

Solar Background Radiation Temperature Calibration of a Pure Rotational Raman Lidar, Jayaweera et al. – a Review

David N. Whiteman, Howard University

General Comments: This is clearly a good idea and the authors are commended on being the first to demonstrate this new calibration idea which many may react to by saying “why didn’t I think of that?”. There will no doubt be some iteration on how to best implement the use of background signal in the two rotational Raman channels for temperature calibration but I suspect that eventually this will become the preferred method of calibrating a RR temperature lidar. The authors’ results certainly point in that direction although I do have some comments and questions that I hope might help to improve the technique and solidify this publication as the seminal one for this new calibration approach.

Major Comments:

1. Line 33: Sentence starting “However ...” this sentence seems to imply that the technique presented here does not carry the uncertainty of the reference sonde but of course it does. Please revise sentence to eliminate any confusion.
2. Line 40: Sentence starting “What sets this approach ...” I disagree that use of a single external reference calibration effectively diminishes the uncertainty stemming from the reference instrument. In fact, use of a single radiosonde for calibration implies that any systematic uncertainties associated with that single sonde will be carried through the entire calibrated time series. This is a distinct weakness in the proposed approach and the reason why an ensemble of sondes is typically used for calibration – the influence of systematic uncertainties due, for example, to sampling issues or sonde/lidar malfunctions will be reduced in the ensemble. Furthermore, use of an ensemble of sondes may permit some characterization of uncertainties in the reference dataset and in the transfer of the calibration to the lidar. Those uncertainties must then be propagated into the uncertainties in the time series. Without an ensemble of sondes, what uncertainty will be attributed to the transfer of the reference calibration? There are at least three terms that must be evaluated to obtain the fully propagated uncertainty of the calibrated time series: 1) the uncertainty in the lidar retrieval, 2) the uncertainty in the radiosonde used for calibration 3) the uncertainty in the transfer of the radiosonde calibration to the lidar time series. How do you evaluate 3) using your current scheme? If an ensemble of sondes is used for calibration, you can use the statistics of these comparisons to estimate the uncertainty in transferring the calibration.
3. Line 104: About the use of 70 deg SZA for solar calibration ... given the spectral proximity of the high and low J channels, is any significant difference in the solar spectrum expected? One of the great attractions of this technique would be if one could simply assume that the solar spectral intensity is essentially identical in the spectral windows of HiJ and LoJ. An interrogation of a high resolution radiative transfer model would answer this question. You also could address this question empirically by trying the method at different SZAs to see if there is any difference in the ratios computed. It may be that this ratio is independent of SZA. It may also be that you could use cloudy data just as easily. The question is whether there is a significant difference in the spectral intensity of the source

(whether the sun, clouds, whatever) at the two closely spaced optical channels. That really should be addressed in this seminal work.

4. Figure 2: plot in Fig 2a seems to show a few clear outliers with temperature differences of 10-20 K in the lowest 10 km. Some filtering should be done to eliminate these from the statistics since they are pretty clearly aberrant comparisons. There also are a couple aberrant comparisons in Fig 2b, the background calibration plot, with differences of close to 30 K. If these are not outliers and should be included then some explanation is needed as to why these comparisons are retained.
5. Table 1. Please include standard deviations with all of these values. I suspect that for many (or perhaps most) there will not be a significant difference, even at the 1-sigma level, between the external and solar background methods (particularly after addressing comment 4 above). If so, some of your discussion will need revising. Realize though that even if after addressing this comment there is not a statistically significant improvement in the results, this new technique has the distinct advantage of tracking calibration changes better. You know this, of course. My point is to not sweat it if the statistical results are not better with the background technique. The ability to calibrate once and then track changes is more than enough to warrant use of this new calibration idea.
6. Line 170: "We have shown the solar background calibration method" You have shown this when used in conjunction with the OEM retrieval method. I suspect most of the community still uses the traditional method a la Behrendt et al. Would it not make your argument more persuasive if you applied your technique also to retrievals using the traditional method? I think you should take the time to perform this work so that you get full credit for establishing the solar background technique as the preferred way to calibrate RR lidar temperature measurements (which you deserve!).

