We would like to thank the referee for the positive assessment and several helpful comments. Below we reply to the issues raised by the referee, where
blue repeats the reviewer's comments,
black is used for our reply,
and green italics is used for modified text and new text added to the manuscript.

Review of Total Column Water Vapour Retrieval from S-5P/TROPOMI in the Visible Blue Spectral Range by Christian Borger et al.

This paper presents a study where the authors have developed an algorithm to retrieve total column water vapour (TCWV) from 'blue' portion of the visible spectrum band on Sentinel 5-P/TROPOMI, and make comparisons with some validation sources.

This is an interesting study, and is a good example of exploiting the potential satellite data beyond that which was intended by the developers of TROPOMI. It is certainly of interest to the AMT community, and I recommend publication, if the following points and concerns are addressed.

We thank the reviewer for the general positive statement.

General Comments

In general, the paper lacks context to the wider TROPOMI picture, and only very briefly states the justification for why the authors attempt TCWV retrieval in the blue part of the visible spectrum. Even here, this discussion is in relation to the 'red' portion of the visible spectrum (where the authors have experience), and not products from other spectral regions (e.g. the SWIR). Especially since the errors can get quite high with these retrievals (up to 50%), I think the authors need to discuss the added benefit of retrievals from the 'blue' band to TROPOMI. The authors very briefly give some discussion to this in the conclusions (processing time, etc), this needs to be expanded upon. The authors identify TCWV products from other spectral regions on other instruments (e.g. TIR from AIRS), but there is no discussion on how TCWV from the 'blue' band compare to the TCWV products from these other instruments and spectral regions. This discussion needs to be included into the paper.

Section 1 – The Introduction.

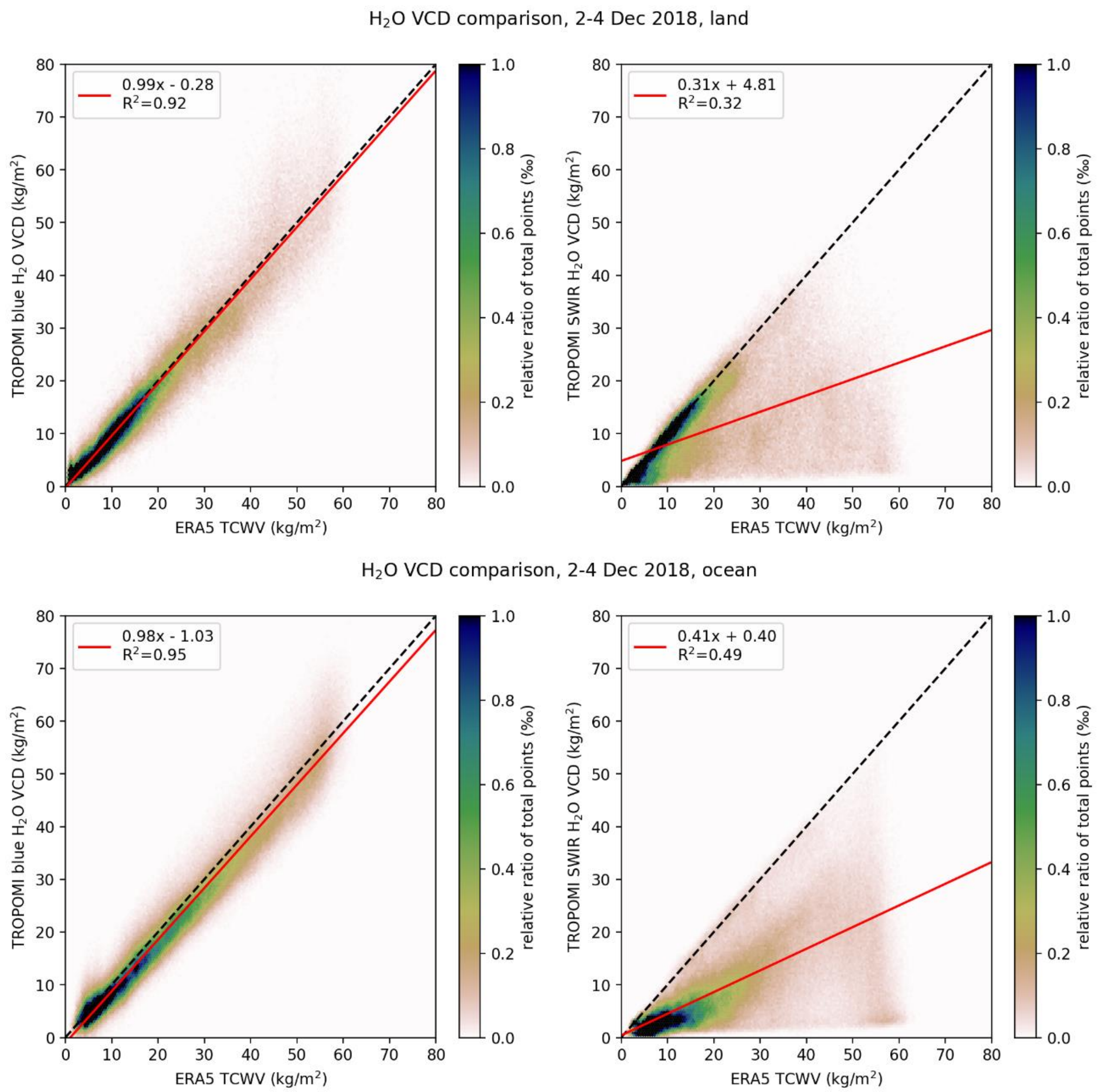
Context is lacking in this section, it is unclear as to the advantages of retrieving TCWV from this particular spectral region. Given the other spectral bands and satellites that the authors mention, and not just the 'red portion' of the spectrum.

