

Response to Reviewer #1

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.

This paper describes an evolution of the algorithm used to categorise SAGE III data, and is based very much on earlier papers by the same group. It should be a straightforward report-like paper describing how the new algorithm works and its impact on GloSSAC. Unfortunately the paper is woefully badly written, with some parts unintelligible, a missing description of a key model, inconsistencies between the figures and the text, and errors in English throughout. In particular, the definite article is omitted time after time ? I have noted some of these cases below but gave up after p.5 l.120 as there are too many examples. Did the co-authors actually read this manuscript before it was submitted? It is their responsibility to check the draft and ensure that it is clear, and grammatically correct. The revised manuscript needs to be checked carefully by a native English speaker. It is not the job of reviewers (or the journal) to do the authors' job for them. Most of my comments below are minor, but a few are more substantial and there are so many of them as to require major revision of the manuscript before it can be accepted.

Specific comments

p.1 l.4-5 the stratosphere ??? the tropopause

Done. Thanks.

p.1 l.10 locating the aerosol centroid

Done. Thanks.

p.1 l.12 identifies

Done. Thanks.

p.1. l.20 the solar occultation

Done. Thanks.

p.3 l.62 the Chappuis band

Done.

p.3 l.66 ? 86 and fig.1. I do not understand the method used here. It seems from fig.1 that negative values above 30 km have been converted to positive values and retained? Is that correct? If so this change of sign must be very carefully justified because it biases any averages that are then made. Furthermore for the blue point at the tropopause on fig 1b four, not three, points have been removed from the set shown in fig 1a ? so the diagram is not consistent with the text.

The extinction profile shown in Figure 1 represents absolute extinction coefficient. All extinction values are plotted as absolute values and negative extinction values are color coded using red (blue) filled circles with $> 50\%$ ($< 50\%$) error, whereas orange symbols represent positive extinction with $> 50\%$ error. This information is now added to the figure caption as well as in the text. Yes, thanks for pointing out the discrepancy in the figure. There was a small bug in the code, which is now corrected and revised. The figure is now revised and shows only three points have been removed. We do not remove any negative extinction coefficients above 25 km as they mostly occur due to noise and errors in the removal of ozone and molecular scattering and thus are retained in the data as shown in Figure 1.

p.4 l.112 centroid and an

Done. The description of empirical model is now moved to supplementary section (S1).

What is this empirical model supposed to be modelling? It turns out to be important (see below) yet you don't say what it is. A proper description is needed.

The empirical model can be fitted to the observed data as shown in Section S1 of the supplementary. We rewrote this section and moved this section to supplementary.

p.4 l.118-9. This sentence needs redrafting. For a start the factor is not 2, it is approximately 2, and that point needs to be made. But as written the sentence does not make sense ? the bit after ?because? is not a valid reason. The reason is given in the next sentence ? so some redrafting is needed here, especially as this sentence also (l.120) does not make sense. And in general you need ?the? before the number (e.g. 756) when it is being used as an adjective (e.g. 756 nm channel). This problem recurs throughout the manuscript, as mentioned above.

We have revised this section and now use 525:1020 nm extinction ratio, same as TV13. The method is explained in detail in section 3.2 .

p.4 l.122 delete 'that'

Done. Thanks.

p.5 l128 Delete the sentence 'Here, 3a' as it is redundant

This section is rewritten and moved to Supplementary (section S1).

p.5. l.128 using the same method

This section is rewritten and moved to Supplementary (section S1).

p.5 l.128-130. This section is confusing. For a start, you've already said that figure 3 shows k_0 computed using the same method as TV13. So what is the point of the next sentence, which just repeats the same thing? And as you haven't said anything about the empirical model previously, other than to mention its existence on l. 112, the sentence 'Following TV13 method, we use an offset of 0.4 for the empirical model ratios' is unintelligible. This model needs a proper description (with equations).

The entire section has been rewritten and we now use flowchart to discuss the method involved. The description of empirical model is now moved to supplementary section (S1) with equations.

p.5 l.131. This appears to be saying that points between the red and blue lines are aerosol/cloud mixtures. So what is the point of the next two sentences? Are they not just repeating what you've just said (other than the last part of the second sentence? This part of the paper is really badly written.

The entire section (now section 3.2) has been rewritten to make things clearer with the help of flowchart (section 3.2).

p.5 l.132 Assuming that the blue line in fig 3 is the empirical model, what significance should be attached to the crossing of the blue and red lines, and why are there so many points to the right of both of them above an extinction ratio of 5?

The vertical red line shows the separation of background and perturbed aerosols. So, data points that fall to the right of red line with ratio greater than 1.4 are considered to be perturbed aerosols in this case. We made it clearer in the revised version. Please note that this section has been moved to supplementary (S1, lines 21-25).

p.5 l.136 relatively

This sentence has been removed.

p.5 l.139. What I think you are saying is that the long tail below an extinction ratio of 2, extending to higher extinction, is actually aerosol. You give no proof at all of this statement? have you examined some individual cases to validate that they are not cloud-affected? If not, how do you know they are aerosol?

