Review of "A Raman lidar at La Reunion (20.8 S, 55.5 E) for monitoring water vapor and cirrus distributions in the subtropical upper troposphere: preliminary analyses and description of a future system", Horeau et al., AMTD

David Whiteman, NASA/GSFC

General Comments

The subject manuscript describes a Rayleigh-Raman lidar used on La Reunion island during the period of 2002 – 2005 for extended measurements of water vapor and cirrus clouds. An upgraded instrument, that will be installed at a new mountain-top observatory on the island and offer significantly improved measurement performance is also described.

The subject material of the manuscript is very interesting and important to be documented as the establishment of a permanent, high-performance, Rayleigh-Raman lidar in a sub-tropical setting will permit very valuable measurements of water vapor, cirrus clouds and temperature to be acquired for the purposes of satellite validation and trend monitoring. This will be an important instrument within the NDACC network. In the current state, however, I find the manuscript to not be ready for publication. Despite the numerous concerns that I point out below, I encourage the authors to continue working on this document as it can become an important one for publication. Part of this effort should include a careful reading by a scientist who is a fluent speaker of English as there are some statements that were unclear due to their unusual structure.

There are in general many figures in this paper and sometimes rather little text devoted to a particular figure. I suggest that the authors decide what the most important points of the paper are, remove some figures and devote more text to those figures remaining. Also, I would appreciate if for all equations the units are provided when the terms are defined.

The paper can become an important contribution with some tightening up. I hope the following detailed comments are helpful toward that end.

Specific Scientific Comments

1. P6452, lines 19-20. You are making a point of how laser power has been increased since early measurements and use our 1992 reference. That reference referred to measurements from the 1980s and the laser power was only 1.5 W. A more appropriate reference would be the following which also discusses results of the MOHAVE-II campaign, one focused on NDACC activities
   1.1. Whiteman, David N., Kurt Rush, Scott Rabenhorst, Wayne Welch, Martin Cadirola, Gerry McIntire, Felicita Russo, Mariana Adam, Demetrius Venable and Rasheen Connell, Igor Veselovskii, Ricardo Forno, Bernd Mielke and Bernhard Stein, Thierry Leblanc and Stuart McDermid, Holger Vömel, Airborne and Ground-based measurements using a High-Performance Raman Lidar, doi:10.1175/2010JTECHA1391.1 (2010).
2. P6453, line 15. After the words "... theoretically possible", I would add appropriate references as there are, as far as I am aware, only three published accounts of first principles calibration of Raman water vapor lidar. They are:
   2.1. G. Vaughan, D. P. Wareing, L. Thomas, V. Mitev, "Humidity measurements in the free troposphere using Raman backscatter", Q. J. R. Meteor. Soc., 114, 1471-1484 (1988).
   2.2. V. Sherlock, A. Hauchecorne, J. Lenoble, "Methodology for the independent

calibration of Raman backscatter water-vapor lidar systems", App. Opt., 38, 27, 5816-5837 (1999).

2.3. D. D. Venable, D. N. Whiteman, M. N. Calhoun, A. O. Dirisu, R. M. Connell, and E. Landulfo, A Lamp Mapping Technique for Independent Determination of the Water Vapor Mixing Ratio Calibration Factor for a Raman Lidar System, Applied Optics Vol. 50, Iss. 23, pp. 4622–4632 (2011)

