

Reply on RC1

Mark T. Richardson et al.

Author comment on "Boundary layer water vapour statistics from high-spatial-resolution spaceborne imaging spectroscopy" by Mark T. Richardson et al., Atmos. Meas. Tech. Discuss., <https://doi.org/10.5194/amt-2021-89-AC1>, 2021

We have bolded our comments and left the reviewer's in standard text.

I congratulate the authors for a good study and a good statistical work.

Thanks for your generous comments and thoughtful review. We also apologise for not responding during the open review period to allow time for a continuous discussion – we simply weren't able to do everything in time. We believe we have addressed all of your comments, however.

I have some minor comments:

* general:

- I miss a lack of comparison with real data. The spatial uncertainties resulting from this study could have been compared with real data under similar conditions as the simulations (e.g. flat-terrain, etc..)

We have added extra text: the abstract now introduces this as an observing system simulation experiment, the introduction makes it clear that this is primarily a model sensitivity study and Section 3.3 adds several paragraphs with comparison numbers but also explains why we can't yet observationally evaluate our σ_x estimates.

Section 3.3 shows our results are consistent with RMSE reported for satellite-surface station comparisons for other studies. We cite OCO-2, MODIS, MERIS, OLCI and Sentinel-2 as examples. We also report some AVIRIS-NG flight statistics, but mention that those also aren't directly comparable because it's a different instrument and those are flights over sunlit ocean.

We also now describe in Section 3.3 and 5 how future measurements, perhaps from airborne campaigns with collocated data, could allow us to do better evaluations of our target statistics, and most importantly σ_x .

We could have grabbed some TCWV retrieval fields from other sensors, but their retrievals might be very different, they might have different sensitivities to surface type etc, and that might require further detailed investigation to

understand. The AVIRIS-NG numbers are the most consistent with ours, so we picked those as our example.

- In such detailed statistical work I miss also an estimation of the order of magnitude of the contribution of the different approximations in the model to the final results.

We struggled to interpret this comment but we think that the current manuscript addresses it. We explain that we can remove random error and identify the main source of error in σ_x as likely due to retrieval assumptions about the atmospheric profile.

Some of this is related to the above point: we lack independent validation data.

- **We cannot include some errors, like spectroscopy or quantify RT assumption errors with this setup. This would require independent measurements from VSWIR + another sensor of the same fields at ~ 50 m resolution..**
- **We have added a prior sensitivity test, showing <0.15 mm mean shift with no effect on σ_x (Section 3.1.1 new text, Supplementary Figure 3).**
- **We show that random retrieval error can be removed almost perfectly (Figure 7)**
- **We say in Figure 7 discussion and later that the biggest uncertainty source relevant for our s_x is the gradient $dTCWV_{ret}/dTCWV$, which we propose may be driven by errors in the atmospheric profile.**

Re-reading, we think that for our main results (i.e. spatial statistics) we make it clear that the gradient a_1 parameter is the most important. A relevant updated Section 5 sentence reads:

"In Isofit, the atmospheric component contributes a bias to $dTCWV_{ret}/dTCWV$ and may be the largest source of our errors in σ_x , which range from -7 % to +34 % of true σ_x ."

* page 1:

- "... upcoming missions such as the Earth Surface Mineral Dust Source Investigation (EMIT) will offer unprecedented horizontal resolutions of order 30 - 80 m..." -> currently there are several missions (VNIR - VSWIR) with this or even better resolution (e.g. multispectral Sentinel-2 (20 m) and hyperspectral DESIS) with bands in the water vapor absorption regions. Actually, wouldn't it be more sound to make this study with the Sentinel-2 20m resolution?.

We now realise that our submitted intro was confusing since we mixed up discussion of our hyperspectral work with later reference to multispectral results from MERIS. We have fixed this by rephrasing throughout; we now use phrasing similar to "modern and upcoming" in the abstract and all sections. We have added a paragraph to Section 1 discussing PRISMA, DESIS and Sentinel-2.

We also added the following justification to the intro and hope this efficiently justifies our choice:

"The purpose of this is a detailed sensitivity study using retrieval code and tools already developed for EMIT. We consider $\Delta x \geq 40$ m since this is appropriate for EMIT and several LES cases in our archive that were run at that resolution."

We didn't make a big deal about it, but Figure 9 shows the $\Delta x=20$ m results for the two simulations that were run at that resolution. We argue that the use of a retrieval code that will be used for sensors at this resolution + the availability of LES simulations justifies our choice here. Finer resolution work would of course be welcome & interesting to us!

The study would also profit of the large amount of real data, which leads me to the first general comment.

We agree, obviously! Please see changes to Section 3.3 and argument as to why the available data are not directly comparable for our results.

* page 7:

- is a plane-parallel atmosphere still a good approximation for SZA = 45?. The effect in σ_{\max} seems to be of the order of 0.025 (figure 12), the same as the difference between 50 - 300 m resolution. This could be included in the second general comment.

