Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-1198-RC2, 2019

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Impacts of Household Sources on Air Pollution at Village and Regional Scales in India" *by* Brigitte Rooney et al.

Anonymous Referee #2

Received and published: 24 January 2019

General Comments: Rooney and coauthors have undertaken a worthwhile modeling study on the impact of residential combustion sources on air quality in rural and urban India. The goal, scope, and methods of the paper are well-suited to ACP. They should also be commended for their efforts processing emissions, meteorology and land use data for this simulation – all challenging tasks. The conclusions drawn are a good start to a conversation about contributions from residential burning to PM2.5 and ozone. However, I find significant areas of improvement possible in the explanation of methods, consideration of underlying assumptions, and discussion of results. Please address the concerns below.

1. Quantitative and statistical analysis. The manuscript includes many statements like "whereas for the end of the month, predicted levels match closely those observed"

Printer-friendly version

Discussion paper

(page 15, line 23), “In general, the degree of agreement between predicted and observed O3 levels in New Delhi over these periods should be considered as reasonable” (page 15, line 27-28) and “the finer resolution computations generally predict somewhat higher PM2.5 concentrations than he coarser computations” (page 16, lines 9-11). These kinds of statements need to be supported with rigorous statistical analysis. I would expect some subset of bias, error, RMSE, fractional bias, fractional error, correlation coefficient, index of agreement, etc. to be provided in any contemporary air modeling study with access to observation data. Aggregate metrics can often fail us if we rely too much on them, but they do help to summarize the overall performance of the model against other studies.

The authors also rely primarily on timeseries plots comparing model predictions to observations for raw concentrations. It is usually useful to look at these comparisons with diurnal plots as well as raw timeseries to see if there are persistent issues at certain times of day; I recommend adding these for the data presented in figures 6-11. It’s also important to add error bars to illustrate variability on all diurnal timeseries plots, including figures 12 and 16.

I also would have expected to see maps of output data for metrics like fractional SOA contribution or fractional residential combustion contribution. Of course, it is just a model result, but I would think experts in air quality issues in India would find such maps interesting to ponder since this data cannot be obtained with existing measurements alone.

2. Species-level evaluation. Are there any observed speciation data that can help evaluate aspects of the model like the POA/SOA split? For example, OC/EC ratios might be instructive. It would also be useful to know if the model is predicting individual inorganic ion components well; do these data exist?

3. Meteorology evaluation. An evaluation of the meteorology fields has been completely omitted. This must be provided (e.g. supplemental information) or referenced

if published somewhere else. The met evaluation would also enhance the value of the timeseries plots of raw concentration data. If performance trends are seen across the month, it is important to find out if they are correlated with biases in any meteorological parameters.

4. Model spinup. The lack of model spinup for the individual scenarios is potentially troubling, considering that the authors are focusing on periods with meteorological conditions that favor accumulation (i.e. low wind speeds, low boundary layers, cold temperatures, etc.). Spin up is happening at the GEOS-Chem domain, so it is not as bad as starting from clean conditions, but there is a resolution change in going from the GEOS-CHem field to the CMAQ parent field and again for the nested field. It would be safer to at least consider omitting some of the timeseries for model spin-up. I recommend one of the following: a) rerun WRF and CMAQ simulations beginning about 10 days before each time period (i.e. August 20 for the Septemeber simulations). For emissions, if you don't have specific emissions for the spinup days, it should be safe to reuse appropriate days from the main simulation. For example, if August 20 is a weekday, make sure you pick a weekday from the main simulation. b) alternatively, cut off the first 7-10 days of each month form the existing results. Based on the timeseries data, it looks like making this revision will sacrifice minimal model-obs pairs for the PM dataset anyway.

5. Impact of Conclusions. Finally, the authors have summarized in the abstract and conclusions sections some important results of their work: a) residential combustion makes a significant contribution to air quality impacts in India, b) the model sometimes performs well, but sometimes does not, and this is a result of meteorological and emissions errors, and c) SOA makes up a significant but minor fraction of the total OA, with varying temporal profile between urban and rural areas. The authors have assembled a wealth of data and numerical predictive techniques to create this dataset; are there any more conclusions that can be drawn from this work if analyzed deeper?

For example, if one population-weights the emission rates of total PM or SOA precur-

sors across the parent domain, what sources stand out as the most impactful? The authors have shown that residential emissions do not matter as much in New Delhi as they do at the rural sites, but how does this translate to people impacted? And what does the emissions dataset say are the most important sources of primary PM2.5, secondary PM2.5 precursors, and Ozone precursors?

Another question to consider (aspects of this are repeated in some specific issues below) is how confident are the authors in their estimate of SOA contributions given the uncertainty in emissions speciation of SVOCs and IVOCs, SOA yields of aromatics, and potential degradation pathways like homogeneous fragmentation or particle-phase reaction/photolysis, among other uncertainties? I'm not calling for a suite of sensitivity studies here, but considering the author list contains unique experts in emissions and chemistry, it would be great to understand and document where this team thinks most of the uncertainty lies at this point for SOA predictions in India. Or, alternatively, what parameters need more investigation before such conclusions can be drawn confidently?

