We would like to thank the referees, Dr. Worden and Dr. Schneider for their expert review. Their comments were judicious and enabled to improve the manuscript. In particular, both referees stressed that a description of the a priori variability as well as the information content was missing. This is adequately addressed in the revised manuscript. All the other comments have been addressed as well. Pleas note that we reply to each referee's comment separately and thus sometimes repeat the same information for completeness.

Please find detailed answers (in blue) to the referee's comments (in gray) below.

Referee # 1, Dr. J. Worden

Comment 1:

Please add a section or statement of the chi2 values typical of a retrieval as well as data selection. Do you use all the data or only data that is less than a threshold chi2 value.

→ Here we understand the chi2 values the referee mentioned as the comparative test between the difference of the fit and the measurement with the experimental error defined by (Rodgers, 2000) as:

$$\chi^2[\mathbf{y} - \mathbf{F}(\mathbf{x}_i)] = [\mathbf{y} - \mathbf{F}(\mathbf{x}_i)]^T \mathbf{S}_\epsilon^{-1} [\mathbf{y} - \mathbf{F}(\mathbf{x}_i)]$$

We indeed use this formula to filter out fits which have a too large residual; and we have now clarified this in the text (section 4). For the information of the reviewer we also show the distribution of chi2 values (normalized by the degrees of freedom for noise) obtained along the latitudinal gradient in Fig 1.

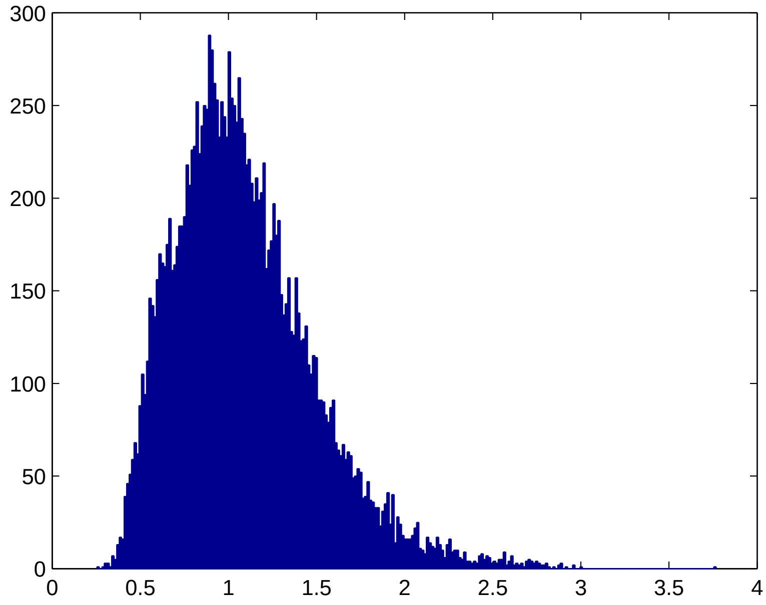

Figure 1: Histogram of normalized Chi2 values obtained in retrievals along a latitudinal gradient.

Comment 2:

Page 13058 Line21: Use of the phrase constrained approach is confusing as all retrievals have some form of constrained approach. What distinguishes this approach from the previous retrieval discussed in Herbin et al. is the use of correlations between(log) HDO and (log) H2O in order to presumably obtain more physically realistic distributions of HDO/H2O estimates. Perhaps call this retrieval the HDO/H2O correlated approach? Make sure to define this jargon early in the document. Page 13058 Line 22: The paragraph beginning with While the constrained. . . is technically correct but confusing to the reader as several disjointed points are being made. I think what you are trying to say is Use of correlations between (log) HDO and (log)H2O helps to constrain the joint HDO/H2O retrieval to physically plausible solutions as demonstrated by Worden et al. and Schneider et al. . . .. ; the choice of cross-correlations is discussed in Section 3.3. The choice of retrieval parameters also affects the vertical resolution and error characteristics of the retrieval; the set of retrieval parameters is discussed in Section 3.2.

