Review of "Impact of Meteorology and Aerosol Sources on PM2.5 and Oxidative Potential Variability and Levels in China" by Liu et al.

The paper describes model calculations of PM2.5 and oxidative potential, which provides information on the impact of PM2.5 on human health. The paper examines the role of meteorology and anthropogenic emissions on PM2.5 concentrations and oxidative potential, finding that meteorological factors contribute more to the predicted surface PM2.5 concentrations and oxidative potential than anthropogenic emissions.

There have been many studies investigating PM2.5 distributions across China, but what is new here is taking those PM2.5 concentrations and estimating the oxidative potential. The results presented in this paper are not especially novel but give an incremental advancement in connecting air pollution dominated by particulate matter to human health. Like many papers, there is much to be clarified in the methodology and interpretation of results. However, a major need of the paper is a discussion of the uncertainties in the results and conclusions. In a sense, the authors address a couple of those uncertainties by performing simulations with different meteorology and different anthropogenic emissions. However, there is no discussion of the uncertainties associated with the assumptions made in their approach.

Major Comments

1. It would be good to see more explanation of what oxidative potential means. What does a value of 2 nmol/min/m3 imply? Oxidative potential is defined as "the ability of PM2.5 to produce reactive oxygen species (ROS) to *in-vitro* that consume intracellular antioxidants". The part in quotes is from line 41 of the paper and does not fully make sense to me. Is it meant that reactive oxygen species consume intracellular antioxidants as observed in lab studies? Does an oxidative potential mean that 2 nmol/m3 of ROS are produced per minute? Is that a lot? Are there thresholds for high oxidative potential versus low oxidative potential?

   I found a much clearer explanation of oxidative potential in Yang et al. (2021). I suggest revising the Introduction of this paper to present a clearer explanation, especially for readers who have not learned about oxidative potential yet.

2. The paper quantifies the annual average PM2.5 concentrations and oxidative potential for all of China and then discusses the spatial distribution of these parameters using maps. Distinct regions are noted in the spatial distribution discussion that I advocate should be quantified on a regional scale. What I mean by "quantify" is to provide average values of PM2.5 and OP for each region as was done in Figure 9a. The authors are the better experts to distinguish these regions, but what stands out in Figure 6 are these regions: northeast China, central and southeast China, southern portion of western China, and northwest China. As northwest China has the poorest agreement with observations, I suggest treating that region separately from the other regions in China.

3. The seasonality evaluation plots were interesting and point to good agreement during April – September and poorer agreement during the winter months. I suggest that the authors also separate their analysis between summer, when there is more confidence in the model results, and winter.
4. What role does PM2.5 from outside China have on PM2.5 concentrations and oxidation potential? For example, does biomass burning PM2.5 from Russia affect PM2.5 in China?
5. The study uses 50 km x 50 km horizontal grid spacing and 29 vertical levels to 100 hPa. What impact does this rather coarse resolution have on the results presented? For example, are the urban regions and their emissions well represented or do the emissions and concentrations get diluted by placing them in a grid box bigger than the urban region? Likewise, does the vertical resolution impact the prediction of the boundary layer height affecting the estimated surface concentrations of PM2.5? Please add a section discussing uncertainties in the study's results including the choice of the model configuration and other assumptions as noted below in the specific comments.

Specific Science Comments

1. Lines 28-29 of abstract: What is the meaning of OP values? Is 1 nmol/min/m3 low, and if so, does that mean there is small potential for health risk? Please explain why it is useful to report the numbers.
2. Section 2.2. It was not clear whether DEHM represents any feedbacks of aerosols on the meteorology. Could this be clearly explained in the model description.
3. Equation 1 (line 151) gives the calculation for oxidative potential as a function of PM2.5 concentrations. If I understand correctly, this equation comes from Liu et al. (2018) cited in the paper and is based on positive matrix factorization performed on samples from coastal cities. Could an explanation be added stating 1) units of the terms, and 2) the reliability of using this equation outside of an industrial coastal region. To me, this is another uncertainty embedded in this study.
4. Line 207. Why is the emissions reduction 30% and not another number? Please justify. Have any tests been done to learn about the response to different assumed emission reduction values? Perhaps this could be included in a section discussing uncertainties of the study's results.
5. Line 207. Could the authors clarify whether only PM2.5 emissions are reduced or if any precursors (e.g. SO2) also have emission reductions. If I understand the methods correctly, it appears that only primary PM2.5 emissions are perturbed for each sector. Since about half the aerosol is produced by chemistry (line 227) and emission controls can be placed on PM2.5 precursors (e.g., SO2 and NOx), it seems that additional calculations for reducing precursor species from different sectors is warranted. Could the authors please justify why only primary PM2.5 anthropogenic emission reductions are investigated, or add results presenting the impact of reducing PM2.5 precursors on the oxidative potential is examined.

