Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., referee comment RC3
https://doi.org/10.5194/amt-2021-212-RC3, 2021

Comment on amt-2021-212

Anonymous Referee #3

Referee 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-RC3, 2021

Review of "Optimization of Aeolus Optical Properties Products by Maximum-Likelihood Estimation" by Ehlers et al.

The article describes a new algorithm for retrieving aerosol optical properties from Aeolus that improves various shortcomings of the existing Aeolus algorithms. Specifically, the new algorithm operates on the direct measurements and their uncertainties (rather than cross-talk corrected pseudo-signals ignoring uncertainty), simultaneously fits both extinction and backscatter (rather than sequentially), and uses explicit constraints to force solutions to have non-negative optical depth and lidar ratio within expected limits (rather than implementing these constraints as filters after the retrieval). The analysis includes the calculation of error bounds that appear to better reflect actual uncertainties than those of the standard algorithm. Comparisons between the new and existing algorithms are presented for simulated scenes and real data scenes where correlative lidar measurements are available.

I have an overall favorable impression of the manuscript and look forward to seeing it published. The improvements over the existing algorithms are useful and justify the publication of this study. An appropriate level of detail is presented and the logical flow of the manuscript is good and helpful for understanding the algorithm and results. However, there are some important aspects which need clarification and revision, particularly those relating to the uncertainty calculations. Since I see these as critically important, I am recommending major revisions, despite the overall high quality of the manuscript.

Major concerns

P3 L77. A positivity constraint on quantities where zero is a valid and common expected value will produce a bias. The authors mention this in the context of the SCA algorithm, but this is also a concern in the new algorithm. It might be better not to include the positivity constraint on backscatter and extinction in the new algorithm. Have the authors tried their MLE method without the positivity constraints?

p10 Box constraints. I think its possible that adding constraints has a smaller impact on improving the results than the other improvements: using raw measurements and uncertaintites in a coupled retrieval. Additionally, the constraints have potential negative consequencs: (1) that they could lead to bias (similar to what happens in SCA) and (2) that they make propagation of the uncertainty very difficult. Have the authors attempted the MLE retrieval without adding the box constraints? It would be useful to compare results with and without using the box constraints. If they are just as good, eliminating the box constraints would eliminate the two issues mentioned above. If they are not as good, the comparison would give a clearer view of the impact of the different improvements.

P19, L469 "This is mainly due". Similar to the previous comment. This statement is made without any analysis to suggest how the different algorithm features dominate the improvements. The authors should assess the impact of the box constraints separately, to support this statement that it is the dominant cause of improved results.

P6 Eqs 2 and 3. P21, Eqs A4 and A5. Very little is said about the cross-talk constants C1, C2, C3 and C4. The top of page 7 seems to imply the constants are known, but in fact they are very challenging to determine and have significant uncertainty associated with them. More information is needed about how these are determined and what are typical values for Aeolus. This is also important for the discussion of the correlated errors and for understanding the magnitude of that problem.

p10-11 L294-304, two approaches to error quantification. While in general I like the idea of analytical propagation of errors using sensitivity analysis, the propagation is seriously challenged by the difficulty of representing the impact of the constraint in the propagation equation. I'm not very convinced by this method of rescaling the variance and assuming the correlation is unchanged (Appendix C, near P23 L575). Can it be explained more clearly why the correlation matrix should remain unchanged? Also, using the lidar ratio constraint to (potentially) reduce both the extinction and backscatter error bounds separately gives an impression of circularity. The authors should add rigor and validation to make this method more convincing.

P25 Figure D1. Although the Monte Carlo methodology is discussed, it seems that the results of the Monte Carlo approach are only shown in Appendix D without any analysis or discussion. These should be promoted to main text and properly analyzed.

Both methods are available for the simulation cases, so I'd like to see a comparison of the analytical (Appendix C) method against the Monte Carlo results for the simulation cases, which might bolster my confidence in the analytic method.

It would also be useful to compare the propagated uncertaities with and without the adjustment for the constraints, to see how big this impact is.

The authors say a Monte Carlo approach is not feasible for the measurement cases, and I think this refers to the fact that they cannot vary the true measurement error. However, it is certainly possible to simulate measurement noise from the measurement error covariance matrix, S_y. Using these simulated measurement errors in a Monte Carlo approach would give an accurate view of how S_y is propagated through the retrieval, more immediately convincing than the method in Appendix C. So, going further, I'd like to see a Monte Carlo propagation of S_y, which could be used to validate or replace the current propagation.

