Interactive comment on “A practical method to remove a priori information from lidar optimal estimation method retrievals” by Ali Jalali et al.

Kgaran

johnkgaran@gmail.com

Received and published: 4 January 2019

Since it took the authors until the very last day of the journals initial deadline to respond to the questions raised in the first comment it was hoped they would have presented the necessary plots and answers. Unfortunately, we see neither and their response to some of the initial comments only raise more question on the validity of the method presented in the paper. Also, some replies were grammatically incorrect and were written in poor English. I do understand that since English may not be the first language of some of the authors of the paper they might have some difficulty in writing, but it is expected that they proofread their replies to the comments before submitting them. Due to the lack of time I couldn’t go over all the responses given by the authors, but I will try to emphasis on the importance of some the raised questions and comment on some of the authors replies which I believe are not satisfactory.

To address the first question about the cost. Cost is one of the MOST IMPORTANT parameters in OEM that is being minimimized to prove an optimum solution. Authors are arguing the cost of the first iteration is huge and at the end of the iteration process it becomes unity. In the authors approach to the coarse grid method in the paper, the normal procedure has been broken by using the a priori in the first iteration, and after that the coarse grid method was used for later iterations. Therefore, the new cost function should be discussed in detail. Questions like how many iterations are used, and what the final cost value is, should be answered. Even if not in the main body of the paper it should be added in a supporting information section. Answers such as: “However, in Sica and Haefele (2016), it was mentioned that the final cost values were close to unity. We do not see how showing the details of the costs is helpful in this scenario” is not an actual answer as the two methods are different. Furthermore, the initial averaging kernel used in this method is the out-put of the OEM first iteration which is 100% dependent on the a priori and the choice of its covariance. The authors are citing Sica and Haefele 2015 and 2016; which in both papers the Levenberg-Marquardt nonlinear iteration method was used to minimize the cost function. As the authors are not willing to show any mathematics I will prove my point by referring to Eq. 5.36 in Rodgers in which in the first iteration (xi = xa) the retrieved state is dependent on both a priori and its covariance. Hence both the cost and more importantly the averaging kernel are certainly dependent on these two terms. The averaging kernel that the authors are using to resample their final retrievals is dependent on the choice of a priori and its covariance. Based on this statement what is the mathematical proof showing that in the final retrieval the effect of a priori is minimized? A common mistake made in answering to this question which has also been made by the authors is “If a coarse grid averaging kernel has a maximum value of unity, by definition (Rogers, 2011) it is completely sensitive to the measurements and has no a priori influence.” However, this is a naïve response as this averaging kernel cannot be used for analyzing and one should transform the (coarse grid) averaging kernel back to the fine grid. I quote parts
of Section 10.2 of the same book (Rodgers, 2011/ which is misspelled by the authors) which states that “Note that in some cases no xa appears to be used. It is usually found on closer examination that in these cases xa = 0 or representation has been used in which A = I. To interpret this equation correctly, the state vector must be a full state vector on a fine grid...”. “If a reduced representation with fewer level being used, then it must be transformed back to the full state vector in order to analyse it.”

To address the response starting with “Additionally, it is unclear if the concern you have is with Von Clarmann and Grabowski (2007) ...” the authors failed to respond to the comment by redirecting the question to another paper. The question is that if an a priori is already available and is used for the first iteration why you should use the re-regularization/resampling scheme and not use the already available a priori for the remaining iterations. The concern is clearly about the current paper. I would like to remind the authors that in their method, in the first iteration they do run OEM using a priori profile and a priori covariance. I am also confused with whether the authors are suggesting that the a priori covariances used by Sica and Hafele (2015, 2016) were not well determined? Is that why the authors have developed this new method? Furthermore, the third paragraph of the response is unclear and poorly written. Could the authors explain their point again?

To address the response starting with “We did not state that these profiles are of poor quality...” to clarify myself, I am not stating that the a priori profiles used in Sica and Hafele 2015 and 2016 are of poor quality, but I am stating that such a statement cannot be made to explain the necessity of the new proposed method. There can be two explanations for the a priori covariance used in the paper. First, is that in Sica and Hafele 2015 and 2016 as well as in Jalali 2018 the a priori covariance was not well determined (which I believe it is not the case) and that is why the authors suggest this new approach using the same a priori. Second, Sica and Hafele (2015 and 2016) used proper a priori covariances which makes the new approach suggested by the authors (using the same a priori and covariance’s) not a valid case study. Why did the authors use a well-defined case instead of choosing a case in which the covariance for the a priori is not well defined? I am repeating myself here, if the a priori and its covariance are well-known and are used in the first iterations why not keep using them considering the new method gives higher uncertainties.

To address the response starting with “The 0.9 and 0.8 cutoff height were investigated by comparing the PCL temperature...”. In Jalali 2018 only the cut-off choice for the temperature retrievals was investigated. It is not correct to generalize that result also for water vapor retrievals. According to Sica and Hafele 2016, the cut-off for the water vapor retrieval is 0.8 and the authors should not change it to 0.9 without any explanation and just because this new cut-off makes their new method look better. Furthermore, referring to Table 4, in Jalali 2018, there exists a difference of 8.3 K at 100-105 Km when comparison is made between the PCL-URB and a difference of 14.2 K at the same region exists when comparison is made between PCL and CSU. These uncertainties cannot be regarded to be too high considering that these uncertainties are about the statistical uncertainty of the retrieval of the OEM alone without considering the uncertainties of CSU and URB. Furthermore, I am sincerely concerned about the authors approach to the same issue, at some point in their response to another comment which they argue that the uncertainty of up to 40% is acceptable. Quoting the response “The definition of accuracy depends on the application of the measurement. For example, WMO guidelines suggest that measurements have at least 40% accuracy in order to be used as inputs for climate models and reanalysis. In this case, with the exception of one data point, all of the measurements are below 40% accuracy and therefore would be considered accurate enough for models which are used for meteorological purposes.” It is therefore surprising that the 10% uncertainty is considered to be too high. I am again repeating myself that by putting the cutoff of the OEM at 0.8, higher altitudes will be gained, and this necessitates to compare the coarse grid method with the OEM when the OEM cutoff is set at 0.8. The main issue is that since, the OEM has considerable lower uncertainty assuming that the same altitudes are reached by putting the cutoff at 0.8 the method presented in this paper won’t hold any advantage.
to the OEM. This is in consideration that the coarse grid method in retrieving the water vapour profile provides an uncertainty of about 30% and higher.