3. line 20. “More independent techniques”. I don't believe there can be a qualification on the use of the term “independent”. A calibration is either independent of an external measurement of water vapor or it's not. I think one can talk usefully about different degrees of dependence of a calibration on external water vapor measurements but there are not different degrees of independence. That being said, of the two references provided (Sherlock 1999b and Leblanc 2008) only the Sherlock reference relates to an independent calibration technique although the more appropriate citation would seem to be Sherlock et al, 1999a. This is the publication where the independent calibration technique based on sky radiance is described so I would add in the 1999a reference here as well.
4. line 25. Reference is made to a need for calibration stability of 10% with no citation. This is a critical number for trend detection purposes. As far as I am aware, the only publication that really addresses the question of the range of tolerable measurement uncertainties for upper tropospheric trend detection is one we have recently published. It demonstrates that random uncertainty budgets of 50% and more are tolerable for the development of a climate quality time series in the upper troposphere. Any calibration changes must fit within that uncertainty budget and, ideally, vary randomly over time. A greater tolerance for uncertainty in the time series relaxes the need for calibration accuracy of Raman lidars, which makes the challenge of calibration perhaps easier to meet.

   4.1. Whiteman, D. N., K. C. Vermeesch, L. D. Oman, and E. C. Weatherhead (2011), The relative importance of random error and observation frequency in detecting trends in upper tropospheric water vapor, J. Geophys. Res., 116, D21118, doi:10.1029/2011JD016610.
5. P6454, Section 2.2.

   5.1. I am confused by this section. In order to calculate the scattering ratio you must use the ratio of the signal backscattered at the laser wavelength and that of Raman nitrogen. So the nitrogen signal exists, but you choose to not use it in the calculation of the cirrus optical depth. Perhaps the signal is too noisy for a good optical depth measurement but please explain why you do not make a direct calculation of cirrus optical depth. If it is for concern with multiple scattering, there are techniques to account for their influence described in references by e.g. Eloranta, Reichardt or Whiteman.

   5.2. The value of LR chosen to invert the SR to optical depth (after converting to backscattering coefficient) is derived from a 1984 publication of Platt. The LIRAD technique over the years gave some very different values for mean LR so I don't think this is the best reference to use. However, analysis based on Raman lidar measurements yield LR values in quite good agreement with the value chosen by the authors (see Reichardt reference). Furthermore, subdivision by cirrus generating mechanism seems to not influence significantly the layer mean LR consistent with the assumption in this study (see Whiteman reference) even though other paramets

      a) Reichardt, J., S. Reichardt, A. Behrendt, and T. J. McGee (2002), Correlations among the optical properties of cirrus-cloud particles: Implications for spaceborne remote sensing, Geophys. Res. Lett., 29(14), 1668, doi:10.1029/2002GL014836.
      b) Whiteman D. N., B. Demoz, Z. Wang (2004), Subtropical cirrus cloud extinction to backscatter ratios measured by Raman Lidar during CAMEX-3, Geophys. Res. Lett., 31, L12105, doi:10.1029/2004GL020003.

