Atmos. Chem. Phys. Discuss., 5, S1061–S1068, 2005
www.atmos-chem-phys.org/acpd/5/S1061/
European Geosciences Union

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Comparisons between SCIAMACHY and ground-based FTIR data for total columns of CO, CH_4, CO_2 and N_2O" *by* B. Dils et al.

Anonymous Referee #1

Received and published: 4 June 2005

General comments

It is very important to perform a proper validation of trace gas measurements from satellites before using these measurements for scientific studies. This is particularly true for greenhouse gas observations, which have only recently become available, and for which highly demanding precision requirements are set. The present study is a first attempt to validate measurements of CO, CH_4, CO_2, and N_2O from the NIR channels of SCIAMACHY by using a network of ground-based FTIR instruments. FTIRs can measure the total column of the respective trace gases, and thus constitute one of the few possibilities to directly validate the (total-column) satellite data. Therefore, a study as presented here is an excellent means to determine the quality of the satellite

measurements. There are, however, a number of problems with the present paper, which should be addressed before it can be published.

Firstly, using FTIR for global validation of satellite measurements of greenhouse gases is not an easy task.

- The FTIRs are often located on high mountains, thus sample a considerably smaller part of the atmospheric column than the co-located satellite measurements.
- The FTIR network is not uniform in the sense that measurements at the different stations are done in a different way, possibly leading to systematic biases between the stations.
- The FTIR network gives a poor global and even latitudinal coverage, so that statements on the global quality of the satellite data are hard to make.

The authors acknowledge these issues, and do some efforts to overcome them. Still, I feel that more work should be done to address the first two points.

A second shortcoming of this study is that the data sets from the three satellite retrieval algorithms cover only part (and different parts) of the target year 2003. This makes intercomparison and discrimination between the algorithms very difficult.

Thirdly, the statistical analysis methods used in the paper should be presented and explained more clearly (e.g. equation 4, see specific comments).

Finally, many of the differences that are found between the satellite data and the FTIRs are washed away as being non-significant. For example: figure 8 shows the biases in CH_4 between the satellite retrievals and the FTIRs, and these are concluded to be 'very small' (P2693, L7). However, if these biases were related to the overall global variation in CH_4 concentration, the conclusion would actually be that they are very large.

ACPD

5, S1061–S1068, 2005

Summarizing: on the basis of this study, I am not convinced that 'the products are useful for qualitative geophysical studies on a global scale' (P2679, L15-16), although I'm not completely sure what this statement means. If the analysis is improved along the lines above, there is a good chance that some more specific conlusions can be drawn.

Specific comments

P2682, L9-10: The inter-hemispheric gradient in CH_4 *columns* should not be compared with the gradient in *surface* concentrations. Also: could the authors give a reference of models which predict a lower gradient?

P2683, L4-7: Is it not better to perform the polynomial fit on the individual data points? When first a daily average is taken, a day with only one measurement will receive the same weight in the fitting as a day with many measurements (which has a more precise daily average).

P2683, eq. (1): I would prefer the use of a different notation, not using 'words' but symbols (see also later comments). For example: y is daily-averaged, column-averaged concentration. y_i^{GB} and y_i^{PF} are the ground-based and polynomial fit concentrations on day i, respectively.

P2683, L28 / P2684, L1: How have these accuracies been estimated?

P2684, L6: The target molecules (except for CO) are well-mixed; their concentration is more or less constant in the entire troposphere. Therefore, the total column scales with surface pressure in first order.

P2684, eq. (2): Why have the authors used a scale height of 7.4 km here, whereas in table 3 a scale height of 8 km is mentioned in relation to the IMAP retrievals?

P2684, L17: I guess this 'auxiliary information' refers to surface pressure. Of course, it would be good to have surface pressure measurements at the FTIR stations. In this way, variability in sea-level pressure could be taken into account. However, for

ACPD

5, S1061–S1068, 2005

the validation of total column measurements (as opposed to column-averaged mixing ratios) from satellite it does not help a lot, since no surface pressure measurements for the satellite pixels are available. The present approach, normalizing to zero altitude, is thus fine, especially because the altitude of the FTIR stations and the mean altitude of the SCIAMACHY pixels are very accurately known. Still, the normalization could be done in a much more sophisticated way than equation (2), by using sea-level pressure and temperature profile information from meteorological models. Such refinements may not be needed for CO, but are probably necessary for the well-mixed greenhouse gases, which need to be measured with at least 1% precision.

P2684, L18: How do you know that the normalisation procedure induces possible errors 'up to 15%'?

P2685, L16-21: Why are data not available for the whole year for all algorithms? The large data gaps (in some cases only 4 months are available) make it even more difficult to draw any conclusions concerning e.g. the ability of the satellite measurements to capture seasonal variations. The authors are encouraged to try and get more complete data sets.

P2686, L13-15: Do I understand correctly that CO_2 is available from the IMAP algorithm? If so, why is it not evaluated?

P2686, L24-26: What is meant with this sentence?

P2687, L5-7: I suppose the 'appropriate correlative measurements' refer simply to surface pressure measurements? Are these really not available from the FTIR stations?

P2688, L4: In two recent papers, Sussmann et al. (2005a and b) compared averaging kernels (AK) of WFM-DOAS and the FTIR at Zugspitze. While for CO the Aks were quite similar, this was not the case for CH_4. Have the authors estimated the impact of these differences between the averaging kernels?

