Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2018-265-AC1, 2018

Atmospheric Measurement Techniques
Open Access
Discussions

Interactive comment on "Joint retrieval of surface reflectance and aerosol properties with continuous variation of the state variables in the solution space: Part 2: Application to geostationary and polar orbiting satellite observations" *by* Marta Luffarelli and Yves Govaerts

Marta Luffarelli and Yves Govaerts

marta.luffarelli@rayference.eu

Received and published: 5 October 2018

1. I think Part II does a rather nice job in introducing the different sources of uncertainties for SEVIRI and PROBA-V (uncertainty assessment - a necessity - is too often neglected). As a general impression though, the writing seems more of a lab "living"

log, i.e. notes accumulated as the work was being carried out. You'll notice how many of my highlighted comments aim at condensing the text as the reader might get lost in details that are often redundantly expressed. Please try and be concise. Every time you start a sentence with "In other words,", ask yourself why you need to re-explain what you just said. As a reader, I had the impression of re-living the struggles to make sense of results, and learning a lot about the things that can go wrong while developing an algorithm, rather than walking confidently away with a message on original and reliable results. This is also reflected in grammatical hurdles. The manuscript should be proofread before submission; see countless instances of 1) "on" instead of "in"; 2) excessive use of "i.e.", "the former" or "the latter", "ones", "it can be seen", "it should be noted", or references to other sections when not really needed; 3) missing plurals; 4) missing articles; 5) "Section" and "Figure" instead of "Sec." and Fig." according to the journal's guidelines; 5) the term "miss-fit".

The paper has been improved thanks to the detailed comments in the annotated PDF which have been implemented in this revised manuscript. Our replies are directly included in that document. Occurrences of "i.e.", "the former", "the latter", ..., has been drastically reduced. The abbreviations are now in agreement with the house standards. The term miss-fit has been replaced by mismatch. The grammatical and styling suggestions have been implemented.

2. Line 116: BRF needs be defined.

TOA BRF are now spelled out separately at line 115 (now 113)

3. Section 2 concludes with "More effort would be needed to demonstrate that the forward RTM is unbiased". This is the kind of sentences disseminated all over Part I that shake confidence in the method. This particular sentence alone gives the impression that the whole method is systematically flawed. Unless the bias is quantified being negligibly small what should the reader take away from this message? As remarked above, the draft goes at quite a length in explaining different sources of uncertainties

smaller than 1%; if this last bias is larger, it would cast quite a different light on the accuracy of the method.

The accuracy of the FASTRE model has been demonstrated only against comparisons with a reference 1D RTM. We underestimated the efforts needed to demonstrate that this model can fit actual satellite data. It would require a detailed characterization of the surface and atmosphere at the overpass time which is currently lacking. It is beyond the resources we had to perform these studies and would probably require a paper of its own. We have therefore decided to remove this paragraph for the time being. We have not found in the literature similar attempts. The assessment of FASTRE uncertainty is now described as in Part I, comparing FASTRE simulations with a reference RTM (RTMOM), and the limitations due to the 2-layer approximation are discussed. Lines 45-46, 500-501, 545-547 and 569-570 have also been removed.

4. Figure 4. This way to depict the subspace of solution is misleading. For example, the way you have things set up now, the magenta triangle does not include the peak of the distribution, with omega>0.98 and gâĹij0.75. Lots of aerosol types are found in this region. How do you deal with this?

The aerosol vertices have been adjusted to include the peak of the distribution. Line 206-207 of the revised paper now read "The selected CISAR vertices defining the solution space cover about the 80% of possible solutions (black triangle)."

5. Fig 6. : merge the two panels into one, since you compare Carpentras with Zinder.

The merged figure looks quite confusing (Fig. 1). Even changing colors and line styles it would still not be very easy to understand. Keeping the separate panels appears a better choice.

6. Line 280-281: this statement is simply not true and has to be reversed. While it is true that the diffraction peak is very sensitive to size, the backscattering contains tons of information (pretty much everything else). We wouldn't be doing space-based

remote sensing otherwise!

Please see answer reviewer 2 regarding line 280.

