The authors thank referee #2 for constructive comments and suggestions for revision. In the following our replies are given in blue, we have revised the manuscript accordingly.

Replies to review #2 comments:

Major point

General comment #1: In my opinion, an ACP paper is supposed to tackle a question/issue/problem of a physical or a chemical process in the atmosphere. That is not done at all in this study. It is (merely) the comparison of three different data products. A lot can be learned from this about retrieval parameters, about effects of spatial and temporal sampling or of using different frequency bands, and so on, but here nothing is learned about the physics or the chemistry of the atmosphere. Hence, I suggest to withdraw the paper from ACP and to submit it e.g. to AMT or a related journal. Those journals are also well-renowned and provide room for rather technical papers like this one here.

General response #1: In the revised version we added more information on atmospheric processes. In particular we reflect more on what the differences among the data sets mean for our ability to learn something about the atmosphere.

Minor points

Specific comment #1: P3L31andL32: "boreal winter and "boreal" summer.

Specific response #1: The text has been changed.

Specific comment #2: P4L7-8: That is very simplified, the issue is a lot more nuanced. Please see Frank et al. 2018, 10.5194/acp-18-9955-2018.

Specific response #2: We agree. The description of the variation of the amount of water vapour produced by oxidation of methane in the stratosphere has been adjusted and the work of le Texier et al. (1988) and Frank et al. (2018) have been referenced.

Specific comment #3: P5L10: There should be more recent literature for this than from 1996.

Specific response #3: Yes, there is. Moyer et al. (1996) is still used as reference since it is the seminal paper regarding satellite observations of deuterated water and their application to study the transport of water vapour from the troposphere into the stratosphere. We have now also added newer references, as Nassar et al. (2007), Payne et al. (2007), Sayres et al. (2010), Steinwagner et al. (2010) and Eichinger et al. (2015).

Specific comment #4: P5L28 and P6L5: Eichinger et al. 2015 (10.5194/acp-15-5537-2015) dealt with this issue in a model-satellite comparison.

Specific response #4: This text passage concerns δD comparisons among satellite data sets. In that sense involving the work of Eichinger et al. (2015), that focuses primarily on model simulations, is not really wanted here. Arguably this work helps to understand the observational discrepancies in the δD tape recorder to some degree. However, we have chosen not to make this aspect a major topic of this work. This will be addressed in much more detail in a different study.

Specific comment #5: P6L16: Explain what LT means.

Specific response #5: LT stands for local time. The abbreviation has now been defined in the text.

Specific comment #6: Sect. 4: Why do you start with showing biases and not the actual profiles first? I find that confusing.

Specific response #6: We agree it would be better to show some profiles before the description of the biases. Showing the actual profiles for all comparisons in a single figure is difficult. With 10 comparisons you end up with 20 profiles, making it hard to distinguish anything. Therefore, we show exemplarily the profiles for one comparison in the main manuscript. The profiles for the remaining 9 comparisons are shown in the supplement.

Specific comment #7: I agree with what Mr. Johnson says that it is hard to say what can actually be learned from this study, since there are so many differences between the different methods, one cannot actually see any causalities. I would also appreciate some sort of conclusion that at least states this product/instrument is better here, and this is worse there. Which product can be trusted more where and/or when for making process studies or model comparisons? And maybe also methodologically, which method is best for what? In a future satellite mission, how would the "best" instrument look like, and how can the retrievals be improved?

Specific response #7: As stated in our general response #3 to the first reviewer all data sets have their strengths and weaknesses and that aspects like application, time, altitude or latitude can influence which data sets is better or worse. In the current situation we rate the ACE-FTS data sets as the most reliable in terms of δD.

