Atmos. Chem. Phys. Discuss., 13, C11993–C12001, 2014
www.atmos-chem-phys-discuss.net/13/C11993/2014/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Fine particulate matter source apportionment using a hybrid chemical transport and receptor model approach" *by* Y. Hu et al.

Y. Hu et al.

yh29@mail.gatech.edu

Received and published: 6 February 2014

Reply to Referee #2

Comment: The manuscript by Y. Hu et al. describes a new approach to estimate source contributions to PM2.5 temporally and spatially based on CMB receptor model and CMAQ chemical transport model. It presents an interesting way to combine strengths of both receptor-based approach and emission-based approach and to give better results of source apportionment of PM2.5 than one model alone. The principles and methodology of the new hybrid approach are formulated clearly and the manuscript

has been relatively well organized for such a topic. Under constraint by observations of PM2.5 and its chemical compositions, the new hybrid approach gives much better model simulation results of PM2.5 and species than original CMAQ prediction, and has reasonable good estimations for contributions of 33 separate sources to ambient PM2.5 at the receptor sites, in which most of the sources are not resolved by receptor model without extra measurement information on unique tracers. The hybrid results from this study are generally consistent with traditional receptor model and also can be used to validate or refine relevant parameters in emission inventory. As a new method of PM2.5 source apportionment combining receptor-based approach and emission-based approach, it merits to be published in ACP. However, more detailed analyses are expected to make it complete and more convincing. In the following, I have a number of comments for the authors to address before publication.

Response: Thanks.

Comment: 1.The CMAQ performance statistics are well within the normal range of current state-of-the-art CTM's (page 26664, line 5). The simulated concentrations are found to be improved substantially compared to the initial simulation after refining source-impact estimates for major individual components and for most of the elements (page 26673, line 12). But it can be seen from Figure 2, both of the initial simulation and refinement are deviated from observations still quite large except for sulfate. Compared to observation, the original prediction is overestimated, but refinement is underestimated. This deviation might have significant influence on final results of the proposed new hybrid approach. The authors should quantitatively assess the impact of this deviation between model prediction and observation on source apportionment results.

Response: An indication of the overall accuracy of the adjustments, i.e. the prediction error can be found using the weighted least square error of all species (i.e.X2 as expressed in Eq. 13). On this regard, our scaled results showed that the refined X2(c,adj) (Eq. 13 with obtained Rj) are much smaller than the corresponding initial X2(c,base) (

Eq. 13 with Rj equal to 1) as seen as scatter dots in Figure 3 (each dot represents one pair of X2(c,base) and X2(c,adj)). The correlation between X2(c,adj) and X2(c,base) reveals a 98% reduction on average from X2(c,base) to X2(c,adj). However, this overall reduction doesn't mean every individual species is substantially better simulated at all times. Note that several elements with very low ambient concentrations (e.g. near the measurement uncertainty) were found to have slightly deteriorated agreement. On the other hand, since X2(c,adj) is the remaining error, and as the referee notes, "refinement(s) are (still) deviated from observations". Because the CTM uses the original source speciation, the error will not go to zero unless the source fingerprints were exactly correct. Further remaining error includes CTM model's other inputs error such as some meteorological bias and/or model limitations, e.g., the uncertainties involved in simulating nitrate or SOA formation. The size of the remaining error itself can be one indicator of the uncertainty of the hybrid results.

Comment: 2.Emission inventory and chemical speciation are essential for CMAQ to simulate PM2.5 and to estimate contributions from different sources. It is understood that emission inventory is not well established usually for some sectors, so scale factors are introduced to refine CMAQ estimations. In this manuscript, authors do not provide detailed description about the inventory used in the model. What is the accuracy of the used inventory in general? Is the same set of emission factors and source profiles applied to all cities nationwide, or different cities use different emission factors and source profiles? How to treat temporal and spatial variation for major emission sources? Those information are very important for explaining scale factors and their temporal and spatial variation.

