2nd Review of deGraaf, et al., 2019.

This is my second review of this manuscript. I thank the authours for taking on-board many of the comments from my and the other review, and for improving the manuscript. It is easier to follow now, although there are still improvements that could be made to make it clearer (see the specific suggestions below, but there are likely more things that could be done). As well as these, a table showing the resolutions of the different instruments would be useful given the frequent re-gridding. Proof reading for grammar is also needed due to a few grammatical errors that hamper the reading somewhat.

Specific suggestions

p.2, L19 – "When these are accounted for, the remaining differences can be completely explained by the higher cloud optical thickness derived from POLDER compared to the other instruments. Additionally, a neglect of AOT at SWIR wavelengths in the method used for SCIAMACHY and OMI-MODIS accounts about a third of the difference between POLDER and OMI-MODIS DRE, which is mainly evident at high values of the aerosol DRE."

- *These two sentences don't seem to agree with each other. Perhaps the first should be modified to "the remaining differences can be almost completely explained", or similar.*

p.3, L56 – "The DRE from SCIAMACHY was compared to Hadley Centre Global Environmental Model version 2 (HadGEM2) climate model simulations, showing that even this GCM, which simulated a large warming over the south-east Atlantic, still fell short in simulating the UV-absorption by smoke (de Graaf et al., 2014)."

- *It would be good to quote the warming from the model, or the degree of underestimate.*

From the first review :- p.5, L2 – 'CER was derived from collocated MODIS measurements.' Would it not be better for POLDER to retrieve the CER? Is this retrieval not possible? Could MODIS CER be biased by the overlying aerosol, or by inhomogeneous clouds, etc.?

> POLDER does not have measurement in the near infrared. MODIS CER is retrieved primarily from the 2.1_m channel over the ocean. It can potentially be biased by the presence of aerosols above clouds. However, in the region of interest, the aerosols typically observed above clouds (i.e. biomass burning aerosols) are characterised by a large Ångström exponent. Therefore, their contribution to the signal at 2.1_m is expected to be negligible. This is the same argument that is used for the (OMI-)MODIS retrievals, except at 1.2_m. At 2.1_m the effect will be much smaller. Regarding the 3D effect, several filters are used on the POLDER AAC products in order to reject inhomogeneous clouds (Waquet et al., 2013b, GRL).

I was thinking of the direct retrieval of CER by POLDER using the separation of the peaks in the polarized scattering phase function (Breon, et al., 2005). This retrieval is considered to bypass some of the potential biases in visible+near-infrared based retrievals (e.g., MODIS, etc.). There are such POLDER CER datasets available, although they may be limited to regions with more homogeneous clouds. A comparisons in some regions should be possible. It may be beyond the scope of this study, but should be mentioned at least.

p.5, L10 – "Such an estimate is often missing."

Better as "Such estimates are rarely given in the literature." , in order to make it clearer that you are referring to previous studies.

p.5, L11 - “Moreover, the correct characterization of the spectral properties of the overlying aerosols is circumvented by DAA.”

Not sure what you mean here. How is it circumvented? This doesn’t seem to fit with the previous sentences.

Section 3.1 – *You should quote the area average for OMI-MODIS when collocated to the POLDER grid (in the text and in Table 1) since this would give an idea of how much difference the restricted retrievals of POLDER makes and would give support to the statements in the text that this makes a big difference (rather than the larger individual values).*

Fig. 2 & Table 2– *In caption for (b) and in Table 2 (for second set of values) it would be better to say that OMI-MODIS and SCHIMACHY were re-gridded to a 6km grid and that averages are only calculated for the 6km gridpoints for which all 3 instruments produced a measurement.*

p.9, L24 – “Additionally, the sampling was checked by gridding the finer POLDER data to the coarser OMI grid”

Please make it clear in the text how this was done – was an average over each OMI-MODIS grid-box taken, or was just one value sampled for each OMI-MODIS box? However, if the latter then this might make the statistics less robust. Doing an average of lots of POLDER values for each box would be better.

p.9, L28 – “loose” -> “lose”

p.10, L38 – “SCIAMACHY DRE is very similar to OMI-MODIS DRE for all pixels sampled by these instruments (not shown).”

This is a strange blanket statement that seems to suggest that SCIAMACY and OMI agree for all pixels, which is surely not the case. This needs to be more quantitative, or show some results (e.g, a scatter plot).

Fig. 5a – *there are 4 lines for each group of COT values, but only two entries in the legend (reff=8 and 12).*

p.11 L21 – “Irrespective of AOT or COT, an error of 20% in COT can lead to an error in DRE of 50 Wm−2.”

This is written in a slightly odd way that makes it seem like the error is always 50 W/m2 no matter what the COT or AOT values are. In fact it looks like the error value ranges from about 30 to 80 W/m2. I suggest :-

“An error of 20% in COT leads to an error in DRE of between ~30 and 80 Wm−2 for the COT values tested (COT=8 and COT=12), irrespective of the AOT.”

p.11, L23 – "A note for DAA is in order here: because the simulated aerosol-free cloud spectrum is computed from the COT and CER retrieved from the measured aerosol-cloud spectrum, the spectra are the same in the SWIR, cancelling retrieval errors in the COT. A test with OMI-MODIS spectra and POLDER COT regridded to the OMI grid yielded erratic DRE, even though the POLDER COT retrieved in the visible is probably superior to the OMI-MODISCOT retrieved in the SWIR. However, in such a case a simulated cloud spectrum using POLDER COT is often different from the spectrum by MODIS in the SWIR, yielding aerosol effects even without aerosols. For the POLDER DRE calculation this effect is different, because the DRE is computed using the scene twice with the same retrieved COT."