Minor Comments:

1. Line 14: delete first appearance of "system" to avoid redundancy
2. Line 28: change "from" to "from the determination of"
3. Line 30: I would introduce the phrase "Optimal Estimation Method (OEM)" in this sentence so that you can refer to OEM moving forward. For most of us, OEM is more likely to be understood as Original Equipment Manufacturer so best to avoid that confusion up front.
4. Line 63: Claim is made that RALMO meets "stringent OSCAR ... requirements ... with an uncertainty of less than 1K". Please note that the OSCAR requirements are specified as "goal", "breakthrough" and "threshold". The 1K uncertainty in temperature that is referred to is the "breakthrough" requirement for High Resolution NWP in the Free Troposphere. It would be best to refer explicitly to this "breakthrough" requirement. *BUT* the breakthrough requirement is not just about the level of uncertainty. In order to meet this requirement, it is also necessary to provide this level of uncertainty at 5 km horizontal resolution every 30 minutes for the entire free troposphere. Please clarify if RALMO meets these other parts of the requirement.
5. Line 67: All lidars require calibration whether internal or external. So I would suggest changing this sentence to something like: "Lidar temperature measurements, including those from Raman lidar studied here, require calibration."
6. Line 71: Equation 1. So that the reader does not have to refer back to the earlier paper, I suggest that you reproduce equation 1 from Mahagammulla Gamage et al., 2019. Then

your equations follow more easily where C_{RR} from M-G becomes either C_{JH} or C_{JL} in your terminology. It would also help to maintain use of R from the earlier paper for the ratio shown in your equation 1 as C^* . If you prefer to use C^* (script R is used for things like aerosol scattering ratio in the literature) then just introduce short text here explaining that your notation is deviation from M-G.

7. Line 72: reference is made to “RALMO’s digital channels”. This seems an awkward construction. The lidar constants refer to optical characteristics which are measured, hopefully accurately, by the digital electronics. Suggest changing “RALMO’s digital channels” to something like “RALMO’s high J and low J rotational Raman channels, respectively.”
8. Line 77: same comment about “digital channels”
9. Line 78: typo “GRAUN”
10. Line 79: you introduce a new concept that equation 2 has a range dependence which you are ignoring. I think you need to expand on this some. What is the range dependence due to? What is the estimated uncertainty due to this omission?
11. Line 93: as stated in the major comments, I disagree that use of a single sonde for calibration significantly reduces uncertainties. Having a single reference for your time series eliminates jumps in calibration, which is definitely a good thing. The issue is whether your single reference should be obtained from one comparison or from an ensemble. The argument in the major comments is that you should use an ensemble of comparisons to derive the calibration.
12. Line 101: Yes, you can choose any point in the time series for establishing the calibration but it seems awkward to me to call it t_0 since this is usually understood to be a start time of an experiment. I would suggest you change this notation to something like t' to avoid this confusion. Or at least I was confused by how you could consider times prior to t_0 ...
13. Line 103: I understand that you are analyzing a historical dataset and thus must avoid the laser return signal. But, our publications have shown Raman nitrogen returns to beyond 55 km (Whiteman et al., 2010 Fig 12) and your rotational Raman signal is likely stronger. So please state what you did to assess that no significant returns were observed from 55 km.
14. Line 109: Regarding the calibration results using the external method. I assume that this work was done as a part of this paper and you are not referring to work performed in an earlier effort. I suggest you make this explicit by, for example, changing the sentence that starts “For the external method...” to “As an illustration of the external calibration method we processed ...”
15. Line 113: Sentence starting “This approach involved ...”. Just to point out that this filtering method could be used to choose a set of profiles to investigate whether the ratio in your equation 1 is the same under cloudy conditions as for solar background. You would still use the values above 55 km as before.
16. Line 119: Suggest rewording sentence starting “The t_0 ...” to “The date selected for the t' calibration was June 5, 2013.”
17. Line 119: You refer to “error values of each external calibration constant”. It is not clear what this refers to. Have you done some kind of RMS difference between lidar and sonde profile to determine a difference statistic? Please describe in more detail but I would avoid the use of the term “error” since this implies a reference to an absolute standard and I don’t think you can make such a claim for a comparison between sonde and lidar where they are likely sampling different volumes. You can talk about differences or uncertainties but error is likely not a proper term here.

