We revised the manuscript in line with the recommendations from the referees. In the following, we will list the most important modifications with respect to the AMTD version of the manuscript. A pdf manuscript compiled with LatexDiff is attached to this "Author's Response".

- We added further explanations of the advantages of the visible "blue" spectral range in the introduction.
- We also provide a brief introduction of the relation between the water vapour profile and cloud screening and surface sensitivity in Section 2.
- In Section 3 we give more information about the COSMIC water vapour profile data in a separate subsection and added few sentences about our procedure of the scale height parameterization at the beginning of the section.
  Also, we calculated AMF from ERA-5 water vapour profiles for the orbit including an atmospheric river and compared the resulting VCDs with the VCDs from the iterative scale height approach.
- We further included error estimations for ISRF biases in Section 5.
- To make the manuscript more concise, we rearranged the figures and moved former Figures 5,6, A3-A6, and A8-14 to the Supplementary Material.
- In the Appendix B we extended the discussion about the choice of the water vapour cross-section by adding plots of the respective absorption cross-section spectra.
- We carefully inspected each figure and modified the labels, legends, etc. if necessary.

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.

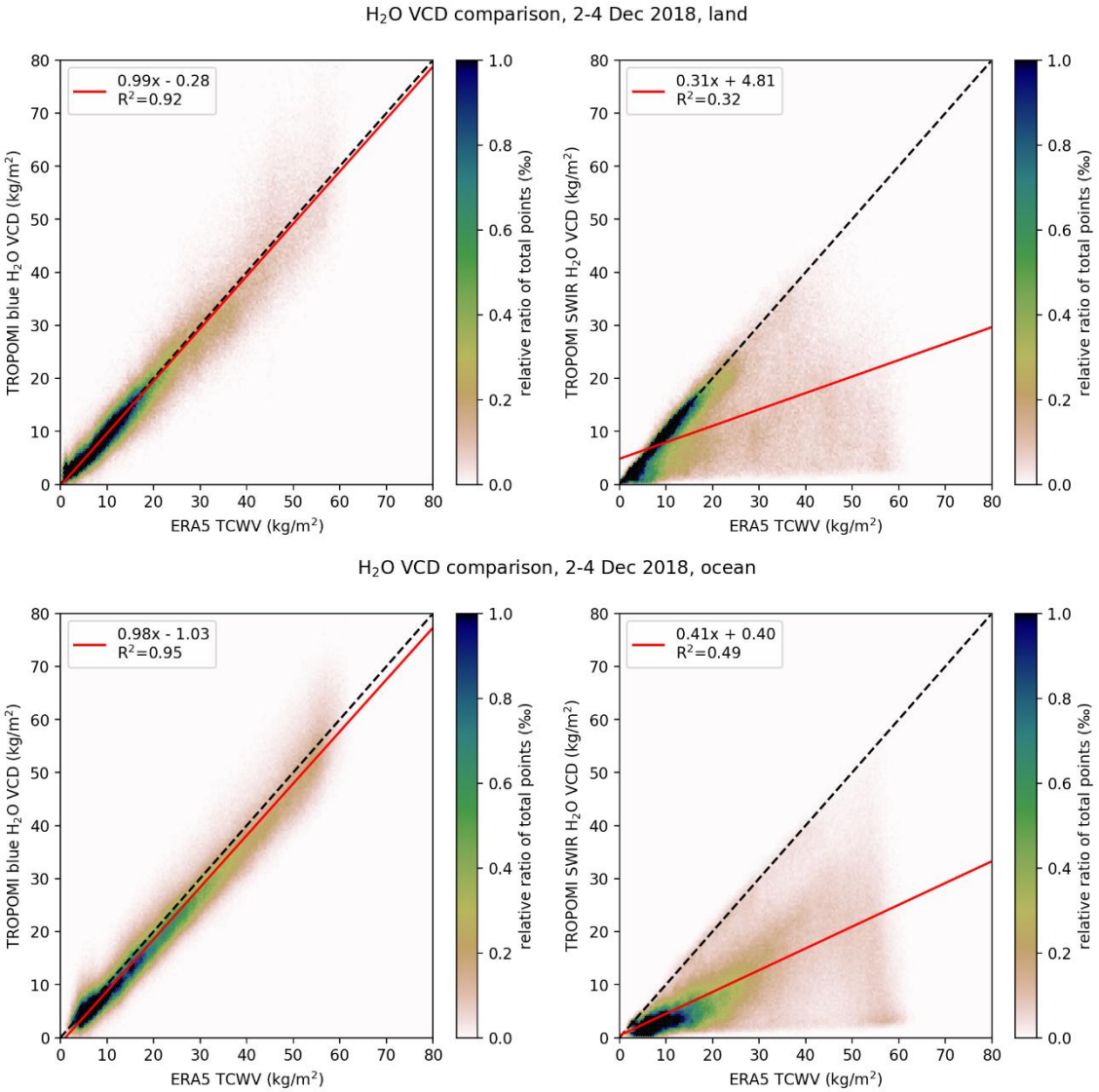
Context is lacking in this section, it is unclear as to the advantages of retrieving TCWV from this particular spectral region. Given the other spectral bands and satellites that the authors mention, and not just the 'red portion' of the spectrum.

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 x 10^6 profiles.

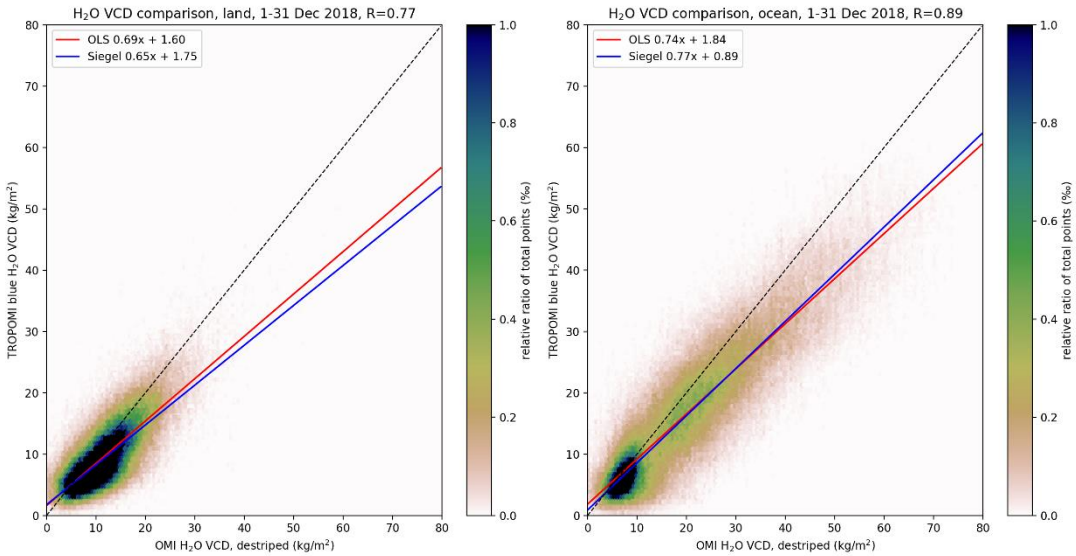
Section 6 – Validation study.

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

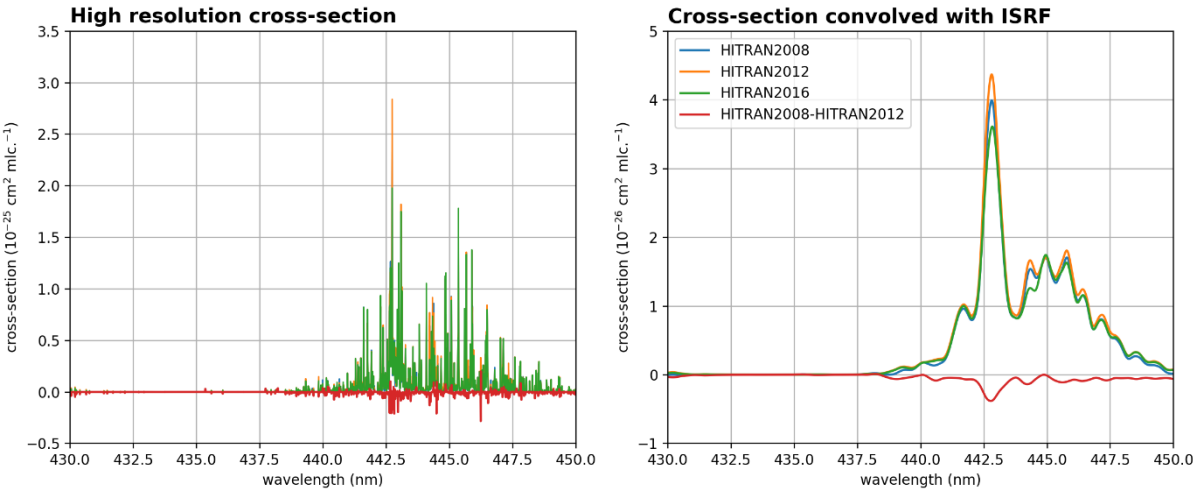
I note that the authors have stated that their dataset is available on request, I would strongly encourage them to place their dataset into an online repository.

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?

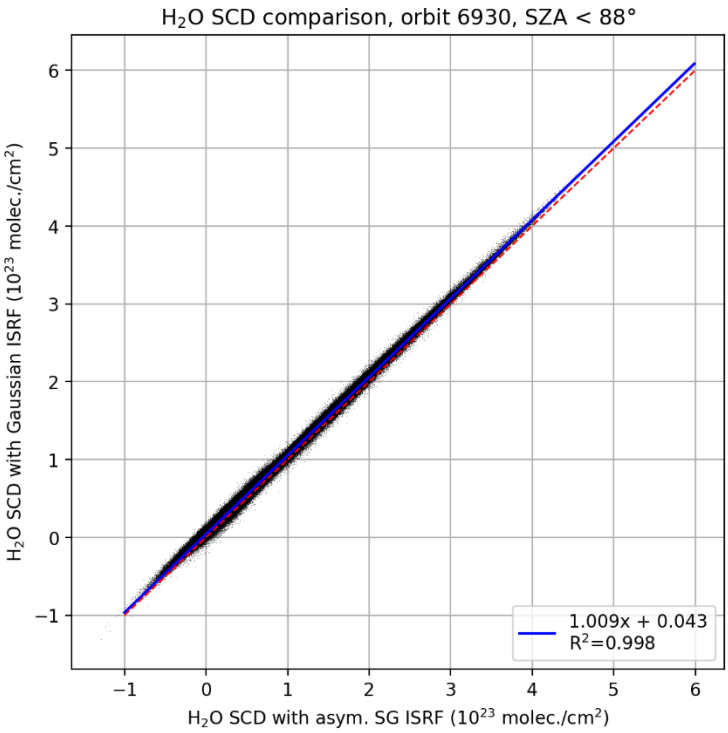
We use a land-sea mask derived from GSHHS coastline data, in which we use the pixel center coordinates for the separation into land and ocean. As the NDVI is not available over lakes, we treat them as ocean.

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.

We would like to thank the referee Rüdiger Lang for the constructive and helpful comments and questions. Below we reply to the issues raised by the referee, where
blue repeats the reviewer's comments,
and black is used for our reply,
and green italics is used for modified text or text added to the manuscript.

The paper by Borger et al., on total column water vapour retrievals from Sentinel-5P (S5P) is demonstrating the large potential of water vapour retrievals in the blue and visible spectral range to yield an accurate estimate of the total water vapour column(TWVC) largely independent from model data and capable to cover all surfaces. This type of TWVC product from instrumentation like GOME-1/2, SCIAMACHY, OMI and TropOMI, therefore serves as an important product for the evaluation of (re-)analysis NWP model data output.

The paper by Borger et al. is overall well written and structured and apart from presenting the very first results of this type of TWVC product from the S5p mission the paper also presents an interesting and novel approach of employing sub-column water vapour profile information to TWVC retrievals, via a parameterization of the water vapour atmospheric scale-height for various conditions, like surface type and different observation geometries. The strong gradient of water vapour in the atmosphere, always required an implicit knowledge of its vertical distribution often “hidden” in the way the conversion from slant to vertical column densities (SCD to VCD) via the calculation or estimation of the air-mass factor (AMF) has been approached.

The paper is an important contribution to this problem, since it approaches this issue for the first time explicitly, and shows convincing improvements, especially when retrieving TWVC in the vicinity of clouds, or evaluating, and improving the performance for various surface reflectance conditions. However the exact relation between cloud coverage, cloud height and retrieval performance remains to this reviewer still - at least to some extent – obscure and while I can highly recommend the paper for publication, I would like the authors to address this and a few other issues before.

We thank the reviewer for his positive general statement.

1. Scale-height parameterization

The paper goes in depth on a specific parametrising of the a priori (better “first guess”) water vapour profile using a parameterization of the water vapour scale height. While the motivation to introduce knowledge on the water vapour profile is in principle clear to any reader familiar with TWVC DOAS-like approaches, for the non-expert reader, the relation to cloud screening and surface sensitivity is not apparent.