6. Lines 218-225. Could information on the fraction of each sub-sector's contribution to E*re* and E*pp* be added?
7. Equation 9. Could the authors explain the 30% term in this equation? If it is the percentage reduction, then wouldn't this term change depending on which source sector is perturbed?
8. Figure 3. Are the results shown for simulation C1? Why not also show the evaluation for simulations C2 and C3? Throughout the manuscript, it would be informative to see the figures for all three scenarios. Please add these figures in either the main manuscript or the supplement.
9. Are the Figure 4 results for the annual mean? If so, please state this and explain if it is subtracting the annual mean of the simulation results from the annual mean of the observations. In contrast, it could be differencing individual time points for each location and then taking an annual mean of the difference.
10. Figure 4 shows differences of up to 18 ug/m3 in central China. How does this compare to the annual mean concentrations? It would be helpful to also see a percent difference map.
11. Line 297. I disagree with the statement, "We conclude that the model performs well ….", which is stated just after noting the poor agreement in northwest China (which is for good reasons). Please qualify this statement. I suggest limiting the remainder of the analysis to all of China except northwest China. See also my comments about performing calculations for specified regions.
12. Lines 309-320. Do the maps of PM2.5 concentration and OP reflect the anthropogenic emissions map of PM2.5? As written, this text does not tell me anything new, but I think the authors were hinting at some useful information in the last few lines. I suggest adding maps of anthropogenic emissions for each sector to support why we see the spatial distribution in Figure 6. Then rewrite these lines to focus on these connections.
13. Section 3.4 provides good conclusions and interesting points but does not complete the analysis of their model output to explain the results. Instead of going into detail on each panel shown in Figure 10, I suggest highlighting what is learned. What I learned is that secondary aerosol formation is the primary contributor to PM2.5, while biomass burning and industrial sources each contribute moderately to PM2.5 and residential and traffic emissions are small (negligible?) contributions. I also learned that residential burning has more of an impact on the oxidative potential because it is weighted more (based on equation1) than the other terms. What I did not learn is why secondary aerosol formation is the largest contributor to PM2.5 and oxidative potential. Are precursor emissions much greater than PM2.5 emissions? What role do oxidants play in controlling secondary aerosol formation? As a consequence of this result, should studies on source attribution to oxidative potential focus on precursor emissions and their source sectors? I did not see any DEHM results focused on carbonaceous aerosols or metals and the oxidative potential equation does not make use of that information. Although these aerosols are key culprits affecting human health, I do not understand

why the text about biomass combustion, coal burning, and traffic emissions discuss these details without supporting information from the DEHM results.

Organization, Clarity, Technical Comments

1. Lines 29-31. Please write this sentence more clearly: meteorological conditions contributed 46% and 65% to PM2.5 concentrations and oxidative potential, respectively, while anthropogenic emissions contributed 54% and 35% to PM2.5 concentrations and oxidative potential, respectively.
2. Line 87. “combing” should be “combining”.
3. Line 105. Please explain further how the DTT measurements are brought into this study. Do you mean via the positive matrix factorization done in a previous study?
4. Section 2. I think it would be good to reorganize the section. Currently section 2.3 interrupts the discussion of model simulations in 2.2 and 2.4. Perhaps the explanation of oxidative potential could be presented first followed by the control and sensitivity simulation configurations.
5. Section 2.2, meteorology setup. Please give more detail. I assume the nested domain at 50 km x 50 km horizontal grid spacing was used for the analysis, but this was not stated explicitly. I also assume that the outer domain provided initial and boundary conditions for the nested domain, but it was not stated. Was there a spin-up period for the simulations before conducting the analysis for 2014? How frequently was the meteorology data updated to reanalysis (or nudged to reanalysis)? Why was ERA5 chosen to drive the WRF model and not another global reanalysis product like MERRA or NCEP?
6. Section 2.2, chemistry configuration. It would be useful to give a description of the gas-phase chemistry and how PM2.5 is formed. For example, what hydrocarbons are included that would contribute to SOA formation? What sulfur chemistry and nitrogen oxides chemistry are represented that make sulfate and nitrate aerosol? Consider including the list of chemical reactions in the supplement. Are the aerosols represented with a bulk aerosol scheme?
7. Section 2.2, emissions. I suggest giving short descriptions of each emissions source: How are biogenic emissions, sea salt emission, lightning emissions calculated and what are they emitting? Is biomass burning from wildfires included? For anthropogenic emissions, it states EDGAR-HTAP is used, but it does not include a description of what that inventory emits and what grid spacing the inventory has. Yet, in section 2.4 there is a paragraph giving that information for the ECLIPSE emissions. It would be good to have similar information about each inventory so that the reader can better understand why there may be differences between simulation C1 and simulation C2. Especially useful would be to report the emission inventories’ annual values for China for PM2.5 and key precursors (e.g. SO2) as this will provide quantitative information on how EDGAR-HTAP and Eclipse differ.

