased approximately 40% high).

~~Probability~~Cumulative distributions were examined for CTM-RM and SH adjustment factors for each source ~~(SI Fig. S3). Cumulative distributions of~~, and adjustment factors were highly correlated for each source~~. The spatial~~ (Fig. S1). 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 demostrated by nearly identical PDFs. The cumulative distribution plots exceed 1.0 (x-axis) for dust, lawn waste burning, prescribed burning, and woodstoves. These sources are highly variable day-to-day, and CMAQ-DDM underestimations are possible in cases where the original emissions missed an actual burn or dust event.

~~3.6 SEARCH and IMPROVE Comparison~~

~~The spatial extension of the hybrid~~Spatial fields of hybrid adjustment factors are presented for dust, on-road diesel and gasoline combustion, and woodstove sources (Fig. 3). Average R_j values over all observation days are also presented for reference (Figure S1). Typically, R_j values 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 are generally near one over most of the US, though R_j values for those sources tend be below one in the southeastern region of the U.S.

Formatted: English (U.K.)

Formatted: Subscript

Formatted: Indent: First line: 0.49"

Formatted: Font color: Auto

In general, for an R_j value less than one, the initial CMAQ-DDM estimate is reduced to be more consistent with observations. In turn, for an R_j value greater than one, the initial CMAQ-DDM estimate is increased to be more consistent with observations. An R_j value of one indicates that no adjustment to the CMAQ-DDM is necessary to improve consistency with observations. As such, after application of the SH method, it was found that while many of the source impacts were adjusted relatively little (i.e., $R_j \approx 1.0$), dust- and biomass burning-related impacts were often biased high in the original CMAQ-DDM simulation and therefore considerably reduced.

The distribution of all R_j values was approximately lognormal, and an analysis was performed to determine whether log-transformation of R_j values prior to the kriging step was necessary to reduce bias in source impact and concentration estimates (Fig. S2). In one approach, we log-transform the R_j values at the monitors before kriging, and then the kriged values are unlogged before use in reconstruction. In the second approach, we do not log-transform before kriging. From the analysis it was determined that lognormal transformation of R_j values was not necessary, as no significant difference was observed in reconstructed concentrations and source impact fields as a result of the transformation.

Additionally for method evaluation, withheld CSN observations were compared with SH concentrations, which were calculated using kriged R_j values and Eq. 3 (Table S3). The mean concentrations of total $PM_{2.5}$ for withheld observation locations 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, Si, Ca, and Fe), 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:

$$Error = \frac{1}{N}\sum_{i=1}^{N}\frac{|obs_i - sim_i|}{obs_i} \qquad (4)$$

In Eq. 4, *i* represents observations and *N* represents the total number of observations withheld for evaluation. 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. The large remaining errors stem from the source profiles leading some elements to being biased consistently high and others low. Further work to optimize source profiles can reduce residual errors.

Performance indicators for some species indicate poorer correlation, such as the beta values for calcium for CMAQ-DDM (beta = 1.22) and SH (beta = 0.16) comparison (Table S4). However, all metrics presented must be taken into account and evaluated holistically. The alpha values for calcium indicate an improvement in performance, as the spatial hybrid value (alpha = 0.044) is closer to 0.0 than the CMAQ-DDM value (alpha = 0.13). Further, mean concentrations at withheld observation locations also indicate better performance of the SH model, where mean calcium concentrations were 0.041 (observed), 0.18 (CMAQ-DDM), and 0.050 (SH) (Table S3). According to the mean concentrations, the SH method performs best. Throughout the analysis, CMAQ-DDM estimates of trace metal concentrations were orders of magnitude too high, while SH results were closer to observations. While some individual metrics indicate better performance of CMAQ-DDM, overall performance of the SH method is most favorable. An important point is that the species where performance is less good are typically those species that have a smaller role in determining source impacts. For example, those species are very trace and/or have high uncertainties in the measurements or source profiles relative to their observed concentrations.

The SH method was further evaluated by comparing simulated concentrations to independent data from the SEARCH and IMPROVE networks (~~See supplemental information).~~Tables S5-6). The mean concentrations over observation days were compared, as well as regression statistics for observations versus modeled results. For the SEARCH network (N~~=8),~~ = 8 monitors), average concentrations of 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 monitors), 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~~Linear regression metrics are also presented for SEARCH and IMPROVE monitors (Tables S7-8). Correlations for all SEARCH and IMPROVE species did not improve, however esimation performance for most trace metals and ions improved.

