Response to Reviewer #2

We thank the reviewer for helpful comments. Our responses to the reviewer's specific comments are listed below. The reviewer's concerns are in bold italicized font and our responses are in regular font. The page numbers and line numbers given in our responses below are in reference to the revised version of the manuscript.

General Comments The paper presents the new version of the GLoSSAC aerosol climatology (version 2), and details all changes brought with respect to the former version 1.1, for wich changes are also briefly described with respect to version 1.0. Prominent changes are the availability of a new version for OSIRIS (version 7.0) with an improved quality, and th release of a standard CALIOP extinction product. Beyonf an improvement of the overall quality, this brings new possibilities to refine the derivation of some GLoSSAC products (e.g. through the use of variable extinction-to-backscatter ratio). Overall, the paper is clear, well written and well structured, in particular the introduction and conclusions. In many places, citation of the first GloSSAC paper (Thomason et al., 2018) is required. I suggest to specify the section to which the citation refers in this paper, in order to ease the reading. A recurrent assuymption is that the SAGE instrument SAGE II and SAGE III/ISS are golden standards, and the key benchmark by which all other data sets have to conform. A very good reason for this is that SAGE instruments are using solar occultation, a technique requiring few assumptions for the data retrieval. SAGE II also has an excellent reputation and was a very long-duration mission. However, doing so ignores the possibility that SAGE II ageing affects the quality of the measurements at the end of the SAGE II, although the use made of SAGE II to calibrate OSIRIS and CALIOP is of critical importance for GLoSSAC. It should be reminded that SAGE II is about 18-years old when OSIRIS and CALIPSO are launched. On the other hand, SAGE III/ISS is recent, and one could miss the broader view on the real quality of this data set. This point of view should also be discussed or at least mentioned, with reference to validation papers giving more insight into the quality of the dataset during the critical period overlapping with the OSIRIS and CALIPSO missions. Overall, the paradigm is that everything is fine tuned to match the two SAGE datasets (and OSIRIS where these datasets are unavailable), but sometimes at all costs, without too much consideration for the consistency or physical significance of the methodology (e.g. different Angstrom exponents used for OSIRIS and CALIOP conversion purposes, ?We do not assume that the derived Angstrom coefficient has any physical meaning it is simply a mean to push OSIRIS toward SAGE II?, L. 22-24, p.9). The fact that instruments (SAGE, OSIRIS, CALIOP) are based on totally different measuring techniques that might have an impact in some altitude or latitude range is hardly considered or discussed, although this might provide an insight into main differences between the data sets. The way CALIOP backscatter coefficient is conformed to GLoSSAC extinction coefficient is also not fully convincing.

Automatically considering a hierarchy of values (?SAGE II is the best instrument?, ?SAGE II/ISS is equal to SAGE II?, ?OSIRIS in the best one after SAGE?) without questioning the physics, the evolving atmospheric state, or any consideration related to aging instruments, has the consequence that SAGE II?s spectre is still hovering on the quantification of the extinction coefficient in the stratosphere as it is about 14 years after SAGE II?s death. This should be questioned or, at the least, discussed.

Specific comments

L. 18-20, p.3: This sentense sound odd. What do the authors mean by ?data (: : :) are made to match or conform with SAGE II?? Do they refer to the transformation of the other source data sets in extinction coefficient profiles at 525 nm/1020 nm at the SAGE II vertical grid? This should be clarified.

Here, for example, we use data from OSIRIS and CALIOP other than SAGE. And, their primary quantity of measurement is different (OSIRIS reports extinction coefficient (after retrieval) at its primary wavelength of 750 nm whereas CALIOP's primary measurement is backscatter coefficient). To match or conform to SAGE II means converting these into SAGE II primary measurements of aerosol extinction at 525 and 1020 nm, either using an Angstrom exponent (for OSIRIS) or using a lidar ratio (for CALIPSO). We included information about primary measurements of OSIRIS and CALIOP as examples in the text (Line 2, page 4).

L. 20-22, p.3: A reference to the first GLoSSAC paper, Thomason et al. (2018) is necessary here to make clear what the authors mean.

Thomason et al. (2018) reference is now included in the text whenever it is needed.

L. 23-27, p.3: Same remark for this discussion: a reference to Thomason et al. (2018) is needed.

Done. (line 18, page 4)

L. 30-32, p.3: Aren?t these differences due to fundamental differences in measurement principles and in such a case, wouldn?t it be a useful way to explore differences and possibly reconcile both techniques?

There are fundamental differences in measurement principles. However, the difference that we see, for example in the conversion of backscatter coefficient to extinction clearly shows some anomalous extinction to backscatter ratios when CALIPSO backscatter coefficient is compared against OSIRIS and/or SAGE extinctions, indicating these differences are due to measurement/conformance deficiencies. These anomalous extinction to backscatter ratios are consistently seen in the entire record of comparison. In addition, we note the differences between OSIRIS and SAGE II/III extinctions and their possible causes were reported earlier (Bourassa et al., 2012; Rieger et al., 2015; Kremser et al., 2016).

L. 33-34, p. 3: What are these changes included in interim version 1.1? Please refer to Section 2.1 where it is described or possibly provide some reference.

The changes made to v 1.1 was mostly to correct how CLAES data is used in the lower stratosphere for the period between July 1991 and April 1993.

We have added a sentence that reads as ”The changes made to interim version 1.1 is described below in Section 2.1.” (Line 26, page 4)

L. 34-35, p.3: If this data set is key, it should at least be cited!

We added a reference in here. The sentence now reads as :

”Within v2.0, a key data set (McCormick et al., 1979) used in the SAGE/SAGE II gap period (1982-1984) has been updated.” (line 27, page 4)

L. 12-13, p.4: This has definitely to be developed and described carefully. Which ground-based lidar product was used, at which location, and which assumptions were used to match them with remote sensing data? Combination of lidar measurements and extinction measurements are not straightforward. Which lidar ratio was used, and how were the data combined?

A detailed description about this is given in section 2.2 (p 8-10) of Thomason et al. (2018) and in section 4.2 of SPARC (2006). The lidar ratios were used from Jäger and Deshler (2002, 2003). We added the reference now and the sentence now reads as:

”CLAES data becomes available in October 1991 but is used in combination with a ground-based lidar product to estimate the aerosol levels from July to September 1991 and is used standalone for a decreasing span of altitude and latitude until the end of its mission in April 1993 (Thomason et al., 2018; SPARC, 2006).”

