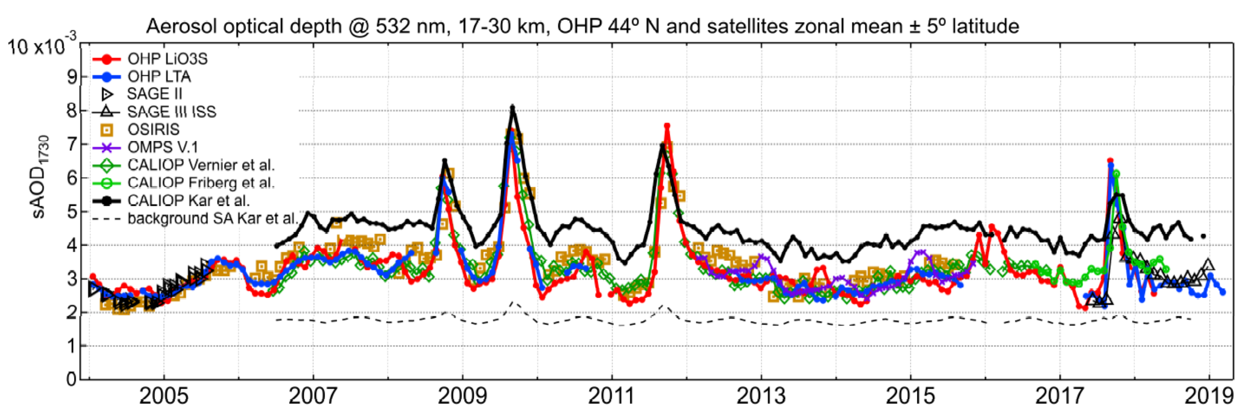
Response to Referee#3

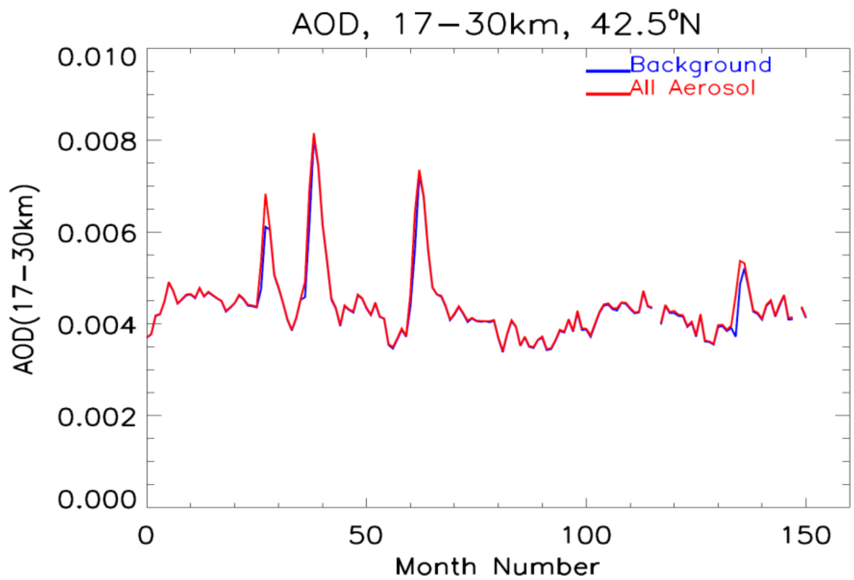
The article by Jayanta Kar and coauthors presents the new stratospheric aerosol (SA) product based on CALIOP measurements, describes the data handling procedure and provides an initial assessment of the data quality through intercomparison with ISS SAGE III measurements of aerosol extinction. With nearly 12 years of continuous operation, CALIOP measurement record represents a valuable source of near-global information on the stratospheric aerosol variability at seasonal to decadal time scales. An obvious advantage of CALIOP measurements is their higher vertical resolution compared to other space-based aerosol sensors (e.g. SAGE, OSIRIS, OMPS etc.) exploiting passive remote sensing techniques. This is why an official release of CALIOP SA data product has been long awaited by stratospheric community and thus the article represents a valuable contribution. The manuscript is well organized and easy to follow, the data retrieval is comprehensively described and the procedure of cloud screening together with the choice of assumption are well discussed. A novel and valuable result is the latitude-height distribution of extinction to backscatter (lidar) ratio inferred by coupling zonally-averaged CALIOP and SAGE III observations.

