Atmos. Chem. Phys. Discuss., 8, S7098–S7106, 2008
www.atmos-chem-phys-discuss.net/8/S7098/2008/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Validation of ozone measurements from the Atmospheric Chemistry Experiment (ACE)" *by* E. Dupuy et al.

E. Dupuy et al.

Received and published: 16 September 2008

We thank Dr. Chi for his interest in the ACE Validation Special Issue and his comments on this paper. We have responded to them below.

The authors have obviously put a great deal of work into this submission. But, the material is probably better suited for publication as a masters thesis or as an ACE technical note. Scientifically, the methodology is weak (being superseded by the approach of Rodgers and Connor, von Clarmann) and the paper is excessively long with little interpretation of the observed ozone differences.

As a positive, the paper is comprehensive in providing a first bias estimate for ACE-FTS/ ACE-MAESTRO ozone profiles by intercomparisons with other remote measurements. Tables 5, 6, 7 adequately summarize the author's results.

As mentioned in the responses to the Referees, this paper was construed as a reference document for the current data products from the two ACE instruments, before the public release of the data, using comparative data from as many instruments as were available. We believe that this study should be published in the open literature to ensure that this information will be available to future data users. While more rigorous approaches for validation methodology have been published recently, simplifications are often necessary to deal with large amounts of data. We have applied the method of Rodgers and Connor (2003), when appropriate averaging kernels were available. Additionally, we have used the approach of von Clarmann (2006) consistently in the calculation of de-biased standard deviations (see also responses to Anonymous Referee #2). This was not explicitly stated in the paper and has been added to the revised manuscript to address the Anonymous Referees' comments.

As noted by von Clarmann (2006): "while in real life it will not always be possible to apply these approaches at full rigorosity, validation scientists certainly will find workarounds and simplifications". We have tried to produce a timely paper (prior to the public release of the data) that will provide thorough bias characterization information for users of the ACE data products.

On the negative, the paper is 144 ACPD pages long (with 20 ACPD pages of citations) and has 47 separate figures. It is worth noting that with the exception of 'Eureka (DIAL)', 'Lauder MWR', and 'Mauna Loa MWR' the authors average the ground-based results when summarizing their work. Accordingly, I question the relevance of tables (2, 3, 4, 5) as they simply describe the geographic location of ground-based research sites. Likewise, many of the 47 figures are repetitive; showing the percent ozone difference between the FTS or MAESTRO with yet another limb profiling instrument. No attempt is made to interpret the differences shown for these individual plots. This paper could be made considerably shorter if the authors removed many of these tables/figures and made them available to interested readers on the ACE website.

The lists of the ozonesonde, lidar and FTIR stations given in Tables 2, 3, 4 were included to give important information on the relevant instruments and to ensure that we included a complete description of the datasets employed in our comparisons. It is customary to describe, for instance, the sensor type and station location for ozonesondes, and to acknowledge the operating institutes (for examples, see Table 1 of Cortesi et al. (2007) or of Nassar et al. (2008)). Table 5, on the other hand, gives the results of the FTIR comparisons and therefore is complementary to Table 7. As mentioned in our responses to the Anonymous Referees, the strength of this work is in using one or more comparison datasets from nearly 20 different instruments or sets of instruments for validating two instruments and showing that the systematic biases in both the instruments are robustly characterized. As such, we feel that all the figures (typically 2 per dataset) provide useful information and can not be reduced further. We also chose to use the same representation (quoting Anonymous Referee #2, “presented in figures with similar layout”) and to show results for each comparison dataset so as not to select only the comparisons most favorable to the ACE instruments. Disregarding some datasets would indeed compromise the significance of the comparison. Finally, as already mentioned, this paper is intended to provide a single reference for the current ACE ozone products. Because the ACE mission will eventually end, providing information on our website is not a suitable long-term method of providing information to the user community.

