Comment on acp-2021-333
Anonymous Referee #2

Referee comment on "Modeled and observed properties related to the direct aerosol radiative effect of biomass burning aerosol over the southeastern Atlantic" by Sarah J. Doherty et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-333-RC2, 2021

This comprehensive study utilizes aircraft observations from the NASA ORACLES campaign during three biomass burning seasons to analyze differences in modeled properties of aerosol plumes over the Southeast Atlantic. The modeling comparison to in situ observations was conducted using two regional models and two global models for the same temporal periods and specified aircraft transects. The work further extends insights into the importance of adequately modeling aerosol, cloud, and optical properties by demonstrating the propagation of biases in parameterized and simulated direct aerosol radiative effect (DARE) which leads to a range of largely positive and marginally negative values. Approaches and findings presented here are compelling and point to modeled parameters that require tuning for improved simulation of biomass burning aerosol on regional and global climate. This manuscript is overall well-written, though ordering and structure of the results sections contribute to some lack of clarity in the manuscript. Nevertheless, this work is worthy of publication in ACP if the results can be presented in such a way that they support the conclusions.

General Comments

- The problem with the presentation of these results is that the authors have not provided calculations with quantitative results or consolidated figures that support their main findings (plumes too diffuse, underestimates in plume properties, COT well simulated). The result is a meandering collection of multi-page graphs that the reader is expected to visually integrate and compare in order to reach the cryptic qualitative assessments described in the results. The authors need to provide summary figures that support their summary statements, preferably with numbers to describe them rather than qualitative descriptors like “low” and “high”. I am not disputing any of these results, but I find it a disservice to readers to not provide them with figures that actually support, in a condensed way, the very generalized conclusions reached. Almost
all of the existing figure panels belong in the SI, as they are not discussed in the text.

- The next big concern for this submission is the presentation of DARE as one of the leading messages of the manuscript. The section as a whole is well-written and findings sound, though its current position is awkward following the summary of the modeling comparisons and reads as an addendum to the manuscript. Given the combined impact of aerosol-cloud-optical properties on DARE (the main assessment of this work), a more appropriate placement for this major section would follow the discussion of cloud fraction and optical thickness biases (Section 4.4) and before the summary and concluding remarks (Section 5), which is neither a summary (as it is long and winding) nor concluding (as it is followed by DARE). Many passages within Section 4.4 and Section 5 repetitively allude to expected findings that are provided in the DARE section (Section 6) and add to the already exhaustive length of the manuscript. Removing these passages or placing them within context of a reordered DARE section would improve the flow from significance of aerosol-cloud-optical properties to climate impacts.

**Minor Corrections:**

Line 102-113: Other than the supplementary modeling/forecasting information provided by UM-UKCA and ALADIN during their respective SE Atlantic campaigns, are there further reasons for using these models? A single regional and global model comparison to observations is a considerable effort. Are these additional models, which in some cases lack some comparison necessary variable fields, used only to expand the comparison discussion? A brief reference to this methodology choice would provide better perspective and support for the length of this manuscript.

Line 250: Provide the typical size range of accumulation mode aerosol, particularly as it pertains to biomass burning aerosol. Redemann et al. (2021)1 provides support of this claim.

Figure 3: Average symbols (black circles) are not clearly identifiable in these panels. Symbols with a bolder weight would improve the presentation and clarity of these figures. This should also be addressed in Figures 7-13 and subsequent supplementary figures (e.g. Figure S.4). This figure (and several others) was paginated as 3 separate pages, making it pretty unwieldy to review as a single figure. The authors should consider a format with less wasted space and higher density of information. Also, the scatter plot approach makes it challenging to see vertical trends for each color; I recommend adding a connecting line.

Fig. 4-14: as with Fig. 3, please rearrange, shrink, or break up into figures that each fit on a single page, with relevant axis labels and legends on the same page and provide captions written upright.
Figure 15: This figure seems superfluous in the main text of the manuscript given the inclusion of vertical profiles of SSA in Figure 14. The statistical inter-model comparison of SSA is interesting, but may be better suited for the supplementary text.

Line 658 “Future studies may want to use a synthesis of modeled and satellite-retrieved properties (e.g. of AOD) for a more robust analysis of sampling representativeness.” Is there a reason this is not done here, i.e. at a minimum a comparison to the avg+SD of the models shown? If the models are too disparate to make this meaningful, then why use them?

Line 967: “UM-UKCA has a significant low bias in BC at plume altitudes in 2016 and a smaller low bias in 2017 (Figure 9).” This statement (and many others like it) are difficult to support from the figure noted (9). This reflects two problems (1) There are no multi-flight-type results presented from which to infer the yearly differences noted, suggesting perhaps that one should be able to eyeball this to see and quantify the low bias? And (2) There is no metric or test presented for significance of the 2016 result vs. 2017. Both of these issues are present throughout the manuscript, with figures showing individual results but text describing trends amalgamated over several graphs.

Of course there are many ways to consolidate these results (by region, by altitude, etc.), and the way chosen will of necessity be limited to the particular aim of this work. But the failure to present any such consolidation results in a mismatch between the text and the figures. A much more useful paper would present the consolidated papers and move the detailed figures and tables to the SI.

Line 1123: This is super interesting; how does it compare to past comparisons, e.g. Heald et al. 2008(?) or similar?

Line 1136: Correct “humification” to “humidification”.

Line 1233: “Except in the 2018 Meridional2 transect, WRF-CAM5 δ□□□□ep is generally 30-40% lower than measured by HSRL-2.” As l.967, where is this shown?

Line 1309: “The bias also has a less consistent dependence on altitude in 2017 (Figure 11). In 2018, the higher ambient RH could be

1310 compensating for some of the low bias in dry aerosol δ□□□□ep.“As l.967, where is this shown?
In the 2016 Diagonal transect this produces significant low model biases for 3-5 km altitude and high biases for 2-3 km and 5-6 km (Figure 12; Table 1), much. What is “significant low” and what is “high”? Which values or plots are referenced?

Lines 1601-1603: This sentence is hard to follow and should be revisited for clarity. Are the authors trying to say a larger number of small particles less than 1000 nm dry diameter as is the case in the biomass burning plume?

Line 2055: Remove “the”

Lines 2201-2202: Suggest “to produce values of DAREavg that is a factor of” to “to produce a DAREavg that is a factor of”.

References:

1https://acp.copernicus.org/articles/21/1507/2021/