Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., author comment AC1
https://doi.org/10.5194/amt-2021-212-AC1, 2021

Reply on RC1

Frithjof Ehlers et al.

Author comment on "Optimization of Aeolus' aerosol optical properties by maximum-likelihood estimation" by Frithjof Ehlers et al., Atmos. Meas. Tech. Discuss., https://doi.org/10.5194/amt-2021-212-AC1, 2021

General comment:

First thing to note is that we were very glad about the thorough feedback of all Referees, who evidently made the effort to go through the details presented in our manuscript. This helped a lot to refine the text and we hope that we could make our statements more concise now. Especially the method of calculating the uncertainties in the MLE method was addressed by more than one of the Referees and required some changes that we want to present to all of Referees.

The procedure outlined in the last paragraph of Appendix C was relying on the simple equation alpha = gamma * beta with extinction coefficient alpha, lidar ratio gamma and backscatter coefficient beta. Motivated by comparing the obtained variances via (C2) with the variances one would get from a standard error propagation with constant lidar ratio sigma_a^2 = gamma^2 * sigma_b^2, we argued e.g. that the uncertainty in alpha cannot exceed 200 sr times the uncertainty of beta in a single bin, whatever the precise value of gamma is in reality, due to the upper bound in lidar ratio.

Now, after the feedback and some graphical considerations, the problem with this statement is, that it leaves aside any uncertainty or spread of the lidar ratio estimate itself, as it can lie somewhere in between 2 sr and 200 sr and hence should have non-vanishing variance sigma_g^2, causing another term to arise from the view of standard error propagation, increasing the uncertainty of alpha proportional to beta itself. This is why we noticed that the presented approximation was not correct, although we observed uncertainty values that matched approximately the SCA mid bin output for the SAL case for backscatter and extinction coefficients, which seemed realistic at first glance.

Following this, we evaluated the errors from equation (C2) directly against the uncertainties calculated by the SCA standard error propagation as in Flamant et al., 2020, and found good agreement, when the same Poisson noise variances are used in the S_y in (C2), see attached figure.

That means, that with (C2) we are only able to provide uncertainties for the unconstrained problem (as SCA solves it), but especially the so-obtained extinction uncertainties are overestimated and not representative of the real uncertainties, as one can see in the comparison of SCA and MLE in the simulation cases.

In lack of a representative alternative for uncertainty estimation we decided to delete the lidar ratio and extinction uncertainties of the MLE solutions in Figures 5 and 6, for now, stating that there is no representative error estimate available yet. The challenge of reporting representative uncertainty estimates is now being stressed in the abstract and conclusion and will require future investigation.

Regarding the comments of Referee 1:

The effort you made to provide such detailed feedback on our manuscript was greatly appreciated. Your greatest concern regards the convergence of the algorithm, on which we elaborate below in the comments and which led us to refine some statements in the manuscript as well. In short: We are currently trying to tune some of the parameters and to change the scaling of the input vector for the future so that the convergence speed is optimised. As this regards a technicality of the processing, rather than the final results (since the final cost function is reasonably low in all cases), we are confident to publish the obtained data.

Below you will find all the individual answers to your comments. All grammar and spelling corrections were taken into account directly, so we dropped them below to focus mostly on the comments regarding formulation and content.

- *In Eqn. 15 you say L_p > 0 but on L281 you say you start iteration at L_p=0, which is then out-of-bounds. Presumably you meant to say L_p >= 0. Regardless, I am surprised you chose to start at one end of your solution space. I would have started at some climatological mean value to keep the number of iterations down by giving the method the ability to increase or decrease L_p in the first step.*

This is correct, we meant to state L_p >= 0. It has been corrected. It is also probably not the best choice to initialise on the boundary of the solution space, but given that the climatological value should be rather small, it does presumably not make a big difference.

- *Forty thousand is a preposterous number of iterations! Are you sure L288 shouldn't say 40? Every optimization routine I've worked with tends to converge in 5 to 10 iterations unless the model is extremely non-linear. If you need more than several dozen iterations to get useable solutions, I would guess you've done something wrong – either a gradient is being miscalculated or an inappropriate optimization method is being used.*

When just one profile is considered, as in problem (14), then the number of iterations is on the order of 50 (roughly 0.5 seconds on an office laptop) to achieve convergence (an average cost function value per bin below 1). In practice, however, we are interested in solving problem (15) for all profiles simultaneously. In this case about 5000 iterations are required (roughly 45 seconds on an office laptop) to let the cost function decrease in the same way for a whole orbit (roughly 450 profiles). This information has been added to the manuscript.

So essentially, solving (15) for the whole orbit saves us time. Though, on second thought there must be a better converging formulation, but we have had no reference on this. We made an attempt to reformulate our model in terms of log(L_p), but the convergence speed did not increase noticeably. It is very likely that a more advantageous scaling exists since the L-BFGS-B algorithm is not scale invariant. Unfortunately we haven't found it and are currently tuning parameters to enhance performance.