We added further explanations of the advantages of the visible "blue" compared to other spectral ranges and satellites. These include 1) similar sensitivity for ocean and land surface allowing for global coverage, 2) possible retrievals under partly-clouded conditions, 3) compared to the thermal IR a much higher sensitivity for the near-surface layers and 4) simple spectral analysis, i.e. no need for forward model calculations. The following text is added to the introduction:

The visible spectral range is particularly interesting for the retrieval of total column water vapour (TCWV): in contrast to the microwave range it has a similar sensitivity for ocean and land surface allowing for global coverage. Also, it is possible to conduct retrievals under partly-clouded conditions and, in comparison to the thermal infrared, it has a much higher sensitivity for the near-surface layers. Furthermore, the spectral analysis is straightforward, i.e. no forward model calculations are necessary.

Indeed, the authors have not referenced the study covering water vapour retrieval from the 2305-2385 nm TROPOMI spectral band (Schneider et al., 2020), which would have been in the AMT discussion forum at time of submission. They have also missed water vapour retrieval studies from the Shortwave Infrared (SWIR) in general, this is very strange since the SWIR gives good sensitivity to the surface e.g. (Trent et al., 2018), and surface sensitivity is clearly one of the key selling points for the 'blue' region.

Thanks for the hint of the missing reference, we added it in the revised manuscript. We had a look at the H_2O data provided by Schneider et al. (2020) and compared them to our dataset for the arbitrary chosen time-range 2 to 4 December 2018. As TROPOMI's SWIR and the vis bands have different pixel sizes and collocations, we gridded the data to a 0.25°x0.25°. We only selected clear-sky observations, i.e. with cloud fractions < 20% in the visible spectral range. As Schneider et al. (2020) do not provide cloud fraction data, we only compare those gridcells that are classified valid for our TCWV data. For a better understanding of the differences, we separated the data into data over ocean and data over land. The results are given in the Figures below (left panel: visible "blue" TCWV; right panel: SWIR TCWV).



The "blue" TCWV shows excellent agreement to ERA-5 over both surface types with slopes around 1.0 and R^2 of 0.92 and 0.95.