Thanks very much for a careful reading of the manuscript and for your useful suggestions.

That said, I have a major concern on the data product as such. The very day I found out about the release of L3 CALIOP stratospheric aerosol product, I incorporated the data into the intercomparison of stratospheric AOD series at NH midlatitudes from various satellites and Haute Provence (OHP) lidars in a way we did it in (Khaykin et al., ACP, 2017). The result came quite surprising to me as the new L3 series were remarkably high-biased with respect to all other data sets, including CALIOP SA data product by Vernier et al. I actually thought that I somehow mistreat these data. However, having read this article I realized that this bias is real and amounts to 30-40% at 45 N (Fig. 12), which is consistent with my estimates. A figure below shows time series of AOD of the 17-30 km layer within a 5 deg. latitude belt centered at 44 N as obtained from OHP lidars and various satellite sensors. It includes the CALIOP SA data product by Vernier et al. as well as a more recent one by Friberg et al., (ACP, 2018). While all the data series - independently of the measurement technique (lidar, solar occultation, limb scattering), data handling and the principal measurand (backscatter or extinction) - are in a good agreement, the new L3 CALIOP series stands out high-biased. With that, the background AOD appears low-biased with respect to the well identified clean periods, e.g. 2013-2014.



We would like to point out two things regarding the intercomparisons plot of AOD you provided. First is the reasonably close match between the CALIPSO L3 product "all aerosol mode" with others for the strong volcanic perturbations. Secondly there is some error in the “background” AOD curve you have plotted. The “background” extinctions do not differ that much from the “all aerosol” mode for quiet conditions. In particular there would be essentially no difference between the two modes above 25 km, since aerosol “layers” hardly ever occur above this altitude and there are no thin cirrus above 25 km. The figure below shows the AODs calculated between 17-30 km from the two modes for the grid cell at 42.5°N:



General remarks.

The article reports the observed bias with respect to SAGE III in an honest and comprehensive way, however the discussion of its possible reasons is not satisfying. Basically, it appeals to inaccurate knowledge of the lidar ratio, cloud screening issues and potential errors in the early version of SAGE III extinction product. However, this can in no way explain the discrepancy with other versions of CALIOP SA products by Vernier et al and Friberg et al., nor with OHP lidar operating at the same wavelength. Obviously, there are other reasons for the positive bias beside the error in lidar ratio or that of SAGE III extinction product. These reasons are neither identified nor hypothesized upon, leaving one wonder about the credibility of the L3 SA product as a whole and strongly limiting its scientific value, particularly for radiative forcing studies. Another missing item is the discussion on the quality of the “background aerosol” product. I suggest that the authors attempt to investigate the possible reasons for the latitude and altitude dependent bias and try to eliminate it if possible or at least sketch the envisaged changes/improvements in the future version of CALIOP L3 SA product, other than refinement of lidar ratios. In order to isolate the lidar ratio issue, the validation of the L3 data product could be done on integrated backscatter available from NDACC ground-based lidars at various latitudes.