Formatted: Indent: First line: 0.4"

3.2 Refined Source Impacts

~~4~~1 Refined dust and biomass burning~~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 PM2.5 for January 2004, based on source emissions, were dust, woodstoves, and coal combustion, estimated at 1275, 5301, and 3407 metric tons per day, respectively (See supplemental information). However, uncertainties associated with dust and woodstove emissions are much higher than most of the other sources, a factor of 10 and 5 respectively (Hu et al. 2014; Hanna et al. 1998; Hanna et al. 2001). 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 PM2.5. 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 PM2.5 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 33rd and 31st, respectively. However, CMAQ-DDM, CTM-RM, and SH hybrid contributions ranked both sources significantly higher (biogenic rankings: 14th, 11th, and 9th, respectively; livestock rankings: 3rd, 1st, and 1st, respectively). Although primary~~ ~~$PM_{2.5}$~~ ~~emissions from these sources are not large, secondary processes lead to high source~~

Formatted: Subscript

~~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 PM2.5.~~

~~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~~estimates were biased very high for these species compared to ~~observed concentrations.~~observations. This is due to the apparently high bias in source impact 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 impact profiles are in error for some species, similar to the findings in Balachandran et al. (2013). ~~Observational data~~Observations for some elemental species (Mg, P, V, Se) were highly influenced by measurement limitations (i.e~~.~~., at or below MDL) and showed the poorest correlation with ~~simulated observations.~~modeled concentrations. Additionally, conversion of observed carbon species between analytical methods, from total optical transmittance to total optical reflectance equivalents, introduced potential bias into concentration comparisons. ~~However, other~~Other studies have shown that conversions may overcorrect observations of carbon species (Balachandran et al. 2013).

~~In summary, the~~Average source contributions to $PM_{2.5}$ at withheld CSN observation locations were ranked from largest to smallest for base CMAQ-DDM, CTM-RM, and SH (Table 1). The top three sources were woodstoves, dust, and livestock emissions for base CMAQ-DDM simulations, the latter source capturing the influence of ammonia emissions on the formation of nitrate. The livestock category includes impacts from agricultural and 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 than for CMAQ-DDM. 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. The order of source contributions at withheld observation locations for the CTM-RM and SH applications

Formatted: Indent: First line: 0.4"

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

Formatted: Superscript

compared well, though often differed greatly from the base CMAQ-DDM rankings. 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 S2). 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; Hanna et al., 2001; Hu et al., 2014). This uncertainty is driven in part by source variability. The 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 includes the secondary formation of sulfate and remains in the top three sources for average SH $PM_{2.5}$ contributions, as its emissions uncertainties are low due to the availability of continuous emission monitoring data. SO_2 emissions are large (Jan. 2004 domain totals: 72924.7 metric tons per day), as are NO_X emissions (74619.7 metric tons per day) (Table S9). During the study period, coal combustion had the highest contribution to SO_2 emissions (35080.3 metric tons/day) and the second highest contribution to NO_X emissions (14250.1 metric tons per day) behind mobile sources. The source impacts found here account for the transformation of these gaseous emissions from coal combustion.

Secondary formation processes increase the impact of coal combustion, biogenic and livestock emissions relative to their initial primary PM contribution. January 2004 primary PM emissions estimates for biogenic and livestock were ranked 33rd and 31st, respectively. However, CMAQ-DDM, CTM-RM, and SH hybrid contributions ranked both sources significantly higher (biogenic rankings: 14th, 11th, and 9th, respectively; livestock rankings: 3rd, 1st, and 1st, respectively). Although primary $PM_{2.5}$ emissions from these sources are not large, secondary processes and emissions from gaseous precursors led to high source contributions (Table S9). Biogenic sources emit large quantities of volatile organic compounds which go on to form secondary organic aerosols. Livestock emissions of gaseous ammonia react with sulfate, nitrate, and other acids to form ammonium salts. Therefore, the SH method captures and refines impacts from sources that contribute precursors of $PM_{2.5}$.

Formatted: Subscript

3.3 Refined Spatial Fields

Base CMAQ-DDM spatial fields were refined by applying R_j fields for each source and on each observation day. An example of the adjustment can be found in Figure 4, where the CMAQ-DDM spatial field of dust impacts is adjusted on January 4, 2004. Sources with high occurrences (~ >50%) of adjustment factors less than 1 include biomass burning, metals processing, and natural gas combustion, and refined spatial fields for these sources are presented in the supplemental information (SI Figs. S5-7). 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 Jan. 4^{th} and 22^{nd} in the eastern U.S. 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 Jan. 4^{th}, and a factor of 5.2 (± 1.0) times higher on Jan. 22^{nd}. Metal processing impacts were reduced for areas highly impacted by smelting and metal works industries including the Ohio River Valley and Mid-Atlantic regions (Fig. S6). On average, the CMAQ-DDM values were 21 (± 21) % higher than SH values on Jan. 4^{th}, and 25 (± 21) % higher on Jan. 22^{nd} for metal processing impacts. Natural gas combustion impacts (area and point sources only) were reduced for the southeastern U.S., the Ohio River Valley Region, the Gulf States, and parts of California and Texas (Fig. S7). On average, CMAQ-DDM levels were 35 (± 14) % higher than SH values on Jan. 4^{th}, and 72 (± 28) % higher on Jan. 22^{nd} for natural gas combustion impacts.

