Response to reviewers for Chang et al. (acp-2022-496)

We thank the reviewers for their constructive comments and suggestions. We are glad to hear that the reviewers find this work well-written and suitable for consideration to be published in ACP. Our responses are in bold-free texts and the reviewers' comments are in bold texts.

<u>Reviewer 2</u>

This well written manuscript does a thorough evaluation of the ability of five ESMs to represent the partitioning of the Aerosol Optical Depth (AOD) between the free troposphere and the marine boundary layer over the southeast Atlantic. It takes advantage of the instruments that were deployed to measure aerosol properties during the two campaigns ORACLES 2016 and ORACLES 2017.

The description of both the models under study and of the instruments used to retrieve the AODs at different heights is well conducted but it leaves the reader wandering about what we have learned about the processes that explaining the discrepancies between model and observations that are described in the conclusions. A complete study requires to consider what are the sources of uncertainties in these models and how they differ from one another both in how they parameterize their processes and also on how they consider the different source inventories for fires that emit these aerosols. A much stronger paper would emerge if the authors took up the (difficult) task to explain why models either underestimate either overestimate the AODs and what processes the modelers should focus on to improve on the results. This extra work will make up for a much better paper.

We thank the reviewer for the constructive comments and suggestions. We are glad to hear that the reviewer finds this work well-written. We significantly expanded Table 1 by elaborating on how emissions, transport and deposition processes are handled in each model. We also added a section called "Discussion on model deficiencies for future investigations" that addresses model deficiencies that may explain the factors for AOD errors for future investigations. The discussions will provide a starting point for the needed step of additional model studies to explore how each factor is affecting AOD. Here is the additional text for that section:

"ESMs are complex and nonlinear systems, so AOD errors are likely caused by numerous factors. Identifying the exact causes of AOD biases is challenging and entails a detailed examination of model source codes. Here, we present aspects of the models that may explain their biases in simulated AOD relative to those measured by airborne lidar, which establishes a starting point for a future in-depth investigation. The assimilation of clear-sky MODIS AODs in the two assimilation systems (i.e., GEOS-FP and MERRA-2) may explain their better performance compared to other models in simulating AODs, especially in August 2017. Despite a lack of MODIS clear-sky AOD retrievals over regions with expansive cloud presence, such as in the austral spring of the SE Atlantic, AOD assimilation is still beneficial for minimizing AOD errors in ESMs.

The mean and median AOD and the AOD fraction in the FT in WRF-FINN generally agree well with those from aircraft measurements. WRF-FINN is also the only model in this study that includes a plume rise parameterization. The importance of the inclusion of a plume rise model for simulating high AODs in this region is unclear since fire emissions in southern Africa already take place at elevated altitudes. Nonetheless, the smoke top heights in the remaining models generally agree with those from lidar measurements (Shinozuka et al., 2020).

The rate of primary organic aerosol (POA) removal and the secondary organic aerosol (SOA) production influences the simulated AOD (Hodzic et al., 2020). For example, the negligible production of SOA in WRF-CAM5, GEOS-FP, MERRA-2, and ALADIN may be contributing to a low bias in simulated AOD. For GEOS-Chem, GEOS-FP and MERRA-2, their aerosol optical properties are assumed to be fixed and do not account for particle evolution during transport. Even though the production of SOA is introduced in the other models, the assumed processes may be oversimplified such that its production is based on precursors at a fixed time-scale without a detailed consideration for chemistry. Moreover, these models do not treat photochemical loss of SOA as shown by its excessive OC according to Shinozuka et al. (2020). Errors in the treatment of aerosol hygroscopicity may also play a crucial role in the aerosol evolution and subsequent AOD biases. Although the AOD fraction in the FT in WRF-CAM5 has a good agreement with lidar measurements, Shinozuka (2020) found that the PBL height of this model was a few hundred meters higher than that in lidar cloud-top measurements in September 2016, possibly leading to overactive entrainment and aerosol removal. While the selection of emission inventory alone impacts simulated AODs (Pan et al. 2020), the use of monthly emission inventory in both EAM-E3SM and ALADIN instead of diurnally-varied emissions as in other models could further be responsible for some of the errors. These deficiencies suggest that AOD errors in each model are likely driven by multiple factors, and a more in-depth model-specific analysis would be needed to investigate model deficiencies that leverages multiple degrees of freedom.”