The biggest concern is probably not to improve the retrieval method itself. There is always the attempt to have the best retrieval with the best characterisation of the instrument. An ideal instrument should have high sensitivity, good coverage spectrally and in space and time, is little influenced by clouds and aerosols, etc. ACE-FTS has a very good signal-to-noise ratio, but a limited temporal and spatial coverage. For MIPAS and SMR coverage is a strength, but the signal-to-noise ratio is clearly smaller than for ACE-FTS. The MIPAS observations are clearly influenced by clouds and aerosols, while the SMR observations are almost insensitive to that. In that sense no single instrument is perfect. What you define as best depends on what you want to achieve.

Specific comment #8: For several reasons the paper is pretty lengthy and that could easily be reduced: The information in Sects. 2 and especially 3 should be reduced to the most important points, technical details and the bulk of the equations should be banished to the supplement. Moreover, the paper is pretty repetitive, e.g. the (first part of the) discussion and the conclusion are not more than summaries of the results. Some restructuring and removing can easily resolve this.

Specific response #8: We prefer to keep data set (Sect. 2) and comparison approach (Sect. 3) description detailed, as done now. Within the water vapour assessment the minor isotopologue data sets are not described in detail elsewhere as done for the main isotopologue (Walker and Stiller et al., in preparation). Given the assessment role, our results should be easily reproducible by others. For that the approach description needs to a have certain level of detail and transparency. Arguably, the calculation approach for δD has become more complicated because of the necessary adjustments of the MIPAS and ACE-FTS data sets to the SMR data set, which do not measure HDO and $H_2O$ simultaneously. However, that should not be brushed over.

Most concerned we are with the conclusion section, which feels quite repetitive with respect to the discussion. We have opted to merge these two sections.

References


Eichinger, R., P. Jöckel, and S. Lossow, Simulation of the isotopic composition of strato- spheric water vapour – Part 2: Investigation of HDO / $H_2O$ variations, Atmospheric Chemistry & Physics, 15, 7003 – 7015, doi:10.5194/acp-15-7003-2015, 2015.

Frank, F., P. Jöckel, S. Gromov, and M. Dameris, Investigating the yield of $H_2O$ and $H_2$ from methane oxidation in the stratosphere, Atmospheric Chemistry & Physics, 18, 9955 – 9973, doi:10.5194/acp-18-9955-2018, 2018.

Le Texier, H., S. Solomon, and R. R. Garcia, The role of molecular hydrogen and methane oxidation in the water vapour budget of the stratosphere, Quarterly Journal of the Royal Meteorological Society, 114, 281 – 295, doi:10.1002/qj.49711448002, 1988.

Moyer, E. J., F. W. Irion, Y. L. Yung, and M. R. Gunson, ATMOS stratospheric deuter- ated water and implications for troposphere-stratosphere transport, Geophysical Research Letters, 23, 2385 – 2388, doi:10.1029/96GL01489, 1996.

Nassar, R., P. F. Bernath, C. D. Boone, A. Gettelman, S. D. McLeod, and C. P. Rinsland, Variability in HDO/$H_2O$ abundance ratios in the tropical tropopause layer, Journal of Geophysical Research, 112(D11), D21,305, doi:10.1029/2007JD008417, 2007.

Payne, V. H., D. Noone, A. Dudhia, C. Piccolo, and R. G. Grainger, Global satellite measurements of HDO and implications for understanding the transport of water vapour into the stratosphere, Quarterly Journal of the Royal Meteorological Society, 133 (627), 1459 – 1471, doi:10.1002/qj.127, 2007.

Sayres, D. S., et al., Influence of convection on the water isotopic composition of the tropical tropopause layer and tropical stratosphere, Journal of Geophysical Research, 115(D14), D00J20, doi:10.1029/2009JD013100, 2010.

Steinwagner, J., S. Fueglistaler, G. P. Stiller, T. von Clarmann, M. Kiefer, P. Borsboom, A. van Delden, and T. Röckmann, Tropical dehydration processes constrained by the seasonality of stratospheric deuterated water, Nature Geoscience, 3, 262 – 266, doi: 10.1038/ngeo822, 2010.

Walker, K. A. and Stiller, G. P.: The SPARC water vapour assessment II: Data set overview, in preparation.