Refined spatial fields of Jan. 2004 averaged source impacts are presented for eight sources: (c,d) dust, (e,f) on-road mobile sources, (g,h) coal combustion, (i,j) sea salt, (k,l) metals-related sources, (m,n) fuel oil combustion, (o,p) biomass burning, and (q,r) agricultural activities (Fig. 5). Total $PM_{2.5}$ concentration fields are also included with overalapped observed concentrations from January 28^{th} (a,b). The CMAQ-DDM spatial field overestimates concentrations in the Eastern U.S., while overlapped concentrations agree more with spatial hybrid results. Modeled concentrations at monitors in mountainous areas, such as Salt Lake City, Utah, are underestimated due to local meterological conditions (Gillies et al., 2010, Kelly et al., 2013). Wintertime temperature inversions, which cause stagnation in air ciruclation and consequently high air pollution episodes in industrial vallies, are challenging to capture in models.

Improved spatial field correlation is reflected in monthly averaged spatial fields (Fig 5). SH dust impacts are greatly reduced domain-wide as compared to CMAQ-DDM. Monthly-averaged refinement of biomass burning, where impacts were also greatly reduced, and metals-related source impact fields are consistent with results previously mentioned for Jan. 4th and Jan. 22nd. Sea salt impacts are localized to coastal areas as expected, and agricultural activity most greatly impact the midwestern U.S., an area dominated by farm lands. Coal and fuel oil combustion impacts are highest in the eastern U.S. and western Mexico (fuel oil only) and were adjusted very little as compared to the original CMAQ-DDM field.

4 Discussion

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~~.~~ (Hu et al. 2015). Kriging is an effective spatial interpolation method for spatially extending the CTM-RM model and generating spatial fields of adjustment factors. ~~Adjusted CMAQ-DDM~~Kriging does not introduce significant error, as the adjusted fields maintain the spatial and temporal variability of the original fields, and this application led to simulated $PM_{2.5}$ mass concentrations being closer to observations. Adjusted 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 PM2.5 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 SH method also improves simulated estimates of crustal and trace metal concentrations.

~~The~~The SH method is being developed both to develop spatio-temporally accurate source impact fields that are consistent with observations, and also provides an approach to increase our understanding of the spatiotemporal characteristics of source impacts in the

Formatted: Indent: First line: 0.49"

United States. We find widespread adjustment to biomass burning and dust impacts (Rj less than one). These source impacts are consistent with observations, emissions estimates, and atmospheric transport and transformation. The SH method is also novel in that, although some sources may not emit a certain pollutant, there still may be some interactions with emissions from other sources leading to those species being part of the source impact. For example, in the case of agricultural fertilizer emissions, although NOx is not directly emitted, the influence on nitrate concentrations is calculated. Although traditionally not quantified in receptor-oriented source apportionment methods, taking into account inter-source interactions is important for determining the primary and secondary impacts of sources on air quality. This hybrid source- and receptor-oriented approach takes this into account and can determine impacts from elusive source interactions. However, this also shows that the formation of secondary species is often dependent upon multiple sources, and the impact of one source is dependent upon other sources, leading to ambiguity in source attribution. The approach here uses the sensitivities at current conditions, though also conducts a mass balance on a species-by-species basis minimizing any overall bias in the source impact attributions.

Spatial hybrid inputs, methods, and results have inherent uncertainties and challenges that are associated with implementation. Input uncertainties include measurement error and challenges are posed with temporal availability and spatial representativeness of concentrations. Emissions inputs for each source are available at different temporal and spatial scales. For instance point source emissions are available at hourly intervals in some cases, while dust emissions are highly variable, both spatially and temporally. Area source emissions are estimated at weekly or monthly intervals and averaged source fingerprints for the primary components of the PM2.5 emissions are used, which removes the consideration of locally-varying source composition. Physical processes in CMAQ-DDM are uncertain as modeling atmospheric behavior is a complex undertaking. Also, first-order sensitivity approaches may not capture all nonlinearities in source-receptor relationships. SH results are also subject to potential systematic bias from the optimization and kriging steps, though our evaluation suggests those biases are minimal.