→ We agree with this comment. This has been corrected and clarified.

Comment 3:

Page 13060 Section 3.3 What is the a priori state vector? Is it output from the LMD model? Does the a priori vary?

→ No the a priori does not vary. We clarified this by adding the following sentence (section 3.3 end of paragraph 3): *To avoid spatial dependency of the results on the a priori profile, a single a priori state vector has been calculated as the mean of an ensemble of HDO and H_2O simulated profiles representative of the whole globe and the whole year. The associated covariances ($\mathbf{S}_{\mathrm{HDO}_L}$,$\mathbf{S}_{\mathrm{H_2O}_L}$,$\mathbf{S}_{(\mathrm{H_2O}_L,\mathrm{HDO}_L)}$) have been computed to build the a priori covariance matrix.* A plot of the a priori profiles of H_2O and δD is now also given (Fig. 2 of the revised manuscript).

Comment 4:

Page 13060 Line Sentence starting with A simple Im surprised that it was necessary to relax the correlations between HDO and H2O in order to improve the residuals between model and measurement as it appears that the constraint is already quite loose. For example, the original TES retrieval had a much tighter constraint (from what I can tell from this paper) and also much higher spectral resolution and the residual of forward model and radiance fits almost to the noise level. Consequently, it is possible that a poor residual is indicative of other parameters not being well fit, such as temperature. Again, some description of the a priori covariance for the HDO/H2O ratio and an improved description of the temperature uncertainties would be useful for understanding this issue better.

→ The authors are grateful to the referee for pointing out this incoherence. Actually the a priori matrix has not been modified to improve the residual of the fit but to facilitate the convergence of the retrieval. Without any modification the number of successful retrievals was very poor, this tuning has been empirically determined to maximize the number of successful retrieval while maintaining a sufficient constraint on the ratio to get meaningful results. This section has been largely modified to correct this and to provide a description of the a priori covariance for the HDO/H_2O ratio.

Further more, you seem to have understood that we retrieve the temperature profile, but we do not, we use the L2 product from EUMETSAT. The latter is thought to be retrieved with an error ranging from 0.6K to 2K, see Section 3.2. We added details on how this error is varying with altitude: *The temperature profiles used are the ones retrieved by the Eumetsat Level 2 processor (...), estimated with an error around 1.5K at the surface, 0.6K between 800 and 300 mb, and 1.5K in tropopause (...).* Finally as for your comment on the constraint: with the different additions in the text and especially the new Figure 2, we believe it is now clear for the reader that the constraint we used is relatively loose compared to what is done by M. Schneider et al. and your retrieval on TES. For a more detailed comparison it would also be necessary to consider the inter-level correlation elements and this is not straightforward.

Comment 5:

Page 13062 Line 9 Sentence beginning with The Construction: : : Could you provide a description of the a priori HDO/H2O covariance? Following Equation 21 in Worden et al. 2006, this covariance should be SR = SHDO SH2O. Also, could you show the square root of the diagonal of the HDO/H2O covariance in Figure 5 for comparison to the a posteriori errors?

→ We think that this comment is addressed in the previous reply: Fig 2 provides now a description of the a priori HDO/H_2O covariance via the square root of its diagonal elements. The covariance can however not be calculated like in Worden et al. 2006 as we don't assume that the fractionation is uncorrelated with the amount of water. Figure 5 only shows observation errors, to keep it readable I would prefer to not plot the square root of the HDO/H2O which is now shown on Figure 2.

Comment 6:
Page 13064 Line 24: Worden et al. 2006 or Schneider et al. 2006 are the correct references here. Risi 2012b used this equation from Worden et al. 2006 for the model/data comparisons. In addition, equation 11 is not technically correct as it does not include the error terms due to noise, temperature etc. The other error terms need to be included here or alternatively indicate to the reader that you are specifically excluding them and to refer to prior equations.