6. P6455, paragraph starting on line 7.
    6.1. Please incorporate the existence of fiber optics in the system in the discussion of this paragraph. You show fiber optics in the system diagram shown in figure 2 and the change away from fiber optics is a detail that you point out about your new system, but there is no discussion of their existence in this paragraph and it seems there should be.
    6.2. Please explain what alpha-epsilon and alpha-omega mean. I am not familiar with this notation.
    6.3. Please introduce figure 2 several sentences earlier in the discussion in the paragraph so that the reader is encouraged to refer to this figure as you describe the spectral separation of the beams.
    6.4. The statement is made that low dark counts (<5 counts per second) are important for detecting the weak N2 signal. I would have thought that for the N2 signal, detector dark counts would not be such a concern. And in any case, it is also important to consider the noise due to background skylight. With a 1 nm wide bandpass filter and a 1 mrad field of view, I suspect that skylight is much, much larger than 5 counts per second. If that is the case, then it really is not necessary to cool the pmt. Perhaps you could discuss the level of skylight background here as justification for the use of a cooled pmt since it is not apparent that it really is useful in this channel. There is also a discussion later in the same paragraph about the need for "very low noise detectors for detecting the very weak water vapor returns." Here, with the much lower water vapor signals, it is much easier to understand that low dark count rate detectors can be a benefit. But only in the case where the skylight has been limited much more due to spectral and spatial filtering. Would you please consider the relative values of the signals being detected, the detector dark count rates and the skylight values and revise this paragraph accordingly?
    6.5. You mention "signal-induced bias associated with the PMT response to an intense luminous pulse". A reference would be helpful for readers not familiar with this phenomenon.
    6.6. You refer to "...maximum count rate for a Poisson signal with exponentially distributed inter-arrival times is 45 MHz." A reference would be helpful here as well.
7. P6457, section 4.1. The authors explain that in the calculation of water vapor mixing ratio no corrections have been applied to account for the different transmission at the two Raman wavelengths. Even though the magnitude of the differential transmission due to molecular scattering is just a few percent, if uncorrected, this introduces a systematic bias in the water vapor profiles. Correcting for known systematic biases is part of good metrological practice and should be done when the magnitude of the effect is known and significant, as here. It's also easy to do. Ideally, differential transmission due to aerosols should be corrected as well. This is particularly important to be done in the region where any comparisons are being made versus another profile such as the radiosonde. Such comparisons typically are done in the lower atmosphere where aerosols are more prominent and the correction will change the shape of the profile where aerosols are present. For the method of calibration used here, a comparison with ECMWF, this is less of a concern and the dis-regard for aerosol scattering is justifiable. But a known systematic bias such as that due to differential transmission by molecules is easy to do and should be corrected for. As support for that position, I quote from the Joint Committee on Guides in Metrology's Guide to Expression of Uncertainty: "It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects."
8. P6459, section 5.1. The authors discuss the number of independent profiles that they acquire during the nighttime measurements. This is based on the technique described in more detail in Hoareau et al, 2009. It is referred to in section 4.1 but I suggest adding more material to section

4.1 to explain this technique in more detail, since at this point in the manuscript I doubt the reader will have a firm idea of what the authors mean by “independent”.

9. P6460 Section 5.2, reference is made to the moist and dry seasons; please describe, in terms of months of the year, how the data were divided for this comparison. In other words, which months are considered to be the moist ones and which months the dry ones?
10. P6461, lines 8 – 10. The authors state that the nighttime measurements show larger variability than the diurnal cycle and refer to Fig 9 as part of the explanation. It is not clear to me exactly what the claim is here nor how Fig 9 supports it. Please expand this discussion for clarity.
11. P6462, The cluster analysis of the cirrus cloud measurements is very interesting and the authors show some well-defined clusters but it is not clear how the analysis has been done. Please expand this discussion for clarity.
12. P6464, line 26. The statement is made “A narrow field of view of 1 mrad is used to reduce as little as possible sky background.” This statement should be revised since several Raman lidar systems (e.g. Payerne, DOE/ARM, HURL, ALVICE) are in operation using 200 – 250 micro-radian fields of view achieving between 16 and 25 times reduction in skylight versus the use of 1 mrad.
13. P6465, line 5. The authors state that their use of fiber optics in the older system permitted them to obtain “constant illumination conditions at the optical fiber output”. Please expand here to describe how this was achieved since the use of short (1-2 m), multi-mode fibers have been shown to preserve a large amount of spatial information from the input to the output. (c.f. Whiteman et al., 2011 that you are currently referencing). It has not been published but tests have been done at GSFC with fibers as long as 30 m where it was not possible to achieve complete scrambling of the signal even with significant bending of the fiber.
    13.1. Regarding fluorescence, I should point out that we are finding the statement from Sherlock et al., 1999a of "...although here absorption of the elastic-backscatter signal occurs in the fiber-optic cable used for signal transfer, it could arise in any optical component.” to be quite correct. We have been testing various optical components: aluminized mirrors, dichroic beamsplitters, lenses, etc and thus far have found that every item tested has at least some fluorescence. So just a word of caution that the concern about fluorescence is not limited to fiber-optics in agreement with Sherlock et al.
14. P. 6465 Section 6.3. please introduce figure 16 much earlier in this discussion so the reader refers to it during the detailed discussion of optical configuration.
15. P 6466 Line 12. the authors state that they will be using gated tubes. I assume that they are not referring to the use of these tubes on the Raman channels or are they? Please clarify. If they will be used on the Raman channels, what will be the minimum altitude of water vapor retrieval?
16. Line 13. Please expand the discussion of what tests you anticipate needing to do on the new pmts and why they will be needed.
17. Line 15. the authors mention two versions of Licel Transient Records. Please explain what the differences are between these two units and why you would choose to use one versus the other.
18. P6466, line 10. The authors state that they will use either R7400-03g or -20g pmts depending on what laser wavelength will be used. Please add a sentence of two to discuss why this is the case and which detector would be used depending on the emitted laser wavelength.
19. P6466 Section 6.4. Paragraph starting on line 18. The authors refer to the “limiting factor for a PMT in photon counting mode is the dark current”. It's not quite clear what the authors mean here but this statement seems to disregard the influence of background skylight which I would speculate, for both old and new systems, is a larger source of noise than the