This sentence has been removed from here to avoid any confusion. The entire section is now rewritten to make things clearer. It is now discussed in section 3.3

p.5 l.143-4 Please point out to the reader where exactly these clusters are on fig 3, as it?s not clear to me. Also, ?there appear? not ?there appears?

This is now discussed in the supplementary section (S1, line 6-7).

p.6 l.163 statistics are

Done. Thanks.

p.7 l.193 points that fall

This sentence has been rewritten.

p.7 l.196 Surely you mean the ratio rather than the extinction coefficient?

Changed it to extinction ratios.

p.7 l.198 there is more than one

Moved to supplementary section S2 (lines 31-35).

p.7 l.211 The first sentence here is redundant ? it just repeats what?s already been said.

This has been removed.

p.7 l.213 These are the same events as you?ve already discussed. Refer to the previous discussion instead of introducing them as completely new events.

This section has been rewritten.

p.7 l.216-8 The horizontal red line also appears in fig.5 but is not described either in the text or the figure caption. Either this description should be moved to the discussion of fig.5 or the horizontal line is redundant in that figure and should be removed. In any case, what does it mean? Are you trying to define a threshold below which the ratio represents a mixture of cloud and aerosol? If so you need to say so explicitly (this issue of course follows on from the lack of explanation of the scattering model on p.4).

This figure and this section have been revised.

p.7 l.219 ?Where? rather than ?Since?, also delete ?then? in l.220

This has been removed and the section is rewritten using a flowchart. .

p.8 l.242. Here you say that the green symbols in fig 6 are categorised as 'enhanced aerosol/tropopause cloud'. In the previous paragraph it is clearly stated (l.221 on) that this category applies to points below the red horizontal line (lower two quadrants). My understanding of this paragraph is that the time series of k_0 after the event was used to define a perturbed period when different rules would apply in this part of the diagram. Yet on l.242 we are told that all the data measured in these periods are categorised as 'enhanced aerosol/tropopause cloud', because all the 20N - 80N points in these panels are green. Surely you have just used green as a geographic label in fig.6?

This section has been rewritten (lines 190-196).

p.8 l.245 What empirical model? I can't see anything that looks like a model in fig.6. Indeed I don't understand what the text from l.245-249 is saying. There is a cloud cluster with a centroid extinction of 10-1 km⁻¹? That's the right-hand axis of the panels in fig 6, where there are no data points.

The empirical model section has been moved to supplementary (section S2).

p.9 l.265 the Coupled

Corrected (line 281).

p.9 l.268 delete 'While other...era'

Removed. Thanks.

p.9 l.271-275. The sentence 'It is known??UTLS region? just repeats earlier text and is redundant ? remove it.

Removed. Thanks.

p.10 l. 296 The sentence 'However, by using the timeframe shown in the monthly time series of k_0 in Figure 4 could alleviate the bias to some extent? is either redundant or needs redrafting ? how does the timeframe 'alleviate the bias'?

Removed. Thanks.

p.10 l.297 ? here you show extinction coefficients at 525 nm, the wavelength where there is a negative bias (l.122), and not the wavelength used for the results presented so far. What is the point of this figure if it doesn't relate to anything shown so far in the paper? The reason, it seems, is that GLOSSAC uses 525 and 1020 nm.

In that case this section must say so, and explain how the SAGE-III data at 525 nm were measured (or calculated). As it stands, you have spent the first part of the paper discussing the analysis of data at 756 and 1544 nm which is not used in the rest of the paper. How does this and subsequent sections relate to the earlier sections?

This has been revised and included explanation on why 525 nm is chosen for GloSSAC (lines 427-434).

p.10 l.300 There is very little difference in the data shown on 8a and 8b, as indeed 8c shows. The most striking difference is that the data in 8b extend to higher latitude. This is completely different to what the text describing this figure says.

While the differences are generally small between the two versions, in the lower stratosphere for the latitudes between 37.5 and 57.5 degrees N, show lower ratios between 0.40 and 0.6 that suggests clear enhancement of extinction in version 5.2. This section has been rewritten (lines 435-440).

p.10 l.308 extended through

Done.

p.12 l.357 delete ?that are?

Deleted and rewritten (lines 356-360).

p.12 l.365. If ?reasonable agreement? is defined as <20% then almost all of fig.10a falls outside this definition. The sentence applies only to 10b.

This has been rewritten to make it clearer (lines 347-350).

p.12 l.367 remove ?in comparison with Figure 10b? (repeats)

Done (line 347).

p.12 l.369. No, the comparison with OSIRIS does not show ?reasonable agreement? in 10a. This whole paragraph is confused, and not consistent with the figure. You must base your argument on the facts you present, not on wishful thinking.

