

Reviewer comments are in italics.

1. *Line 83 – “using version 1.5 of the retrieval algorithm”, reference on this version?*

Citation to Shephard et al. 2020 added here.

2. *Line 218 – “apply it an emissions inversion” to “apply it to an emissions inversion”*

Changed.

3. *Line 292 – “temporal profiles as described in Section 2.3”: do you mean that you used the temporal profiles from SMOKE to allocate the monthly disturbed emissions for GEM. For an ensemble of 100 members (Line 242) , I assume that you run SMOKE and GEM-MACH for 100 times at ~10km resolution, right? I am wondering whether resources and time required for the 100 times SMOKE and GEM-MATCH simulations impose hurdles in the operational air quality modeling or forecast, particularly when you try to shorten the time lag and reduce inversion time window from a month to 1-2 weeks as mentioned in the conclusion.*

Here we mean that the intra-month temporal distribution of the emissions are the same for all emissions sets. These temporal profiles are set via the input profiles to SMOKE and the processing done by SMOKE, but since the same profiles are used for every emissions set, SMOKE only has to be run once. In practice, after the original set of emissions was created by SMOKE, temporal profiles for each grid point for each month were computed. When a new monthly mean emission set was computed, these pre-computed profiles were used to create the hourly emission files. The application of the pre-computed profiles takes relatively little computing resources.

4. *Line 359 – “In summary, the CrIS retrieval is compared to the GEM-MACH model by computing the difference between the total column”. For ground level emission perturbation, it affects the surface (or ground) level (highest pressure level) NH3 concentration most (Lines 380 – 381) . I would think comparing surface level GEM-MACH to the surface level CrIS makes more sense. Thus, you can avoid upper level concentrations resulting from transport. Do you have any comments on the reason for choosing the whole column NH3 for comparison instead of the surface NH3?*

Comparing only the surface level concentrations would likely give reasonable results. However, although many retrievals have their highest sensitivity at the surface, this is not true for all retrievals. Some retrievals have averaging kernels that peak much higher above the surface (most peak by 700 hPa, but a few peak even higher above the surface). So using these retrievals in an inversion that only compares the surface values would not be optimal. One could simply not use these retrievals in the inversion, but we wanted to maximize the information used in the inversions. For retrievals that do peak at or near the surface, the averaging kernels dampen the influence of levels farther above the surface in the inversion, so I wouldn't expect much of a difference in effect that these retrievals have on the inversion if only the surface was used instead.

5. *For the Fig 6 plots on the right, I am suggesting that % of sites with positive value (percent of improved*

sites – reduced RMSE) is placed on the lower right corner of each plot (like Fig 4 plots) as it is hard to tell overall performance based on 1- RMSE/RMSE plots now.

Done.

6. Lines 397-398 – “displayed in Figure 7, which shows the monthly mean column values within 0.5. x0.5 longitude/latitude bins ...” Does this mean that the ammonia emissions inversions are conducted at 0.5X0.5 long/lat bin resolution? If yes, I do not recall this information is stated explicitly in the paper. Does the 0.5 bin selected have something to do with 40km mentioned in Line 297? The GEM-MACH simulation is at ~10km resolution and CrIS resolution is ~14k at nadir. If inversion is conducted at 0.5 bin, what does it mean to emissions of cells within 0.5 bin for GEM-MACH?

The 0.5 x 0.5 longitude/latitude bins were used here for display purposes only and not used anywhere in the inversion. Looking back at this plot, I can see that the choice of this binning for this plot may have been confusing. I have changed this figure so that the data is binned using the GEM-MACH horizontal grid instead. There were a few other plots that used the 0.5 x 0.5 longitude/latitude bins, which I have changed as well to reduce the potential of confusion.

7. Results from 4.1, 4.2, 4.3, 4.4 subsections seem to be all mixed together, such as that figures in 4.1 show results for 4.3. Here are the suggested sub-sections with deposition at the end instead of being between the two PM subsections:

4.1 -- Emissions Inversions: discuss results for fig 7 and 3. I don't think 4.4 should be in the result section because it is related to your emission inversion approach development described in Section 3. If you want to demonstrate the impacts of the selections on emissions, it is probably better to have it in this sub-section.

4.2 - Effect of Inversions on Model Ammonia: Combine your original 4.1 Model Ammonia Performance Without Inversion into this subsection to reduce redundance. In analyzing the effect of the inversion, the model results with/without the inversions are compared and evaluated.

4.3 - Impacts on PM Formation

4.4 - Impact on PM Size Distribution

4.5 - Impact on Deposition

The results section was reorganized to a structure similar to that suggested.