PMTs themselves for the case of UTLS water vapor measurements, which are the focus here. Please clarify and relate to the sky background.

20. P 6467 Section 6.5. If I understand the discussion here, the radiosonde will be located 20 km away and the authors state that ideally the GPS system should be located near the radiosonde. But if the goal is to calibrate the lidar system based on total column water, it would seem important to have an IPW measurement collocated with the lidar system. Please clarify if that is the plan.
21. P6476, line 7. The statement is made that the new design will extend measurements down to the ground. Please clarify and reconcile with earlier statement about the use of gated PMTs. Also, it is unlikely that the lidar will measure usefully all the way to the ground so you might choose some different wording such as "down to close to the ground".
22. Line 17. Reference is made to "NDACC future recommendations". It is possible that some variant of the hybrid technique may be recommended by NDACC in the future after careful consideration of failure modes of the technique. But no recommendation was made regarding the use of the hybrid technique at the Raman lidar calibration workshop held in May 2010.
23. P6469, line 8. The ratio of detector noise to sky noise is given as 2 for a measurement made in 2005. I guess this was with the apd detector for the water vapor channel. Please give the dark count rate of the detector and sky background as separate values, if possible, since this seems to imply a very high dark count rate for the detector.
24. P6469, line 22. The statement is made that the decrease in the performance ratio discussed (I think of the future system versus the current system) is "essentially due to the altitude squared dependence in the return signal." The altitude-squared dependence is the same for both current and future systems. It seems to me that this change in the performance ratio, which I take to be the ratio of water vapor signals between the current and future systems, must be due to something else. A difference in signal to noise between the two configurations would be expected as the signal decreases if the skylight contribution of one system is different from the other. But (I think) the authors are just comparing water vapor signal strength between the old and new systems so I don't think this should occur. Please investigate and clarify what exactly the ratio is and why it changes with altitude.
25. P6470 Line 7. The simulated performance is a factor of 4 higher than the actual performance. Please provide an explanation of what this might be due to and whether the simulations of the future system have been adjusted by the same factor of 4 to better correspond with real-world expectations. And if not, why not.

(Mostly) Scientific Comments concerning Figures

1. There are several items identified in the figure that are not described in the text. Pplease either describe these items in the text or remove the labels from the figure.
2. OK
3. In the caption, it is stated that the lidar profile has been calibrated with respect to the ECMWF profile and then compared with the CFH. Please add this to the text and explain how it was done (based on IPW or by profile comparison between a certain pair of altitudes) since the impression one gets from this comparison is that the calibration was actually done by direct comparison with CFH, which is a perfectly acceptable thing to do but something different than stated. Also, please explain what is shown in the left part of the figure. This is some display of the calibration values but it is not clear if this value pertains just to the data shown on the right or to a larger ensemble of measurements.. Finally, I suggest using color on this figure for ease of interpretation.

