This manuscript presents a thorough evaluation of the performance of the GEOS global atmospheric chemistry model with the GOCART aerosol parameterization. The model output is compared with in situ, airborne measurements of aerosol compositional, microphysical, and optical properties made in the vicinity of the Philippines as part of the CAMP2EX project in 2019.

The output from the model is compared with the observations over project-wide averages, and for some specific flights in biomass burning smoke, and includes both extensive and some intensive parameters. The importance of assimilation of AOD values is examined, as are the effects of a newer convective parameterization scheme. While not comprehensive, this analysis is enough to identify discrepancies in the assumptions regarding aerosol composition, hygroscopicity, and size distribution in the GOCART parameterization. These discrepancies are large enough to be a concern, and indicate that perhaps the GOCART approach, with its one-size-fits-all static parameterization of the properties of different aerosol components, should be replaced with a more physically based and interactive aerosol parameterization.

This manuscript is suitable for publication in ACP following revisions that fall somewhere between major and minor (I've described them as minor). I have one significant comment below that I would definitely like to see addressed prior to publication, followed by a number of minor comments. Despite the number of comments, this is a generally well written paper, and an interesting and enjoyable one to read.

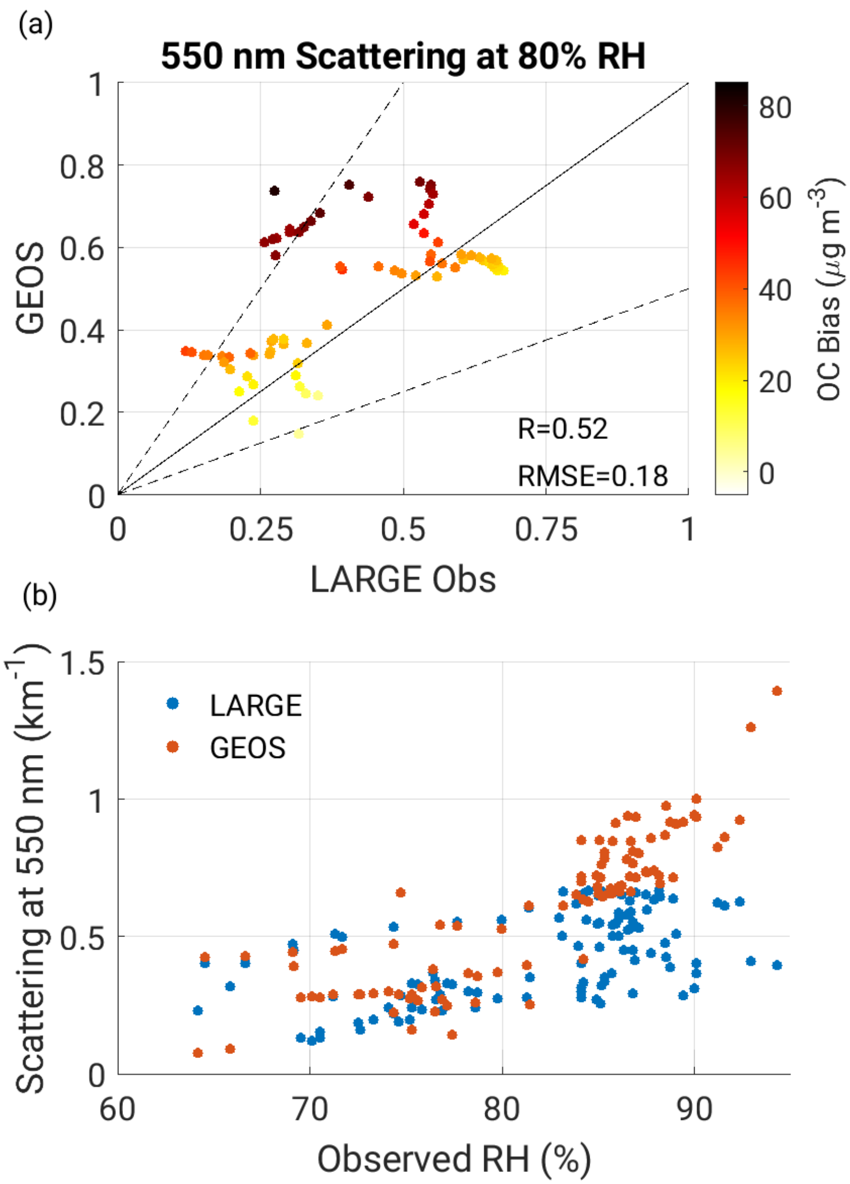
Thank you for taking the time to read our manuscript and provide feedback! Please find our responses to the comments written in red below. Of most importance, a new figure has been added to the manuscript to further investigate scattering at 80% relative humidity and the extrapolation that is performed in the observations to arrive at ambient scattering under humid conditions.