7. Line 283-284: what was the retrieved optical depth for this day? AND AT WHAT WAVELENGTH? This is an essential piece of information. How would the figure change if the AOT is 0.05 or 0.8? A discussion on the linearity of the AOT Jacobians is due in the text.

The retrieved optical thickness at 0.55 μm is now shown in the plot (Fig. 8) at each observation. A more detailed discussion on the AOT Jacobians magnitude can be found in Luffarelli et al. 2016 (the reference has been added in the text).

8. The "Principle" in Sec. 5.1 needs to be explained better. Please re-elaborate lines 320- 330. I simply couldn't get why the number of could-free pixels should be proportional to the quadratic sum of the mismatch between simulation and observation. Even in the rest of the section, I lost the logical thread. The QI/p tests part is very mysterious, I just did not get it. "QI" is not even defined, and there's no explanation of its range of values. Please review the whole text and try to make it more understandable. Also, "miss-fit" is not a correct terminology; change to "mismatch" or something else. Little to no guidance is offered for the comprehension of Fig. 11. WHEN IS A RETRIEVAL DEEMED SUCCESFULL?

The whole Sect. 5 has been rewritten and it is now organized as follows: 5.1 Review of existing methods 5.2 Overview 5.3 Quality indicator tests 5.4 Quality indicator computation Section 5.3 now includes most of what was described in Sect. 5.4 which is now much shorter and, hopefully, readable. QI was defined in the introduction and according to the house standard does not have to be repeated. The term miss-fit has been replaced by mismatch. The QI/p tests part has been simplified, removing the qi definitions. Lines 346-348 (now 304-306) commenting Figure 11 (now Fig. 10) read now: "Figure 10 shows an example of the evaluation of the retrieved AOT against AERONET data for the mismatch test (3). As the mismatch increases, the correlation decreases, while the RMSE shows opposite behaviour."

9. In both manuscripts, it's never clear if CISAR can be applied to water and land indifferently. This should be made more clear throughout.

Line 66 (previously 69) now reads "These targets span different geometries and land cover types (vegetation, urban, bare areas, water, mixed)". Table 1 includes both water and land cover type (it was already the case). Part I (lines 209-210) states that surface reflectance simulations over water are performed with the Cox-Munk model. However, in that case, surface reflectance is not retrieved but calculated on the basis of the surface wind considered as a model parameter.

10. The approximation of a two-layer atmosphere is not discussed. In fact, it could be a reason for the algorithm failure in many cases.

The two-layer approximation has been inherited from the approach proposed by Pinty et al. (2000) and Govaerts et al. (2010). Section 2.5 now discusses the limitations of the two-layer approximation. Lines 162-170 of the revised paper read: "The forward model uncertainty is lower than 3% in all processed bands, presenting its largest value in the SEVIRI VIS0.8 band, the most affected by water vapour absorption (Table 4). The FASTRE two-layer approximation of the atmosphere does not allow a correct discretisation of the water vapour vertical profile and, thus, a correct characterisation of its interaction with the scattering particles. Moreover, the two-layer approximation assumes that the scattering particles are only present in the lower layer. Given the spectral behaviour of the AOT, this assumption leads to a higher uncertainty at wavelengths shorter than 0.4 μm (Seidel et al., 2010). Despite the limitations associated to the two-layer approximation, FASTRE uncertainty is in the range of 1% - 3% (Table 6), which is smaller or equal to the instrument radiometric noise."

11. Overlap graphs in Figs. 9 and 10 so as to make one figure only

Fig. 9 and 10 are merged in one figure with 2 panels showing, for both satellites, the

entropy related to the AOT and to the RPV parameters respectively.

12. Discussion following Eq. 10: it has to made clear if you're talking about "entropy" or "entropy difference" between pre and post retrieval.

The concept of entropy difference is never mentioned nor used in the paper. What it is used is the entropy, computed as in Eq. 10 after Rodgers (2000). The entropy is mathematically defined as the logarithmic ratio between the prior uncertainty and the posterior uncertainty. It thus measures the uncertainty reduction from the prior to the posterior.