We thank the reviewer for this note and added further explanations on the relation between the water vapour profile and cloud screening and surface sensitivity in Section 2.2 in the revised manuscript. The following text and Figure have been added to the revised manuscript:

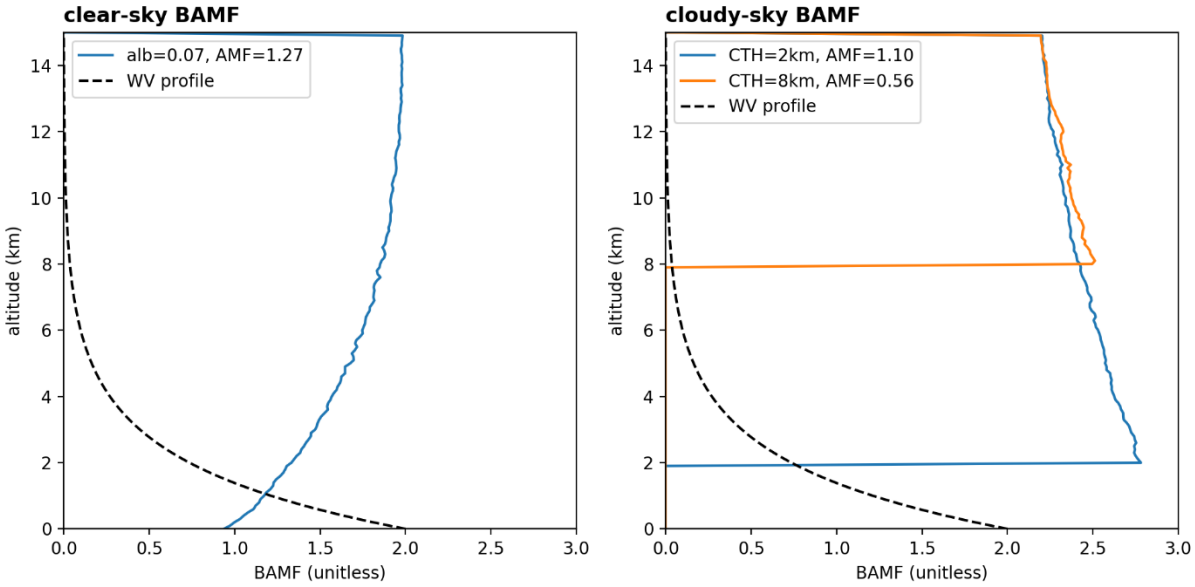
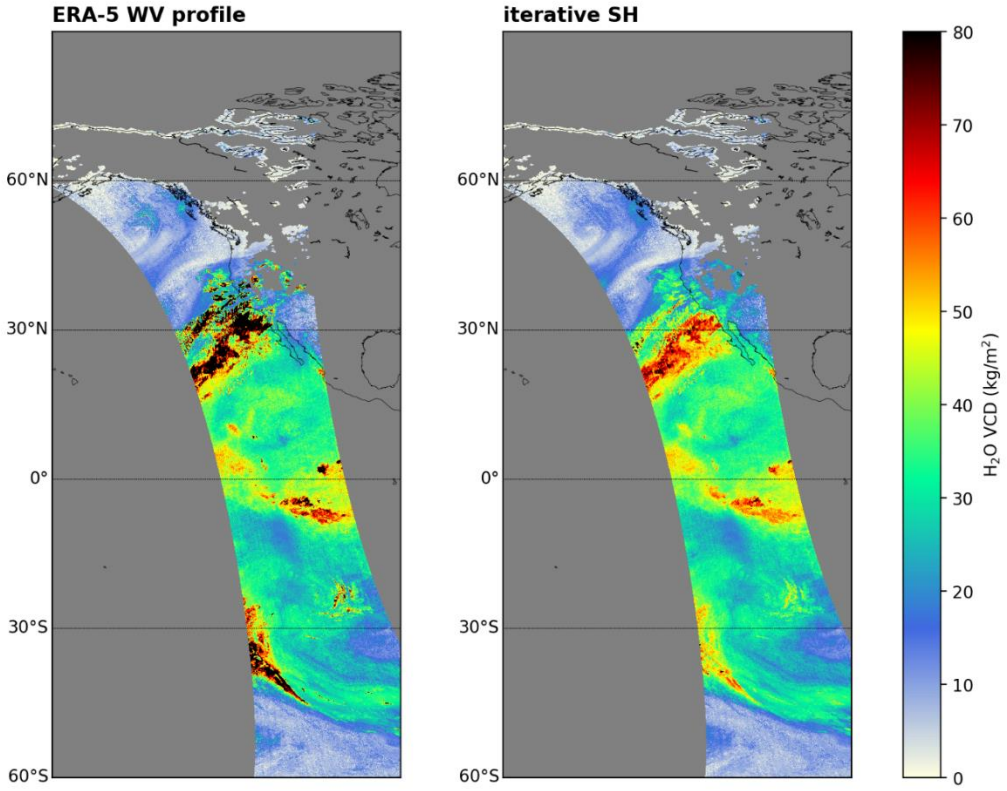


Figure 3 depicts typical examples of BAMF profiles for different clear- and cloudy-sky scenarios. The AMFs for the cloudy-sky scenarios were calculated assuming a surface albedo of 7% and an effective cloud fraction of 20%. For the clear-sky scenario (left panel) the sensitivity decreases towards the surface. For the cloudy-sky scenarios (right panel) the BAMF profiles slightly increase towards the (bright) cloud top surface of the respective scenario. Below the cloud, the sensitivity is 0, because the atmosphere is shielded. Since high clouds shield large fractions of the atmosphere and hence also of the water vapour column below the cloud (see black dashed curve), the AMF has to be corrected correspondingly and thus decreases for increasing cloud top heights.

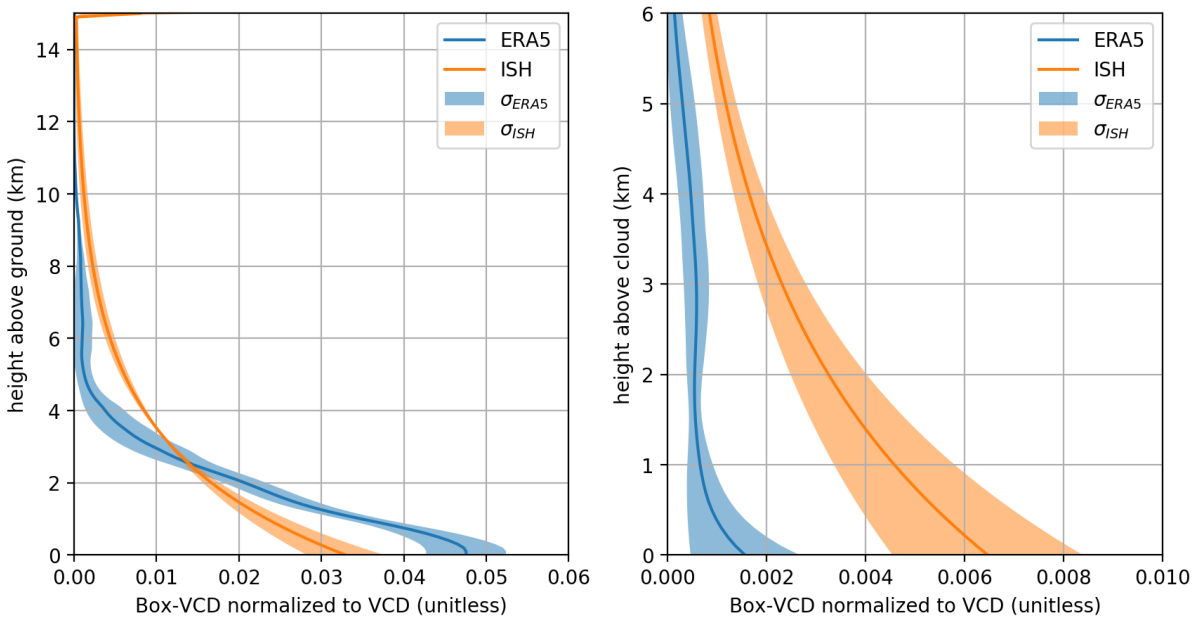
The simplest solution to use a “first-guess” water vapour profile from NWP re-analysis data is simply excluded with a reference to Wang. Then GPS climatologies are used to derive a scale-height parametrization. However, it is not made explicit or clear that the rational to use a scale-height (instead of the probably already quite realistic NWP full profile) is probably the simple need to regularize, and therefore constrain the retrieval problem. Since instruments measuring in the UV-visible range, will not be able to retrieve water vapour with significantly more than 1 to 2 independent pieces of information in the vertical. Adjusting VCD sub-columns by changing a single parameter, ie. scale height, therefor serves the need to constrain the problem. Since otherwise using re-analysis data as a-priori or first guess (depending on the inversion approach) and adjusting multiple layers in the retrieval would clearly have been the better approach. Such a parameterization using a scaled full profile, then however, also has the tendency (or advantage) to compensate for missing information (e.g. below the cloud), which is essentially missing in the measurements. To which extend this happens here is not very clear, and leads to the next issues concerning the treatment of cloud-coverage.

We agree with the reviewer that the simplest solution for an a priori water vapour profile is the usage of NWP profiles. Nevertheless, our goal was to be independent from any model data, as there are some issues to consider when using profiles from

NWP or reanalysis: Models could be affected by a systematic bias in their simulations and, for the case of reanalysis, are also affected by the varying numbers of observations. Furthermore, current global models typically have a spatial resolution that is significantly larger than a TROPOMI pixel (3.6x5.6km^2 vs 25x25km^2). In addition, the temporal resolution is also limited. As a result, the model is unable to adequately reproduce sub-scale/sub-grid processes (e.g. cloud cover/height). This also raises the question to what extent the modelled atmosphere is trustworthy, e.g. if the modelled WV profile does not coincide with observations.



To illustrate this conflict of potentially wrong WV profiles, we calculated the AMF using water vapor profiles from the reanalysis model ERA5 (hourly timestep, 0.25x0.25° grid, 60 vertical levels) for the same orbit as in Figure 13 and compared the resulting VCDs. It can be seen that, especially in cloudy areas, the VCDs from ERA5-AMFs clearly overestimate compared to the VCDs from iterative scale height AMF (e.g. in the area of the Atmospheric River by 30°N).



Further analyses of the water vapour profiles in these particular regions reveal that ERA5 underestimates the VCD above the cloud in comparison to the iterative scale-height (ISH) method. In addition, the variation of the above-cloud mean WV profiles of ERA5 is much smaller than those of the ISH method which could indicate that for these cases ERA5 tends more to its a priori information, i.e. climatological mean.

Nevertheless further investigations are beyond the scope of this paper and should be addressed in later studies.

We included the VCD from the ERA5 profiles as well as the figure of the profiles in the revised manuscript and added the following text to Section 3.4:

Taking a closer a look at the reasons for the deviations of results retrieved for the ERA-5 profiles, Fig. 13 depicts the mean of the normalized water vapour profiles of ERA-5 and the iterative scale height approach for the AR region (around 30°N). The left panel of Fig. 13 shows the water vapour profile from ground up to 15km. In comparison to the iterative approach, ERA-5 is much drier above approximately 2.5km for these particular cases, indicating that ERA-5 might systematically underestimate the water vapour content above the cloud within the region of the atmospheric river. This finding is further supported by the right panel of Fig.13 which illustrates the normalized water vapour profiles above the cloud top: ERA-5 profiles are close to 0 and show only small variations, whereas the profiles of the iterative approach indicate higher water vapour concentrations along with a much higher variability. One potential reason for the discrepancies of ERA-5 could be the missing of observational input data for the reanalysis: without observations, the reanalysis model is dominated by its a priori information (e.g. a climatological mean), so that it can be systematically distorted from the real atmosphere. However, further investigations of possible ERA-5 biases are beyond the scope of this paper.

2. Treatment of clouds

The treatment of partially (or fully) cloudy scenes (as observed and expressed in geometrical cloud coverage (cloud fraction CF) with the use of collocated imager data) is critical for TWVC retrievals, since clouds may shield or amplify (through scattering) the true total column value. Up to the validation section, (Section 2 to 5), the paper discusses the conversion of SCD to VCD for any level of cloud fractions, using the independent pixel approximation. The AMF error analysis however seems to be carried out for a CF of up to 50%. The impressive results shown in Figure 13, present the results for CF<20% and for all-sky (CF up to 100%), arguing for the usage of the presented scale-height method. These results seem to indicate that the method even works for CF>20%. In the introduction to the validation section (Section 6) it is then however stated that the validation is carried out for CF of up to 20% ("clear-sky") only and the paper validates the results with SSMI in dependence of CTH in Figure 21. Itis assumed that this evaluation is also for CF<20%. Otherwise, the reader would expect significant underestimations of the results for large CF with high CTH. So Figure A10 to 13 adds to the confusion since, the results presented there seem to indicated the opposite: high cloud fractions for high cloud top levels lead to overestimations with respect to SSMI.

In Figures 20, 21, 23, 24, and 26 we investigated the CTH dependence for CF<20%, which would correspond to the first two columns of Figures A10 to 13. To avoid confusion we added further explanations in the captions of Figures 20, 21, 23, 24, and 26. We added the following phrase to the Section's introduction:

For the sake of completeness, we also briefly investigate higher cloud fractions at the end of each subsection and provide the results in the Supplementary Material.

How is this result interpreted with respect to the used combination of independent pixel approach, WV profile scale-height parameterization method, and the evaluation of the AMF and its error?

In very special cases problems with using the IPA can occur due to TROPOMI's small pixel size. However, in general the same effects also occur in cloud retrieval and in cloud correction of the AMF. These two effects largely compensate each other.

What is eventually seen by the authors as the final product and at which CF? The CF threshold has been key to all previously published retrieval methods from the UV/visible to NIR spectrometers. Therefore one would expect a clear statement even up-front in the introduction and for sure in the conclusions, if the product, with its novel scale-height approach, wants to target a specific cloud-coverage threshold or is proposing one for final use.

In fact, the results imply a possible use of the data up to cloud fractions of 100%, but it must be taken into account that for high cloud fractions no information below the cloud can be gained. Also the input parameters (clouds & albedo) are currently still subject to relatively large uncertainties and are continuously improved. For example,

the OMI albedo is still used for the calculation of the cloud fraction, so that if there will be an update to an improved albedo data set derived from TROPOMI, significant changes in the cloud fraction can be expected. Such changes in CF also lead to changes of the cloud height. Furthermore various specific cloud height updates are also being implemented at the moment. Due to the continuous changes/improvements of the input parameters we refer the expert user to the results of the error estimation depicted in Figures 17 and 18 and summarized in Table 4 and 6. We also added a clear statement in the revised manuscript and recommend the non-expert user to only use VCDs with CF<20% and AMF>0.1, which represents a good compromise between coverage and retrieval accuracy. The following can be found now in the summary and abstract:

For the general purpose we recommend to only use VCDs with cloud fraction < 20% and AMF > 0.1, which represents a good compromise between spatial coverage and retrieval accuracy.

Minor comments:

l.57 p.2: Is the wavelength alignment carried out for all solar measurements provided by S5p, or only once for all retrievals.

The wavelength calibration is performed for each day with the latest available daily irradiance. We added this information to Section 2.

l.73, p3: Are ISRF changes in width found with the WV retrieval over the orbit for S5p?

We do not find significant ISRF changes along the orbit as reported by Beirle et al. (2017) for the GOME-2 instrument which is probably due to the TROPOMI's better cooling system compared to GOME-2. Nevertheless we see changes of the ISRF width over cloudy scenes which might indicate differences in the pixel illumination.

l. 84, p.3: the difference between ground and spectral pixel should be made clear to avoid confusion.

This point was made clearer in the updated version.

l.101, p.4: In this equation "beta" is not defined and it is not clear how I0 from the solar irradiance is used here.

The respective explanations were added. To avoid confusion with other sections, we changed the variable names of beta to mu and from VCD_i to c_k. The text is now:

These simulations yield a Jacobian vector J=dln I / dμ (with the absorption coefficient μ and the simulated intensity I at TOA normalised by the solar spectrum I_0) defined at each grid box k.

These BAMF profiles have to be combined with the partial vertical columns c_k of an a priori water vapour profile: …

Eq4, p.5: Is the refractivity equation relevant here? I guess the important point to be made is the use of COSMIC water vapour climatologies in contrast to the already smoothed model data. It is still very puzzling why the former should be better for this purpose than using actual reanalysis data. Since GPS (and hyper-spectral TIR) profile data is meanwhile an essential component in NWP data, and the model helps in reducing the vertical information content towards the one from S5p.

As Reviewer #1 also states that this equation only creates confusion, we removed it in the revision process.

p. 4: The VCD equation on page 4 is not numbered and the usage of I is confusing here, since I guess it refers to the iteration step instead of the previously used sub-column layer number. Otherwise some clearer explanation would be needed.

We numbered all equations and modified them to avoid for potential conflicts between the different sections of the paper.

Literature

Beirle, S., Lampel, J., Lerot, C., Sihler, H., and Wagner, T.: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives, Atmos. Meas. Tech., 10, 581–598, https://doi.org/10.5194/amt-10-581-2017, 2017.

