General Comment:

The paper by Worden et al. presents a new version of Aura TES retrievals of CH4, HDO, H2O, (and HDO/H2O), and N2O profiles. Furthermore, the authors present an interesting method for an a posteriori correction of the estimated CH4 profiles using the co-estimated N2O profiles.
The authors claim that the new retrieval version produces profiles with a better vertical resolution if compared to former versions. The better vertical resolution is achieved by applying the following modifications to the retrieval algorithm:

(1) They fit a broad spectral region between 1170 and 1330 cm^{-1} instead of a few small spectral microwindows as has been done in former retrieval versions. In addition they now simultaneously retrieve all the absorbers with signatures in the broad spectral region (in former versions each absorber was fitted separately applying absorber specific small microwindows).
(2) They reduce the hard constraint. Therefore they increase the number of retrieval levels (altitude/pressure levels) if compared to former retrieval versions.
(3) They reduce the a priori constraint applied for the optimal estimation inversion procedure. For the new version they reduce the a priori assumed correlation length scales between the mixing layer and the lower troposphere. They argue that such weaker a priori constraints are suggested by several recent in-situ and satellite-based observations as well as by new model simulations.

The paper is of high interest for the space-based thermal-nadir infrared remote sensing community and I recommend its publication in AMT. For me it is particularly interesting since I am working in the same field but with IASI instead of TES. However, I would also like to recommend some revisions (see my following major comments) before publication.

Major Comments:

I: Applying a broad spectral region and fitting all absorbers simultaneously instead of the former microwindow approach is a good idea. I well believe that it produces profiles with an increased vertical resolution. However, the by most important modification is the changed a priori constraint. Weakening the constraint naturally increases the theoretical profiling capability of the remote sensing system. I wonder whether the effect of applying a broad spectral region can be completely neglected if compared to the effect of reducing the a priori constraint.
Did the authors simulate the effect of applying a broad spectral region? By how much increase the DOFs if you apply the broad spectral region instead of the microwindows but keep the constraint constant? A table where the effects of the different modifications are documented would be nice: change to broad spectral region means additional xx DOFs; change of hard constraint means additional xx DOFs; change of soft constraint means additional xx DOFs. I have the feeling that the change of the soft constraint is clearly dominating. If so, it should be made clear in the manuscript!

II: By weakening the a-priori constraints one assumes that the real atmosphere is more variable than has been assumed before. Theoretically the quality of an optimal estimation retrieval will be best, if one uses the actual atmospheric covariance as the a-priori covariance. The actual atmospheric covariance can only be determined by real atmospheric data or by very reliable model data. The authors state that they use new measurements and new model data for constructing the new a-priori constraint. However, in this context I found the subsequent discussion about the theoretical uncertainties a bit misleading. With this discussion the authors cause the impression that their theoretical calculations can document the superiority of the retrieval with reduced constraint: e.g., page 6690, line 1-4: "[…] there is a net increase in the error in the boundary layer due to temperature and noise of approximately 3%. On the other hand, the total error for the HDO/H2O ratio in the free troposphere has decreased because the increased vertical resolution reduces the smoothing error". The authors calculate the so-called smoothing error by:

$$(A_{xx} - I) * S_a * (A_{xx} - I)^T \qquad \text{(see Equation 3 of the manuscript)}$$

It would be very important for the reader to know what S_a has been used by the authors for this error estimation. The "smoothing error" strongly depends on the assumed a-priori covariance! If they assume the S_a that has been used for calculating the new weakened a-priori constraint it is no surprise that they estimate that the new retrieval will provide profiles with reduced total error. However, if the new S_a assumption is wrong and the former S_a assumption was right, the former S_a should be applied when calculating the "smoothing" error. Then the error estimation will look different and the new retrieval will probably perform poorer.

In summary: The error estimation is made for the new S_a (obtained from the new measurements and model calculation cited by the authors). It is only valid if the new S_a is the right S_a. I think this should be made clearer in the error discussion. Since the improvement of the retrieval depends on the validity of the new S_a I would also like to recommend changing the title to something like: "Using new a priori assumption for producing profiles of CH4, HDO, H2O, and N2O with improved lower tropospheric vertical resolution from Aura TES radiances"

III: The manuscript motivated me to investigate a reduction of the a-priori constraint for our IASI H2O profile retrieval (for more details see Schneider and Hase, 2011, in the following referred to as SH 2011). Therefore, I compared our IASI H2O profiles with coincident Vaisala RS92 radiosondes profiles. I think that the results are interesting for the authors:

