Review of «Calibration of a Water Vapour Lidar using a Radiosonde Trajectory Method » Shannon Hicks-Jalali et al.

General comments

I think this study is interesting. The implementation of the methodology is justified. The method is now well described and allows members of the lidar community to test it if they wish. The results show a relative improvement and details under which conditions. As the authors point out, any positive improvement is important when it comes to conducting trend studies in UTLS. The effort on the calculation of uncertainty is to be stressed and essential in view of this perspective. I welcome the great effort made by the authors to respond to my first report. The work done has been substantial. The article is much better structured, the speech is clearer and all the figures commented. There are still a few typos and sentences to be clarified that I point out in the "Specific comments" section of this report.

I still have one more question mark left. The authors clearly state that their method is "more accurate" (p1, l 23) and justify it in the summary by citing the 20% improvement for heterogeneous nights (p1, l 19). However, this « 20% » refers to a result on a single profile and on one specific altitude (p16, l 4, 18.07.2012, at 3.8 km). In Section 5, when the authors comments the bias for the heterogeneous nights, there are a lot of different values: "15-20 %" (p17, l 9), "5 %" (p17, l 10), "10 %" (p18, l 3), "reduced from -25 % to 10 % … from +20 % to +10 %" (p18, lines 5 and 6), "up to 15 %" (p18, l 9), "by 5-10 %" (p25, l 2). It is difficult to conclude with this section which value should be statistically retained and this "20%" in the abstract seems to be overstated. On the other hand, a 4% uncertainty (p1, l 23) is presented for the trajectory technique but there is no uncertainty stated for the traditional method for comparison. While I have noted that the authors have made it clear to me that the term "more accurate" does not refer to the comparative uncertainties of the two methods, this conclusion in the abstract appears directly after quantifying the uncertainty of the trajectory (p1, l 23). This might confuse the reader.

After careful reading of the article, here are the other main comparative elements between the two methods:

- there is a theoretical improvement. The idea of taking into account the movement of air masses is indeed in theory an improvement but it remains qualitative.
- the average difference between calibration coefficients calculated with both methods is 0.4 % for homognenous nights (p24, l 23) and 2 % for heterogeneous nights (p24, l 24).
- the validation of this method (Sect. 5) by looking at the average bias between the lidar profile and the "reference" profile shows that it is about 1% for the trajectory method against 0.7% (p19, l 4) for the traditional method (considering the absolute value) (p19, l 2). It is therefore lower with the traditional method.
- "The standard deviation of the ensemble of percent difference profile between both calibration methods and the trajectory radiosonde shows that the trajectory method has less variability with respect to the radiosonde profile above 2km" (p19, l 5 to 7).
- The σ_{err} on the calibration coefficient is 4.5% for both methods (p24, l 8 and 9).

This means that in the article, statistically, the only parameter that shows an improvement is the standard deviation for the trajectory method which is lower than the standard deviation for traditional method and it is for an altitude range between 2.5 and 4 km (p19, l 6, and Fig. 8). The authors say "less variability" but it is not quantified.

Therefore, it seems that the authors overstate a little bit the added value of this methodology in the abstract and when it comes to concluding. I think that this could be reviewed by the authors. That being said, all the quantified data are present in the article, it could also be up to the readers to evaluate their contributions. But the authors take the risk that the reader will share this view.

After correcting the points listed below, I think the publication of the article in AMT will be relevant.

Specific comments

Page 1 - Abstract
Line 10: Could you clarify the term "frontal boundaries"?
Line 19: You should specify what this "up to 20%" corresponds to so that this value can be found in the text.

Page 2
Line 27: I suggest replacing the word "GPS" with "GNSS".

Page 3
Lines 4 and 5: Do not the references only refer to RS92S? In this case, this should be specified.
Lines 22 "travelled" and 23 "traveled": Please standardize.

Page 4
Line 8: Please add a reference.

Page 5
Line 16 "a fog,clouds": A space is missing.

Page 8
Lines 6 and 7 "The radiosonde relative humidity […] 2014).": This should appear in Sect. 2.1.

Page 14
Line 11 "30 minute": I would have written 30-minute.

Page 15
Line 4 "will produce": Did not you mean "should"?
Line 4 "homogeneous nights": Figure 5 shows the 2012-07-27 time series. According to Table 1, this night corresponds to a heterogeneous night. Please reconcile.

Page 16
Lines 7 to 9 "is reduced to oscillate around 0 %": I do not understand what you mean. Please clarify.
Line 14: Please refer to Table 1 after "1 %".

Page 19
Line 6 "less variability": Please quantify
Line 7: I suggest adding a small conclusion on this comparative part (Sect. 5). The improvements shown refers to the altitude range 2.5 - 4 km and highlight the values that quantify it.

Page 21
Line 10 "mixing ratio ,": There is a space to delete.
Line 20 ".1 and .3 K": I think that some zeros are missing. Please correct it.

Page 22
Lines 6 and 7: There are ‘’ and ". Please standardize.

Page 23
Line 15: The average total accuracy is stated to be 4 % for both method. The dominant uncertainty is the radiosonde uncertainty of 4%. Table 2 shows that there are other uncertainties. Are you sure that the total is 4%.

Page 24
Lines 26 to 28: Why the discussion on these two uncertainties is at this point when the other uncertainties are discussed in the list that begins l 4 on p25?

Page 25
Line 2 "by 5-10 %": Would not it be better to use this 5-10% value in the abstract instead of "up to 20 %"?
Line 8: You should precise the average difference for the traditional method.
Line 10: This should be precised that it is between 2 and 4 km.
Line 12: In Sect. 6, 4.5% corresponds to the standard deviation (p24, l 9) and 4% to the total uncertainty on calibration (p23, l 15). In this summary and Table 1, it is the uncertainty on lidar photo counting and radiosonde mixing ratios whereas the authors find a total uncertainty of 4% (p23, l 7). Please reconcile.
Line 20: I have the impression that these values appear for the first time in Sect. 7, why were they not brought into the dedicated part, namely Sect. 6?

Page 26
Table 2: You should put the total uncertainty. You could differentiate the average uncertainties according to the calibration method.

Page 27
Lines 12 to 14: You say that the trajectory method "removes the representative uncertainty" and then that this uncertainty is considered "to be small". This means that it is not deleted. Please reconcile.
Line 19 & line 24: Be careful you have inverted the letters from GRUAN to "GRAUN".

Page 28
Line 5: This should be precised that it is between 2.5 and 4 km and not only "below 4 km".
Line 12: I am puzzled by the term "reference instrument". For me, the reference instrument would be a third instrument considered as the reference that would be used to validate this method. The profile calibrated with RS92S would then be compared to the profile measured by this reference instrument. I suggest replacing this term with another that would mean: the instrument used to calibrate.