18. Line 122: concerning "using a Licel detection system". What is the significance that the data acquisition electronics were manufactured by Licel? Also, a "detection system" would describe both optical and electrical components I would think. So in reference to Licel I would just call it the data acquisition system.
19. Line 123: Sentence starting "This change ..." This is indeed a very nice demonstration of how your background method tracks calibration changes. But I would appreciate some more discussion on why adjusting a coaxial cable would make any changes in the calibration at all. In order to have a dramatic change in the ratio as you show, an effective change in optical/electrical efficiency of one of the channels would have to occur. So please describe briefly how adjusting the coax cable results in such a change. It would be much easier to see this large change in ratio as being due to a change in some optical component, for example, or in a change in a discriminator setting and not just due to changing a cable.
20. Figure 2: typo "GRAUN"
21. Line 128: concerning "much more stable after 2013". Can you offer an explanation for why this is so?
22. Line 130: as noted earlier I suggest introducing the abbreviation OEM when first mentioned.
23. Line 130: "Only digital channel measurements were used ..." I suggest that you add somewhere a brief explanation that the Licel electronics offer two simultaneous measurements of the same signal - one using photon counting, the other using voltage measurements determined through an analog to digital converter. Please note that both of these signals are "digital signals" so these signals are usually referred to as photon counting and analog signals.
24. Line 132: "1 instead" is this a typo?
25. Line 134: I would delete the mention of Licel and reword to not use "using" twice in rapid succession. For example, "The measurements for the years 2011 to 2014 were processed first using the externally determined calibration constants and secondly ..."
26. Line 135: Awkward construction. Suggest rewording sentence starting "During the data processing employing external calibration coefficients, these coefficients were interpolated ..." to "The externally determined calibration coefficients were interpolated ..."
27. Line 138: You state "...the calibration constant $C^*(t_0)$ was computed by averaging over the 5 to 8 km altitude range." I am confused. You are referring to the background calibration technique in this sentence and the value of C^* is determined above 55 km, right? It is for the external calibration technique that you use a comparison between 5-8 km. Correct?
28. Line 139: you refer to the "cost associated with the OEM retrieval process". Please add some explanation for what this cost function is.
29. Line 140: Ditto. What are these cost values and why do they imply they should be eliminated. Since you are relying so much on this earlier OEM retrieval work, I think you should add a section describing the OEM details enough so that readers understand the technique along with the cost function and its range of values.
30. Line 140: Instead of "Each dataset" can you name the datasets to avoid confusion? Do you mean externally calibrated and solar background calibrated?
31. Line 142: reference is made to a "measurement response function". Please explain what this is.
32. Line 144: reference is made to "homogenized radiosonde measurements". What are these? Please explain.

33. Line 144: you also refer to “GRUAN certified radiosondes”. This is a little misleading it seems to me. GRUAN reprocesses the standard radiosonde products to create a GRUAN certified data product. The radiosondes themselves are the same either way. So the radiosondes are not certified but rather the final processed profiles are certified. You also state that these are “special soundings”. I don’t think this is right. The difference is that you have sent the standard data products to GRUAN and received re-processed data products. Right? I would make all of this more clear when you explain the difference between homogenized and GRUAN soundings.
34. Line 147: I think instead of “correlating” you mean “corresponding”?
35. Line 147: “data set”. Please be consistent throughout the paper either “dataset” or “data set”. Personally, I don’t care which.
36. Line 151: the statement “...an underestimation of the lidar temperatures in this range” is unclear to me. Suggest : “an underestimation of the lidar-derived temperatures with respect to the radiosonde measured temperatures”. Also, I wonder if after including standard deviation statistics in Table 1 this underestimation will be statistically significant...same statement pertains to other assessments of underestimations later in the paper.
37. Line 161: Suggest change “(and unanticipated)” to “(and possibly unanticipated)”