I made two IASI H2O retrievals: a first as presented in our SH 2011 paper, and a second for which I reduced the a-priori correlation length below 800 hPa from 2.5 km to 1 km. I guess that this second constraint is similar to the new TES retrieval constraint. Figure 1 depicts on the left panel the H2O averaging kernels for the retrieval as presented in SH 2011 and on the right panel the retrieval with the "relaxed" constraints as suggested by the authors. Naturally, one observes an increase in the DOF (from 3.81 to 4.54) and a better vertical resolution in the boundary layer and the lower troposphere. However, Fig. 2 and 3 indicate that applying the "relaxed" constraints increases the disagreement with the coincident RS92 measurements. Please note that for the comparison the vertically highly-resolved RS92 data have been smoothed by IASI's averaging kernels, i.e., the left and right panels show comparisons for profiles with different vertical resolution.

In summary: This brief study empirically documents that the increased resolution is on cost of the precision. At least for our IASI retrievals we will leave the constraints as is: poorer vertical resolution but higher precision.

By the way: Our Schneider and Hase (2011) IASI thermal nadir retrieval is very similar to the new setup reported by the authors. We fit a broad spectral region (1190-1400cm^{-1}) and simultaneously retrieve H2O, HDO, (and HDO/H2O), CH4, N2O, HNO3, and CO2. The paper is on APCD since May 2011 (http://www.atmos-chem-phys-discuss.net/11/16107/2011/), and the authors might just have overlooked it. Today it should go online on ACP and due to the similarity to the author's "improved retrieval setup" I think it our paper should be cited.

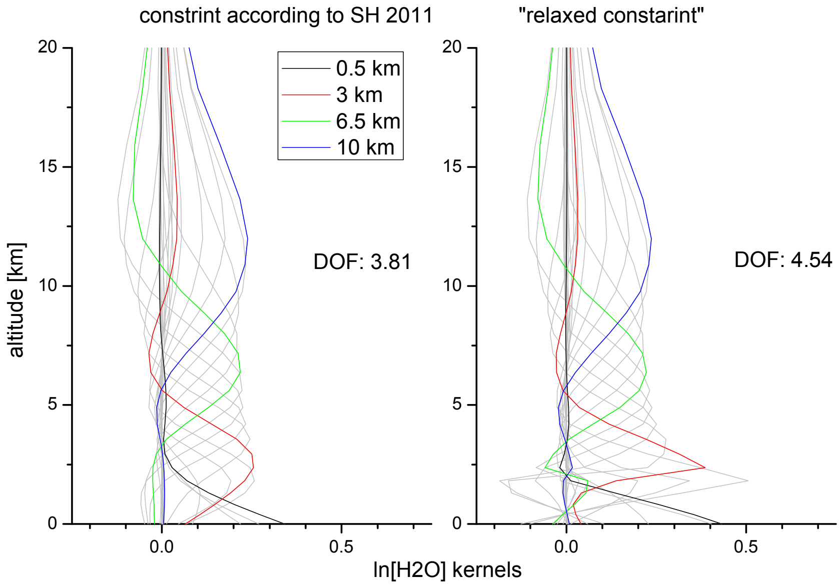

Figure 1: Averaging kernels for ln[H2O]. Left panel: for constraint according to Schneider and Hase (2011); Right panel: for the "relaxed" constraint according to the manuscript under revision.