In contrast, though the large majority of the SWIR TCWV data are concentrated along the 1-to-1 diagonal over land, we also find a large point cloud which negatively influences the linear fit. Over ocean, a similar point cloud can be observed, and the majority of the points are not concentrated along the 1-to-1 diagonal. Reasons for these major discrepancies could be the missing cloud information within the SWIR data and maybe also saturation effects of the retrieval. Additionally, the low sensitivity of the SWIR retrieval for near-surface layers over the dark ocean surface could lead to potential discrepancies.

Section 3 – A priori water vapour

While the aim of this section is clear, the story of how the methods that are used are a little unclear, and the general message gets a little lost in the details. Here in section 3, I think a few sentences that summarises the processes that are undertaken in sections 3.1 and 3.2 would be useful.

We added further explanations of our steps within the Section's introduction The following text is added:

We proceed as follows: First, we evaluate how well the method used to calculate the water vapour scale height can reproduce the COSMIC profiles via an AMF comparison. Then we examine how the scale height can be parameterized globally and investigate for a parameterization over ocean and land separately. Finally, we implement the parameterization in an iterative retrieval scheme and evaluate the new estimates of the H_2O VCD.

We also rearranged the introduction and moved the description of the COSMIC data into a separate subsection.

Readers may well be unfamiliar with COSMIC and ROMSAF, and some background would be beneficial here. Given the importance of the a priori profile on the results, some discussion on the biases associated with using COSMIC as a base would be welcome.

We added further details on the used COSMIC data and brief discussion on possible biases associated with using COSMIC within the new subsection in Section 3:

3.1 COSMIC water vapour profiles

For our investigations we use profile data retrieved from measurements of the Constellation Observing System for Meteorology, Ionosphere, and Climate (COSMIC, Anthes et al., 2008) program provided by the Radio Occultation Meteorology Satellite Application Facility (ROMSAF). The COSMIC data are based on the GPS radio occultation (RO) technique, which provides high resolution vertical profiles of bending angles (Hajj et al., 2002) that can be used to retrieve the atmospheric refractivity. Since the atmospheric refractivity is dependent on the air pressure, the air temperature, and the water vapor pressure (Smith and Weinstraub, 1953), GPS RO allows for the retrieval of profile information under all-weather conditions with a high vertical resolution of approximately 100m in the lower troposphere up to 1km in the stratosphere (Anthes et al., 2011) and an accuracy of around 1g/kg (Heise et al., 2006; Ho et al., 2010b) while having an almost uniform global distribution (Ho et al., 2010a).

The ROMSAF profiles have been retrieved via a 1D-VAR scheme within a reprocessing initiative for creating climate data record (CDR) v1.0. Given the strict product requirements and the validation studies with ERA-Interim and radiosondes (Nielsen et al., 2018), biases associated with using COSMIC should be of secondary order.

We use data retrieved between 2013 and 2016, which accumulates to approximately 1.6×10^6 profiles.

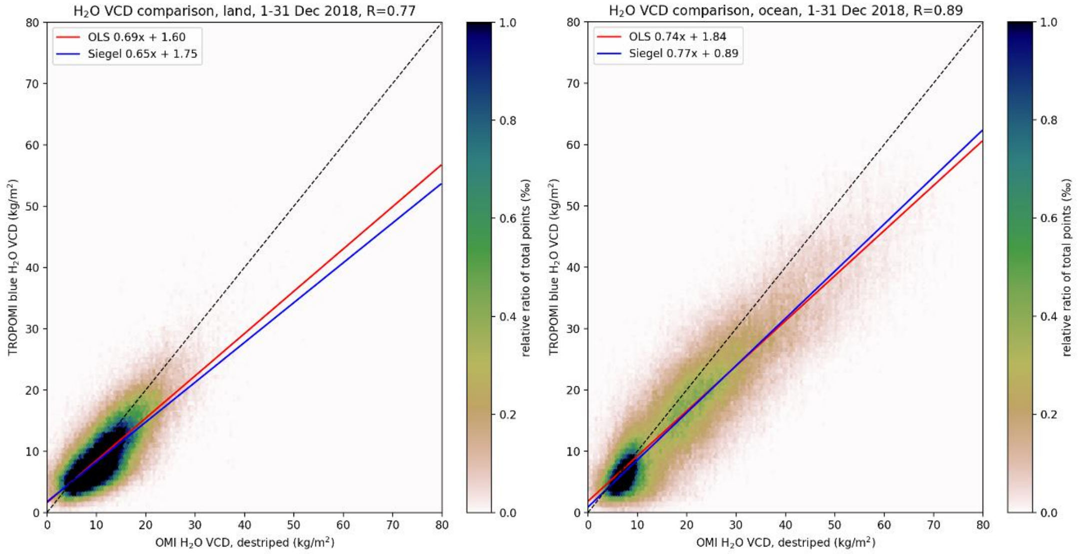
Section 6 – Validation study.

