Atmospheric Measurement Techniques Discussions

Atmos. Meas. Tech. Discuss., author comment AC2
https://doi.org/10.5194/amt-2021-212-AC2, 2021

Reply on RC2

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-AC2, 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 comments of Referee 2

We thank for the thorough review of our manuscript and are hopefully able to clarify on most aspects. The main concerns of Referee 2 regard the comparability of the MLE method versus a purely algebraic inversion scheme as benchmark. We can clarify here, that the forward model of SCA and MLE is virtually identical (besides some equivalent variable transforms outlined in section 3.1). Hence, the only difference between SCA and MLE is how the forward model equations are solved/inverted. That implies, that any bias due to the forward model formulation should be shared, while any other differences in bias must then either regard the zero-flooring (only in SCA extinction) or the interaction of non-linearities and noise levels (in SCA and SCA MB). Because this point was not made entirely clear in the manuscript, we modified the main text as outlined in the individual comments below.
All grammar and spelling errors were adopted or corrected without explicit mentioning.

Lines 81 to 83: In the references that are cited, the work were specific to developing inversion methods for ground based lidar instrument. The authors are proposing a spaceborne inversion HSRL method. This distinction matters because for ground based lidar instruments inaccuracies in the geometric overlap function calibration parameter can introduce biases in the extinction coefficient. Hence, the backscatter coefficient is inferred as a separate first step to isolate any geometric overlap function induced biases to the extinction coefficient. With spaceborne HSRL measurements and depending on the instrument, there is not necessarily a geometric overlap function which are confirmed by equations 2 and 3 in this paper. Thus, a spaceborne HSRL allows for the simultaneous inference of the backscatter and extinction coefficients that are not biased to due to the geometric overlap function.

For the sake of understanding how this paper's method relates to the references that are cited, the reviewer suggests that the authors mention that the geometric overlap function calibration parameter is a limiting factor for ground-based HSRL instrument when trying to simultaneously infer both the backscatter and extinction coefficients.

We acknowledge the reviewers point of view and suggest to add the sentence: "It is important to note that such a simultaneous retrieval with ground based lidars would require the additional geometric overlap function calibration parameter. Hence, it is often preferred to retrieve backscatter coefficients independent of extinction coefficients to mitigate biases on the former due to calibration errors."

Line 198: It is unclear where in (Flamant et al 2017) equations 9 and 10 are defined. Since equations 9 and 10 are esoteric, it would be helpful for the reader to know where in (Flamant et al 2017) equations 9 and 10 are defined since (Flamant et al 2017) has 124 pages.

We are aware that these equations are rather unappealing, but felt the need to state them once in the exact form they are used. We now refer to the equivalent equations by number in the cited document: "[...] see equations 6.35-6.36 in Flamant et al. (2017) and definitions above." and also refer to equations (5) and (6) just above.

Lines 199-200: Equations 9 and 10 do not make sense. For example, consider equation 9. Since L_{p,i} > 0, we have that -2L_{p,i} < 0. Hence, X_{i} < 0, since the rest of the *terms in equation 9 are positive. However, this a contradiction, since X(r) in equation 5 is strictly positive.*

Indeed the minus in the denominator was incorrect so it was removed in in (9) and (10) and (A2) and (A3). For your readability: The function (1-exp(-2x))/(2x) is almost equal to exp(-x) for small x.

Line 224: Should J be J_{obs}? Should S_y not be the inverse of the measurement error covariance matrix? Refer to equations 2.21 and 5.3 of (Rodgers, 2000).

Thanks for spotting these mistakes, we corrected for them throughout the whole manuscript and added the suggested references after line 224.

Line 287: The reviewer does not agree with the statement "MLE estimates usually suffer from overfitting and noise amplification.", because it is the forward model along with the MLE that determines whether there is overfitting or "noise amplification". E.g., the averaging operator (or just averaging) can be derived via MLE and averaging operator with a sufficient number of samples does not overfit. In other words, whether a MLE overfits or does "noise amplification" depends on the MLE's parameterization with the forward model.

This is a good point and was certainly expressed with too little care in the manuscript. We now changed the wording to "MLE estimates may suffer from overfitting and noise amplification in lidar retrieval problems, so an implicit regularization is often achieved by optimal choice of the number of iterations [...]", while keeping the existing references, in which more information can be found.