This is difficult to follow and understand – it could do with a rewrite and a more careful explanation of what you mean.

p.11, L41 – "However, it is assumed to have negligible effect in the SWIR from about 1.2 µm, which may be an underestimation (in AOT)."

This sentence doesn't really make sense. The first part of the sentence seems to say that AOT errors have a negligible effect, but does not say on what. The second part suggests there is an underestimation of AOT, but it is not clear.

p.12, L26 – "Note that POLDER COT is retrieved at 0.87 µm, while COT from OMI-MODIS is retrieved at 1.2 µm, which effectively is the MODIS channel."

MODIS has lots of channels, so I'm not sure I understand the last part of the sentence. Also, MODIS can retrieve COT using the 0.86um channel (as for POLDER), although combined with the 2.1um or 3.7um channel. However, most of the signal comes from the 0.86um channel. Or does the possibility of the above-cloud aerosol prevent the use of this?

p.12, L38 – "Here, POLDER COT is regridded, while the MODIS radiances are averaged in the OMI footprint and one COT is retrieved for that OMI pixel."

It's not clear at this point in the manuscript why this is being done. It is interesting to compare the effect of averaging the (1km resolution, or higher? Please state) MODIS radiances to the OMI resolution before doing the COT retrieval. But it needs some introduction at the start of the paragraph about the reasons for doing it. Is this is what is done to get the OMI-MODIS COT values used in the DRE retrieval? Or are the 1km MODIS COT values averaged over each OMI gridbox? It should be mentioned in Section 2.3 how this is done. If the MODIS COT values are averaged, you should also show a scatter plot of the POLDER vs OMI-MODIS COT values where the POLDER values are averaged to the OMI grid and the MODIS COT values are averaged. Also, please state what you do with the POLDER COT values. Do you average them over the OMI grid-boxes?

Table 3 – *it would be good to add that the POLDER data is gridded to the coarser OMI-MODIS grid in the caption. And whether the POLDER data was averaged over the OMI boxes, or interpolated/sampled.*

p.13, L30 - "This approach removes issues related to selecting high positive DRE values by filtering on COT and CF, which introduce large differences in the average DRE. Even if the same filtering is used for the CF and COT for all instruments, different areas will be sampled, because the CF and COT retrieved by the different instruments may be different."

My issue here from the 1st review was that it wasn't clear what filtering by CF and COT was done. Looking through the manuscript I see that this probably refers to the requirement of minimum CF and COT values for a DRE retrieval in Sections 2.1-2.3. However, since this information is in the methods section and not mentioned since, it wasn't clear here what was being referred to. Before talking about CF and COT filtering here it would be good to reiterate that different minimum CF and COT values were used for the different instruments.

Also, I think here you mean "This approach removes issues related to filtering based on COT and CF, which can select high positive DRE values and lead to large differences in the average DRE."

p.13, L38 – "This difference was removed by gridding POLDER to the coarser OMI grid, improving the comparison between OMI-MODIS DRE and POLDER DRE."

This wouldn't completely remove the effect because the effect of averaging the reflectances from the coarser resolution instrument (and then doing the retrieval) would still remain – this causes differences due to the non-linearity between the reflectances and retrieved product. I.e., it's not just a case of averaging the POLDER values to a coarser grid.

p.13, L62 – "The underestimation of the AOT for high values can explain about a third of the difference in DRE between POLDER and OMI-MODIS om 12 August 2006, an overestimation of AOT by POLDER is difficult to establish."

The last part of this sentence (after the comma) doesn't fit with the rest. Also "om 12 August".

p.13, L69 – "Normally, MODIS COT retrievals at 0.8 and 2.1 μm ~~retrievals~~ are close to POLDER COT for fully clouded scenes with liquid water clouds (Zeng et al., 2012) (not considering overlying smoke). However, to avoid biases from smoke absorption, the MODIS channels at 1.2 and 2.1 μm are used to derive COT and CER for OMI-MODIS DRE retrievals, which may further influence the results."

- This information should also be in the methods section to explain why the usual MODIS channels are not used for COT. Also, the second "retrievals" word should be removed.

p.13, L76 – "The difference between COT from OMI-MODIS and POLDER on 12 August 2006 can explain about 80% of the difference in DRE on that day."

- Please state whether this is before or after the footprints have been collocated.

p.13, L85 – “However, a test using this approach using DDA on OMI-MODIS spectra using POLDER COT yielded very erratic results.” -> “However, a test of this approach using DDA on OMI-MODIS spectra but with POLDER COT yielded very erratic results.” (too many instances of “using”).

p. 13, L87 – “clouds spectrum” -> “cloud spectrum”.

References

Bréon, F. M., & Doutriaux-Boucher, M. (2005). A comparison of cloud droplet radii measured from space. IEEE Transactions on Geoscience and Remote Sensing, 43(8), 1796–1805. https://doi.org/10.1109/TGRS.2005.852838