Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-224-RC3, 2020

Open Access Earth System Science Data Discussions

Interactive comment on "Validation of GRASP algorithm product from POLDER/PARASOL data and assessment of multi-angular polarimetry potential for aerosol monitoring" *by* Cheng Chen et al.

Gregory L. Schuster (Referee)

gregory.l.schuster@nasa.gov

Received and published: 17 September 2020

This paper applies several versions of a relatively new retrieval algorithm (GRASP/HP, GRASP/Optimized, GRASP/Models) to an existing satellite measurement archive (POLDER/PARASOL). The authors then compare the results from these algorithms to several legacy retrieval algorithms, including MODIS Dark Target (MODIS/DT), Deep Blue (MODIS/DB), Multi-Angle Implementation of Atmospheric Correction (MAIAC), the operational PARASOL product (PARASOL/Operational), and AERONET. The authors provide a large number of maps and statistics that not only inform the reader about

the performance of the GRASP products, but they also inform the reader about the performance of the legacy aerosol retrievals (e.g., Tables 9 & 10). The paper is clear and well written and I find it suitable for publication.

The paper is also quite long (48 pages of text, 23 tables, and 28 figures) and probably won't be carefully read in its entirety by anyone except the reviewers. One could paraphrase the paper as "here are some new data products, and here is how they compare to similar data products as well as the gold standard (AERONET)." Nobody will learn the machinery behind the retrievals from this paper, but there are other papers for that. One reviewer pointed out that the statistical parameters chosen for this paper are not ideal, but the authors use statistical parameters that are familiar to many readers (correlation coefficient, bias, RMSE, etc.) and common in many satellite/AERONET comparison papers. Unfortunately, the aerosol remote sensing community has not yet adopted a "skill score" for comparisons, as is sometimes used in the modeling community (Taylor, 2001).

It is important to have all of this material in one place, in my opinion, so that readers can quickly assess the relative performance of the different algorithms for the various parameters. However, hyperlinks to tables, figures, citations, and section headings would greatly improve the readability of the paper. A bookmarked Table of Contents in the sidebar would also be helpful. I had to keep two copies of the paper open on my screen – one for the text, and another for reading tables and figures. Otherwise, I would have spent as much time scrolling as reading for this paper! Hyperlinks would allow the reader to go directly to a table, and then return to the text with the "previous view" buttons. Hopefully this is something that can be accommodated in the typesetting process.

Taylor, K.: Summarizing multiple aspects of model performance in a single diagram, J. Geophys. Res., 106, 7183–7192, 2001.

SIGNIFICANT ISSUES

I noticed that the data volume for GRASP/Models is much different than the data volume for GRASP/HP and GRASP/Optimized. For instance, GRASP/Optimized shows 41,268 AOD comparisons with AERONET over land in Table 3, but GRASP/Models only shows 27,551 comparisons. However, GRASP/Model comparisons are greater than GRASP/Optimized over ocean (2064 vs 1495). These large discrepancies appear elsewhere in the paper as well. I found this quite odd, since all three retrievals use the same instrument (PARASOL). I imagine that the cloud screening procedure is identical for all three algorithms, so I suspect that GRASP/Model fails to provide a retrieval much more frequently over land than than the other two GRASP algorithms (and that GRASP/Optimized, GRASP/HP fail more frequently than GRASP/Models over ocean). This should be discussed, since GRASP/Models is lauded for its ability to retrieve AOD(550) (e.g., line 1180). The success rate of a retrieval is important to readers, too.

It is also curious that GRASP/Models did so well for AOD at nearly all wavelengths over both land and ocean (Table 3), but the AEs for GRASP/Models is significantly worse than the other GRASPs. Since AE is derived from AOD, I would have thought the retrieval that produced the best AOD at multiple wavelengths would also produce the best AE. A comment about this would be helpful.

Is there a reason for comparing AODf and AODc to the SDA extinction-based retrievals instead of using the sky scan retrievals? The sky scans are probably more accurate. Many readers (most?) won't know the methodology behind SDA and may incorrectly assume that it is derived from the sky scans. The SDA papers use the AERONET almucantar scans as a performance benchmark, so why not use the same benchmark? You're already using the AERONET sky-scans for SSA and AAOD.

Line 588:
The authors say that they are using AERONET L2 inversions, but which version of AERONET (i.e., Version 2 or Version 3)?

Lines 590-595:
This paragraph will probably confuse some people. Line 590 says that AERONET L2 inversion products require $AOD(440) > 0.4$, but the PARASOL/GRASP filtering includes much lower values, especially over ocean (the authors require PARASOL/GRASP $AOD(443) > 0.3$ over land and $AOD(443) > 0.02$ over ocean). However, since AERONET L2 requires $AOD(440) > 0.4$, many of the low PARASOL/GRASP AODs won't actually appear in the comparisons anyways. . . . Unless the authors using Level 1.5 AERONET inversions at low AOD, like some other authors? If so, what are the Level 1.5 constraints?

MINOR ISSUES
Line 35:
The links do not take me directly to the data products. www.icare.univ-lille.fr takes me to the main page, and the 2nd link on that line tries to take me to www.grasp-35, but that is a dead end.

Line 1125
Do you mean GRASP/Models instead of PARASOL/Models?

It would be interesting to repeat the AOD comparisons using AERONET's "coincident" AOD (that is, using only the AODs that are used during the sky scans). This would be interesting because the cloud screening for the sky-scan products is more comprehensive than for the direct AOD measurements, and it is possible that satellite (and model) AOD performance comparisons wrt AERONET will differ for these two datasets. If the coincident AODs comparisons are different than the "all AODs" comparisons, this could assist our thinking wrt the other sky-scan products. You are probably already set up to do this. I include this as a minor issue, though, because the paper is already too long and this should really be a topic for another day.

Interactive comment on Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2020-224,

2020.

C5