5 Conclusion

The spatial hybrid model is an effective approach for reducing the error in simulated source impacts spatial fields through statistical optimization, instead of rerunning CMAQ-DDM which is more computationally expensive. ~~Moreover, the methods presented generate~~

Formatted: Indent: First line: 0.49"

Formatted: Font color: Auto

~~daily, spatially complete fields that can be utilized by atmospheric scientists, air quality managers, and epidemiologists in health-related analyses.~~Despite the several points of uncertainty, SH source apportionment can provide daily, spatially complete source impacts across a large domain over a long time period. The SH technique does not necessarily isolate specific atmospheric processes, as it is not a chemistry or physics model. It is a model based on statistics with the assumption that by incorporating observations (truth) and modeled atmospheric processes (prediction), two results can be statistically combined together to yield a better approximation of source impacts. Efforts are continual for reducing uncertainties, increasing the time span of available results, and evaluating estimations with other data sources, such as satellite imagery and independent field measurements. In future studies, the model will be extended temporally to generate daily, adjusted spatial fields for the continental U.S. for multiple years and to develop improved source profiles for emissions characterization. Results from SH implementation are beneficial to policy makers, public health analysts, and other air quality scientists that use spatially and temporally complete source impact data in studies where outcomes influence human welfare.

Formatted: English (U.K.)

Formatted: Indent: First line: 0.3"

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

Baker, K. R., Simon, H., and Kelly, J. T.: Challenges to Modeling "Cold Pool" Meteorology Associated with High Pollution Episodes, Environ Sci Technol, 45, 7118-7119, Doi 10.1021/Es202705v, 2011.

Balachandran, S., Pachon, J. E., Hu, Y. T., 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, DOI 10.1016/j.atmosenv.2012.07.031, 2012.

Balachandran, S., Chang, H. H., Pachon, J. E., Holmes, H. A., Mulholland, J. A., and Russell, A. G.: Bayesian-Based Ensemble Source Apportionment of $PM_{2.5}$, Environ Sci Technol, 47, 13511-13518, Doi 10.1021/Es4020647, 2013.

Bell, M. L. The use of ambient air quality modeling to estimate individual and population exposure for human health research: A case study of ozone in the Northern Georgia Region of the United States. Environ Int 32(5): 586-593, 2006.

Binkowski, F. S., and Roselle, S. J. Models-3 Community Multi-scale Air Quality (CMAQ) model aerosol component: 1. Model description, Journal of Geophysical Research, 108(D6), 4183, doi:4110.1029/2001JD001409, 2003.

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, Doi 10.1115/1.2128636, 2006.

Campen, M. J., Nolan, J. P., Schladweiler, M. C. J., Kodavanti, U. P., Evansky, P. A., Costa, D. L., and Watkinson, W. P.: Cardiovascular and thermoregulatory effects of inhaled PM-associated transition metals: A potential interaction between nickel and vanadium sulfate, Toxicol Sci, 64, 243-252, DOI 10.1093/toxsci/64.2.243, 2001.

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.

Chow, J. C., Watson, J. G., Pritchett, L. C., Pierson, W. R., Frazier, C. A., and Purcell, R. G.: The Dri Thermal Optical Reflectance Carbon Analysis System - Description, Evaluation and

Formatted: Subscript

Applications in United-States Air-Quality Studies, Atmos Environ a-Gen, 27, 1185-1201, Doi 10.1016/0960-1686(93)90245-T, 1993.

Darrow, L. A., Klein, M., Strickland, M. J., Mulholland, J. A., and Tolbert, P. E.: Ambient Air Pollution and Birth Weight in Full-Term Infants in Atlanta, 1994-2004, Environ Health Persp, 119, 731-737, Doi 10.1289/Ehp.1002785, 2011.

Dominici, F., Peng, R. D., Barr, C. D., and Bell, M. L.: Protecting Human Health From Air Pollution Shifting From a Single-pollutant to a Multipollutant Approach, Epidemiology, 21, 187-194, Doi 10.1097/Ede.0b013e3181cc86e8, 2010.

Dunker, A. M.: Efficient Calculation of Sensitivity Coefficients for Complex Atmospheric Models, Atmospheric Environment, 15, 1155-1161, Doi 10.1016/0004-6981(81)90305-X, 1981.

Dunker, A. M.: The Decoupled Direct Method for Calculating Sensitivity Coefficients in Chemical-Kinetics, J Chem Phys, 81, 2385-2393, Doi 10.1063/1.447938, 1984.

Fletcher, R.: Practical Methods of Optimization, John Wiley and Sons, 1987.

Gill, P. E., Murray, W., and Wright, M. H.: Practical Optimization, Academic Press, London, 1981.

Gillies R. R., Wang S. Y., Booth M. R. Atmospheric Scale Interaction on Wintertime Intermountain West Low-Level Inversions. Weather Forecast 25(4): 1196-1210, 2010.

Grell, G., Dudhia, J., and Stauffer, D.: A description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5), NCAR Technical Note: NCAR/TN-398+STRNCAR/TN 398+STR, 1994.

