Reply to Reviewer 2

Z. Chen et al.
zhong.chen@ssaihq.com

In this paper, the authors use a gamma distribution fit to a sectional aerosol model run (coupled to a GCM) to derive aerosol scattering phase functions. These are used in the radiative transfer forward model for the retrieval of aerosol extinction coefficient profiles from the limb scattering measurements made by the OMPS Limb Profiler. The assumption of the aerosol size distribution in the radiative transfer model for aerosol extinction retrieval from scattered light measurements is a long standing problem, and it boils down a basic lack of information in the remote sensing measurement to make a bias free retrieval. Other groups working on similar limb measurements with SCIAMACHY and OSIRIS have tried other forms of the size distribution with no real agreement or even criteria for what is best.

We appreciate the referee's comments and provide point-to-point responses below in regular font.

In general this is an insightful paper and the sample data sets that are presented show improvement over the previous version of the OMPS retrieval that used a bi-modal lognormal distribution. However, there are two related major points of concern. The first is that the main source of proof for improvement presented in the paper is analysis of a one month test data set (plus the six month time period used for SAGE III intercomparison). It does indeed seem that the gamma distribution is an improvement over the previous bi-modal assumption from V1.0; however, as the authors point out, the actual aerosol size distribution is a strong function of time, latitude, volcanic perturbation, etc., and it could very well be the case that the retrieval is worse at other time periods that are not analyzed.

Authors: We now use a one-year test data set covering all of 2017 for the ASI residual analysis shown in Fig. 7 and Fig. 8, in addition to the 1-month data set following the Calbuco eruption used for Fig. 5 and Fig. 6.

The V1.0 bi-modal assumption certainly had difficulties (choice of 5 free parameters, uncertainty in fitting OPC data) and the gamma function is demonstrably better, however, what about the simple unimodal size distribution assumed by the OSIRIS and SCIAMACHY algorithms? These are also simple 2 parameter distributions that roughly match the (measured, not modelled) background aerosol state. The corresponding phase functions for these distributions should be compared to the gamma distribution used here, and a clear case made for the use of gamma distribution.

Authors: We now include the parameters for the OSIRIS and SCIAMACHY size distributions in Table 1, and show the corresponding phase functions in Figure 2. While we did not create test data sets using those size distributions, we note that their differences from the V1.0 phase function in Fig. 2 are in the same direction as the gamma distribution (i.e. lower value at backscattered angles), but smaller in magnitude. So we would expect that processing LP data with one of these unimodal size distributions would yield less change relative to our V1.0 product than the gamma distribution adopted for V1.5. The improved agreement with SAGE III data for V1.5 extinction data shown in Fig. 10-13 suggests that we would not want to adopt a size distribution that produces less change in extinction.

Since the bias is such a strong function of solar scattering angle, which for OMPS is essentially a latitude dependence, it might be the case that a "better" choice for OMPS is not a better choice for an instrument in a different orbit. Overall, users of limb scatter aerosol products would benefit from uniformity in the algorithm choices between the various groups, or at least publications that show how/why the assumptions are different.

Authors: We agree with the reviewer's comment regarding the instrument-specific nature of our choice of size distribution, and have added the following text at the end of Sect. 5.

"Because of the large variation of phase function with scattering angle (Fig. 2) and the strong dependence between scattering angle and latitude for OMPS LP (Fig. 4), the size distribution determined here is not necessarily the optimum choice for a satellite instrument with a different measurement geometry resulting from a different orbit."

The second major point is a more philosophical point about the use of model data in the retrieval. The authors are not yet using space and time dependent model size distribution, but they allude to this work as the first step towards that plan. To do this the authors must make a convincing case that the information folded into the retrieval from the model size distribution makes the result substantially better in a way that is quantifiable.

Authors: We agree with the reviewer's point. We feel that internal validation comparisons such as Fig. 7 and Fig. 8 provide a quantitative demonstration of the large-scale improvement in LP retrievals from the use of the gamma function size distribution. We anticipate using similar comparisons to validate refinements to this size distribution that incorporate variations as a function of latitude, altitude, or season.

The bias resulting from uncertainty in the aerosol size distribution is a second order effect that can be understood and characterized in a relatively simple way. But now will introducing a complicated spatially/temporally varying model distribution make enough improvement to push this uncertainty to a third order effect, or will it just modify the results so that the second order effect is harder to understand and characterize due to the complex nature of the input assumption? Again, I realize this is not the case for this paper, but anticipation of this as an obvious next step is worrisome. Some of these issues should at least be discussed and approached with caution, especially as with this paper they have chosen to move away from using in-situ measurements to using model output.

Authors: We understand and appreciate the reviewer's concern regarding the potential challenge of implementing a variable size distribution for LP aerosol retrievals. We anticipate moving only incrementally towards this concept, evaluating only one parameter (e.g. latitude dependence) in each step, and carefully examining possible impacts such as discontinuities in retrieved extinction profiles at the boundary between different size distribution functions.