More detail about the SSMIS and SuomiNet data are needed here, fundamentally why did the authors choose these measurements for their validation, and why are they appropriate for TROPOMI inter-comparisons?

SSMIS and SuomiNet measure in the microwave and radio spectral range, respectively, which allows for the retrieval of TCWV under all-sky conditions with high accuracy (~1kg/m^2 and ~2kg/m^2, respectively). Thus both datasets are widely considered as gold standard and used as references within our study. These informations are given in the introduction of every subsection of the validation section.

In the introduction the authors identify some specific advantages of retrieving TCWV from the blue part of the spectrum as opposed to the red, which has been done previously. To me, the validation section would have been a good opportunity to compare against TCWV data from the red part of the spectrum, not necessarily from TROPOMI, but from other satellites. Indeed comparisons from other 'blue' TCWV measurements from OMI as identified by the authors would have been useful, to identify any potential differences between the spectral regions. Or, please justify their exclusion from the study. Validation is largely done using TCWV retrieved from microwave instruments, are there any particular biases associated with TCWV retrievals, when compared against the visible band?

There a lot of TCWV datasets available that could be used for validation studies. We decided to restrict ourselves to SSMIS and SuomiNet, because both data sets are based on measurements which are insensitive to clouds and are known for their proven high accuracy. In the case of OMI, it has to be taken into account that the instrument is affected by a row anomaly since 2007 and hence almost half of the data have to be used with caution. Also, OMI does not cover the red spectral range. Nevertheless we had a look at the OMI TCWV data provided by Wang et al. (2019). We applied the required filter criteria (cloud fraction < 5%, RMS < 0.001, MDQFL=0, cloud pressure > 750hPa, and TCWV < 75kg/m^2) and also destriped the OMI H_2O SCD. For our TROPOMI data we only applied the proposed cloud fraction filter. As comparison time range we selected December 2018. Since OMI has a much larger pixel size than TROPOMI, we gridded both data to 0.25°x0.25° grid. For a better understanding of the differences, we separated the data into data over ocean and data over land. The Figures below show the results of the comparison.



The 2D histograms reveal that for both surface types the products do not agree well: the OMI VCDs are much larger than our TROPOMI VCD, by almost 30%. However, considering the higher signal-to-noise ratio of TROPOMI and the findings of our validation study using even looser filter criteria than Wang et al. (2019), we conclude that the TROPOMI VCDs are more reliable than the OMI VCDs.

Other

With regards to the English, the paper could be improved given a review by a native English speaker. I will not go into detailed specifics in this review, but this paper would benefit from a greater use of punctuation (i.e. commas).

We agree that our English (grammar, spelling, punctuation) is not perfect as we are no native speakers. We carefully checked the manuscript with respect to English language and punctuation. In addition, Copernicus will provide a thorough language editing before final publication.

There are a large number of figures in this paper, and while I applaud the efforts of the authors for the detail they have gone into, 26 figures + 14 in the appendices is unneeded for a journal paper, which should only be showing the key highlights. I suggest to the authors to place some of these figures in supplementary materials.