- *I don't agree with you at L291 that artificially extending the number of iterations produces a fairer comparison to the SCA. This sounds like an attempt to avoid discussing the method of identifying convergence because those are contentious. I honestly don't care how you do it, but I do think the paper should explain how you intend to produce data outside of the context of this validation and, ideally, give some idea of the magnitude of difference. I suspect it makes very little difference, which is why you ran a set number of iterations, but that should be stated.*

You are right that for an operational use, the presented approach is not handy. So we decided to add at the end of the paragraph: "For operational use we plan to tune the number of iterations in an ad-hoc-fashion, based on if the average cost function value per bin has fallen below a value of 1."

In contrast to OEM and regularized MLE, the only contribution to the MLE cost function is how well the raw signals y are represented by F(x*). Once they are within the estimated limits, we can essentially stop iterating, because we cannot expect to gain any additional information content from the measurements. So this will determine convergence.

- *Something similar comes up again at L357 when you say the first guess "contaminates" your mean. A first guess is not a prior; it should not affect your final solution. Exceptions are where the method fails (and so you throw away the solution) or there are multiple minima (which would require a more advanced minimisation routine). Could you explain what you mean here? Is it just that you aren't doing any quality control and so failed retrievals are present in the output? If so, you should do quality control! You don't need a value from every single pixel.*

So the problem is the following: Whenever the retrieved aerosol optical depth L_p is very low, the influence of the lidar ratio estimate on the cost function becomes increasingly insignificant, until L_p becomes zero and no lidar ratio can be provided at all (these are not included in the statistics). Wherever L_p is very low, the lidar ratio has a high error estimate and some may remain close to the first guess, which then seems to act as an 'implicit a priori' (see https://doi-org.tudelft.idm.oclc.org/10.1029/JD095iD05p05587, end of section 7). What our paragraph wants to say is that, if lidar ratios were averaged without weights, then the first guess would contaminate the statistics. In the shown simulation case, this would only affect the statistics when only bins with retrieved backscatter coefficients below 5*10^(-2) (Mm sr)^(-1) are considered.

So eventually we propose to replace the sentence "Otherwise, the first guess of $\gamma_{||,p}=60$ sr would contaminate the statistics for MLE, e.g., $\text{mean}(\alpha_p/\beta_{||,p}) $ would be biased towards the first guess." with "[...], in order to disregard the influence of bins with nearly vanishing aerosol optical depth, for which no reliable lidar ratio can be retrieved."

Additionally, lines 360-361 and lines 380-381 are dropped, because they are more confusing than helpful.

- T*he text in figures should aim to be approximately the same size as the text in its caption. If it is possible, all the figures in this paper would benefit from being regenerated with a smaller page size, such that the font is larger relative to the image.*

The font size of figures 2,3,4 and D1, D2 has been adjusted.

- *L68) I would say backscatter is 'measured' rather than 'known' with higher precision.*

corrected

- *L77) Is there a reason you preferred to constrain L to be positive rather than retrieve its logarithm? I could see this being useful in your proposed future work to produce consistent regularization in variables that span several orders of magnitude. Further, aerosol optical depth (which is what you actually retrieve) is known to be log-normally distributed, such that the log of L is a more natural basis in which to define a regularization.*

There is actually no elaborate reason for the choice of L instead of log(L), but it might contribute to proper scaling of the variables, indeed. We will investigate this aspect in future investigation, also regarding the convergence speed.

- *L103) I feel like this paper demonstrates that such a retrieval is possible rather than it being something you need to assert in advance.*

The possibilities were discussed pre-launch and the SCA has been implemented already before. The paper's scope is improving the precision of the retrieval.

- *L204) I've not encountered this use of a backslash before. Did you mean to say "a ratio"?*

No, we meant "or". It has now been dropped for convenience.

- *L230) While true, lidar signals are often far from this limit.*

Considering also the comment of Reviewer 2, we changed the sentence to

"[...] because the discrete Poisson noise distribution can already be well fitted by a smooth Gaussian with identical mean and variance for very low (photon) counts and the aforementioned additional noise sources and their corrections, e.g., subtraction of measured solar background, will naturally smear out the discrete nature of the Poisson noise."

- *L261) I think that the two uncertainty ranges shown in Fig. 5 come from the two sources of uncertainty you mention here – a simple Poisson assumption and the variance of the downlinked profiles – but it would be useful for that to be stated clearly somewhere. This sentence currently implies the Poisson approximation is not used in the remainder of the paper, but you go on to mention it several times.*

Thanks for underlining this source of incovenience. The Poisson approximations for MLE in Fig 5 and 6 have only been used for the uncertainty estimation and not in the minimization. Following the feedback of Reviewer 2 and 3, we noticed that our implementation of the box-constraints in the error calculations was flawed. In the lack of an alternative we removed what we initially claimed to be an upper bound in Figures 5 and 6.