Dear Eamon Conway,

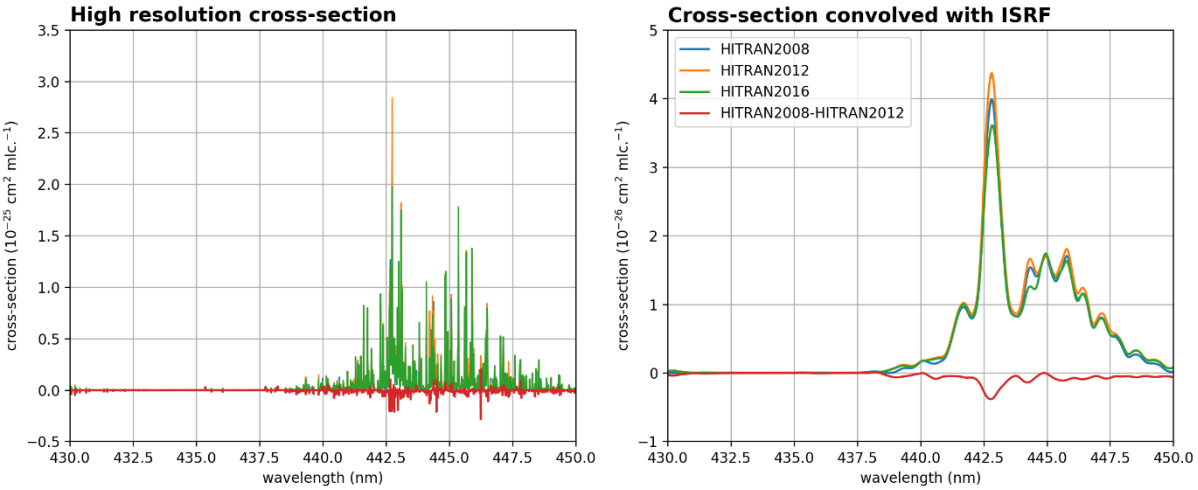
Please find below our replies to the raised issues, where
blue repeats your comments,
and black is used for our reply.

We have a couple of comments on the way different editions of the HITRAN database are represented in this paper.

1. Regarding statement in Line 75. "The molecular absorption by water vapour within our fit window is relatively weak and hence the modelled line lists vary strongly from HITRAN 2008 to HITRAN 2012 (Rothman et al., 2013) and to HITRAN 2016 (Gordon et al., 2017)."

Please see attached figure, which shows the absolute absorption cross-sections ofHITRAN2008 and HITRAN2012 plotted on a linear scale, all isotopologues included. The cross-sections were all calculated using the HITRAN API with a temperature of 288K, 1 atm, 0.001 cm^{-1} resolution and a 50 cm^{-1} wing. The data sets appear similar and do not seem to "vary strongly". The residuals at 446.2 nm, 444.2 nm and 444.5 nm are due to improvements in calculated line positions in HITRAN2012. The line centers are slightly shifted, but they are present in each set. The HITRAN2012 line list is also more complete than HITRAN2008, note the extra HITRAN2012 weak absorption around 440 nm. Can the authors provide any evidence that suggest the data sets "vary strongly" in 430-450 nm?

We agree that looking at the high-resolved cross-sections, the differences between the HITRAN versions are small. However, for atmospheric remote sensing applications the high-resolved cross-sections have to be convolved with the instrumental spectral response function (ISRF) of the instrument. The figure below depicts the high-resolved H_2O cross sections of HITRAN 2008, 2012, and 2016 (all at 296K) in the left panel and the corresponding ("low-resolution") cross-sections convolved with a typical ISRF of TROPOMI in the right panel. For the ISRF we assumed a symmetric Super-Gaussian with parameters taken from Beirle et al. (2017).



After the convolution one observes that the strengths of peaks of the cross-sections differ distinctively: HITRAN2012 is 7-9% higher and HITRAN2016 is 7-9% lower than

HITRAN2008. Ergo we find an alternating pattern between the cross-section versions, which is why we conclude that the versions “vary systematically” among themselves. In the revised manuscript, we added the figure to better clarify our view.

2. Regarding statement in Line 77. "Lampel et al. (2015) found out that HITRAN2012 underestimates the water vapour concentration derived from Long Path DOAS observations by approximately 10% and that the previous version HITRAN2008 agrees better to the reference measurements."

Within the discussion section in Lampel et al. (2015b), they found residuals in all their windows to be reduced by going from HITRAN2009 to HITEMP2010/HITRAN2012 cross-sections. (HITRAN2008 is HITRAN2009: the article was online in 2009 but the edition of the database is HITRAN2008.) Lampel et al. (2015b) also say "development of water vapour absorption compilations from HITRAN 2009 to HITEMP/HITRAN 2012 results in a better fit of the measurement data". This is not in line with what is claimed here? Can the authors please verify?

In our above mentioned statement we referred to the comparisons of LP-DOAS measurements to meteorological data made by Lampel et al. (2015). Their findings are summarized in Table 8 of their paper which gives regression results. The regressions yield a slope of 1.004 for HITRAN2008 and 0.918 for HITEMP/HITRAN2012, i.e. HITRAN2008 agrees better to the meteorological observations than HITRAN2012.

We did not discuss the fit quality and it is true that Lampel et al. (2015) found residuals decreasing by switching from HITRAN2008 to 2012. However, looking at the relative fit errors of the fit window around the strongest line (see Table 6 in their paper, W3) the error decreases from 1.04% to 0.98% for LP-DOAS.

Thus, in view of the observational evidence with almost negligible changes in the relative fit error, we see our statement confirmed.

3. Appendix B. The authors state that Lampel et al. (2015b) found the HITRAN2012 cross-sections to underestimate water vapor mixing ratios by 8% in 430-450 nm, while for HITRAN2008, the results are in excellent agreement with the meteorological station. The meteorological station is quoted by Lampel et al. (2015b) to have a 5% uncertainty on their value of humidity and a 2% error on the temperature. Lampel et al. (2015b) then proceeds to state "absolute differences of the cross-sections shown in Table 3 cannot be absolutely validated with sufficient precision". It therefore may not be true that HITRAN2008 is ’superior’ to HITRAN2012 and 8% deviation is very close to the uncertainty of the measurement. Does this provide sufficient evidence to support the use of HITRAN2008 over HITRAN2012?

In Appendix B in our paper we present further LP-DOAS measurements conducted by Johannes Lampel and Stefan Schmitt at the CESAR Tower in Cabauw (Netherlands) during the CINDI-2 campaign. The H_2O SCDs have been derived using HITRAN2012 cross-sections and the resulting water vapour volume mixing ratios have been compared to the meteorological measurements at different altitudes of the tower (see Figure A2 in our paper). From these further measurements we find underestimations

of 17% during day and 11% during night. These additional observations, combined with the findings of Lampel et al. (2015), provide sufficient evidence to use HITRAN2008 over HITRAN2012.

Literature

Beirle, S., Lampel, J., Lerot, C., Sihler, H., and Wagner, T.: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives, Atmos. Meas. Tech., 10, 581–598, https://doi.org/10.5194/amt-10-581-2017, 2017.

Lampel, J., Pöhler, D., Tschritter, J., Frieß, U., and Platt, U.: On the relative absorption strengths of water vapour in the blue wavelength range, Atmos. Meas. Tech., 8, 4329–4346, https://doi.org/10.5194/amt-8-4329-2015, 2015.

Total Column Water Vapour Retrieval from S-5P/TROPOMI in the Visible Blue Spectral Range

Christian Borger, Steffen Beirle, Steffen Dörner, Holger Sihler, and Thomas Wagner
Satellite Remote Sensing Group, Max Planck Institute for Chemistry, Mainz, Germany
Correspondence: C. Borger (christian.borger@mpic.de) and T. Wagner (thomas.wagner@mpic.de)

Abstract. Total column water vapour has been retrieved from TROPOMI measurements in the visible blue spectral range and compared to a variety of different reference data sets for clear-sky conditions during boreal summer and winter. The retrieval consists of the common two-step DOAS approach: first the spectral analysis is performed within a linearized scheme and then the retrieved slant column densities are converted to vertical columns using an iterative scheme for the water vapour a ~~apriori profile shape~~priori profile shape, which is based on an empirical parameterization of the water vapour scale height. Moreover, a modified albedo map was used combining the OMI LER albedo and scaled MODIS albedo map. The use of the alternative albedo is especially important over regions with very low albedo and high probability of clouds like the Amazon region.

The errors of the TCWV retrieval have been theoretically estimated considering the contribution of a variety of different uncertainty sources. For observations during clear-sky conditions, over ocean surface, and at low solar zenith angles the error typically is around values of 10-20% and during cloudy-sky conditions, over land surface, and at high solar zenith angles it reaches values around 20-50%.

In the framework of a validation study the retrieval demonstrates that it can well capture the global water vapour distribution: the retrieved H_2O VCDs show very good agreement to the reference data sets over ocean for boreal summer and winter whereby the modified albedo map substantially improves the retrieval's consistency to the reference data sets, in particular over tropical landmasses. However, over land the retrieval underestimates the VCD by about 10%, particularly during summertime. Our investigations show that this underestimation is likely caused by uncertainties within the surface albedo and the cloud input data: Low level clouds cause an underestimation, but for mid to high level clouds good agreement is found. In addition, our investigations indicate that these biases can probably be further reduced by the use of ~~updated~~ improved cloud input data. For the general purpose we recommend to only use VCDs with cloud fraction < 20% and AMF > 0.1, which represents a good compromise between spatial coverage and retrieval accuracy.

The TCWV retrieval can be easily applied to further satellite sensors (e.g. GOME-2 or OMI) for creating uniform measurement data sets on longterm, which is particularly interesting for climate and trend studies of water vapour.

1 Introduction

Water vapour is the most important natural greenhouse gas in the atmosphere and plays a key role in the atmospheric energy balance via radiative effects and latent heat transport (Held and Soden, 2000). Due to its high spatiotemporal variability on

all atmospheric scales, accurate knowledge of the amount and distribution of water vapour is essential for numerical weather prediction and climate monitoring.

Several in situ and remote sensing measurement techniques have been developed in the past decades, enabling to observe the water vapour distribution from platforms like radiosondes, balloons, aircrafts and satellites. The particular absorption properties of water vapour allow to retrieve the water vapour content via satellites for several different spectral ranges from the radio (Kursinski et al., 1997), microwave, e.g. AMSU (Rosenkranz, 2001), thermal infrared, e.g. AIRS (Susskind et al., 2003), near-infrarednear- and shortwave-infrared, e.g. MODIS (Gao and Kaufman, 2003)and MERIS (Bennartz and Fischer, 2001) , MERIS (Bennartz and Fischer, 2001), and TROPOMI (Schneider et al., 2020) to the visible, e.g. GOME (Noël et al., 1999; Wagner et al., 2003; Lang et al., 2007), SCIAMACHY (Noël et al., 2004), and GOME-2 (Grossi et al., 2015).

In the The visible spectral range is particularly interesting for the retrieval of total column water vapour (TCWV)has so far : 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.

So far TCWV has been retrieved mostly in the visible "red" spectral range because the absorption is strongest there. However, for this spectral range the ocean surface albedo is relatively low, leading to a low sensitivity for the lowermost troposphere, where the highest water vapour concentrations occur. In addition, current and past satellite sensors can not resolve the fine absorption structure of water vapour in this spectral range causing non-linear absorption effects (e.g. saturation) which have to be accounted for in post-processing. Thus, Wagner et al. (2013) suggested to apply retrievals in the "blue" spectral range (around 442 nm) where the absorption is much weaker than in the red making the retrieval problem quasi-linear. In addition, the ocean surface albedo is much higher leading to a higher sensitivity of the near-surface layers. First operational analyses of a similar approach have been performed by Wang et al. (2019) for measurements of the Ozone Monitoring Instrument (OMI, Levelt et al., 2006).

In October 2017 the TROPOspheric Monitoring Instrument (TROPOMI, Veefkind et al., 2012) onboard ESA's Sentinel-5 Precursor (S-5P) satellite was launched in a sun-synchronous polar orbit with an equator crossing time of 13:30 local time. TROPOMI is a UV-Vis-NIR push-broom spectrometer and consists of 450 detectors/rows covering a swath width of 2600 km. The outstanding property of TROPOMI is that its spectral bands in the visible combine a high signal to noise with an unprecedented spatial resolution of $3.5 \times 7.5\,\mathrm{km}^2$ (and $3.5 \times 5.6\,\mathrm{km}^2$ since August 2019; Rozemeijer and Kleipool, 2019) at nadir which allows to perform spectral analyses at a never seen before accuracy even on small spatial scale.

In this paper we introduce a TCWV retrieval based on the spectral analysis approach of Wagner et al. (2013) to S-5P/TROPOMI observations. The paper is organized as follows: In Sect. 2 we give an overview of the retrieval describing general retrieval principles and presenting the retrieval set-up. In Sect. 3 we present an empirical parameterization of the a priori water vapour profile shape and an iterative scheme making use of the relation between the water vapour profile shape and TCWV. In Sect. 4 we evaluate different input albedo products and in Sect. 5 we perform a detailed uncertainty analysis including a variety of different error sources. In Sect. 6 we present first TCWV results retrieved from TROPOMI measurements and perform a vali-

dation study using data sets from satellite, ground-based measurements, and reanalysis models as reference. In Sect. 7 we draw conclusions and summarize the outcomes of our investigations.

2 Retrieval principles

2.1 Wavelength calibration and spectral analysis

In a first step the wavelength alignment of the daily measured irradiance is calibrated for each of the 450 TROPOMI detectors/rows via a nonlinear least-squares fit in intensity space using the solar spectrum from Kurucz (1984) as reference. Simultaneously, the instrumental spectral response function (ISRF) is approximated assuming an asymmetric Super-Gaussian following the definition of Beirle et al. (2017):

$$S_{asym}(x) = \begin{cases} \exp\left(-\left|\frac{x}{w-a_w}\right|^k\right) & \text{for x} \leq 0 \\ \exp\left(-\left|\frac{x}{w+a_w}\right|^k\right) & \text{for x} > 0 \end{cases} \quad (1)$$

Next, we perform a spectral analysis using the differential optical absorption spectroscopy (DOAS; Platt and Stutz, 2008) scheme in which the attenuation along the light path is calculated via Beer-Lambert's law in optical depth space:

$$\ln\left(\frac{I}{I_0}\right) = \tau \approx -\sum_i \sigma_i(\lambda) \cdot \mathrm{SCD}_i + \Phi \quad (2)$$

where i denotes the index of a trace gas of interest, $\sigma_i(\lambda)$ its respective absorption cross section, $\mathrm{SCD_i} = \int_s c_i ds$ its concentration integrated along the light path s (the so called slant column density), and Φ a closure polynomial accounting for Mie and Rayleigh scattering as well as low-frequency contributions.