→ We corrected the reference. The purpose of using this model/data comparison equation was to get rid of the smoothing error to assess the error due to the measurement noise and the error due to temperature profile uncertainty.

Comment 7:
Page 13065 Lines 11 through 25 starting with First, we performed I would strongly encourage the authors to remove this section or move it to an Appendix with additional explanation. Basically, you do not need to do these kinds of tests unless you have reason to believe that your retrieval algorithm is not mechanically robust (that is, the retrieval has software bugs or incorrect equations). Instead, use (or calculate) the posteriori temperature covariance from the temperature retrieval and apply the gain matrix from the HDO/H2O retrieval to this covariance to obtain the impact of temperature on the HDO/H2O estimate.

→ The forward simulations approach was chosen to obtain an estimate of the error due to the measurement noise and the uncertainties in the temperature profile in a metric that would speak to the reader, that is to say in per mil. If the error on log(HDO/H2O) is approximated by the fractional error and then transformed in per mil, the approximation can lead to significant overestimation which increases with the error. For example see new Fig. 2 panel c) of the revised manuscript: we plotted the true error on HDO/H2O (computed from the model data) and the error on HDO/H2O computed from the log(HDO/H2O) error both expressed in per mil. At 5 km where the error is maximal, the true error is 117 per mil while the approximated one is 141 per mil. Moreover, as we do not retrieve temperature profile, we can not compute systematically the derivatives. The construction of an a priori covariance error for the temperature is also not straightforward as we rely on the L2 profiles retrieved independently by EUMETSAT processor. For these reasons we preferred keeping the forward simulations as a tool to estimate the error and we think the results are reliable.

Comment 8:
Page 13065 Line 27 paragraph starting with To determine: : : This paragraph is confusing because you say not in agreement which implies that the error calculations contradict each other. Perhaps state The altitude region where the retrieval has the most sensitivity is inferred by the reduction in smoothing error shown in Figure 4. The total error is shown in Figure 5 and includes uncertainties from noise, temperature etc., is reduced throughout. However, the IAIS HDO/H2O estimates cannot distinguish the HDO/H2O variability in the lowermost troposphere from that in the middle troposphere where sensitivity peaks.

→ Thanks for your suggestion, we changed the text according to this.

Comment 9:
Page 13066 Line 17. Do you need to augment the previous sentence to indicate that average of the averaging kernels is robust for grid boxes that include land and ocean points? Otherwise your conclusion in Line 16-17 is confusing.
→ We removed the confusing sentence.

Comment 10:
Page 13066 Line 19: What is a pattern correspondence? I think you mean comparison?

→ Pattern correspondence is also confusing, we changed this sentence to: *To evaluate the differences between observations and model (..)..*

Comment 11:
Page 13066 Line 3 Sentence starting with This can be explained Again, you need to show the a priori covariance. If I look at the total error in Figure 5 I would expect that the a priori covariance is quite loose (¿ 100 per mil) at the surface. Please explain.

→ Yes this was definitely missing and yes the a priori covariance is quite loose relatively to TES and Schneider & Hase 2011 as explained above. At the surface and also in the rest of the troposphere. In the modified section 3.3 we explain that this loose constraint is a direct consequence of the lowering of the inter species correlation elements used to facilitate convergence.

Comment 12:
Page 13066 The advantage of comparing the IASI data to the LMD model is that the LMD model has been compared to the TES and SCIAMACHY data, as discussed in Risi et al. 2012a. Consequently, it would be useful to the reader for you to review the results of 2012a at the start of Section 4 so that it is more clear to the reader that you are using the LMD model to indirectly compare IASI data with TES and SCIAMACHY data. Thereafter, you can include a more in-depth discussion on whether the differences between LMD and IASI are consistent or inconsistent with LMD and TES, when applicable; this discussion would then address my criticism during the initial review about using the LMD model as a transfer between TES and IASI.

