Atmos. Chem. Phys. Discuss., 7, S3642–S3645, 2007
www.atmos-chem-phys-discuss.net/7/S3642/2007/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Intercomparison of ground-based ozone and NO_2 measurements during the MANTRA 2004 campaign" *by* A. Fraser et al.

H. Roscoe (Referee)

hkro@bas.ac.uk

Received and published: 2 August 2007

This is a worthy technical paper that is important in our quest to improve and standardise the various instruments making ground-based measurements of stratospheric trace gases. As such it should clearly be published after revision. Unfortunately, I fear that some of the revision may be major.

Major comments:

1. The authors use the Windoas package for spectral analysis, and acknowledge Caroline Fayt and Michel Van Roozendael for supplying it, but they do not recall giving any help in setting up Windoas for this project (Van Roozendael, personal communication,

August 2007). Hence it was used in some standard form.

This is probably why the results in Figure 1 are poor at the lower light levels of SZA$>$92° for the SPS instrument, and for SAOZ at SZA$>$94° for ozone and SZA$>$93° for NO2. This behaviour is typical if the analysis used too many windows in the wavelength and slit-function calibration. If 10 are used the S/N ratio must be very good, and as few as 5 might be necessary at low S/N. If too many are used the wavelength calibration is poor, which results in large errors near Fraunhofer lines and so in large scatter.

The large offset shown for SAOZ in Figure 1a may also be caused by non-optimal settings in Windoas, though it more difficult to see why. It is conspicuous that all three analyses with SAOZ in Figure 3 are negative and have large standard deviations. Is it possible that wavelength calibrations were run on reference only and not spectra+ reference in Windoas? Or that the slit-function fitting was switched off? Or that a Gaussian was used for SAOZ which has a strong asymmetry in its slit function in its modern incarnation of a grating with 360 gr/mm? The manuscript gives no information about such details of Windoas implementation for this project.

I urge the authors to engage in a dialogue with Caroline Fayt and Michel Van Roozendael about these issues, with a sample of SAOZ data if not also SPS. They are already committed to enabling better use of Windoas with SAOZ data as part of the EU-funded project GEOMON, so significant co-operation is likely. The results of some test analyses should then either allow the authors to refute my speculation above, or convince them that they must reanalyse the whole campaign with different Windoas settings. The latter would be a lot of work, but I would hope the authors would embrace the task with good grace if some tests showed it to significantly improve the tendency towards NDACC certification.

2. The figure, text and table discussing slopes and intercepts of DSCD regressions use standard deviations when standard errors are surely more relevant. The purpose of these regressions is the search for bias in sensitivity or offset. The slopes and inter-

cepts averaged for the campaign must not exceed the bounds for NDACC certification cited on p10215 lines 11 to 16, so it is the error in the campaign average that matters. The standard deviation could only be relevant if the NDACC rules said they must not exceed the NDACC bounds even on one day, in which case it is the extreme values which matter (about ś2.5 standard deviations) and all instruments overwhelmingly fail NDACC certification.

Furthermore, the discussion of ozone seems to suggest that, for example, a regression slope of 1.13 ś0.15 falls below 1.03 and so qualifies. This is not obvious. In this example, assuming the quoted error is 1-sigma, the probability is only about 25% that 1.13 ś0.15 falls below 1.03, remembering that this is a one-sided t-test, not the usual two-sided. The NDACC rules as quoted do not discuss this point, but I would prefer a 68% probability (the usual 1-sigma value), if not 95%. Table 1 seems to embrace this notion for NO2 but without specifying percentages, whereas the text about ozone seems not to embrace it.

Minor comments:

1. p10215 line11 - surely diverging DSCDs would cause a slope of other than unity in the regression. Residuals that increase or decrease with SZA are a sign of a non-linear error in one of the data sets, such as that of SAOZ data at SZA $> 93^\circ$ in Figure 1a. In this case it should be conspicuous in the regression of SAOZ versus MAESTRO, which is unfortunately not the example chosen for display in Figure 2.

2. The alternating error in Figure 1b from UT-GBS at SZA 80 to 84° is most odd. It cannot simply be the small amounts of NO2 as the values from SZA 75 to 79° have similarly low scatter to those of SZA$>85^\circ$. Furthermore the alternating nature is systematic, not random, almost as though some spectral files have been wrongly labelled. The authors should check this again as it would greatly help Table 1 to include the results at these smaller SZAs.

3. The ultimate purpose of the NDACC methodology, of regressions of results from all

ACPD

7, S3642–S3645, 2007

EGU

pairs of instruments, is to show at least one pair with unity slope, negligible intercept and small residuals. Both of those instruments are then almost certainly of high quality. This then gives more meaning to the comparisons of others with either of them, in that the bias of the others can be found. This is not discussed in the manuscript, probably because no such pair exists for ozone, and because there are only two instruments for NO2. But mention should be made, and perhaps the reanalyses suggested above might create a high-quality pair of instruments for ozone so that a useful discussion of causes of bias in the others can follow.

4. The captions to Figures 3, 4, 5 and Table 1 do not define the error bars shown (you have to search it out in the text) and do not say if they are 1-sigma or 2-sigma (even in the text). They also use the normal symbols for standard error (error bars and ś), not for standard deviation. Table 2 correctly specifies standard deviation in unambiguous notation.

Editorial comments:

p10208 line19 - “issues” is personnel-speak for “problems” - why not be straightforward when the instruments are unlikely to read the manuscript and thereby take offence?

p10210 line2 - U of T FTS is not really an acronym, and is inconsistent with the UT-GBS of p10208. What is wrong with UT-FTS? If the UT team members cannot agree a common style of acronym, what hope is there for a common approach to statistics?

p10210 line5 - indium & mercury should be lower case, antimonide is spelt wrongly.

p10213 line6-10 - a table would be much better than this wealth of numbers in text.

p10213 line18 - the acronym “OEM” is widely used in electronics construction for something other than optimal estimation, and the acronym is avoided later (p10214 line11). It should be removed.

Interactive comment on Atmos. Chem. Phys. Discuss., 7, 10205, 2007.