Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., referee comment RC1
https://doi.org/10.5194/acp-2021-256-RC1, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on acp-2021-256

Anonymous Referee #1

Referee comment on "Combining POLDER-3 satellite observations and WRF-Chem numerical simulations to derive biomass burning aerosol properties over the southeast Atlantic region" by Alexandre Siméon et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-256-RC1, 2021

Review of

Combining POLDER-3 satellite observations and WRF-Chem numerical simulations to derive biomass burning aerosol properties over the Southeast Atlantic region

By Alexandre Siméon et al.

This manuscript explores aerosol absorption in the South East Atlantic Ocean (SEA), both in cloud free conditions and above clouds, using WRF-Chem constrained by POLDER-3 observations. While the time frame is limited (half of one month), this manuscript is a valuable investigations of that snapshot in time, and illustration of the strengths and weaknesses of WRF-Chem for these conditions. I would rate the scientific significance of this (timely) manuscript to be excellent. There's a few things I would do differently regarding scientific approach and presentation, so I rate those as 'good' but am willing to rate higher if convinced otherwise. Overall, I believe this is a paper that should be published after addressing a few issues I'll describe below.

- In this review, I'll reveal myself as a remote sensing specialist with less expertise in models such as WRF-Chem. With that in mind, I'd like to question the two step method of constraining WRF-Chem with POLDER and other data. Is this common practice? What is the benefit compared to constraining in one step, in other words adjusting WRF-Chem all at once based on POLDER and met data? I could understand differentiating between constraints that come from the POLDER retrievals and other data. But that is not the case – AOD from POLDER is used in step 1, microphysical properties in step 2. The problem is that the POLDER algorithms work by simultaneously retrieving all of those parameters, which I would think should be considered when using those data to

adjust WRF-Chem. For example, if one were to impose an incorrect AOD in a POLDER retrieval, there would potentially be an incorrect retrieval of aerosol absorption (and vice versa). In a similar fashion, by initially modifying emissions to match AOD, and then later constraining aerosol optical properties might mean AOD no longer matches. I wonder if the POLDER/WRF-Chem match for AOD in Figure 16 and the ACAOD in Figure 17 could be further improved by constraining emissions a second time at this point. Which begs the question: why not just do it all at the same time. Am I missing something here?

1b. As an aside, I find the nomenclature for S1 "reference configuration" to be confusing. "Reference," to me, implies some external (to this effort) source of data, while instead S1 represents the first stage of constrained WRF-Chem results. Alternatives that may be better:

"partially constrained configuration" or "emissions adjusted configuration" or something else more informative. Furthermore, Figure 3 could be made more useful by also listing what is constrained in WRF-Chem at each step.

- Continuing the discussion on constraint methodology: why is it so approximate? In section 3.2.1 you adjust the APIFLAMEv1 BB emissions inventory by a factor of 1.5, while the ratio of WRF-Chem to POLDER-3/GRASP AOD(565) is 0.49/1.10 = 2.24. I imagine one reason could be that AOD does not directly relate to emissions estimates, but a discussion of how you go from that ratio value to 1.5 is needed. The same could be said for the choice of the Williams et al. (2007) BC refractive index, BCx2, OC/2.5 and 2.5% BrOC: is this the best match you can make? Could you do some more fine tuning to get a better constraint, or are the computational needs too much?

2b. By the way, I think table 4 would be more informative if the 'score' you used in section 3.3 were also included. And regarding how that score is calculated: why not use statistically appropriate parameters such as the Mean Absolute Error (MAE) or Root-Mean Squared Error (RMSE)?

- Another issue relates to how you visualize and assess the differences between POLDER and the various versions of WRF-Chem. I like the images in, for example, Fig. 7, 8, 11, etc. since they give a good intuitive understanding of the spatial context in comparing the two. However, it is qualitative. Since you've aggregated the POLDER data to the coarser WRF-Chem model grid, you have the ability to do a pixel by pixel comparison – in which case you can calculate and visualize more statistics. In most cases, you present a mean difference between POLDER and WRF-Chem, when instead you could show a histogram of differences, and make a mean-bias ("Bland-Altman") or some other sort of plot. Please just don't use the (unfortunately common) scatterplot and calculate correlation coefficients, as that would be statistically inappropriate (see for example, Altman and Bland, 1983, https://doi.org/10.2307/2987937). One of the reasons this matters is because AOD (and presumably ACAOD) is not normally

distributed (Sayer and Knobelspiesse, 2019, https://doi.org/10.5194/acp-19-15023-2019).

- Essentially, you're only assessing half of one month of one year. I realize that many of your choices are driven by computational resources, but how do you expect these results to fit within the annual cycle of BB in southern Africa, and how relevant is it considering inter-annual variability? I was expecting some discussion on this. These results need to be expressed in terms of their general usefulness. Perhaps a few words on computation resources needed for this analysis would be useful too, should somebody want to scale up to a full year or years.
- Is the RETRO anthropogenic emissions inventory, which doesn't cover the year of your analysis, the best source for that information?
- When considering POLDER measurement uncertainty, please recall that this often refers to per pixel estimates. In that sense, are the error bars in Figure 14 appropriate, since they are applied to geographically averaged values?

More minor points:

- Why are the reference wavelengths for SSA and AOD different for the POLDER/GRASP and above cloud products different? I agree that they're close enough to not matter, but it should be explained.
- Page 2, line 45: Clarify if "BBA direct radiative forcing" refers or global or SEAO.
- Page 14, line 125: I think you mean "specialty" not "specificity"
- Table 2 and discussion in text. I'm a little confused what you mean by "range of uncertainties" here. Does this really mean range of the values reported in the literature?
- Page 21, lines 507-8: Is the really the primary source of the differences, or could it also be due to the nature of your WRF-Chem constraint as well?
- Page 25, line 577: Is the score calculated for AOD, SSA, AND the top of atmosphere spectral observations? If so, the latter are used to derive the former, so the logic seems a bit circular.
- Page 29, line 657: You mention here and elsewhere a "spatial correlation coefficient" but give no description of what you mean by this or how it is calculated. Similarly on Page 30, line 687. Are those error metrics what I'm calling for in 3? Regardless, they need to be explained.