Major Comment:

Much of the analysis relies on comparison of the model output with extinction values at ambient RH. These are referred to as "measurements", but in fact they are derived values that rely implicitly on an underlying assumption regarding aerosol hygroscopicity. As documented in the headers of the data files from the LARGE measurements for this campaign, scattering (which vastly dominates extinction here) was measured at two fixed RH values: ~20% and ~80% RH. The ambient scattering is then calculated by assuming a functional form (the gamma parameterization), which is just a power law fit. In other words, the ambient scattering (hence extinction) is based on fitting a parameter to the two measurement points, then extrapolating to the ambient RH. The error in this approach is probably not large for RH values lying between or close to the two measurement RH values (20% and 80%). However, in the CAMP2EX profiles, the ambient RH exceeds 90% in the upper half of the boundary layer, and this is where the greatest contribution to extinction lies. In this high-RH region, the power-law hygroscopic growth curve pitches very sharply upwards, and small errors in the assumed shape of this growth curve can amplify to very large errors in calculated ambient extinction. So this comparison is not optimal, because the ambient extinction values will have large, and unknown, uncertainties in the high-RH region that dominates AOD.

One of the goals of the comparison of the measured and modeled extinction is to determine if the model hygroscopic growth is consistent with the measurements. To accomplish this, it might make more sense to compare the model to the measurements for the actually measured extinction values at the 20% ("dry") and 80% ("wet") conditions. That way the comparison is not between the modeled values and those derived from the measurements with an assumed shape to the hygroscopic growth curve. Alternatively, you could plot the full f(RH) curve the model would produce (for a given location and time), from 20% RH to ~95% RH, and compare that with the same curve provided by the gamma parameter calculated from the LARGE measurements. This would be informative, and would allow you to compare the extinction at the measured 20% and 80% RH values as well as look at the response at higher RH values. The ratio of these two curves would indicate where there might be relative biases (although whether these lie in the model or in the gamma parameterization assumption might not be so clear).

Thank you for pointing out the extrapolation for scattering above 80% RH, which is indeed the case. The contribution of absorption to the total ambient extinction does not account for hygroscopic growth as indicated by the header in the data files. As suggested, a new figure has been added to the manuscript (and copied below) that shows scattering at 80% from the observations and the model. The top panel demonstrates that there is decent agreement in scattering between LARGE and GOCART2G when the overestimation of organic carbon is not exceptionally large. Otherwise, there are data points that indicate an overestimation in scattering, somewhat like the overestimation in extinction when using the bias corrected relative humidity shown in Figure 10. The bottom panel illustrates that there is a stronger relationship between scattering and relative humidity in the model when the observed relative humidity is above 80%. Given the extrapolation for the observations, it is difficult to determine whether the model or observations are responsible for all or some of the bias.



Minor Comments:

There are a few places where there are typos or where the manuscript could be edited for clarity.

1) Lines 81-85. The FIMS and HSRL2 are introduced by name, but the text doesn't say what they measure. That information appears in lines 100-104. I was trying to figure out what a FIMS measured throughout the text between these lines. Suggest moving the latter section up to lines 81-85.

Clarification has been added to the line in question, which now reads as “Here, we make use of the NASA Langley Aerosol Research Group Experiment (LARGE) suite of instruments, particle size distribution from the Fast Integrated Mobility Spectrometer (FIMS; Kulkarni and Wang, 2006; Wang et al., 2017), and aerosol scattering profiles from the 2nd generation High Spectral Resolution Lidar (HSRL2) (Burton et al., 2018) as summarized in Table 1."