Hanna, S. R., Chang, J. C., and Fernau, M. E.: Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables, Atmospheric Environment, 32, 3619-3628, Doi 10.1016/S1352-2310(97)00419-6, 1998.

Hanna, S. R., Lu, Z. G., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., and Hansen, D. A.: Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain, Atmospheric Environment, 35, 891-903, Doi 10.1016/S1352-2310(00)00367-8, 2001.

Hansen, D. A., Edgerton, E. S., Hartsell, B. E., Jansen, J. J., Kandasamy, N., Hidy, G. M., and Blanchard, C. L.: The southeastern aerosol research and characterization study: Part 1-overview, Journal of the Air & Waste Management Association, 53, 1460-1471, 2003.

Hansen, D. A., Edgerton, E., Hartsell, B., Jansen, J., Burge, H., Koutrakis, P., Rogers, C., Suh, H., Chow, J., Zielinska, B., McMurry, P., Mulholland, J., Russell, A., and Rasmussen, R.: Air quality measurements for the aerosol research and inhalation epidemiology study, Journal of the Air & Waste Management Association, 56, 1445-1458, 2006.

Henry R. C. Duality in multivariate receptor models. Chemometr Intell Lab 77(1-2): 59-63, 2005.

Hu Y., Odman, M. T., and Russell, A. G. Mass conservation in the Community Multiscale Air Quality model, Atmospheric Environment, 40, 1199-1204, 2006.

Hu, Y., Balachandran, S., Pachon, J. E., Baek, J., Ivey, C., Holmes, H., Odman, M. T., Mulholland, J. A., and Russell, A. G.: Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach, Atmospheric Chemistry and Physics, 14, 5415-5431, DOI 10.5194/acp-14-5415-2014, 2014.

Hu, Y., Odman, M. T., Chang, M. E., and Russell, A. G. Operational forecasting of source impacts for dynamic air quality management, Atmospheric Environment, http://dx.doi.org/10.1016/j.atmosenv.2015.04.061, 2015.

Kelly, K. E, Kotchenruther, R., Kuprov R, Silcox, G. D. Receptor model source attributions for Utah's Salt Lake City airshed and the impacts of wintertime secondary ammonium nitrate and ammonium chloride aerosol. Journal of the Air & Waste Management Association 63(5): 575-590, 2013.

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(17): 6669-6675, 2009.

Laden, F., Neas, L. M., Dockery, D. W., and Schwartz, J.: Association of fine particulate matter from different sources with daily mortality in six US cities, Environ Health Persp, 108, 941-947, Doi 10.1289/Ehp.00108941, 2000.

Lim, S. S., Vos, T., Flaxman, A. D., Danaei, G., Shibuya, K., Adair-Rohani, H., Amann, M., Anderson, H. R., Andrews, K. G., Aryee, M., Atkinson, C., Bacchus, L. J., Bahalim, A. N., Balakrishnan, K., Balmes, J., Barker-Collo, S., Baxter, A., Bell, M. L., Blore, J. D., Blyth, F., Bonner, C., Borges, G., Bourne, R., Boussinesq, M., Brauer, M., Brooks, P., Bruce, N. G., Brunekreef, B., Bryan-Hancock, C., Bucello, C., Buchbinder, R., Bull, F., Burnett, R. T., Byers, T. E., Calabria, B., Carapetis, J., Carnahan, E., Chafe, Z., Charlson, F., Chen, H. L., Chen, J. S., Cheng, A. T. A., Child, J. C., Cohen, A., Colson, K. E., Cowie, B. C., Darby, S., Darling, S., Davis, A., Degenhardt, L., Dentener, F., Des Jarlais, D. C., Devries, K., Dherani, M., Ding, E. L., Dorsey, E. R., Driscoll, T., Edmond, K., Ali, S. E., Engell, R. E., Erwin, P. J., Fahimi, S., Falder, G., Farzadfar, F., Ferrari, A., Finucane, M. M., Flaxman, S., Fowkes, F. G R., Freedman, G., Freeman, M. K., Gakidou, E., Ghosh, S., Giovannucci, E., Gmel, G., Graham, K., Grainger, R., Grant, B., Gunnell, D., Gutierrez, H. R., Hall, W., Hoek, H. W., Hogan, A., Hosgood, H. D., Hoy, D., Hu, H., Hubbell, B. J., Hutchings, S. J., Ibeanusi, S. E., Jacklyn, G. L., Jasrasaria, R., Jonas, J. B., Kan, H. D., Kanis, J. A., Kassebaum, N., Kawakami, N., Khang, Y. H., Khatibzadeh, S., Khoo, J. P., Kok, C., Laden, F., Lalloo, R., Lan, Q., Lathlean, T., Leasher, J. L., Leigh, J., Li, Y., Lin, J. K., Lipshultz, S. E., London, S., Lozano, R., Lu, Y., Mak, J., Malekzadeh, R., Mallinger, L., Marcenes, W., March, L., Marks, R., Martin, R., McGale, P., McGrath, J., Mehta, S., Mensah, G. A., Merriman, T. R., Micha, R., Michaud, C., Mishra, V., Hanafiah, K. M., Mokdad, A. A., Morawska, L., Mozaffarian, D., Murphy, T., Naghavi, M., Neal, B., Nelson, P. K., Nolla, J. M., Norman, R., Olives, C., Omer, S. B., Orchard, J., Osborne, R., Ostro, B., Page, A., Pandey, K. D., Parry, C. D. H., Passmore, E., Patra, J., Pearce, N., Pelizzari, P. M., Petzold, M., Phillips, M. R., Pope, D., Pope, C. A., Powles, J., Rao, M., Razavi, H., Rehfuess, E. A., Rehm, J. T., Ritz, B., Rivara, F. P., Roberts, T., Robinson, C., Rodriguez-Portales, J. A., Romieu, I., Room, R., Rosenfeld, L. C., Roy, A., Rushton, L., Salomon, J. A., Sampson, U., Sanchez-Riera, L., Sanman, E., Sapkota, A., Seedat, S., Shi, P. L., Shield, K., Shivakoti, R., Singh, G. M., Sleet, D. A., Smith, E., Smith, K. R., Stapelberg, N. J. C., Steenland, K., Stockl, H., Stovner, L. J., Straif, K., Straney, L., Thurston, G. D., Tran, J. H., Van Dingenen, R., van Donkelaar, A., Veerman, J. L., Vijayakumar, L., Weintraub, R., Weissman, M. M., White, R. A., Whiteford, H., Wiersma, S. T., Wilkinson, J. D., Williams, H. C., Williams, W., Wilson, N., Woolf, A. D., Yip, P., Zielinski, J. M., Lopez, A. D., Murray, C. J. L., and Ezzati, M.: A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010, Lancet, 380, 2224-2260, 2012.