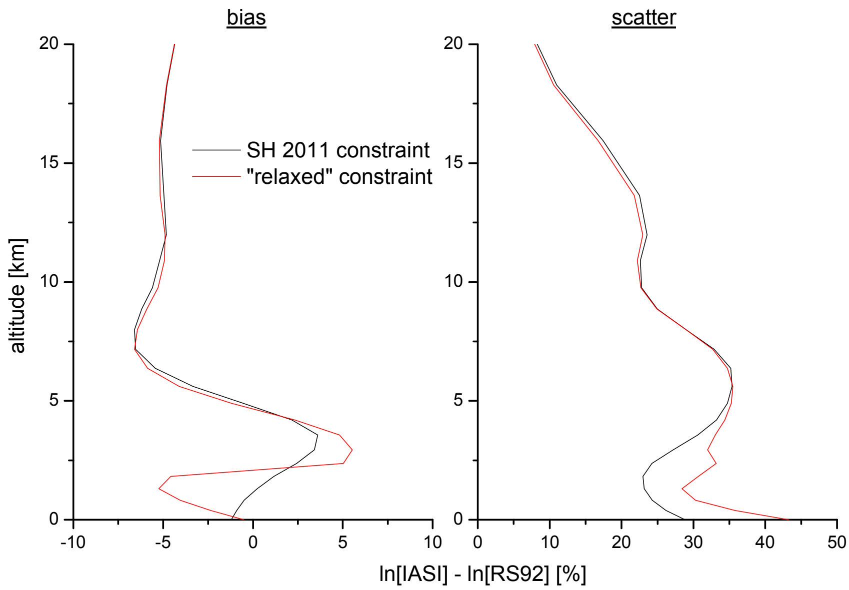

Figure 2: IASI-RS92 intercomparison study based on a large number of IASI and RS92 coincidences (for more details please refer to SH 2011). Left panel: systematic difference; Right panel: Std of the difference.

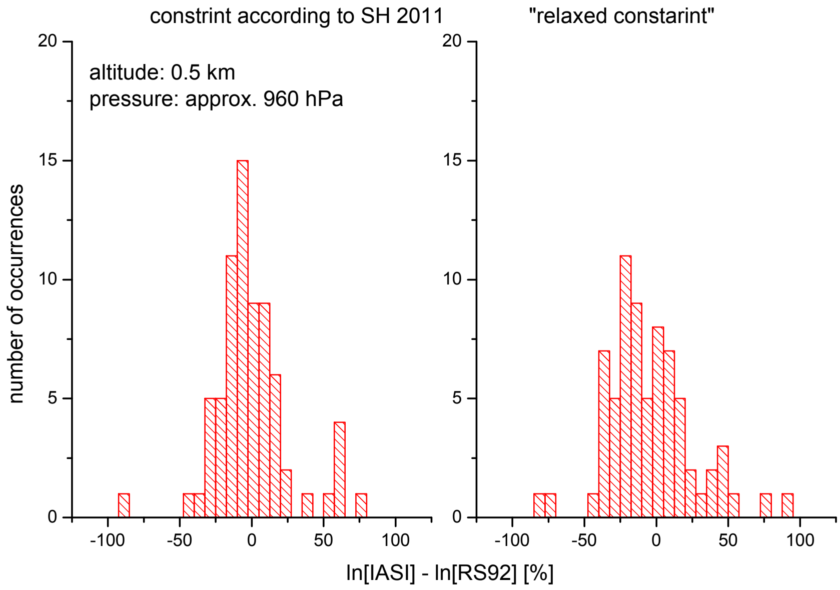

Figure 3: Histograms for IASI-RS92 in the boundary layer (altitude 500m a.s.l.; approx. 960 hPa). Left panel: for constraint according to Schneider and Hase (2011); Right panel: for the "relaxed" constraints.

IV: Sensitivity with respect to HDO/H2O: In Section 4.2 the authors state that "[…] the HDO averaging kernel best describes the vertical sensitivity for the HDO/H2O estimate […]". I disagree! In the following I will show that using the HDO kernels as a proxy for the HDO/H2O kernel significantly overestimates the HDO/H2O sensitivity. Actually the HDO/H2O sensitivity is smaller than the HDO sensitivity: The reason is that the space spanned by the HDO kernels is no sub-space of the space spanned by the H2O kernels.

Figure 4 shows typical kernels of our IASI H2O, HDO, HDO/H2O retrieval (SH 2011). Since we perform an inter-species constraint the H2O and HDO kernels cannot be considered independently. Instead we have to consider the full H2O-HDO kernel system consisting of A_{HH} (describing how a change in the atmospheric H2O is reflected in the retrieved H2O), A_{DH} (describing how a change in the atmospheric HDO is reflected in the retrieved H2O), A_{HD} (describing how a change in the atmospheric H2O is reflected in the retrieved HDO), and A_{DD} (describing how a change in the atmospheric HDO is reflected in the retrieved HDO). We can see that not only the H2O variability affects the HDO retrieval, instead there is also a vice versa response from HDO to H2O. The amount of information that HDO contributes to the H2O retrieval is not negligible. The respective kernel is A_{DH} and has a trace of 0.22 (DOF value).

Since tropospheric H2O and HDO change almost in parallel the [A_{HH} – A_{DH}] kernel documents the actual H2O sensitivity of the H2O-HDO remote sensing system. The [A_{HH} – A_{DH}] kernel is depicted in Figure 5.

