Thank you for a rigorous and challenging review. We disagree with rejecting our paper and respond now to allow discussion and reviewer guidance before the AMTD deadline, in the hope that we can persuade you that there have been technical misunderstandings and that our work is robust and valuable. We see how our current structure causes these misunderstandings and will modify the paper in the meantime.

At the beginning of this project we agreed with your concerns regarding the use of an emulator, and we agree that it is "not justified by this study". To justify it, we performed an OSSE as you describe it, using radiative transfer forward modelling and an optimal-estimation inverse retrieval with realistic instrument uncertainty. These results are described in the published companion study Richardson et al. (2021, https://doi.org/10.5194/amt-14-5555-2021). That paper was very, very dense (26 figures inc. supplementary) and we aimed to build on this work with a briefer summary of a result we found very promising for a practical new atmospheric measurement technique.

Your review has made us realise we erred too far on the side of brevity, and simply referencing the prior study did not provide the necessary context. A more explicit title would be something like "Boundary layer water vapour statistics from high-spatial-resolution spaceborne imaging spectroscopy Part II: a new sampling strategy to remove biases in sub-km vapour scaling statistics introduced due to the horizontal component of the sunlight path through a horizontally varying water vapour field." But that's ridiculously long! We are open to a better title suggestion but for now intend to add paper content that better describes the title's meaning. We will explicitly link to the OSSE results on which our emulator is based (perhaps with an example figure), better explain terminology and illustrate the solar-smearing bias.

Methodological details
We went back and forth about terminology. All VSWIR retrievals obtain path-integrated water vapour *IWV*, but literature phrasing varies and our target physical variable is *TCWV*, referring specifically to the column above a surface footprint. Despite this, much of the literature uses *TCWV* to refer to *IWV* retrievals from VSWIR instruments, e.g. for MERIS[1], OLCI[2], MODIS[3] or TROPOMI[4].

Ultimately we selected *TCWV* with $TCWV_{ret}$ representing the retrieved value and differences from TCWV include those due to typical errors (e.g. retrieval error) and representation differences (e.g. the fact that it is actually path integrated water).

This is very important and we were not clear in the submission, we will address that with specific text.

Next, we must justify the linear emulator. We are also changing the text to do so, but provide an extended discussion here to hopefully allay the reviewer's legitimate concerns. It is worth

---

[1] https://doi.org/10.5194/amt-5-631-2012 - "1D-Var retrieval of daytime total columnar water vapour from MERIS measurements"
[2] https://doi.org/10.3390/rs13050932 - "Retrieval of Daytime Total Column Water Vapour from OLCI Measurements over Land Surfaces"
[3] https://amt.copernicus.org/articles/8/823/2015/ - Retrieval of daytime total columnar water vapour from MODIS measurements over land surfaces
[4] https://doi.org/10.5194/amt-13-2751-2020 - Total column water vapour retrieval from S-5P/TROPOMI in the visible blue spectral range

noting that our key result – the use of solar-relative geometry to calculate this spatial statistic of TCWV – is not dependent on using ISOFIT, in principle many retrievals may behave similarly. Minor differences in retrieval performance or emulator fidelity should not alter the fundamental idea that structure function exponents are best calculated perpendicular to the slanted sun-ground path.

That said, we will modify the text to include discussion of previous work justifying the emulator. Firstly, the Richardson et al. (2021) OSSE, on which our emulator is based, used the MODTRAN6 (forward), and ISOFIT (inverse) models so followed the reviewer's standard OSSE description.

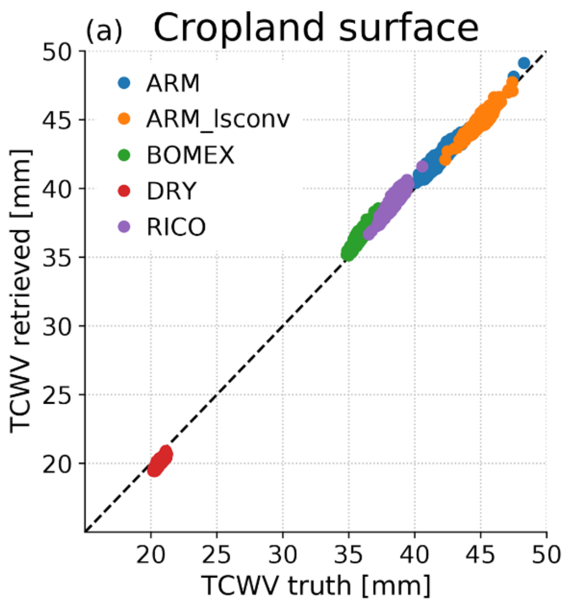
Forward and inverse simulations were done for a range of surfaces, aerosol optical depths, and solar zenith angles across each of the LES snapshots. $TCWV_{ret}$ was well-predicted by a linear fit to the forward-model TCWV.