→ The results we document here are presented in a way to illustrate the high spatial and temporal sampling resolution of IASI. By comparing our results with LMDz we aim to illustrate the consistency of our results. Of course, like you say, LMDz is a very interesting model because it has already been compared to TES, SCIAMACHY, and also ground based FTIR. We think however, that an inter-comparison between instruments/retrieval methods should be done considering similar areas and periods. This would be of high interest but it is currently out of the scope of the paper. As you suggested it, the advantage to use LMDz has been added as well as a short summary of the 2012a results.

Comment 13:
Page 13068 Line 27 Sentence beginning with At Darwin. Im not yet convinced that IASI captures the short-term variability as represented by the LMD model (Also, does short-term variability mean approximately monthly variability?). If I examine the figures I would conclude by eye that IASI delta-d values generally captures the seasonal variability and sometimes it captures variations at monthly time-scales and sometimes not. The metrics used for this conclusion are insufficient because it is not obvious that these metrics are or are not driven by the overall agreement to the seasonal variability versus short-term variability. If you want to make this a result for the paper you will need a way to filter the seasonal variability from this comparison. Thereafter, you could then quantify the difference between the seasonally detrended time series and compare to the averaged mean variability of the IASI data. Also, is it possible that you are capturing the LMDz variability because you are using the LMDz model as a priori ? (see previous query about your a priori state vector choice)

→ We thank you for this useful comment. It is true that the metrics we presented are influenced by the seasonal and intra-seasonal variability. Following your suggestion, trends have been fitted for each time series to obtain a seasonally detrended residual. The function parameters used were different for the LMDz time series and for the IASI time series as they exhibit different behavior (for example, deltaD time series at Izana presents local maxima in March and in August while LMDz is relatively flat and maxima do not occur at the same time). The results are given in a new figure (Fig. 9). We find they are even more convincing in showing the ability of IASI to capture the short-term variability (by this we mean day-to-day variability) of deltaD. The text (subsection 4.1.2) has been modified accordingly.

Comment 14:
Page 13069 Line 18 Sentence starting with Note: : : Im not convinced of this argument. Application of the averaging kernel to the model field will take into account the variability allowed by the constraint if the retrieval is well characterized. In addition, you are making an inference about the ability of TES data to capture short term variability relative to the IsoGCM model by comparing the IASI data to the LMD model; for this conclusion to be robust you need to now compare the LMD model to the IsoGCM model. Consequently, you either need to perform this additional IsoGCM/LMDz model comparison or remove this conclusion.

→ We agree that the statement was insufficiently supported and that a detailed satellite inter-comparison and satellite/model comparison would be needed to properly address this. We removed this conclusion.

Comment 15:
Page 13072 Line 9 Sentence starting with More generally This is a relative statement. You need to say very good performance relative to something else (e.g. previous retrieval algorithms?).

→ We removed the relative statement. It now reads: *More generally, results presented here highlight further the exceptional potential of IASI to contribute to the understanding of hydrological processes.*

Comment 16:
Page 13079 Figure 1. Please add another panel or figure showing the comparison between the radiance, forward model and noise.

→ Done.

Comment 17:
Page 13082 Figure 4. If I interpret this figure as described it says that the smoothing error is 2% of the a priori covariance for the HDO/H2O ratio. I am guessing that this is not the case
→ You are right, that was not correct, thank you for pointing this out. The figure has been corrected.

Comment 18:
It might be useful for the authors to plot the DOFS for the HDO component of the retrieval because HDO is generally the limiting component (with respect to the sensitivity) to the HDO/H2O retrieval. However, please review the language in Worden et al. 2012 AMT as the use of HDO DOFS as a proxy for the HDO/H2O retrieval is not completely robust as pointed out by Dr. Matthias Schneider during review of the Worden et al. 2012 paper.

→ We added an entire subsection to document the DOFS values of our retrieval (Section 3.5), as well as a new figure (Fig. 6).

