Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-56-RC2, 2020

Interactive comment on "A Global Space-based Stratospheric Aerosol Climatology (Version 2.0): 1979–2018" *by* Mahesh Kovilakam et al.

Anonymous Referee #2

Received and published: 6 July 2020

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.

Printer-friendly version

Discussion paper

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 Angström 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.

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.

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

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?

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.

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

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?

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.

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?

L. 19-20, p.4: If these changes are qualified as “important”, they should definitely deserve an appropriate discussion.

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.

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.

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?

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

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?)?

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

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?

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

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.

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).

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

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?

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

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

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

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.2017.06.002, 2017], a key issue CMIP6 is intended to address.

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”?

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.

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

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

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.

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.

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.

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

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.

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

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.

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.

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.

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.

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 Angström 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.

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.

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.

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.

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

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.

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.

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.

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.

Technical corrections

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

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

L. 3, p.5: missing period (“.”).

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

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

L. 3, p.5: “its”.

L. 20, p.7: incorrect sentence: “can transition”.

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

L. 16, p.9: “Extinction”.

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.

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

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

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

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-56,

2020.