Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., author comment AC1
https://doi.org/10.5194/acp-2021-549-AC1, 2021

Reply on RC1

Michael Sitwell et al.

Author comment on "An ensemble-variational inversion system for the estimation of ammonia emissions using CrIS satellite ammonia retrievals" by Michael Sitwell et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2021-549-AC1, 2021

Thank you for your comments. Reviewer comments are in italics.

Major comments

1) Evaluation

Mean bias evaluation can be misleading without also evaluating absolute error (ME or RMSE) due to the possibility of positive and negative biases canceling each other

This concern is already addressed in the paper as the standard deviation of differences were computed for these statistics, displayed in Tables S1-S5, and commented on lines 440-445, 553-556, and 664-665 of the paper. As discussed at the end of Section 2.2, the bias, standard deviation of differences, and correlation coefficients were computed for all data sets, all of which are displayed in Tables S1-S5. Any cancelling of errors will be reflected in the standard deviation of differences. Note that the RMSE can easily be computed by the reader by adding the NMB and NSTD in quadrature and then taking the square root. I had included the RMSE in Tables S1-S5 in an earlier draft of the paper, but decided to remove it because the tables were too big to fit in the page, which were already in landscape. Since the RMSE is redundant information if the bias and standard deviation are already given, I decided to remove the RMSE from the tables (although it would have been nice to display).

Throughout the results, total NMB (e.g. Figures 7, 11, 13, 14,16) from all sites are used to demonstrate the impacts of the updated ammonia emissions using the inversion approach… RMSE should be presented in the paper with NMB figures (7, 11, 13, 14, 16).

The reason more emphasis is given to the NMB as compared to NSTD (or RMSE, as well as the correlation coefficient) is that the changes in the NSTD were statistically insignificant for all cases examined, with only one exception (comparison with the log-space operator in June for AMoN). All differences between the NSTD of original and updated hybrid cases were statistically insignificant, which can be seen by looking at the 'sig' column of the NSTD in Tables S1-S5. This was discussed on lines 440-445, 553-556, and 664-665 of the paper. For this reason, I chose not to include the NSTD in Figures 7, 11, 13, 14, and 16, with the thought that the descriptions in the lines referenced above would be sufficient considering the results.

If RMSE is plotted, the differences between in the RMSE are a mixture between the statistically significant differences in the biases and the statistically insignificant differences in the NSTD. This is why NMB and NSTD were displayed separately. Although displaying the RMSE can be nice, as it yields a single number for comparison, it is redundant with the bias and NSTD taken together. As including plots of either the RMSE or NSTD would greatly increase the number of plots in the paper, given that the changes in NSTD were statistically insignificant and already described in the text, we chose not to include these plots.

With comments made on lines 440-445, 553-556, and 664-665 directing the reader to the NSTD results, it should be clear that taking the NMB and NSTD results together constitutes an evaluation of the absolute error and that the emphasis given to the NMB is simply due to the statistical significance of these results.

2) Sensitivity of constants

Before discussing this point, I'd like to address a misreading of Figs. S3, S5, S6, S7, and S8.

For instance, based on Figure S3, sites in the western central U.S. (around Colorado) tend to have worse performance

…

Why do all RMSE (updated) / RMSE (original) ratio figures (S3, S5…, S8) in supplement have negative values?

Note that the right columns in Figs. S3, S5, S6, S7, and S8 are labeled as '1 – RMSE(updated)/RMSE(original)'. I think the '1 - ' might have been missed. So the sites around Colorado show better performance, not worse. I also assume that in the comment made on evaluation that

RMSE is presented in the supplement figures (Figures S3, S5, S6, S7, S8) and it seems that many sites show worse performance.

that the same mistake has been made. I'm not sure how much correcting this misreading will change these comments, but I will try to respond to these comments the best I can given the situation. I have reformatted this title to try to make it more legible and fit better within the column. The new version of Fig. S3 is attached.

Sensitivity analysis on the constants used in the hybrid approach seems to be important and useful for the "ideal" constant selection.

In this study, the results do not appear to be very sensitive to the chosen parameters for the hybrid method. I tried lowering the value of X_min by a factor of 10, which only changed about 0.02% of retrieval comparisons for May-August 2016. The locations of these retrievals were also reasonably spread out over the model domain, so this change is unlikely to have much of an influence in the inversion at any location. I tried lowering X_min by another factor of 10, which showed almost the same differences of 0.02%. I only tried lowering X_min since increasing it, say by a factor of 10, would start to label some non-negligible profiles as negligible, which is not desirable. When lowering c_min by a factor of 10 (keeping X_min at its original value), 0.7% of retrieval comparisons change, again spread out over the model domain. Although this change effects more profiles, it is still a small number of retrievals, and not localized in any particular location. I have added this text at the end of Section 3.3:

"For the time period and locations examined in this study, the hybrid comparison method does not appear to be particularly sensitive to the values chosen for X_min and c_min for values smaller than those chosen here. Reducing X_min and c_min by an order of magnitude only changes the operator selected for less than 1% of retrieval/model pairs, which were spread out throughout the model domain. While reducing the values of X_min and c_min yielded little difference in the retrieval-to-model comparison, selecting significantly higher values for these parameters would result in classifying some non-negligible profiles as negligible, and so must be done with caution."

