Review of deGraaf, et al., 2019.

This paper examines satellite retrievals of the radiative effect of absorbing aerosols that overlie clouds (here termed the DRE). Retrievals from OMI+MODIS, POLDER and SCIAMACHY are compared. The latter can observe at many different wavelengths, but has low resolution. POLDER can observe the degree of polarization of the reflected light, which allows extra information about the aerosol and cloud to be obtained and minimizes the retrieval assumptions that need to be made. It is found that OMI+MODIS and SCIAMACHY agree reasonably well, but that POLDER produces larger DRE and cloud optical thicknesses (COT). Some of this difference is attributed to sampling issues (mainly arising from the different resolutions of the instruments) and some due to the larger optical depths retrieved by POLDER.

The study should be useful to other researchers since it would be useful to know how large this warming effect is (can it offset a significant amount of aerosol-cloud cooling?) and whether the models get it right. It also seems like the POLDER approach has some promise, particularly if it can be combined with more conventional instruments on e.g., the upcoming METOP-SG 3MI platform. As such I think it should be published after the suggested revisions.

However, the arguments are often a bit muddled and it would be good to see the reasons for the larger POLDER COT values explored a little more, as well as some more investigation into the effect of the low resolution retrievals from the other instruments. The paper talks a lot about "sampling errors" for OMI+MODIS and SCIAMACHY, but this seems to assume that all such errors are just from averaging of the final DRE or COT values, whereas it seems likely that some retrieval errors may be introduced by the averaging effects of the reflectances to low resolution, particularly if the relationship between the reflectances and the retrieved quantities are non-linear. Such effects occur for MODIS retrievals of effective radius and COD for example (Zhang; doi:10.1029/2012JD017655, 2012). It would be good to discuss this and to look into this possibility. It would even be possible to test what effects the averaging of reflectances to lower resolutions might have using synthetic higher resolution reflectances. On a similar note – considering just the "sampling effect" (i.e., just the effect of averaging the retrieved quantities, rather than the reflectances), it should be possible to quantify this effect by degrading the POLDER retrievals to the coarser grids, rather than the other way round, as is currently done.

Section 3.2.3 needs some checking as some of the statements regarding the POLDER optical depth being smaller seemed to contradict the results. The explanations were also not clear.

Specific statements

p.1 L15 – "Aerosol-cloud-radiation interactions currently present the largest uncertainty in our understanding of Earth's climate (Boucher et al., 2013). The effects of atmospheric aerosols are especially uncertain."

The second sentence here reiterates the first and does not really make sense. It should be removed, or else made more clear what it is referring to. Do you mean that the effects of aerosols alone are especially uncertain (compared to cloud-aerosol interactions)? However, I think that it is hard to argue that this case.

p. 1 L18 – "The presence of clouds has a strong influence on the DRE from the light absorbing species in smoke at TOA."

It's hard to understand what you mean here. I think you mean something like this :-

"The DRE (at TOA) due to the light absorbing species in smoke is strongly affected by the presence of clouds." Although maybe it would be good to introduce the idea of light absorption (rather than just scattering) affecting the DRE before this sentence. Or maybe this sentence isn't necessary given what follows?

p. 1 L20 – "Over clouds, on the other hand, scattering by aerosols is negligible" – this is not quite correct I think. The scattering due to aerosols overlying cloud would be quite high – it is the cloud that is doing less scattering in this case because of this. I think you mean that the addition of aerosols above a cloud has negligible extra impact on scattering relative to that which the cloud is already causing.

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.?

p.5 L30 – "MODIS, on-board the Aqua satellite, flies in formation with Aqua in the A-Train, leading Aqua by about 15 minutes"

Should this be MODIS flies in formation with and leads Aura?

p.6 L14 - "Note however, that such an estimate is often missing, while methods other than DAA are moreover highly uncertain due to their dependence on the correct characterization of the spectral properties of the overlying aerosols."

This doesn't quite make sense. Do you mean that often such an error estimate is not made in other studies (does this only apply to those that use DAA)? Please correct if so. The part after should probably be a separate sentence.

p.6 L16 – "Other minor error sources for the DAA method are the uncertainty in input parameters, the influence of the smoke on the estimated cloud fraction, cloud optical thickness and cloud droplet effective radius, an uncertainty in the anisotropy factor (de Graaf et al., 2019), and the uncertainty of estimating the COT and CER at SWIR wavelengths."

In Section 3.2.3 you say that the DRE depends very strongly on the COT. So, wouldn't the COT uncertainty be likely to have a larger contribution to the error than indicated here? Also, this sentence needs to use semi-colons to make it clearer to become :-

"Other minor error sources for the DAA method are the uncertainty in input parameters; the influence of the smoke on the estimated cloud fraction, cloud optical thickness and cloud droplet effective radius; an uncertainty in the anisotropy factor (de Graaf et al., 2019); and the uncertainty of estimating the COT and CER at SWIR wavelengths."

p. 8, L2 – "the instantaneous aerosol DRE over clouds was normalized by dividing by the cosine of the solar zenith angle."

Have you checked whether the DREs scale linearly with the cosine of the angle (presumably a proxy for the incoming SW)? This could be checked with a radiative transfer code. If not then this might introduce some bias. Presumably there is a lower limit for the solar zenith angle allowed?

p. 8, L26 – “The main reason for the much larger area-averaged POLDER DRE on 12 August 2006 is the smaller coverage of the area by POLDER, compared to that by OMI/MODIS, due to a smaller swath. This was illustrated in Figures 1d and 1e.”