13. Sec. 5.2 is "Theoretical Concept" and comes after Sec. 5.1, i.e., "Principle". I see no point in fragmenting the text this way. Please condense the sections.

Please see anwer to comment #8

14. Line 360: it remains a mystery why a cloud mask is not applied.

"Cloud contamination" has been changed to "cloud mask omission errors". A cloud mask is indeed applied, but omission errors might be present as discussed in Sect. 6.2 (previously 6.1).

15. Line 456-459. This is one of my most important comments. After the manuscript goes to a great length in describing a very elaborate way to aid the retrievals with "tests", the results presented in Fig. 14 are clearly not satisfactory (a look at the correlation coefficients immediately tells that the algorithm is not retrieving appropriate AOTs). Then it is commented that at high AOTs the algorithm might fail (then why all the tests?), but that's not too worrisome since it is better if it performs accurately at low optical depths, which are more typical. I might agree with that, but then I have to ask 1) how do you deal with the fact that the 1:1 correlation is as poor at low optical depths; and 2) why the only AOT used for testing was 0.4 in part 1.

The bias between the CISAR retrieval and the AERONET data is shown in Fig. 2, which shows different performances for SEVIRI and PROBA-V. These differences show that the bias does not only depend on the CISAR algorithm itself, but also on the quality of the processed data. The green histogram shows the AERONET AOT distribution for each bin associated with the CISAR applied SEVIRI AOT product. It can be seen there only few points correspond to AOT>0.8 (less than 5% of the total number of observations), affecting the reliability of the statistics for high values of AOT. The histograms have been added in Fig. 14.

The CISAR AOT product shows overestimation at low AOT and underestimation at large AOT values. The overestimation rapidly decreases as the AOT approaches values of about 0.2. The retrieval is within the GCOS requirement (dashed lines) for 0.2 <AOT< 0.75. For SEVIRI, two factors might explain the overestimation of the retrieved AOT below 0.2. Firstly, most of the selected AERONET stations are located in Europe as can be seen on Fig. 1 of the revised paper, where the SEVIRI pixel resolution is about 5 x 8 km (as opposed to 3x3km at the subsatellite point) which is compared to AERONET point measurement. The probability of residual cloud contamination at this scale might thus explain part of the overestimation (Henderson and Chylek 2005, https://ieeexplore.ieee.org/document/1499014/authors#authors). Secondly (and most likely explanation), it should be reminded here that SEVIRI shortest spectral band is 0.67. At low optical thickness, e.g., 0.1, the sensitivity to aerosol at 0.67 is about 2 times smaller than in the blue spectral regions and 1.5 smaller than in the red. A preliminary analysis revealed that the sensitivity of the TOA BRF to an increase of the AOT from 0.05 to 0.15 is responsible over dark surface to a change comparable with the magnitude of the radiometric uncertainty in the 0.67 μm band. Consequently, the retrieval in these cases essentially relies on the prior information despite the very large associated uncertainty (1.0 for the fine mode, 2.0 for the coarse mode). The prior AOT magnitude is taken from the climatology proposed by Kinne et al., 2013 (doi:10.1002/jame.20035.), which exhibit typical mean values around 0.12 in the SEVIRI disk.

As concern the underestimation at large AOT, very high AOT normally correspond

to local events, especially in Europe (e.g. plume, fire), therefore the AOT obtained by the retrieval from the satellite pixel containing the AERONET station will be lower than the one measured by the AERONET tower (Jiang et al., 2006, https://doi.org/10.1016/j.rse.2006.06.022). The processing of more data would be necessary to increase the number of points with large AOT. Regarding PROBA-V, since the spatial resolution is one km and it has a blue band, overestimation at low AOT should not be present in the data set as is the case for SEVIRI. The retrieval from PROBA-V observations is affected by additional problems: The poorer radiometric performances which decreases the importance of the information derived from the observations The lack of a thermal channel that leads to an unreliable cloud mask

We acknowledge that fact that there is an issue with these results that underperform AOT retrieval with respect to other algorithms retrieving AOT from other instruments. However we are not aware of any algorithm capable of delivering a good AOT product from PROBA-V over land surfaces. Within the PV-LAC project, the CISAR benefit compared to the current operational method has been proven (https://earth.esa.int/web/sppa/activities/instrument-characterization-studies/pv-lac-atmo/about).

