Response to the Interactive Comments During the Discussion Phase

Manuscript Ref: acp-2016-790

Comparing multiple model-derived aerosol optical properties to spatially collocated ground-based and satellite measurements

I. B. Ocko and P. A. Ginoux

We sincerely appreciate the thoughtful reviews of our manuscript, and thank the referees and Editor for their time. The suggestions have undoubtedly and considerably enhanced the manuscript.

Specifically, we have improved the analysis by (i) developing a new section with regional evaluation of model performance with two new instruments and three new figures, (ii) providing quantification of results throughout the text, and (iii) converting AERONET AOD to 550 nm from 440 nm. We have also refined and reformatted the conclusions, and added 12 new references.

Below, we have responded point-by-point to comments and provided information on the modifications in the text.

Responses to Interactive Referee Comment #1 (Anonymous Referee #2):

Comment 1: In this manuscript, "Comparing multiple model-derived aerosol optical properties to collocated ground-based and satellite measurements" the authors compare two different versions of the NOAA GFDL model with measurements of aerosol optical properties. They demonstrate the importance of looking at more than just the AOD when assessing the model performance and highlight deficiencies in the model representation of aerosol, such as biomass burning aerosol not lofted high enough in either model. The research clearly highlights the difficulties in modeling basic aerosol seasonality and loading in polluted regions. However, most AEROCOM studies do look at more than just the AOD when assessing the aerosol in models (e.g. Kinne et al., 2006, Huneeus et al., 2011). Therefore, I'm not sure how novel the multiple-metric approach truly is, a point that is highlighted in the abstract and throughout the work. The research presented is valuable but some aspects of the research need revisiting and the conclusions need improving.

> **Response:** We thank the referee for their time in thoughtfully reviewing our manuscript. While AeroCom studies do look at different variables to perform regional analysis, this is different than comparing numerous aerosol parameters with data at one location. We have clarified this in the introduction (lines 2.22-2.24): *"However, most studies do not take advantage of all available datasets beyond regional analysis (Kinne et al., 2006; Huneeus et al., 2011), even though a multi-dataset approach can provide a more comprehensive picture (Miller et al., 2011)."*
>
> Our method is a step toward closing model uncertainties, where all parameters are constrained with observation. As these properties vary spatially and temporally, it is more appropriate to use spatially collocated data rather than a regional average. On the other hand, comparing at a larger scale has its own merit and we have therefore revised our manuscript by adding a new section with broad regional analysis (Section 4.2.1), three new figures (Figures 3, 4, and 5), and comparing with data from two more instruments (MODIS and MISR).

Comment 2: I'm left feeling that the model representation of aerosols is generally poor in the regions compared, and that this might be a combination of emissions (definitely for biomass burning), spatial resolution, potentially optical properties, aerosol size distribution etc. While the authors show that comparing multiple metrics with observations can provide more insight, there is little in the way of concrete evidence that the those insights have helped improve the understanding of the discrepancies between model and observations. I don't mean to be overly critical, and realize the simulations are time consuming, but I think the authors must justify their choice to stop at the point of speculation and not perform further simulations to try understand which of the many plausible causes actually contribute to the observed discrepancy. Key findings should be presented more concisely if possible, and more from the viewpoint of the underlying causes rather than the models being X% higher or Y% lower than the observations which is of limited use to the reader.

Response: We thank the referee for their feedback. To show that overall model performance is not poor, we have added a new section (Section 4.2.1) with three new figures (Figures 3, 4, and 5) to evaluate model performance on regional scales. This analysis shows that the models' representations of aerosols is satisfactory in relatively clean regions, and provides more context for why the seven locations were chosen as key regions for further evaluation.

This work is an attempt to analyze as best as possible the strengths and weaknesses of the aerosol properties which are forcing two prominent, related climate models. Similar to many climate centers, we follow precise model setup and emissions scenarios as guided by the IPCC so that model comparisons can be made. While it is not our role to test emission schemes or other climate tuning parameters, we are able to provide feedback to improve emission scenarios or aerosol properties. We have clarified this in the text (lines 2.29-2.31): *"By characterizing model strengths and weaknesses, we are able to provide feedback to improve emission scenarios and aerosol properties for future model generations."*

Further, we have presented our key findings more concisely in the conclusions section by restructuring the text, providing more meaning to the results, and adding future research directions based on our findings.

Comment 3: If I understand correctly, the AERONET observations used are for 440nm whereas the model is at 550nm. This will cause a general high bias in the AERONET AOD relative to the models. The difference may be small where coarse aerosol dominates but this will increase up to maybe 25% in regions with fresh, fine aerosol, such as biomass burning regions. I don't think the current comparison is rigorous and recommend converting AERONET AOD to 550nm. AERONET provides AOD at multiple wavelengths (and the Angstrom Exponent) so it is trivial to calculate the AERONET AOD at 550nm.