We rearranged our figures in the revise manuscript. We have moved former Figures 5,6, A3-A6, and A8-14 to the Supplementary Material.

In addition, on some of the Figures the axis labels and legends are quite small (e.g. Figure 5), I recommend that the authors increase the font size.

Thanks for the hint. We carefully inspected every figure and modified the labels, legends, etc. if necessary.

Additional thoughts

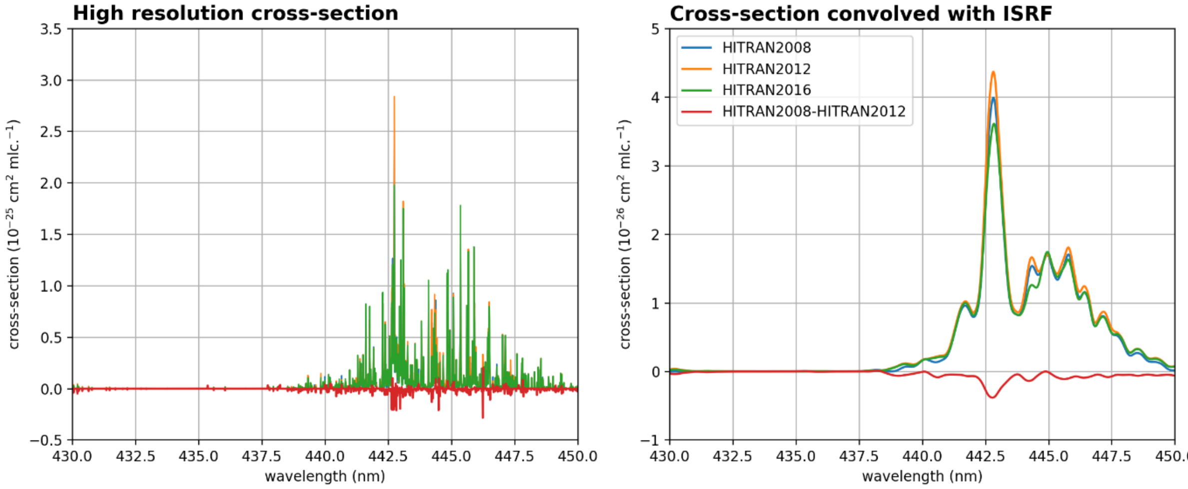
I note that the authors have stated that their dataset is available on request, I would strongly encourage them to place their dataset into an online repository.

Currently, we are working on the improvement of the input parameters of the retrieval, e.g. we are working on a standalone cloud fraction and cloud top height product. Both are major tasks and will take several months to be available. Thus we cannot include the improved cloud products in the current study. However, in light of the expected improvements to come, we consider the presented TCWV product not as the final version and refrain from an explicit publication of this dataset But we intend to provide the updated data set after these improvements have been implemented. The current data will be made available on request.

Specific comments.

On p3, line 75. The authors state that the absorption is weak in the fit window, and hence the line lists vary. I am not convinced by this argument, to me just because absorption is weak the uncertainty of the spectral lines shouldn't necessarily be higher, does HITRAN state this? What exactly is different between the databases, is it the position of the lines, the number of lines?

A similar question has also been risen in the Short Comment by Eamon Conway. What we meant by "variation" of the line list is the distinctive disagreement/change of the strength of the peak absorption between the different HITRAN versions. Please see the figure below: The left panel depicts the high-resolution cross-sections and the right panel the same cross-section convolved with the TROPOMI ISRF. Though the high-resolution cross-sections seem to be quite similar, the convolved cross-sections reveal an alternating pattern going from HITRAN 2008 to 2016.