L. 14-16, p.4: Again, how was this combination (here: CLAES-HALOE with SAGE II) implemented? If only a few points are considered, the possible impact of biases may be high? These aspcts should be carefully discussed.

Please note that this is described in detail in section 2.2 of Thomason et al. (2018). Actual coincidences, by usual standards is pretty low, so we are using some binned comparisons but it still ends up averaging hundreds of data points. The uncertainty in the fits is a part of the error budget for these parameters. In addition, figures 6,7, and 8 of Thomason et al. (2018) show how the combinations were used.

L. 18-19, p.4: ?A few defects missed in v1.0?: which kind of defects and what were the consequences of these defects? Is there any publication or technical report where

these modifications could be found?

While it did not make much difference between the two approaches, we decided that the more conservative use of CLAES and HALOE data preserved the sampling pattern of SAGE II that appears in the entire SAGE II part of the record (1984-2005). Unfortunately, we do not have any publication on this. There is a product quality summary document available from ASDC for v 1.1 and 2.0.

L. 21, p.4: How did this outlier removal occur? Smooth curves may be esthetically more satisfactory, but at risk of leaving out minor events of interest, and possibly of importance for the climate modelling applications the authors want to serve. Also, outlier removal may imply the use of poorly controlled data manipulation and of changes in values very dificult to trace. How did the authors deal with this difficulty? See also comment on L.8, p.7.

We do not think it adversely affects the data by applying an outlier filter as we note that the filtering has only very minimal impact on 452, 525 and 1020 nm extinction coefficients but it impacts 386 nm extinction by reducing noise- a channel with increased noise (Thomason et al., 2008) and should be used cautiously due to the noise in the data which we note in the manuscript and in the previous version (Thomason et al., 2018).

L. 24-26, p.4: This sentence is useful for readers not familiar with the SAGE II dataset. Please provide a citation where this issue is discussed.

We added a couple of references here and the sentence now reads as:

"Users should continue to use caution using the SAGE II 386 nm aerosol extinction coefficient data as a low bias is evident in this data in the lower and upper stratosphere and at all altitudes as aerosol extinction coefficient magnitudes approach background levels (Thomason et al., 2008, 2018). "

L. 26-29, p.4: In Thomason et al. (2018), (at least) two kinds of interpolation mechanisms are used for gap filling. One is a linear intepolation in time (but not in latitude and altitude), and another one is the use of an empirical relationship between the 1020 nm and 525 nm extinction coefficient values defined from a statistical analysis of pairs of (1020 nm, 525 nm) extinction coefficient values retrieved from SAGE II observations (Fig. 8 of this paper). Which one is meant here by the authors?

We used the empirical relationship between 1020 and 525 nm extinction coefficient values.

L. 30, p.4: The concept of equivalent latitude is unclear for a possible ?new reader?. Please provide a reference.

We have now added reference and the sentence now reads as:

"In addition, with the apparent success of filling the high latitudes using the equivalent latitude/latitude mechanism developed based on Manney et al. (2007) for v1.0 (Thomason et al., 2008), we have reduced the role of simple linear interpolation at high latitudes and allow the new equivalent latitude/latitude mechanism fill more of the missing data at high latitudes."

L. 31, p.4: ?The new filling mechanism? is unclear. Do the authors mean: ?the filling mechanism by use of equivalent latitude?? (or ?new more elaborate mechanism? that might be distinguished from ?simple mechanism?). Also, ?the simple intepolation?: do the authors mean ?a linear interpolation? (with respect to time?)?

We have revised the sentence now and it reads as:

"In addition, with the apparent success of filling the high latitudes using the equivalent latitude/latitude mechanism developed based on Manney et al. (2007) for v1.0 (Thomason et al., 2008), we have reduced the role of simple linear interpolation at high latitudes and allow the new equivalent latitude/latitude mechanism fill more of the missing data at high latitudes."

A detailed description as to how the interpolation is done is given in Thomason et al. (2018).

L. 33, p.4: Please be specific to ease the reading: ?the simple (linear?) interpolation process??

We are not sure what the reviewer is referring to ? If it is line 31, we now changed to "linear" interpolation.

L. 1, p.5: It might be useful to specify that these quantities reflect the natural variability and the instrumental error, respectively. In Thomason et al. (2018), an increased value of zonal standard deviation is described when averaging by latitude is used, especially at the boundary of the polar vortex. Is it observed accordingly here that the more extensive use of the equivalent latitude results in a decreased zonal standard deviation?

The use of equivalent latitude reduces the zonal standard deviation in areas where strong zonal gradients occur. The standard deviations we report are always a combination of measurement noise and natural zonal variability.

L. 1-2, p.5: Again, this sentence requires a citation.

We are not sure about the citation of Line 1-2. If it is about CLAES and HALOE data sets, we have now included references to those data sets. The sentence now reads as : "The CLAES (Massie et al., 1996) and HALOE (Thomason, 2012) data sets now include zonal standard deviation and median reported measurement uncertainty following the approach used for SAGE II data. The conversion of CLAES and HALOE data follow the methodology described in Thomason et al. (2018)."

L. 3-5, p.5: This sentence is particularly unclear. Please rephrase, and specify

sections or figures in Thomason et al. (2018) that may ease the understanding of the method.

The conversion of CLAES and HALOE data is described in detail in Thomason et al. (2018). We, think that including the same method here will be a repetition.

Title 2.1 and l. 6-16, p. 5: I suggest to keep the structure and similar titles as in Thomason et al. (2018) by splitting this section is a 2.1 ?The SAGE II period? and ?The pre-SAGE II period?. This should ease the reading, and a possible combined reading of both paper in parallel (and e.g. the comparison of methods used, such as interpolatin methods).

Done. We now use two subsections under Section 2.1.

L. 10, p.5: ?the results?: Do the authors mean ?the values extended along isentropic surfaces between Nov. 1981 and Oct. 1984??

Yes. The sentence now reads as:

"While the revised file was nominally created in same way as the existing data file, the values extended along the isentropic surfaces can be significantly smaller at times than those used in v1.0 particularly in the Southern Hemisphere during the Spring."

