Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2018-265-AC2, 2018

Atmospheric Measurement Techniques 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. My issue is that I see no evidence that this algorithm currently produces acceptable results. Fig. 14 is not good. It's not bad enough to imply your technique is without merit, but if that's the only plot you're going to provide, you will struggle to attract interest

in this algorithm as your correlation, bias and RMSE are worse than most products I've encountered. At the very least, you need to find some circumstances where your retrieval's ability to mix aerosol types produces a better retrieval than a more developed product (e.g. MODIS collection 6.1 or the Swansea University product from Aerosol CCI). Maybe biomass burning emissions from Africa or the industrial regions of China?

Please refer to the reply to Reviewer #1 for Fig. 14.

2. Also, the heritage of the algorithm and the plots in the supplement imply this method is a much better retrieval of the surface than of aerosol. I would warm to the paper more if it was arguing that you made a slightly better aerosol retrieval without harming the surface product rather than the current structure, which implies you were trying to make an aerosol retrieval and skims over the significant limitations in your current results.

Results on the surface BHR are now shown in Section 6.1 to present them prior the AOT. The following lines (506-511) are added to the discussion "The CISAR surface albedo is compared with the MODIS product, showing a correlation higher than 0.74 in all processed bands (to the exception of the NIR PROBA-V band). The better performances of CISAR in retrieving the surface reflectance rather than the AOT are explained by the larger contribution to the TOA BRF at the satellite of the surface. The little variance of the surface reflectance on a short time scale allows a good prior definition based on the previous CISAR retrievals."

3. §4 Though I'm pleased to see a discussion of information content in an atmospheric science paper, yours is rather unusual. You're using the magnitude of the Jacobian to argue which terms are the most important. However, the Jacobian has units and so the magnitude of different terms isn't direct comparable. To illustrate, consider Fig. 5, which you use to argue that 0 is a more dominant driver of changes in TOA radiance than θ. A small change in surface reflectance could be of order $10-3$, which would produce a change of about $10-3$ in y (as the Jacobian is approximately unity). A small

change in viewing angle could be 1 âŮę and, if the Jacobian shown was in units of degrees, that would imply a change of -0.2 in y, which is much larger than that for 0. (The change is still larger if the units are radians.) The value of the Jacobian must be scaled by an appropriately small change to be compared to other values. Optimal estimation already has a mechanism to evaluate this. It's called the averaging kernel and Eq. 2.78 of Rodgers (2000) defines it as, A = KTS -1 K + S -1 a -1 KTS -1 K. You likely already calculate this when determining the entropy (see Eq. 2.80). A row of the averaging kernel summarises the contributions of each state vector element to the retrieval of each other variable while the diagonal elements quantify the reliance on the prior. (Things are slightly complicated by the addition of smoothing, H, terms to your cost function. The difference is subtle; ask Oleg Dubovik about it.) For your retrieval, I would expect the diagonal of A for 0 to be close to one and k to be closer to zero. It would also illustrate the interdependence of the different terms. I don't know if the average reader would find such an analysis easier to understand. Averaging kernels, though very powerful, are confusing. I tend to put them in supplementary material for people that care to find. If you don't switch to averaging kernels, label your plot axes as derivatives rather than Jacobians (e.g. the x-axes on Fig. 6 is dy dτ) so readers have some chance of understanding what's being plotted. More practically, I'd say a superior test to use in §5.2.4 would be the number of degrees of freedom for noise (e.g. n $-$ tr A)

Thanks for your suggestions. The analysis of the information content is now performed on the Jacobians scaled on the variability range of each variable, to account for the different units. Figure 5 now shows the scaled Jacobians, and the axes are labelled accordingly. We prefer not to switch to averaging kernels as they are confusing, as explained by the reviewer.

4. §4 More generally, I'm not sure why this section is so long. It's worthwhile to point out that the retrieval's sensitivity is a function of what is observed, but there must be a more efficient way to show that the retrieval has minimal sensitivity at some times of

day/year.

It is very important to discuss the challenges associated with retrieving information from satellite observations and the difficulty to get a retrieval with constant retrieval with time as the magnitude and sign of the Jacobians can change.