We added the figure above and the following text to the appendix section of the water vapour cross-section:

Figure A3 compares the absorption cross-sections of the different HITRAN versions. For the high-resolved cross-section (left panel) the differences between the versions are hardly visible, however, after the convolution with the TROPOMI ISRF (right *panel), distinctive differences in the peak absorption are clearly visible: in comparison to HITRAN2008, the absorption peak of HITRAN2012 is approximately 7-9% higher than HITRAN2008 and the absorption peak of HITRAN2016 is approximately 7-9% lower than HITRAN2008.*

P4, line 91. Why are the data shown from Figures 1 and 2 based on different orbits? It'd be more consistent to show the same orbit results.

We changed Figure 1 accordingly and took an example from the orbit depicted in Figure 2.

P5, Equation (4) – The calculation of atmospheric refractivity is highly dependent of air pressure and air temperature. Therefore the scale height for the water vapour profiles is highly dependent on knowledge of these factors. How accurate is this knowledge, and how sensitive is the calculation of scale height to errors in the knowledge of air pressure and air temperature?

ROMSAF requests in their product requirements a target pressure accuracy less than 1 hPa from 0-50km and a target temperature accuracy of 1-2K from 0-5km and 1K from 5-30km. The provided geopotential height thus has an accuracy of around 50-500gpm in the troposphere which is approximately 5-50m in geometric height (increasing with altitude). Considering typical specific humidity errors of around 1g/kg, errors in the altitude grid are negligible within the scale height calculations.

P5, line 143. Why 63% (assuming this is total water vapour up to 150 hPa)? Is this because there is not much water vapour above this point? If so, this should be stated.

63% is the fraction of total vertical column that should be encountered within the sub-column between ground and first scale height: 1-1/e = 0.63 → 63%

P5, line 147. Why 7%?

7% is a typical value for the ocean surface albedo (Tilstra et al., 2017).

P6, I think Figures 5 and 6 could be moved to the appendices or supplementary material. For Figure 5, it is very unclear to me exactly what causes the 'bad' profiles, as opposed to the good profiles, this needs to be expanded upon.

We have moved Figure 5 to the Supplementary Material. The "bad" profiles occur if a sharp decrease of the water vapour concentration with altitude exists. Such profiles occur when a moist boundary layer is topped by a dry free atmosphere. The "good" profiles are associated with a well-mixed troposphere and thus decrease with altitude following an exponential decay. This is clarified in Section 3.2 of the revised manuscript as follows:

In general, bad agreement (left column) occurs for profile shapes in which a sharp gradient is observed in the lower troposphere and from that quasi-constant values with altitude. Such profiles usually occur when a moist boundary layer is topped by a dry free atmosphere. Nevertheless the maximal absolute relative AMF-deviations only have values around 15%. In contrast, good agreement (right column) is found for profile shapes following an exponential decay with altitude, which indicates a well-mixed troposphere.

P7, line 187. How are ocean and land differentiated? What about heterogeneous scenes or lakes?

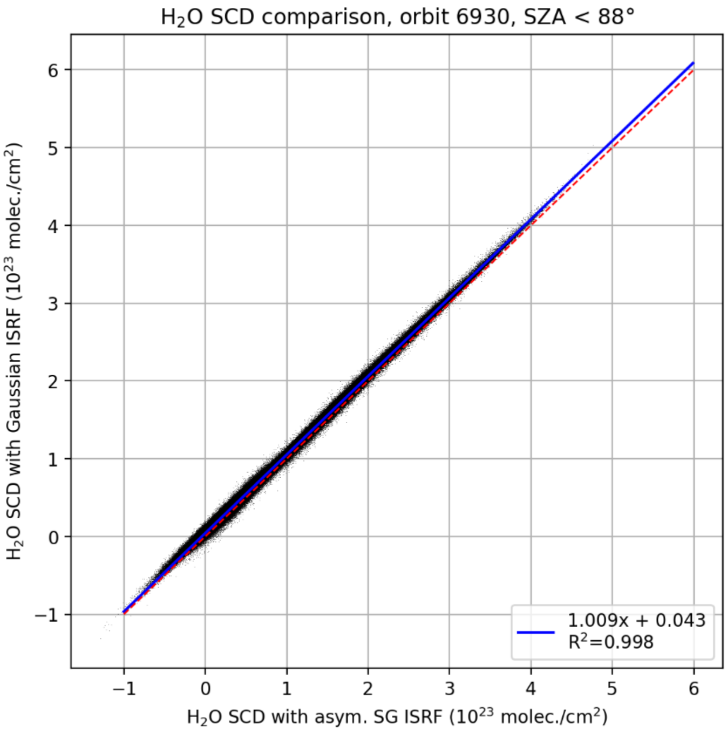
We use a land-sea mask derived from GSHHS coastline data, in which we use the pixel center coordinates for the separation into land and ocean. As the NDVI is not available over lakes, we treat them as ocean.