This is a tricky point – we only have plane-parallel RT output.

We don't think the Figure 12 results should be very sensitive to this though, since this atmosphere isn't strongly scattering except for potential influence from cloud 3D radiative effects. This is a really complex problem that needs different (expensive!) radiative transfer tools to address. We have discussed doing this if future time and funding permits, but it would really need a whole additional study.

Successful error budget closure in past AVIRIS Isofit work gives us some confidence that non-plane-parallel effects aren't too important, but we have added text to draw some extra attention to this, e.g. Section 3.1 and Section 5 (added text in *italic*):

Section 3.1 : "We first remind readers that "retrieval error" here only includes errors present in these synthetic retrievals, and excludes several real-world sources, such as how the true atmosphere is not plane-parallel as assume in our radiative transfer"

Section 5: "Future work could address uncertainties that are ignored here, such as topography or cloud 3-D radiative effects *via 3-D radiative transfer simulations which avoid several of our assumptions, such as a plane parallel atmosphere*"

- I did not find the Supplementary figures

Apologies for the confusion, these were uploaded as a separate file at

<https://amt.copernicus.org/preprints/amt-2021-89/amt-2021-89-supplement.pdf>

Our review response files will include an updated SI to include a new table on SZA change effects and on a TCWV prior sensitivity effect.

* Page 10:

- The difference between different surface brightness seems to be of the same order of magnitude of the dispersion within the same surface brightness. And there is a much

larger offset between the retrieved TCWV and the true one. The offset seems to be smaller for for brighter surfaces, but there is still 1mm difference for 50% retrieved surface reflectance.

This is one of many results we clipped to stop the paper from getting even longer. In hindsight it deserves comment so we have added the following text to the Section 3.2.1 paragraphs discussing Figure 4:

“Regardless of the surface, a bias of order ~ 1 mm remains, which is similar to the largest difference introduced by surface type and may be related to other retrieval errors such as inappropriate atmospheric profile shapes assumed in the LUT. However, the derived spatial statistics we are interested in here are not affected by any mean bias.”

This is another handy reminder that we are interested in s_x , not absolute biases here, and so feeds into our new Section 3.3 discussion on observational comparisons.

* Page 15:

- which is the typical error of Isofit with respect to WV true measurements?

The AERONET results from Thompson et al. (2021) are now mentioned in the new Section 3.3.

- which depend the TCWV variability of -7% to 34% of?. E.g. Is it a function of the true TCWV?

We have added “of true σ_x ”.

This comes from our gradient calculation. Running through the maths it’s not dependent on true TCWV but simple scales the standard deviation.

- I have missed some conclusions for smaller spatial resolutions sensors.

We are not sure how to interpret this comment. We *think* we have addressed it with our added text.

Section 1: introduces Sentinel-2 and explains why we’re not doing 20 m.

Section 3.3: discusses errors

Section 5: Added text:

“Finally, this work could be extended to other sensors, such as MSI on Sentinel-2, which is not hyperspectral but provides exceptionally fine Δx of approximately 20 m. Additional high-resolution analysis may be required for this, since Figure 9(a,b) imply increases in retrieved σ_x at $\Delta x=20$ m for the two simulations that were run at that resolution.”

If some of my comments are already explained somewhere in the text, I would thank the authors to point me to the section containing the explanations.

Unfortunately I did not have the chance to read the article in a row and I might have missed some of the explanations to my comments.

Once again, we appreciate your efforts for this review and understand going

through this paper piecemeal. We had to address a *lot* of potential issues since this is intended to lay the groundwork for a lot of future analysis.

ADDITIONAL COMMENTS TO BOTH REVIEWERS

Please note that we found and corrected some figure, table and notation errors, these do not affect results.

Figure 5 – removed y ticklabels from (c,f). Added axis labels to edge subplots.

Figure 9 – x-axis label added to panel (c).

Figure 11 – axis labels added.

Table 1 – we accidentally included an old version calculated from bugged code, and have now corrected rows (vii)–(ix). The LES vertical grid definitions differ between simulations (either Arakawa A-grid or C-grid). The biggest issue was that for C-grids our early code smeared upper-LES qv into the reanalysis layer above, resulting in way too much TCWV. This was fixed before submission except in Table 1. For example, RICO TCWV was originally 49.6–49.7 mm in Table 1, but the new ~ 37 mm agrees with Figure 8. The BOMEX LES only includes ~ 86 % of the total TCWV so P4L17 has been changed from “...show that the LES capture >90 % of total TCWV...” to “...show that the LES capture >85 % of total TCWV...”.

Equation (4) had a mix of r and Δr we changed to consistent use of Δr to emphasise it's a gap between points.

We also had feedback that the error estimation in Sec. 3.1.4 could be confusing. To our knowledge it is new, so we have rephrased some text in Section 3.1.4 and referred to new Supplementary Figures 9–11. These figures show a step-by-step guide to our method, and compare results with a standard spatial filtering method.