Specific Issues and Typos:

1. Page 1: consider mentioning residential heating as well as cooking since it does become a significant part of the discussion in the results section.

2. Page 2, line 4: consider replacing "pollution species" with "pollutants" since I wouldn't consider PM2.5 to be a species, technically.

3. Page 3, Introduction: There's little to no mention of ozone here, which seems strange given its importance in the discussion later. Consider highlighting it here in parallel to PM.

4. Page 4, line 16: What does this sentence mean? There is no "prediction" for SOA in India? Do you mean there is no inferred measurement of it? Or that there is a dearth of reported model predictions to compare to? Please be more precise.

5. Page 7, line 14: This statement asserts that there are no emissions from heating in the study, but Figure 2 clearly shows parameters used to inform the heating emissions. Please explain the heating emissions more clearly.

6. Page 7, line 15: This sentence makes it sound like there are no solvent emissions because the simulation is in August. I suggest rewriting and stating why there are not any solvent emissions. I assume it's from a lack of input data?

7. Page 8, line 5: Burning of biomass fuel can give compounds with quite high OM:OC. How robust are your results to this simplifying assumption?

8. Page 8, line 6: PNCOM, PMC and PH2O are ignored in the emissions. Do you have references for the validity of assuming they are negligible? I understand omitting coarse PM since you are studying PM2.5. However, studies from airsheds like LA have shown that how a model treats inorganic aerosol in the coarse mode can be important for fine mode predictions. What is your basis for assuming negligible effect?

9. Page 8, line 9-10: How much of the PM2.5 emissions is carbon? Shouldn't it be included in this equation? Is it 100

10. CMAQv5.2 assumes one volatility distribution for all POA, while this study considers POA from a variety of fuels and burning methods. What uncertainties are introduced here and what impact do they have on conclusions?

11. The resolutions of the coarse and fine grids should be mentioned in the emissions methods discussion as it would help the reader understand right away the challenge involved in specifying emissions for this simulation.

12. Page 11, line 9: Did you use 4-D data assimilation to 'nudge' the met model to observations within the domain? If not, how often did you reinitialize the met model with realistic fields? One of these approaches should probably have been used to keep the met model realistic over a month-long simulation. If neither were used, it makes a met evaluation all the more important to demonstrate or reference to better understand

Printer-friendly version

Discussion paper

any growing biases.

13. Page 13: Could the authors please discuss why they feel it is appropriate to omit the pcSOA species, which is based on the SIMPLE method of Hayes et al. (2015)? Is this because the speciation studies used for this residential emissions are believed to account for all of the particle and vapor phase mass that may be missed in inventories like the United States NEI? Page 13, line 11 says SVOCs could not be quantified. This mass is potentially important and represents part of what the pcSOA is designed to address. Can the authors be more explicit about what classes of organic compounds are accounted for by their emissions speciation and which might be underrepresented? Discussing this in the context of volatility would suffice. For example, are IVOC precursors for SOA missing as well?

14. Page 14: Much of the content in the first four paragraphs of section "ozone" is better-suited for the introduction section than the results.

15. Page 16, line 12: Recommend replacing "closeness" with "similarity". How did you verify this statement? Perhaps make a timeseries of the emission rate of residential primary PM (median and error bars) across the child domain for the 4 km and 1 km cases?

16. Page 16, lines 17-20: Is this statement true at all sites? Or is there variation in the importance of daily-varying emissions like agricultural burning?

17. Page 18, lines 19-21: This sentence confused me. It says the emission inventories are the same, but that the main differences between scenarios is explained by differences in POA? I'm probably not understanding what scenarios the authors are comparing (e.g. seasons or resolution or sensitivities). It would be helpful to be more precise here.

18. Page 19, lines 11-18: I'm confused again. Which conclusions are from Fleming et al., and which are new to this study? If they're all from Fleming et al., how do they

compare with the contribution CMAQ is predicting for this work? Again, if aromatics are currently the dominant source of SOA, how much could be coming from IVOCs and potentially being missed?

19. Figure 1: 4 km grid cell should read 4 km grid domain.

20. Figure 4: Recommend adding some visual cues for the flow of information from the parent domain all the way through to the nested domain. For example, add "4 km resolution" to the "Gridded Airborne Concentrations" box. Then duplicate that box, but with "1 km resolution" instead. You could then color the arrows based on whether they are relevant for the 1 km or 4 km simulation. For example, purple for 4 km and orange for 1 km. So the arrow from GEOS-Chem to IC-BCs would be purple. And the arrow labeled "Nested Domain" would be orange. The arrows connecting IC-BCs, Emissions and Met to CMAQ would be both orange and purple. Also, shouldn't there be an arrow from GEOS-Chem to the Met box? And from the Met box to Emissions? Maybe this is more detail than you want to include? You probably don't want to use the phrase "Parent Domain" in the GEOS-Chem box since I assume this was a global simulation? CMAQv5.2 Chemical Transport "Module" should read "Model".

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-1198, 2018.

Printer-friendly version

Discussion paper