Response to Anonymous Referee #2; Dec 2018

Black: Referee's comments
Green: Author's reply

We would like to thank Reviewer #2 for taking the time to review our manuscript. We believe the revised manuscript has improved thanks to thorough and thoughtful comments provided.

1 General Comments

The paper deals with two main topics:

a) first the authors assess the limitations in retrieving the real Water Vapor (WV) vertical variability from the boundary layer to the upper troposphere - lower stratosphere, with a standard inversion of FT solar absorption measurements in the middle-infrared. The study includes the validation of WV profiles retrieved from round based FTS measurements operated from Boulder (Colorado) and Mauna- Loa (Hawaii), via intercomparison with WV profiles measured by state-of-the-art Frost Point Hygrometers (FPH) operated from balloons.

b) Secondly, a sensitivity study is presented, showing the error on retrieved HCN, CO, and C2H6 VMRs due to assuming a less than perfect WV vertical profile.

The subject of the paper is clearly within the scope of AMT. The methods used are scientifically sound, the presentation is sufficiently concise, however it could be improved by rephrasing a few sentences as outlined in the specific comments reported below. The paper does not introduce novel concepts or ideas, however the results of the study will be useful for other scientists using the data presented or data deriving from similar measurements. For this reason I recommend this paper for publication in AMT, after some revisions as outlined below.

My main comment or criticism is about the strategy the authors adopt to deal with the Averaging Kernels (AKs). I agree that the AKs may not be a sufficiently accurate tool to evaluate the smoothing error of the retrieved WV profiles. This is due both to the fact that AKs are only a "linear" approximation of the vertical response function of the measuring system (instrument plus retrieval algorithm), and to the fact that it is generally hard to setup a covariance matrix which represents properly the variability of WV from ground to the Upper-Troposphere / Lower Stratosphere (UTLS). To show the limitations (in your test case) of the smoothing error as derived from AKs and the Rodgers (2000) approach, rather than moving the AKs analysis to the supplemental material, I would have compared, in the main paper, the actual smoothing error (obtained via intercomparisons with FPH) with the estimate of the same error obtained from AKs.

We agree that adding the comparison of smoothed FPH profiles in the main text improves the quality of the manuscript. In the revised manuscript we have included smoothed FPH profiles in Figures 5 and 6, instead of showing them in the supplemental material. Additionally, we have included in the main text results of comparisons between FTIR retrievals with both un-smoothed and smoothed FPH. We kindly refer the reviewer to the revised manuscript for additional/modified text and figures. In particular, table 3 summarizes the finding of both comparisons.

The second general comment I have is about the sensitivity analysis presented in Sect. 5. To my opinion it would be worth to better explain why, after the analysis presented in the first part of the paper, then you start to study a quite different subject, such as the mapping of WV errors on subsequent VMR retrieval of other gases.

While the first section of the paper is focused in the retrieval/comparison of water vapor, we do not consider that section 5 is completely different. The first section considers FPH water vapor as reference and the water vapor sensitivity in the retrieval of other gases also considers FPH as reference, considering that is rare to have fully resolved coincident measurements we believe this is the right place to show both. We do agree that further details might be needed to better explain this second section. In the revised manuscript we slightly have expanded the description of this second part, mainly in the introduction.

More-over, since your measurements cover the middle-infrared, I also expect a sensitivity of the retrieved VMRs to the temperature error. This is already shown in Fig. 3 for WV. What about the error on HCN, CO, and C2H6 VMRs due to the temperature error ? Do you suggest to retrieve also the temperature profile from the same measurements or you are satisfied with the temperature profiles taken from NCEP at NDACC ?

As shown in the manuscript the importance of pre-retrieving water and use it in the retrieval of other gases is important but can be gas and site specific. The effect of using a daily NCEP temperature profile versus a more refined temporal temperature profile (or joint gas/temperature profile) can also be site/gas specific. Similar as WV, full error analysis is considered HCN, CO, and C_2H_6, i.e., temperature profile uncertainty is considered. In the revised manuscript we added a sentence explaining this.