Maybe the values used for linear-log cutoff should be variable spatially or temporally depending on the ground sources.

The parameters used in the hybrid method are used to detect model profiles with non-negligible amounts of ammonia that have been 'zeroed out' by the log-space averaging kernel. As such, the method's parameters X_min and c_min, are used to define a minimum non-negligible profile. I'm not quite sure what the motivation would be to have these parameters vary in space or time given their physical interpretation. Having them varying in time or space would imply that what you consider to be the minimum non-negligible profile varies in space or time as well, and am not quite sure why this would be a desirable property. However, since the hybrid method does not seem to be very sensitive to the chosen values for X_min and c_min to begin with, this might be a moot point.

Minor comments

1) *what does "The number of degrees of freedom for this retrieval is 0.956" mean*

I'm assuming this question means "What does 'degree of freedom' of a retrieval mean"? The degrees of freedom for a signal is a very frequently used diagnostic quantity of a retrieval, which is the number of independent pieces of information that could be measured in the retrieval process.

2) May to August 2016 study. Since this approach is developed for the GEM-MACH air quality forecasting model, probably it is important to evaluate how this approach performs in other seasons with cooler temperature and low ammonia emissions as well

We agree. For this initial study, demonstrating the proof-of-concept for the NH3 inversion method, as well as the model-to-retrieval comparison method, we focused on the warmer months across North America as these conditions are more favourable for infrared satellite ammonia retrievals (higher concentrations of ammonia and greater thermal contrast between the surface and the atmosphere). More evaluations are planned for the future that will cover the whole year, including the cooler seasons with less ammonia emissions.

3) In reality, fires exist, and fire emissions are included in forecasting. Is this approach appropriate if weekly updates are applied for emissions under fire conditions?

How fires are handled depends on context. For instance, if the inversions are going to be used to update emissions to be used for a different year, then if a fire significantly impacts the inversion, then fires from one year may effect the prescribed emissions in a different year, which may not be desirable. If instead the inversions are only being used for the same time period, then having the fire emissions significantly influence the inversion could be desirable. I have added these lines in Section 2.1:

"… due to forest fires with other emission sources. While we seek to minimize the effect of forest fires on the emissions inversions in this work, in other contexts this might not be necessary or desirable. For example, if the emissions are only used for the time period

when the fire occurred, having the fires affect the inversion may be advantageous."

4) What is magnitude of the issue related to the non-detection of ammonia discovered on the quality of CrIS data which affects non-source regions in the domain?

As the focus of this study is on source regions, this issue related to non-detects does not have a significant effect on the CrIS retrievals used in this study. However, at high northern latitudes far away from significant sources, this becomes an important issue. So the impact of this on the inversions performed for this work is small, but could be an important issue if instead we focused on remote non-source regions. As the newer version of the CrIS NH3 retrievals included non-detects, this allows for the possibility of focussing more on remote non-source regions for future work.

5)70–85% of the retrievals used in the inversions coming from daytime retrievals. What causes nighttime retrievals to have low quality?

During the development of the version of the CrIS NH3 retrieval product used, it was found that performing retrievals over areas with temperature inversions near the surface was challenging. This could be in part due to not having adequate a priori profiles for these situations. For this reason, a quality flag was added to filter these retrievals. As these situations occur more frequently during the night, this quality flag removed a large fraction of the nighttime retrievals, while removing a much smaller fraction of the daytime retrievals. Also, as the NH3 signal is generally higher during the daytime, as non-detects were not included in this product version, more of the retained retrievals were for the daytime.

6) Figure S1: hard to tell difference among 0 to 50 color scale in plots.

I have rescaled the colour bar from 0 to 100 so that the 0 to 50 portion is easier to read. I tried out different colour maps to see if it made it easier to read, but didn't find that they improved the readability much. I also tried a log scale, but it then made the higher end of the colour scale harder to read. New figure is attached.

7) Figure 9a – NH3 value higher than graph horizontal range.

The x-axis range has been extended. New figure is attached.

8) Why do all RMSE (updated) / RMSE (original) ratio figures (S3, S5…, S8) in supplement have negative values?

Response to this comment is in (2) of the major comments section

9) The inclusion of the critical load exceedances seems to be out of the focus for this paper although updating ammonia emissions affects N deposition. Given the purpose of the paper and more thorough evaluation needed for this new approach, it is recommended to remove the critical load results from the paper

The inclusion of the deposition of NHx is within the general overall focus of the paper as the inversion can have a significant effect on NHx deposition (as mentioned by the reviewer). Including the critical load provides context to the deposition results by relating it to the health of an ecosystem. That being said, I recognize that the paper is on the long side. As a middle-ground approach, would the reviewer be satisfied if the two paragraphs describing the critical load model were moved to an appendix? I could also move Fig. 17 to the Supplement, but would at least like to reference the critical load in the main text of the paper.

Please also note the supplement to this comment:
https://acp.copernicus.org/preprints/acp-2021-549/acp-2021-549-AC1-supplement.pdf