Atmospheric
Chemistry
and Physics
Discussions

Interactive comment on "Retrospective analysis of 2015–2017 winter-time PM$_{2.5}$ in China: response to emission regulations and the role of meteorology" *by* Dan Chen et al.

Anonymous Referee #2

Received and published: 9 January 2019

The manuscript by Chen et al. evaluates PM2.5 in January in 2015 - 2017 in China. Data assimilation is used to construct a reanalysis that well matches the available observations of surface-level PM2.5 concentrations. The difference between this and a non-forced simulation is cleverly used to parse out the separate roles of emissions from meteorology in variability of PM2.5 concentrations, which are significant across these 3 years (although they authors refer to these as trends – which I think is better described as inter annual variability). This is important and valuable work, because it helps identify reasons why emissions control strategies may or may not have an immediately visible impact. As such the topic and scope are suitable for ACP. The manuscript includes some additional analysis of AOD, but it is somewhat secondhand, not as well

supported in terms of model accuracy, and doesn't particularly add to the focus of the paper, which I suggest remain on surface-level PM2.5 concentrations. This would provide space for the authors to provide more details on other aspects of their modeling and assimilation methods, which in several places are too abbreviated or presented without sufficient background or justification. The major scientific weakness in this work is likely the use of out-dated emissions (from 2010), given the rate at which species like SO2 and NOx are known to be changing since 2010 in China. Below I provide more detailed comments on these and other aspects to address prior to publication in ACP.

Comments: The grammar needs work throughout. This might mean adding another co-author or hiring an editing service. The number of corrections needed (nearly every sentence) are far too extensive for me to detail here.

The abstract would be improved by considering some of the following suggestions:

- lead with a broader, introductory statement

- avoid jargon when possible

- try to provide a mix of qualitative and quantitative results. Presently, only qualitative results are reported

- end with a statement about the bigger impacts of this work

- reduce the overall length – currently it goes too much into details of the methods, without quantitatively summarizing the most important conclusions.

4.22: I believe that this statement regarding the inorganic PM2.5 in this region has been known for some time.

7.1-3: There is considerable debate in the literature regarding the reactions that may be leading to high concentrations in haze events. The ones included here are based on assumptions of pH that may not be correct. Other recent works have e.g. suggested HCHO may play a role. In short, I recommend the authors review the relative literature

on this topic (which should be easy to find, several very high-profile papers). Even if they decide to stick with their current mechanism (which I think is acceptable, given their is yet to be scientific consensus on this issue), discussion is warranted, both in the introduction and consideration of sources of errors towards the end of the paper.

Figure 2: It's not clear what this shows. What are "PM2.5 emissions"? Since many of the species contributing to PM2.5 are gas-phase precursors, which don't necessarily completely transform into aerosol, I'm quite puzzled. Perhaps it is a plot of a subset of PM2.5 precursor emission species, but that isn't made clear.

8.1 and other locations (e.g., 14.11): The emissions used in this work are quite out-dated. It is documented in several studies of emissions in China that SO2 emissions have been decreasing since around 2009, and NOx emissions since around 2011 or 2012. These previous studies need to be cited, and considered in the present work. It is certainly a significant source of error worth considering.

9.2: Not sure why capital pi notation is used here, as that means product, where the definition is in terms of a sum.

9.10-12: I don't understand how this works – can it be explained further? The way in which measurements of total PM2.5 are used to adjust concentrations of specific species should be clarified, even though it comes from an earlier work, if only briefly.

9.20: What is the origin of this assumption regarding error? The relative error component of 0.75% seems very small. Other parameters such as gamma and L, the maximum concentration threshold (500) or analysis increment (120) are similarly intro-duced without much explanation. I recognize these values have been used before, but still more explanation here would be appreciated.

10.12: Omission of cross-correlation between species like ammonium and nitrate seems critical – how does this impact the results?

Section 2.5: The method used for estimating the impact of meteorology separately from

emissions seems sound; I'm not sure the extended explanation (by way of comparison to radiative forcing calculations, etc.) is needed and suggest simply removing the first paragraph of Section 2.5 and jumping directly into the statement of how this study was carried out.

Fig 4: The model performance seems good after the assimilation. One small question though – it seems like the residual bias in the CONC_DA case is most often negative. Is there a reason for this? I would have expected that, given the initial simulation is biased high, the analysis would not necessarily be unbiased, but would similarly have a small slight high bias (owing to the constraint term in the cost function).

Fig 5: Figs 5b and 5c could be omitted or moved to supporting information.

15.24: Why not use level 2 data?

16.7: I didn't follow nor agree with the logic behind this statement. Why not include statistics of the AOD comparison in a table, or directly on the figures themselves (there is room in the white space). Also, it seems that DA only serve to decrease the AOD, not increase it, even in areas.

16.14: Let's be honest – the DA didn't "didn't correct the bias" – it made the bias even worse, for 6 out of the 9 sites. This won't go un-noticed by the readers, so it might was well be presented fairly.

Overall, the AOD analysis was on the weaker side, the connections to the policy and haze questions not as clear, and the model performance not as good (especially for AERONET). I would suggest the authors consider dropping AOD related content en-tirely, unless some more satisfactory explanations can be included.

Section 4: Evaluation of concentrations in January alone of 3 consecutive years is not sufficient enough time range for a "trends" analysis. We do get some sense of interannual variability though, which is interesting. It is just mislabeled. For example, every place that says something like "decreasing trend from 2015 to 2016" should

say "decrease from 2015 to 2016", as the long-term trend hasn't been determined. If the authors really do wish to study trends, they should have considered years such as 2005, 2010 and 2015. Or if they can't go back that far, owing to data availability, 2012, 2014, and 2017. That would start to be close to enough years to make a trend analysis. My overall suggestion would actually be to remove section 4.1 entirely, at least the second paragraph on AOD. I also wonder if any of these years happened to be impacted by PM2.5 transport more than others, e.g. from fires?

18.13: I don't understand what is meant by the sentence beginning "Thus only. . ."

Fig 8: Suggest removing this figure; it is barely discussed, and doesn't add much. The analysis of PM2.5 surface concentrations is sufficient and also more convincing, since the model performance is better.

20.1: I think this is an important point – if the authors are using this approach to separate emissions impacts from meteorology in the observed dataset, then it is critical the relative changes in the total assimilation experiment (row b) match the observations.

21.14: Change "verified" to "evaluated", here and throughout, since in the strict sense of the word, the model has certainly not been verified.

22.2: I note the authors stop short of including anything regarding the AERONET evaluation in the conclusions – an indication that this part of the manuscript could be removed without impact.