2) Line 117. Change "hydroscopic" to "hygroscopic".

This should have been hydrophobic and has been corrected.

3) Line 123. Define sigma as the geometric standard deviation.

This has been updated.

4) Line 130. Change to, "The underlying meteorology from GEOS is used for horizontal and vertical transport and deposition of all of the aerosol species, as well as wind-driven emissions of dust and sea salt."

This has been changed as suggested.

5) Line 193. Define "lidar ratio".

Lidar ratio is now defined.

6) Lines 211 and 223. "Optical array" has a particular meaning in optical design. Suggest changing to "optical properties instruments" or something similar.

Array has been changed to instrument suite.

7) Line 213. Change to "fine particles that are efficiently sampled".

Fixed.

8) Lines 222-234. Here is where you could expand on the model/measurement comparison of extinction as a function of RH and focus on the actual full frh curve or compare the extinctions at 20% and 80% RH.

Thank you for the suggestion. We opted to discuss the scattering at 80% RH in Section 3.2 since that is where f(RH) is discussed.

9) Line 241. I don't understand this sentence. "Analysis increment"?

Analysis increment quantifies the adjustment in aerosol mass that occurs in response to the assimilation of aerosol optical depth. This sentence has been modified to "The assimilation of aerosol optical depth results in an increase of black carbon mass and a doubling of the positive bias in the mass concentration within the boundary layer."

10) Line 258. Please add a comma between "region" and "or".

Done.

11) Line 273. Please change to, ". . .represented that percentage of organic carbon well; however, it struggled with. . . ."

Fixed.

12) Line 281. In the heading for Table 2, I believe this in not "total aerosol mass" because it excludes BC.

Yes, “total” has been removed from the heading for Table 2 (now Table 3).

13) Line 299. I believe the values for the extinction measurements are 20% and 80% RH, at least according to the file headers.

Yes, correct. Gamma is found using 40% however f(RH) is determined by using gamma at 20% and 80%, which is now noted in the text.

14) Line 313. Saying the GEOS values for SSA are "below 0.96" doesn't do justice to the magnitude of the discrepancy, which is quite large. Maybe say they range from "0.9 to 0.96". The co-albedos disagree by a factor of 2-4; this is quite large in the context of direct radiative effects.

This sentence has been updated to “Nearly all points for the observations have a SSA greater than 0.98 while the SSA in GEOS ranges from ~0.9 to ~0.96”.

15) Line 323-324. It's true that the extinction could be juiced up by making the modal diameter larger and the width wider, but the assumed standard deviation of ~2+ is already quite a bit larger than literature values would support. You might want to lead into the next sentence by saying, "We will examine the in situ measurements to see if such changes could be justified. We begin by looking at . . . ."

This sentence has been modified to “We begin by looking at two flights, RF9, the flight examined in Section 2.2, and RF10, which also captured aged biomass burning aerosol but downstream in the Philippine Sea on the following day, to determine if changes to the particle size distribution are justified.”

16) Line 337. The sentence beginning "The primary peak" is confusing. Please define what you mean by the "primary peak? It's being "shifted toward a larger radius" than what? Are these the measured or GEOS size distributions you're talking about? Please clarify.

This sentence has been modified to “The primary peak in the size distribution, centered at ~0.015 μm in the observations, is shifted towards a larger radius in GEOS.”

17) Line 368. I'm not sure what "mass piling up at the top of the PBL" means. I don't think mass (of air or of aerosol) can "pile up", at least not without increasing air density a lot!

This sentence has been updated to “Using aerosol mass concentration as a tracer indicates an overestimate of aerosol within the boundary layer in GEOS, particularly at the top of the PBL such that the aerosol is unable to sufficiently penetrate into the free troposphere.”

18) Lines 391-395. These sentences are unclear. What "concern" about the size distribution could be "rectified"? Please be specific, e.g., "agreement between modeled and measured extinction at high RH could be improved if the modeled size distribution had a larger modal radius and/or a larger standard deviation. However, these adjustments are not supported by the FIMS size distribution measurements. In fact, the observed mode radius is in excellent agreement. . . ."

