Author comments on:

"Validation of SCIAMACHY HDO/H2O measurements using the TCCON and NDACC-MUSICA networks"

By

R. A. Scheepmaker, C. Frankenberg, N. M. Deutscher, M. Schneider, S. Barthlott, T. Blumenstock, O. E. Garcia, F. Hase, N. Jones, E. Mahieu, J. Notholt, V. Velazco, J. Landgraf, and I. Aben

Referee comments are in blue italics.
Author's responses are in black plain text.

Referee comments by K.G. Gribanov

The paper by Scheepmaker et al. presents a detailed study of creating consistent HDO/H2O data set from different sources covering a substantial period of time. Such a data set can be extended in the future using new satellite and ground-based measurements and can be used for GCM validation studies. In my opinion, the paper is well written. Only small technical corrections are needed.
Despite the fact that I do not insist on even minor revision of the paper, I have a couple of comments regarding the gaps that can be filled in a separate study.

We thank the referee for his kind words and his expeditious review of the paper.

(1) The section 2.1 (starting from line 14) contains that "Unfortunately, full averaging kernels (including the cross-correlations) are not accessible...". In my opinion, despite initial data set does not contain averaging kernels for every individual retrieval, authors of this paper could reproduce at least some or typical full averaging kernel. It is very interesting to have a look at the structure of full averaging kernel.

We agree that it is interesting to have a look at the structure of the full averaging kernels. Therefore, we have repeated a few retrievals, writing out the full averaging kernel matrices. From these, we created **Figure 1** at the bottom of these Author Comments. As is also explained in the caption of the figure, it shows how sensitive the retrieval of H_2O and HDO in the lowest layer of the atmopshere (the only layer which we allow to vary) is to real variations of H_2O and HDO in all the layers. It shows that the retrieval of H_2O and HDO in the bottom layer is also sensitive to variations of respectively H_2O and HDO in the higher atmosphere. This is good, and shows that the entire atmosphere contributes to the total column measurement. The figure also shows that the retrieval of H_2O is *not* affected by real variations in HDO but, on the other hand, the retrieval of HDO *is* slightly sensitive to real variations in H_2O in the higher layer. This "bump" in AK_{DH} at $P_{layer}/P_{surface}=0.4$ gives rise to the cross-correlations between H_2O and HDO that are mentioned in the paper.

We will also rephrase this text in the paper: *"Unfortunately, full averagingkernels (including the cross-correlations) are not accessible...".* Full averaging kernels are accessible (as we showed with Figure 1 below), but not part of the standard output provided to the users, because of data storage limitations.

(2) The section 2.3 (starting from line 9) contains statement that "a posteriori δD contain new information...". I think that this idea should be described more carefully, because quantitative measure of information can be represented by the difference of entropy of probability distribution before measurement and afterwards. Looking at Fig. 2 with wider distribution after measurement, one can think that information is negative.

There are different arguments to show that information has been added by the retrieval, without the need for advanced methods to calculate the information content of the system.
First of all, the structure in the prior δD vs H2O relationship (Fig. 2a, slightly decreasing δD with increasing H2O) is completely artificial. It's a result of our choice of a fixed prior δD profile, in combination with slight changes in the prior H2O profiles (such as an increased scale-height for the profiles with the largest total H2O columns). However, the structure in the posterior δD vs H2O relationship (Fig. 2b) resembles the natural, expected slope of a Rayleigh distillation curve (decreasing δD with decreasing H2O). We have tested that this posterior structure remains the same, even if we assume a prior relationship with no δD depletion at all. Of course the posterior covariance (1-sigma error) on δD is rather large (typically 80 ‰) and widens the distribution, but not enough to remove the natural relationship.
Secondly, although the prior distribution of δD vs H2O is confined to a smaller area compared to the wider posterior distribution, the prior covariances are actually much larger than the posterior covariances. For the retrieval we assumed a prior covariance of 10000% for the amount of HDO and H2O in the lowest layer of the atmosphere, which would lead to a much larger prior covariance for δD than the posterior 80 ‰ and which is also proof that information has been added by the retrieval.

Technical corrections. Authors referred panels of figures in the paper main text as "lefthanded" of "right-handed", but panels on all figures are grouped vertically and have to be referred as "top" or "bottom" as it correctly done in figure legends. Please make these references consistent with actual figures.

Thanks for pointing this out. We will make sure that the text and captions are consistent with the orientation of the figures in the final published version.

Referee comments by N. Rokotyan

The paper describes a newly available SCIAMACHY delD dataset, calculated aposteriori from the columnar H2O and HDO SCIAMACHY retrievals. The Authors did a great job by extending the dataset up to 2007 and validating it against ground-based FTIR stations from the TCCON and NDACC networks. The manuscript has a good structure and is well prepared. The paper can be publish after minor technical corrections specified in the attached PDF document.

I would like to thank the Authors for a job they did. The presented dataset offers an important contribution to the atmospheric sciences community.

We thank the referee for his kind words and for taking the time to review our work.

For their future research I would like to encourage the Authors:
1) To use averaging kernels into account when comparing to other observations. At least in TCCON retrievals, the vertical sensitivity for HDO and H2O is appreciably different and it is varying from measurement to measurement;

The referee is right that ideally these sensitivity differences should be taken into account. But, as the referee already mentions, the averaging kernels vary from measurement to measurement, which makes a proper implementation of this ideal practically impossible (e.g. due to data storage and other practical limitations the averaging kernels are not generally available). But we are encouraged to improve this in future work. Until then, we rely on the fact that the total column averaging kernels for the ground-based FTS measurements and the SCIAMACHY measurements are actually quite similar in the most important lowest layer of the atmosphere.