L. 5-7, p.6: I guess the two ?potential sources of bias? are basically a single one. Please rephrase. This source of bias is not ?potential?, but real and potentially quite significant. In Thomason et al. (2018) the lidar ration was equal to 50. Why this change, and what are the effects of this change?

The sentence now reads as: "Finally, the conversion from backscatter coefficient to extinction coefficient presents a source of bias; as this process depends on details of an unknown aerosol composition and size distribution (Kar et al., 2019) that is a another potential source of bias."

While the method described in Thomason et al. (2018) is dependent on a median value that is obtained from a relationship between CALIOP backscatter coefficient and OSIRIS extinction, the lidar ratio of 53 is from Kar et al. (2019). We do not think that it makes much of a difference between using 50 and 53 in the lidar ratios. Also, the method used in Thomason et al. (2018) for the CALIOP data is different compared to Kar et al. (2019) method. We also see significant difference in the backscatter coefficient data between these two versions in the lower stratosphere and also at higher latitudes. Therefore, comparing these two data set based on lidar ratio difference may not be a direct comparison.

L. 22-23 p.6: Please provide some explanation or a reference for the PSC identification.

Different methods have been used for detecting PSCs in different data sets. For SAGE measurements, we use a threshold temperature of 200 K, meaning if the temperature is below 200 K between tropopause and 25 km, then those measurements are eliminated as PSCs. For OSIRIS data, additional constraints are used in addition to temperature based PSC detection (Rieger et al., 2019). For CALIPSO, a method employed by Pitts et al. (2009) is used which is also a temperature based approach.

L. 19, p.6: ?found in the lower stratosphere?: at all latitudes?

Occurrences of clouds have often reported in the Upper Troposphere Lower Stratosphere (UTLS) region in the tropics (± 20^0) and midlatitudes (± 40-60^0) in addition to the PSC's in the polar latitudes.

L. 5-7, p.7: See comment on L. 21 p.4.

Again, we do not think it adversely affects the data by applying an outlier filter. And, as mentioned in the paper, we are not using the conservative IQR method which appears to remove some of the enhanced aerosol extinction data. We also ensured that we are not removing any peak data points that occurs due to any volcanic/fire events in the data as it is evident from the time series plots (Figure 10, 12 and 17) that show "before" and "after" conformance of OSIRIS data. These plots clearly show peaks associated with any volcanic/fire events. Again, we are using only a very minimal outlier removal as it can be seen from Figure 1.

L. 8, p.7: Being resigned to accept this fact is harmful because it is known that the accumulation of medium eruptions plays an important role in the correct assessment of the aerosol radiative forcing [Vernier et al., Geophys. Res. Letter, doi:10.1029/2011GL047563, 2011; Bingen et al., Remote Sensing Env., doi: 10.1016/j.rse.2 2017], a key issue CMIP6 is intended to address.

We are not sure what the reviewer is referring to. If the reviewer is concerned about the depiction of volcanic events in the GloSSAC data, it is clear from the time series plot (Figures 10, 12, and 17) that almost all volcanic events can be identified with the peak in the data and the cloud filtering method has a minimal impact on them.

L. 13-15, p.7: ?the extreme outlier was effective at identifying outliers in the aerosol distribution?: The formulation is confusing, please revise. ?outliers in the aerosol distribution?: do the authors mean ?outlying data possibly related to medium volcanic/ pyrocumulonimbus events??)

The sentence now reads as:

"We found that the conservative outlier appeared to remove many enhanced aerosol measurements particularly when stratosphere is perturbed due to volcanic/pyrocumulus events, whereas the extreme outlier was effective at identifying outliers in the density distribution. Therefore we use the extreme outlier to clear cloud-affected observations from the data set."

L. 7, p.8: Using a constant Angstrom exponent implies the assumption that the particle size distribution is constant. This is potentially a rough assumption impacting the accuracy of the values of the extinction coefficient at 525 nm used in GLoSSAC.

This method has been employed in previous studies (e.g. Bourassa et al., 2012; Rieger et al., 2015). And, the conversion factor and the difference between OSIRIS and SAGE II measurements were noted in (Rieger et al., 2015) which led to a scaling of extinction based on OSIRIS to SAGEII extinction ratio (Rieger et al., 2015). As described in the following sections of the paper, we are using a climatological Angstrom exponent and currently we do not have any other way to address this issue. May be a possible transient Angstrom exponent can be implemented in a future version.

L. 11, p.8: What is a ?strong aerosol measurement wavelength??

What we mean by strongest aerosol wavelengths is with least uncertainty and increased accuracy in measurements. For SAGE II and SAGE III/ISS, there are two strongest wavelengths channel in common which are 525 and 1020 nm. We revise the sentence now as:

" Since the SAGE III/ISS instruments operates in a manner similar to SAGE II, the expectation is that there would be minimal bias between these instruments at least at the strongest aerosol measurement wavelengths of 525 and 1020 nm."

L. 18, p.8: What do the authors mean by ?a rather benign October 2004?.

The sentence is now revised as:

"In figures 4a and 4b, it is apparent that for much of the stratosphere the difference between SAGE II and OSIRIS is less than 10% particularly in Figure 4a for a rather benign (less affected by volcanic/fire events) October 2004."

L. 19 and 32, p.8: These estimates are particularly optimistic. Following the color bars, the differences often exceed 50% in both cases.

The sentence now reads as:

"However, it is also clear that OSIRIS extinction is consistently higher than SAGE II in the lower stratosphere with percentage difference exceeding 50% near the tropopause. Another departure is shown in Figure 4b for March 2005 that shows similar features as October 2004. However, in the tropical low and middle stratosphere there is a difference of about 50% in an enhanced aerosol layer associated with the eruption of Manan in January 2005."

L. 1-3, p.9: The assumptions made for the conversion of OSIRIS extinction coefficient from 750 nm to 525 nm seems an obvious cause of deficiency, which is confirmed by the result of the revision of the conversion factor as illustrated in

Figure 5c (and the end of Section 2.4). See comment on L. 7, p.8..