Referee # 2, Dr. M. Schneider

Comment 1:
Mention the difficulties of isotopologue ratio measurements already in the introduction section: Isotopologue ratio remote sensing is very demanding. I think the additional difficulties of water vapour isotopologue ratio remote sensing if compared to the remote sensing of atmospheric molecules should be mentioned in the introduction section. I propose to add something like: Tropospheric water vapour concentrations are very variable. Compared to this large variability the mid-tropospheric HDO/H2O ratio is rather stable. For measuring water vapour isotopologue ratios, we need a technique that is, firstly, sensitive over a large dynamic range, and secondly at the same time, very precise. Since it is very difficult for any measurement technique to optimally meet both requirements, tropospheric water

vapour isotopologue ratio measurements are very difficult

→ We thank you for your suggestion. We have added the sentences as recommended (end of the introduction section).

Comment 2:
Page 13058, line 25ff: [] it is anticipated that the retrieval will greatly depend on the choice of retrieval panrameters, and in particular on the choice of the a priori information. For instance, Worden et al. (2012) show an improvement in the sensitivity of TES to dD after a change of the retrieval parameters. Worden et al. (2012) improved the interpretation of the TES spectra by using a finer gridded model atmosphere and by fitting a broad spectral microwindow instead of small spectral microwindows. The same strategy has already been applied by Schneider and Hase (2011). However, the retrieval method of the authors does not apply these improvements: they use a sparse model atmosphere (limited to 10 levels below 10km) and rather small microwindows. I think this needs to be motivated. Why do the authors not build on the improvements introduced by Schneider and Hase (2011) and Worden et al. (2012)? I think they should briefly explain their motivation. For instance, if their main motivation is a very fast retrieval method they should mention it.

→ This comment and some of the next ones go in the same direction and express the need of explanations concerning the motivations of our retrieval choices compared with what have already been done by Schneider et al on IASI spectra and on TES spectra. Having a fast retrieval method is one of the motivations for the choice of the retrieval spectral range and of the state vector to be retrieved. We have added a few sentences at the end of section 2 as well as in section 3.2 to clarify this. Of course we acknowledge the development made recently by you and the TES team but we had started our work prior to these publications. We are furthermore limited by some of the features of the retrieval algorithm (e.g. not fitting the temperature profile).

Comment 3:
Section 3.2, model gridding: The authors limit the model atmosphere to the lowermost 10km of the atmosphere, use a gridding of 1km, and state that variations of the state vector [at other altitudes] do not significantly affect the measurements. As already mentioned in my previous comment, I think it needs to be explained why the authors do not build on the Schneider and Hase (2011) and Worden at al. (2012) experiences: Schneider and Hase (2011) used a fine model grid and a model atmosphere up to the upper stratosphere for their IASI retrieval. Their H2O averaging kernels show still good sensitivity at 15km (sum along row of kernel matrix at 15km is still 0.5). Worden et al. (2012) found it important to increase the number of grid points for the new TES data version in order to increase the information content. Furthermore, in Fig. 3 the authors seem to indicate that there is still some information above 10km. I think that these discrepancies need to be discussed.

→ Part of your comment comes from missing information concerning the altitude of the modeled atmosphere. The atmosphere is modeled in the forward model up to 24 km but we only retrieve the 10 first kilometers. From 10 to 24 km we fixed the water amount from the EUMETSAT L2 water product. We have checked that the expected error at those altitude do not significantly affect the measurement within the spectral region used.