Lines 402-409 of the revised paper read now:

“The GCOS requirements are a useful tool to compare different algorithms’ performances. However, it should be considered that both SEVIRI and PROBA-V missions were not originally designed for AOT retrieval. GCOS requirement of 0.03 for low optical thickness translates into a radiometric noise requirement much better than 2 (1)% at 0.4 (0.6) μm, i.e., way below the radiometric performance of the SEVIRI and PROBA-V instruments (Table 3). The duration of the corresponding missions provides however a decisive advantage for the generation of AOT datasets from these instruments. In the following, the GCOS requirements are evaluated in terms of percentage of retrievals satisfying them.”

Lines 410-414 of the revised paper read:

“This is in accordance with the poor radiometric performances of the polar orbiting instrument and with the outcome of the information content analysis performed in Sect. 4. The boxplots in Fig. 14 show an overestimation of the retrieval for low AOT and an underestimation for large AOT.”

Lines 412-440 of the revised paper read now: “Additionally, very high AOT normally correspond to local events, especially in Europe e.g. plume, fire), therefore the AOT obtained by the retrieval from the satellite pixel containing the AERONET station will be lower than the one measured by the AERONET tower (Jiang et al., 2007). The histograms in Fig. 14 show that AOT values larger than 0.8 represent less than 5% of the total number AERONET observations, affecting the reliability of the statistics for high values of AOT. The processing of more data would be necessary to increase the confidence in results for high AOT values. Some examples of CISAR's ability to detect high AOT are shown in the Supplement. The overestimation of low AOT might originate from the different spatial scale between the satel- lite observations and the ground measurements. Most of the selected AERONET stations are located in Europe (Fig. 1), where the SEVIRI pixel resolution is about 5x8 km (as opposed to 3x3 km at the subsatellite point), which is compared to AERONET point measurement. The probability of residual cloud contamination at this scale might thus explain part of the overestimation (Henderson and Chylek (2005), Chand et al. (2012)). Furthermore, the shortest SEVIRI spectral band is centred at 0.67 μm, where the sensitivity to low optical thickness is about 2 times smaller than in the blue spectral region. Consequently, the retrieval in these cases essentially relies on the prior information regardless the very large associated uncertainty. Despite the presence of a blue band and a better spatial resolution (1 km), the retrievals from PROBA-V observations still show overestimation at low AOT, due to the poor radiometric performances which decrease the importance of the information derived from the observations and to the lack of a thermal channel that leads to an unreliable cloud mask.”

16. Line 477-480. I don't understand these comments about Fig. 17. CISAR/SEVIRI is in very good agreement? As CISAR/PROBA-V, it misses the peak of the distribution. Also, CISAR/PROBA-V is said to be underestimating the fraction but so does CISAR/SEVIRI. The significance of the ratio should also be discussed. What are typical ranges?

“Very good agreement” has been removed and replaced by “It can been seen that the distribution related to CISAR retrievals from SEVIRI and PROBA-V observations underestimate the fine mode concentration for τF/τC >3.”. I'm not sure I understand if the reviewer is referring to typical ranges of the fine/coarse mode ratio. In this case, the AERONET data can be taken as reference.

17. The relative magnitude of those"spikes" in Figs. 19 and 20 are worrisome. For the causes you attribute, shouldn't they confirm that your choice of the three vertices is inadequate?

The aerosol vertices have been adjusted as suggested by the reviewer. With the new vertices the magnitude of the spikes strongly decreases. The percentage of points falling on these values is reported in Table 1. The percentages in Table 1 are in agreement with the solution space encompassing about the 80% of the AERONET data.