P17, Figure 5 (also figure 6). I can't understand which error bounds are shown in Figure 5 and Figure 6. The descriptions "lower error bound" and "upper error bound" are confusing because they don't match any description in the methodolgy, and they are possibly inaccurate as well. Another terminology should be seletected, preferably one that reflects the descriptions in the methodology section. The two sets don't seem to be the "two approaches" introduced on P10-11. Are they instead two formulations of the analytic propagation of errors using different characterizations of the measurement error S_y? Is the "upper error bound" the propagated uncertainty from Appendix C? (If so, calling it an upper error bound is particularly problematic since the Cramer-Rao inequality characterizes a lower bound, not an upper bound). And how is the "lower error bound (Poisson)" calculated? I don't see that in the manuscript.

If indeed the one labeled "upper error bound" is the one produced in this analysis, and the one labeled "lower error bound" is the standard one for SCA and midbin-SCA, then another emergent theme of this analysis is that the uncertainty of the standard algorithms is inaccurate as well. This should be highlighted in the manuscript as another primary impact of the new algorithm.

Minor comments

P1 L8 "algebraic inversion scheme to a (partly) ill-posed problem and therefore sensitive to measurement noise". The sentence should probably be reworked. "(partly) ill-posed" is not really informative, and even a well-posed inversion is sensitive to measurement noise.

P2 L53-54 I think this point about there being no resolution between these regimes is particularly well articulated.

P3, L78-79. I think the sentence should be split up into multiple sentences that clearly list the factors that led to poor results in the SCA (there are at least two, I think: the correlated noise and the incorrect removal of negative extinctions) and the factors that lead to improvements (coupled retrieval, using the measurement error covariance matrix in the retrieval, box constraints), and to explain their causal relationships. In particular, I believe it is incorrect to suggest that the box constraints are responsible for eliminating the correlated noise issue, as this sentence seems to say. Instead, I think the bias is avoided by using the measurement channels and their error covariance matrix in a coupled retrieval.

P4, L130. How is the grid spacing chosen? What are the typical values?

p7 Eqs 9 and 10. The minus sign in the denominator does not appear in Flament et al 2017. The rest of the equation looks similar but with a few more algebraic steps. I think it is otherwise correct, but I would like the authors to double-check to be sure.

p7 Eqs. 9 and 10. It would be helpful to have a sentence or two summarizing the derivation of the quoted equations; for example, an explanation that the equation includes an expression that explictly integrates the transmission over an extended vertical bin.

p8 near the the end of section 3.1, it would be good to have a sentence or two discussing the anti-correlated noise that's mentioned in the Appendix. This seems like an important point that is referred to both in the introduction and later, so it should be described in the main text with enough detail so a reader knows what it is about.

P9 L232. I'd like to know more about how the box constraint is implemented and specifically whether it may negatively impact results. Could it potentially bias the results (by perhaps producing a disproportionate number of solutions near the boundaries) or affect their precision of the solution (via a variable transform that changes the gradients near the solution and near the boundaries)?

p9 242. "future developments". This is really more of a response to another reviewer, but I just wanted to say that I support the authors' decision to perform single profile retrievals without including horizontal scene smoothing. It makes sense to explore the simpler solution first. Furthermore, not requiring scene continuity for the algorithm means that scene continuity can be used as a check on algorithm performance.

p9 L250 As I mentioned, I believe that using the measurements and their appropriate measurement error covariance matrix is largely responsible for avoiding the bias due to the correlated errors in the cross-talk corrected signals. Do the authors agree? If so, this paragraph might be a good place to highlight that.

p10 L272-275. Is there anything in the current algorithm that addresses the bias from partially filled bins? If not, is the current algorithm compatible with the strategy implemented by Flament et al. 2017?

p10 L290-292. There appear to be three rather weak arguments here to justify not using an a priori covariance matrix such as called for by OE. The weakest part of the argument, in my opinion, is saying that the algorithm has no prior information or regularization. In fact, the constraint is prior information and does provide regularization. If the authors believe the influence of the constraint is significant (which I think they do, because they attribute the improved results largely to implementing the constraint), then even suggesting that this is a weak constraint would not be justified. Therefore, the authors should acknowledge that the constraint plays the role of prior information in the retreival, but they chose not to cast this constraint in the terms required by optimal estimation, because that would require a different (probabalistic) form for the prior that isn't compatible with the desired form of the constraint. In my opinion, that is a reasonable reason to use constrained MLE rather than OE. (However, the choice leads to difficulties in working out the correct way to calculate the impact of the constraint on the uncertainties, which I discussed in the "Major" section).