Comment 4:
Section 3.2, applied spectral microwindow: The authors use very small spectral microwindows. This assures a high speed of the inversion algorithm. However, a small microwindows theoretically limits the sensitivity of the retrieval. Schneider and Hase (2011) use a broad microwindow. Worden et al. (2012) also changed to a broad microwindow for the new TES retrieval and argued that it is important, firstly to increase the sensitivity with respect to H2O and dD, and secondly to reduce interferences from N2O and CH4. So the authors method of using a small microwindow might be very helpful to speed up the retrieval process but theoretically it will lead to reduced sensitivity and increased interference errors. In this context it would be very helpful if the authors discuss the information content that they obtain form their microwindows. I think it would be very important that they mention the degree of freedom

of signal (DOF) they obtain for H2O and dD. They should discuss their values with the values obtained by Wordens TES retrieval and Schneider and Hases IASI retrieval.

→. We added a new section (section 3.5) and a new figure (Fig. 6 of the revised manuscript) to address this issue on the DOFS, that we agree was missing to describe our retrieval. We also briefly discuss their values with the values obtained with the last version of the TES retrievals and yours. Regarding the spectral region used, our DOFS are quite high and it is likely due to the a priori covariance matrices that exhibits high diagonal elements but also to the high inter level correlation which also increases the DOFS (Rodgers). For info, the approximated correlation length of our Sa is about 9 km.

Comment 5:
Section 3.2, cloud filtering: How was the cloud filtering performed? Please specify. In case the authors use the EUMETSAT cloud flags, they should specify the flag used for excluding cloud contaminated observations.

→ Yes we use EUMETSAT cloud flag, the specification of the flag has been added : *Only spectra not contaminated by clouds have been considered in this study (defined as EUMETSAT's level 2 cloud fraction below 10%)..*

Comment 6:
Section 3.3, choice of a priori information, page 13061, line 24ff: We found that the covariance matrix constructed in this way [applying LMDz model data] constrained the retrieval too much. This was evident by looking on the retrieval residuals y-F(x,b), which were found to be much larger than the instrumental noise. By overly constraining the retrieval, the retrieval remains too close to the a priori, and not all available information is extracted from the spectrum. [] H2O varies [] 260% at 7.5km.. I find Fig. 2 not really informative. In my opinion here some details on the dD a priori data are needed. Instead of showing the covariance between HDO and H2O it would be much more interesting to know the covariance of dD (or ln[HDO]-ln[H2O]). What is the covariance of ln[HDO]-ln[H2O] you obtain from the model and what is the covariance finally used for the constraints, please give numbers! Maybe show a plot of the LMDz statistics: mean and standard deviation of ln[H2O] and ln[HDO]- ln[HDO]. - Is the mean LMDz profile used as your a priori profile? - Do you use a single/fixed a priori profile for all your IASI retrievals or is there some kind of latitudinal or seasonal dependence in your a priori? This should be clearly stated!

→ Your comment was also raised by Dr. Worden. In the revised manuscript we are fully addressing this. We have replaced Fig. 2 by more informative plots: we plotted the a priori deltad profile as well as the a priori H2O profile. The variabilities (1σ) are presented as well.

Furthermore, yes, the mean LMDz profile is used as our a priori and yes we use a single a priori profile. This is now clearly stated: *To avoid spatial dependency of the results on the a priori profile, a single a priori state vector has been calculated as the mean of an ensemble of HDO and H_2O simulated profiles representative of the whole globe and the whole year. The associated covariances ($\mathbf{S}_{\mathrm{HDO}_L}$,$\mathbf{S}_{\mathrm{H_2O}_L}$,$\mathbf{S}_{(\mathrm{H_2O}_L,\mathrm{HDO}_L)}$) have been computed to build the a priori covariance matrix.*

The approach of tuning the constraint in order to minimize the residual in the spectral fit can be dangerous. Part of the residuals above the noise level might be due to inconsistencies in the spectroscopic line parameters or due to an error in the applied temperature profiles. Addressing these residuals by reducing the dD constraint means that the retrieval misinterprets line parameter uncertainties or temperature errors as an atmospheric dD signal! Furthermore, I am wondering if there might be a connection between the residuals you are talking about and your restricted model atmosphere. For instance, if you limit your model atmosphere to the first 10 km of the atmosphere, doesnt this mean that variability at higher altitudes is seen as residual in your simulated spectra? Concerning H2O: To my knowledge a 260% variability for middle tropospheric H2O is rather large. At least I have never observed such high variabilities in radiosonde climatologies. I think the authors need to comment on that.

