

Responses to the Manuscript ACP-2022-216: Estimating global ammonia (NH₃) emissions based on IASI observations from 2008 to 2018

Dear Editor-in-Chief:

We hereby submit the revised version of our manuscript (ACP-2022-216).

We greatly appreciate you and referee #2 for providing highly insightful and constructive suggestions, which have substantially improved the clarity of our manuscript. We have carefully addressed all these comments, please see below our point-to-point **responses in blue** and **red** text and refer to the **revised manuscript**.

We hope you find our manuscript suitable for publication and look forward to hearing from you

Yours sincerely, Zhenqi Luo, Yuzhong Zhang, Wei Chen, Martin van Damme, Pierre-François Coheur, Lieven Clarisse

Email: zhangyuzhong@westlake.edu.cn, zl725@cornell.edu

1. RESPONSES TO REFEREE, REFEREE #2

Comment #1: In regards to the comparison with ground-based observations:

- a. Could you add a line or two to very briefly describe each observation data set (e.g. instrument type, active or passive sampler, integration time, observation frequency, etc...)?

We add the text in **Line 345-352**.

- b. For the comparison of the RB and RMSE stats between BUE1 and TDE, could you quantify the statistical significance of these differences? Without uncertainty estimates on these stats, it's hard to tell if, for example, the difference in FB values for JFM are meaningful or not.

We add the Mann-Whitney U test for the comparison between surface measurements with simulation driven by BUE1 and TDE (**Line 363**).

- c. On lines 342-347, you discuss the limited impact of the top-down estimate on the comparison with the surface observations and say one possible explanation is 'systematic differences between satellite and surface measurements'. Could the weighting in Eq. (1) and/or (3) also be a cause of this? i.e. If you increase the weighting of the IASI data would the FB values

move closer to zero (at least for JFM and AMJ where TDE shows a slight improvement in FB over BUE1)? If this is not the case and the top-down emissions are being moved to mean values away from those near the surface observations because there is an underlying bias between the surface observations and IASI, could you quantify the magnitude of this bias?

Our consistency evaluation (Fig. S5) shows that simulations driven by either BUE1 or TDE show relatively small fractional biases (<0.5) in regions where we have these surface observations (i.e., North America, Europe, and south-eastern Asia), suggesting that our estimates are already in line with the IASI data. It is unlikely that further pushing the estimates towards IASI data can greatly improve agreement with surface observations.

However, we acknowledge that there may be reasons other than systematic observation biases (or inconsistency as total column and surface measurements are not directly comparable) to explain this discrepancy. In general, it is not straightforward to evaluate whether total column and surface measurements are consistent. We now remove the sentence to avoid the impression that we are sure this being the major reason. This information was added into the section labeled 'Uncertainty evaluation'. I'm not sure this title describes this information. Probably better to put this information into its own subsection.

We add a new subsection titled "**Comparison to independent surface networks**" (**Line 344**)

Comment #2: Lines 129-134: Thank you for adding this information. However, I think some of these sentences need to be clarified.

- a. 'We note that the ANNI-NH3-v3R retrieval does not provide averaging kernels (Whitburn et al., 2016; Van Damme et al., 2021).' Looking at Whitburn et al. 2016, the neural network based retrieval algorithm does not use an averaging kernel to make the retrieval. Is this correct? If so, from the wording of this sentence it sounds like averaging kernels were used but not included in the data product. Could you clarify this point?
- b. Sentences in lines 130 and 131 begin with 'However' and 'Besides', which link these sentences together. But I'm a bit confused why these sentences are being linked together based on their content. Could you reword or further clarify this?

We have revised the sentences to avoid the confusion (**Line 129-132**).

Comment #3&4:

- a. Line 325: You write on this line 'ranges of perturbation tests divided by their averages'. Are these ranges then divided by 2? i.e. $(\text{upper-lower})/2$?

We revised the relevant sentences to improve clarity (**Line 327**).

- b. Fig. 7: It might be good to add a plot of the absolute uncertainty (units of $Tg\ a^{-1}$ per area or similar) so you can better discern what's going on in regions with relative uncertainty bigger than 100%. If these regions all have small absolute uncertainty, then we can easily see that the large relative uncertainties were caused by small average values. But there are quite a few regions with the relative uncertainty larger than 100%, so it would be good to know if this is the case with all of these regions or if there are regions with non-negligible average values and large uncertainties.

Thank you for your suggestion. We add a panel in **Fig.7** to show absolute uncertainty derived from the perturbation tests and briefly discuss the results in the main text.

Comment #5: Thanks for adding the sensitivity tests for the SO₂ trend values. I see the description in the lines you mentioned in your reply, but I can't see where the results of this sensitivity test is. I might have just missed this in the text, in which case could you direct me to these lines, and if not could you add some text about these results.

Thanks for pointing this out. We now report the results of this sensitivity test in **Line 286-287**.