We agree that using a constant Angstrom exponent may not work well during volcanic events and we state that in the manuscript as one of the caveats. Again, this method has been employed in previous studies (e.g. Bourassa et al., 2012; Rieger et al., 2015). And, the conversion factor and the difference between OSIRIS and SAGE II measurements were noted in Rieger et al. (2015) which led to a scaling of extinction based on OSIRIS to SAGEII extinction ratio (Rieger et al., 2015). As described in the following sections of the manuscript, we are using a climatological Angstrom exponent and at this point we do not have any other way to address this issue. May be, a possible transient Angstrom exponent could be implemented in a future version.

L. 17-20, p.9: Did the authors compare the results obtained only with SAGE II, and only with SAGE III? This seems important to assess possible differences, either between the two SAGE instruments, or between both periods.

Since the SAGE III/ISS instruments operates in a manner similar to SAGE II and SAGE III meteor, the expectation is that there would be minimal bias between these instruments at least at 525 and 1020 nm. Some previous studies have compared/validated SAGE II with SAGE III meteor (Thomason et al., 2010; Damadeo et al., 2013). While the differences between SAGE II and SAGE III meteor aerosol extinction coefficient are relatively smaller, previous studies (Thomason et al., 2010; Damadeo et al., 2013) reported a small bias between SAGE II (v 7.0) and SAGE III (v 4.0) meteor that are within $\pm$ 10% for measurement wavelengths at 525 and 1020 nm for the altitudes between 7 and 25 km.

L. 22-24, p.9: This statement is particularly strange! The Angstrom exponent does have a physical meaning, since it reflects the size properties of the aerosol population. Pursuing as sole purpose the replication of one data set at all costs (even one supposed to be good, although its comparison with the real truth is impossible ? this should always be kept in mind!) and getting rid of any concern about the correct quantification of known underlying effects at this aim, looks problematic to me.

We have revised text and replaced Angstrom exponent with "Pseudo Angstrom exponent". Please note that we are conforming the data based on OSIRIS and SAGE II/III extinction comparisons as is described in the manuscript. The pseudo angstrom exponent we use here is merely a scaling factor that conforms OSIRIS data to SAGE II/SAGEIII. We do agree that size changes during volcanic events do matter and we currently do not have a way to address this issue. However, we are working on developing a method as to how various volcanic eruptions affect the particle sizes (Thomason et al., 2020), which might help us understand better about the process. We plan to implement a possible correction on size dependence particularly following a volcanic/fire event in a future version of GloSSAC.

L. 24-25, p.9: ?Angstrom exponent values?.

Yes. It is now corrected.

Caption Figure 4, 5, and 8: The quantity provide should be precisely mentioned, e.g.: ?OSIRIS and SAGE II extinction coefficient at 525 nm?. In caption of Figure 5, ?for at? is not correct and ?for? should be removed. In caption of Figure 8, ?Altitude versus Latitude of percent difference.? is meaningless. Difference in what? The authors should also clearly mention the period covered by this plot.

Done.

L. 27, p.9: I suggest to stick to the naming ?Angstrom exponent?. Please check the whole document.

Done. We now checked for consistency and only "Angstrom exponent" is used throughout the document.

L. 32, p.9: After using an Angstrom exponent of 2.33 to convert OSIRIS extinction coefficient from 750 nm to 525 nm (cf. L. 7, p. 8), another value of the same Angstrom exponent, 1.50, is used to convert the CALIOP extinction coefficient from 532 nm to 525 nm. Why such a difference? This incoherence should be discussed or justified.

The Angstrom exponent of 1.5 is typically used for CALIOP conversion of aerosol extinction (Vernier et al., 2011). We, however, changed that to 2.33 to be consistent with the values used in the manuscript. We have updated Figure 8a, b that use Angstrom exponent to convert CALIOP extinction.This method is just for the comparison purpose (between this version and the conformed version of CALIOP extinction). And, we note that this conversion does not matter as far as the GloSSAC data is concerned as we are not using extinction computed by this method in GloSSAC. We instead use the conformed aerosol extinction that have been computed using scale factor (Figure 9) as described in the manuscript.

L. 7-8, p.10: Smaller eruptions also occurred during the SAGE II mission (1984-2005). Is there any similar observations by SAGE II that might support such tendency? This might help depicting if such effect is real, or is the reflect of some limitation either of the OSIRIS instrument, or of the OSIRIS retrieval.

Yes. There is another paper in ACP which is in review (https://doi.org/10.5194/acp-2020-480) that discusses how various volcanic eruptions impact aerosol sizes that occurred during SAGE II mission and also in the current SAGE III/ISS mission.

L. 25, p.10: ?roughly consistent with values for sulfuric aerosol in the stratosphere?: The extinction-to-backscatter ratio shows much variability in the stratosphere (See

for example Vernier et al., Geophys. Res. Lett., 38, L12807, doi:10.1029/2011GL047563, 2011), and the size characteristics also play a role in the variability of this parameter. Hence, I think that this statement is not very relevant.

We revised the sentence and it reads as:

"This value, 53 sr, is roughly consistent with the extinction-to-backscatter ratio used within CALIOP data processing (50 sr)."

L. 29, p.10: ?As a result?? This sentence is the transition between considerations about version 1.0, and work around version 2. This should be made clear by an adequate introduction. Furthermore, at this stage, it would ease the reading to remind that the CALIOP extinction coefficient product by Kar et al. (2019) is the one used in GLoSSAC, as mentioned in L. 3-4, p.3.

Not sure what the reviewer's comment is. We have already cited Kar et al. (2019) in here where we mention standard stratospheric aerosol product.

L. 31-32, p.10: Why are the authors using now another value of the Angstrom exponent (1.50) for the conversion CALIOP, while a value of 2.33 was used before for OSIRIS extinction conversion? This is quite confusing and increase the level of incoherence between the data sets.

The Angstrom exponent of 1.5 is typically used for CALIOP conversion of aerosol extinction (Vernier et al., 2011). We, however, changed that to 2.33 to be consistent with the values used in the manuscript. We have updated Figure 8a, b that use Angstrom exponent to convert CALIOP extinction.This method is just for the comparison purpose (between this version and the conformed version of CALIOP extinction). And, we note that this conversion does not matter as far as the GloSSAC data is concerned as we are not using extinction computed by this method in GloSSAC. We instead use the conformed aerosol extinction that have been computed using scale factor (Figure 9) as described in the paper.