Finally as far as I can tell, this paper is essentially a revision of Chen et al., 2018, which was not submitted for final publication. It seems than that this paper should stand independently and not reference the previous discussion paper, although the editor should weigh in on this.

Authors: A previous version of this paper was submitted to this journal as manuscript amt-2018-4. We prepared responses to the original reviews and returned the revised manuscript to the journal. The associate editor recommended further revision, and stated "… you might think about re-writing the manuscript (possibly including the comparison with SAGE data and the new

OPC data) and submitting anew to AMTD". The current manuscript (amt-2018-221) has been substantially revised in accordance with this editorial guidance. Specifically, we expanded on amt-2018-4 to provide extensive comparisons with SAGE III/ISS data that were not available at the time that manuscript was submitted. The present journal editor for this manuscript also recommended that "you cite the old submission in the new submission". We have added a brief explanation of the relationship between this manuscript and amt-2018-4 at the end of Sect. 1 to clarify the situation for the reader.

Other minor comments:
Page 2, 3rd paragraph: The SCIAMACHY and OSIRIS work must be better referenced and discussed to put this work in context. These groups have done much more work, especially with regard to the size distribution, since the papers that are briefly mentioned here.
Authors: We have added more recent references to OSIRIS and SCIAMACHY results in Sect. 1. We have also added text to discuss those results in Sect. 3.

Several places throughout the paper refer to results from internal validation tests. It seems that some of these are shown and some are not. Is this simply referring to testing of the algorithm performance without validation data from other instruments? If not, this language is frustrating and it leaves the reader wondering what is behind the scenes.
Authors: We have revised the text to clarify that the phrase "internal evaluation" does indeed refer to tests that do not require external data sets.

Why choose the CARMA simulation for no volcanic eruptions? If the goal is really to use representative model data, why not run as realistic a simulation as possible for the OMPS mission time frame (which is definitely influenced by small volcanic eruptions) and then choose the median or average distribution? Why do the authors then choose to analyze a one month period that is perturbed by the Calbuco eruption? What does a 10% change in the gamma distribution parameters mean in terms of particle size? Is this realistic for a moderate volcanic eruption?
Authors: To understand the quality of the present aerosol size distribution and to estimate the uncertainty associated with the retrieved aerosol extinction, we first perform the aerosol retrieval code runs for conditions without a significant volcanic eruption. This provides a baseline situation. To evaluate the performance of the presented aerosol size distribution, aerosol extinction profiles were retrieved from OMPS/LP measurements before and after the Calbuco volcano eruption to see if the volcanic eruption can be captured by the new model. Creating an averaged size distribution from a multi-year model data set (covering the full OMPS LP mission) would give us less confidence that the retrieved extinction profiles are representative of a specific situation, and less understanding of how to interpret changes in retrieved extinction in terms of size distribution changes.
We added the following text to address the effect of gamma distribution parameter changes on particle size: "Examination of the corresponding differential distribution curves (not shown) indicates that increasing α produces an increase in the peak $dN/d\log r$ value, whereas increasing β shifts this peak to larger values of r."

Is there any potential for stray light or calibration effects in the interpretation of the spectral residuals?

Authors: Potential errors due to stray light or absolute calibration bias are addressed in part through the use of altitude-normalized radiances in constructing the ASI measurement vector. Jaross et al. (2014) discuss possible remaining altitude-dependent errors from these sources.

Why does the southern hemisphere look extremely good, where most of the differences in the phase functions seem to be for backscatter angles?
Authors: Absolute ASI values are lower in the Southern Hemisphere because of the smaller phase function values at backscattered angles and the LP measurement geometry, as shown in Fig. 2 and Fig. 4. This difference can be a factor of 5-10 at 20.5 km (see Fig. 16 of Loughman et al. (2018)), which will generally produce smaller absolute residuals. Relative residual values (normalized by zonal mean ASI) are more similar between hemispheres.

The refractive index should be representative of hydrated sulfuric acid and referenced.
Authors: We added text: "for hydrated sulfuric acid (Palmer and Williams, 1975)" with the reference.

The symbol E for aerosol extinction is not standard. Why not k? Then E switches to x in Equation 3. "Extinction ratio" is usually used for the ratio of aerosol to molecular extinction.
Authors: The symbol "*E*" for aerosol extinction has been changed to "β_a", as suggested. You are right, "extinction ratio" is usually used for the ratio of aerosol to molecular extinction. For the sake of convenience in comparing, we use "extinction ratio" in this paper.

What is the point of the discussion of the non-linearity of the ratio of the data versions with reflectivity?
Authors: Figure 5 shows a complex behavior that the ratios of extinction are smaller than the ratio of phase function and the ratios of extinction vary with altitude even though the ratios of phase function do not. The point of the discussion is to explain that the effective reflectivity causes this complex behavior.