As described in the text, we rigorously solve the lidar equation from 36 km downwards to 1 km below the tropopause using the same extinction retrieval module as for the standard CALIPSO tropospheric retrieval. In contrast, the Vernier et al. product does not solve the lidar equation, but instead compute the extinction by multiplying the gridded attenuated backscatter by a lidar ratio of 50 sr. As a result, they do not correct the attenuated backscatter for the particulate attenuation at the successive range bins. This may not make significant difference in the retrieved extinction above 20 km under low aerosol loading. However the attenuation can quickly build up below 20 km and under high aerosol loading leading to a significant low bias in the Vernier et al. extinction profiles. Friberg et al. (2018) recognized this and pointed out the large underestimation in AOD that can result from this effect, particularly in the lowermost stratosphere. Friberg et al.(2018) did apply an approximate particulate attenuation correction to their product, nonetheless, even they did not rigorously solve the lidar equation.

In the manuscript, we have pointed out the high bias in our extinction values in mid-high latitude. Some of this can come from cloud clearance issues, either from the inadequacies of our CAD algorithm or from difficulties in complete removal of thin cirrus clouds, the latter mostly affecting a few kilometers above the tropopause in the tropics. We have now added a new figure (Figure 4) and discussed possible cloud contamination issues in both the modes. However, what needs to be appreciated is the retrieval of extremely low values of extinction ($\sim 10^{-4}$ - 10^{-5} km^{-1}) that is being attempted using the CALIOP backscatter measurements with significantly lower SNR than in the troposphere. We believe we have discussed the possible issues impacting the differences with SAGE III to the best of our knowledge at this time. However, we do believe that the more significant contributor to this bias at mid-high latitudes is the lidar ratio, as also becomes quite clear from Figure 13 (now Figure 14) in the manuscript. In any case we shall continue to investigate the reasons for these biases further using stratospheric extinction retrievals from other satellites and NDACC ground based lidars at various locations (as you have suggested). However that is beyond the scope of this manuscript; here we are primarily describing the algorithm involved in retrieving the extinction profiles.

Specific remarks

Figure 4. A strong signal above southern high latitudes is certainly due to PSC and I believe these are type Ib PSC (supercooled ternary solution, STS), which are non depolarizing and thus may be aliased as stratospheric aerosol. The interpretation in p.12 l.5-7 (signature of particles in the process of becoming PSCS) is thus incorrect. I wonder, why the PSCs could not be screened out using temperature threshold for PSC formation, which are relatively well known.

As mentioned in the text, we have used the level 1 CALIOP data for this product while using the level 2 products to filter out the cloud layers. Globally we have used the level 2 clouds as detected by our CAD algorithm and for consistency we decided to use the level 2 CALIPSO PSC mask product to remove the PSCs. Clearly the backscatter is from residual material below the detection limit of the PSC algorithm and likely to be the background aerosol particles which are transitioning into full blown detectable PSC layers. There is only a fine line between background stratospheric aerosol and when it starts growing by HNO_3 uptake and thus may be considered a PSC particle. We have modified the sentence in line 5, page 12 as:

"Since all PSC layers detected by the dedicated PSC detection algorithm were removed, what remains are the signatures of only those particles below the detectability threshold of the PSC mask data product."

Also, using the temperatures from MERRA-2 at these high latitudes as a threshold for detecting PSCs may have its own caveats. For this first version of the product, we have not taken that approach.

Figure 8 and 9, both showing latitude-height sections. Is there a particular reason why the former reports the attenuated scattering ratio, whereas the latter reports extinction coefficient? It would be easier to compare them had they presented the same units.

The attenuated scattering ratio and the retrieved extinction coefficient are both part of the product and we wanted to show examples from both the variables.

Figure 9. What causes the strong signal around the tropopause at midlatitudes? If this is cloud contamination, this should be carefully discussed.

We have added the following:

"Note the high extinction values near 50^{o}N-60^{o}N in the lower stratosphere (~10-15km). These are similar to the summer rise in extinctions at these latitudes as discussed earlier (Figure 7) and are possibly due to biomass burning effects but could also be related to possible cloud clearance issues. As also mentioned above, the high extinctions at high southern latitudes could be related to scattering from particles below the PSC detection threshold as well as to transported volcanic material from Kelud."