Table 1 Percentage of SSA and Asymmetry factor retrievals falling on the spikes in Fig. 17 and 18 w0 g 0.6μm 0.8μm 0.6μm 0.8μm SEVIRI 20% 23% 8% 7% PROBA-V 15% 31% 5% 4%

18. Line 487: I take the chance here to expand on previous comments. “Coarse mode characterization” is very far-fetched. The algorithm is not so much retrieving surface and aerosol properties, as much as two aerosol radiative properties and a set of RPV parameters white variability has not been ascertained. Even here, you've already got problems with unreliable retrievals of fine-to-coarse ratio, so much that you focus on the ratio being less or larger than 1. For these reasons, the title sounds a bit pretentious and should be adjusted accordingly. Omega and g are properties but based on the

current title nowadays most readers would expect an extended set of microphysical and optical properties

Indeed, CISAR retrieves the Single Scattering Albedo and the phase function for the aerosols and the RPV parameters for the surface. As described in Part I, each of the surface parameters controls the BRF differently, describing its magnitude, shape, anisotropy and hot spot. Any previously present reference to micro-physical aerosol properties was erroneous and has been removed. The title is therefore consistent with what the algorithm retrieves.

19. Sec. 6.3: how about Carpentras?

The timeseries is shown in Fig. 3 where the MODIS data have been filtered according to their associated quality flag (https://lpdaac.usgs.gov/sites/default/files/public/modis/docs/MODIS_LP_BRDF+Albedo_QA 4.pdf). It can be seen that the MODIS timeseries shows some issues and cannot be considered reliable. This might also partially explain the scattering in the BHR density plots in the supplements. Using MODIS to simulate satellite observations in the attempt of proving the FASTRE capability of correctly characterise the satellite observations we underestimated the effort to collect ground truth RPV parameters.

20. Line 545-547: This is either too obvious or a concept I don't get. You don't describe state variables, you retrieve them, so isn't just that the algorithm fails?

Following comment 3, this sentence has been removed.

21. The manuscript should report complete statistics on the number of analyzed scenes, so that the retrievals can be put in context. I'm not sure this is what happens in Table 11.

The concept of "report complete statistics on the number of analyzed scenes" is not clear. Unfortunately, we cannot answer this comment.

Please also note the supplement to this comment:
https://www.atmos-meas-tech-discuss.net/amt-2018-265/amt-2018-265-AC1-supplement.pdf

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2018-265, 2018.

