Atmos. Meas. Tech. Discuss.,
doi:10.5194/amt-2017-135-RC2, 2017

Atmospheric Measurement Techniques Discussions

Interactive comment on "Quality assessment of integrated water vapour measurements at St. Petersburg site, Russia: FTIR vs. MW and GPS techniques" *by* Yana A. Virolainen et al.

Anonymous Referee #2

Received and published: 7 September 2017

The paper is well written and organized with a very good background discussion that helps put the work into historical perspective with past efforts of a similar nature. The work itself is well done and follows similar past work. This is the weakness of the paper because the work follows so closely to Virolainen et al., 2016 only with a different locale that allowed inclusion of the GPS data. No new methods for the comparisons are proposed in the current work nor additional insight into the problem is noted. Thus, the paper is a modest extension of past work. I recommend the work be published with the hope that the authors will consider including a few statements clearly highlighting how the current work goes beyond the current literature.

The comments below are suggestions I think would help the paper, but I leave it to the authors to decide whether they need to be implemented.

The abstract does a good job of summarizing the work including quantitative results.

The paper's background is very thorough and the requirements for IWV retrieval accuracies and limitations of the various methods help give the paper perspective.

The paper does an excellent job of referencing past work that helps put the current work into perspective. In fact, too good of a job since the current paper follows very closely to Virolainen et al. 2016.

Methods are well described and follow past techniques, thus no technical issues in the current work.

In figure 5, I disagree that the comparison plot needs to be in log scale in order to cover IWV from 1-30 mm. The log scale will always make the fit look appreciably better than it is. At the very least, there should be some indication of the goodness of fit to straight line. This appears to some extent in the tables and follow-on discussions but it would be better to include this information with the figure. The 0.02 standard deviation of the fit and the large number of data indicate that the slope is statistically different from the a 1:1 line, but a direct comment to this effect would be helpful. Also included in the statistical discussion should be whether the slope and standard deviation of the current method are truly different from the 1.06 slope retrieved by Buehler. At first glance, there does not seem to be a statistical difference.

An interpretation of the log-normal result for the A to M retrieval difference should be given as well as the logic for using only a multiplicative correction given that the offset of the fit is 0.14 mm. It does not seem as though using the full fit would increase the complexity of an A to M harmonization.

Again, I disagree with the use of a log-log scale for Figure 7. More important is that the labeling on Figure 7 does not make it clear which IWVs are being compared. The text allows me to infer that the lower right is MW versus GPS but the other two are not easily figured out from the text. It would be better to label both the x and y axes.

From the results presented, it is not clear that that the FTIR is the clear better answer for low IWVs. It is clear that the GPS disagrees with the MW and FTIR data, but it could also be concluded that the MW data performs better at low IWV since the variability of both the FTIR-MW and FTIR-GPS data show similar values. This could imply that the FTIR is the root of cause of the large variability. The authors should provide clearer justification for why they conclude FTIR is the method of choice for low IWV.

The approach used in Section 3.3 is not providing an accuracy assessment as much as a relative uncertainty between the three methods. The terminology used in the summary as it being an empirically based upperbound of statistical measurement errors is much more correct.

For the uncertainty assessment, selection of dates for which the IWV varied by small amounts (<1% for example) would imply a spatial homogeneity and would provide data for which the assumption of spatial homogeneity would be better than simply neglecting the misalignment error. The authors should point out if there are not enough dates with little variability or why the uncertainties from those dates are not suitable for the assessment of the overall uncertainty.

The paper provides a wealth of information, both from the current work and reference to past efforts. The difficulty for the reader is locating the key important points of the new results. For instance, are the authors recommending that FTIR users correct A and M retrieval differences and do they propose their approach as the best method? Are the estimated uncertainties acceptable, do they agree with past work, are there recommendations for which method is more suitable? These questions are covered with the results in the paper, but are not easily found.

One last comment is that the authors have an opportunity to discuss whether all three methods can be viewed as equally suitable for IWV and do not. That is, do the authors recommend that a network of IWV instrumentation rely on a single measurment approach or can it use any of the three methods? Is the capability of MW and GPS measuring under cloudy conditions worth the added scatter in the measurments? Is a zenith-viewing MW measuremnt with higher scatter better than a solar-path based FTIR measurement with lower scatter? Including such insight into the paper would improve its relevance to the broader community and take it past being only another collection of IWV retrieval results.

Interactive comment on Atmos. Meas. Tech. Discuss., doi:10.5194/amt-2017-135, 2017.