Previous studies, e.g., Schneider and Hase (2008); Schneider et al. (2008), have shown that a joint retrieval of temperature profiles significantly improves the quality of O_3. In general, we do expect similar results for other gases. The temperature sensitivity is out of the scope of the manuscript but have added a brief description of the joint approach from previous studies in the revised manuscript.

In addition, just for reference, below there is a description of our current approach in the estimation of the error in the temperature profiles.

Errors in the temperature profile can have both systematic and random components. In our sites, we quantify these components by comparing radiosonde temperature profiles at or near each site with the daily NCEP temperature profiles. Both radiosonde and NCEP temperature profiles are interpolated onto the retrieval input grid for each site. The mean and the standard deviation of the differences between the NCEP and radiosonde temperature profiles are calculated. The mean of the difference can be as the systematic component of the error, while the standard deviation of the difference can be viewed as the random component. This is carried out for several years. Then, we include this onto the forward model parameters errors to obtain the covariance matrix for the forward model parameters. Ideally, the same can be done for WV, however high quality and coincident measurements of WV for all sites is rare, hence the importance of the sensitivity analysis given in the second part of the manuscript.

2 Specific Comments

P4 L27,28: Constraining is important to select the solution which, among the possible solutions of the ill-posed inversion, is the most likely on the basis of our prior knowledge.

We edited text and included the reviewer's suggestion.

P4 Eq.1: Here it is not clear if your retrieval performs only a single or several Gauss- Newton iterations, because you don't have an iteration index in the Equation. Please, also define clearly the meaning of K. Do you compute it at each iteration ? I guess K is the Jacobian of the forward model with respect to the retrieval parameters, therefore it should be re-computed at each iteration and should show an iteration index.

Thanks for catching this up. We have modified the equation and text accordingly.

P5 L2: It would be interesting here to know which is the spectral range covered by the spectrometers used, and which is the rationale behind the selection of the listed micro-windows (e.g. minimum retrieval error ?).

Measurements at BLD and MLO follow standard measurement protocols of the InfraRed Working Group (IRWG) of NDACC. Several optical band pass filters are used to maximize the signal to noise in the middle and near infrared (~ 700 – 5000 cm^{-1}), described briefly in section 2.1. As mentioned in the manuscript, we do not aim to optimize a retrieval strategy of WV but rather to use a retrieval approach that we have been following in the past year. Past studies have shown that multiple micro-windows can be used. For example, Schneider et al. (2006) applied signatures between 700–1400 cm^{-1}; Schneider et al. (2012) used several micro-windows to retrieve water vapor isotopologues for the FTIR MUSICA retrievals. In particular for WV they used the ~ 2800 – 2900 cm^{-1} range. Later, Barthlott et al. (2017) removed strong micro-windows used in Schneider et al. (2012) with strong absorption and added weaker lines in the 2700 cm^{-1} range. An extensive number of micro-windows can be used to retrieve water. In our case, the micro-windows we use cover a similar range (~2800 – 2840 cm^{-1}) and were tested to maximize the information content and minimize total error and they give consistent results across a wide range of WV columns. We added the following sentence in section 3: *"These micro-windows have been chosen to maximize the information content and minimize total error."*

P5 L8,9: I got an idea of what the authors would like to say, however I suggest to re-phrase more clearly this sentence.

After reviewing section 3, we believe the following sentence:

"*The SNR determines how much influence the spectra has in each micro-window versus the a priori, as well as to characterize the measurement error described in section 3.1"* can be removed since the SNR is primarily used in the error analysis section 3.1.

P5 L10: Which are the "relaxed covariance matrices" that induce oscillations ? Please clarify.