Line 230 to 232: Could the authors provide a reference to support the claim that the averaged signal's noise can be approximated via a Gaussian distribution? For example, are there sufficient photon counts during the nighttime at 30km altitude that are accumulated and normalized in order for the average signal's noise (at 30km) to be accurately approximated by a Gaussian distribution?

The reference to the central limit theorem was misleading. What we meant to say is that besides its discrete nature and asymmetry, the Poisson distribution can be decently well approximated by a Gaussian with same mean and variance. So instead we propose to change the old wording

"[...] because the Poisson noise distribution becomes indistinguishable from a Gaussian given sufficient signal accumulation (central limit theorem)."

into

"[...] because the discrete Poisson noise distribution can already be decently well approximated 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."

Lines 259 to 261: Poisson noise is not additive and therefore it is confusing to read \hat{sigma}_{s} = \sqrt{s + \epsilon_{s}}. For Poisson noisy observations the noise standard deviation is \sqrt{s + b_s + b_d} where b_s is the solar background radiation and b_d is the dark counts.

The mathematical formulation was meant to support the understanding, but can indeed be misleading here as it was inspired by the perspective of Gaussian noise. We propose to remove this part and include "because a single draw does in general not equal the true mean and the true variance" as a reason for this bias from using the signal samples themselves.

Lines 291 to 292: The statement "the estimate should fit as close as possible to the signal data and only solve the physical contradictions" is contradictory compared to the previous sentences in this paragraph. If the estimate should fit as close as possible to the observations (signal data), then why not let the L-BFGS-B converge to a solution and why bother introduce an early stopping criteria of 40 000 iterations? In other words, in this sentence the authors are suggesting that the estimator should "overfit".

There is no contradiction if read the following way: Since the number of 40.000 iterations is really high (see other reviewer's remarks), we try to maximise convergence by brute force and hence aim to eradicate any implicit regularization additional to the constraints.

Lines 431 to 433: Please add a reference that shows the range of lidar ratio values for desert dust at wavelength 355nm.

This is certainly helpful for the flow of reading. We added "(Wandinger, 2015)".

Lines 457 to 460: Are the authors reporting co-polarized lidar ratios or BERs? If BERs are reported, how were the co-polarized lidar ratios converted to BERs?

No BER is calculated here, but we refer explicitly to the co-polarized lidar ratio now.

Sections 4.1 and 4.2: The authors should validate the uncertainty quantification as described between lines 300 and 305, since it is unclear how accurate the proposed uncertainty quantification is for observational data.

see statement of the author's

Section 4 & conclusion: The authors do not thoroughly explain why SCA methods produce biased results compared to the MLE method, and it will be insightful to know why SCA methods produce biased results. Could it be that the X and Y (equations 7 and 8) are modeled by SCA (equations 9 and 10) could introduces biases in the estimates of the backscatter and extinction coefficients relative to the MLE method, since the MLE method employs equations 7 and 8?

The forward model for SCA and MLE is equivalent. The MLE also relies on eq. 9 and 10, see line 212. So this is not the reason for biases.

The biases occur in our opinion mostly due to the non-linear operations in equations (9) and (10): Take as an example the ratio operation and arbitrary random variables x and y such that mean(x)/mean(y) = 1. Now, depending on the noise amplitudes on x and y, mean(x/y) can be biased high to any value, solely because of the varying noise amplitude. Considering the non-linear character of equations (9) and (10), it becomes explainable that biases in SCA and SCA MB can be (partly) mitigated with MLE by suppressing noise. This is now also discussed in more detail in sections 4.1 and 4.2, aided by better visualisation of the significant aerosol below 2 km with modified versions of graphs D1 and D2. We added:

"For the origin of this bias we can think of two causes: Firstly, the violation of the hypothesis of uniformly filled bins due to the strong gradient in aerosol concentration and, secondly, the non-linear way the backscatter coefficient is calculated from $\beta_{||,p,i}=Y_i\beta_{m,i}/X_i$, because here $mean(\beta_{||,p,i})$ will become biased high increasingly with increasing uncertainty of X_i."   and   "The high backscatter coefficient bias in the lowermost 2 km [in simulation case II] is more pronounced in all algorithms as compared to atmospheric simulation case I, supporting the hypothesis that it originates from the non-linear ratio operation rather than the gradient of the aerosol concentration. Consequently, due to its noise suppression capabilities, that is likely why MLE shows less bias in the lowermost 2 km. A similar statements can be made for the extinction, because it is essentially calculated from a ratio of subsequent pure molecular signal values X_i."

