Earth Syst. Sci. Data Discuss.,
https://doi.org/10.5194/essd-2020-224-RC2, 2020

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.

Kirk Knobelspiesse (Referee)

kirk.d.knobelspiesse@nasa.gov

Received and published: 9 September 2020

Review of https://doi.org/10.5194/essd-2020-224 Chen et al., Validation of GRASP algorithm product from POLDER/PARASOL data and assessment of multi-angular polarimetry potential for aerosol monitoring

I commend the authors for what is clearly a comprehensive analysis of POLDER/PARASOL aerosol data products. Some version of this paper should be published, but I have a number of somewhat serious comments and concerns.

Printer-friendly version

Discussion paper

First and foremost, I'm not sure this manuscript is within the scope of ESSD. The "Aims and scope" portion of the ESSD website (https://www.earth-system-science-data.net/about/aims_and_scope.html) says:

"Earth System Science Data (ESSD) is an international, interdisciplinary journal for the publication of articles on original research data (sets), furthering the reuse of high-quality data of benefit to Earth system sciences. The editors encourage submissions on original data or data collections which are of sufficient quality and have the potential to contribute to these aims.

The journal maintains sections for regular-length articles, brief communications (e.g. on additions to data sets) and commentaries, as well as review articles and special issues.

Articles in the data section may pertain to the planning, instrumentation, and execution of experiments or collection of data. Any interpretation of data is outside the scope of regular articles. Articles on methods describe nontrivial statistical and other methods employed (e.g. to filter, normalize, or convert raw data to primary published data) as well as nontrivial instrumentation or operational methods. Any comparison to other methods is beyond the scope of regular articles."

I would think that an ESSD style manuscript on this topic would simply describe the various GRASP algorithms (briefly) and where they are archived. The majority of the paper is indeed comparisons to other data sets. I think it is far more appropriate for Atmospheric Measurement Techniques, for example (to pick another Copernicus Journal). I imagine this is something the editor needs to weigh in upon, but I wouldn't look to ESSD for this type of manuscript.

Secondly, about the length. The manuscript review version is 108 pages long with 23 tables and 28 figures. You've chosen an extensive set of data to compare and contrast the various versions of POLDER data to the various versions of MODIS and then again to AERONET. What I'm missing is a concise set of objectives and how those are met.

Printer-friendly version

Discussion paper

While I'm sure they are in the manuscript, they are lost among the noise. Because of the scale of the task, you need to be creative in finding ways to condense all of this analysis into something that easily and simply supports your work.

Given the above two points, you may consider splitting this manuscript. For example, you could make an ESSD manuscript that describes basics of the dataset and its creation and archive location. It would point to one (or more) manuscripts that contain analysis. One could be on continuity with MODIS (and VIIRS?) and another on the full set of PARASOL products and comparison to AERONET. This is just a suggestion.

Since the core of this manuscript is comparisons of other datasets, the choice of statistical metrics is very important. Table 3, for example, shows the use of nine different metrics, although Pearson's linear correlation coefficient is most commonly employed in the text. Unfortunately, some of these metrics, and especially the linear correlation coefficient, are not suitable for use on non-gaussian distributed data. The correlation coefficient is an expression of association, not agreement (Altman and Bland, 1983 and Bland and Altman, 1986), and is subject to numerical distribution, outliers, and sample range. So, for example, the R values for SSA are lower, but is that because of the lower success of the PARASOL retrieval (what you want to know) or because of the truncated numerical distribution of SSA which makes it non-gaussian distributed? Additionally, what threshold of R can be considered a success? To that end, I think your metric for percentage within the GCOS requirements is a much more appropriate measure, and should instead be emphasized, although you need to take care to account for measurement uncertainty in both POLDER and MODIS or AERONET. Seegers et al. 2018 is a nice overview of these issues in ocean color data products, but equally appropriate here. They also identify regression slope and root mean square error as problematic, while noting that mean bias (which you use) and mean absolute error as appropriate. Bland and Altman suggest something similar, also recommending the pairwise mean bias and the "limits of agreement" which is similar to the mean absolute error. Variable measurement uncertainty can also be incorporated into these techniques (Knobelspiesse et al, 2019) which addresses the salient question “Do measurements agree to within stated uncertainties?” Ultimately, you should revise (and perhaps simplify) the metrics you use to assess your results.

Altman, D. G. and Bland, J. M.: Measurement in medicine: the analysis of method comparison studies, The statistician, 307–317 , 1983.

Bland, J. M. and Altman, D.: Statistical methods for assessing agreement between two methods of clinical measurement, The lancet, 327(8476), 307–310 , 1986.

Knobelspiesse, K., Tan, Q., Bruegge, C., Cairns, B., Chowdhary, J., van Diedenhoven, B., Diner, D., Ferrare, R., van Harten, G., Jovanovic, V., Ottaviani, M., Redemann, J., Seidel, F., and Sinclair, K.: Intercomparison of airborne multi-angle polarimeter observations from the Polarimeter Definition Experiment, Appl. Optics, 58(3), 650–669 , 2019.

Seegers, B. N., Stumpf, R. P., Schaeffer, B. A., Loftin, K. A., and Werdell, P. J.: Performance metrics for the assessment of satellite data products: an ocean color case study, Optics express, 26(6), 7404–7422 , 2018.