p10 L295. Some clarification would be helpful here about the simulated measurement noise. Is it simulated as gaussian with the variance determined from Appendix B, or is it simulated with various separate error components? Which components are included?

p12 L347-349. While I certainly agree that filtering negative results would cause a high bias that is worse in bins with low SNR, I can't clearly follow the explanation in this part of the text and I don't see a particularly clear indicator that they are triggered by the shift from coarser high-altitude bins to finer low-altitude bins as implied by the text. Can the explanation be clarified?

p12 L354. The indicator from the averaging kernal is a good idea. Is there a quality flag in the data product related to this? It would be nice if this indicator were included in the plots to show which bins are not trustworthy. This would be particularly useful in Figure 4, for instance, where I am curious to know if the bins below the lofted aerosol are reliably retrieved. This curiosity is fueled by the fact that CALIPSO's 532 nm data frequently misses the aerosol at the bottom of attached layers due to attenuation, making them look lofted. The HSRL capability should act to prevent this problem given sufficient SNR, but on the other hand, attenuation at the shorter wavelength of 355 nm would be worse. So, to be sure, it would be good to see some indication from the retrieval algorithm that the lack of aerosol below the apparent plume bottom is reliable.

p12 L354. "The extinction bin closest to the ground cannot be well retrieved." Is this true in general, or specific to this case?

p12 L357 "Otherwise". I'm confused by this sentence, and not sure if I'm confused by content or just the wording. Does "otherwise" indicate the strategy of taking the mean lidar ratio? I agree that taking the ratio of the mean extinction and backscatter is a better strategy than the mean lidar ratio and will give potentially different results, since lidar ratio for small extinction and backscatter is not as reliable as when the SNR is higher. By why is the mean lidar ratio contaminated by the first guess? And how does taking the ratio of mean extinction and backscatter avoid the influence of the first guess?

Figure 2. What is the explanation for the low bias in the median lidar ratio from the MLE for the lowest bins?

Figure 2-4 and D1: The line plots are so small that I can't see important information. The data are particularly important below about 2.5 km where there is significant aerosol, but this is a very small portion of the plot and not readable due to its size and the closely spaced grid lines. Specifically, the lowest bin is called out in the text, but it is so small I cannot see the data or error bar for that lowest bin in the line plots. Please include inset boxes to show the line plot data in the lowest 2.5 km, or a second set of line plot figures that show only the lowest 2.5 km, or in some other way improve the visualization of the lowest 2.5 km.

p14 L362. The reduced uncertainty from the new retrieval compared to the SCA should be discussed and quantified. I think the new retrieval probably produces usefully smaller uncertainty, but the fact that the figures are so tight makes it nearly impossible to see the region where there is significant aerosol below ~2.5 km, and I can't even tell if the error bars are smaller than 100%. A discussion of uncertainty results is just as important as the mean tendency of the profile, because, for example, a profile that "looks right" but is indistiguishable from 0 everywhere due to its uncertainty would be rather useless.

P14, L366. Is this number a typo? exp(-2*0.4) = 0.45 not 0.38.

P14, L369-379. It seems that it's not just delayed (i.e. slow decay below the cloud) but the cloud is smoothed into the regions both below and above. In other words, it looks like the effective resolution of the extinction is much less than the backscatter, which make sense since it takes at least two measurements to calculate a derivative. Is there any attempt to calculate backscatter on the same coarser resolution as extinction to produce the lidar ratio?

P15, Figure 3. It appears that the bias in the mean backscatter below the cloud is actually worse in the MLE than in the SCA and SCA-midbin results. This should be discussed.

P19, L449, what causes missing values in SCA midbin?

P19, L458, Lidar ratios aren't shown in the figure for the Real Data Case II, as in other cases. Better to show them, if possible.

P19, L459-460, Is this comparison between copolarized lidar ratio from AEOLUS to total (co- and cross-polarized) lidar ratio from Polly? This is not the best option. Since Polly is also sensitive to polarization, wouldn't it be possible to calculate copolarized lidar ratio from Polly for a more direct comparison? In any case, it should be clearly stated what's being compared.

P19, L461. It seems strange that the quoted uncertainties are from the Poisson method that the authors are hoping to replace. It would be better to quote uncertainties from the method that the authors think are more represenatative.