Several considerations drove us to an emulator rather than a full retrieval. Our radiative transfer has to adequately represent the spectral detail of water vapour absorption across >200 channels from λ=380—2500 nm. We also need sufficient atmospheric layers to capture the vertical structure of absorption line broadening. MODTRAN6 can meet these requirements, but for the 11.8 billion combinations of footprint-SZA-surface-aerosol we originally targeted, serial processing time is many millennia. Our requirement for detailed absorption band spectroscopy makes the computations more expensive than some others that use less detailed spectral/absorption information. Even after turning off outputs including the atmospheric correction data, MODTRAN writes out enough that parallel processing hit a write bottleneck, keeping *parallel* processing time in the millennium range with our resources. Fortunately, we have demonstrated that our optimal estimation retrieval's TCWV performance can be captured very well by a linear fit, which was developed from a subset of footprint combinations and is what we use in this paper. Since the scientific results only care about the level of accuracy of the methods, and our emulator does a good job of representing all of the issues that are relevant for this structure-function analysis, we choose to use it.

Our argument begins with saying that the along-path water vapour is $q(z_{path})$, which is distinguished from the real vertical profile $q(z)$.

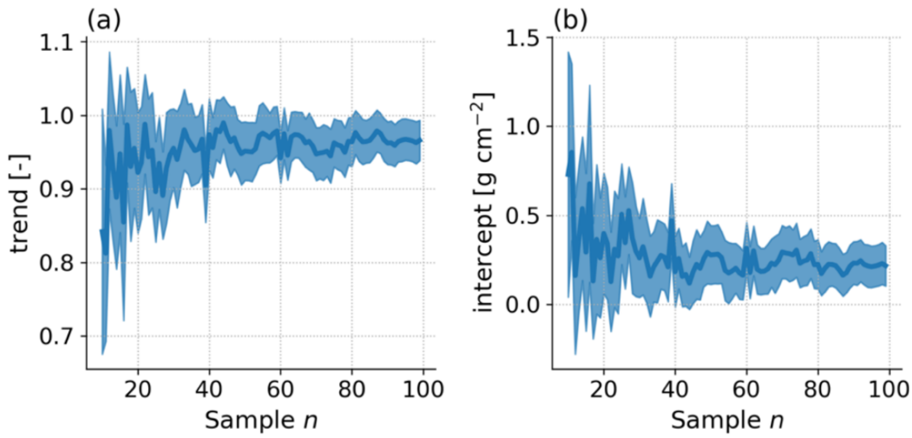
MODTRAN is plane-parallel so assumes the same atmosphere on the downward and upward paths, albeit the light passes through each at different angles. We found that *regardless of the LES water vapour profile q(z)*, and regardless of the $q(z_{path})$ that implied for the plane parallel simulations, the relationship between the true TCWV and $TCWV_{ret}$ was captured by a linear emulator. If this is true for the arbitrary set of profiles we included where differences between $q(z_{path})$ and $q(z)$ are simply related to SZA, it follow that $q(z_{path})$-$q(z)$ differences due to horizontal variability shouldn't destroy the existence of a linear relationship between TCWV and $TCWV_{ret}$ either. We anticipate that they will affect the *parameters* that describe that relationship, but our structure-function results are immune to changes in a_2 (intercept doesn't change variance) and ϵ (we reliably identify and subtract it, Figure 3a). The slope parameter a_1 is important, and it is the largest error source in estimating spatial standard deviation (a key result from Richardson et al., 2021), but our structure function parameters are actually quite insensitive to realistic values of a_1 (Figure 3b).