The following paragraph:

"In order to prevent sporadic vertical profile oscillations due to relaxed covariance matrices we implement ad hoc diagonal elements of Sa with a maximum variability of 50% at the surface and exponentially decreasing by altitude with a inter-layer thickness correlation coefficient. A Gaussian correlation with a length of 25 km is used for the off-diagonal elements of Sa."

has been replaced by:

"*The Sa matrix is specified at each layer as a fraction of the a priori profile, which allows for a linear scaled retrieval. We adopted a maximum variability of 50 % in the diagonal covariance and exponentially decreasing by altitude. In order to prevent sporadic vertical profile oscillations, we include a Gaussian correlation length of 25 km in the off-diagonal elements of Sa. This Sa has been optimized in order to obtain similar information content for all a priori presented in section 4.3, a requirement for efficient processing of decades of NDACC spectra."*

P5 L14: If apodization is not used, how broad is the ILS used in the forward model to emulate the instrument effect ?

The spectra are recorded at an OPD of 250 cm giving an unapodized spectral resolution of nominally 0.004 cm^{-1}. At this resolution all atmospheric spectral features are fully resolved in this MIR region. We have specified this OPD in the text to be clear.

P5 L25,26: Here it is not clear how the 0.5% rms error "on the fit" maps onto the retrieved WV. Does this rms error refer to the "residuals of the fit" or directly to the retrieved WV ?

To make it clear, the description of this has been modified accordingly. The revised paragraph is:

"We examined the effect of using more temporally refined temperature profiles. In general, the six hourly temperature profile from the ERA-I reanalysis model, produced by the European Center for Medium-Range Weather Forecasts (ECMWF) (Dee, et al., 2011), follows the daily average temperature profile shape very well for both sites. The root mean square error (rmse) between the six hourly data of ERA-I and daily average temperature is less than 0.5% using 2013 data for both BLD and MLO and the biases are less than 0.25 % for BLD and less than 0.1 % for MLO. These results suggest daily mean temperature should be adequate for retrievals but we further investigated the sensitivity of water vapor to this variability and found that water vapor agrees within 1 % if using the daily average profile. The temperature profile uncertainty is considered in the error analysis in section 3.1."

P6 Fig.1: It would be better to show CH_4 and N_2O absorption contributions∂ with different line colours.

In the updated figure species are identified with the same color for the different micro-windows.

P6 L14-16: This bias is with respect to the a-priori state vector which, in turn, will probably have some bias with respect to the real profile. Note that if the bias of the retrieval was known both in sign and amplitude, then it would be possible to correct for it...

The text has been updated with the above description.

P7 L1-5: Here I would state clearly which is your retrieval vector. Do you retrieve a WV profile using the discretization mentioned in the colour scale of Fig.2a ? Do you include further fitting parameters ? (Such as atmospheric continuum, for example).

The color code scale in Fig. 2a illustrates the kernels on the retrieval grid in the troposphere but the retrieval grid does go up to 120 km. Atmospheric continuum is not considered.

P7 L7-9: Smoothing the high-vertical-resolution profiles via the averaging kernels of the coarse-vertical-resolution experiment is not mandatory, especially if you attribute a "smoothing error" to the profile differences or if you want to try characterizing the smoothing error itself. Therefore I would simply state your choice here, without trying to find a justification, which is also rather fuzzy to my view.

This description has been removed and we further improve this description in section 4. We kindly refer the reviewer to the updated text, especially in section 4.

P7 L18: Off-diagonal elements of Se may play a very important role if the spectrum is oversampled (wrt interferogram) and/or if apodization is used. Please state explicitly that, apparently, this is not your case.

In the case of very high spectral resolution spectra that fully resolve all absorption features where no apodization is used off diagonal elements are in general rendered moot. The interferogram processing may zero fill for speed but no excessive zero filling is used.

P8 L8: Please define also the symbol Kb.

Done

P8 L17-ff: In Fig.3a the error due to interfering species is also shown. Which are the considered interfering species that are not simultaneously retrieved with WV ?

