Review of Chazette et al., "To what extent are IASI water vapor profiles representative of the conditions in the autumn before the HPE? Lessons learned from the WaLiNeAs campaign"

David N. Whiteman, Howard University

General Comments

My apologies for taking so long to complete this review.

It is gratifying to see Raman water vapor lidars used in the manner of the subject manuscript. It describes the WaLiNeAs campaign, a large coordinated effort involving 8 lidar stations from 4 different countries where the measurements are used to compare with IASI water vapor profile inversions. I raise several questions regarding how Raman water vapor lidar data should be reduced in order to best be used as a validation reference as done here. The paper is well-written, interestingly detailed and well-cited and should be published after significant revision to address the major and minor items listed below.

Major Comments

1. Consistency between lidar calibrations.
    1. From the text, I was left wondering if the various lidar systems are calculating the particle component of the differential transmission calculation. It is not quite clear from the text even though it is addressed at some level. Is it true that the analysis of data from all systems has neglected to calculate the particle component of differential transmission? Please address this explicitly. Even though the AOT may not have been large (something that needs more clarification as noted later), if this calculation is not included in the data reduction it results in a bias in the derived products, not a random uncertainty that could be reduced by averaging. The point is made in the manuscript that neglect of the particle component of differential transmission results in a bias but it seems to be small compared with the magnitude of differences that are found in the analysis. But when lidars are being used as a reference, should any component of the data reduction be neglected? I would argue that it is long past time that the full differential transmission, including both molecules and particles, is routinely accounted for in the analysis of Raman water vapor lidar data and reviewer comments such as this become a thing of the past. Please see the recent paper by Diaz-Zurita et al. (Remote Sensing, 2025), for an example of a way to automate the calculation of the particle contribution to the differential transmission.
        1. To address this, please be explicit if the differential transmission due to particles has been included in the data reduction and, if not, state the estimated magnitude of the bias that can result. Better yet, of course, would be for the groups to include this in their analysis but I realize the difficulty of getting that done in a timely manner with 8 different lidar groups since some sort of approximation is likely needed such as done in Diaz-Zurita et al mentioned above. So a clear statement of the estimated bias, lidar by lidar, would suffice although I sincerely hope the community makes progress toward routinely including this in their analysis of Raman water vapor lidar data.
2. Not addressed in the paper is the influence of the temperature sensitivity of Raman scattering in the evaluation of the mixing ratio equation. Since these systems offer daytime measurements one can assume that the water vapor filters are quite narrow (this is quite important information for the consideration of this temperature sensitivity so should be included for all systems). I could only find the filter width of 0.3 nm listed for the RL2 system. This filter width (assuming optimal positioning of the filter) was analyzed in Whiteman, 2003 and found to result in ~4%

increase in mixing ratio between the surface and 11km. So for the purposes of the evaluation done here the effect might be about half of that given that comparisons are done up to an altitude of 5 km. It is very important to note that this temperature effect also results in a bias that is in the same direction as the particle contribution to differential transmission. In other words, these effects are additive and both create an apparent increase in mixing ratio with height. Given that the authors state that neglecting the particle component of differential transmission can result in biases up to ~4% and that the temperature dependence adds another ~2% to that, the combination of the two can create more than a 5% bias in the lidar profiles, which are treated here as reference profiles. Should reference profiles contain 5% biases when it is well known how to do the calculations to reduce or eliminate these biases?

    1. To address this, please consider the filter widths of the various systems in use and give the reader information on the magnitude of the unevaluated temperature sensitivity of the Raman lidar measurement of water vapor mixing ratio.
        1. Better yet, of course, would be for the community to also routinely perform these temperature dependence calculations which are not more difficult than those of rotational Raman scattering. I understand that this is also unlikely to be accomplished in a timely fashion when dealing with 8 different lidar groups. But, as with the particle component of differential transmission, the community should strive to make these calculations a routine part of the data reduction so reviewers cannot pick at incomplete data analysis in papers like this.

3. An argument is made that due to uncertainties in the mixing ratio that are calculated to be <= 0.1 g/kg, and due to regressions between lidar and IASI that yield coefficients of determination of 0.71 – 0.91, that in calculating the consistency factors using comparisons of lidar and ERA5, only the slopes of the regressions between lidar and ERA5 need to be considered. It is also stated that the biases of the regressions are more likely to be due to modeling errors than to instrument errors thus supporting dropping the calculated biases (offsets). I don’t believe that either of these justifications is convincing.
    1. I believe the authors need to take a different approach with this part of the analysis. Is any filtering of the datapoints done prior to regression? Common sense considerations can be used to yield a more representative population of points. For example Whiteman et al., AMT, 2012 show a technique where outliers in the regression between sonde and lidar are filtered out by considering how geometrically similar portions of the profile are. In the 2012 paper, choosing geometrically similar portions of the curve helped to deal with the issue of the radiosonde drifting with the wind and thus sampling a different atmosphere than the lidar. Perhaps part of the disagreement in the comparisons as shown in Table 2 comes from the sampling of different atmospheres between the lidar and IASI. A filtering procedure could remove portions of the profiles that show poorer agreement with each other thus improving the regression.
        1. BUT, and this is very important, the calibration factor determined from such a comparison is not just the slope of the regression obtained. The calibration factor can be taken as the mean (or median if there remain some outliers in the population) of the ratio of the points selected for regression by the filtering technique. I would argue that the same type of procedure should be performed here in this analysis so that there are no questions about selectively choosing the slope alone from the regressions of lidar/ERA5 and neglecting the offsets.
            1. Please redo this analysis and either use the full regression equation, including offset, to compare systems (not recommended) or use the mean/median of the ratio of points selected by a careful filtering algorithm (recommended).

