Atmos. Chem. Phys. Discuss., 8, S5296–S5298, 2008
www.atmos-chem-phys-discuss.net/8/S5296/2008/

Atmospheric Chemistry and Physics Discussions

Interactive comment on “Long-term field performance of a tunable diode laser absorption spectrometer for analysis of carbon isotopes of CO_2 in forest air” *by* S. M. Schaeffer et al.

S. M. Schaeffer et al.

Received and published: 25 July 2008

We thank the referee for helpful comments on the manuscript.

This methodological paper demonstrates that TDL, and in fact all field-based trace gas measurements, must clearly present the method of estimation of precision and accuracy. The use of an “unknown” cylinder for QC is extremely valuable. This is an important contribution. There are some details that I questioned.

1. It seems misleading to state (page 9543, line 11) that your measurements are 1 second observations, when those are actually 1 second means of many observations per second. Can you clarify exactly how you are averaging up your data?

Text was changed to “Each reported measurement is the mean of 10 (or at a minimum 6, see below) one-second observations from the TDLAS (X10), and each 1-sec value was calculated internally by the instrument as the mean of ten 100-msec values”

2. Same page, please explain why you are willing to toss out 4 of your one second observations but not 3 or 5. It seems arbitrary. This paper helps our science because it helps us take TDL QC away from arbitrary decisions, but this decision appears arbitrary. Also, why is one std deviation your cutoff? Why not 1.5 or 2.0 std dev? Also, (line 24, same page) why was this processing applied to the calculated CO2 and 13C values rather than the isotopologues? This paper will be used as a methods reference in the future so clear explanation of how and why you made these choices will increase the value of the paper to future TDL users.

Our procedure is entirely arbitrary – there are no guidelines for how to do this, we just tried a few things which provided better performance. We were (arbitrarily) uncomfortable rejecting more than 4 out of 10 measurements, but did allow 0-4 outliers to be rejected as discussed (the original text reads “Values were discarded starting with the greatest absolute difference from the mean, and continuing up to a maximum of four values”). We used CO2 and d13C values because it was simplest to code up (this scheme was added to many thousands of lines of existing data processing code that we have developed over several years). While others may use our approach as a general guide, we would be very surprised if others adopted our exact procedure and not some other procedure that they prefer for their own reasons.

3. Page 9544, lines 10-15, this is a very nice explanation of why RMS error is a useful first-cut measure of bad performance but not of good performance.

Thanks – no changes needed here.

4. It seems there can be debate about cut-offs for tossing data and for calculating precision. For example, choosing sigma10 cutoffs of 0.55 for CO2 (fig. 2a,c) results

ACPD

8, S5296–S5298, 2008

Interactive Comment

in a somewhat low reportable precision but includes almost all of the data. Using a smaller value, e.g. 0.4 or 0.3, results in a higher reportable precision and tosses out more data. I would like to see more discussion and explanation for the basis of your decision beyond that 0.55 (for CO_2 in this case) is where the data flattens out. Is there an engineering or statistical basis for this? The same holds for the unknown cylinders (Figure 3). Why not toss more data by having a tighter threshold than 0.5 (difference from actual)? It could help the reader to understand the basis of these decisions and the background on what are the costs and benefits of choosing different values for cutoffs.

Again, there are no guidelines (that we are aware of) for how to make high-quality measurements. In this paper we are using our experience as scientists to evaluate how well an instrument performs for a particular application that is useful to us. Here we have shown the data and discussed exactly what we did and how we did it. We just selected thresholds that rejected only about 1% of the data (0.55 for CO_2, etc) – the text is clear about this. Of course we need to treat all the data the same way (the unknown air and the QC cylinders) to provide an unbiased measure of how well we’re doing for the unknowns. If we selected different criteria for the QC tanks than the unknowns, we’d be cheating.

We added the following text to acknowledge the arbitrary decisions we have made “Exact values of the thresholds (described below) are arbitrary and likely to be different for different TDLAS instruments and different applications. Hopefully these can serve as a guide for other researchers to develop their own criteria for data quality. ”

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 9531, 2008.

ACPD

8, S5296–S5298, 2008

Interactive Comment