To address the response “We believe that the answer to your previous comment addresses this one. We purposely chose a different date for this paper to avoid concerns of using the same nights in multiple papers and to illustrate that the method is applicable to any night. We would prefer not to change the night in this paper or add an additional night.” I am surprised that the authors are not willing to show their results for the night of 24 May 2012. This night is essential in comparing the method introduced in this paper with OEM introduced in Sica and Haefele (2015) since the night has been investigated in detail. Even if you prefer not to add it to the actual body of the paper you can add it in a supporting information section and it is customary to show the plot as your response instead of replying you don’t see the reason to show more nights. Since I believe the OEM is robust and fast, thus showing the results for few more nights would not take any time, but in return would be a good supplementary for this paper. To address the response starting with “Perhaps this sentence is too general…” The new paragraph is still too ambiguous. What is the valid altitude range? Can you specify? What does it mean more similar to radiosonde? Does this mean a similar trend within plus or minus some error? These statements should all be quantified and cannot be written in a scientific paper in this manner.

To address the response starting with “It would be helpful to know what your definition of …”. The authors have not addressed the raised concern. To clarify myself since the paper claims that using the new method the retrieval for day time extends for 2 km which means that the retrieval extends till about 4.5 km and the averaging kernel is 1. Since the authors have not yet shown the actual measurements by responding in another section “We have not included the measurements for daytime and night time as we did not deem them necessary and we already had many figures in the paper. If the editor deems it necessary, we would be happy to add in the figures.” My concern is that by judging from the OEM averaging kernel, above 2.7 km the SNR is too low and thus at altitudes higher than 2.7 Km the high value for the averaging kernel is artificial. Therefore, the actual measurements for daytime and night time should be shown.

To address the second paragraph of the response starting with “These statements are not incorrect or misleading.” In the day time the retrieval goes 2 km higher than the OEM but the same effect is not seen in the night time. Therefore, the response is still confusing as in both cases (day and night) the background is removed in the same way. The only difference is that in day-time the background is larger than the night time. Meaning the noise in day time is contributing in the proposed retrieval of the new method.

To address the response starting with “While it is true that the degrees of freedom of the retrieval cannot be increased”. Quoting von Clarmann and Grabowski (2007): “Since re-gridding implies further degradation of the data and thus causes additional loss of information, a re-regularization scheme has been developed which allows resampling without additional loss of information.” However, in the method presented in this new paper the re-regularization has not been implemented and yet the acceptable retrieved height has increased meaning that more information is gained. Such a claim needs mathematical proof. It is expected that the authors know the relation between the degree of freedom, the information content of the retrieval, and the relation of these to the accepted height of retrieval. Thus, the authors clearly failed to address the comment scientifically. The terms such as the original retrieval altitude and the final retrieval height that are used are not introduced in the paper and neither are they defined in the response. Moreover, referring to the mentioned figures are rather more confusing. Besides, quoting a section of the response: “While it is true that the degree of freedom of the retrieval cannot be increased, the accuracy of your retrieval can increase when it is compared to another instrument and the a priori information is removed.” Yet another claim with no proof. How come does the accuracy increase? To address the second paragraph of the response starting with “We would like to point out that the major advantage…”. What is the “physical meaning of the difference” can
you please explain? Furthermore, the large uncertainty that is seen in the resampling method is because of the error in the extrapolation term and it cannot be claimed that the uncertainty is more conservative.

To address the response starting with “The examples used in this paper . . .”. It seems that the introduced method is not consistent with everyday lidar measurements and thus the method is not robust. Quoting from the response of the authors “The example used in this paper was used to contrast the daytime result and show a case when the coarse grid method does not work as effectively”. So out of the only two case studies provided for water vapor retrievals only one works. I am really confused that under what conditions this method provides the advantages which are claimed by the authors. Using terms such as low SNR and high SNR will not provide any useful information. Moreover, the authors are implementing that they have a new method that might or might not work, because in their own words quoting here “Using the a priori removal technique may be helpful and did not say definitively whether it would not” and “This meant to be future work . . .”. Thus, if the authors are not convinced in what conditions this method is useful this paper should be a work in progress. To validate this whole new method, it is necessary to examine the conditions in which this method is working, and it cannot be referred to as future work.

To address the response starting with “It is true that the dependence on the a priori should remove the entire”. I fail to follow how the nonlinearity can have any effect. Such a claim should be followed with related mathematical proof statements. Furthermore, replying “We have tested this scenario and differences between the two retrieved profiles still remain if Sa^-1 is set to 0 in the first run. Therefore, we are confident that the difference of zero is a result of the a priori removal method” is not a satisfactory answer. For making such a statement authors should either provide their codes or the mathematical proof showing such a thing. Replying that you have done it without showing anything for it is not an answer.

Thanks, John Kgaran

C7