The variability in ln[HDO] - ln[H2O] is a good proxy for the variability in δD (see Eq. (6) of SH 2011) and we can use ln[HDO] - ln[H2O] kernels interchangeably to δD kernels. ln[HDO] - ln[H2O] can change by a change in ln[H2O], by a change in ln[HDO], or by a simultaneous change in both ln[H2O] and ln[HDO]. Figure 6 present the respective ln[HDO] - ln[H2O] kernels. The three kernels are very similar. For our typical example we get a DOF of the ln[HDO] - ln[H2O] kernels of about 0.7. Please note that these kernels are significantly different from the A_{DD} kernel (bottom right panel of Fig. 4) that is used by the authors as δD kernel.

In summary: The authors use the A_{DD} kernel as an approximation for the δD kernel. This approximation overestimates the remote sensing system's sensitivity with respect to δD. In our example we found that the trace of A_{DD} is 0.9, whereas the trace of the ln[HDO] - ln[H2O] kernels is only 0.7. I recommend changing the text in the last paragraph of Section 4.2 accordingly.

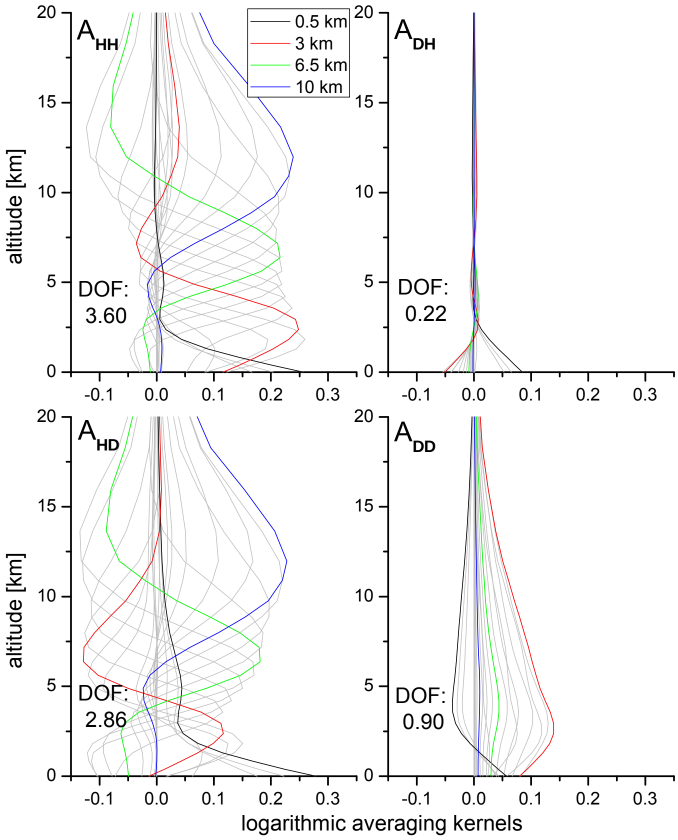

Figure 4: The column kernels for the combined IASI H2O-HDO retrieval. The grey lines show all kernels, the different colours highlight particular kernels (see legend).

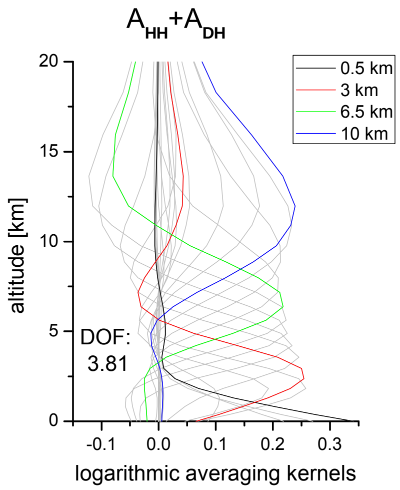

Figure 5: ln[H2O] column averaging kernel.