Table 1 summarizes the fit setup of the retrieval's spectral analysis. The retrieval's fit window ranges from 430 nm to 450 nm and accounts for molecular absorption by water vapour (HITRAN 2008, Rothman et al., 2009), NO_2 at 220K (Vandaele et al., 1998), ozone (Serdyuchenko et al., 2014) and the O_2-O_2 dimer (Thalman and Volkamer, 2013). In order to account for the Ring effect we include two Ring spectra (Wagner et al., 2009) and for Φ we use a 5th order polynomial. Furthermore, we include pseudo-absorbers accounting for intensity offset, for shift and stretch effects (Beirle et al., 2013) and, for ISRF changes along the orbit (Beirle et al., 2017) for the ISRF parameters w and k in Eq. (1). All molecular absorption cross sections are convolved with the ISRF of the corresponding TROPOMI row/detector determined during the calibration process.

The molecular absorption by water vapour within our fit window is relatively weak and hence the modelled line lists vary ~~strongly~~ systematically from HITRAN 2008 to HITRAN 2012 (Rothman et al., 2013) and to HITRAN 2016 (Gordon et al., 2017). Thus, the choice of line list is afflicted by a high degree of uncertainty. Lampel et al. (2015) found out that HITRAN 2012 underestimates the water vapour concentration derived from Long Path DOAS observations by approximately 10% and that the previous version HITRAN 2008 agrees better to the reference measurements. Further Long Path DOAS measurements taken during the CINDI-2 campaign also confirm the findings from Lampel et al. (2015) (see Appendix B for more details). ~~Thus~~ Hence, combining the findings from Lampel et al. (2015) and Wang et al. (2019) we conclude that HITRAN 2008 fits

best our needs and is superior to the most recent version of the HITRAN line lists (HITRAN 2016).

Due to the high daily data volume of the TROPOMI L1B radiances ~~(about 40 gigabyte per day for Band 4),~~ 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 400-499 nm, generates about 40 gigabyte per day. Therefore, we implemented a weighted linear least squares fit for our retrieval, in which the weights are the fractional coverage of the spectral pixel within the fit window (details in Appendix A). This weighting of the outermost pixels of the fit window avoids "jumps" of pixels included in the DOAS fit, as it would occur for a fixed fit window due to the changing pixel-to-wavelength mapping across track. Thus, across track "stripes" in the SCDs are avoided. 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).

Figure 1 illustrates a typical example of such a spectral analysis of a TROPOMI measurement spectrum in which the absorption structures of water vapour, NO_2, and the Ring effect can be well identified and the residual spectrum showing a mainly noisy structure. Figure 2 depicts the distribution of the H_2O SCD from one TROPOMI orbit (orbit number 6930) on 13th February 2019. It demonstrates that the TROPOMI retrieval is able to capture the meso- to macro-scale water vapour patterns like convective updrafts in the tropics and atmospheric rivers in the midlatitudes, whereby the small H_2O SCD values in the tropics are caused by cloud shielding.

2.2 VCD conversion and Box-AMF simulations

To convert the slant column density to a vertical column density (VCD), we apply the so called airmass factor (AMF):

$$\mathrm{VCD} = \frac{\mathrm{SCD}}{\mathrm{AMF}}$$

$$\mathrm{VCD} = \frac{\mathrm{SCD}}{\mathrm{AMF}} \tag{3}$$

The airmass factor accounts for the non-trivial effects of the atmospheric radiative transfer and is usually based on radiative transfer model (RTM) simulations. In our case we used the 3D Monte Carlo RTM McArtim (Deutschmann et al., 2011) and performed simulations at a wavelength of 442 nm for different retrieval scenarios (summarized in Tab. 2) assuming an aerosol-free atmosphere. These simulations yield a Jacobian vector ~~$J = \frac{\partial \ln I}{\partial \beta}$~~ $J = \frac{\partial \ln I}{\partial \mu}$ (with the absorption coefficient μ and the simulated intensity I at TOA normalised by the solar spectrum I_0) defined at each grid box ~~i and the~~ k. The altitude-dependent AMFs (BAMF) can then be calculated according to the formula:

$$\mathrm{BAMF}_i = -\frac{J_i}{I \Delta h}$$

~~with the simulated intensity I normalised by the solar spectrum and the~~

$$\mathrm{BAMF}_k = -\frac{J_k}{I \Delta h} \tag{4}$$

with the box thickness Δh. These BAMF profiles have to be combined with the partial vertical columns ~~VCD_i~~ c_k of an a priori water vapour profile:

$$\mathrm{AMF} = \frac{\sum_i \mathrm{BAMF}_i \cdot \mathrm{VCD}_i}{\sum_i \mathrm{VCD}_i} \frac{\sum_k \mathrm{BAMF}_k \cdot c_k}{\sum_k c_k} \tag{5}$$

with $\sum_k c_k = \mathrm{VCD}$. For the case of a cloud contaminated pixel we assume that the cloud is a Lambertian reflector with an albedo of 80% and use the cloud top height as surface altitude input for the AMF. Under the assumption of the independent pixel approximation the resulting cloud-affected AMF can then be calculated as a linear combination of the AMF for a clear-sky scenario and the AMF for a cloudy-sky scenario weighted by the respective simulated intensities I and the effective cloud fractions ζ as follows:

~~$$\mathrm{AMF} = \frac{(1-\zeta)\, I_{clear} \mathrm{AMF}_{clear} + \zeta \cdot I_{cloud} \mathrm{AMF}_{cloud}}{(1-\zeta)\, I_{clear} + \zeta \cdot I_{cloud}}$$~~

$$\mathrm{AMF} = \frac{(1-\zeta)\, I_{clear} \mathrm{AMF}_{clear} + \zeta \cdot I_{cloud} \mathrm{AMF}_{cloud}}{(1-\zeta)\, I_{clear} + \zeta \cdot I_{cloud}} \tag{6}$$

Figure 3 depicts typical examples of BAMF profiles for different clear- and cloudy-sky scenarios. The AMFs for the cloudy-sky scenarios were calculated assuming a surface albedo of 7% and an effective cloud fraction of 20%. For the clear-sky scenario (left panel) the sensitivity decreases towards the surface. For the cloudy-sky scenarios (right panel) the BAMF profiles slightly increase towards the (bright) cloud top surface of the respective scenario. Below the cloud, the sensitivity is 0, because the atmosphere is shielded. Since high clouds shield large fractions of the atmosphere and hence also of the water vapour column below the cloud (see black dashed curve), the AMF has to be corrected correspondingly and thus decreases for increasing cloud top heights.

3 A priori water vapour profile shape

As described in Sect. 2.2 and Eq. (5), knowledge of the a priori water vapour profile shape is necessary for accurate calculations of the AMF from BAMF profile. However, simply assuming the same a priori profile shape for the whole globe might cause biases because it can not account for the atmospheric variability of water vapour, such as latitudinal variation, seasonal cycles, or different profile shapes over maritime and continental regions due to different water vapour ~~source~~ sources (e.g. evapotranspiration by plants). Also, simply using profiles from numerical weather models is not uncritical: for instance Wang et al. (2019) found out that their calculated AMF change strongly depends on which reanalysis model data they were using. Weaver and Ramanathan (1995) approximated the water vapour profile by an exponential decay with altitude~~and a corresponding scale height defined as~~

~~$$H_{vapor} = \frac{R_v \langle T \rangle^2}{L \langle \Gamma \rangle}$$~~

:

$$n_v(z) = n_0 e^{-z/H_v} \tag{7}$$

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

$$H_v = \frac{R_v \langle T \rangle^2}{L \langle \Gamma \rangle} \tag{8}$$

where $\langle T \rangle$ denotes the mean air temperature within an atmospheric column, $\langle \Gamma \rangle$ the mean lapse rate within the same atmospheric column, R_v the gas constant of water vapour, and L the specific latent heat. However, this definition requires knowledge of the mean air temperature and/or the lapse rate and that the relative humidity is constant with altitude. The former can be only estimated using numerical weather models and the latter is very unlikely to occur in the atmosphere.

Thus, we investigate to find an empirical parameterization of the scale height and thereby focus on its dependency on the H_2O VCD and the aforementioned atmospheric variabilities, i.e. dependencies of latitude, seasonal cycle and surface properties (such as vegetation effects).

~~For these~~ 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.

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 ~~ROMSAF~~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~~N.~~. Since the atmospheric refractivity ~~N is given by (Smith and Weintraub, 1953):~~

$$N = 77.6\frac{p}{T} + 3.73 \times 10^5 \frac{E}{T^2}$$

~~with~~ is dependent on the air pressure~~p~~, the air temperature~~T~~, and the water vapor pressure ~~E~~(Smith and Weintraub, 1953), GPS RO allows ~~to retrieve~~ for the retrieval of profile information under all-weather conditions with a high vertical resolution of approximately 100 m in the lower troposphere up to 1 km in the stratosphere (Anthes, 2011) and an accuracy of around 1 g/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.

3.2 Calculation of scale height

For the calculation of the scale height we highsample the COSMIC profile to a 100 m grid up to 14 km or rather only consider profile data below 150 hPa (close to the tropopause height). Then we sum up all the partial columns of the COSMIC profile data from ground up to a (scale-) height H_{sum} where ~~63% of~~ the H_2O VCD ~~are reached:~~

$$\frac{\int_0^{H_{sum}} n(z)dz}{\int_0^{\mathrm{TOA}} n(z)dz} > 1 - \frac{1}{e} \approx 63\%$$

~~In order to~~ reaches $1 - \frac{1}{e}$:

$$\frac{\int_0^{H_{sum}} n(z)dz}{\int_0^{\mathrm{TOA}} n(z)dz} > 1 - \frac{1}{e} \approx 63\% \tag{9}$$

To evaluate this scale height approach, we performed a synthetic study in which we compared AMFs calculated for the original COSMIC water vapour profile measurements with AMFs for an exponential profile ~~with~~ using the corresponding calculated scale height H_{sum}. For the simulation of the BAMF profiles we assume an albedo of 7%, which is representative value for the ocean surface albedo (Tilstra et al., 2017). The solar zenith angle is calculated for the location of the COSMIC profile assuming an hour angle of 90° and the line of sight angles are prescribed for −90°, −70° and −50°.

The results of the intercomparison are given in Figure 4. The 2D histograms reveal that the AMFs derived with the exponential profile agree well with the AMFs calculated directly from the COSMIC profiles indicating that the chosen method can well reproduce the shapes of the COSMIC profiles. This good agreement can be also observed in the histograms of Fig. 5 which illustrate distributions of relative deviation between the AMFs for for selected latitude bins. These distributions have a sharpe shape and peak around values of 0% indicating that the AMFs from the exponential shape are almost unbiased to the reference AMFs. In addition, Fig. S1 shows examplary profiles for cases of good and bad agreement to the reference AMFs for the same selected latitude bins as in Figure 5. In general~~the~~ , bad agreement (left column) occurs for profile shapes in which a ~~distinctively strong~~ 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.

The results of the intercomparison for prescribed cloudy-sky conditions and nadir viewing geometry are illustrated in Fig. S2 in which the panels show histograms of the relative AMF deviation for the same selected latitude bins as in Fig. 5, but for different cloud fraction (10%, 20% and 50%; left to right column) and cloud top height (1 km, 2 km and 5 km; top to bottom row) scenarios. For a cloud top height of 1 km the AMFs calculated from the exponential profiles are generally biased negative for all cloud fractions, in particular for the latitude bin of -30° to -20°N. However, for higher clouds the AMFs agree well with the reference AMFs for almost all cloud scenarios except the extreme case with a cloud fraction of 50% and a cloud top height of 5 km or more.

Alternative methods for calculating the scale height yielded systematic overestimations of the AMF for clear-sky conditions

(Fig. A4) and higher scatter within the AMF for cloudy-sky conditions (Fig. S8) in comparison to the sum method, as shown in detail in Appendix C.

3.3 Parameterization of scale height

Figure 6 depicts the distribution of the calculated COSMIC scale height H_{sum} against the COSMIC TCWV for boreal summer over ocean for latitude bins of 10°. The regression fits (solid red lines) are based on orthogonal distance regression (ODR) using the "scipy.odr" package built on ODRPACK (Boggs et al., 1992). For low latitudes (tropics and subtropics) the scale height shows a high linear correlation to the H_2O VCD with slopes around 0.04 and Pearson correlation coefficients R of 70% and above. In contrast, for high latitudes the slope increases up to 0.1 and also the scatter increases distinctively, i.e. the correlation coefficient only reaches values of around 0.3 in the polar regions. This decrease in linear agreement is likely caused by the higher atmospheric variability due to higher atmospheric dynamics in the midlatitudes. Also, the uncertainty is higher in COSMIC profile because a drier atmosphere leads to a smaller sensitivity of the COSMIC profile retrieval to water vapour concentrations ~~(compare Eq. (??) and~~ (compare Kursinski et al. 1997).

Figure 7 illustrates the same panels as Fig. 6 but for data over land. In general, the scatter for all latutude bins has increased distinctively, resulting in an inferior linear agreement between the H_2O VCD and the scale height compared to the data over ocean, especially for deserts and northern polar regions. Fortunately, the surface albedo of these regions is usually high and thus the AMF is less dependent on the a priori profile shape. In addition, these regions are governed by an arid climate and thus the retrieved H_2O VCDs are expected to be small. Correspondingly, the absolute H_2O VCD errors due to uncertainties in the AMF are still relatively small.

In the following we investigate a parameterization of the scale height with respect to H_2O VCD, latitude, season for ocean and land separately. To distinguish between ocean and land surface, we use a land-sea mask derived from GSHHS coastline data (Wessel and Smith, 1996).

3.3.1 Ocean

The regression line parameters of the ODR fit results between COSMIC TCWV and COSMIC scale height for each latitude bin for each month for data over ocean are illustrated in Figure 8. The values for the fitted slopes (left panel in Fig. 8) indicate a quadratic dependency with latitude and reveal a seasonal shift towards higher latitudes during July, August, and September. Also the values for the fitted intercept vary with latitude and season.

Thus, the scale height over ocean H_{ocean} can be approximated as follows:

$$H_{ocean}(\mathrm{VCD},\theta,t) = \alpha(\theta,t)\cdot \mathrm{VCD} + \beta(\theta,t) \tag{10}$$

with

$$\alpha(\theta,t) = a_0(t) + a_1(t)\cdot\theta + a_2(t)\cdot\theta^2$$
$$\beta(\theta,t) = b_0(t) + b_1(t)\cdot\theta + b_2(t)\cdot|\theta - \theta_0(t)| \tag{11}$$

with the latitude θ and the day of year t. The annual variation of the function parameters a_i, b_i and θ_0 from Eq. (11) fitted for the monthly data sets (illustrated in Fig. 8) is depicted in Figure S9. Most function parameters reveal an annual and semi-annual cycle over the year. Hence, these function parameters can be approximated by a superposition of two simple cosine functions with prescribed frequencies:

$$a_i(t) = a_{i1} \cdot \cos(a_{i2} + \omega t) + a_{i3} \cdot \cos(a_{i4} + 2\omega t) + a_{i5}$$

$$a_i(t) = a_{i1} \cdot \cos(a_{i2} + \omega t) + a_{i3} \cdot \cos(a_{i4} + 2\omega t) + a_{i5} \quad (12)$$

with t as the day of year and $\omega = \frac{2\pi}{365}$. Such functions have also been fitted and illustrated for the monthly data in Fig. S9 (solid orange lines) whereby we assumed that the day of year representing the month is the first day of the month. For most function parameters the fits coincide well with the data points and in the cases of suboptimal fit results the annual variation of the data is relatively small, indicating that our choice of parameterization is valid.