This section of the conclusions has been re-written as copied below:

“Agreement between the modelled and observed extinction could be improved by adjusting the assumed particle size distribution for brown carbon to have a larger mode radius or a wider distribution. A comparison of the assumed particle size distribution to observations from FIMS indicated that the current mode radius is too small. The observed mode radius is in excellent agreement with observations from past field campaigns, however the modal width for CAMP^{2}Ex is larger. A limitation of GOCART, however, is that the particle size distribution cannot vary in time or space for a given aerosol species. This contradicts the variability seen within particle size distribution during the low-level flight segment from CAMP^{2}Ex's RF9 as well results from other field campaigns that indicate the particle size distribution of biomass burning aerosol changes with respect to median diameter and modal width as smoke ages (June et al., 2022). If the assumed particle size distribution for brown carbon in GEOS were to be modified such that the mode radius is larger, it would need to be thoroughly evaluated as a change in the particle distribution would also impact fields such as the single scattering albedo and Angstrom exponent. Ultimately, a more physically based aerosol module would need to be used for GEOS to accurately represent the variability in the particle size distribution."

19) References. Please ensure that all references are compliant with Copernicus formatting guidelines. For example, some of the references (e.g., Burton et al., 2012) have capitalized titles, while most do not. This is a result of reference management software, which always needs to be thoroughly checked manually.

The citations have been standardized such that titles are no longer capitalized.

20) References. I'm not sure Schill et al. is citable--it's an unpublished conference presentation.

A section has been added to the supplemental material to further describe the particle size distribution developed by Schill et al. in lieu of the citation, which has been copied below.

Merging the Particle Analysis by Laser Mass Spectrometry (PALMS) number fractions with an independently measured, quantitative size distribution to make species-specific (e.g., biomass burning, dust, sea salt) size distributions has been detailed in several publications (e.g., Froyd et al., 2019; Brock et al., 2021). The biomass-burning-only size distributions from the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (Toon et al., 2016) and NASA Atmospheric Tomography (Thompson et al., 2022) missions are used. Using both missions allows us to assess the biomass-burning-only size distributions from plumes <1 day old to biomass burning aerosol that have aged in the atmosphere ~30 days. We find that, regardless of age, the biomass burning size distribution can be described by a single lognormal mode. In this work, we use the size distributions from fires <5 days old. The rn, median and σ values from these size distributions are 0.1175 μm and 1.3, respectively.

Brock, C. A., Froyd, K. D., Dollner, M., Williamson, C. J., Schill, G., Murphy, D. M., . . . Wofsy, S. C. (2021). Ambient aerosol properties in the remote atmosphere from global-scale in situ measurements. Atmospheric Chemistry and Physics, 21 (19), 15023–15063. doi: 10.5194/acp-21-15023-2021.

Froyd, K. D., Murphy, D. M., Brock, C. A., Campuzano-Jost, P., Dibb, J. E., Jimenez, J.-L., . . . Ziemba, L. D. (2019). A new method to quantify mineral dust and other aerosol species from aircraft platforms using single-particle mass spectrometry. Atmospheric Measurement Techniques, 12 (11), 6209–6239. doi: 10.5194/amt-12-6209-2019.