Figure 14. The authors claim that there is a distinct separation of the H2O VCD between land and ocean, which vary between albedo versions. I did see it eventually, but it is quite subtle, this might be represented better if focused upon? Also, I don't feel that the inclusion of this Figure is particular beneficial to the paper, and can be placed in supplementary material.

We changed Figure 14 and highlighted the areas of distinct separation between land and ocean with black circles in each subfigure.

P10, section 5. There is no discussion on instrumentation errors such as ILSF biases and radiometric errors. Can you comment on the impact of these?

Radiometric errors are already included within the fit error of the DOAS analysis. Considering ISRF/ILSF biases we investigated the impact of using a Gaussian ISRF instead of an asymmetric Super-Gaussian and compared the resulting H_2O SCDs (see Figure below) for the same orbit as in Figure 2 of our paper.



In principle the SCDs using the Gaussian ISRF are slightly higher (by about 1%) than the SCDs using the asym. Super-Gaussian ISRF. We added this Figure to the revised manuscript and also added the following text to Section 5.1:

To estimate errors associated with ISRF biases, we calculated the H_2O SCD using a Gaussian ISRF (instead of an asymmetric Super-Gaussian) for orbit 6930 and compared them to the SCDs from the "standard" retrieval setup for a SZA < 88°. The comparison depicted in Figure S3 reveals that the SCDs using the Gaussian ISRF highly correlate with the "standard" SCDs and only differ by approximately 1%. Considering the much higher fit errors, errors due to biases in the ISRF are negligible.

P11, line 316. Is the spectroscopic uncertainty not as significant as any of these other factors? It would be useful to state the impact of this uncertainty in relation to the others. These are mentioned in the summary, but would be useful in the main text as well.

The statement about errors of spectroscopic uncertainty is now included in the description of “Uncertainties in the slant column density” (Section 5.1.):

Considering the LP-DOAS comparisons (see Sect. 2.1 and Appendix B) we estimate these errors to be around 5% for this study.

P15, summary and conclusion. Can you comment on the uncertainty of the retrievals (you say up to 50%) in comparison to TCWV retrievals from other spectral bands.

The uncertainty of our retrieval is around 10-20% for favourable and 20-50% for unfavourable observation conditions. Most uncertainty estimates of retrievals in other studies are based on validation studies rather than theoretical calculations and thus only provide absolute values of the error. For example Bennartz and Fischer (2001) estimated an error of 2.5kg/m^2 for their MERIS TCWV product, but most of their reference measurements were taken under dry atmospheric conditions of 20-25kg/m^2 (Bauer 2009). Hence (and taking into account our additional comparison studies) we conclude that our estimated uncertainty is within or even better than the typical range of NIR and IR TCWV retrievals.

P15, lines 469-471. Here the authors state that the retrieval allows for a fast execution of large datasets. This to me is one of the key benefits of the retrievals in this waveband, however this is the first time that it has been mentioned in the manuscript. A brief discussion on processing times, and comparisons with other TCWV products would be beneficial.

As we are using a linearized retrieval scheme, the spectral analysis of one TROPOMI orbit takes approximately 1 min (depending on the IT system). As a rule of thumb Beirle et al. (2013) found an increase of speed of 3 orders of magnitude by going from non-linear to linear fit for their MATLAB routine (see Table 3 in their paper). We added the following text to Section 2.1:

According to Beirle et al. (2013) the computational speed increases by 3 orders of magnitude by going from non-linear to linear fit for their MATLAB routine (see Table 3 in their paper).