Altogether, we have to fit 35 parameters to the complete data set of calculated COSMIC scale heights for the parameterization of the scale height over ocean. The goodness of the parameterization in approximating the scale height is illustrated in Fig.9 for different latitude zones. For the latitude zones including the tropics (-15° to 15°N) and subtropics (-35° to 35°N) we find a good agreement between the parameterization and the calculated COSMIC scale height with R^2 of 0.72 and 0.60 respectively. However, including higher latitudes in the evaluation, i.e. midlatitudes (-60°to 60°N) and polar regions (-90°to 90°N), leads to an increased scatter and a worsening of the parameterization (R^2 of 0.45 and 0.44 respectively). This inferior agreement is likely caused by the larger atmospheric variability in the midlatitudes (e.g. higher atmospheric dynamics) as well as an increased uncertainty in the COSMIC water vapour profile measurements due to lower water vapour concentrations.

3.3.2 Land

Figure 7 already revealed much larger scatter in the distribution of COSMIC TCWV and COSMIC scale height for data over land, indicating that the water vapour profile shape over land surface is less homogeneous than over ocean likely due to further heterogenously distributed water vapour sources, such as evapotranspiration by plants and soil. Thus, the H_2O VCD and scale height are likely to be dependent on the amount of vegetation, i.e. high vegetation is associated with high evapotranspiration and high water vapour concentrations near the ground and thus the scale height should be close to the scale height over ocean. In contrast~~to that~~, low amounts of vegetation are associated with less evapotranspiration and a usually drier atmosphere indicating that the scale height should be higher than over ocean.

For quantifying the amount of vegetation we use the Normalized Difference Vegetation Index (NDVI) where a value of 1.0 indicates high vegetation and a value around 0.0 indicates low vegetation. As data source for the NDVI we use data within the MODIS Aqua MYDC13C2 Version 6 product (Didan et al., 2015) and continue as follows: First, we calculate the parameterized scale height H_{ocean} assuming an ocean surface globally. Then we calculate the ratio of the calculated COSMIC scale height over land H_{land} and the parameterised scale height H_{ocean}.

Figure 10 shows the ratio H_{land}/H_{ocean} as a function of the NDVI for data sets filtered by different landcover types and the solid red lines represent the robust regression results (summarized in Tab. 3) using the model from Siegel (1982). The left panel depicts the distribution for which no filter is applied. Except for low NDVI values a linear relation between ratio and NDVI is observable, however for NDVI values around 0.1 the ratio varies strongly between 0.7 and 3.0. In the center panel we use the landcover classification from the MODIS Aqua MCD12C1 Version 6 product (Sulla-Menashe et al., 2019) to filter measurements for locations classified as landcover type 15 (corresponding to a desert). With this filter the ratio now only varies between 0.7 and 1.5 with a weak dependence on the NDVI. If we further filter locations of landcover type 7 (corresponding to open shrublands), the fit results of the robust regression change only slightly compared to the first filtered data set.

Hence the scale height over land H_{land} can be approximated as the scale height over land H_{ocean} multiplied by a first order polynomial of the NDVI:

$$H_{land} = H_{ocean}(\mathrm{VCD}, \theta, t) \cdot (\gamma_{land} + \delta_{land} \cdot \mathrm{NDVI}) \tag{13}$$

whereby in the following we use the results for the data set filtered for landcover type 7 and 15 globally. Since regions of landcover type 7 or 15 are usually arid, the retrieved H_2O VCD is small and thus the error due to an inadequate parameterization of the AMF is much smaller than the fit error of the spectral analysis.

3.4 Iterative retrieval scheme

For the calculation of the H_2O VCD we precomputed AMF look-up tables (LUT) for the different water vapour profile shapes with scale heights ranging from 0.5 km to 5.0 km. These LUTs can then be used within a fixed-point iteration. In our case the iterative retrieval scheme is based on a fixed-point iteration according to Steffensen's method (Steffensen, 1933; Wendland and Steinbach, 2015):

$$\mathrm{VCD}_{i+1} = \mathrm{VCD}_i - \frac{\left(f(\mathrm{VCD}_i) - \mathrm{VCD}_i\right)^2}{f(f(\mathrm{VCD}_i)) - 2 \cdot f(\mathrm{VCD}_i) + \mathrm{VCD}_i}$$

$$\mathrm{VCD}_{i+1} = \mathrm{VCD}_i - \frac{\left(f(\mathrm{VCD}_i) - \mathrm{VCD}_i\right)^2}{f(f(\mathrm{VCD}_i)) - 2 \cdot f(\mathrm{VCD}_i) + \mathrm{VCD}_i} \tag{14}$$

where f is a function calculating the scale height for a given VCD using Eq. (10) and (13), applying it to the precomputed AMF look-up tables and from that returning a new VCD. The advantage of Steffensen's method is that it does not need a derivative and is able to determine the fixed-point even for the case of a non-contractive function (Wendland and Steinbach, 2015). For the first guess we derived the initial VCD from the SCD using a geometric AMF ($\mathrm{AMF}_{\mathrm{geo}} = \frac{1}{\cos(\mathrm{SZA})} + \frac{1}{\cos(\mathrm{VZA})}$) and stop the iteration as soon as the logarithmic difference between two consecutive results is smaller than 5% (approximately $1\,\mathrm{kg\,m^{-2}}$ assuming an average H_2O VCD of $20\,\mathrm{kg\,m^{-2}}$) or after six iteration steps. We also checked other values for the first guess and could confirm that the convergence of the iterative scheme is independent from them.

Figure 11 illustrates a comparison of H_2O VCD distributions for the cases of using a global constant a priori water vapour

profile shape (left panel) with a scale height of 2 km (in accordance to Weaver and Ramanathan (1995)) and using the iterative scale height approach (center panel) for all-sky conditions (i.e. no cloud filter applied) during an atmospheric river event at the Western coast of the US on 13th February 2019. The right panel of Fig. 11 depicts the distribution of the water vapour scale height yielded during the iterative VCD conversion. The water vapour scale height varies a lot along the orbit and differs distinctively from 2 km, causing large deviations between the two approaches, particularly at pixels with high TCWV values and for clouded pixels. Yet, in contrast to the approach with a constant scale height the iterative approach is still able to give reasonable TCWV results and does not exceed values higher than 80 kg m^{-2}.

~~Similiar results can be found in Fig. 12 which~~ Figure 12 illustrates the H_2O VCD distributions ~~of both approaches~~ from calculations using constant, ERA-5 and iterative profile shapes for the same scenario for clear-sky (effective cloud fraction CF<20%, top row) and all-sky (CF≤100%, bottom row) conditions. ~~In addition~~ For ERA-5 we used the data provided by Copernicus Climate Change Service (2017) on an hourly 0.25° × 0.25° grid and interpolated the model profile data to the TROPOMI pixel center coordinates. In addition to the TROPOMI H_2O VCDs, Fig. 12 also depicts the TCWV distribution from microwave sensor SSMIS f16 which has a temporal difference of around +2.3 hours.

For the clear-sky case (top row) the VCD distributions between ~~both~~ all profile approaches are almost identical whereby for the constant scale height approach (~~left panel~~ first panel from the left) very high VCDs (exceeding values higher than 80 kg m^{-2}) can be observed at the edges of the cloudy regions in the Northern subtropics. For the all-sky case (bottom row) the differences between ~~the two~~ all approaches are largest in cloudy regions~~, but as already mentioned before~~ : for instance in the region of the atmospheric river, the VCDs from the constant and ERA-5 profiles distinctively overestimate the VCD and exceed values higher than 80 kg m^{-2}. In contrast, even under these unfavourable observation conditions the iterative approach is still able to give reasonable VCD values~~, whereas the constant~~ . Furthermore, the iterative approach shows an overall good agreement to the SSMIS observations.

Taking a closer a look at the reasons for the deviations of the results retrieved for the ERA-5 profiles, Fig. 13 depicts the mean of the normalized water vapour profiles of ERA-5 and the iterative scale height approach ~~distinctively overestimates the VCD. In addition~~ for the AR region (around 30°N). The left panel of Fig. 13 shows the water vapour profile from ground up to 15 km . In comparison to the iterative approach~~shows an overall good agreement to the SSMIS observations~~, ERA-5 is much drier above approximately 2.5 km for these particular cases, indicating that ERA-5 might systematically underestimate the water vapour content above the cloud within the region of the atmospheric river. This finding is further supported by the right panel of Fig. 13 which illustrates the normalized water vapour profiles above the cloud top: ERA-5 profiles are close to 0 and show only small variations, whereas the profiles of the iterative approach indicate higher water vapour concentrations along with a much higher variability. One potential reason for the discrepancies of ERA-5 could be the missing of observational input data for the reanalysis: without observations, the reanalysis model is dominated by its a priori information (e.g. a climatological mean), so that it can be systematically distorted from the real atmosphere. However, further investigations of possible ERA-5 biases are beyond the scope of this paper.

4 Evaluation of different surface albedo input data

The surface albedo has a strong impact on the radiative transfer and thus also on the AMF. Hence we investigated the impact of different albedo products on the TCWV retrieval: the OMI monthly a) mean and b) minimum Lambertian equivalent reflectance (LER) at 442 nm from Kleipool et al. (2008) and c) MODIS Aqua blue surface reflectance from the MODIS MYD13C2 Version 6 product (Didan et al., 2015). The MODIS reflectance covers a broad spectral window from 459 nm to 479 nm. Thus to account for the different spectral windows of the albedo products we scale the MODIS albedo by factor of 0.9. This factor was estimated by calculating the ratio between 472 nm/442 nm of the OMI yearly minimum LER over parts of Australia where cloud contamination is generally low and hence the OMI LER has reasonably accurate values.

Figure 14 illustrates the global mean H_2O VCD of boreal summer 2018 for the different albedo input data over land (top row: monthly mean OMI LER, middle row: monthly minimum OMI LER, bottom row: scaled monthly MODIS Aqua blue surface reflectance). In the tropical and subtropical regions the OMI albedos cause a distinctive separation of the VCDs between land and ocean, in particular at the coasts of South America, Africa and Indonesia. These aforementioned regions are often affected by cloud cover, which might cause that the OMI albedo statistics can not filter cloudy cases correctly, so that cloud-contaminated observations are used within the albedo calculations. As a consequence, the values in the OMI albedo are too high and lead to an overestimation of the AMF which in turn causes an underestimation of the H_2O VCD.

In contrast~~to that~~, MODIS pixels have a much higher spatial resolution and MODIS' NIR channels are more sensitive to cloud contimination, yielding a higher sample size and allowing for correct cloud filtering. Hence, the H_2O VCD distribution using the MODIS surface reflectance results in a much smoother transition from ocean to land and in general much higher VCD values over land along the equator. Thus, in the following we use a combination of the MODIS and OMI albedos: the scaled MODIS Aqua blue surface reflectance over land and the monthly minimum OMI albedo over ocean.

5 Uncertainty estimation

The error budget of the H_2O VCD is determined by the propagation of the main error sources of the fitted SCD and the precalculated AMF. Errors in the SCD are mainly caused by random errors like the photon noise, and systematic errors, e.g. the uncertainty of the absorption cross section whereas errors in the AMF are mostly systematic with random contributions.

5.1 Uncertainties in the slant column density

Table 4 summarizes the different error sources for the H_2O SCD and the corresponding estimated uncertainties. As demonstrated in Sect. 2.1 and Appendix B the water vapour absorption cross section varies ~~distinctively~~ systematically between the different HITRAN versions. Hence, we assumed that the uncertainty of the water vapour cross section is of the same order of magnitude as the changes between the cross section versions, i.e. approximately around 10%. Considering the LP-DOAS comparisons (see Sect. 2.1 and Appendix B) we estimate these errors to be around 5% for this study.

The retrieval's spectral analysis directly yields the 1σ standard fit error of the H_2O SCD which is usually dominated by noise.

For a better understanding of these fit errors we separated them into data for small/large solar zenith angles (SZA<20° and 70°<SZA≤90°, respectively), low/high surface albedo (<3% and >15%, respectively), and clear-/cloudy-sky observation conditions (CF<5% and CF>20%, respectively). The distributions of the standard and relative fit errors of the spectral analysis are given in Fig. 15 and 16, respectively. The median values in Fig. 15 indicate that the standard errors for high SZA (around 0.3×10^{23} molec cm^{-2}) are twice as high as for small SZA (around 0.15×10^{23} molec cm^{-2}). Under clear-sky conditions the standard error for small surface albedo values is larger than for high surface albedo, but for cloudy conditions it does not depend on the surface albedo.

Figure 16 reveals that the relative fit errors for high SZA are higher than for low SZAs. However, the locations of maximal probability density and the medians also indicate that the distributions are positive skewed in particular for high SZA scenarios: for these scenarios the relative errors easily exceed values of 100%. Nevertheless, using the locations of maximal probability density as a rule-of-thumb estimate, relative fit errors have values around 10% for low SZAs and approximately 30% for high SZAs.

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.

5.2 Uncertainties in the AMF

The uncertainty in the AMF depends on the uncertainty of its input parameters. Because the parameters of the viewing geometry (i.e. solar zenith angle, line of sight angle, and solar relative azimuth angle) are known with high accuracy, the most important uncertainties are uncertainties of the surface albedo, cloud fraction, cloud top height, and water vapour profile shape. In order to estimate the contribution of each input parameter to the overall AMF uncertainty we define standard scenarios (summarized in Tab. 5) for which we calculate the AMF from the precalculated LUT ~~and then~~. Then we vary the input parameter for each scenario according to its uncertainty assumption listed in Table 6. The uncertainties of the water vapour scale height have been derived from the fit results of the intercomparisons between the measured COSMIC scale height and the parameterized scale height over ocean (see Fig. 9) and land (see Fig. 10).

Figure 17 depicts box-whisker plots of the relative AMF error due to uncertainties in surface albedo and scale height for the standard clear-sky scenarios of surface albedo, solar zenith angle, and scale height. It reveals that uncertainties in surface albedo and scale height over low vegetation have strongest impact on the AMF and can cause AMF errors larger than 30%, in particular for scenarios with low surface albedo or high solar zenith angle. On average the median values of the AMF errors typically vary around approximately 10%.

Figure 18 illustrates box-whisker plots of the relative AMF error due to uncertainties in surface albedo, scale height, cloud fraction. and cloud top height for all standard scenarios listed in Table 5. In contrast to the clear-sky scenarios the impact of the surface albedo uncertainties has strongly decreased, but in general the contributions of all AMF errors have increased distinctively. The main source for the AMF errors is still the uncertainty of the scale height over low vegetation whose median

values varies between 20-50%, but can also cause AMF errors larger than 60%.