8. Line 152. Please add *se* and its definition.
9. Line 155. “Industry source is primarily from specific industry processes” is not providing any insight as to what kind of industry or specific types of emissions. Please give more information.
10. Section 2.4.1. There are a number of reanalysis datasets available (e.g., ERA5, MERRA, NCEP FNL). Why were ERA5 and CESM chosen?
11. Line 167. “reanalyse” should be “reanalysis”
12. Line 172. “was first ran” should be “was first run”
13. Line 195. I suggest putting the sentence, “abs represents the absolute value” at the end of the paragraph, and rewrite to something more readable, e.g. “In the equations, the abs function represents the absolute value of the quantity in parentheses.”
14. Line 203. I suggest using “described” instead of “proposed”.
15. Line 225. Change “More and more” to “Previous”.
16. Line 257. Please maintain the same verb tense. I suggest “are mainly … are limited”.
17. Line 263. Change “were” to “are”.
18. Section 3.1. Please specify which model domain is being evaluated.
19. Line 275. Is it MME or MEE?
20. Diff_si-ob is used to express the difference between simulated and observed values. The way this term is written implies that it equals the simulation value minus the observations value. However, the values in Figure 4 appear to be observations minus simulated values. Could the authors clean up the terminology please.
21. Line 293. Change “are” to “were.
22. Line 293-294. Remove “with Figure S1a … December”.
23. Line 306-307. First sentence needs to be written better to something like: To learn about the spatial distributions of PM2.5 concentrations and OP, we plot maps of surface PM2.5 and OP for scenario C1 (Figure 6).
24. Line 308. The sentence, “The findings … and OP” is not needed.
25. Line 313. Is the term “urban areas” for low OP meant? Or is this area more rural?
26. Line 318. “northern residents in China right region” does not make sense to me.
27. Lines 336-340. I do not think so many significant digits are needed. I suggest using 85% instead of 84.8%, and likewise for the other numbers used here.
28. Line 349. Change to “illustrates”. Line 351. Change to “presents”.
29. Line 350-351. The sentence is not needed as it repeats the figure caption.
30. Line 392. Change “are” to “is”.

31. Line 430 and line 432. I recommend reducing the number of significant digits.

Figures and Tables

1. In all figure captions that show results, please include information on the time period shown (e.g., annual average) and spatial region shown (where appropriate).
2. Figure 3, figure caption. Please state what parameter is being plotted. I assume PM2.5, but it should be explicitly stated.
3. Figure 3 would benefit from having less white space. I suggest changing the maximum value to 120 or 150 ug/m3.
4. Figure 3. Are the points shown in panel b for the Dalhousie dataset for the same locations as the MEE observations? Or are there more points taking advantage of the gridded dataset?
5. Figure 4. Adjust the colorbar so that the whitest color is zero. That makes it easier to see differences between positive and negative values.
6. Figure 5. Are the observations shown in the figure from the MEE data or the Dalhousie reanalysis? Please note this in the figure caption.
7. Figure 5. It is difficult to discern the dashed and solid horizontal lines because the solid line does not extend from one edge of the colored region to the other. Is it possible to fix this?
8. Figure 8. Using 2 rows and 3 columns makes for smaller panels. I suggest using 3 rows and 2 columns (transposing the panels). I also suggest adding titles for each column, “PM2.5 (units)” and “OP (units)” and then panel labels that simply are the simulation name.
9. Figure 8. What are the insets showing in the bottom right of each panel? They are not discussed, so I suggest removing them.
10. Figure 10. Like Figure 8, I suggest using 5 rows and 2 columns instead of 4 rows and 3 columns. I also suggest adding titles for each column, “PM2.5 (units)” and “OP (units)” and then panel labels are the sector source (e.g., residential heating).
11. Figure 11. A more complete figure caption is needed: Percent contribution of different anthropogenic sources (traffic, industry, secondary aerosol formation, biomass burning, coal combustion) to total PM2.5 concentration and oxidation potential.

References

Yang, F., C. Liu, H. Qian, Comparison of indoor and outdoor oxidative potential of PM2.5: pollution levels, temporal patterns, and key constituents, *Environment International*, **155**, 2021, https://doi.org/10.1016/j.envint.2021.106684.