The interfering species are those simultaneously retrieved. We wish to take into account the uncertainty in those lines on the retrieval of the target species. This is typically a small contributor.

P9 L1: This sentence is not clear and may be questionable. Does this mean that your AKs are not a good estimate of the vertical response function of your system (instrument plus inversion scheme) ? Why ?

The paragraph of this section has been updated, based on previous suggestions. As mentioned before the revised manuscript shows smoothed FPH profiles and quantitative comparisons of FTIR with both un-smoothed and smoothed FPH profiles. In general, the biases are lower when compared with un-smoothed profiles pointing out that FTIR AKs may not be adequate tool to use in the comparisons. This is pointed out in the updated conclusions.

P9 L7: Please note that re-gridding via interpolation is, on its own, an arbitrary smooth- ing. So, I do not fully understand why you do not want to use AKs to smooth and re- sample high-resolution profiles prior to intercomparison (as it seems you already did in the plots presented in the supplemental material).

The revised text of this section includes the comparison with smoothed profiles, as suggested in the first general comment. We kindly refer the reviewer to the updated text

P9 L12-14: As shown in Fig. 2c, the FTS retrievals have less than 3 DOFs therefore, why using so many layers for the intercomparison ? The risk is to find biases of different sign in adjacent layers.

This might be true for most gases. In fact, we do normally use total and/or partial columns depending on the number of independent pieces of information. However, a goal of this paper is to assess the ability to retrieve water vapor gradients. Hence, we decided to use these layers. Furthermore, as stated before the averaging kernels might not be a proper tool to assess sensitivity.

P10 Sect. 4.1: The underlying idea is good, however, please note that the variability evaluated here could underestimate the real WV variability, due to the constraint of the retrievals towards the a-priori state vactor. I would have estimated from measurements the variability of the spectrum vs time and would have derived the corresponding “time- mismatch” error covariance matrices relating to the individual WV profiles, using Eq. 3. I suggest to include a comment on this regard (or change approach...).

In the manuscript is already stated that the real variability might be greater because of potential lost variability during retrieval smoothing. In general, we see the stability of the spectrum by means of the signal to noise ratio (SNR), which does not change significantly, however the we see changes in WV spectral features so we see a change in amount of WV. Note that we also aim to see variability of water vapor at several altitude layers and analyzing the spectrum would not give us that information. In the revised manuscript we have removed the last two layers since we do not have enough sensitivity.

P11 Sect. 4.2: Here I did not understand if, from this analysis, you also derive an estimate of the error component to be attributed to the difference between FTS and sonde WV profiles, due to the spatial mismatch of the measurements. I agree that it is hard to derive such an error estimate however, lacking this estimate, I do not see very much the usefulness of this section. Please explain.

As described in this section, we aim to assess the spatial mismatch between the sonde at various altitudes and the maximum sensitivity location of the FTIR. This assessment is already complex, and to our knowledge the first time applied at various altitudes. We actually believe contains great value because as mentioned at the end of this section the spatial mismatch depends on the complex convective dynamics and not only in the coincidence time interval. It is already mentioned that a thorough assessment of the spatial variability would require measurements of an extensive area simultaneously and at different altitudes. However, in the revised text we are explaining further that an error due to spatial mismatch is not derived and only an assessment of the spatial mismatch is aimed.

P13 L2,3: The second effect of a-priori WV on the solution is not clear. Did-you mean that the a-priori WV influences the solution also because it is used as initial guess for the Gauss-Newton iterations ? (This latter effect should be negligible if the retrieval converges properly).

We have re-phrased the sentence as follow:

"The optimal estimation method is influenced by the a priori profile because it may bias the solution of equation 1."

P16, Sect. 5: the link between this Section and the work presented in the previous Sections of the paper is non very clear. Maybe you could state at the beginning of this Section how the work you are going to present is linked with the analysis presented earlier.