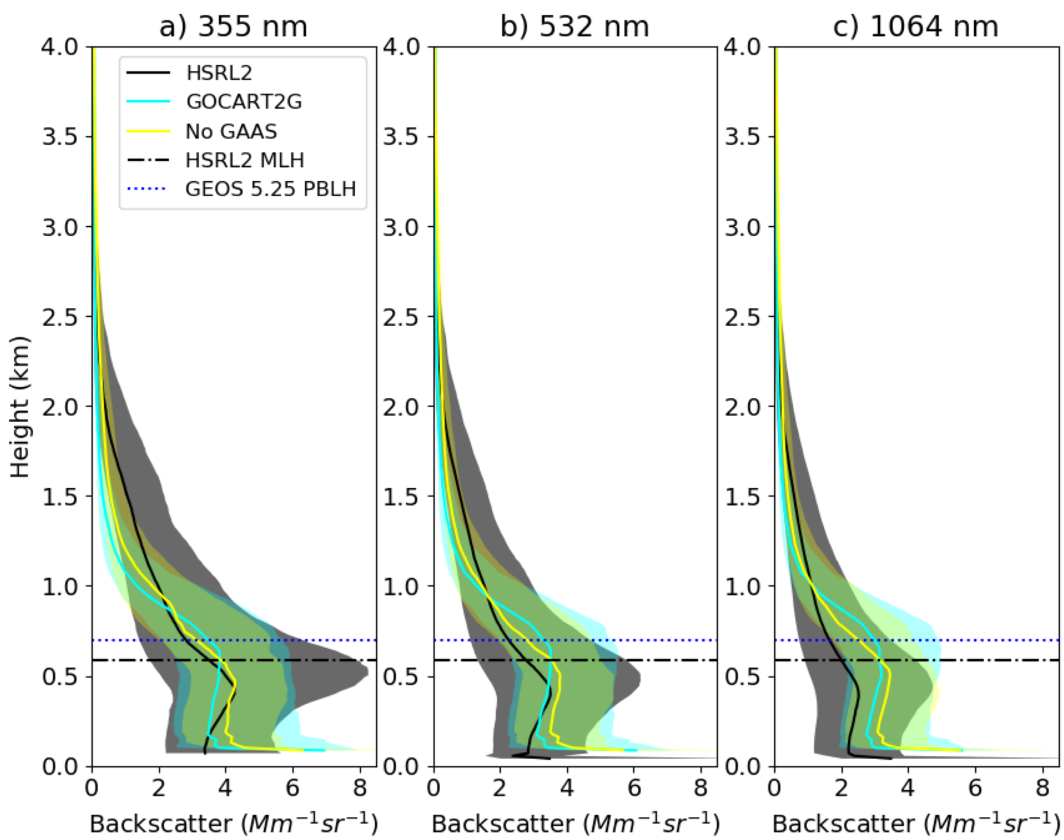
Thompson, C. R., Wofsy, S. C., Prather, M. J., Newman, P. A., Hanisco, T. F., Ryerson, T. B., . . . Zeng, L. (2022). The NASA Atmospheric Tomography (ATom) Mission: Imaging the Chemistry of the Global

Atmosphere. Bulletin of the American Meteorological Society, 103 (3), E761–E790. doi: 10.1175/BAMS-D-20-0315.1.

Toon, O. B., Maring, H., Dibb, J., Ferrare, R., Jacob, D. J., Jensen, E. J., . . . Pszenny, A. (2016). Planning, implementation, and scientific goals of the Studies of Emissions and Atmospheric Composition, Clouds and Climate Coupling by Regional Surveys (SEAC 4 RS) field mission. Journal of Geophysical Research: Atmospheres, 121 (9), 4967–5009. doi: 10.1002/2015JD024297

21) Figures. The figures are generally very nicely done, clearly labeled. However, I'm not sure of the ability of those with color impairment to read the shaded vertical profiles (e.g., Figs. 3-6). Also, in these figures the horizontal lines indicating HSRL2 MLH and GOES PBLH are difficult to discern; the black and blue colors are quite close. Could one line be made dotted? In Fig. 4 you might need a 3rd line type for clarity.

The figures were uploaded to https://www.color-blindness.com/coblis-color-blindness-simulator/ and checked for clarity for those with a color impairment. Modeled PBL height now uses a dotted line as shown below.



22) Figures 13, 14. It might be nice to fit a lognormal to the FIMS peaks, then you could directly compare the mode diameter and standard deviation with the model. My eye says that the standard deviation for

the measurements is less than the very broad modeled values, but I can't be sure without fitting. I much prefer the aspect ratio of Fig. 14 to that of Fig. 13; could Fig. 13 be made with side-by-side plots, rather than vertically stacked ones? This would make Fig. 13 and 14 look more similar.

The sub-flight variability and averaging make it difficult to accurately fit a lognormal distribution to the observations in Figure 13, but a lognormal fit was performed for the flight segment shown in Figure 14. The low-level segment through the smoke plume indicated a median radius of 0.0995 μm and a modal width of 1.77, which would be a narrower distribution with a larger radius than what is assumed by GEOS. The aspect ratio of Figure 13 was modified and is shown below.