Table 6 summarizes the results of the different error sources considered in the AMF uncertainty for clear- and cloudy-sky conditions. For clear-sky conditions one can typically assume a relative AMF error around 10-15% and for cloudy-sky conditions around 10-25%.

5.3 Total H_2O VCD uncertainty

The total relative H_2O VCD uncertainty can be approximated by

$$\frac{\Delta \mathrm{VCD}}{\mathrm{VCD}} = \sqrt{\left(\frac{\Delta \mathrm{AMF}}{\mathrm{AMF}}\right)^2 + \left(\frac{\Delta \mathrm{SCD}}{\mathrm{SCD}}\right)^2} \tag{15}$$

With our findings of typical relative AMF and H_2O SCD uncertainties the total relative VCD uncertainty is typically around 10-20% for observations during clear-sky conditions, over ocean surface, and at low solar zenith angles. During partly clouded-sky conditions, over land surface, and at high solar zenith angles the error reaches values of approximately 20-50%.

6 Validation study

In order to evaluate the retrieval's performance we conducted a validation study for the time ranges of boreal summer (June, July, and August) 2018 and boreal winter (December, January, February) 2018/2019 whereby we only include clear-sky observations (i.e. pixels with a effective cloud fraction smaller than 20%) and ice- and snowfree pixels. To avoid extreme outliers, we only include observations with an AMF > 0.1. As reference data for the validation we use TCWV from the Special Sensor Microwave Imager/Sounder (SSMIS), from the reanalysis model ERA-5 and from the ground-based GPS network SuomiNet. For the sake of completeness, we also briefly investigate higher cloud fractions at the end of each subsection and provide the results in the Supplementary Material.

As cloud input data we use the cloud information (effective cloud fraction at 440 nm and cloud top height) as well as the surface altitude from the TROPOMI L2 NO_2 product (Van Geffen et al., 2019) and as surface albedo input data we use the combination of the modified MODIS and OMI albedo described in Section 4. To distinguish between ocean and land surface, we use a land-sea mask derived from GSHHS coastline data (Wessel and Smith, 1996), 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.

6.1 SSMIS comparison

For the evaluation we use measurements from SSMIS onboard NOAA's f16 and f17 satellite processed by Remote Sensing Systems (RSS) and provided by NASA Global Hydrology Resource Center on a daily $0.25° \times 0.25°$ grid. SSMIS can observe the TCWV distribution under all-sky conditions over ocean with an accuracy of around $1\,\mathrm{kg\,m^{-2}}$ (Wentz, 1997; Mears et al., 2015). Since SSMIS changes its equator crossing time (ECT) we only include SSMIS observations whose ECT is within 3 hours (and 5 hours for f17, respectively) with respect to TROPOMI's ECT of 13:30 LT. For the intercomparison we only include SSMIS measurements that are not affected by rain.

Figure 19 depicts the comparison between SSMIS (f16, top row and f17, bottom row) and TROPOMI for boreal summer (left column) and winter (right column). For f16 (top row) the scatter is distributed closely along the 1-to-1 diagonal (dashed lines) for both seasons and the fitted regression lines (red solid lines) indicate a very good agreement between both data with slopes around 0.96, intercepts around $-1.6\,\mathrm{kg\,m^{-2}}$ for summer and $-1.7\,\mathrm{kg\,m^{-2}}$ for winter and coefficients of determination of $\mathrm{R}^2 = 0.91$. For f17 the comparison reveals similar agreement with slopes around 0.97 and intercepts around $-1.5\,\mathrm{kg\,m^{-2}}$ with $\mathrm{R}^2 = 0.89$ for both seasons. Overall, considering the differences in collocation time (3 hours and 5 hours for f16 and f17, respectively), the comparison shows that the TROPOMI TCWV retrieval can well capture the water vapour distribution over ocean.

To investigate the influence of clouds on our retrieval, we plot the difference (top row) and relative difference (bottom row) between TROPOMI and SSMIS as a function of the input cloud top height (CTH) in Figure 20 and 21 for f16 and f17, respectively. The median over the whole CTH range (blue dashed line) indicates an underestimation of the TROPOMI H_2O VCD of approximately 12-13% ($2.6\,\mathrm{kg\,m^{-2}}$). However, the large majority of data points is distributed within the CTH bin between 0-1 km revealing that the underestimation of the TROPOMI TCWV is mainly caused by low clouds. For mid clouds the median difference almost cancels out, whereas for high clouds it first increases and then remains almost constant with cloud top height.

Further validation results for SSMIS f16 and f17 separated into different cloud fraction and cloud top height bins for July 2018 are given in Fig.S10 and Fig. S11 respectively. The results indicate that there is no dependency with cloud fraction but a distinctive dependency with cloud top height: The retrieval underestimates for clouds below 1 km, is in very good agreement for mid level clouds (1 km to 4 km) and overestimates for higher clouds.

6.2 ERA-5 comparison

For the intercomparison between the reanalysis model ERA-5 and TROPOMI we use ERA-5 TCWV data provided by Copernicus Climate Change Service (2017) on a $0.25° \times 0.25°$ grid. We only take into account values which are within 1 hour with respect starting sensing time of the TROPOMI orbit and separate the data into data over ocean and data over land.

The results of the intercomparison are summarized in Figure 22. Over ocean (top row in Fig. 22) the results are similar to the results from the comparison between TROPOMI and SSMIS: Apart from slopes close to 0.95 and intercepts close to zero, the linear regression yields R^2 of 94% for summer and 95% for winter, respectively. Over land the linear regression still yields high values of the coefficient of determination R^2, but the TROPOMI retrieval generally underestimates the H_2O VCD by approximately 12% during summer (and 7% during winter). Since the values of the correlation coefficient are still high and the values over ocean coincide very well with the reference data sets, we assume that this underestimation has to be caused by a systematic uncertainty within the input parameters for our retrieval.

The influence of the cloud top height input is illustrated in Fig. 23 for data over ocean. The median is around $-1.6\,\mathrm{kg\,m^{-2}}$ (-7.1%) and $-1.3\,\mathrm{kg\,m^{-2}}$ (-6.7%) during summer and winter, respectively, whereby similar to SSMIS these underestimations are caused by the majority of data points within the 0-1 km CTH bin. For increasing CTH the deviation from the reference increases and leads to an overestimation. For data over land (Fig. 24) the CTH variability is much larger than over ocean,

i.e. most data points are now distributed between 0 to 3 km and the median is around values of $-1.5\,\mathrm{kg\,m^{-2}}$ (−10.3%) and $-0.4\,\mathrm{kg\,m^{-2}}$ (−4.0%) during summer and winter, respectively. Furthermore, low clouds still cause an underestimation and for mid to high clouds the deviations almost cancel out, but one can also observe an increasing scatter for winter data.

All these findings reveal that the combination of albedo uncertainties and uncertainties in the cloud properties (cloud fraction and cloud top height) as well as in the scale height parameterization have a distinctive influence on the AMF. The cloud products from TROPOMI rely on the OMI albedo which, as we have demonstrated in Sect. 4, has several problems over land surface. In addition, the uncertainty of the OMI albedo over land surface is higher than over ocean due to a highly spatiotemporal variability of the scenery and the differences between the monthly minimum and the monthly mean albedo are higher over land than over ocean. Furthermore, the cloud top height is calculated via the cloud top pressure and has to be combined with the surface pressure. Thus, the uncertainty of the cloud top height over land is higher than over ocean since over ocean the topography is much simpler.

Nevertheless, the complex radiative interactions between albedo and clouds might amplify or cancel out these deviations and thus make it difficult to draw clear conclusions.

As for the SSMIS comparison, further validation results for ERA-5 over ocean and land separated into different cloud fraction and cloud top height bins for July 2018 are given in Fig. S12 and Figure S13. Similar to SSMIS, the results over ocean reveal an underestimation for low clouds and an overestimation for high clouds that there is almost no dependency with cloud fraction. Over land low clouds still cause an underestimation, however for cloud top heights above 2 km the retrieval shows very good agreement to ERA-5, indicating that the input cloud top height for our retrieval is too low.

6.3 SuomiNet/GPS comparison

For the intercomparison with TCWV from ground-based GPS we use data from the network SuomiNet (Ware et al., 2000) provided by UCAR. SuomiNet stations are distributed over North and Central America and provide data every 30 min with a typical accuracy of $2\,\mathrm{kg\,m^{-2}}$ (Duan et al., 1996; Fang et al., 1998). Thus, we only take into account TROPOMI pixels within a distance of 0.1° to the GPS station and within 2 hours with respect to the GPS measurement.

Figure 25 illustrates scatter plots of the intercomparison between TROPOMI and SuomiNet for boreal summer and winter. For both seasons the robust regression indicates an underestimation of around 20% (i.e. slopes of 0.82 and 0.84) with high Pearson correlation coefficients of 88%. In order to investigate the influence of clouds on our retrieval, we plot the difference (top row) and the relative difference (bottom row) between TROPOMI and Suominet as a function of the input cloud top height (CTH) in Figure 26. The median over the whole CTH range (blue dashed line) indicates an underestimation of the TROPOMI H_2O VCD of approximately 14% ($3.5\,\mathrm{kg\,m^{-2}}$) during summer and of 8% ($0.8\,\mathrm{kg\,m^{-2}}$) during winter. However, during summer the median values for each 1 km CTH bin (blue dots) reveal that the underestimation is mainly caused by low clouds, whereas for mid and high clouds the median difference almost cancels out. During winter this pattern is not clearly observable due to much larger scatter, but also here low clouds mainly cause the underestimation in TCWV, whereby the difference is generally within the range of accuracy of the SuomiNet retrieval.

Figure S14 depicts further validation results separated into different cloud fraction and cloud top height bins for boreal summer

2018. Though the sample size is much smaller, similar results as for SSMIS and ERA-5 are obtained: ~~indenpendent~~ independent from the cloud fraction low clouds cause an underestimation of around 15-20% whereas for mid clouds the TROPOMI H_2O VCDs show much better agreement to the SuomiNet TCWV and for high clouds TROPOMI overestimates by around 10%.

7 Summary and conclusions

In this paper, we introduce a total column water vapour retrieval from TROPOMI spectra in the visible blue spectral range using an iterative vertical column conversion scheme and provide a detailed characterization of our retrieved H_2O VCD by performing a detailed uncertainty analysis and intercomparisons to reference data sets from the microwave sensor SSMIS, from the reanalysis model ERA-5 and from the ground-based measurements GPS network SuomiNet.

For the iteration scheme we describe the a priori water vapour profile as an exponential decay with a scale height H and developed an empirical parameterization for this scale height. This parameterization is based on COSMIC water vapour profile data and relates the a priori water vapour profile shape to the H_2O VCD, the seasonal cycle, the latitude and the vegetation (and NDVI, respectively). We demonstrate that we can correctly reproduce the scale heights, in particular for data at low latitudes (tropics and subtropics). However, we also observe an increasing scatter if higher latitudes are included in the comparison, likely because of the higher variability in H_2O VCD due to midlatitudinal cyclone dynamics and a general higher uncertainty in the COSMIC profile data for drier atmospheric conditions. Overall, the retrieved profile heights are very reasonable and we obtain a substantial improvement using the new parameterisation compared to the use of a prescribed constant water vapour profile.

For the uncertainty analysis we investigated the impact of several error sources on the H_2O SCD and AMF like clouds, surface albedo, profile shape and instrument properties. The error estimation reveals that the main SCD uncertainty is the fit error of the spectral analysis and that the main AMF uncertainties are caused by uncertainties in the surface albedo and water vapour profile shape. For the H_2O VCD we estimated a typical total relative error of around 10-20% for observations during clear-sky conditions, over ocean surface, and at low solar zenith angles. For observations during cloudy-sky conditions, over land surface, and high solar zenith angles the error reaches values of approximately 20-50%. Thus, the theoretically estimated errors are of the same order of magnitude as the deviations found during the retrieval's evaluation. However, uncertainties in the absorption cross section of water vapour are a further systematic error source that can additionally contribute up to 10%. Based on the LP-DOAS comparisons we estimate these errors to be around 5% for this study, so that they are negligible compared to the other error sources.

In the validation study we demonstrate that for clear-sky conditions the retrieved TROPOMI H_2O VCDs over ocean are in very good agreement to the reference data sets and can correctly capture the global water vapour distribution. Over land the TROPOMI retrieval can reproduce the TCWV distribution, however we also observe a distinctive underestimation of around 10% in particular during boreal summer.

Nevertheless, these underestimations might be caused by the uncertainties of the external input data for the retrieval: For instance the OMI LERs from Kleipool et al. (2008) are too high over tropical landmasses likely due to incorrect cloud filtering

which causes too high AMFs leading to too low H_2O VCDs. Although we tried to overcome this issue by using a surface reflectance product from MODIS Aqua, the cloud products from the TROPOMI L2 NO_2 product still rely on the OMI LER for calculating the effective cloud fraction and cloud top height and thus also have a large uncertainty. The intercomparisons to the reference data sets show that these uncertainties in the cloud products have a substantial impact on the H_2O VCD: Our investigations reveal that the input cloud top height is probably too low, which in turn leads to higher AMFs and consequently to an underestimation in TCWV. Yet, one has to consider that the radiative properties of the cloud and albedo products interact at a high degree of complexity so that a clear explanation or suggestion on how to overcome these issues is beyond the scope of this paper. Because of all these uncertainties we recommend for the general purpose to only use VCDs with an effective cloud fraction < 20% and AMF > 0.1, which represents a good compromise between spatial coverage and retrieval accuracy.

Overall, the successful application of the TCWV retrieval in the visible blue spectral range on TROPOMI measurement is very promising for further investigations including application to further satellite sensors such as OMI, SCIAMACHY, and GOME-1/2 or the upcoming Sentinel-4 ~~instrument~~ and Sentinel-5 instruments and expanding the retrieval to measurements contaminated by higher cloud fractions. As the retrieval allows for a fast execution of large data sets, investigations of longterm trends using a TCWV data set of merged timeseries of different satellite sensors are easily possible. However, since these data sets have to be uniform, they require consistent input data across the different satellite sensors, in particular for cloud products.

Data availability. The TROPOMI TCWV data presented here are available upon request.

Appendix A: Weighted linear least squares fit for spectral analysis

To handle the daily high data volume of TROPOMI and to avoid "jumps" of pixels included in the fit window, we implemented a weighted linear least squares fit for the DOAS analysis. The weights W are the fractional coverage of the pixel within the fit window (see also Fig. A1):

$$W(\lambda) = \begin{cases} 1 - \frac{|\lambda - \lambda_{low}|}{\Delta\lambda} & \frac{|\lambda - \lambda_{low}|}{\Delta\lambda} < 1 \wedge \lambda - \lambda_{low} < 0 \\ 1 & \lambda_{low} < \lambda < \lambda_{up} \\ 1 - \frac{|\lambda - \lambda_{up}|}{\Delta\lambda} & \frac{|\lambda - \lambda_{up}|}{\Delta\lambda} < 1 \wedge \lambda - \lambda_{up} > 0 \\ 0 & \text{else} \end{cases} \tag{A1}$$

with λ_{low} and λ_{up} the lower and upper boundary of the fit window and $\Delta\lambda$ the average wavelength increment within the fit window. The elements of the weight matrix are then given as $w_{ii} = \sqrt{W_{ii}(\lambda_i)}$. Hence Eq. (2) can be solved by simple linear algebra:

$$\mathbf{y'} = \mathbf{M'x}$$

$$\hat{\mathbf{x}} = \left(\mathbf{M'^T M'}\right)^{-1} \mathbf{M'^T y}$$

$$\mathbf{S} = \left(\mathbf{M'^T M'}\right)^{-1} \chi^2$$

$$\beta_i = \sqrt{S_{ii}}$$

with the solution of the linear problem $\hat{\mathbf{x}}$ containing the SCDs, $\mathbf{y'} = \mathrm{diag}(\mathbf{w})y$ the weighted measurement spectrum, $\mathbf{M'} = \mathrm{diag}(\mathbf{w})\mathbf{M}$ the weighted absorption structures to fit, β_i being the estimated 1σ fit error of the results for each fitted parameter, and χ^2 the reduced chi-square.