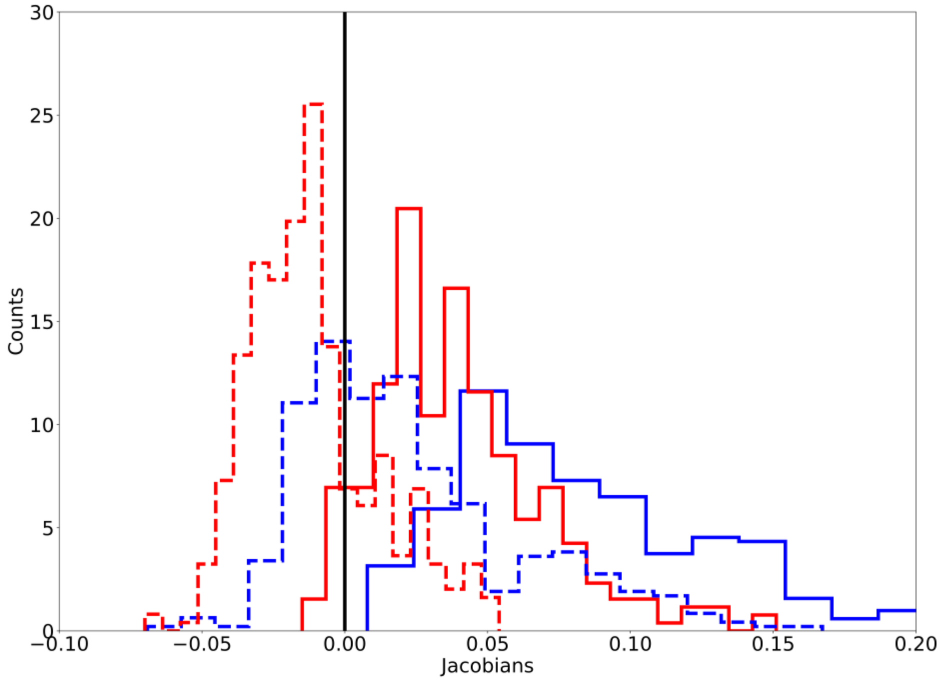

Fig. 1. Merged Figure 6

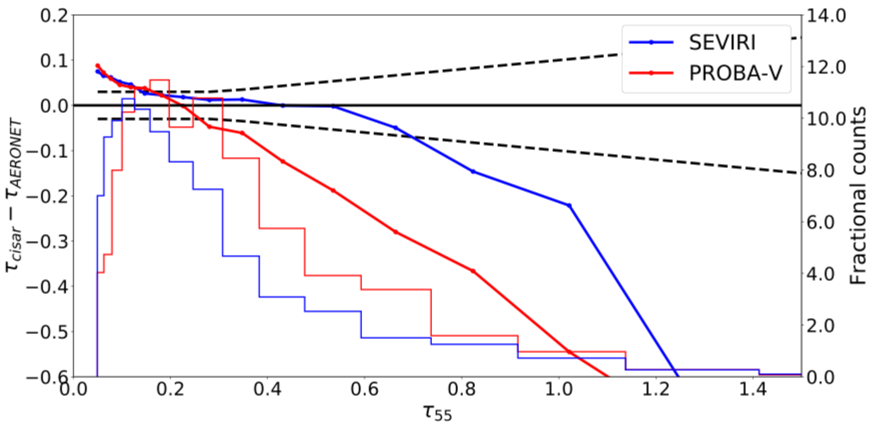

Fig. 2. Bias between CISAR retrieved AOT from SEVIRI (blue) and PROBA-V (red) and AERONET data. The histograms show the distribution of the AERONET data.

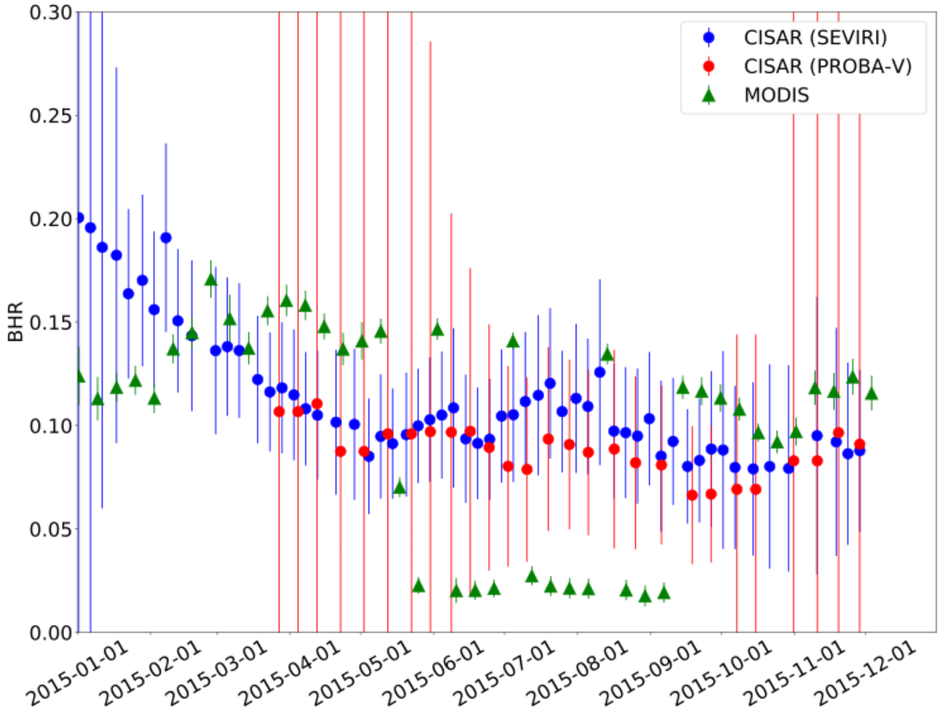

Fig. 3. BHR timeseries at 0.6um over Carpentras