This is part of the reason, but it seems that generally POLDER gives considerably higher values for the same regions. You could help demonstrate the magnitude of the differences caused by the different swath area vs that of POLDER values being higher by giving the collocated averages in Table 1. It would be better to change “The main reason” to “One of the reasons”.

p. 8, L33 – “This means that a (dense) plume may be sampled once by a far off-center pixel, or by 15 nadir pixels, all of them receiving the same high values, depending on the satellite track.”

For this to have an effect on the average it would require that the values retrieved from the mean reflectances over the larger pixel did not produce the correct average DRE value – i.e., there is a non-linear relationship between reflectance and the retrieved products, so that the result is dependent on the averaging scale (pixel size). It would be worth nothing this here. Also, the sentence would be clearer without “all of them receiving the same high values,”.

p.10, L20 – “This issue could be resolved if all values were regridded to the coarsest available. However, since this is the SCIAMACHY grid, not many grid cells would remain.”

Although you could do it for the OMI grid vs POLDER, which would be useful?

p.11 L19 – “A comparison of SCIAMACHY, OMI/MODIS and POLDER COT histograms (not shown) revealed a slightly higher COT from SCIAMACHY and OMI/MODIS compared to POLDER (up to 42 for POLDER and 48 for OMI/MODIS (Schulte, 2016)), but the maximum of POLDER is restricted due to LUT limits.”

It’s not clear here where or when these histograms apply to. I see that it is likely to refer to the 19th August case (Table 3), but it needs to be mentioned in the text. Also, “a slightly higher COT from SCIAMACHY” should be changed to “a slightly higher maximum COT from SCIAMACHY” since it otherwise it sounds like you are referring to mean values. However, visually it looks from Figure 5 like POLDER has higher maxima in general? You should also explain the part about the LUT limits in the context of the statement on p.10 L24 (“The POLDER DRE is dependent on the retrieved AOT and COT, which in principle are both unbounded.”).

p. 11 L28 – “Even though the OMI/MODIS data are regridded to a high resolution grid, the values are obviously still more smoothed compared to the COT on the native high resolution POLDER grid. Therefore, even though POLDER COT and POLDER DRE are generally smaller than from OMI/MODIS on average, the extreme values and averages are higher.”

The second sentence seems to contradict the rest of the paper – from the tables and figures POLDER has a generally larger DRE and COT?

Table 3 – it should be made clear in the table caption that the DRE was calculated using the POLDER AOT in both cases.

p.12 L8 – “It shows that the difference between these two quantities disappears completely for these instruments, and the slope is even reversed.”

- It has reduced a lot, but not disappeared completely! Plus, saying that the slope has reversed is a bit unclear. Perhaps better to say it went from <1 to >1.

p. 12 L20 – “The aerosol DRE from POLDER is completely independent. It correlates well with SCIAMACHY and OMI/MODIS DRE for moderate values, but is larger than SCIAMACHY and OMI/MODIS DRE for high values. This is caused by a larger COT retrieved by POLDER, and to a lesser degree by an underestimation of the aerosol DRE using DAA, which by definition assumes a zero AOT at SWIR wavelengths.

The largest contribution to the difference between SCIAMACHY, OMI/MODIS and POLDER DRE are sampling issues.”

- It seems that the last sentence contradicts the ones before where it says that larger COT retrievals by POLDER are the cause. Is it the COT differences or the sampling issues that are most important? Or are they equally important? See also p.13 L8. Also, L9 in the abstract says that sampling issues are the most important – is this actually the case and can you point to the evidence that shows that it such errors are larger than the COT errors?

p.13 L5 – “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.”

- It’s not clear what you are referring to here regarding filtering of COT and CF – is this a method that has been suggested in the literature (please say so and give a reference if so). Or from this paper – again this needs to me made clear.

p. 13 L13 – “Normally, MODIS COT retrievals at 0.8 and 1.2 microns retrievals”

- Doesn’t the usual MODIS retrieval over oceans use the 0.86 and 2.1um bands?

Figures

Fig. 2 – The linewidths of the monthly mean lines need to be quite a bit thicker for the colour and dash style to be visible.

Fig.3 – the legend lines need to be thicker to be able to see the different colours.

Typos

The word “microns” is used a lot, but also the symbol “µm”. I think that the latter is the ACP standard for units.

p. 7, L20 – “when the comparison between the instrument is worst” -> “when the comparison between the instruments are worst”

p. 7, L30 – “Here, we show the effect of ignoring the sampling effect, even of area averages of, in this case, aerosol DRE over clouds.”

– this doesn’t quite make sense. How about something more simple like “Here we show the effect of ignoring the sampling differences between instruments”?

p.10 L23 – “possibly” -> “possible”.

p. 11 L12 – “This way, an AOT at 1.2 µm can be found between 0.15 and 0.35” -> “In this way an AOT at 1.2um of between 0.15 and 0.35 can be found”

p. 11, L22 – “However, the spectral variation in COT is very small. Only for very small cloud droplets the COT at 0.87microns is about 4% smaller than the COT at 1.2microns for cloud droplet effective radii of 4 microns, and this reduces for larger droplets.”

- This would be better as :-

“However, the spectral variation in COT is very small and is only significant for very small droplets. For example, for cloud droplet effective radii of 4 microns the COT at 0.87microns is about 4% smaller than the COT at 1.2microns and this reduces for larger droplets.”

p. 13 L23 – “Comparing AOT over clouds POLDER with MODIS and CALIOP, showed POLDER to be high, but not necessarily overestimated” – insert “from” between “clouds” and “POLDER”.

p.13 L23 “om”.