Technical P2, line 45 – A reference should be added for the increase in TROPOMI spatial resolution to 3.5x5.6.

We added a reference to the TROPOMI L1b Product Readme File (Rozemeijer and Kleipool, 2019).

P3, line 83 – Band 4 is mentioned for the first time, please identify what TROPOMI Band 4 is.

We provided further information on Band 4. We changed the text in Section 2.1 accordingly:

Due to the high daily data volume of the TROPOMI L1B radiances the execution of a non-linear fit without high performance infrastructure is demanding in computation time. For instance TROPOMI's UVIS Band 4, which covers the spectral range of 400-499nm, generates about 40 gigabyte per day. Therefore, we implemented a weighted linear least squares fit …

A number of the equations appear to be missing equation numbers, e.g. the AMD calculations p4, lines 106, 111.

We carefully revised the numbering of the equations.

P5, line 120 – Could you provide an explicit statement for what is meant by scale height please.

We assume a water vapour profile following an exponential decay with altitude, i.e. $\rho_w(z) \sim \exp(-z/H)$ where H is the scale height of water vapour. Thus, we define H as the altitude (from ground) at which 1-1/e=63% of the TCWV are accumulated. The text in the introduction in Section 3 is now:

Weaver and Ramanathan (1995) approximated the water vapour profile by an exponential decay with altitude:

$n_v(z) = n_0$ e^{$-z/H_v$}

where H_v is the scale height of water vapour, which they defined as: …

P5, line 130 – Please define ROMSAF.

We added the full name of the acronym.

P5, line 135 – GPS RO allows to retrieve profile information -> GPS RO allows for the retrieval of profile information

Thanks for this hint.

P15, line 468 – and Sentinel 5

We also added Sentinel 5.

P19, line 534 – greak -> great

Thanks for the hint.

References

Schneider, A. et al. (2020) 'First data set of H2O/HDO columns from the Tropospheric Monitoring Instrument (TROPOMI)', Atmospheric Measurement Techniques, 13(1), pp.85–100. doi: 10.5194/amt-13-85-2020.

Trent, T. et al. (2018) 'Observing Water Vapour in the Planetary Boundary Layer from the Short-Wave Infrared', Remote Sensing, 10(9), p. 1469. doi: 10.3390/rs10091469

Literature:

Bauer, P.: 4D-Var assimilation of MERIS total column water-vapour retrievals over land. Q.J.R. Meteorol. Soc., 135: 1852-1862. doi:10.1002/qj.509, 2009.

Beirle, S., Sihler, H., and Wagner, T.: Linearisation of the effects of spectral shift and stretch in DOAS analysis, Atmos. Meas. Tech., 6, 661–675, https://doi.org/10.5194/amt-6-661-2013, 2013.

Bennartz, R, and Fischer, J.: Retrieval of columnar water vapour over land from backscattered solar radiation using the Medium Resolution Imaging Spectrometer, Remote sensing of Environment, 78.3, 274-283, 2001.

Nielsen, J., Gleisner, H., and Lauritsen, K.: Validation Report: Reprocessed Level 2B and 2C 1D-Var products, Tech. Rep. SAF/ROM/DMI/REP/1DVAR/001, ROMSAF, https://www.romsaf.org/product_documents.php, product version 1.0, 2018.

Rozemeijer, N. and Kleipool, Q.: S5P Mission Performance Centre Level 1b Readme, Tech. Rep. S5P-MPC-KNMI-PRF-L1B, https://sentinel.esa.int/documents/247904/3541451/Sentinel-5P-Level-1b-Product-Readme-File, product version V01.00.00, 2019.

Tilstra, L. G., Tuinder, O. N. E., Wang, P., and Stammes, P. (2017), Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY, J. Geophys. Res. Atmos., 122, 4084– 4111, doi:10.1002/2016JD025940.