P 20, Figure 6. Is there an uncertainty bound available for the "ground truth" (which is also a retrieval)?

P20, Eqs A2 and A3 are difficult to mentally derive from the previous step. It would be helpful if more intermediate steps were added to make it clearer how the equation is derived. The appendix is a good place to do this.

P22, L538. This point about anticorrelated error is interesting and informative. However, some parts of the discussion are confusing. For instance, Eqs A7 and A8 show that an error spike in one measured channel gets distributed in an anticorrelated way into the corrected signals. But how does it follow that there is correlation (or anticorrelation) in the errors in backscatter and extintion? It seems logical that if the errors in the two corrected channels are anticorrelated, the ratio would tend to be biased low, so the backscatter would tend to be biased low. However, the error in the extinction would not be correlated with it, because extinction derives from gradients in just one of these corrected channels.

P23, L557, I suggest replacing "in spirit" with something more informative, such as "except with an adaptation to account for the impact of the constraint"

P23, L561. It's not true that no a priori knowledge is imposed. The box constraints are a priori knowledge.

P25 Figure D1. "Bias" and "standard error" should be defined in equations.

Grammar and word usage

I found that, although the overall flow and logic are very quite good, in some spots the word choice made the writing difficult to understand. I've marked several below. My list is probably not complete, however, so a round of editing for English usage (not just spelling and grammar) would be helpful.

Title. The title would be more informative if it contained the word "aerosol". For example: Optimization of Aeolus Aerosol Optical Properties Products by Maximum-Likelihood Estimation

P1 L1. Typically "embarks" is not used this way, and used only for the much narrower circumstance of a person getting on or off a vessel or starting a journey. It could be replaced with "includes"

P1 L5. Being an HSRL

P1 L9. Consider replacing "rephrase" with "rework" or "revise"

P1 L12. Delete "equally". I think most writers would use "also" rather than "equally" here, but it really isn't necessary at all.

P1 L14. Consider moving the phrase "to consolidate and illustrate the improvements" to the beginning of the sentence. I find it easier to follow when phrases that modify the verb are near the verb.

P2 L32. "Because of high uncertainties" and "regarding the indirect effect" (add "of" and "ing")

p2 L34. Delete "aspect of"

p2 L34. Consider inserting a paragraph break after "coverage"

P2 L53. Insert "For Aeolus" at the start of the sentence that begins "There is no suitable resolution"

p3 L64. I suggest deleting the parentheses around "particle"

p3 L69-70. I'm not sure I understand what is meant by "well located". Does this mean retrieved at finer resolution? (Perhaps not, because if so, it still doesn't quite make sense. I understand that extinction and backscatter are retrieved simultaneously on the same grid, but the effective resolution of extinction and lidar ratio will always be less than the finest possible backscatter resolution because it takes at least two measurements to determine a derivative.)

p4 L94. Delete "means of"

p4 L96. emitted

P4, L130. Consider adding "irregular grid" to this description. Perhaps here: "in steps of 250m, to produce an irregular grid with a total number"

p5 Figure caption. "Example" rather than "exemplary"

p7 L179. "Extensive use" rather than "Excessive use"

p7 L180. Replace "lightened" with "simplified" p7 Equations 5-8. I think the variable names and subscripts could be chosen to better indicate which set of variables are the raw signals and which are the cross-talk corrected signals. It's particularly confusing that the pair with mixed Rayleigh and Mie components are subscripted "ray" and "mie" while the pair where they are actually separate is generic with no mnemonic. Using p and m subscripts for the corrected measurements in 5 and 6 might help. I would also suggest renaming the variables in 7 and 8 to avoid the ray and mie subscripts on the raw channels (although I know that suggestion might be more controversial because it's based on historical precedent with this kind of instrument).

p8 L203 Delete "can be rephrased into". The slash notation is confusing and unnecessary. In this case "are equivalent to" is the better phrase.

p8 L209 Appendix, not Annex

p8 L227, I believe the meaning will be clearer if you remove "to account for" and instead say "because additional noise contributions, such as ..., are likely to dominate over the Poisson noise"

p9 L246, "In theory" instead of "Principally"

p9 L252, "simplify" instead of "lighten"

p9-10, L261-263, I had trouble understanding this sentence until I read the Appendix. It might be clearer to move "scaled" and break the sentence into two: "Here we use the variance measured at 2.9 km resolution, scaled to approximate the noise level in the 87 km bins. This approximation assumes the scene is homogeneous so that all the variability is due to measurement error."