5. Tab. 2 This is a substantial problem. You should be more upfront about the current limitations of your method and outline in more detail what you intend to do about them. There's nothing wrong with incremental progress. This also affects L568.

The FASTRE validation is now presented in a different way. The comparison between simulations and actual observations has been removed. Now FASTRE is only evaluated against a much more accurate radiative transfer model (RTMOM) is SEVIRI and PROBA-V bands, as introduced in Part I. Please refer to the answer to the comment #3 and #19 of Reviewer #1.

6. Fig. 4 I agree with the other reviewer in wondering why you selected vertices that exclude a significant population of observed aerosols.

The aerosol vertices have been adjusted in order to encompass a wider area and include the peak of the distribution.

7. L299 I strongly suspect that there is less information content in the polar data because you ascribed more uncertainty to it (σc and $\sigma\theta$), not because of anything intrinsically advantageous to the geostationary view. This affects your conclusions on L555 and L561. (My opinion is that geostationary data is superior when you need temporal resolution and polar data superior when you need global coverage.)

The reviewer is right, given the larger radiometric uncertainty, PROBA-V data carry less information than SEVIRI ones. Lines 262-264 (previously 297) now read: "The distribution of the surface and AOT entropy related to SEVIRI observations exhibits higher values compared to the one related to PROBA-V observations, given the larger radiometric uncertainty associated to the observations acquired by the polar orbiting

satellite.". Lines 306-310 are eliminated. Lines 549-551 (now 491-493) have been changed to: "Though the PROBA-V instrument has one blue channel which is not present on SEVIRI, the better radiometric performances of the geostationary satellite provide more information for the retrieval of surface reflectance and aerosol properties than the polar orbiting instrument.". Line 561 (currently 501-502) reads now "These differences are explained by the different information content associated to the observations acquired by the two satellites".

8. L321 Do you mean that the magnitude of the cost increases with the number of observations because there are, well, more observations? – L297 of Part 1 addressed something similar by putting a scaling into the cost function; you could do that. – The cost function is (theoretically) a χ 2 distribution with a number of degrees of freedom equal to the number of observations. Using that model, the cost can be converted into a probability that the fit is coincidental and a threshold for retrieval quality defined in terms of that (for example, keeping only retrievals with less than a 5 % probability of being the result of chance). – Regardless, I agree that filtering by cost alone will not identify retrievals with minimal sensitivity.

Thanks for your comment. Indeed, the cost function could be converted in some form of probability and used in the quality indicator computation. However, this test would be performed on the entire accumulation period rather than on a single observation. In CISAR a different QI for each observation is computed to proceed as in test 3.

9. §5 This section is very difficult to follow and needs redrafting with help from someone unfamiliar with the method. Switching between p, q, and QI doesn't help, especially when 1 is a good value for one while 0 is a good value for the other. It would be substantially easier to follow if you provided a decision tree.

The whole section has been rewritten. The term q has been removed. The terms pi now represent the different tests. Section 5.2 now incorporates part of Sect. 5.3, leaving the latter much simpler. Good values are associated with 1, bad values with 0.

It was already the case, but probably it was not very clear. Line 204 reads "Each test pi can assume values 305 between 0 (bad quality) and 1 (good quality)." and line 369 reads "The final QI(ti) ranges from 0 to 1, where 0 designate a poor quality retrieval and 1 indicates a reliable solution.". Please refer also to comment #8 of reviewer #1.

10. §5.2.3 Though I understand the motivation behind this test, I should point out that ym$-$y0 σ0 is normally distributed. As such, 31.8 % of observations would be expected to fail your test by simple chance.

The reviewer is indeed right. The choice of this test and the relative thresholds derives from the choice of being more or less conservative.

11. L360 I agree with the other reviewer that the lack of discussion of a cloud masking is surprising. PROBA-V lacks thermal channels, making it difficult, but you have no problems on SEVIRI.

"Cloud contamination" has been replaced by "cloud mask omission errors". An external cloud mask is applied (Sect. 2.3), however some clouds might not be detected and lead to the overestimation of the AOT.

12. L425 This extra test should have been mentioned back in §5.2.5. More justification of this work around is necessary.

This extra test is now mentioned in Sect. 5.2 and the following sentences (lines 362-364) are added: "Low entropy might be due to a reliable prior information, with a low associated uncertainty. Similarly, the uncertainty reduction would be very large in case of prior information with a very large uncertainty on the state variable."