→ This comment has also been made by Dr. Worden. As said in response of comment 4 of Dr. Worden, the modification of the Sa have been done to facilitate the convergence of the retrieval. In the previous version this was not well explained and has been clarified in section 3.3.

The 260 % you quote from the previous manuscript was misused. 260 is the variability expressed as the variance (σ^2) erroneously expressed in %. This has been removed, the variability (1σ) is instead plotted on Fig. 2.

Comment 7:
Section 3.4.3, error estimation, page 13065: The authors argue that Schneider and Hase (2011) found that the interference errors are not really contributing to the total error budget and that the same should apply for their retrieval. I think they should be a bit more careful: Schneider and Hase (2011) apply a broad spectral microwindow and retrieve the interference species simultaneously. The authors retrieval is different and the errors might also be different.

→ We do also retrieve a CH4 total column and have avoided major interferences from N2O. This therefore leads to low errors as expected from these parameters. The following statement has been added: *we assume that the interference errors are small as we avoid the part of the spectra where major CH4 and N2O interferences occur.*

Comment 8:
Validating measurements with a nudged model (LMDz): In principle a good idea, because there are few data available for validation. Applying the averaging kernels on the model data is important but at the same time dangerous, since it destroys the independency between the model and the measurement. For example: what if there is a large H2O interference on dD? It means that the retrieved dD is not independent from the retrieved H2O. Then both the retrieved dD and H2O will mainly reflect real atmospheric H2O variability (there will be no real information on dD). This interference will also be present in the averaging kernels. Consequently, the smoothed model data will well agree for dD, but what you actually compare is not dD, but H2O. I think the problems of comparing to model data, which are not independent to the observational data (because of the smoothing with the averaging kernels), should be briefly discussed. Did you also compare to unsmoothed model data? This would not compare apples with apples but at least you would compare independent datasets. I think both comparisons are important: comparison to smoothed and unsmoothed model data.

→ We agree that both types of comparisons have their pros/cons. Other reviewers may never agree to see a comparison between models and unsmoothed data. In our work, we did the comparison for both and we do not find high differences. This means that the partial columns presented in this work is in fact pretty well measured with only a small part of prior information. Because this difference is rather small, we prefer not to extend the discussion in the paper and keep only the comparison to the smoothed model data.

Comment 9:
Additional subsection in Section 3 for discussing the differences to other retrievals?: Some of the aforementioned issues can be addressed within a subsection where the differences of the authors retrieval with the work of the Schneider and Hase (2011) are documented. Some examples of the differences (advantages/drawbacks) to Schneider and Hase that I quickly identified: - The authors fit a rather small spectral microwindow. Advantage: short retrieval time; Drawback: less sensitivity? - This constraint is different to Schneider and Hase. We assume a 100% variability for H2O and a 80% variability for dD for the whole troposphere (deduced from radiosonde climatology and Ehhalt et al., 1974, respectively). - The Schneider and Hase retrieval provides data with much smaller errors, right? - The authors present a retrieval that already works over land and ocean (Schneider and Hase so far only have presented retrievals for ocean scenes). - The Schneider and Hase retrieval is empirically validated, while the authors compare their retrieval to model data. - etc.

→ See also replies to comment 2. With the sentences added (and the section on DOFS) we hope that we have clarified some of the issues. We believe that going much further in listing advantages and drawbacks of one or the other method would require a detailed algorithm inter-comparison exercise. We would very much be in favor of this but this is obviously beyond the scope of the present paper.

Comment 10:
Page 13069, line 19: It is IsoGSM instead of IsoGCM.

→ This sentence has been removed in response of Dr. Worden's comment (#15). Thank you anyway.