Richardson et al. (2021) results to support the linear emulator now follow. The first question is: how are realistic profiles of vapour with "TCWV truth" related to retrieved $TCWV_{ret}$? Linearly:

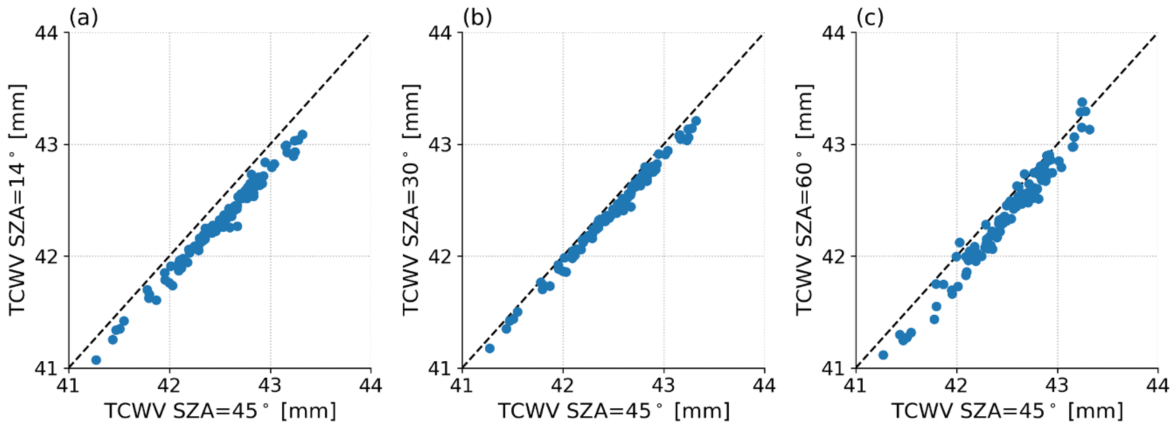


Each of the colours represent different LES cases, and they may have gradients $a_1 \neq 1$ and which may differ from each other. We hypothesise (with some evidence from retrieval tests) that this is partly due to differences between the mean profiles in each LES, which results in different changes in line broadening as TCWV changes. For this reason we use separate emulators with different parameters for each case.

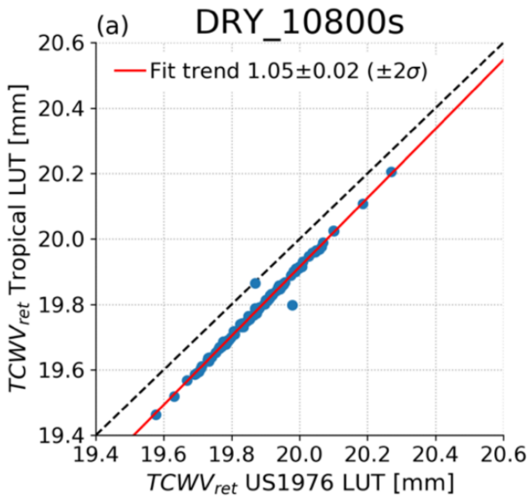
The linear fit can be characterised with stable parameters using just ≥50 footprints, provided we sample footprints spanning the LES TCWV range. Here are linear fits with ±2σ confidence intervals using different numbers of footprints. In our current paper we used 303—707 footprints per case. Way beyond the number required to obtain a stable fit.



Next, we changed SZA input to the plane-parallel RT and retrieval. For ≤45° things are very similar, small (non-significant) changes in parameters occur for SZA=60°, these can be captured by changing the linear parameters, and our structure function results are only weakly sensitive to these changes. SZA=60° may have a slight change in a_1 (again, it's not significant in this case), but this actually has very little effect on derived ζ_2 (Figure 3b):



We also checked what happened when we assumed a different vertical profile in the retrieval with the same integrated water vapour. This changes parameters but not the linear relationship: And it is for this reason (one example shown below) that we argue that the a_1 parameter is at least partially related to errors in the shape of $q(z_{path})$:



Other tests in Richardson et al. (2021) found a few other caveats and limitations and e.g. are the basis of our requirement that structure function properties only be calculated over *mixed vegetation* or *mixed urban/mineral* surfaces, rather than a combination of those two groups.

