Atmos. Meas. Tech. Discuss., 7, C1857–C1859, 2014
www.atmos-meas-tech-discuss.net/7/C1857/2014/

Atmospheric Measurement Techniques
Discussions

Interactive comment on "The AquaVIT-1 intercomparison of atmospheric water vapor measurement techniques" *by* D. W. Fahey et al.

Anonymous Referee #2

Received and published: 24 July 2014

The manuscript as written discusses an important ground facility based blind intercomparison activity held at the AIDA chamber in Karlsruhe, Germany. In this intercomparison, many different instruments were intercompared, including a "core" set which have extensive heritage and experience measuring water vapor in the UT/LS and TTL, from airborne platforms. The impetus for this intercomparison was, in large measure, the large disagreements among these instruments in prior airborne campaigns. One primary result of the intercomparison study, reported for the first time in this manuscript, is that the differences observed in flight cannot be explained by the performance in the lab setting at the AIDA chamber.

The paper is well written and clear. The appendices, which describe each instrument, add valuable information to the understanding of this activity. Many instruments were

required to operate, by the nature of the facility, under conditions not identical to those experienced in flight. Conditions within the chamber ranged from those found in the UT/LS/TTL to some that are not generally found in the earth's atmosphere. For example, conditions included water vapor mixing ratios below 1 ppm at pressures exceeding 150 hPa.

A statement is made at the end of section 6.3.1, "Although mixing ratios in this range occur rarely in the UT/LS, these measurements help define the detection limits and performance limits of the instruments." This statement may not be true in general, and in some specific cases during this intercomparison. Many instruments are optimized for performance in certain regimes while possibly sacrificing performance in other, less important, regimes. To force all instruments to make measurements under conditions never encountered does not necessarily generate information on their detection limits, accuracy, precision, and performance under more realistic conditions. And such a strategy may well penalize some instruments that are so optimized (for example, optical measurements) with regard to instruments that are not optimized. I suggest rewording or removing the statement.

The lack of an agreed-upon reference causes additional concerns, as each instrument is being asked the question how well does it agree with the others, not the more important question – how accurate is the measurement being made. It is understandable that the collective reference strategy used in this intercomparison was the only fair choice, but it should be made very clear that agreement with the ensemble is not intended to imply accuracy in the strictest sense.

Regarding the collective reference method, the method for determining precision is not as clear as it could be. At the end of section 6.2 (in subsections d and e) a method is described which uses the ensemble mean slope vs time as a reference. From the difference between an individual instrument's measurement during that segment and the ensemble mean, one would expect to see a contribution not only from the Gaussian width of the distribution but also from the varying offset (unless all slopes are the same). This additional difference can't, in general, be lumped into the Gaussian, but it is not clear how the authors handled that time-varying offset. The actual equation used is given in the caption for Figure 4 and it doesn't seem to include the offset. I suggest making these sub-sections more clear and perhaps inserting the appropriate equation(s) in the main text body.

Some of the figures (notably 4-6) are too small to be easily read and understood. I suggest remaking those figures with readability in mind.

Interactive comment on Atmos. Meas. Tech. Discuss., 7, 3159, 2014.