We also edited the last paragraph of 3.2 to explain that the uncertainties of SCA and SCA MB are calculated as in section 6 of the Algorithm Theoretical Baseline Document for the L2A and changed the name of this uncertainty to "Poisson error estimate" in Figures 5 and 6.

- *The first paragraph of Section 4 unnecessarily repeats the preceding paragraph.*

The paragraph 3.3 has now been deleted.

- *I would mention the existence of Appendix D at the end of L343 as that's when I asked "what's the RMS deviation"?*

According to the feedback of Reviewer 3 we promoted a modified version of figure D1 (with zoomed in boxes for better visibility below 2 km altitudes) to the main text and added a more detailed discussion of the biases and relative errors in the regime of most interest below 2 km altitude.

- *In the caption for Fig. 4, don't you mean the west coast? Also, the description of 4(b) implies that three versions of the feature mask are shown. I think you meant to say that the rows show backscatter, extinction and lidar ratio, with identical features masks shown in each frame of the rightmost column.*

We corrected for both errors.

- *L418) I'd argue that the image shows the advantage of forcing L>0 rather than robustness.*

This sentence lacked a concise message indeed, so we changed it to say "[...] which demonstrates the advantage of the box-constraints."

- *L493) Is there any intention to release this data? Any possibility of funding to process the full record or become an operational product?*

The MLE is being considered for operational implementation.

- *Appendix A largely repeats Section 3.1.*

This was considered to be more reader friendly, since section 3.1 presents the bare minimum of derivations.

- *L516) ECMWF data would typically be interpolated rather than averaged as the values represent behaviour at a point rather than a grid cell average. What do you mean by "mean" here?*

We write "by means of" in the sence of "through" or "with the help of" and do not refer to the statistical mean.

- *L32) addressing high uncertainties in climate change modelling due to the indirect*

We changed this sentence to "[...] for the largest single cause of uncertainty in antropogenic radiative forcing has been reported to be from the indirect effect of aerosols on clouds (Illingworth ...)"

- *L93) Traditionally, "Earth" is the planet and "earth" is dirt so I don't see why you use both versions here.*

Although amusing in its current form, we corrected the non-capital "earth"s.

- *I don't think "Products" is necessary in the title of the paper. Then again, I would have called it "Optimization of Aerosol Optical Properties from Aeolus Profiles by Maximum-Likelihood Estimation" because I'm concerned that someone might think the "Optical Properties" are possessed by Aeolus rather than by particles in the atmosphere.*

This is a very good point, also considering that the Journal's scope can be on both measurement systems and the data analysis. So we agree to change the title.

- *Remove "(partly)" from L8. A problem is either ill posed or it isn't; there is no partly.*

We removed ill-posed as what we mean is "sensitive to or ambiguous under noise influence".

- *L52) A white paper under preparation in the UK is hoping to call these "representation errors" rather than worry about 'representivity' vs 'representativeness'. Obviously there's no obligation to agree with us but I think it's a cleaner phrase.*

Fair enough, we follow this advice since the term "representativeness" seems to bulky anyhow.

- *Fig 1) While this image is technically "Exemplary" in that you are using it as an example, English has corrupted the word to typically mean "Outstanding". I'd say "Illustrative" or "Example of" instead.*

Thanks! We did indeed not mean to exaggeratedly underline the presented data.

- *Eq 2-3) I know you're using T for transmission but it hurts slightly to see temperature denoted t. I would have used \mathcal{T} for transmission.*

I agree but this way the difference is clearer and comes without unnecessary flourish. If we had used temperature frequently, we would have changed it for sure.

- *L370) "...reacts after the backscatter coefficient, resulting in the attenuation of the first 500 m of the cloud being captured incorrectly with consequences..." I struggled to understand this sentence and have made my best guess at what you were trying to say.*

We changed it to "the SCA extinction coefficient reacts delayed compared to the backscatter coefficient, leading to that the attenuation by the first 500m of the cloud is not captured correctly".

- *Fig 4) Is it necessary to use very similar colours for the two lines? If you moved away from a rainbow colour bar, you would have more options so I don't have to strain to see if the line has little dots in it or not.*

Our excuses, but this data has been plotted through online services. I completely agree that rainbow color bars are no optimal choice.

- *L463) Perhaps "refactored" rather than "rephrased"? The latter usually refers to words rather than actions.*

We chose "reformulated" instead.

- *L479) Considering you go on to critique the use of a feature mask, you might want to say "it is possible to" rather than "it would be advantageous to".*

Thanks for making this statement more consistent.

- *Eq C2) I have never seen the notation K^{-T} to indicate an inverse transpose. Probably because it's ambiguous if I transpose the inverse or inverse the transpose. (K^T S_Y^{-1} K)^{-1} is a more traditional way to write this equation.*

Indeed true, thanks. We changed it to (K^{-1})^T instead.

Please also note the supplement to this comment:
https://amt.copernicus.org/preprints/amt-2021-212/amt-2021-212-AC1-supplement.zip