L. 7-10, p.11: I don?t understand what the authors intend here. In 3, p.10, it is explained that the CALIOP extinction used in GLoSSAC is the CALIOP extinction product (Kar et al., 2019) at 532 nm, converted to 525 nm based on an Angstrom exponent of 1.50. Why do they use now the CALIOP 532 nm backscatter converted using an empirical scaling factor, with some kind of warning that this scaling factor will also reflect ?any kind of biases?? This is extremely confusing.

We compared the standard CALIOP extinction product (after using angstrom exponent of 2.33 to convert from 532 nm to 525 nm) with conformed OSIRIS and SAGE III/ISS data. We however, consistently see an enhanced aerosol extinction in the lower stratosphere and also at higher latitudes (poleward of 40 N/S), indicating that the CALIOP data is biased high. We, therefore thought it is appropriate not to use standard CALIOP extinction, but use CALIOP backscatter and the scale factor (based on OSIRIS 525 nm extinction and CALIOP 532 backscatter) to convert backscatter coefficient to extinction. This scale factor is like a pseudo lidar ratio which has altitude-latitude dependence. We note that Kar et al. (2019) has also pointed out increased extinction levels in the lower stratosphere and higher latitudes in the standard CALIOP data and also computed a lidar ratio (Figure 13 of Kar et al. (2019)) that are retrieved using SAGE III/ISS extinction and CALIOP backscatter measurements which also shows that lidar ratios are variable.

L. 13-27, p.11: I don?t really understand what the authors are doing here. The CALIOP backscatter is the primary quantity measured by CALIOP. What is the interest of rederiving the primary measured quantity from the CALIOP extinction (derived with a simplified assumption of a constant lidar ratio equal to 50), using an empirical scaling factor taking into account all possible problems (?aerosol-related effects and bias between the two data sets?), based on modified (?bias-corrected?) OSIRIS extinctions at another wavelength with some rough approximation about the atmospheric transmission (mentioned as ?clearly not correct? by the authors themselves) , and a simplified formula to account for the scattering ratio and molecular backscatter. And from the conclusion that ?it does not matter a great deal whether we use the standard CALIOP stratospheric backscatter product or the alternative alternative?, the authors choose using this hazardous construction of alternative backscatter product! This is extremely strange and confusing, and if the aim is ? again ? to ?match? at all costs CALIOP with OSIRIS, the methodology used is, at the least, questionable.

We initially thought of using the standard particulate backscatter product. We later realized that the particulate backscatter in the Level 3 data file is retrieved using a lidar ratio 50 Sr. So, if we use the retrieved particulate backscatter for computing scale factor (SF) which is based OSIRIS extinction to CALIPSO backscatter ratio, we are in fact using a SF (which is similar to a lidar ratio) on a product that was already retrieved using a constant lidar ratio of 50. We, therefore used an alternate method that does not use any fixed value for lidar ratio. Please note that a similar method has been used earlier for retrieving backscatter measurements (Vernier et al., 2009). Below is a formulation that we used to derive particulate backscatter as described in the current version of the manuscript. We start with the scattering ratio which is defined as the ratio of total backscatter coefficient to molecular backscatter coefficient. We then assume the transmission of atmosphere ($T_{[\lambda,p]}(z)^2$) is close to 1 as shown in the second step of the formulation. The particulate backscatter ($\beta_{[\lambda,p]}(z)$) is then derived using scattering ratio and molecular backscatter.

$$\text{Scattering Ratio} \equiv SR \equiv \frac{\text{Total Attenuated Backscatter}}{\text{Molecular Attenuated Backscatter}} \tag{1}$$

$$SR = \frac{\left(\beta_{\lambda,m}(z) + \beta_{\lambda,p}(z)\right) T^2_{\lambda,m}(z) T^2_{\lambda,oz}(z) \cancelto{1}{T^2_{\lambda,p}(z)}}{\beta_{\lambda,m}(z) T^2_{\lambda,m}(z) T^2_{\lambda,oz}(z)} \tag{2}$$

$$SR = \frac{\beta_{\lambda,m}(z) + \beta_{\lambda,p}(z)}{\beta_{\lambda,m}(z)} \tag{3}$$

$$SR = 1 + \frac{\beta_{\lambda,p}(z)}{\beta_{\lambda,m}(z)} \tag{4}$$

$$SR \cdot \beta_{\lambda,m}(z) = \beta_{\lambda,p}(z) + \beta_{\lambda,m}(z) \tag{5}$$

$$\beta_{\lambda,p}(z) = (SR \cdot \beta_{\lambda,m}(z)) - \beta_{\lambda,m}(z) \tag{6}$$

, where $\beta_{[\lambda,m]}(z)$, $\beta_{[\lambda,p]}(z)$, $T_{[\lambda,oz]}(z)^2$, and $T_{[\lambda,p]}(z)^2$ are molecular backscatter, particulate backscatter , ozone and particulate transmittance respectively.

We have computed a percent difference between the standard retrieved backscatter coefficient and the backscatter coefficient computed using the alternate method (inferred backscatter). As shown below in Figure 1a, the percent difference computed between retrieved and inferred backscatter for March 2007. At altitudes above 18 km, the percent difference is below ±10%, while the percent difference increases to about ±30% near below 15 km. While there is increased difference below 18 km, it does not really matter much as we scale those differences away in the conformance process by using OSIRIS extinction to CALIOP backscatter, defined as scale factor (SF) in the manuscript. We then computed the ratio of 525 nm OSIRIS extinction to 532 nm CALIOP backscatter coefficient using both the retrieved and the inferred CALIOP backscatter. Figure 1b,c show the ratio of OSIRIS extinction to retrieved and inferred CALIOP backscatter coefficient respectively for March 2007. There are differences between the two methods particularly below 18 km, where they match with the increased percent difference shown in Figure 1a. While the SF computed using retrieved CALIOP backscatter shows values below 30 sr below 18 km (Figure 1b), the SF using inferred CALIOP backscatter shows a higher SF which is around 40 sr (Figure 1c). Generally, below 18 km the retrieved backscatter coefficient is larger than inferred backscatter coefficient. However, these differences are scaled away in the conformance process where we use OSIRIS extinction to CALIOP backscatter ratios (SF) as they are evident from Figure 1. We, therefore believe that our alternate method to infer backscatter coefficient is not a hazardous construction of backscatter.