*With regards to the analysis, the approach is rather simplistic and, in many respects, lacks the rigor one would hope to see in a modern and comprehensive review of ozone measurements. As the authors note, ozone is a rather important trace gas and even small changes in the VMR profile are significant. Their method can effectively be summarized as repetitively computing (x1 - x2)/x2 * 100% where x1 is an interpolated FTS/MAESTRO profile and x2 is some other interpolated profile. There is no discussion of how outliers are handled. More importantly, there is little discussion of the FTS/MAESTRO measurement characterization, i.e. (1) vertical resolution, (2) the ex-*

pected random errors, (3) the expected systematic errors (see Rodgers, World Scientific, 2000; Rodgers and Connor, JGR, 2003).

As mentioned above, we thought it prudent to treat all the comparisons in a simplified yet consistent manner so as to be able to characterize the biases in an unambiguous way. The filtering criteria used to exclude unphysical profiles have been described in page 2528, l.12-24 of our ACPD paper. As stated already in the responses to the Anonymous Referees, the ACE retrievals do not rely on optimal estimation and cannot be characterized in a straightforward manner by application of the concepts described in the references given above. Using the fields-of-view of the instruments, the vertical resolution of the ACE-FTS and ACE-MAESTRO measurements have been discussed on pages 2520 (l.15-19) and 2521 (l.15-17), respectively, of our ACPD paper. As described in the responses to the Anonymous Referees, the discussion of the available error estimates has been expanded. Further, the expected random errors are only needed if the precision estimates are to be validated. Bias validation is possible without consideration of ex ante estimates of random errors (von Clarmann (2006), Eq. 32 and following line).

Considering the work cited above, many authors have adopted Rodgers' view of validation,
"A full error analysis and characterization is needed as a basis for any comparisons to be made..."
and intercomparisons,
"the purpose of an intercomparison is to determine whether different observing systems agree within their known limitations."

The detailed error budgets for either of the ACE instruments are not available for the current data products and they will be the focus of future validation efforts. In any case, the knowledge of the bias of an instrument is useful regardless of whether the systematic uncertainty had been predicted or not. The information needed to judge

ACPD
8, S7098–S7106, 2008

the significance of the bias is included in the standard error of the mean differences (plotted and described in the paper). Precision validation certainly needs random error estimates, but precision validation is not the main focus of our work. For clarification, we have added statements in the abstract and in the introduction and methodology sections of the revised paper explicitly mentioning the focus (bias determination) of our work.

It is therefore disappointing that an neither an error analysis nor characterization were presented in this work. While very well cited, there is surprisingly little interpretation of the comparison results. For each instrument, VMR and difference plots are shown and explicitly described in the text (including both seems redundant unless you are highlighting a particular region for subsequent interpretation/discussion). A generic (1 page) summary suggests systematic and random errors (smoothing, interfering species, diurnal changes, spectroscopic linelists, atmospheric variability, etc) that may have contributed to the observed difference or the standard deviation but it is purely speculative. I am left to assume that this is a consequence of an inadequate characterization of the systematic and random error terms. The authors are quick to attribute the errors to spectroscopy and/or the retrieved temperature errors but offer no quantification or plots showing the sensitivity of ozone retrievals to these fields. All the statistics given in this paper arise from considering the ensemble mean and standard deviation (implicitly assuming a Gaussian PDF) for the given spatial/temporal coincidence sets. A more comprehensive definition of bias (i.e. one that includes the ex ante estimate of the systematic and random terms) is given by von Clarmann (ACP, 6, 4311-4320, 2006).

I have similar concerns regarding the interpolation, smoothing/convolution being applied to the retrieved profiles, and the comparison of profiles that contain explicit a priori information.

