Atmos. Chem. Phys. Discuss.,
https://doi.org/10.5194/acp-2018-833-RC2, 2018

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Ultrafine Particulate Matter Source Contributions across the Continental United States" *by* Melissa A. Venecek et al.

Anonymous Referee #1

Received and published: 6 October 2018

Review of Veneek et al., "Ultrafine Particulate Matter Source Contriutions across the Contentnal United States

In this paper, Veneek et al., have applied the UCD model to simulate PM2.5 in 39 different cities in the United States to identify source contributions. The paper addresses an interesting and potentially important area, i.e., what are the sources of ultrafine particles. A similar paper was recently published in ACP by the Pandis group. First, just to be more clear, the title of the article should include "simulated" or "modeled". What makes this article rather unique is the great lack of explanation of many important aspects. First, they choose 39 cities? What was the basis for the choice of cities? They

Printer-friendly version

Discussion paper

appear to be most of the large cities, but it would still be good to know why those were chosen and others were not (fundamentally, did they choose the largest 39 cities, and if not, why were others excluded? Having a research rational is important.). The next very, very important issue is that they chose very limited time chunks in each of the cities, some overlapping, but a very odd collection. What motivated such a choice? If it was model-performance driven, than any statement about good model performance is not so relevant, or at least should be taken very carefully and explained, as the model performance metrics have not been designed for allowing the ability to choose specific locations and times when the model is performing adequately. (This would be akin to letting epidemiologists to go back and choose periods that specifically do or do not find associations. That would be viewed as bad practice and not acceptable.) It is my understanding, and I have checked this with colleagues, that the typical modeling with a state-of-the-science model like CAMx (which is also a regional chemical transport model, similar to UCD), it is applied over the domain over a chosen period, and then you look at all of the results for a model evaluation. (When applied in SIPS, both the location and time period are predetermined.) Thus, what is appears is done here, but not stated, is that it was applied over the continental US (or some similar domain) for some period(s), again not explicitly stated. The manuscript should state what was the actual modeling period used, or did they choose the specific simulation dates a priori for each city and just simulate those? This must be stated if it is the case. If that is indeed, the case, was the model started on the beginning date chosen, or was there a period allowed for the initial conditions to be minimized? Also, might the authors better justify choosing episodes beginning in March or Octdober. The former seems a bit early, the latter a bit late. Were those the peak episodes that year? Also, the very limited time periods chosen limit the importance of the manuscript. Are we to take a 3-4 day period of the year to represent the source impacts for the whole year? I would expect a rather different set of sources in the winter than the summer. This whole area is not explored or discussed. I would have thought that that the choice might be driven by the availability of observational data. I realize ultrafine observations are rare, but

it would have been preferred to choose time periods when they were available. Why choose 2010 if UFP measurements are available other years? They might look at the recent Pandis paper (https://www.atmos-chem-phys.net/18/13639/2018/) doing a similar exercise over Europe. The choice of just peak ozone events may very well bias the assessment of source impacts on PM0.1 and PM2.5. Why not choose peak PM2.5 events? That would seem more natural. The model performance part is also rather opaque and requires references. First the model performance should be brought in to the main manuscript, e.g., as done in most modeling papers, showing overall performance across the entire domain and modeling period (not just selected locations and periods). See Simon et al., (2014) for the metrics typically provided and a more complete discussion of model performance evaluation. The working of the paragraph beginning on line 220 is also rather strange, it says “”. . .MFB values lower than 0.15 and MFE values lower than 0.35 are considered the goal or “excellent” in model performance.” Then they go on to say they do not meet them. First, I don’t think I have seen EPA have “excellent” as a performance description associated with those levels. I was looking for a citation here (a citation to the specific EPA evaluation metrics is required, as well as adherence to the terminologies used. My understanding is that the current EPA guidance is found at https://www3.epa.gov/ttn/scram/guidance/guide/Draft_O3-PM-RH_Modeling_Guidance-2014.pdf, and looking through that document, I don’t see them use the term excellent in terms of performance associated with any metric. (Indeed, a search of that document for the term “excellent” found only two occurrences, one in terms of protocol, another in terms of conceptual model.) That said, if the model does not meet the specific guidance levels, what does it meet? If the guidance was developed for regional scale modeling, without allowing selection by location and time, what does that imply here?). Second, if you do not meet them, what does that mean? Unlike the authors, I took the evaluation as not “building confidence in the accuracy of the model results. . .”, but left me questioning it. I would very much recommend that the authors follow the EPA guidance (or other recent articles, e.g., those by the AQMEII initiative or Ramboll: Emery et al.) in terms of how to conduct,