Appendix B: Evaluation of the water vapour absorption cross section

Figure A2 depicts intercomparisons between LP-DOAS and meteorological measurements of water vapour volume mixing ratios (WVMR) at different altitudes (10 m, 40 m and 200 m) at the CESAR Tower for day- and nighttime during the Cabauw Intercomparison of Nitrogen Dioxide Measuring Instruments 2 (CINDI-2) campaign. The results of the regression methods

indicate that for every altitude the LP-DOAS underestimates WVMR by around 17% during day and 11% during night. These findings independently confirm the results of further LP-DOAS measurements taken at the Cape Verde Atmospheric Observatory, for which Lampel et al. (2015) observed an underestimation of around 8% when using the water vapour line lists from HITRAN 2012. However, when using the water vapour line lists from HITRAN 2008 Lampel et al. (2015) observe an excellent agreement to the reference meteorological measurements at the observatory (see Table 8 in their paper). ~~With~~ 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 HITRAN 2008, the absorption peak of HITRAN 2012 is approximately 7-9% higher than HITRAN 2008 and the absorption peak of HITRAN 2016 is approximately 7-9% lower than HITRAN 2008. Combining these findings with the shortcomings of HITRAN 2016 indicated by Wang et al. (2019) and the observational evidence from the LP-DOAS measurements, we conclude that it is most adequate to use the water vapour line list from HITRAN 2008.

Appendix C: Evaluation of methods for calculating water vapour scale height

The water vapour scale height can be calculated in different ways. Here, we compare two different approaches: The first method is the calculation of the scale height via a weighted non-linear fit:

$$\min \sum_i \frac{\left(y_i - f(z_i, n_0, H_{nl})\right)^2}{\sigma_i^2}$$

$$f(z, n_0, H_{nl}) = n_0 e^{-\frac{z}{H_{nl}}}$$

where y_i are the COSMIC profile data points, $f(z, n_0, H)$ is the approximation of the exponential function, and σ_i is the inverse of the layer thickness at the observation y_i. The second method consists of summing up all the partial columns of the COSMIC profile data until a defined threshold is reached, which in our case is 63% of the H_2O VCD:

$$\frac{\int_0^{H_{sum}} n(z)dz}{\int_0^{\mathrm{TOA}} n(z)dz} > 1 - \frac{1}{e} \approx 63\%$$

$$\frac{\int_0^{H_{sum}} n(z)dz}{\int_0^{\mathrm{TOA}} n(z)dz} > 1 - \frac{1}{e} \approx 63\% \tag{C1}$$

Figure S4 depicts the mean profile shapes calculated using both methods as well as the mean profile shape of the COSMIC data for different latitude bins for the year 2013 for which the sample size is largest. Further statistics of goodness are given in Fig. S5 (bias), S6 (mean absolute error) and S7 (standard deviation). In general, the profile shapes of both methods agree well

with the COSMIC measurements, however Fig. S5 and Fig. S6 also reveal that the largest deviations occur in the lowermost troposphere, in particular for the southern polar regions. Nevertheless, the profiles of standard deviations in Fig. S7 also demonstrate that both methods are able to well capture the vertical and temporal variations in the water vapour profile shape and that these variations are within the same range of the variation of the COSMIC profile data.

Figure A4 depicts histograms of the relative AMF deviation for both methods for selected latitude regions assuming nadir viewing geometry and clear-sky conditions (like in Sect. 3.2 and Fig. 5). The peaks of the histograms for the sum method are close to the 0% line indicating very good agreement with AMF calculated from the COSMIC profiles. In contrast, the histograms for the non-linear fit peak at values around 2% and show a broader distribution than the histogram of the sum method, thus revealing an inferior agreement to the reference AMFs. For cloudy-sky conditions (see Fig. S8), both methods are biased to smaller AMF-values (deviations of around -5%) for a cloud top height of 1 km, but for higher clouds both methods show similar good agreement to the reference AMFs. Yet, the variance in the AMFs for the sum method is much smaller than in the AMFs for the non-linear fit.

In summary, the sum method is to be preferred because it provides more consistent results for clear-sky and cloudy-sky scenarios than the non-linear fit.

Author contributions. CB performed all calculations for this work and prepared the manuscript together with SB and TW and in collaboration with all coauthors. SB developed the concept of the linearized retrieval scheme and CB and SB implemented most of the retrieval code. SD helped with the McArtim calculations and HS helped with the tessellation of the TROPOMI H_2O VCD orbit data to a regular grid. TW supervised this study.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We would like to thank ESA and the S-5P/TROPOMI level 1 and level 2 teams for their ~~greak~~ great work on initiating and realizing TROPOMI and for providing the respective data sets. We also thank NASA for providing MODIS and SSMI data and ECMWF for providing reanalysis data. Furthermore, we acknowledge UCAR and ROMSAF for providing SuomiNet and COSMIC data. We also would like to thank Stefan Schmitt and Johannes Lampel from the Institute of Environmental Physics at the University of Heidelberg for performing the analysis of the LP-DOAS measurements during CINDI-2 and for providing the WVMR results in a very useful format.

References

Anthes, R. A.: Exploring Earth's atmosphere with radio occultation: contributions to weather, climate and space weather, Atmospheric Measurement Techniques, 4, 1077–1103, https://doi.org/10.5194/amt-4-1077-2011, https://www.atmos-meas-tech.net/4/1077/2011/, 2011.

Anthes, R. A., Bernhardt, P., Chen, Y., Cucurull, L., Dymond, K., Ector, D., Healy, S., Ho, S.-P., Hunt, D., Kuo, Y.-H., et al.: The COSMIC/FORMOSAT-3 mission: Early results, Bulletin of the American Meteorological Society, 89, 313–334, 2008.

Beirle, S., Sihler, H., and Wagner, T.: Linearisation of the effects of spectral shift and stretch in DOAS analysis, Atmospheric Measurement Techniques, 6, 661–675, https://doi.org/10.5194/amt-6-661-2013, https://www.atmos-meas-tech.net/6/661/2013/, 2013.

Beirle, S., Lampel, J., Lerot, C., Sihler, H., and Wagner, T.: Parameterizing the instrumental spectral response function and its changes by a super-Gaussian and its derivatives, Atmospheric Measurement Techniques, 10, 581–598, https://doi.org/10.5194/amt-10-581-2017, https://www.atmos-meas-tech.net/10/581/2017/, 2017.

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, 274–283, 2001.

Boggs, P. T., Boggs, P. T., Rogers, J. E., and Schnabel, R. B.: User's reference guide for odrpack version 2.01: Software for weighted orthogonal distance regression, Tech. Rep. NISTIR 4834, U.S. Department of Commerce, 1992.

Copernicus Climate Change Service: ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate, 2017.

Deutschmann, T., Beirle, S., Frieß, U., Grzegorski, M., Kern, C., Kritten, L., Platt, U., Prados-Román, C., Puķı, J., Wagner, T., et al.: The Monte Carlo atmospheric radiative transfer model McArtim: Introduction and validation of Jacobians and 3D features, Journal of Quantitative Spectroscopy and Radiative Transfer, 112, 1119–1137, 2011.

Didan, K., Munoz, A. B., Solano, R., and Huete, A.: MODIS vegetation index user's guide (MOD13 series), Tech. rep., Vegetation Index and PhenologyLab, https://doi.org/10.5067/MODIS/MYD13C2.006, https://lpdaac.usgs.gov/documents/103/MOD13_User_Guide_V6.pdf, 2015.

Duan, J., Bevis, M., Fang, P., Bock, Y., Chiswell, S., Businger, S., Rocken, C., Solheim, F., van Hove, T., Ware, R., McClusky, S., Herring, T. A., and King, R. W.: GPS Meteorology: Direct Estimation of the Absolute Value of Precipitable Water, Journal of Applied Meteorology, 35, 830–838, https://doi.org/10.1175/1520-0450(1996)035<0830:GMDEOT>2.0.CO;2, https://doi.org/10.1175/1520-0450(1996)035<0830:GMDEOT>2.0.CO;2, 1996.

Fang, P., Bevis, M., Bock, Y., Gutman, S., and Wolfe, D.: GPS meteorology: Reducing systematic errors in geodetic estimates for zenith delay, Geophysical Research Letters, 25, 3583–3586, https://doi.org/10.1029/98GL02755, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/98GL02755, 1998.

Gao, B.-C. and Kaufman, Y. J.: Water vapor retrievals using Moderate Resolution Imaging Spectroradiometer (MODIS) near-infrared channels, Journal of Geophysical Research: Atmospheres, 108, 2003.

Gordon, I., Rothman, L., Hill, C., Kochanov, R., Tan, Y., Bernath, P., Birk, M., Boudon, V., Campargue, A., Chance, K., Drouin, B., Flaud, J.-M., Gamache, R., Hodges, J., Jacquemart, D., Perevalov, V., Perrin, A., Shine, K., Smith, M.-A., Tennyson, J., Toon, G., Tran, H., Tyuterev, V., Barbe, A., Császár, A., Devi, V., Furtenbacher, T., Harrison, J., Hartmann, J.-M., Jolly, A., Johnson, T., Karman, T., Kleiner, I., Kyuberis, A., Loos, J., Lyulin, O., Massie, S., Mikhailenko, S., Moazzen-Ahmadi, N., Müller, H., Naumenko, O., Nikitin, A., Polyansky, O., Rey, M., Rotger, M., Sharpe, S., Sung, K., Starikova, E., Tashkun, S., Auwera, J. V., Wagner, G., Wilzewski, J., Wcisło, P., Yu, S., and Zak, E.: The HITRAN2016 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 203,

3 – 69, https://doi.org/https://doi.org/10.1016/j.jqsrt.2017.06.038, http://www.sciencedirect.com/science/article/pii/S0022407317301073, hITRAN2016 Special Issue, 2017.

Grossi, M., Valks, P., Loyola, D., Aberle, B., Slijkhuis, S., Wagner, T., Beirle, S., and Lang, R.: Total column water vapour measurements from GOME-2 MetOp-A and MetOp-B, Atmospheric Measurement Techniques, 8, 1111–1133, https://doi.org/10.5194/amt-8-1111-2015, https://www.atmos-meas-tech.net/8/1111/2015/, 2015.

Hajj, G., Kursinski, E., Romans, L., Bertiger, W., and Leroy, S.: A technical description of atmospheric sounding by GPS occultation, Journal of Atmospheric and Solar-Terrestrial Physics, 64, 451 – 469, https://doi.org/https://doi.org/10.1016/S1364-6826(01)00114-6, http://www.sciencedirect.com/science/article/pii/S1364682601001146, 2002.

Heise, S., Wickert, J., Beyerle, G., Schmidt, T., and Reigber, C.: Global monitoring of tropospheric water vapor with GPS radio occultation aboard CHAMP, Advances in Space Research, 37, 2222 – 2227, https://doi.org/https://doi.org/10.1016/j.asr.2005.06.066, http://www.sciencedirect.com/science/article/pii/S0273117705008379, atmospheric Remote Sensing: Earth's Surface, Troposphere, Stratosphere and Mesosphere - II, 2006.

Held, I. M. and Soden, B. J.: Water Vapor Feedback and Global Warming, Annual Review of Energy and the Environment, 25, 441–475, https://doi.org/10.1146/annurev.energy.25.1.441, https://doi.org/10.1146/annurev.energy.25.1.441, 2000.

Ho, S.-p., Kuo, Y.-H., Schreiner, W., and Zhou, X.: Using SI-traceable global positioning system radio occultation measurements for climate monitoring [In" State of the Climate in 2009"], Bulletin of the American Meteorological Society, 91, S36–S37, 2010a.

Ho, S.-p., Zhou, X., Kuo, Y.-H., Hunt, D., and Wang, J.-h.: Global Evaluation of Radiosonde Water Vapor Systematic Biases using GPS Radio Occultation from COSMIC and ECMWF Analysis, Remote Sensing, 2, 1320–1330, https://doi.org/10.3390/rs2051320, https://www.mdpi.com/2072-4292/2/5/1320, 2010b.

Kleipool, Q., Dobber, M., de Haan, J., and Levelt, P.: Earth surface reflectance climatology from 3 years of OMI data, Journal of Geophysical Research: Atmospheres, 113, 2008.

Kursinski, E., Hajj, G., Schofield, J., Linfield, R., and Hardy, K. R.: Observing Earth's atmosphere with radio occultation measurements using the Global Positioning System, Journal of Geophysical Research: Atmospheres, 102, 23 429–23 465, 1997.

Kurucz, R. L.: Solar Flux Atlas from 296 to 1300 nm, National Solar Observatory Atlas, 1, 1984.

Lampel, J., Pöhler, D., Tschritter, J., Frieß, U., and Platt, U.: On the relative absorption strengths of water vapour in the blue wavelength range, Atmospheric Measurement Techniques, 8, 4329–4346, https://doi.org/10.5194/amt-8-4329-2015, https://www.atmos-meas-tech.net/8/4329/2015/, 2015.

Lang, R., Casadio, S., Maurellis, A. N., and Lawrence, M. G.: Evaluation of the GOME Water Vapor Climatology 1995–2002, Journal of Geophysical Research: Atmospheres, 112, https://doi.org/10.1029/2006JD008246, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2006JD008246, 2007.

Levelt, P. F., van den Oord, G. H., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O., and Saari, H.: The ozone monitoring instrument, IEEE Transactions on geoscience and remote sensing, 44, 1093–1101, 2006.

Mears, C. A., Wang, J., Smith, D., and Wentz, F. J.: Intercomparison of total precipitable water measurements made by satellite-borne microwave radiometers and ground-based GPS instruments, Journal of Geophysical Research: Atmospheres, 120, 2492–2504, 2015.

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.

Noël, S., Buchwitz, M., Bovensmann, H., Hoogen, R., and Burrows, J. P.: Atmospheric water vapor amounts retrieved from GOME satellite data, Geophysical Research Letters, 26, 1841–1844, https://doi.org/10.1029/1999GL900437, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/1999GL900437, 1999.

Noël, S., Buchwitz, M., and Burrows, J.: First retrieval of global water vapour column amounts from SCIAMACHY measurements, Atmospheric Chemistry and Physics, 4, 111–125, 2004.