Primarily, the discussion of the comparison results has been done in Sec. 7 (Summary

- discussion) to highlight the very good consistency of the mean relative differences from the various datasets. As stated in the text, investigations of the sources of the observed differences are on-going as part of the development of the next version of the ACE data products. Therefore preliminary results have been presented. The standard deviation and ensemble mean are defined for arbitrary distributions. Since we have not concluded that, for instance, 68% of the data are within plus minus one standard deviation, we have not assumed a Gaussian PDF. In the cited paper by von Clarmann (2006), the definition of the bias does not include the ex ante estimate of the systematic and random terms. It is explicitly mentioned that the ex ante estimates of the random error are not necessary to evaluate the bias and its significance. The ex ante estimates of the systematic error terms are only necessary to check if the actual bias is significantly larger than the estimated systematic errors (von Clarmann, 2006). Discussion of the interpolation and smoothing techniques used and the influence of the a priori information on the ACE profiles was provided in detail in the response to Anonymous Referee #2.

For limb comparisons, the FTS/MAESTRO profiles are interpolated quadratically (Boone et al., AO, 2005) and other profiles are linearly interpolated without any consideration of the measurement covariance (Migliorini et al., JGR, 2004; Calisesi et al., JGR, 2005). Likewise, when comparing with column amounts, FTS/MAESTRO profiles are integrated for a slant-column without any discussion/comment regarding the precision of the resulting quantity.

The ACE data products are provided on the as-measured tangent height grid and interpolated onto an even 1 km or 0.5 km grid for ACE-FTS or ACE-MAESTRO, respectively. Piecewise quadratic interpolation is used for ACE-FTS (Boone et al., 2005) and linear interpolation is used for ACE-MAESTRO (McElroy et al., 2007). These interpolation steps are done as part of the retrieval process. Linear interpolation is often used in validation studies for comparing data products with similar vertical resolutions with limited loss in significance of the results. The analysis techniques used in the papers cited above require information that was not available for all comparison datasets (e.g. measurement covariances). Also, these techniques are time-consuming and therefore not practical to apply routinely in a study of this size. Uncertainty estimates can be made for ACE-FTS and ACE-MAESTRO partial columns based on the fitting errors. However, these are difficult to compare with the complete error estimates of the ground-based FTIR partial columns.

With regards to smoothing, the formalism for dealing with these effects is well established and used within the NDSC community (Connor, JGR, 1994; Rodgers & Connor,JGR, 2003). Ironically, the authors correctly smooth the ground-based FTIR/MW observations but do not feel compelled to rigorously handle the FTS/MAESTRO profiles. Vertical sampling is described as varying with a beta angle (?) from 1.5 to 6 km but a constant 3 km smoothing is somehow deemed appropriate. A variety of techniques can be used to estimate the actual vertical smoothing for FTS/MAESTRO (i.e. perturbation methods as done on MIPAS, http://www.ifac.cnr.it/retrieval/documents/AK_report.pdf).

As discussed in the response to Anonymous Referee #2, the higher vertical resolution ACE retrievals have been degraded to the lower resolution of the ground-based observations (Rodgers and Connor, 2003), as described in Section 4. However, to do this correctly, the ACE-FTS and ACE-MAESTRO profiles have been smoothed rather than the ground-based FTIR and MWR results. The technique of Rodgers and Connor (2003) could not be applied for comparisons between ACE and higher vertical resolution observations because averaging kernels are not available for the ACE-FTS and ACE-MAESTRO measurements. We have used the geometry of the ACE measurements to estimate the smoothing. For ACE-FTS, the field-of-view of the instrument provides a maximum vertical resolution of $\sim$3-4 km at the limb and it is this vertical resolution we used for determining the width of the triangular convolution functions.

ACPD

8, S7098–S7106, 2008

The vertical spacing of the ACE-FTS measurements depends on the beta angle for the occultation (defined on page 2527, l.4 of our ACPD paper) and varies between 1.5 and 6 km. This provided the fixed grid for the ACE-FTS convolution functions. Perturbation studies are not necessary: the averaging kernel of an unconstrained retrieval is the gridwidth (as easily demonstrated by applying the formalism of Rodgers (2000) with a numerical constraint R=0) and this is what we have used. Perturbation studies would only be necessary if an averaging kernel reported on a grid finer than that of the retrieval was needed (i.e. response to delta perturbations). For more detail see also the responses to Anonymous Referee #2.