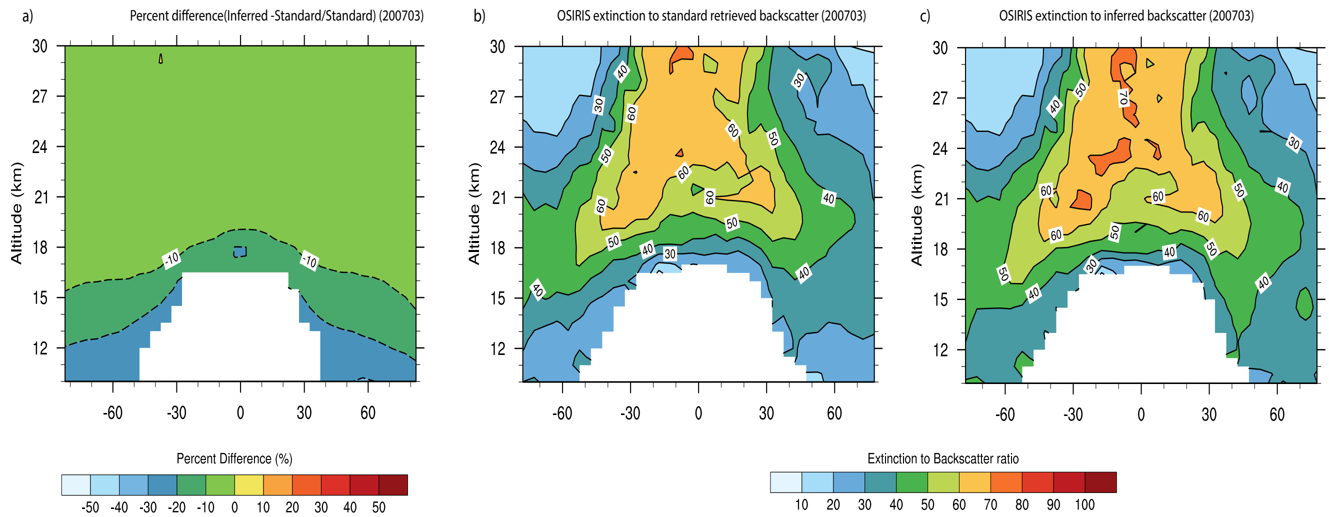

Figure 1: Percent difference and extinction to backscatter ratios for 200703. (a) percent difference between standard retrieved and inferred CALIOP backscatter coefficient computed as (Inferred-Standard/Standard)*100, (b) 525 nm OSIRIS extinction to retrieved 532 nm CALIOP backscatter and (c) 525 nm OSIRIS extinction to inferred 532 nm CALIOP backscatter.

Additionally, as pointed out by reviewer 3, Kar et al. (2019) reported large differences ($> \pm 50\%$) with SAGE III/ISS below 20 km, which could be due to the presence of clouds. These differences are consistent with OSIRIS data as well. This was the reason we decided to use the conformance process based on the extinction to backscatter ratio as shown in Figure 9 of the manuscript, also shown in Figure 1 above.

L. 29-30, p.11: The SF values varying between 25 and 65 might reflect the objective to get rid of the fixed 50-value of the lidar ration used by Kar et al. (2019) to better match local aerosol features. If this indeed is the case, the authors should completely revise this discussion to make it clear, and they should justify why they expect improvement with respect to Kar et al. (2019), see previous comment.

Our response to the previous comment in detail, answers this comments as well.

Figure 5: It is very strange to mix both SAGE II and SAGE II/ISS overlap periods as if these two SAGE sensors were one single data set or mission. SAGE II and SAGE III/ISS are two different instruments measuring different situations in very different conditions. Assimilating the SAGE II and SAGE III/ISS to one single perfect data set looks excessive, and at least, results for both data sets should also be shown (or quantified in some way) to justify that just mixing both is appropriate.

Since the SAGE III/ISS instruments operates in a manner similar to SAGE II and SAGE III meteor, the expectation is that there would be minimal bias between these instruments at least at 525 and 1020 nm. Some previous studies have compared/validated SAGE II with SAGE III meteor (Thomason et al., 2010; Damadeo et al., 2013). While the differences between SAGE II and SAGE III meteor aerosol extinction coefficient are relatively smaller, previous studies (Thomason et al., 2010; Damadeo et al., 2013) reported a small bias between SAGE II (v 7.0) and SAGE III (v 4.0) meteor that are within ± 10% for measurement wavelengths at 525 and 1020 nm for the altitudes between 7 and 25 km. We are not sure if any quantification can be done with SAGE III/ISS and SAGE II as they differ in measurement time period. The best we can do is to compare these measurements with OSIRIS that have overlap measurements with these two instruments, which we have done in the manuscript. We do not think assimilating SAGE III/ISS and SAGE II will create any bias in the monthly climatology of pseudo angstrom exponent, followed by the conformance of the data.

L. 1-4, p.12: The methodology used here is expected to provide more variations of the extinction-to-backscatter ratio than the fixed one assumed by Kar et al. (2019). However, the question is to know it the whole construction with a succession of more or less coarse assumptions used here provide a better estimate of this parameter. See also comment on L. 29-30, p.11

Please see a detailed response to an earlier comment L 13-27, p. 11.

Additionally, as pointed out by reviewer 3, Kar et al. (2019) reported large differences (> ±50%) with SAGE III/ISS below 20 km, which could be due to the presence of clouds in CALIOP data. These differences are consistent with OSIRIS data as well. We also note the pattern of the lidar ratio based on SAGE III/ISS extinction and CALIOP backscatter (Figure 13 of Kar et al. (2019)) is more or less consistent with our SF in Figure 9a.

L. 20, p.13: Is the linear interpolation implemented only in the time dimension? What about the possible use of equivalent latitudes? This should be specified.

When no data is available, grids are filled using linear interpolation in time.

The sentence is now revised and it reads as:

"It should also be noted that, in some cases (particularly for February 2016 and October 2018) when no CALIOP data is available, we linearly interpolate CALIOP data in time between January (September) and March (November) of 2016 (2018) to fill in the missing monthly data following methods used in interpolating SAGE II data."