4. This is a useful plot. Presumably the lidar data are calibrated using the best median values shown in the figure for the periods shown. Please clarify. Also, please explain why the numbers shown here have such different values than the calibration values shown on the left of figure 3.
5. In the text, the reason given for the larger number of profiles in the Feb-Mar and Oct-Nov periods is due to the longer nighttimes. But aren't the nights longer at 20S latitude in the June-July period? What is the influence of cloudiness on these statistics? Is that a more likely explanation? Perhaps I misunderstand the written text, but it would seem that the number of independent measurements acquired per month is not explained by the length of the nighttime period.
6. OK
7. the references to left and right panels are reversed. Also, what is displayed on the right is the ratio of the two seasonal profiles, not the relative difference as stated in the figure caption. Please correct.
8. I think the bimodal distribution shown here is a bi-product of subsidence of the moist layer. I'm not sure how significant it is to refer to this as a bi-modal distribution of water vapor. I think if a different height range (e.g. 3.25 km) or interval (3.0 – 3.5 km) had been chosen a quite different impression of this distribution could have been obtained. Please consider these points. Also, a pink line is referred to in the caption here but I cannot see a pink line. And, "On the right panel" should be changed to "On the bottom panel".
9. Are these data from lidar measurements? Please clarify. Also, I do not really understand how these measurements relate to the statements about diurnal variation. Please expand the text concerning these measurements to clarify their significance to your argument.
10. OK
11. OK
12. Should the confidence interval value here be 95% instead of 5%. 5% confidence would not show such a strong clustering, I think.
13. OK
14. I'm not sure this figure is needed. If it is to be retained, please label the figure to clarify the points made in the text.
15. OK
16. please make the labels more legible here.
17. These are very interesting simulations but there are some details that I do not understand.
    1. If I understand which line is which (I suggest changing the plot to use color), why does the future system have a higher noise floor under the moonlit conditions than the current system but a lower noise floor under moonless conditions? Perhaps I have not thought deeply enough about it but it is a curious result that could use some clarification.
    2. Also, I think it would be interesting to see what altitude increase could be obtained if a telescope fov of, for example, 0.25 mrad and an interference bandwidth of 0.25 nm were used. Several Raman lidar systems with parameters close to these values already exist. The combined use of these two values would increase by a factor of 64 the signal to noise ratio of the future system over what is simulated. This would both significantly decrease the random error budget of the UTLS measurements and increase the upper altitude limit of useful water vapor measurements. These would seem to strongly address the goal of UTLS water vapor monitoring. In terms of increasing the upper level performance, such changes would have much more effect than doubling the laser power which would increase the S/N by a factor of 2. S/N optimization for improvement of UTLS water vapor measurements by consideration of the noise term (instead of the signal term as the laser power addresses) is discussed in Whiteman et al, 2010 referenced earlier.

Technical Corrections

1. P6451, line 20. Partial sentence; either eliminate or complete "This requires an accurate determination"
2. line 24. Sentence starting with "Water vapor can be considered ..." is overly long and rather confusing. Please rephrase and perhaps split in two in order that your point is clear.
3. P6457, line 15. Reference is made to "an a-priori period of quasi-stationary conditions..." Please explain more of what you are doing here and why. Supplying a reference would be helpful.
4. Line 25. The statement is made that the CFH measures to the "middle atmosphere". Please change to "lower stratosphere". As phrased it gives the impression that the instrument is used to much greater altitudes.
5. P 6458, line 16. The integration time is provided. Please also provide the vertical resolution of the data as displayed in the figure.
6. P6459, line 10. what does the phrase "on the whole of medians" mean? Please explain.
7. P 6460, line 10. The qualifier "almost likely" is used. Please re-phrase.
8. P6463, line 6. Please remove incomplete sentence starting "with based on ..."