Marmur A., Park S. K., Mulholland J. A., Tolbert P. E., 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. Atmospheric Environment 40(14): 2533-2551, 2006.

Maykut N. N., Lewtas J., Kim E., Larson T. V. Source apportionment of PM2.5 at an urban IMPROVE site in Seattle, Washington. Environ Sci Technol 37(22): 5135-5142, 2003.

Napelenok, S. L., Cohan, D. S., Hu, Y. T., and Russell, A. G.: Decoupled direct 3D sensitivity analysis for particulate matter (DDM-3D/PM), Atmospheric Environment, 40, 6112-6121, DOI 10.1016/j.atmosenv.2006.05.039, 2006.

Oberdorster, G., Oberdorster, E., and Oberdorster, J.: Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles, Environ Health Persp, 113, 823-839, Doi 10.1289/Ehp.7339, 2005.

Pattero, P. and Tapper, U. Positive Matrix Factorization - a Nonnegative Factor Model with Optimal Utilization of Error-Estimates of Data Values, Environmetrics, 5(2), 111-126, 1994.

Peel, J. L., Klein, M., Flanders, W. D., Mulholland, J. A., Freed, G., and Tolbert, P. E.: Ambient Air Pollution and Apnea and Bradycardia in High-Risk Infants on Home Monitors, Environ Health Persp, 119, 1321-1327, Doi 10.1289/Ehp.1002739, 2011.

Peters, A., Wichmann, H. E., Tuch, T., Heinrich, J., and Heyder, J.: Respiratory effects are associated with the number of ultrafine particles, Am J Resp Crit Care, 155, 1376-1383, 1997.

Pleim, J. E., and Xiu, A.: Development and Testing of a Surface Flux and Planetary Boundary-Layer Model for Application in Mesoscale Models, J Appl Meteorol, 34, 16-32, Doi 10.1175/1520-0450-34.1.16, 1995.

Pope, C. A., Burnett, R. T., Thun, M. J., Calle, E. E., Krewski, D., Ito, K., and Thurston, G. D.: Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, Jama-J Am Med Assoc, 287, 1132-1141, DOI 10.1001/jama.287.9.1132, 2002.

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, Doi 10.1021/Es802930x, 2009.

Formatted: Subscript

Samet, J. M., Dominici, F., Curriero, F. C., Coursac, I., and Zeger, S. L.: Fine particulate air pollution and mortality in 20 US Cities, 1987-1994., New Engl J Med, 343, 1742-1749, Doi 10.1056/Nejm200012143432401, 2000.