Figure 15: The choice of dynamic range for the color scale of pannels 15(a), (b), (d), and (e) is particularly poor. Same for Figure 16. Differences mentioned in L. 2-3, p.14 are hardly visible, and the ?substantially smaller enhancement? in 2005 in version 2.0 with respect to version 1.1, is just invisible in both cases to me.

Done.

L. 4-11, p.14: Could several latitudinal dependence and hemispheric dependences possibly be explained by differences in data coverage and/or in intrumental techniques? This possibility has not been discussed.

Some differences may be the result of the different performances of the instruments and that those differences may well be of latitude dependence (e.g. like the scattering angle effect for limb scatter instruments).

L. 22-27, p.14: I think, indeed, that in view of all efforts made to force some data sets to fit in as much as possible some other one, any discussion about trends Is absolutely premature.

Please note that the conformance was done based on studying each data set carefully. Many previous studies have reported the difference between SAGE and OSIRIS measurements (e.g. Bourassa et al., 2012; Rieger et al., 2015) in the first place and for CALIOP, with the standard stratospheric aerosol product we clearly see an enhancement in the lower stratospheric/higher latitude data which is also pointed out by Kar et al. (2019). So, the conformance process was based on studying each data set carefully. While the conformance process applied here merely force the data toward SAGE II/ III-ISS based on intercomparison of individual data sets, we do not believe any signatures of volcanic/fire aerosol is compromised as they are evident in the entire record. Additionally, these signatures are evident from figures 10, 12, 15, 16, 17 and 18.

L. 24, p.2: incorrect sentence: ?whose accuracy? should be removed.

Done.

L. 25, p.2: ?Which this change?? (Or another change?)

The sentence is revised and now reads as:

"With these changes, the retrieved extinction coefficient at 750 nm is in better agreement with observations by SAGE II and SAGE III/ISS than the version used in GloSSAC v1.0 (v5.07)."

L. 3, p.5: missing period (?.?).

Done.

L. 16-17, p.4: odd sentence.

We are not sure what the author is referring to. The sentence seem to be correct.

P. 25, p.6: New sentence starting with ?However, ? ?

Revised the sentence and now reads as:

”While the OSIRIS version 7.0 aerosol data product is similarly cloud screened (Rieger et al., 2019),we, however found some additional clearing was beneficial to the analysis.”

L. 3, p.5: ?its?.

We do not see any ”its” in this line though.

L. 20, p.7: incorrect sentence: ?can transition?.

The sentence is now revised and reads as: ”Cloud identification is complicated by mixed fields of view where observations transition between mostly cloudy extinctions and extinction ratios and those more typical of purely aerosol.”

L. 24, p.7: incorrect reference: should be ?Thomason and Vernier (2013)?.

Done. The sentence now reads as: ”Various techniques to parse these mixed measurements have been developed and GloSSAC makes use of the technique developed by Thomason and Vernier (2013). ”

L. 16, p.9: ?Extinction?.

The sentence is now revised and reads as:

”where, $k_{525[t,m,i,j]}$, and $k_{750[t,m,i,j]}$ are extinctions at 525 nm and 750 nm respectively, $\eta_{[m,i,j]}$ is the pseudo Angstrom exponent while the indices $[t, m, i, j]$represent year, month, latitude, and altitude respectively. $(\frac{\lambda_{525}}{\lambda_{750}})$ represents ratio of wavelengths at 525 and 750 nm. All data are gridded to 5 degree latitude and 0.5 km altitude resolution. ”

Caption Figure 9: The authors should be more explicit: ?(b) Relative standard deviation of the extinction-to-backscatter ratio shown in (a)?. ?deviation of (a) in percent? is unclear.

Done. It now reads as

”(b) Relative standard deviation of (a) is computed at each grid point with respect to the median value in percent”

L. 30, p. 12: ?We use? with capital letter.

Done.

L. 31-32, p.12, L. 18, p.15, and caption Figure 11: ?SAGE II and SAGE III/ISS?.

Revised as ”between OSIRIS and SAGE II/SAGE III-ISS”.

L. 19-21, p.19: Rieger et al. (2019) is published, and the reference should be adapted.
Done.

References

Bourassa, A., Rieger, L., Lloyd, N., and Degenstein, D.: Odin-OSIRIS stratospheric aerosol data product and SAGE III intercomparison, Atmos. Chem. Phys., 12, 605–614, 2012.

Damadeo, R. P., Zawodny, J. M., Thomason, L. W., and Iyer, N.: SAGE version 7.0 algorithm: application to SAGE II, Atmos. Meas. Tech., 6, 3539–3561, https://doi.org/10.5194/amt-6-3539-2013, 2013.

Jäger, H. and Deshler, T.: Lidar backscatter to extinction, mass and area conversions for stratospheric aerosols based on midlatitude balloonborne size distribution measurements, Geophys. Res. Lett., 29, 35–1–35–4, https://doi.org/10.1029/2002GL015609, 1929, 2002.

Jäger, H. and Deshler, T.: Correction to "Lidar backscatter to extinction, mass and area conversions for stratospheric aerosols based on midlatitude balloonborne size distribution measurements", Geophys. Res. Lett., 30, n/a–n/a, https://doi.org/10.1029/2003GL017189, 1382, 2003.

Kar, J., Lee, K.-P., Vaughan, M. A., Tackett, J. L., Trepte, C. R., Winker, D. M., Lucker, P. L., and Getzewich, B. J.: CALIPSO level 3 stratospheric aerosol profile product: version 1.00 algorithm description and initial assessment, Atmospheric Measurement Techniques, 12, 6173–6191, https://doi.org/10.5194/amt-12-6173-2019, URL `https://www.atmos-meas-tech.net/12/6173/2019/`, 2019.