Solar smearing bias

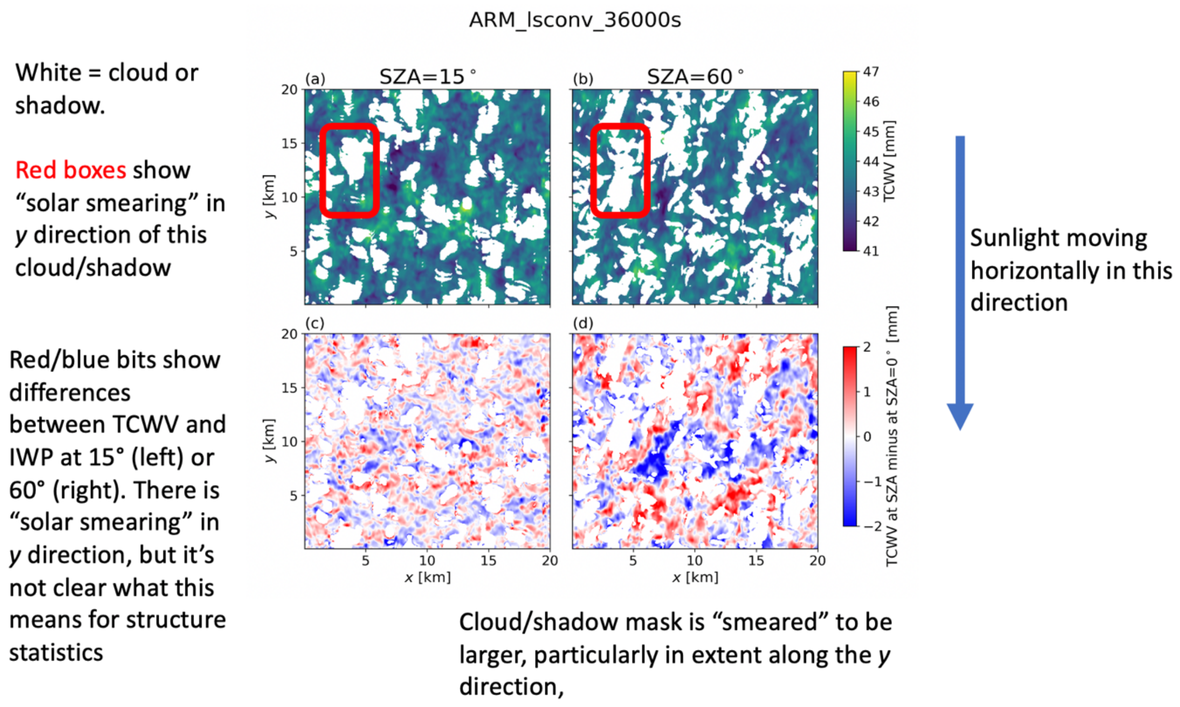
This is our own term and is the shortest name we came up with, but we are open to other suggestions. We are editing the paper to describe this in more detail, and perhaps add a schematic or two. Reviewer input on whether our descriptions are clear to them would be helpful, and in particular which details are necessary for wider understanding.

The reviewer rightly notes that the horizontal variability in the q field leads to differences between $q(z)$ and $q(z_{path})$. Since sunlight comes from a particular azimuth, when the field is measured from the nadir perspective the water vapour measurement is blurred in the horizontal direction of the solar ray, "smeared" due to the slanted path of downwelling photons. This changes the IWV and therefore $TCWV_{ret}$ of each footprint, but importantly for our argument this directional blurring occurs in the direction of the solar azimuth.

To test, we traced the ray path and obtained its IWV, then used our emulator to reliably predict a realistic $TCWV_{ret}$ for that footprint. We then repeat this for every footprint to obtain a map of $TCWV_{ret}$ given SZA.

In the figure below we have the solar azimuth at 12 o'clock and the sunlight's horizontal path being directly in the negative *y* direction. White bits are the combined cloud/shadow mask and the comparison of the two bits in the red boxes makes the *y*-direction smearing obvious, since there is larger *y*-extent of the cloud/shadow mask as SZA increases.

The patterns for IWV and therefore $TCWV_{ret}$ are more complex visually, but we hypothesised that provided that we calculate structure function in the *x* direction, that results would be less affected by the smearing in the *y* direction. By calculating the exponents in each direction (Figures 4/5), we confirmed that the exponents are indeed biased when calculating in the solar-parallel *y* direction but not in the perpendicular *x* direction. Despite all the complicated messiness of the 3D field, our sampling strategy works in LES-simulated low-convective boundary layers.



This is, to our knowledge, a completely novel application and we phrased things with caution and suggested a way to test this in the real world using airborne measurements. We believe we have been very clear about the limitations of this study, but we appreciate the reviewer's response that made it obvious that our decision to exclude prior technical results for brevity was unhelpful.

We hope that we have convinced the reviewer that the technical details are sound. We are updating the paper and are unlikely to finish before the open discussion ends, but we would appreciate any further commentary so we can integrate and address everything as clearly and completely as possible.