Comment about the methods: It will also be useful know what is the performance of a method that directly algebraically solve for the lidar ratio and backscatter coefficient via equations 7 and 8. In other words, what is the performance of the MLE method without constraints? The reason why this comparison will be useful, is to gain an understanding of the low performance of the SCA methods. Specifically, is the loss of performance due to the formulation in equations 9 and 10, or because of the lack of physical constraints?

As stated above, both SCA and MLE rely on the same equations (9) and (10). So the SCA is the direct, algebraical solver for the MLE problem, as suggested. Hence, if the MLE had no constraints and was initialised with the same condition of vanishing optical depth in the first bin (as SCA is) then outputs of SCA (without zero flooring of extinction, see e.g. figure 10 in https://earth.esa.int/eogateway/documents/20142/37627/Aeolus-Data-Innovation-Science-Cluster-DISC-Level-2A-user-guide.pdf for an example) and MLE would be virtually identical. Indeed we see very similar noise patterns to this figure 10 when applying no constraints in MLE. We added the following to the text after equation (14):

"It is important to mention that it is only the a-priori knowledge in form of the box-constraints that makes the MLE solution distinct from the algebraic SCA solution (without zero flooring, see section 6.2.2.1 in Flamant et al. 2020), because this algebraic solution corresponds to the exact signal values in y and therefore to J_{obs}=0, which is the global minimum to the unconstrained counterpart of problem (14)."

Line 6: It will be helpful to the reader to know that ALADIN does not make cross-polariation measurements (lines 107 & 108) and therefore ALADIN is not able to make direct lidar ratio measurements.

We agree that this might be confusing for the lidar community, hence we added the remark "(the cross-polarized return signal is not measured)".

Line 8: Do the authors mean that the inversion problem is statistically ill-posed? If the inversion problem is ill-posed, then it would not be possible to infer the extinction coefficients from ALADIN measurements without using lidar ratios.

This statement was based on the work of Pornsawad et al., 2008: "But From a mathematical point of view Eq. (1) is a Volterra integral equation of the first kind, which is ill-posed." We find this wording misleading as well and thus decided to drop "ill-posed" throughout our text.

Line 33: It is unclear what the authors mean by "because high uncertainties in climate change modelling regard the indirect effect of aerosols on clouds and anthropogenic radiative forcing (Illingworth et al., 2015).". Are the authors saying there are high uncertainties of indirect effects of aerosols on clouds in climate change modelling?

That sentence was a bit bulky. The actual statement in Illingworth et al. reads: "The largest single cause of uncertainty in anthropogenic radiative forcing is from the indirect effect of aerosols on clouds." So we reformulated our text to say "[...]; the largest single cause of uncertainty in antropogenic radiative forcing has been reported to be from the indirect effect of aerosols on clouds (Illingworth ...)."

Line 41-42: Lower SNRs of what? This part of the sentence is vague.

This refers to the actual lidar signals. So the SNR of the signal. In order not to write the "signal's signal-to-noise ratio" we added "in the receive channels" for clarity.

Line 47: Do the authors mean that the inversion problem is statistically ill-posed? If the inversion problem is ill-posed, then it would not be possible to infer the extinction coefficients from ALADIN measurements without using lidar ratios.

See answer to comment on line 8.

Line 69: Fine resolution of what? Image resolution?

We mean the effective resolution in the retrieval, which is not clear. Likewise, any explanation would inflate the text here, so we kept only the statement about the higher precision.

Line 133: Do the authors mean detector random errors instead of wind random errors?

No, the primary concerns for Aeolus were the wind random errors.

Line 135: Is it unclear what is meant by Basic Repeat Cycle (BRC) or Observation. Are these terminologies that are used in the field of wind lidar? If so, would it be helpful to add a reference? Does BRC mean that every 30 consecutive vertical profiles are accumulated?