13. L453 A factor of two is not a 'slight' overestimation and the fact that your retrieval was this bad eight years ago does not forgive it's failure now.

Please see answer to comment #15 of Reviewer #1.

14. L478 That isn't good agreement. A good agreement can be seen between the red

and green lines in Fig. 18(a).

The comment on Fig. 17 (now Fig. 15) reads now (lines 443-445): “It can been seen that the distribution related to CISAR retrievals from SEVIRI and PROBA-V observations seem to underestimate the fine mode concentration for τF/τC > 3.”

15. §6.2 These comparison look good! Why not give us a version of Fig. 14 for SSA and g? Considering they’re what you retrieve, I wouldn’t be surprised if you could estimate them better than you could AOT. Wouldn’t make me think the product was any better as most users want AOT, but they aren’t many global SSA and if you could provide one, even if it’s very uncertain, that would be something worth writing about.

The correlation is strongly dependents on the amount of variability in the datasets (Goodwin et al., 2006 https://pdfs.semanticscholar.org/b6cf/001cbab0375a96c370585462dd3c163669af.pdf). As the variability range of the aerosol single scattering properties is very limited (about 10%), we don’t find it useful to show the same kind of plot as Fig. 14.

16. L116 There are many potential calibration methods for SEVIRI. If you’re using IMPF or GSICS, could that be mentioned explicitly? If you’re using something in-house, a citation would be appreciated.

The calibration method used within this study is the one proposed by Govaerts et al. (2013), as specified in Sect. 2.2. GSICS provides routinely correction factors from IMPF values only for the thermal channels, not for the solar ones.

17. L145 Why make this approximation? Is the calculation of the other terms computationally expensive?

Yes, the calculation of the other term is computationally expensive as it implies the calculation of additional partial derivatives.

18. Eq.6 This seems a strange choice. Why not the standard deviation or interquartile range or a constant value based on climatology?

The range in which they vary is less conservative than the standard deviation or the interquartile. We don’t want to impose a too strong prior. We are not using any climatology for the surface and we do not intend to.

19. §2.4(1) What’s the value of Nmin? Why increase the uncertainty by 5 % per day rather than any other amount?

Nmin has been added is Table 8. The prior uncertainty is increased by the arbitrary value of 5% per day in order not to rely on a solution retrieved too far away in time from the current inversion. This value has been empirically adjusted.

20. §2.5 I’d actually prefer to see a thorough sensitivity study of bias as a function of the various parameters rather than the simple 1 - 3 % uncertainty you’ve added, but that can be in a third paper.

Thanks for the suggestion, we might consider this for a future study. For the time being it has been implemented in this way for efficiency purposes.

21. P12L2 In my experience, the first guess is set to reduce the number of iterations needed to reach a solution. Avoiding local minima involves checking the shape of state space around the final solution or annealing (i.e. running multiple retrievals on the same data).

Indeed, one alternative solution to avoid local minima is to run multiple retrievals on the same data. However, this is also computationally expensive. The idea behind alternated first guess is to simulate the annealing running the inversion starting from different first guess for each observation, rather than repeating N times the same inversion.

22. Eq.8 So you’re using a different first guess for even and odd numbered time steps? That’s peculiar and, on its own, I don’t see how it avoids local minima.

The first guess of the RPV parameters is not defined to minimize the probability of falling into a local minima as for the AOT. As empirical results showed that even a

slight overestimation or underestimation of the surface can lead to larger bias in the AOT retrieval, the different first guess is set to not get stuck in a over/under-estimation situation.

23. §4 The third paragraph covers four pages. Perhaps split it up.

This has been done.

24. P17L1 As the sensitivity drops through the day, I would expect the uncertainty to increase.

The AOT retrieval uncertainty depends not only on the Jacobians, but also on the temporal and spectral smoothness constraints and the quality of the surface. However Fig. 1 shows that for high Jacobians the retrieval uncertainty decreases.

25. L351What is the maximum number of iterations?

It is 20, this has been added to Table 8.

