Atmos. Meas. Tech. Discuss., 4, C2543–C2546, 2012
www.atmos-meas-tech-discuss.net/4/C2543/2012/

Atmospheric Measurement Techniques Discussions

Interactive comment on "Global and long-term comparison of SCIAMACHY limb ozone profiles with correlative satellite data (2002–2008)" *by* S. Mieruch et al.

S. Mieruch et al.

weber@uni-bremen.de

Received and published: 2 January 2012

Reviewer #3: When validating data it is important to take the error estimates of the data into account. This is particularly important when discussing the significance of the results. In the manuscript the error estimates (random systematic) of SCIAMACHY data are not discussed at all. I would also like to see some discussion about the error estimates vs natural variability.

This is a good point. A full error characterization of the SCIAMACHY ozone profiles are in preparation and a paper is in preparation. We mention that the full error covariance

Full Screen / Esc
Printer-friendly Version
Interactive Discussion
Discussion Paper

(page 4881, line 1) is not available. From Eq. 8 it can be seen that the error used in the hypothesis testing is given by the standard deviation and includes all kinds of errors, random and systematic

Reviewer #3: The resolution of the measurements vary from instrument to instrument and this should be taken into account in the comparisons or at least commented if it does not play major role. The resolution of the solar occultation measurements seems to be much better than SCIAMACHY

The vertical resolutions between instruments vary between 2 and 4 km and do not differ significantly enough to warrant resolution matching. The use of averaging kernels to match vertical resolution is mostly appropriate when comparing data with markedly different vertical resolutions, e.g. ozone sondes and satellite data. Therefore, no resolution matching was applied here.

Reviewer #3: The statistical significance tests should be performed using extra attention in order o avoid false conclusions. In general, I find Bayesian statistical approach, with proper error characterization, more solid background for making conclusions than various statistical tests (but this is only my personal view). It is in particular important to distinguish what we are looking for: in the data validation we are typically more interested about finding systematic biases than 'statistically significant results'. As an example, bias at one altitude might not be statistically significant, but when the same behavior can be seen at several altitudes the systematic behavior is clear and may indicate some instrument/retrieval issues. Would it be possible to discuss the purpose of the significance tests in this light? In addition, the error estimates should be taken into account when applying significance tests.

Indeed, Bayesian statistics have a solid fundament, however they need prior information, which leads to a certain subjectivity. Further, the computational effort is sometimes enormous and sophisticated algorithms like MCMC are needed to estimate the posteriors. Statistically significant results are systematic, because the chance of being

AMTD

4, C2543–C2546, 2012

random is low. The problem of similar biases at several heights are due to correlations between altitudes and related to the smoothing errors in connection with regularization constraints. To answer such questions, geo-statistical methods like variogram estimation and Kriging are needed, which is beyond the scope of this study.

Reviewer #3: I find the conclusions that statistically significant trend is observed at 2002-05, but not at time period 2004-08 a bit confusing as the time periods overlap with 2 years.

The early and late period are defined by the availability of the correlative data (HALOE and SAGE in the earlier period and MLS and ACE-FTS in the later period). For trend estimates one should always use the longest available period available for the pair of measurements.

Reviewer #3: I don't think that for modern computers it is time consuming to find co-located measurements. In my opinion, the work could be motivated by obtaining better statistics when using zonally averaged comparisons. At high latitudes during eg. ozone hole conditions this is more complicated and representativeness might play role. This could also be commented.

We agree with the reviewer that computing time is not really a critical issue, however, direct comparisons of zonal means is much simpler and easier to do.

Reviewer #3: Difference in local time: what does it mean for ozone in the stratosphere? Is it better to compare night-day? (p. 4874, l. 6)

Diurnal variations are negligible in the lower to middle stratosphere, above 40 km there can be differences of a few percent. For this reason a tight time constraint (4h) for co-located measurements is applied here, except for SABER and MLS. In case of SABER, there are some retrieval issues due to the different wavelengths and approaches used in the air glow retrieval (daytime) and thermal IR retrieval (nighttime). We added the following sentence "The SABER retrieval in the thermal IR is considered more reliable."

AMTD

4, C2543–C2546, 2012

Reviewer #3: P4881 eq 12 – this seems to be variability and not errors. The error estimates of the data are not included as far as I can see.

Actually it contains both random errors and natural variability as pointed out later (see p. 4882, l. 11).

Reviewer #3: p.4882 – l8. Here m deg of freedom, but above month. Please, use other letter here for consistency.

The number of months is very close to the degree of freedom

Reviewer #3: The impact of solar cycle should be discussed in the context of observed/not observed trends.

We clearly stated that we are not interested in the scientific interpretation of trend contributions and we think for trend comparisons between different instruments it is sufficient to assume that solar cycle contributions are part of the linear trend term (descending phase of solar cycle 23).

Reviewer #3: Figures 9-11 are rather unclear ...

The significance of single alt/lat grid cells has little to do with visual patterns.

Reviewer #3: I would like to encourage the authors to include the other Envisat instruments (GOMOS and MIPAS) to the comparison since they have the exactly same time period of measurements. This would probably improve the interpretation of the results.

As mentioned in the conclusion it is planned to include more instruments in the comparisons for the validation of updated SCIAMACHY limb profiles in the near future.

Interactive comment on Atmos. Meas. Tech. Discuss., 4, 4867, 2011.