Interactive comment on "Technical Note: A trace gas climatology derived from the Atmospheric Chemistry Experiment Fourier Transform Spectrometer dataset" by A. Jones et al.

A. Jones et al.

kwalker@atmosp.physics.utoronto.ca

Received and published: 23 April 2012

We thank reviewer #2 for his/her helpful comments. We have addressed the comments as listed below.

In this manuscript Jones et al. present a climatological dataset of the middle atmosphere which has been generated for 14 atmospheric trace gases on basis of ACE-FTS measurements. For a technical note the description of the dataset is adequate since it is concise and covers all necessary information. However, the applied scheme for data selection and averaging is not convincing

and needs further elaboration/explanation before publication in ACP is possible. The central part of the paper is the description of the 'filters' which have been applied to the original dataset together with the method of calculating the mean values from the selected data. On pages 29857 and 29858 it is described that the selection criteria are based on the data themselves. In principle a selection of data should be made on basis of other information than their values themselves (e.g. instrumental parameters indicating a problem, a bad fit of the spectra or convergence problems during the retrieval).

In detail:
'Unrealistically large or small (for example, very negative) VMRs are then removed.'
Could you specifiy this and give examples what this means. Which limits have been set here in positive and negative direction?

Extreme values were excluded to remove values that are well beyond what has been observed in the atmosphere. For example, ozone values that were less than -2 ppmv or greater than 20 ppmv have been removed. We estimate that this filtering removed data from less than 20 profiles (of more than 19,000) overall. We have clarified the following sentence.

"Unrealistically large or small VMRs are then removed, for example O_3 values that were less than -2 ppmv or greater than +20 ppmv. Overall, this has removed data from less than 20 profiles from the data set."

'...any measurement for which the fitting uncertainty is larger than retrieved value is ignored...'
Why has this been done? Also from many noisy data you can get realistic results by averaging.

No flags are provided with the current version of the ACE data to provide data quality information for these profiles. We use this method based on the absolute fitting uncertainty to identify issues that may be related to poor quality spectral fits or instrumental issues affecting the profiles. This filtering method has been employed with other ACE studies (see Dupuy et al., Atmos. Chem. Phys., 9, 287–343, 2009). The fitting uncertainty provides insight into the retrieval precision and a related exclusion method was employed in the HALOE climatology described in Grooß and Russell (Atmos. Chem. Phys., 5, 2797–2807, 2005). To clarify this, we have added the following sentence and corrected the text to indicate that we are using the absolute fitting uncertainties.

"A related exclusion method was used in the HALOE climatology (Grooß and Russell, 2005) and this particular filtering method has been employed in other ACE studies (e.g. Dupuy et al., 2009). "

'…any measurement with a fitting uncertainty that is smaller than 0.01% of the retrieved value is also ignored…'
The reason for this has to be explained.

We have found that in the retrievals of a few occultations the fitting procedure does not converge, yielding a profile that is very similar to the initial guess. These values are associated with very low measurement uncertainties. Since we do not have a flag to indicate if the fit converged, this filter serves to eliminate profiles with this convergence issue.

'Final filtering is performed by using a statistical technique, namely the median absolute deviation (or MAD).'
Why, after all these previous steps, is a further filter necessary?

The previous filtering steps have been employed to remove data that we believe are due to instrumental or retrieval issues. This use of the median absolute deviation or MAD is based on work of Toohey et al. (2010). These authors found that the MAD was a more robust method of reducing the impact of outliers when using measurements from ACE-FTS. Therefore we have added the following sentence.

"Toohey et al. (2010) have found the MAD to be an effective method of reducing the impact of outliers when using measurements from ACE-FTS."

Toohey, M., Strong, K., Bernath, P. F., Boone, C. D., Walker, K. A., Jonsson, A. I., and Shepherd, T. G.: Validating the reported random errors of ACE?FTS measurements, J. Geophys. Res., 115, D20304, doi:10.1029/2010JD014185, 2010.

'Since a climatology is the most probable state of the atmosphere for a given time and location, we must be able to remove observations that are considered non-representative of the most probable state.'
I do not agree with this reasoning. In fact, the ACE-FTS sampling is due to it's observation geometry rather sparse. If one measures even with this sparse sampling an 'unusual' profile, one could also argue that such a profile is in reality not as 'rare' as it seems through the ACE-FTS perspective. So this is not a reason to discard it. To substantiate that the applied filters are all really necessary it would be instructive to show as an example how the climatologies differ when the selection is stepwise switched on.

The MAD technique was used to produce the ACE-FTS NOy climatology (Jones et al., 2011). It was found that the influence of energetic particle precipitation produces large ranges of NOy values on an annual basis. The NOy zonal averages produced in regions of high EPP are still enhanced even after filtering was included. In addition,

the variance was extremely large. Here, MAD and other filtering should be included as the influence of a few outliers (even if they are real) can produce variances that may be deemed unrealistic compared to the most probably state for a given time and place. As noted above, we have used the approach of Toohey et al. (2010) in employing the MAD filtering technique.

Concerning the calculation of the mean of the remaining data the following is said:
'In making the climatologies, we take into account the quality of the data such that we weight each measurement by the inverse of their fitting uncertainty when calculating the mean.'
How has this weighting been done? In an Gaussian optimal sense, by weighting each data by the inverse of the square of its precision (here called 'fitting uncertainty') or really weighting by its precision only, as the text indicates? If the latter is the case, I would question this approach unless there are good reasons for its application.

We weight each measurement by the inverse of their "absolute" fitting uncertainty when calculating the mean. We have compared this method to using an un-weighted mean calculation and see that the former method is the more favorable technique. This is because the use of weighting provides more confidence in the final mean value with a lower fitting uncertainty, although not totally neglecting those measurements with a larger uncertainty. This is similar to the weighting approach taken by Grooß and Russell (2005) in their HALOE climatology. However, Grooß and Russell used the accuracies to calculate these weights. This has been clarified in the text.

"In their HALOE climatology, Grooß and Russell (2005) used the inverse of the accuracy to account for the quality of the data in their calculations."

C16536

'The total number of observations per grid cell can be used in these cases to determine the degree of statistical robustness.'
It would be very instructive to show, at least for one example, in the same format as the other plots (Fig. 4-11) the number of observations entering each bin.

Because the total number of observations per grid bin can vary significantly by species, season and altitude, it would be difficult to produce plots that would be representative of the entire climatology. As such, and because this information is available to the user in the data files, we would prefer not to include these additional plots in the paper.

Technical: Fig. 10: the figure numbering (ABCD) is missing.

This has been corrected.

———————————————