26. L352 Could you clearly state that p0 = 1 in all other circumstances. I wasn't certain of that till I got to Eq. 15.

This is now clearly stated. Line 352 (now 311) reads "When the maximum number of iteration is reached p0 is equal to 0, otherwise p0 = 1."

27. §5.2.2 Did you ever explore using the a priori cost for this test (i.e. the difference between the retrieval and the prior)?

Thanks for the suggestion. We might explore this option in the future.

28. Eq. 11 Aren't the y terms vectors? If so, wouldn't this require some sort of sum?

The formula is now more explicitly written. As I'm actually considering the maximum mismatch among the different bands within this test (this was not explicitly written earlier), those terms are vector components.

29. L371 Didn't you have to calculate the full Jacobian to perform your inversion? I see

your point, but this is a lot of explanation for why you don't use something you should already have

I do have the full Jacobians, but considering them would require even more tests and manipulations. Anyways, as suggested from Reviewer #1, this part has been shortened

30. L379 I assume that if I ask for a justification of this statement, I will be told to go look at your papers from 2010 so I will make this sarcastic remark instead.

Indeed, in Wagner et al. (2010) the impact of the surface prior update on the covariance matrix is analysed. Furthermore, in Luffarelli et al. (2017) the effect of the updating mechanism on the retrieval is also analysed (https://ieeexplore.ieee.org/document/8035227).

31. Eq.15 For the sake of future readers' comprehension, please restrict qi to the range [0, 1] and make QI a simple product rather than use the difficult to comprehend $1 - \max(q, 1)$ construction.

The range of qi has been restricted to [0,1]. However the QI construction cannot be replaced by a simple product as it would give the same results.

32. L409 Please specify this sigmoid function (or at least give it's width).

The width of the sigmoid function is now specified in lines 333-335 "When the mismatch assumes values within the range defined by T1 and T2, thresholds excluded, a value between a minimum m and 1 is assigned to the test 3 through a sigmoid function with width equal to $10/(T2 - T1)$ (Fig. 11)."

33. Fig. 14 Can we please have a version of this plot as a 2D histogram in the supplement, similar to the ones already there for the BRF?

We are not sure what the reviewer refers to, as there are no histograms here for the BRF.

34. The y-axis of Figs. 5, 6, 9, 10, 15, 17, 18, 19 should probably be 'Fractional counts'

considering they clearly have non-integer steps.

This has been done.

35. Fig.16 (b) and (c) aren't that interesting or helpful. Perhaps make (c) an inset in (a).

The figure has been removed.

All grammatical suggestions have been implemented.

Please also note the supplement to this comment:
https://www.atmos-meas-tech-discuss.net/amt-2018-265/amt-2018-265-AC2-supplement.pdf

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

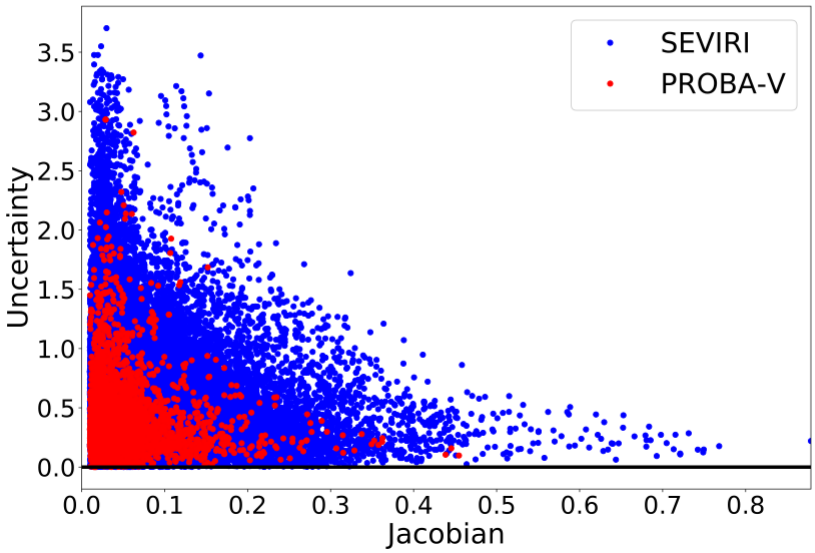

Fig. 1. AOT retrieval uncertainty as a function of the Jacobians