4. It is difficult to assess the magnitude of the effect of the neglected height dependent calculations that are being requested above since nowhere in the paper nor in the supplement are the mean lidar, IASI, ERA5 profiles of water vapor mixing ratio shown. Please add mean mixing ratio profiles to figures 4 and 5 (and to figs S18 – S25 in the supplement), perhaps as a third panel. This will greatly aid understanding the measurements in general.

Minor Comments

1. Line 2: please spell out the meaning of HPE in the title. It is not defined in the paper until line 284 so I also suggest mentioning this in the abstract if the relationship of the measurement period to when HPEs usually occur is important to the paper. In fact, it is not clear from the text of the paper that this is an important consideration and the authors may want to consider changing the title of the paper to more agree with the work presented in the body of the manuscript.
2. Lines 39-48: a description of IASI is given. It is described that IASI radiances are assimilated into forecast models but the focus here is on validating the IASI inversions. It would be very helpful to know what IASI profiles of water vapor are currently being used for as it would help to motivate the current work.
3. Line 109: statement is made "...the planetary boundary layer is therefore less directly constrained by IASI measurements." From the discussion of the paragraph, you are referring to products that use IASI radiances as one of many inputs. But the paper is about inversion products such as the profile of water vapor which are based mainly on IASI radiances, I believe (perhaps with some machine learning to improve the boundary layer?). So is this statement about the use of channels that peak above 2 km pertinent to the main subject of this paper which is comparing IASI and lidar profiles?
4. Line 151: Figure 2. Gradients in IPW appear to be strongly correlated with land-ocean boundaries raising the question of whether the IASI retrieval may have emissivity dependencies that are influencing the images shown in Fig 2.
    1. Can the authors supply any references that serve to address that concern about the IASI inversions?
    2. Are there known issues in the inversions over land vs water? If those inversion biases are known to exist, do they influence the comparisons shown here since all stations are located near the land/water boundary? If so, can you estimate how much? Can you speculate if this intercomparison might have yielded better results (reduced IASI dry bias) if the lidar systems had been situated away from the coastlines?
5. Line 168: corrections covering a range of 17% are referred to as "small". In the context of a validation campaign I would not describe such a range as small. Please use more appropriate wording.
6. Line 170: Statement is made: "In some cases, these may be associated with imperfect corrections for molecular and particulate contributions in the lidar data." I believe using the term "corrections" here is misleading. A correction is something that is done because of a known error/mistake. The error or mistake in this case would be to *not* calculate the molecular or particle transmission. (Authors state that the molecular component of differential transmission is being calculated the same by all groups. The major question above addresses what the different groups are doing regarding particle transmission.) So calculating the differential transmission is not a correction; it is simply a calculation of what is explicitly in the lidar equations. I suggest instead referring to "calculations of the differential transmission due to molecules or particles". Full disclosure ... in the past I have also inappropriately used the term "correction" in this context..

7. Line 173: “...either because aerosol optical thicknesses (AOTs) values are below 0.15”
   1. please make an explicit statement that at all sites for the entire campaign the AOT remained below 0.15. Please specify the wavelength at which this statement pertains and what the source of information is for AOT.
   2. Note the issue in the quoted text if it remains: either strike “values” or change “thicknesses” to “thickness”
8. Lines 177-180: Very good. This is helpful. But as noted above please be explicit about which systems have made a correction for particle transmission. It is not optimum for the reference systems (lidars in this case) to have biases of 4% if they are to be used for assessing IASI profiles.
   1. Please note that here you refer to a “typical AOT of 0.2 (355 nm)” but earlier you stated that AOT is always below 0.15. Please reconcile these statements as they bear exactly on the issue of uncalculated differential transmission.
   2. Please provide a citation for the statement “...aerosols are primarily desert dust with Angstrom exponents less than 0.5”.
9. Line 191: Table 2. I don't understand what you have done here. Please provide an equation so it is clear how this correction is applied. It is also confusing (and not really correct) to give a correction factor in units of percent. I assume that what you mean is if you have a correction of +10%, the factor is 1.1 and one of -7% is 0.93. An equation showing how the correction works with the actual factors provided (not percentages) would help clear up this confusion.
10. Line 200: “The number of IASI pixels …. ensuring minimal terrain obstruction.” Please expand on this by giving an example. Are pixels completely removed or is the retrieval only used above mountain height?
11. Line 206: “These metrics are used …”. Please tell the reader how many points are used at each z to determine MB, RMSD, etc.
12. Equation 3: Please add z dependence to RMSD and MB.
13. Equation 5: MB and RMSD are not continuous functions. Please express these equations using summations instead of integrals.
14. Line 225: “...temporal dynamics differ between stations.” I believe that “among” is more appropriate than “between”.
15. Line 227: “Low clouds can easily … lidar measurements.” This sentence seems to not have a real place in the flow of the text. Can you prepare it better with sentences before or after?
16. Line 239: “The vertical profiles of the three statistical parameters…” Please revise to add the figure numbers to which this statement pertains.
17. Line 250: “This is due both to … within the PBL.” I suspect this is also influenced by the fact that ERA5 was used as part of the cross calibration of the lidars. This should be mentioned here.
18. Line 351: Table A1. Reception channels for vibrational N2 are noted as 387.5 nm. This disagrees with the text in the paper and is likely just a typo. Also, there was confusion earlier about how many of these systems are using the RR signal to normalize the water vapor in the MR calculation. Please make sure the paper is consistent in this regard between the paper text and this Table.
19. Line 360: N2 wavelength is given as 387.5. Please correct.

Supplemental material

1. Figures S1 - S9: please provide regression equation as an inset to the figure itself. This will greatly aid ease of interpretation.
2. Figure S18 – S25: as a reminder, please add the pertinent mean mixing ratio profiles perhaps as a third panel to all figures.