The first paragraph of this section has been modified in order to make clear the connection with previous findings.

P16 L1,2: What is the meaning of "expecting WV" in this context ? Usually an optimized microwindow selection scheme tries to avoid spectral interferences from WV and re- lated isotopologues. From the second sentence, however, I understand the opposite.

"expected" has been replaced by *"present"* in the revised text.

P17 Table 3: Which is the rationale for the adopted sorting of interfering species in the Table ? Perhaps their relative importance ? In this case one should assess the retrieval errors due to the interference of ozone. In this spectral region I also expect temperature knowledge to be of importance (see general comments above). Moreover, I understand that the micro-windows used were selected in already publishes papers, however you could at least mention here the rationale with which they were selected / optimized (e.g. with the aim of minimizing the total retrieval error of the gas to be retrieved).

We added the following sentence in the revised text:

"Table 3 presents the interfering species with strong and/or weak absorption signatures within each micro-window for all target gases. In all cases, the selected settings have been chosen in order to maximize the information content and minimize the total error in the retrieval."

Furthermore, the uncertainty due to spectroscopic absorption of ozone and other species (interfering species) is considered for the final error analysis. Since this is described for water vapor, we just added the following sentence:

"Similar as WV, full error analysis is performed, i.e., mainly considering measurement noise error and forward model parameter error (see Sect. 3.1)."

Please see also our comment above regarding the temperature and error analysis.

3 Technical Corrections

P2 L26: Remove double comma between "Zugspitze" and "Germany".

Double comma removed.

P4 L3: Description... described elsewhere (rewording needed).

The sentence has been changed to:

"A thorough description of the FPH measurement technique has been described in Hurst et al. (2011b) and Hall et al. (2016)."

P7 L3: but as explained before, one of the goals...

Done

P11 Caption of Fig. 5: Same as Fig. 4 but for MLO.

Corrected.

P12 L9: As mentioned above, the initial spatial difference...

Corrected.

P13 L17: I do not see the usefulness of putting some figures in the supplemental material when these are recalled and described in the text of the main paper. I would put all the figures in the main paper file (of course only if this operation does not cost too much!)

As mentioned above, two of the initial supplemental figures have been removed (since the smoothed profiles have been included the main text). We decided to keep three figures in the supplement for at least two reasons: (1) we would like to provide a complete and clear process during the analysis; (2) these images are not critical part in the main text, but of course they are mentioned in the main text so the reader can check them out if they want.

P16 Fig. 9: in the vertical axis labels “ppmv” has a small “v”, is this an intentional choice ?

The subscript has been removed.

References

Barthlott, S., Schneider, M., Hase, F., Blumenstock, T., Kiel, M., Dubravica, D., García, O. E., Sepúlveda, E., Mengistu Tsidu, G., Takele Kenea, S., Grutter, M., Plaza-Medina, E. F., Stremme, W., Strong, K., Weaver, D., Palm, M., Warneke, T., Notholt, J., Mahieu, E., Servais, C., Jones, N., Griffith, D. W. T., Smale, D., and Robinson, J.: Tropospheric water vapour isotopologue data ($H^{16}O$, $H^{18}O$, and $HD^{16}O$) as obtained from NDACC/FTIR solar absorption spectra, Earth System Science Data, 9, 15–29, doi:10.5194/essd-9-15-2017.
Schneider, M. and Hase, F.: Technical note: Recipe for monitoring of total ozone with a precision of 1 DU applying mid-infrared solar absorption spectra, Atmos. Chem. Phys., 8, 63–71, 2008, http://www.atmos-chem-phys.net/8/63/2008/.

Schneider, M., Redondas, A., Hase, F., Guirado, C., Blumenstock, T., and Cuevas, E.: Comparison of ground-based Brewer and FTIR total column O3 monitoring techniques, Atmos. Chem. Phys., 8, 5535–5550, 2008.