Sarnat, J. A., Marmur, A., Klein, M., Kim, E., Russell, A. G., Sarnat, S. E., Mulholland, J. A., Hopke, P. K., and Tolbert, P. E.: Fine particle sources and cardiorespiratory morbidity: An application of chemical mass balance and factor analytical source-apportionment methods, Environ Health Persp, 116, 459-466, Doi 10.1289/Ehp.10873, 2008.

Schauer J. J., Cass G. R. Source apportionment of wintertime gas-phase and particle-phase air pollutants using organic compounds as tracers. Environ Sci Technol, 34(9): 1821-1832, 2000.

Seigneur, C., Pun, B., Pai, P., Louis, J. F., Solomon, P., Emery, C., Morris, R., Zahniser, M., Worsnop, D., Koutrakis, P., White, W., and Tombach, I.: Guidance for the performance evaluation of three-dimensional air quality modeling systems for particulate matter and visibility, Journal of the Air & Waste Management Association, 50, 588-599, 2000.

Thurston, G. D., Ito, K., Mar, T., Christensen, W. F., Eatough, D. J., Henry, R. C., Kim, E., Laden, F., Lall, R., Larson, T. V., Liu, H., Neas, L., Pinto, J., Stolzel, M., Suh, H., and Hopke, P. K.: Workgroup report: Workshop on source apportionment of particulate matter health effects - Intercomparison of results and implications, Environ Health Persp, 113, 1768-1774, Doi 10.1289/Ehp.7989, 2005.

Tolbert, P. E., Klein, M., Peel, J. L., Sarnat, S. E., and Sarnat, J. A.: Multipollutant modeling issues in a study of ambient air quality and emergency department visits in Atlanta, J Expo Sci Env Epid, 17, S29-S35, DOI 10.1038/sj.jes.7500625, 2007.

Wang, X. B., Ding, H., Ryan, L., and Xu, X. P.: Association between air pollution and low birth weight: A community-based study, Environ Health Persp, 105, 514-520, Doi 10.1289/Ehp.97105514, 1997.

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, Atmospheric Environment, 18, 1347-1355, Doi 10.1016/0004-6981(84)90043-X, 1984.

Watson J. G., Chow J. C., Fujita E. M. Review of volatile organic compound source apportionment by chemical mass balance. Atmospheric Environment 35(9): 1567-1584, 2001.

Xiu, A. J., and Pleim, J. E.: Development of a land surface model. Part I: Application in a mesoscale meteorological model, J Appl Meteorol, 40, 192-209, Doi 10.1175/1520-0450(2001)040<0192:Doalsm>2.0.Co;2, 2001.

~~(CEP, 2003)~~

Zhang W., Capps S. L., Hu Y., Nenes A., Napelenok S. L., and Russell A. G. Development of the high-order decoupled direct method in three dimensions for particulate matter: enabling advanced sensitivity analysis in air quality models. Geosci Model Dev 5(2): 355-368, 2012.

Tables

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~~observation locations (N= = 75 ~~monitors~~observations) for January 2004. Note: All averages and standard deviations are expressed in µg m^{-3}. Average total mass ~~over~~of withheld ~~monitors for~~ observations, and corresponsing CMAQ-DDM, CTM-RM, and SH ~~was~~estimates were 11.7 (± 8.3), 16.3 (± 11), 8.59 ± 4.7, and 9.2 (± 5.7) µg m^{-3}. NR = Nonroad, CM = Combustion.

Formatted: Subscript

Source Categories	Abbreviation	CMAQ-DDM			CTM-RM			SH Hybrid		
		Avg.	**St. Dev.**	**Rank**	**Avg.**	**St. Dev.**	**Rank**	**Avg.**	**St. Dev.**	**Rank**
Agricultural Burning	AGRIBURN	0.0040	0.003	25	0.0016	0.011	26	0.0012	0.0052	28
Aircraft Emissions	AIRCRAFT	0.0038	0.013	26	0.0037	0.013	25	0.0038	0.013	25
Biogenic Emissions	BIOGENIC	0.074	0.22	14	0.069	0.22	11	0.074	0.22	9
Coal CM	COALCMB	0.16	0.39	9	0.15	0.38	4	0.15	0.38	3
Diesel CM.	DIESELCM	0.00060	0.0017	30	0.0006	0.0017	30	0.0006	0.0017	30
Dust	DUST	0.36	0.095	2	0.061	0.22	13	0.048	0.12	14
Fuel Oil CM	FUELOILC	0.14	0.54	12	0.14	0.62	6	0.14	0.63	5
Livestock Emissions	LIVEST2	0.31	0.89	3	0.31	0.85	1	0.31	0.88	1
Liquid Petroleum Gas CM	LPGCMB	0.0043	0.013	24	0.0043	0.013	24	0.0043	0.013	24
Lawn Waste Burning	LWASTEBU	0.10	0.032	13	0.018	0.067	21	0.010	0.026	22