p10 L269 replace "artificially increased" with "increased"

p10 L272-275. I found these few sentences very confusing. I think the authors are saying the extinction bias due to underfilled bins tends to decrease the measured range of lidar ratio, but that this is not a compelling reason to reduce the lidar ratio upper bound in the algorithm, so they end up ignoring the bias found by Flament et al. 2017. If the effect is ignored, then these sentences are somewhat of a distraction and could simply be deleted. If the authors feel that it is important to keep these sentences about the bias due to underfilled bins, (and if my interpretation is correct), I think the readability could be improved by (1) signalling the contrasting thought using "On the other hand" (or "In contrast") instead of "Additionally", (2) using "aerosol partially filling a range bin" instead of "different hypothesis on the distribution of aerosol layers within a range bin" (3) specifying "underestimate" instead of "alter" and (4) correcting "co-polarized particle backscatter coefficients" to "co-polarized measurements". That is, the bias due to underfilled bins is a bias to extinction, not to particle backscatter coefficients.

p10 L288. Forty or forty thousand? Forty seems more likely, in which case three digits after the decimal point for an integer is a strange instance of false precision (should be just "40"). Or if forty thousand, it should be a comma not a point "40,000" - but in that case, 40,000 is a crazily large number of iterations. What is the typical number of iterations actually required for convergence?

p10 L289. "many fewer" instead of "much less"

p11 L305 Consider inserting Equations C2 here. The sentence could read "This relation is inverted to produce Eq. C2", and deleting the terminology "Moore-Penrose pseudoinverse" in the main text, since it's more understandable in the Appendix where it is explained more completely.

p12 L339. "optically thin" rather than "thin"

p12 L350. "at the price of" instead of "to the price of"

Figures: I hope Figures 2-4 will be larger with proportionately larger text. They are quite difficult to read.

P14, L361 "either" not "wither"

P14, L381 "either" not "wither"

P14, L383, I'm not sure what "locks extinction and backscatter to appear colocated" means. Does it mean something like "MLE retrieves extinction and backscatter at the same effective resolution"?

Figure 4 caption. "CALIPSO Feature Mask". Better to use the official CALIPSO product name and be specific: is it the "Vertical Feature Mask (VFM)" or the "Atmospheric Volume

Descriptor (AVD)"?

Figure 4 caption, the sentence starting "The rightmost column shows the Calipso Featuremask" is confusing. This could be rearranged to first describe the rows as backscatter, extinction, and lidar ratio. Then in a separate sentence, say the CALIPSO feature mask is repeated on each row.

Figure 5 caption, "error bounds" rather than "errors"

P17 L408-409. Delete "and it might be horizontally and vertically". It is disrupted, not might be.

P18 L419. I suggest using "copolarized" rather than "co-polar", to be consistent with how you have described it earlier.

P18 L427. Explicitly include copolarized lidar ratio in this sentence: "expected values of copolarized lidar ratio of 80sr - 120sr for desert dust". It would be good to avoid the possibility of these expected values being taken out of context and mistaken for the more usual non-polarized lidar ratio.

P18 L431. Replace "all values" with "expected values" and replace "the estimated error ranges" with "the error ranges estimated from Eqs C2 and C3" (or some other phrase to make it clear that it's the error ranges calculated from the new methodology. They do not fall within the "poisson" error bounds.)

P19 L449. ":Hence, " is not the right connector since it implies causality, and consequently I'm not exactly sure I know what is meant. I think the authors are only explaining what was meant in the first part, so perhaps replace with "; that is, ".

P19 L470. Replace "with this" with "in addition". Again "with this" is usually used idiomatically to indicate a causal relation.

P19 L472-474. The flow of this paragraph is interrupted. I suggest moving these sentences about how the coupled retrieval improves backscatter to the earlier spot at L469 immediately after the statement that backscatter is improved along with extinction. The information about the constraints and the anti-correlated noise are separate thoughts.

P20 L483. "suitable" rather than "suited"

P21 L499. "the properties above" doesn't make sense here, since no properties have been defined yet in the Appendix.

P21 L515. Keep "are equivalent to" and delete "can be rephrased into"

P21 L521, I suggest replacing "optical depth/extinction" with "layer optical depth". Dropping "extinction from this sentence is no loss since the relationship with extinction is described in the very next sentence.