Response: Lines 11-29 on page 26663 describe the "a prior" inventory used in our case study: "Emissions inputs used were developed from a 2004 inventory that was projected from the 2002 National Emissions Inventory (NEI2002, obtained from http://www.epa.gov/ttn/chief/emch/index.html#2002). Projection of the 2002 inventory to 2004 was conducted using growth factors obtained from the Economic Growth Anal-

ysis System (EGAS) Version 4.0, and control efficiency data obtained from EPA for the existing federal and local control strategies. In addition, the US emissions from large NOx and SO2 point sources for 2004 were obtained from the continuous emissions monitoring (CEM) database (http://ampd.epa.gov/ampd/). The inventory has emissions of seven criteria pollutants including PM2.5. PM2.5 emissions were split into major components (sulfate, nitrate, EC, OC and other) using source-specific speciation profiles from the SPECIATE program (Simon et al., 2010). The component historically called "unidentified" in the emissions modeling process, is called "other" here because this portion of PM2.5 is derived from measurements that provide the composition of the emissions, and includes metallic species which can be used to track source specific impacts on primary PM2.5. We group the emissions into 33 integrated source categories (a simple description of the source categories are in Table 1 and further detailed grouping information using source classification code (SCC) can be found in Table S2). The Sparse Matrix Operator Kernel for Emissions (SMOKE) model (CEP, 2003) is used to prepare gridded, CMAQ-ready emissions inputs. "

We also listed detailed sector-separate emissions information surrounding the six select CSN sites in the Table S8. Sufficiently detailed information about the preparation of the NEI2002 inventory is provided at the website cited above. Note that we used 2004 CEM data for emissions of large point sources in the US. These CEM emissions data are more accurate than other sources. Due to this, SO2 emissions are the most accurate. We also derived emissions uncertainties from literature (Hanna et al., 1998; Hanna et al., 2005; Hanna et al., 2001) for each sector by considering the daily emissions estimates uncertainties for each source and listed them in Table S2.

We used SMOKE to process the emissions inventory and allocate criteria pollutants' emissions, chemically to model species, spatially to grid cells and temporally to hours. The source-specific speciation profiles from the SPECIATE program were used in SMOKE. The same set of source profiles has been used nationwide with some exceptions. Spatial surrogates provided by the US EPA, derived from census and geog-

raphy information such as population, household, roads, railroad, landuse etc., were used in SMOKE for spatially distributing different emission subcategories according to their SCC codes. Monthly, weekly and diurnal temporal profiles were used to allocate emission to hours. Many temporal profiles were used nationwide, but dozens of state-specific temporal profiles were also applied. For example, different diurnal profiles have been developed for prescribed burning emissions from different states. While the emissions at each location are based on locally-specific information, the source specific uncertainties used in Eq. (14) are the same, since we have no knowledge of how they might be different in different locations. We will expand our description of the "a prior" emissions inventory.

Comment: 3.Particulate source apportionment technique is available in CMAQ (TSSA) and CAMx (PSAT), which have been applied in some research projects. Because TSSA-CMAQ for source apportionment has the same problem as concentration prediction by CMAQ due to the uncertainty of inventory, the new hybrid approach might provide more reliable results by using observation as constraint. TSSA or PSAT should be discussed in the "Introduction" section or the "Results" section. Of course, it will be nice if the authors can show comparison results between TSSA and hybrid model in some cities during same time periods either from literature or from author's work.

Response: We will add discussions on the PSAT application work of Burr and Zhang (2011) and conduct a general comparison with their January 2002 results at specific urban sites.

Comment: 4.In page 26666, line 18, an effective fij* is defined, which is more or less similar as source profile used in traditional receptor model CMB. The fij* directly accounts for secondary formation of PM2.5 and nonlinearities in pollutant transformations. It is true for secondary aerosol, but not for elements in PM2.5 because there are not much chemical transformation for elements and their mass should be conservative in air if dry and wet deposition processes are not significant. Thus, fij* could be calculated and then be compared with source profiles to validate emission inventory and its

chemical speciation. It can also be used to check the uncertainty of CMAQ modeling as well as the scale factors due to transport process.

Response: We formulated fij* so that it includes not only primary fraction and but also secondary fractions if applicable. They are similar to emissions source profiles but have secondary fractions if the emissions take part in chemical transformations. These suggestions will be incorporated into the revised manuscript. Comparing the calculated fij* back to original fij could reveal some useful information. However, not much can be said about the accuracy of the emissions inventory as the primary components are subject to similar atmospheric processing.