P2688, eq. (3): Continuing on my previous remark on notation, $SCIA$ might be re-

ACPD

5, S1061–S1068, 2005

placed by $x_{i,j}^{\text{SCIA}}$, where j is the measurement counter on a particular day i.

P2689, L2: It would be helpful to mention these scaling factors explicitly in the paper.

P2689, eq. (4): This equation is hard to read, mainly because of the notation. Following up on above suggestions, a daily average of SCIA measurements (termed SCIAday in the paper) would become:

$$y_i^{\text{SCIA}} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{i,j}^{\text{SCIA}}, \tag{1}$$

where n_i is the number of SCIAMACHY measurements on day i. Similarly, 'Biasday' would become:

$$b = \frac{1}{N} \sum_{i=1}^{N} \left(\frac{y_i^{\text{SCIA}} - y_i^{\text{PF}}}{y_i^{\text{PF}}} \right), \tag{2}$$

where N is the number of days. It is confusing to call this 'a daily bias'; better would be something like 'bias of the daily-averaged measurements'. Equation (4) can now be written as:

$$\sigma_{\text{scat}} = \text{stdv} \left(\frac{y_i^{\text{SCIA}} - (1+b) y_i^{\text{PF}}}{(1+b) y_i^{\text{PF}}} \right). \tag{3}$$

Is this correct? If yes, what does this equation mean? If the authors' purpose is to 'evaluate the scatter in SCIAMACHY measurements themselves', why do they not follow the same procedure as for the FTIR, namely calculate the standard deviation of the SCIAMACHY data with respect to a polynomial fit through these data? Moreover, equation (4) refers to scatter in *daily-averaged* measurements. In practice, this daily averaging is spatial averaging over a grid box around the FTIR station. The scatter in *individual* SCIAMACHY measurements will be larger.

P2689, L15-17: Please give references for the desired target precisions.

P2692, L2-4: Could the authors spend some words on explaining the incredibly low P-values for all cases, even if, from R, there appears to be absolutely no correlation?

ACPD

5, S1061–S1068, 2005

P2692, L12-23: Do the authors have any clue as to the possible reason of this intriguing dependence of the bias on the total column?

P2693, L7: As noted in the general comments, biases of the order of 5-10% are large for CH_4. Therefore, the characterization 'very small' should be removed.

P2693, L8-10: The fact that the FTIRs exhibit a variation of 3.3% with respect to the polynomial fit, implies that SCIAMACHY data cannot be validated beyond this 3.3% using the present validation method. This is a serious problem considering that a target precision of around 1% is desired, and should be mentioned more clearly.

P2693, L12-13: Is 'closely followed' the correct terminology for a $(5.03-3.6)/3.6 = 40\%$ difference? Is $(6.56-5.03)/5.03 = 30\%$ 'only slightly larger'? What does the 'factor 2.3' refer to?

P2693, L21: Again, the differences between the retrieval algorithms are by no means 'minimal'.

P2694, L12: Do the authors mean 'statistically significant', or 'considerable'? Same question for lines 18/19.

P2694, L16: Why are the biases for N_2O 'not significant'? They are at least large: of the same order as the global variation.

P2694, L23-25: There seems to be a clear seasonality in the bias of N_2O in Lauder, with a maximum in January and a minimum in June/July. Can the authors comment on this?

P2695, L3-7: I am not convinced by this study that the variability of in particular CH_4, CO_2, and N_2O can be detected by SCIAMACHY.

P2695, L23-24: I do not believe that statements on the precision of the CH_4 measurements can be made on the basis of this study. Similarly for the other gases.

P2696, L11-16: One would expect that spatial averaging *reduces* the scatter in the

ACPD

5, S1061–S1068, 2005

Interactive Comment

measurements. Thus, there should be less variability in the large-grid data than in the small-grid data.

P2697, L1-5: This is a weak point of the study. Of course, the conclusion of an inter-comparison like this may be that the FTIR data do not allow to discriminate between the respective satellite retrieval algorithms. But the authors should at least try to get data sets that can be compared (cover the same time period, etc.). Otherwise they might as well focus on the evaluation of one single retrieval algorithm. Besides, I do think that some differences in the performance of the algorithms were detected. For example: the bias in CO dependent on CO itself was found for the WFM-DOAS and IMAP algorithms but not for the IMLM algorithm.

Table 3: 'Over land (altitude > 0)': does this suggest that land is characterized by an altitude > 0?

Table 3: 'weighted variance': what kind of weighting is meant here?

References:

Sussmann, R. and Buchwitz, M., Validation of ENVISAT/SCIAMACHY columnar CO by FTIR profile retrievals at the Ground-Truthing Station Zugspitze, ACPD, 5, 557-572, 2005a.

Sussmann, R. , Stremme, W., Buchwitz, M. and de Beek, R., Validation of ENVISAT/SCIAMACHY columnar methane by solar FTIR spectrometry at the Ground-Truthing Station Zugspitze, ACPD, 5, 2269-2295, 2005b.

Technical corrections

P2683, L24: 'hereabove' –> 'above'

P2691, L21: 'more positive' –> 'higher'

P2692, L15/16: 'lesser underestimation': is this correct English?

ACPD

5, S1061–S1068, 2005

P2696, L26: ‘associated to’ –> ‘associated with’

Table 1: ‘Coordinates’ –> ‘coordinates’

Table 4: ’SCIALACHY’ –> ’SCIAMACHY’

Figure 9: ’... incremented by 50%’ –> ’... incremented by 25%’

Interactive comment on Atmos. Chem. Phys. Discuss., 5, 2677, 2005.