Kremser, S., Thomason, L. W., von Hobe, M., Hermann, M., Deshler, T., Timmreck, C., Toohey, M., Stenke, A., Schwarz, J. P., Weigel, R., Fueglistaler, S., Prata, F. J., Vernier, J.-P., Schlager, H., Barnes, J. E., Antuña-Marrero, J.-C., Fairlie, D., Palm, M., Mahieu, E., Notholt, J., Rex, M., Bingen, C., Vanhellemont, F., Bourassa, A., Plane, J. M. C., Klocke, D., Carn, S. A., Clarisse, L., Trickl, T., Neely, R., James, A. D., Rieger, L., Wilson, J. C., and Meland, B.: Stratospheric aerosol—Observations, processes, and impact on climate, Reviews of Geophysics, 54, 278–335, https://doi.org/10.1002/2015RG000511, 2016.

Manney, G. L., Daffer, W. H., Zawodny, J. M., Bernath, P. F., Hoppel, K. W., Walker, K. A., Knosp, B. W., Boone, C., Remsberg, E. E., Santee, M. L., Harvey, V. L., Pawson, S., Jackson, D. R., Deaver, L., McElroy, C. T., McLinden, C. A., Drummond, J. R., Pumphrey, H. C., Lambert, A., Schwartz, M. J., Froidevaux, L., McLeod, S., Takacs, L. L., Suarez, M. J., Trepte, C. R., Cuddy, D. C., Livesey, N. J., Harwood, R. S., and Waters, J. W.: Solar occultation satellite data and derived meteorological products: Sampling issues and comparisons with Aura

Microwave Limb Sounder, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2007JD008709, 2007.

Massie, S. T., Gille, J. C., Edwards, D. P., Bailey, P. L., Lyjak, L. V., Craig, C. A., Cavanaugh, C. P., Mergenthaler, J. L., Roche, A. E., Kumer, J. B., Lambert, A., Grainger, R. G., Rodgers, C. D., Taylor, F. W., Russell III, J. M., Park, J. H., Deshler, T., Hervig, M. E., Fishbein, E. F., Waters, J. W., and Lahoz, W. A.: Validation studies using multiwavelength Cryogenic Limb Array Etalon Spectrometer (CLAES) observations of stratospheric aerosol, Journal of Geophysical Research: Atmospheres, 101, 9757–9773, https://doi.org/10.1029/95JD03225, 1996.

McCormick, M., Hamill, P., Chu, W., Swissler, T., McMaster, L., and Pepin, T.: Satellite studies of the stratospheric aerosol, Bulletin of the American Meteorological Society, 60, 1038–1046, 1979.

Pitts, M. C., Poole, L. R., and Thomason, L. W.: CALIPSO polar stratospheric cloud observations: second-generation detection algorithm and composition discrimination, Atmospheric Chemistry and Physics, 9, 7577–7589, https://doi.org/10.5194/acp-9-7577-2009, URL `https://www.atmos-chem-phys.net/9/7577/2009/`, 2009.

Rieger, L. A., Bourassa, A. E., and Degenstein, D. A.: Merging the OSIRIS and SAGE II stratospheric aerosol records, J. Geophys. Res. Atmos., 120, 8890–8904, https://doi.org/10.1002/2015JD023133, 2015.

Rieger, L. A., Zawada, D. J., Bourassa, A. E., and Degenstein, D. A.: A Multiwavelength Retrieval Approach for Improved OSIRIS Aerosol Extinction Retrievals, Journal of Geophysical Research: Atmospheres, 124, 7286–7307, https://doi.org/10.1029/2018JD029897, URL `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018JD029897`, 2019.

SPARC: Assessment of Stratospheric Aerosol Properties (ASAP), Tech. rep., SPARC Report,WCRP-124,WMO/TD-No. 1295,SPARC Report No. 4, 348 pp., 2006.

Thomason, L. W.: Toward a combined SAGE II-HALOE aerosol climatology: an evaluation of HALOE version 19 stratospheric aerosol extinction coefficient observations, Atmospheric Chemistry and Physics, 12, 8177–8188, https://doi.org/10.5194/acp-12-8177-2012, URL `https://www.atmos-chem-phys.net/12/8177/2012/`, 2012.

Thomason, L. W. and Vernier, J.-P.: Improved SAGE II cloud/aerosol categorization and observations of the Asian tropopause aerosol layer: 1989-2005, Atmospheric Chemistry and Physics, 13, 4605–4616, https://doi.org/10.5194/acp-13-4605-2013, 2013.

Thomason, L. W., Burton, S. P., Luo, B.-P., and Peter, T.: SAGE II measurements of stratospheric aerosol properties at non-volcanic levels, Atmos. Chem. Phys., 8, 983–995, 2008.

Thomason, L. W., Moore, J. R., Pitts, M. C., Zawodny, J. M., and Chiou, E. W.: An evaluation of the SAGE III version 4 aerosol extinction coefficient and water vapor data products, Atmospheric Chemistry and Physics, 10, 2159–2173, https://doi.org/10.5194/acp-10-2159-2010, URL `https://www.atmos-chem-phys.net/10/2159/2010/`, 2010.

Thomason, L. W., Ernest, N., Millán, L., Rieger, L., Bourassa, A., Vernier, J.-P., Manney, G., Luo, B., Arfeuille, F., and Peter, T.: A global space-based stratospheric aerosol climatology: 1979–2016, Earth System Science Data, 10, 469–492, https://doi.org/10.5194/essd-10-469-2018, 2018.

Thomason, L. W., Kovilakam, M., Schmidt, A., von Savigny, C., Knepp, T., and Rieger, L.: Evidence for the predictability of changes in the stratospheric aerosol size following volcanic eruptions of diverse magnitudes using space-based instruments, Atmospheric Chemistry and Physics Discussions, 2020, 1–28, https://doi.org/10.5194/acp-2020-480, URL https://www.atmos-chem-phys-discuss.net/acp-2020-480/, 2020.

Vernier, J.-P., Pommereau, J.-P., Garnier, A., Pelon, J., Larsen, N., Nielsen, J., Christensen, T., Cairo, F., Thomason, L., Leblanc, T., et al.: Tropical stratospheric aerosol layer from CALIPSO lidar observations, J. Geophys. Res. Atmos, 114, 2009.

Vernier, J.-P., Thomason, L., Pommereau, J.-P., Bourassa, A., Pelon, J., Garnier, A., Hauchecorne, A., Blanot, L., Trepte, C., Degenstein, D., et al.: Major influence of tropical volcanic eruptions on the stratospheric aerosol layer during the last decade, Geophys. Res. Lett., 38, 2011.