Finally, no attempt is made to remove a priori information when comparing with profiles retrieved using optimal estimation. Again, a variety of techniques exist (i.e. von Clarmann and Grabowski, ACP, 2007).

This technique is only applicable if the full retrieval covariance matrices have been stored and made available. This is usually not the case for satellite measurements, considering the amount of data that this would represent. In cases such as comparisons with MIPAS or MLS, the vertical resolutions of the ACE and comparison instruments are similar enough that neglect of resolution issues is justified. The consistency of the comparison results from instruments using optimal estimation with those that employ other retrieval methods also suggests that the impact of the a priori information is minimal.

Again, without a more comprehensive error characterization, it is difficult to assess whether the above methods are appropriate.

Please see above explanation and refer to the responses to the Anonymous Referees for clarification on the error characterization issue.

ACPD

8, S7098–S7106, 2008

Interactive Comment

References

Boone, C. D., Nassar, R., Walker, K. A., Rochon, Y., McLeod, S. D., Rinsland, C. P., and Bernath, P. F.: Retrievals for the Atmospheric Chemistry Experiment Fourier-transform spectrometer, Appl. Opt., 44(33), 7218–7231, 2005.

Cortesi, U., Lambert, J.-C., De Clercq, C., Bianchini, G., Blumenstock, T., Bracher, A., Castelli, E., Catoire, V., Chance, K. V., De Mazière, M., Demoulin, P., Godin-Beekmann, S., Jones, N., Jucks, K., Keim, C., Kerzenmacher, T., Kuellmann, H., Kuttippurath, J., Iarlori, M., Liu, G. Y., Liu, Y., McDermid, I. S., Meijer, Y. J., Mencaraglia, F., Mikuteit, S., Oelhaf, H., Piccolo, C., Pirre, M., Raspollini, P., Ravegnani, F., Reburn, W. J., Redaelli, G., Remedios, J. J., Sembhi, H., Smale, D., Steck, T., Taddei, A., Varotsos, C., Vigouroux, C., Waterfall, A., Wetze, G., and Wood, S.: Geophysical validation of MIPAS-ENVISAT operational ozone data, Atmos. Chem. Phys., 7, 4807–4867, 2007.

McElroy, C. T., Nowlan, C. R., Drummond, J. R., Bernath, P. F., Barton, D. V., Dufour, D. G., Midwinter, C., Hall, R. B., Ogyu, A., Ullberg, A., Wardle, D. I., Kar, J., Zou, J., Nichitiu, F., Boone, C. D., Walker, K. A., and Rowlands, N.: The ACE-MAESTRO instrument on SCISAT: Description, performance, and preliminary results, Appl. Opt., 46(20), 4341–4356, 2007.

Nassar, R., Logan, J. A., Worden, H. M., Megretskaia, I. A., Bowman, K. W., Osterman, G. B., Thompson, A. M., Tarasick, D. W., Austin, S., Claude, H., Dubey, M. K., Hocking, W. K., Johnson, B. J., Joseph, E., Merrill, J., Morris, G. A., Newchurch, M., Oltmans, S. J., Posny, F., Schmidlin, F. J., Vömel, H., Whiteman, D. N., and Witte, J. C.: Validation of Tropospheric Emission Spectrometer (TES) nadir ozone profiles using ozonesonde measurements, J. Geophys. Res., 113, D15S17, doi:10.1029/2007JD008819, 2008.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding – Theory and Practise, World Scientific (publisher), Vol. 2, *in: Series on Atmospheric, Oceanic and Planetary Physics, 2000*.

Rodgers, C. D., and Connor, B. J.: Intercomparison of remote sounding instruments, J. Geophys. Res., 108(D3), 4116, doi:10.1029/2002JD002299, 2003.

von Clarmann, T.: Validation of remotely sensed profiles of atmospheric state variables: strategies and terminology, Atmos. Chem. Phys., 6, 4311ñ-4320, 2006.

Interactive comment on Atmos. Chem. Phys. Discuss., 8, 2513, 2008.