Co author Sayer has a publication about the numerical distribution of AOD and its impact on averages of data which is relevant to your matchup methodology and your use of Level 3 products. Were you calculating arithmetic or geometric means? I assume they are arithmetic since you don’t mention otherwise, however this can in some cases cause an artificial bias in the results.

Sayer, A. M. and Knobelspiesse, K. D.: How should we aggregate data? Methods accounting for the numerical distributions, with an assessment of aerosol optical depth, Atmos. Chem. Phys., 19(23), 15023–15048 , https://doi.org/10.5194/acp-19-15023-2019, 2019.

I’m a little surprised that you make no mention of the retrieval algorithms for PARASOL (Hasekamp et al. 2011). In the beginning of section 2.1 you mention “A unique aspect of GRASP is that it can perform radiative transfer (RT) computations fully accounting for multiple interactions of the scattered solar light in the atmosphere, and that it can perform it online without the use of traditional LUTs." You show that GRASP is an extremely power retrieval algorithm, but it is certainly not unique in its use of iterative RT computations, and I find it problematic that you make this claim and do not mention similar algorithms. Without acknowledging that there are other algorithms, even for POLDER, the manuscript sounds more like a sales pitch for GRASP and less an dispassionate piece of peer reviewed literature. In addition to Hasekamp et al, 2011, which is applied to POLDER/PARASOL, many others come to mind including those listed below.

Di Noia, A., Hasekamp, O. P., Wu, L., van Diedenhoven, B., Cairns, B., and Yorks, J. E.: Combined neural network/Phillips–Tikhonov approach to aerosol retrievals over land from the NASA Research Scanning Polarimeter, Atmos. Meas. Tech., 10(11), 4235–4252 , https://doi.org/10.5194/amt-10-4235-2017, 2017.

Fu, G. and Hasekamp, O.: Retrieval of aerosol microphysical and optical properties over land using a multimode approach, Atmos. Meas. Tech., 11(12), 6627–6650 , https://doi.org/10.5194/amt-11-6627-2018, 2018.

Gao, M., Zhai, P.-W., Franz, B., Hu, Y., Knobelspiesse, K., Werdell, P. J., Ibrahim, A., Xu, F., and Cairns, B.: Retrieval of aerosol properties and water-leaving reflectance from multi-angular polarimetric measurements over coastal waters, Optics express, 26(7), 8968–8989 , 2018.

Hasekamp, O. P. and Landgraf, J.: Retrieval of aerosol properties over land surfaces: capabilities of multiple-viewing-angle intensity and polarization measurements, Appl. Optics, 46(16), 3332–3344 , 2007.

Hasekamp, O. P., Litvinov, P., and Butz, A.: Aerosol properties over the ocean from PARASOL multiangle photopolarimetric measurements, J. Geophys. Res, 116(D14), D14204 , 2011.

Stamnes, S., Hostetler, C., Ferrare, R., Burton, S., Liu, X., Hair, J., Hu, Y., Wasilewski, A., Martin, W., van Diedenhoven, B., Chowdhary, J., Cetinic, I., Berg, L. K., Stamnes, K., and Cairns, B.: Simultaneous polarimeter retrievals of microphysical aerosol and ocean color parameters from the MAPP algorithm with comparison to high-spectral-resolution lidar aerosol and ocean products, Appl. Optics, 57(10), 2394–2413 , https://doi.org/10.1364/AO.57.002394, 2018.

Xu, F., Dubovik, O., Zhai, P.-W., Diner, D. J., Kalashnikova, O. V., Seidel, F. C., Litvinov, P., Bovchaliuk, A., Garay, M. J., van Harten, G., and Davis, A. B.: Joint retrieval of aerosol and water-leaving radiance from multi-spectral, multi-angular and polarimetric measurements over ocean, Atmospheric Measurement Techniques Discussions, 2016, 1–90 , https://doi.org/10.5194/amt-2015-394, 2016.

Xu, F., Diner, D. J., Dubovik, O., and Schechner, Y.: A Correlated Multi-Pixel Inversion Approach for Aerosol Remote Sensing, Remote Sensing, 11(7) , https://doi.org/10.3390/rs11070746, 2019.

Regarding the different versions of GRASP, I'm curious what the difference between GRASP optimized and GRASP high precision. What specific parameters are changed?

Section 2.2: PARASOL and Aqua were in the same orbit for a subset of the total PARASOL lifetime, please note this.

Paragraph starting at Line 435: so, if you find both valid ocean and land pixels in the grid box are they averaged for the estimate? If so please state this more directly.

Paragraph at line 605: For comparisons over ocean to AERONET, I presume you're using AERONET-OC (since AERONET-MAN does not include SSA retrievals). I think the distinctions of the AERONET subsets should be made more clear, and also the case that AERONET-OC is restricted to platforms near shore. So the validation against AERONET does not include deep ocean scenes. While this is similar to other studies, it should be noted.

Why is aerosol absorption optical depth included in the assessment when it is derived from two other parameters (AOD and SSA) that are already assessed?

So, the data in Figure 7 represent a subset of what is plotted in Figure 2? Could this be a bit redundant?

Do GRASP/Optimized, GRASP/HP and GRASP/Models use all the same cloud screening, quality flagging and goodness of fit metrics? I'm trying to understand the differences between the number of retrieved cases in, for example, Table 3.

I find Figure 18 (and most of the subsequent maps) to be too small to see properly. Making them larger, however, would be overwhelming. Could this be condensed to fewer maps, say by plotting sGCOS fraction alone?

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