Comment: 5.Scale factor Rj is introduced to refine initial source apportionment results by CMAQ under constraint of observed PM2.5 and species. In principle, constant scale factors Rj for the same source should be found without temporal and spatial variation. However, it is partly true in this manuscript. Temporal and spatial diversities of Rj for the same source are still large, as seen in Fig S1 and Table S9. More explanation is needed for this diversity or variation. Is it caused by unified emission factor and source profile nationwide without area specific character, meteorological bias, or model bias? If Rj is mostly related to source uncertainty, I suggest that Rj value should be evaluated quantitatively by using the most recent emission factor and source profile.

Response: As noted above, emissions uncertainties and biases are not expected to be spatially or temporally uniform, especially on a daily basis. These are the major reasons for the various Rj being found. Meteorological bias and model formulation biases can be confounders, but their impacts are minimized by using a meteorological model, and corresponding fields, that is well evaluated, and the use of a well evaluated air quality model that represents the most important physical and chemical processes.

We did use the most recent source profiles provided from the US EPA's SPECIATE program (Simon et al. 2010) in our case study. We also used the most appropriate a priori emissions inventory for our case study period January 2004, which was

compiled based on the most appropriate emissions factors, activity database and projection factors for the year. First, we used 2004 Continuous Emissions Monitoring (CEM) data for emissions of large point sources in the US. The reason we didn't use NEI2005 (as base for projection to 2004) was because "EPA developed the 2005 NEI v2 based on a reduced level of effort. Part of this reduced effort involved using some NEI 2002 v3 data in the NEI 2005 v2 as surrogates for emissions data representing 2005" (http://www.epa.gov/ttnchie1/net/2005inventory.html). It would be inappropriate and difficult to do emissions projection based upon an inventory mixing emissions from different years. EPA has issued warning for the use of NEI2005 on their website: "Prior to using the 2005 NEI for analyses, users should consider whether the use of 2002 data as a surrogate for the selected sources documented above materially affects their analysis."

Comment: 6.The manuscript separated the primary and secondary contributions in the aggregated source impacts and merged the secondary portions correspondingly into ammonium sulfates, ammonium nitrate, and secondary organic carbon. It is worthwhile that authors provide some results for source contributions from different primary sectors to secondary species such as sulfate and SOA.

Response: Thanks for the suggestion. We will add a table to show this information.

Comment: 7. The comparisons with traditional receptor model should be during same time period at the same site. In Table 6, this is only true for the Atlanta site while other five sites use literature results in different time periods. For these five sites, it is suggested that authors reanalyze the dataset of same time periods using traditional receptor model.

Response: We agree that for the other five sites, comparing hybrid results with traditional receptor model results for the same period would be preferred, but they were not available. Our comparison at these five sites with longer time period literature results is still meaningful because we didn't compare the results in a sense of absolute-values.

Instead we compared the results in their major features such as what sources being resolved and the relative contributions between certain sources such as gasoline vs. diesel vehicles etc. Those features were captured by those selected RM method studies with longer simulation periods. Indeed, when we have discussed this method with others, they view that the more we can compare our results with other study results, the better. We agree. If we had found results that disagreed with prior studies, that would be of interest.

References: Burr, M. J., and Y. Zhang (2011b), Source apportionment of fine particulate matter over the Eastern U.S. Part I: source sensitivity simulations using CMAQ with the Brute Force method, Atmospheric Pollution Research, 2(3), 300-317, doi: 310.5094/APR.2011.5036. Hanna, S., J. Chang, and M. Fernau (1998), 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. Hanna, S. R., and R. Yang (2001), Evaluations of mesoscale models' simulations of near-surface winds, temperature gradients, and mixing depths, Journal of Applied Meteorology, 40, 1095-1104. Hanna, S. R., A. G. Russell, J. Wilkinson, J. Vukovich, and D. A. Hansen (2005), Monte Carlo estimation of uncertainties in BEIS3 emission outputs and their effects on uncertainties in Chemical Transport Model predictions, Journal of Geophysical Research, 110, D01302, doi: 01310.01029/02004JD004986. Simon, H., L. Beck, P. V. Bhave, F. Divita, Y. Hsu, D. Luecken, J. D. Mobley, G. A. Pouliot, A. Reff, G. Sarwar, and M. Strum (2010), The development and uses of EPA's SPECIATE database, Atmospheric Pollution Research, 1(4), 196-206.

Interactive comment on Atmos. Chem. Phys. Discuss., 13, 26657, 2013.

Equation 13:

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

Fig. 1.