Metal Processing	MEATALPR	0.18	0.16	7	0.12	0.70	7	0.064	0.22	12
Meat Cooking	MEATCOOK	0.034	0.089	19	0.034	0.10	16	0.032	0.10	17
Mexican CM	MEXCMB_M	0.00070	0.0028	29	0.0007	0.0028	29	0.0007	0.0028	29
Mineral Processing	MINERALP	0.030	0.062	21	0.026	0.075	19	0.024	0.076	19
Natural Gas CM	NAGASCMB	0.17	0.21	8	0.11	0.36	8	0.078	0.20	8
NR Diesel CM	NRDIESEL	0.14	0.48	11	0.14	0.73	5	0.14	0.73	4
NR Fuel Oil CM	NRFUELOI	0.010	0.036	23	0.010	0.041	23	0.010	0.039	23
NR Gasoline CM	NRGASOL	0.063	0.22	16	0.061	0.23	14	0.064	0.23	13
NR Liquid Petroleum Gas CM	NRLPG	0.0014	0.0056	28	0.0014	0.0056	27	0.0014	0.0056	26
NR Natural Gas CM	NRNAGAS	0.0005	0.0014	31	0.0005	0.0014	31	0.0005	0.0014	31
Other NR Sources	NROTHERS	0.0005	0.0012	32	0.0005	0.0012	32	0.0005	0.0012	32
Open Fires	OPENFIRE	0.15	0.099	10	0.021	0.11	20	0.017	0.10	20
Onroad Diesel CM	ORDIESEL	0.070	0.17	15	0.066	0.19	12	0.068	0.19	11
Onroad Gasoline CM	ORGASOL	0.27	0.60	4	0.20	0.54	2	0.24	0.62	2
Other CM Sources	OTHERCMB	0.040	0.072	18	0.029	0.14	18	0.026	0.11	18
Other PM Sources	OTHERS2	0.18	0.22	6	0.10	0.28	9	0.10	0.28	7
Prescribed Burning	PRESCRBU	0.032	0.054	20	0.031	0.24	17	0.032	0.24	16

Railroad Emissions	RAILROAD	0.013	0.046	22	0.013	0.046	22	0.013	0.045	21
Seasalt	SEASALT	0.0001	0.0005	33	0.0001	0.0005	33	0.00	0.0	33
Solvent Emissions	SOLVENT	0.051	0.094	17	0.044	0.14	15	0.040	0.13	15
Wildfires	WILDFIRE	0.0018	0.0034	27	0.0012	0.0033	28	0.0013	0.00	27
Woodfuel Burning	WOODFUEL	0.22	0.28	5	0.20	1.3	3	0.12	0.90	6
Woodstoves	WOODSTOV	0.62	0.44	1	0.083	0.29	10	0.069	0.28	10

Figure Captions

Figure 1. Modeling domain (dotted, red line) and CSN, SEARCH, and IMPROVE monitors used for model development, application, and evaluation.

Formatted: Font color: Black, German (Germany)

Formatted: Don't suppress line numbers

Figure 2. CTM-RM vs. Spatial Hybrid adjustment factors for withheld CSN observations. Regression statistics: intercept, $\alpha = 0.14 \pm 0.02$; slope, $\beta = 0.84 \pm 0.02$; and correlation coefficient, $r = 0.89$.

Figure 3. Spatial fields of kriged adjustment factors (R_j^{SH}) for dust, on-road diesel combustion, on-road gasoline combustion, and woodstove sources for January 4, 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.

Formatted: German (Germany)

Formatted: Don't suppress line numbers

Figure ~~3~~4. Hybrid-kriging adjustment of the dust impacts on $PM_{2.5}$ on January 22, 2004. (a) Original CMAQ-DDM simulation of dust source impacts. (b) Spatial field of hybrid adjustment factors for dust (R_j^{SH}). (c) Adjusted spatial field of dust source impacts.

Formatted: German (Germany)

Formatted: Don't suppress line numbers

Figure ~~4.~~ 5. Average CMAQ-DDM and spatial hybrid source impacts on $PM_{2.5}$ for observation days in Jan. 2004 for eight source categories. Total $PM_{2.5}$ with overlapped $PM_{2.5}$ observations for Jan. 28th (a,b). Impact of (c,d) soil/crustal material, (e,f) traffic-related sources, (g,h) coal combustion, (i,j) sea salt aerosol, (k,l) metals-related sources, (m,n) fuel oil combustion, (o,p) biomass burning, and agricultural activities (q,r). ~~CTM-RM vs. Spatial Hybrid adjustment factors for withheld CSN monitors. Regression statistics: intercept, α = 0.14 ± 0.02; slope, β = 0.84 ± 0.02; and correlation coefficient, r = 0.885.~~

Formatted: Justified