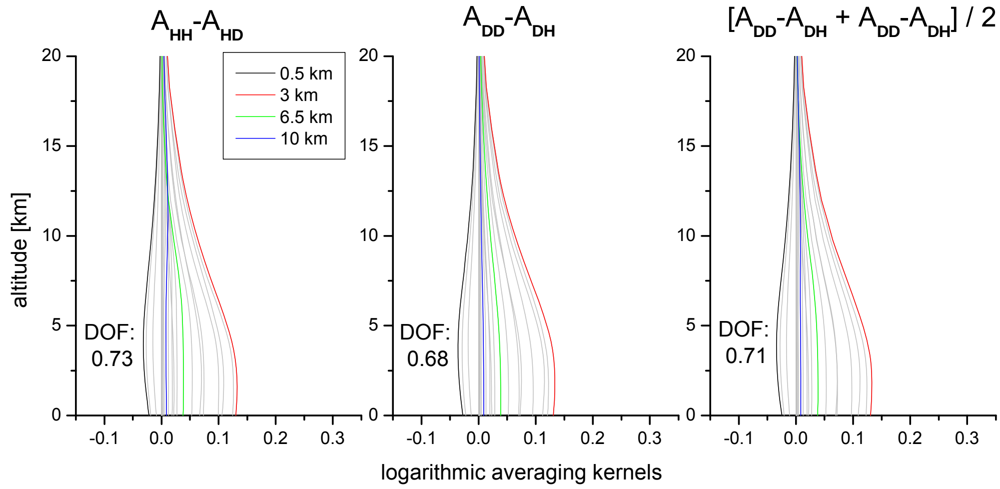

Figure 6: Different ln[HDO] – ln[H2O] column kernels. From the left to the right: for a change in ln[H2O], for a change in ln[HDO], and for a simultaneous change in both ln[H2O] and ln[HDO].

V: Section 4.3.4, a posteriori correction of CH4 by the retrieved N2O: This is very interesting. In addition I think it can be further improved. Instead of correcting the CH4 a posteriori you could a priori introduce a ln[CH4]-ln[N2O] inter-species constraint thereby constraining against a CH4/N2O ratio similar to what is done for HDO/H2O. I am not sure but maybe this will reduce the jumps you are talking about. You might think about mentioning such ln[CH4]-ln[N2O] inter-species constraint retrieval and say that it would be a interesting future development of the CH4 retrieval.

<u>Minor Comments:</u>

Page 6683, Eq. (1): a "+" is missing in the third term

Page 6684, Eq. (2) and (3): what is the difference between Λ_z and S_a^{-1}? From explanation in the text I got the impression that you apply as constraint S_a^{-1}, i.e., $\Lambda_z=S_a^{-1}$.

Page 6684, Eq. (3): the third term should be S_M or $G_z * S_m * G_z^T$. Please correct.

Page 6685, line 1 and 2: I would relate S_{x1} and S_{x2} to S_a and S_{tot}: $S_{x1}= S_a$ and $S_{x2}=S_{tot}$, right?

Page 6685, line 16 and 17: two times "illustrates"

Page 6685, line 23: remove "the"

Page 6687, line 13: "[…] covariances from these models are not typically invertible.": this is no good argument because one can perform a "pseudo" inversion via a singular vector decomposition.

Page 6688, line 7: what are "observation covariances"? Please define. It is explained in the caption of Fig. 3, but I think it should also be explained in the text.

Page 6689, line 18: Worden at al. (2010) does not appear in the reference list.

Page 6691, line 12: Worden et al. (2011) does not appear in the reference list.

Page 6692, line 16-19: "[…] this increased sensitivity to the lower and middle troposphere is due to use the methane lines around 1230 cm^{-1}.": This is interesting! For methane changing from the microwindow approach to fitting the broad spectral region significantly increases the sensitivity, whereas for the H2O and HDO the increase in the sensitivity (or vertical resolution) is mainly due to the "relaxed" soft constraints. As already mentioned in my major comment (I), I think that a table describing how the DOFs for the different absorbers change due to the different modification (broad spectral region, hard, soft constraints) would be very useful for the reader.

Page 6692, line 24: I guess you mean here Fig. 2 instead of Fig. 5.

Page 6692, line 25: isn't an assumed a priori variability of methane of 5% a bit too large? The peak-to-peak amplitude of the seasonal cycle is only about 2%, right?

Page 6694, you mention that the bias might be caused by an anti-correlation between upper and lower/middle tropospheric methane. You say such an anti-correlation is suggested by the kernels of the new retrieval (negative values of lower/middle tropospheric kernels in the upper troposphere, right panel of Fig. 8). In old kernels there are no negative values of the lower/middle tropospheric kernels. In consequence there should be no bias? Is this the case?

Page 6696, Eq. (11): I think you should write this Equation similar to Eq. (3). Writing it different is an unnecessary source of confusion. Therefore, I suggest modifying Eq. (3) a bit. Change the last two terms of Eq. (3) to:
$G_R * S_m * G_R^T + G_R * (\Sigma_i K_i * S_b^i * K_i^T) * G_R^T$

Then you can also mention that writing here G_R=G_C-G_N instead of G_C (or G_z) makes the difference.

Page 6696, last line: I do not understand what you mean with "[…] the bias error described in Eq. (9). Do you mean "[…] work for correcting the bias error shown in Fig. 11."?