These are Aeolus mission terminologies, which are to find in the Aeolus documentation referenced in this section. But maybe it was misleading that we used "integrated"? Instead, we mean "accumulated", which should clarify.

Line 162: The reviewer presume that the measurement geometry is the so called geometric overlap function. It is unclear why the authors define O(r) which it is not used in equations 2 and 3?

We wanted to give a general formulation, but it is true that we don't use it. We added "For Aeolus the range overlap function $O(r)$ equals 1."

Line 163: The range of T(r) is defined but not for \beta(r). To be consistent it will be helpful to also define the range for \beta(r) e.g. \beta(r) > 0.

This was added.

Line 168: If the ALADIN instrument is measuring the co-polarized backscattered energy, shouldn't the molecular backscatter coefficient also be labeled as co-polarized?

To my knowledge, it is assumed here that the molecules are a non-depolarizing target, making both the same.

Line 203: What is 'this' referring to? Is 'this' referring to L_{p,sat}?

We now explicitly referenced the equations we intend to refer to.

Line 205: Did the authors mean to say "or" instead of "/"?

Yes, we corrected this.

Line 221: It is unclear what the "It" refers tp. Does the "It" refer to the second term in equation 12?

Yes, also corrected

Line 222: Should J be J_{obs}?

Here it is equivalent, but we clarified this in the text.

Line 225: In regards to S_y, refer to comment of line 224.

corrected

Line 226: Should "Lidar" be "lidar"?

corrected for all occurrences

Lines 236 to 245: The paragraph is superfluous since 1) the authors are not employing a constraint or penalty term and 2) the paragraph does not add value to the current text. If the authors at a later stage employ a constraint or penalty term and report results in a next publication, the next publication can include this paragraph. Furthermore, the conclusion does discuss employing a regularization term.

The reviewer suggests replacing this paragraph with one line saying that equation 14 is applied on each averaged vertical profile of measurements.

We acknowledge the reviewer's point of view, but are bound to report our method the way it is implemented. Since the actual implementation hands the problem (15) to the L-BFGS-B algorithm and not (14) in a for loop over single profiles, we prefer to keep the paragraph.

Line 256: The last sentence on this line is unclear. To what does "following" refer to?

We changed it to "As in SCA, this overlap is not considered in this work."

Lines 263 to 264: Are the 2.9km and 87km numbers referring to different horizontal resolutions?

Yes, exactly, see Fig. 1 for clarification (horizontal length values added).

Line 268: The citation style of Wandinger at at. 2015 is not consitent with the citation style of Illingworth et al. 2015.

corrected

Line 288: It is unclear why new sentence starts with "But".

Pardon, that is a remnant of my German. We coupled the sentences to clarify our message.

Lines 306 to 313: Large portions of sub-section 3.3 are repeated in the first paragraph of section 4. Therefore this sub-section can be removed.

Many thanks for spotting this mistake. We removed sub-section 3.3.

Line 362+382: It is unclear what the authors mean by "wither".

We meant to say “either”; corrected

Figures 2 and 3: For the curtain plots; what is the horizontal axis? Profle number? Seconds?

This is the profile number, but in Aeolus’ mission-specific vocabulary. See Instrument section: “[...] a total of 30 measurements are accumulated to one so-called Basic Repeat Cycle (BRC) or Observation (Aeolus mission terminology), equivalent to approximately 87 km along track distance on ground.”

Line 427: Can the authors elobrate on how the cross-polarized lidar ratio is transformed into BER (the actual lidar ratio)?

The cross-polarized BER is meant. This is added to the text now.

Line 434: More robust compared to what?

As compared to the reference algorithms. “than SCA and SCA MB” has been added.

Line 473 to 475: See the comments of lines 81 to 83.

We added: “The choice of dependent and independent retrieval of backscatter and extinction coefficients is a trade-off between improved precision and potential biases. A coupled retrieval may improve the precision of the retrieved backscatter coefficients, but it relies on a potentially erroneous calibration as input (geometric overlap function and cross-talk correction).”

Please also note the supplement to this comment:
https://amt.copernicus.org/preprints/amt-2021-212/amt-2021-212-AC2-supplement.zip