This has been rewritten.

p.12 l.380. Recently, it has been shown??.. eruptions were manifest in the SAGEII?..

It now reads as "Recently, it has been shown that many small to moderate eruptions were manifest during SAGE II and III/ISS data " (line 377)

p.12 l.381 ?show a decrease? (remove ?that?)

Done.

p.12 l.380-388 What do you mean by ?inferring aerosol size .. for the post-SAGEII period .. deficient?? From lines 330-331, the data from 2005-17 came from OSIRIS (750 nm) and CALIOP (532 nm). Do you mean that these wavelengths are too close together to provide a reliable size estimate? Converting them to 525/1050 with some arbitrary Angstrom coefficient is a smokescreen. This paragraph doesn't come to any useful conclusion ? are we to take it that the size estimates for 2005-17 are no good?

We are using measurements to estimate Angstrom exponent climatology as showed in Figure 12. As mentioned in lines 372-376, the angstrom exponent conversion technique does not work effectively when the stratosphere is in the perturbed state. The challenges in converting extinction from one wavelength to another, particularly with a single wavelength measurement is deficient in inferring aerosols following perturbed events. For OSIRIS, we only have extinction at 750 nm and for CALIPSO it is 532 nm backscatter. So, both these measurements provide extinction/backscatter coefficient at a single wavelength that is not adequate to infer size information. So, the challenge we have here is to account for evolving size changes following any perturbed event and fill in the gaps effectively when multiwavelength measurements are not available. We do this by using the conversion technique mentioned in the manuscript and also in Kovilakam et al. (2020). So, we are being transparent here by stating that there is deficiency in this method, particularly for the period between 2005 and 2017, where we only have single wavelength measurement available from OSIRIS/CALIPSO. Therefore, with the data we have, this is the maximum we could do and it is not a smokescreen.

p.13 l.399. The data in fig 11 cover -70-+27⁰ latitude and are unlikely to be much affected by PyroCb at 50-60⁰N.

It is now rewritten as "While we can attribute some of these differences in the northern higher latitudes to PyroCb event associated with Canadian wildfire" (lines 397-398).

p.13 l.406 in the same way

Done.

p.14 l.427-430. First of all, there is another spurious ?that? in this sentence, between v2.2 and apparently. Secondly, the statement made in this sentence is not evident to me, looking at the figures. Please point out more clearly where on the figures you are referring to ? not the difference $\pm 60^0$ it seems, as you go on to discuss that separately. And thirdly, how can you have ?clearly? and ?apparently? in the same sentence? Which is it?

Corrected the sentence and rewritten the paragraph (lines 449-455).

p.14 l.431. This time ?that? has been omitted (changes that occurred).

Done. Thanks.

p.14 l.431 Fig 14 can be removed as this is the only (fleeting) reference to it in the paper.

We kept this figure and removed Figure 13 as we would like to show version changes in the individual measurements.

p.15 l.455 are otherwise

Done.

p.15 l.475 what does ?relatively the same? mean? How is this different to ?the same??

Removed "relatively".

p.16 l.485. This is the first mention of CLAES so where does the ?presumably underestimating aerosol extinction? comment come from? And why is it relevant?

We have now revised this part and it now reads as: "We are currently revisiting this method to identify smoke events for SAGE II. In light of the new insights in the development of this new technique, we will likely revisit cloud detection used for the SAGE II in the production of the GloSSAC data set."

Fig 1 caption l.1 remove ?how?

Done.

Figs 2, 3, 5, 6, 7 increase size of labels on the plots: they should be readable on a laptop screen without zooming right in

Done.

Fig 3 caption ? explain what the red and blue lines are

Done. This section has been moved to supplementary.

Fig 5 caption ? what is the horizontal red line? Explain which panels show data from which period, and point out the difference in height for the two events. Try to

help the reader understand your paper!

Removed those lines from figures and revised the caption. Thanks.

Fig 6 caption ? point out the differences between this and fig.5 e.g. different latitude bands and algorithms. Explain clearly what all the lines and points mean.

This section has been rewritten using a flowchart and figure. Thanks.

Fig 8 caption: log to base 10

Done.

Fig 9 increase size of axis labels and contour labels. In caption, state the period of the climatology

Done. Thanks.

Fig 11 caption says that ?CALIOP data in (c)and (d) are bias corrected using the scale factor (SF) showed in Figure 9a.? But fig 9 shows Angstrom coefficients not scale factors, and according to the text (l.401-406) the bias correction comes from another paper by this group. You must ensure consistency between text and figures.

Corrected. Thanks.

References

Kovilakam, M., Thomason, L. W., Ernest, N., Rieger, L., Bourassa, A., and Millán, L.: The Global Space-based Stratospheric Aerosol Climatology (version 2.0): 1979–2018, Earth System Science Data, 12, 2607–2634, <https://doi.org/10.5194/essd-12-2607-2020>, 2020.