Printer-friendly version

Discussion paper

and report, model performance. Having now looked at the Ramboll study (JAWMA, doi/full/10.1080/10962247.2016.1265027), their ozone and PM performance do not meet the "goal", nor fully meet the recommended levels for "criteria". One could also follow the approaches recommended by the AQMEII initiative. The authors should definitely provide current citations for model performance for which they are using as their benchmarks. Thinking more holistically, the proper approach here would be to apply the model to a whole year, or, if that is computationally infeasible (which should be stated), be applied to one month periods in each season, and the results from each of those months be given. If they only want to consider the peak photochemically-active periods, they should choose a three month period (or more, preferably) that will capture events in all the cities for that year. I was a bit puzzled by the explanation given for Fig. 1. A 4 km grid is pretty fine, and the mobile sources in a typical urban area like LA are pretty ubiquitous. There are a number of monitors in LA: do any of them not go to zero at night? If not, that might suggest a different issue The mismatch in the evening needs a bit more discussion and justification. They commit to providing the outdoor exposure fields. They should also provide the model and its inputs. I assume this is journal policy, but the authors should likewise commit. A main conclusion of the paper is that natural gas is the main contributor to population-weighted exposure. This is a rather unique result and certainly requires more justification and discussion in light of the work that has shown via careful experiments that mobile sources and air craft are major sources of natural gas. Where is the empirical evidence of natural gas being a main contributor and can they show that they have captured the contribution of those other two sources? Their explanation of why Lane et al., or Posner and Pandis is not sufficient to argue that the current results are reliable. How well does the current study capture the spatial dynamics found by the groups from USC (Sioutas), Harvard (Spengler) and UW (e.g., Atmospheric Environment Volume 139, August 2016, Pages 20-29) which tend to point the finger at mobile sources and aircraft, so much so, the latter claim that ultrafine particle counts can be used as a tracer for aircraft turbine emissions. Given the lack of empirical evidence, compounded with their not having done any evaluation of the

ultrafine results against observations nationally, the speculative nature of this section suggests it should be removed, or couched in very different terms (i.e., noting the limitations, with a statement of the speculative nature). The question should be asked if there is sufficient evidence to support controls on a source based on the current analysis. The performance evaluation example (e.g., in reference to the EPA goal/excellence criteria) is not the only place where a citation is needed. They state that a number of other locations have PM0.1 levels above 2 ug/m3 (line 418). Then they go on to say there were sharper gradients in the observations. This would seem to contradict their findings/interpretation. More discussion needed. The entire last paragraph seems to contradict the findings of this paper. Again, they should compare their findings more directly to observations. In the summary, they state their analysis was for "peak photochemical periods during the year 2010." That is a rather strange way to characterize periods during which, in Atlanta, the 8-hr maximum ozone shown reached only about 50 ppb; I visually averaged, the actual value would be useful) in Cincinnati, about 60, in Los Angeles about 90, in New York about 85 ppb. I am not sure about the other locations, but I think the design value for LA in 2010 was about 120. In Cincinnati, it was about 0.079 (https://www3.epa.gov/airquality/greenbook/hbtcw.html), In Atlanta, about 80 ppb. The chosen periods would not appear to be "peak photochemical smog" periods. How different would Fig. 8 be if you simply used the local emissions? It is not apparent what this analysis adds and how it might be useful beyond simply using the inventoried sources. Line 425 uses the word "consistent" after saying that the model could not resolve the observed mobile source peaks. While one can see what they may be trying to say, it should be said differently, and more precisely. One could easily have said, experiments have found peak ultrafine particle levels were tied to on-road mobile sources and aircraft emissions, though those three sources combined account for only 22%, and are thus inconsistent with the results here that identify natural gas combustion. Minor: First line: should be "concentrations" L 156: Missing ")"

4 km stated in both of the first lines of the Abstract.

Fig. 2. Are these the individual, daily values for each site? If so, why are there not more circles? A bit more information in the figure caption would be useful.

Figure 3. What do they mean by “air pollution event”? Is the event simply when they compared their results to the observations? It would be good to know when, and where, the maximums occurred. In summary, at present the manuscript should not be accepted in to ACP. There are a number of critical shortcomings that may be addressed as discussed above. Those include, but are not limited to:

1. A much more straightforward and evenly presented model evaluation is needed. It should be conducted over the entire modeling domain for the entire modeling period, minus spin up days (which should be provided). Sub-regional evaluations can be presented, but they should be in addition to an overall evaluation. If subregional analysis are done for reduced periods, the specific criteria for the choice of period should be given. Saying they were “peak air pollution events” needs more definition (i.e., were they the highest ozone events in those cities that year?) It should not refer to excellent performance without appropriate citation. If it does not reach criteria or goal levels as laid out by either agency or peer-reviewed documents (preferably recent), they need to address this issue fully. 2. Precisely how the 39 cities were chosen should be specified. 3. In general, the text should be more precise over all. As noted above, “excellent” is one example. “peak photochemical period” is another. How the term ”confidence in the accuracy of the model results” is used is another. 4. The discussion of how the model results align with empirical studies should be extended, and the differences not dismissed. Greater justification of the finding of natural gas being the major contributor is needed in light of past studies. If the goal really is to provide an analysis of UFP over the US, using a set of one 4-day period for each city without a traceable reason for the specific choice of period, does not accomplish the objective.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2018-833, 2018.