2) To consider the uncertainty of the coefficient of temperature dependence of airbroadened half width, which can be an important source of a bias;

Interesting point. We agree that uncertainties in the spectroscopic parameters (or differences between the line lists used for SCIAMACHY, TCCON and MUSICA) remain an important source of possible biases. We were already considering taking these uncertainties into account in a general way in future work, but we thank the referee for his suggestion to also specifically look at the temperature coefficient. We might incorporate this into a future sensitivity study.

3) To give more attention to a priori information. How do you construct an a priori profile for a measurement at certain time? In TCCON retrievals only one a priori profile is used for a whole day of measurements, but the temperature and humidity can change a lot during the day introducing an error to the retrievals.

We have already described this briefly at the start of Sect. 2.3. We use a priori profiles (pressure, temperature, humidity) that vary for every single measurement, by taking them from ECMWF and interpolating them to the location and time of every SCIAMACHY ground pixel. The a priori HDO profile is a scaled version of the H2O profile, by assuming a fixed depletion profile of delta-D = -100 ‰ at the lowest layer, increasing to delta-D = -500 ‰ at the highest layer. Subsequently, we took special care to show the column-averaged a priori delta-D values in as many figures as possible, including the a prioris from TCCON and MUSICA, which were considerably different. This way we showed that the conclusions do not depend strongly on the choice of the a priori information, for example in Figures 10 and 11 and the discussion in Sect. 6.2. We agree with the referee that such attention to the a priori information is important and we thank him for the advice to keep this in mind in any future work.

Please also note the supplement to this comment.

Thanks. We have updated the paper accordingly. "microwindows" was changed into "windows" and the "number" on the color scale in the densitity distribution figures has been explained in the captions.

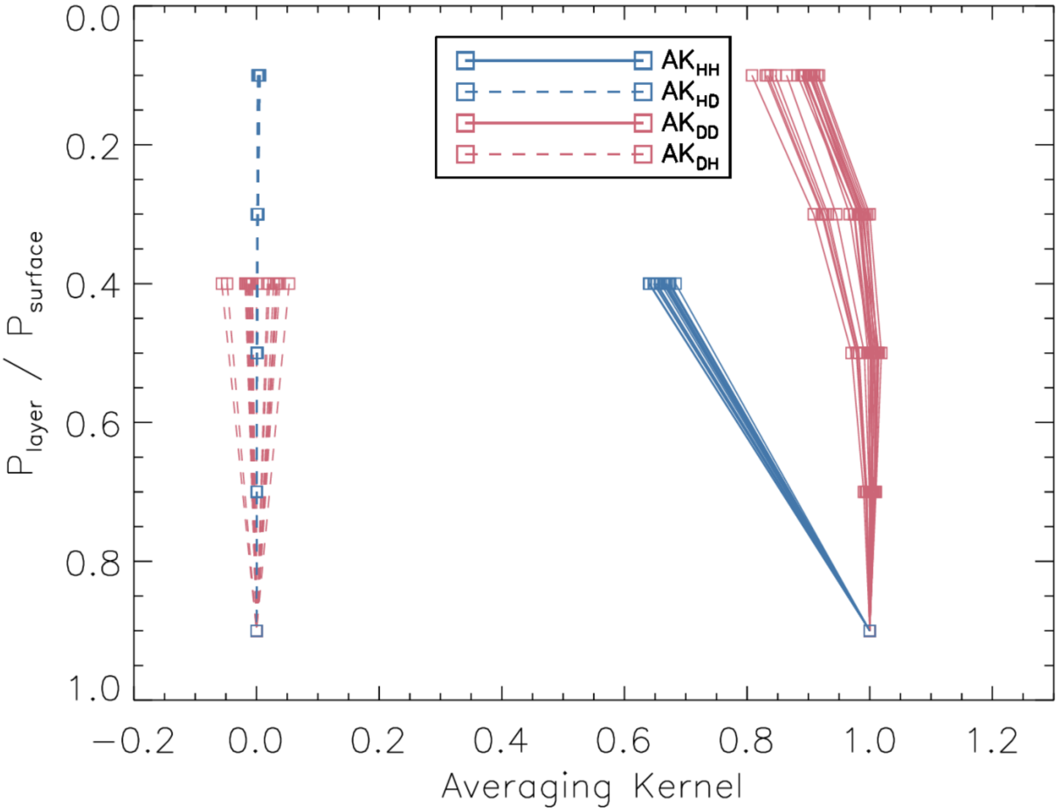

Figure 1 Full averaging kernels for 21 measurements above the Sahara in July 2007. AK_{HH} shows the sensitivity of H_2O ("H") in the lowest layer (pressure P_{layer} between 80-100% of the surface pressure $P_{surface}$) to real variations in H_2O ("H") in all other layers (there was only 1 other H_2O layer used for the top 80% of the atmosphere). AK_{HD} shows the sensitivity of H2O in the lowest layer to real variations in HDO ("D") in all other layers (4 HDO layers were used for the higher atmosphere, but the H_2O sensitivity to HDO is zero). AK_{DD} shows the sensitivity of HDO in the lowest layer to real variations in HDO in all other layers and AK_{DH} shows the sensitivity of HDO in the lowest layer to real variations in H_2O in the higher layer. The nonzero values for AK_{DH} at $P_{layer}/P_{surface}=0.4$ point to some cross-dependencies between H_2O variations at higher layers, influencing the retrieved HDO in the bottom layer.