Review of "High accuracy measurements of dry mole fractions of carbon dioxide and methane in humid air" by Rella et al.

Submitted to Atmospheric Measurement Techniques

Summary:

This paper presents a number of tests of the Picarro cavity ring down spectrometer (CRDS) instruments to characterize the transferability and stability of the water vapor correction applied to humid air measurements in order to calculate dry mole fractions. Five different institutions contributed data from their lab and field tests using several different methods. The authors conclude that at relatively low water vapor concentrations (<=1%), the manufacturer-supplied correction factors are sufficient to meet WMO/GAW standards and that by doing additional characterization on individual instruments, periodically over their lifetime, this range is extended to >4%.

General comments:

Overall, it is a much-needed publication regarding the use of a popular CO_2 and CH_4 analyzer. Everyone wants to know if they have to dry the humid air, or if not, how often do they need to test the water vapor correction on their individual analyzer. This paper addresses those questions and makes recommendations for instrument calibration based on the humidity levels it is operated in. The subject of this paper is appropriate for publication in AMT and this special issue.

However, because so many research groups contributed to this effort, the presentation feels pieced together. The paper could benefit from standardizing the data presentation in the figures and the evaluation metrics used to make it easier to read. The authors may also consider including a table of the correction coefficients determined for each analyzer. That way other users may easily determine whether their own water vapor correction experiment yielded results that are in the right ballpark.

More importantly, there are a few questions that they are poised to answer that they don't.
1. What is the variability in calibration coefficients from instrument-to-instrument? Or, if one applies the Chen et al 2010 coefficients to any instrument, how big can we expect the errors to be? This should be included in the conclusions section.
2. Also, there is one other source of noise that is not really addressed by the authors. That is the uncertainty in the coefficients 'a' and 'b' because of the noise in CO_2 and CH_4 measurements. How well can we expect to measure 'a' and 'b' on any given instrument?
3. How do these recommendations change if the goal is 0.05ppm CO_2 (Southern Hemisphere GAW goal)?
4. If the correction coefficients do change over time, are they linear? If not, how does one correct the time series?

Specific comments:

P5825, L6:   Southern hemisphere target = 0.05ppm CO2.
P5829, L4:   Please define 'cross-talk'.
P5831, L21:   'Similarly, the range of water vapor…'
P5834, L12:   'The values for r1 and r2 in ambient air were…'
P5835, L11:   'The correction coefficients for Eqn 4 determined…'
P5835, L18:   I believe there is an error in the noise equation. It should not decrease with water vapor level.
$$\sigma_{corr} = \sqrt{[-C_{wet}(a+2bH)\sigma_H]^2}$$
Also note what ambient values of CO2, CH4 and H2O are used in this estimate. Assuming 400 ppm CO2, 1900 ppb CH4 and 3% H2O, I get values the same as, or very close, to those you list in the text.
What about the noise in the correction coefficients due to instrument noise in CO2 and CH4 during the water vapor correction experiment?
P5837, L6:   Figure shows that no switching between wet/dry streams is needed. There are 2 analyzers.
P5837, L10:   What H2O levels were tested?
P5837, L19:   Need to mention wet/dry stream switching here.
P5837, L25:   Are 1 or 2 analyzers used in this NOAA implementation?
P5838, L17:   LSCE and NOAA used 1 analyzer, MPI used 2.
P5838, L20:   Was the dew point generator water acidified?
P5839, L16:   Has the wet hydrophobic particulate filter ever been tested for CO2 effects? Any wet surface has the potential to modify CO2.
P5818, L18:   How were all the instruments adjusted to the same water scale? Were they compared against the 'golden instrument' used in Chen et al?
P5847, L23:   This is why you should note the concentrations of CO2 and CH4 used in the water vapor correction experiments presented in this paper.
P5850, L6:   Comment on the offsets and variability of the comparison.
P5850, L28:   Drift or noise?
P5851, L9:   Not biasing under these humidity conditions. Most of these time series comparisons are only done in the 0-1.5% H2O range.
P5851, L23:   Note the temperature and pressure conditions of the Nafion dryer.
P5851, L27:   What water vapor correction coefficients were used in this experiment? Dilution by H2O = 0.04-0.15% still needs to be accounted for.
P5853, L14:   This is why the data should be H2O corrected.
P5855, L3:   This is a big spread. Comment on the scatter of the observations > WMO.
P5857, L21:   Within compatibility targets up to ~1.2% H2O.

Supplementary material:
        Needs some introduction on what you are testing, pH, CO2/CH4 degassing and why. Also where these tests were done. Picarro or NOAA?
        Fig S2, right and left panel descriptions are reversed. Explain the addition of dry air and the step change in each droplet experiment. Point out CO2 degassing on the figure panel.

Fig 1: List labs that used this setup. MPI, NOAA, LSCE.

Fig 2: Need legend. Pick a different symbol for the injection T. Same as excess flow now. List labs (MPI/NOAA, empa, lsce, picarro)

Fig 5: Standardize the label for the CDRS, CFADS, Picarro.

Fig 7 & 8: CO_2 and CH_4 concentrations used in these experiments?

Fig 8: '(without rescaling water vapor)'

Fig 9: Standardize y-axis labels. LSCE results. Over 20 days and 5 replications.

Fig 11: Relative to Chen et al would make these comparable with previous plots.

Fig 12: Comparison of flask measurements with what?

Fig 13: Flow rates?

Fig 14: Why is there so much noise in the target tank? 'as well as the difference (grey…'

Fig 15: In histograms, mark zero in red, not the mean. Plot difference vs H_2O?

Fig 16: PAL color is hard to see.

Fig 17: CO_2?

Fig 18: How does this look if the PAL instrument is corrected every 9.5 h or daily, etc. for drift like the WCC instrument?

Fig 19: Mention GAW target for CH_4.

Fig 23: legend?

Fig 24: Hourly averages? Dried minus wet or wet minus dried? Scale makes it hard to see trends. Can zoom in on y-axis.

Fig 25: Why not average both at 30s?

Fig 26: The slopes, not the medians are important here. Medians are the same in all figures (vs H_2O, CO_2, time), no extra information is presented this way.