Response to Anonymous Referee #1 Comment on amt-2021-229

The authors would like to thank both reviewers for their thoughtful and detailed feedback, after which the manuscript has undoubtedly improved. Reviewer comments below are in green, while responses are in black.

Addressing “*Detailed Comments”*:

Line 120: Because of issues with the dropsondes, quantitative water vapor profiles between the DIAL and radiosondes could not be done. This is a major weakness of this paper. Can the authors speak to future plans for DIAL/Sonde comparisons?

This comment aligns with remarks from reviewer #2 about the ability to adequately validate with the given datasets. We agree with the issues the reviewers raise, and to reflect this the end of the manuscript title will be changed from “validation” to “preliminary results”. Verbiage throughout the text will be similarly updated.

Additional sonde comparisons yielding robust validation statistics will be forthcoming with data collected from the CPEX-AW campaign in the tropical Atlantic/Caribbean (Aug-Sept 2021) and the upcoming IMPACTS mission in the northeast U.S> (Jan-Feb 2022). Preliminary comparisons between HALO and dropsondes from the CPEX-AW campaign show favorable agreement, but cannot be shared publicly until spring 2022 when the campaign data are posted to a publicly available archive.

Line 185, the DIAL equation.

The references:

A. Theopold and J. Bosenberg, "Differential Absorption Lidar Measurements of Atmospheric-Temperature Profiles - Theory and Experiment," J Atmos Ocean Tech 10, 165-179 (1993).

Bosenberg, "Ground-based differential absorption lidar for water-vapor and temperature profiling: methodology," Appl Optics 37, 3845-3860 (1998).

developed a methodology for the DIAL technique that accounts for the Doppler-Broadened Rayleigh scattering. In this development, one can not simply use a modified cross section to account for the Doppler broaded molecular scattering. How do the authors justify using the DIAL equation with the modified cross section and how does that compare with the results from the above two references?

We do employ the methodology outlined in those papers. It is apparent from both reviewer’s comments that the original description of the approach in this manuscript was unclear, so the Doppler correction section has been rewritten.

As a side note, challenges faced in the Theopold & Bosenberg reference drew the conclusion that “high quality information on the different contributions to the backscattered signal is required,” namely the contributions of particulate versus molecular backscatter. This is a distinct advantage of HALO having HSRL + DIAL, as noted in the manuscript.

Line 204 What is the physical reason for the 1.5 exponent in the delta r term?

As derived from Poisson statistics, random uncorrelated noise can be described by sqrt(n) where n is the number of received photons (or in the case of analog systems, number of digitizer counts multiplied by the noise scale factor). Increasing the number of along track accumulated shots or the vertical range bin size increases the received counts linearly and decreases the noise in the observed water vapor retrieval by $\Delta t^{-.5}$ and $\Delta r^{-.5}$, respectively. The additional 1.0 for the Δr term is present because the DIAL technique relies on along-path optical depth where the random noise decreases linearly with accumulated optical depth. This Δr originates in the extinction term of the lidar equation (second term in equation equation 2 from Nehrir et al. 2009, JTECH) and persists through to DIAL uncertainty, as shown in Ismail & Browell 1989. This reference has been made more clear in the text.

Line 240: How was the random error plot generated i.e. what random errors are included here.
The following sentence has been added to the text: "Fig. 1c is a manually drawn estimate of the DAOD and Poisson statistics contributions (shot noise) to the WV uncertainty, meant to illustrate the typical sensitivity of each wavelength pair."

Line 264: If detection is limited to 15 m due to the electronics, how can you get a 1.5 m resolution?
"Signals" in that sentence has been changed to "detection chain" to help clarify.
The detection chain for the WV DIAL is separate from the detection chain for HSRL. As discussed earlier in that paragraph, the 15m limit is for the WV DIAL detection chain.

Line 284: Having the instrument paper would be a useful compliment to this paper. I am not sure about the reference here to an as yet unpublished paper. Could you say more about the status of this instrument paper?
The instrument paper is planned to be written this winter, but as it is still months away from the public forum, mentions of this paper have been trimmed in the text and no specific authors are named.

Line 314: Please add a reference for self-broadening calculations.
This is not a special consideration, since a proper broadening calculation should consider every molecule in the atmosphere. It just so happens that our species of interest (WV) can make up to ~3-5% of the atmosphere, and is thus a non-negligible trace gas. Reference to the general broadening discussions of Ismail & Browell (1989) has been added and the text has been slightly reworded for clarity.

Line 379: How do the choice of the OD theshold values affect the retrieval? Will non-experts in lidar be able to do this manual inspection and can this be automated?
The OD choice does not need to be extremely precise, since the sensitivity ranges of the HALO wavelengths typically overlap as shown with Fig. 1. If an OD value is chosen that is too high or too low, it will typically be very apparent in manual inspection as insufficient DAOD in one of the wavelength pairs results in a very noisy retrieval. HALO WV profiles are archived with this splicing already done, so end users of the data don't need to do it. As for a non-expert in lidar would be capable of, that is hard to comment on as it would vary based on scene and relevant experience…
The following sentence has been added to the text: "Results are typically not affected by threshold changes within ±0.1 of chosen values because the overlap region where both wavelength pairs perform well is sufficiently broad."

Line 413: Do you have any evidence that using the surface reflection over land or the reflection for cloud tops can be used to retrieve the water vapor mixing ratio in the lowest range bin? With what accaurcy can you expect to retrieve water vapor inthe lowest range bin over land or cloud top?
The same principle holds true for making the retrieval over land/cloud as water, though as the text denotes, "Topographic and albedo variability of cloud and land will require a more detailed treatment to ensure surface-related changes in signal are separated from atmospheric OD variation." We have not done enough work on this additional capability yet to make a quantitative projection on performance. Text has been added to that paragraph to be more conservative about future efforts and results, since sufficiently complex surfaces/scenes may prohibit an accurate retrieval.

Line 436: How do errors in the assumed temerpature profile affect the accuarcy of the Doppler correction factor and overall water vapor retrieval?

The magnitude of the Doppler corrections to the HALO WV profiles from the Aeolus Cal/Val campaign datasets were 3.5% at most (this has been added to the text), with the largest magnitudes in the UT/LS and generally <2% in the lower troposphere.

We performed a sensitivity test by changing MERRA-2 temperature by ±10 K and found roughly 10% changes to the Doppler correction factor, at most. Because this is a small perturbation to an already small Doppler correction, and errors in MERRA-2 should generally be much less than 10 K, we neglect any compounding uncertainty that this may cause.

Line 475: When reporting the accuacy of the water vapor retrieval near the surace of less than 15% -- was this only over water? If so please state.

Yes. This has been added to the text.

Unfortantely, the issue with the dropsondes results in the validation consisting of one profile comparison with the DLH profile that resulted from a spiral desent of the aircraft. The comparison of the first range bin of the HALO with the DHL and the comparison of the pericitable water vapor show good agreement, but comparison with the DHL and HALO have a 400 m offset, coparison with the pwv from satellite based measurements will result in a dry bias for the HALO instrument since it misses the pwv above the aircraft, and comparsion of the satellite profiles and HALO are qualitative at best due to the spatial reolution of the passive sensor. This results in a non ideal set of measurments for validation. While this paper is still publishable and, I feel, should be published, I hope to see follow on validation efforts based on sonde profiles.

We share the reviewer's perspective on this and will take advantage of future datasets for a more rigorous validation of HALO WV. As mentioned above, the manuscript title and some related phrasing within the text have been changed to avoid conclusive "validation" terminology.