Response: We thank the referee for this suggestion, and have converted the AERONET AOD from 440nm to 550nm using the Angstrom component. Figures 6 and 7 have been modified to show the AERONET data for 550nm, and all comparisons within the text have been updated. We have also calculated the correlation coefficients between AERONET and the models for all sites and parameters to provide more quantitative assessment. Overall, converting AERONET AOD from 440 nm to 550 nm lowered total AOD in all industrial and biomass burning sites.

Comment 4: Also regarding the comparison with AERONET, is the comparison of the closest grid box to the AERONET site, or has the model grid been interpolated to the exact site location? Lack of interpolation may make a substantial difference where there are strong gradients in aerosol.

> **Response:** The comparison between the model and AERONET is indeed the closest grid box, with no interpolation. This was written in the former Section 2.1, but we have now removed it based on feedback from another referee that says to their knowledge hardly any model interpolates grid box data when doing comparisons because the model uncertainties are often larger than the concentration gradient in the grid box. We have instead provided clarity and discussion of this in Section 3 (lines 7.5-7.9): *"Lack of interpolation of model data in polluted regions may introduce a bias in locations with strong aerosol gradients; however, interpolation is rarely employed for comparisons with observations because the model uncertainties are often larger than the concentration gradient in the grid box."*

Comment 5: With the CM3 model, it is difficult to understand how much of the discrepancy with observations might arise from the climate model meteorology (rather than using reanalysis fields). The authors do average over a 5-year period using the model, but it would be useful to see the interannual variability of the models on Figure 4 & 5 and some understanding of the interannual variability in the CALIOP observations.

> **Response:** This is a good point, as the discrepancy may be in large part due to climate meteorology. In addition, CM2.1 and CM3 have different physics and produce different climates (cf. Donner et al., 2011). It is also important to note that these comparisons are made with climate models which are unable to reproduce specific synoptic events. We add in the text the following (lines 17.31-18.2): *"While some of the discrepancy between CM2.1 and CM3 is due to different meteorology (Donner et al., 2011), differences between model and observations also arise because the climate models are unable to reproduce specific synoptic events."*

Comment 6: It would be interesting to use the difference between the model and the observations to understand how the error in the models translates into uncertainties in the radiative effects and the interhemispheric forcing asymmetry. These are discussed qualitatively, but is it possible to expand this into some quantitative assessment using other model output fields (surface and TOA radiative effect, etc.)?

> **Response:** We have looked at the clear-sky downward shortwave radiation, and it is generally larger in CM3 than CM2.1 and closer to observations from the Baseline Surface Radiation Network (Donner et al., 2011). The increases in clear-sky downward shortwave radiation are due to reduced aerosol AOD in CM3. Although correlation of AOD decreases with CM3, from a climate perspective Donner et al. (2011) showed an improved agreement of CM3 simulations of downward clear-sky surface shortwave radiation, optical depths, and coalbedo with BSRN and AERONET. These improvements made the authors conclude that the direct effects of aerosols are more realistically simulated in CM3. A quantitative assessment of how model biases translate into radiative forcing uncertainties is currently beyond the scope of this paper, but an excellent idea for

a future paper, and we appreciate the suggestion. We have added this as a possible future research direction in the conclusions section.

Comment 7: I do not think the bullet-point conclusion format works well when the results are not concise. Splitting some of the conclusions into bullet points while others remain in paragraph form sees arbitrary. Please consider revising the fragmented conclusions into a more holistic discussion of the findings and how future research should proceed based on these findings.

Response: We thank the referee for this feedback. We have greatly improved the conclusions section by restructuring the text and consolidating the results. We have also included discussion on future research based on our findings.

Comment 8: Minor Comments

pg1 ln 29 Aerosol can travel 1000s of km in a week, so I wouldn't say it is localized around sources. Perhaps more localized than GHGs.
pg4 ln 9 Include a reference for the optical properties of BC and dust discussed.
pg5 ln 14 Add "(see Section 3.1)" regarding "computed offline" to let the reader know this will be explained.
pg5 ln 29 Remove extra period.
pg 8 please add to the description how SOA formation is treated. This is simplified and often underestimated in many models so is a potential source of discrepancy between the observations and the models.
pg 9 ln 8 Make it clear to the reader why using different years is not expected to be an issue.
pg11 ln 31 "have better magnitudes" - please rephrase.
pg12 ln 9 Remove extra punctuation
pg18 ln7 "Very nice job", please reword.
pg19 ln29 "poor emissions databases" this is very vague. Are any of the examples given included or not?
Figures 4 & 5
-in the caption, please state what the error bars represent.
-I may have missed it in the text, but the reason for missing data at Alta Floresta and other sites should be stated. I assume it is the lack of high enough AOD during that season for SSA retrieval?
-is it possible to add CALIOP AOD to these? This would be helpful when AERONET and CALIOP are often compared qualitatively in the text.

Response: We thank the referee for helping us improve the manuscript by clarifying major and minor points and tightening the text. We have made all of the above modifications. However, while we have plotted the CALIOP AOD in comparison to the models and AERONET, we have ultimately decided not to include it in the former Figs. 4 and 5 because it is known to be very problematic when the extinctions are integrated vertically, and thus may provide misleading information.