Platt, U. and Stutz, J.: Differential Optical Absorption Spectroscopy: Principles and Applications, Physics of Earth and Space Environments, Springer Berlin Heidelberg, https://books.google.de/books?id=y90g9yF1QncC, 2008.

Rosenkranz, P. W.: Retrieval of temperature and moisture profiles from AMSU-A and AMSU-B measurements, IEEE Transactions on Geoscience and Remote Sensing, 39, 2429–2435, 2001.

Rothman, L., Gordon, I., Babikov, Y., Barbe, A., Benner, D. C., Bernath, P., Birk, M., Bizzocchi, L., Boudon, V., Brown, L., Campargue, A., Chance, K., Cohen, E., Coudert, L., Devi, V., Drouin, B., Fayt, A., Flaud, J.-M., Gamache, R., Harrison, J., Hartmann, J.-M., Hill, C., Hodges, J., Jacquemart, D., Jolly, A., Lamouroux, J., Roy, R. L., Li, G., Long, D., Lyulin, O., Mackie, C., Massie, S., Mikhailenko, S., Müller, H., Naumenko, O., Nikitin, A., Orphal, J., Perevalov, V., Perrin, A., Polovtseva, E., Richard, C., Smith, M., Starikova, E., Sung, K., Tashkun, S., Tennyson, J., Toon, G., Tyuterev, V., and Wagner, G.: The HITRAN2012 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 130, 4 – 50, https://doi.org/https://doi.org/10.1016/j.jqsrt.2013.07.002, http://www.sciencedirect.com/science/article/pii/S0022407313002859, hITRAN2012 special issue, 2013.

Rothman, L. S., Gordon, I. E., Barbe, A., Benner, D. C., Bernath, P. F., Birk, M., Boudon, V., Brown, L. R., Campargue, A., Champion, J.-P., et al.: The HITRAN 2008 molecular spectroscopic database, Journal of Quantitative Spectroscopy and Radiative Transfer, 110, 533–572, 2009.

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.

Schneider, A., Borsdorff, T., aan de Brugh, J., Aemisegger, F., Feist, D. G., Kivi, R., Hase, F., Schneider, M., and Landgraf, J.: First data set of H_2O/HDO columns from the Tropospheric Monitoring Instrument (TROPOMI), Atmospheric Measurement Techniques, 13, 85–100, https://doi.org/10.5194/amt-13-85-2020, https://www.atmos-meas-tech.net/13/85/2020/, 2020.

Serdyuchenko, A., Gorshelev, V., Weber, M., Chehade, W., and Burrows, J. P.: High spectral resolution ozone absorption cross-sections – Part 2: Temperature dependence, Atmospheric Measurement Techniques, 7, 625–636, https://doi.org/10.5194/amt-7-625-2014, https://www.atmos-meas-tech.net/7/625/2014/, 2014.

Siegel, A. F.: Robust regression using repeated medians, Biometrika, 69, 242–244, https://doi.org/10.1093/biomet/69.1.242, https://doi.org/10.1093/biomet/69.1.242, 1982.

Smith, E. K. and Weintraub, S.: The Constants in the Equation for Atmospheric Refractive Index at Radio Frequencies, Proceedings of the IRE, 41, 1035–1037, https://doi.org/10.1109/JRPROC.1953.274297, 1953.

Steffensen, J. F.: Remarks on iteration., Scandinavian Actuarial Journal, 1933, 64–72, https://doi.org/10.1080/03461238.1933.10419209, https://doi.org/10.1080/03461238.1933.10419209, 1933.

Sulla-Menashe, D., Gray, J. M., Abercrombie, S. P., and Friedl, M. A.: Hierarchical mapping of annual global land cover 2001 to present: The MODIS Collection 6 Land Cover product, Remote Sensing of Environment, 222, 183 – 194, https://doi.org/https://doi.org/10.1016/j.rse.2018.12.013, http://www.sciencedirect.com/science/article/pii/S0034425718305686, 2019.

Susskind, J., Barnet, C. D., and Blaisdell, J. M.: Retrieval of atmospheric and surface parameters from AIRS/AMSU/HSB data in the presence of clouds, IEEE Transactions on Geoscience and Remote Sensing, 41, 390–409, 2003.

Thalman, R. and Volkamer, R.: Temperature dependent absorption cross-sections of O2-O2 collision pairs between 340 and 630 nm and at atmospherically relevant pressure, Phys. Chem. Chem. Phys., 15, 15 371–15 381, https://doi.org/10.1039/C3CP50968K, http://dx.doi.org/10.1039/C3CP50968K, 2013.

Tilstra, L. G., Tuinder, O. N. E., Wang, P., and Stammes, P.: Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY, Journal of Geophysical Research: Atmospheres, 122, 4084–4111, https://doi.org/10.1002/2016JD025940, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2016JD025940, 2017.

Van Geffen, J., Boersma, K., Eskes, H., Maasakkers, J., and Veefkind, J.: TROPOMI ATBD of the total and tropospheric NO2 data products, Tech. Rep. S5P-KNMI-L2-0005-RP, Royal Netherlands Meteorological Institute, https://sentinel.esa.int/documents/247904/2476257/Sentinel-5P-TROPOMI-ATBD-NO2-data-products, 2019.

Vandaele, A., Hermans, C., Simon, P., Carleer, M., Colin, R., Fally, S., Merienne, M., Jenouvrier, A., and Coquart, B.: Measurements of the NO2 absorption cross-section from 42 000 cm- 1 to 10 000 cm- 1 (238–1000 nm) at 220 K and 294 K, Journal of Quantitative Spectroscopy and Radiative Transfer, 59, 171 – 184, https://doi.org/https://doi.org/10.1016/S0022-4073(97)00168-4, http://www.sciencedirect.com/science/article/pii/S0022407397001684, atmospheric Spectroscopy Applications 96, 1998.

Veefkind, J., Aben, I., McMullan, K., Förster, H., de Vries, J., Otter, G., Claas, J., Eskes, H., de Haan, J., Kleipool, Q., van Weele, M., Hasekamp, O., Hoogeveen, R., Landgraf, J., Snel, R., Tol, P., Ingmann, P., Voors, R., Kruizinga, B., Vink, R., Visser, H., and Levelt, P.: TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications, Remote Sensing of Environment, 120, 70 – 83, https://doi.org/https://doi.org/10.1016/j.rse.2011.09.027, http://www.sciencedirect.com/science/article/pii/S0034425712000661, the Sentinel Missions - New Opportunities for Science, 2012.

Veefkind, J. P., de Haan, J. F., Sneep, M., and Levelt, P. F.: Improvements to the OMI O_2–O_2 operational cloud algorithm and comparisons with ground-based radar–lidar observations, Atmospheric Measurement Techniques, 9, 6035–6049, https://doi.org/10.5194/amt-9-6035-2016, https://www.atmos-meas-tech.net/9/6035/2016/, 2016.

Wagner, T., Heland, J., Zöger, M., and Platt, U.: A fast H_2O total column density product from GOME - Validation with in-situ aircraft measurements, Atmospheric Chemistry and Physics, 3, 651–663, https://doi.org/10.5194/acp-3-651-2003, https://www.atmos-chem-phys.net/3/651/2003/, 2003.

Wagner, T., Beirle, S., and Deutschmann, T.: Three-dimensional simulation of the Ring effect in observations of scattered sun light using Monte Carlo radiative transfer models, Atmospheric Measurement Techniques, 2, 113–124, https://doi.org/10.5194/amt-2-113-2009, https://www.atmos-meas-tech.net/2/113/2009/, 2009.

Wagner, T., Beirle, S., Sihler, H., and Mies, K.: A feasibility study for the retrieval of the total column precipitable water vapour from satellite observations in the blue spectral range, Atmospheric Measurement Techniques, 6, 2593–2605, https://doi.org/10.5194/amt-6-2593-2013, https://www.atmos-meas-tech.net/6/2593/2013/, 2013.

Wang, H., Souri, A. H., González Abad, G., Liu, X., and Chance, K.: Ozone Monitoring Instrument (OMI) Total Column Water Vapor version 4 validation and applications, Atmospheric Measurement Techniques, 12, 5183–5199, https://doi.org/10.5194/amt-12-5183-2019, https://www.atmos-meas-tech.net/12/5183/2019/, 2019.

Ware, R. H., Fulker, D. W., Stein, S. A., Anderson, D. N., Avery, S. K., Clark, R. D., Droegemeier, K. K., Kuettner, J. P., Minster, J. B., and Sorooshian, S.: SuomiNet: A real-time national GPS network for atmospheric research and education, Bulletin of the American Meteorological Society, 81, 677–694, 2000.

Weaver, C. and Ramanathan, V.: Deductions from a simple climate model: Factors governing surface temperature and atmospheric thermal structure, Journal of Geophysical Research: Atmospheres, 100, 11 585–11 591, 1995.

Wendland, W. and Steinbach, O.: Analysis: Integral- und Differentialrechnung, gewöhnliche Differentialgleichungen, komplexe Funktionentheorie, Lehrbuch : Mathematik, Vieweg+Teubner Verlag, https://books.google.de/books?id=ZbWIBwAAQBAJ, 2015.

Wentz, F. J.: A well-calibrated ocean algorithm for special sensor microwave/imager, Journal of Geophysical Research: Oceans, 102, 8703–8718, 1997.

Wessel, P. and Smith, W. H. F.: A global, self-consistent, hierarchical, high-resolution shoreline database, Journal of Geophysical Research: Solid Earth, 101, 8741–8743, https://doi.org/10.1029/96JB00104, https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JB00104, 1996.

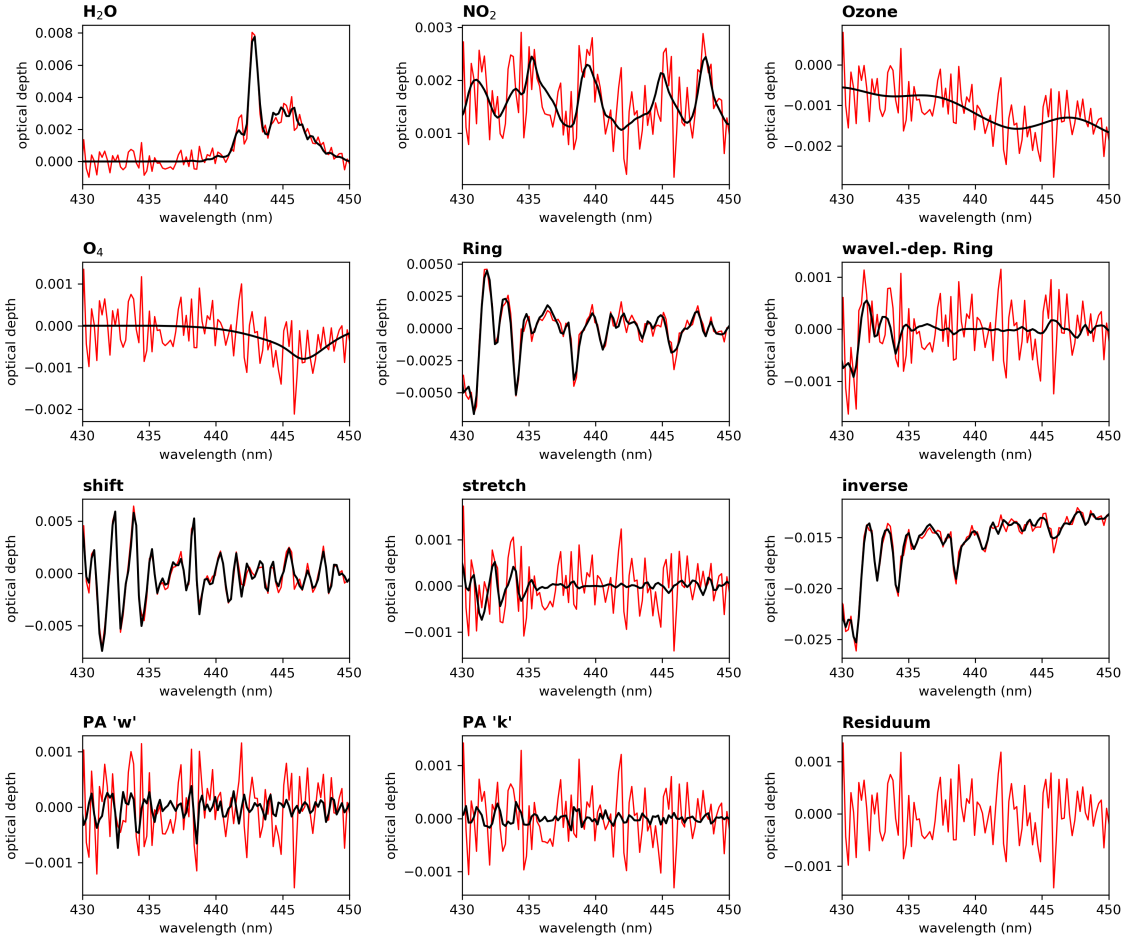


Figure 1. Example of a typical spectral analysis of a TROPOMI measurement spectrum (RMS: ~~0.6‰~~0.5‰, orbit: ~~4088~~6930, ~~14.88~~–7.41°N, ~~-2.59~~–111.97°E). The black line indicates the fit result for the respective trace gas and the red line indicates the residual spectrum and residual noise for each constituent.

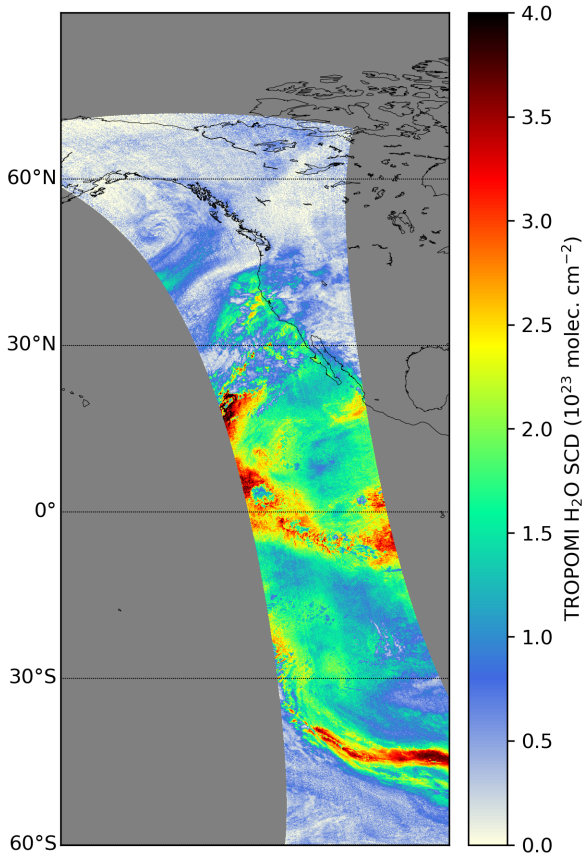


Figure 2. H_2O SCD distribution retrieved from one TROPOMI measurements (orbit: 6930) on 13th February 2019 during an atmospheric river event at the Western US coast.

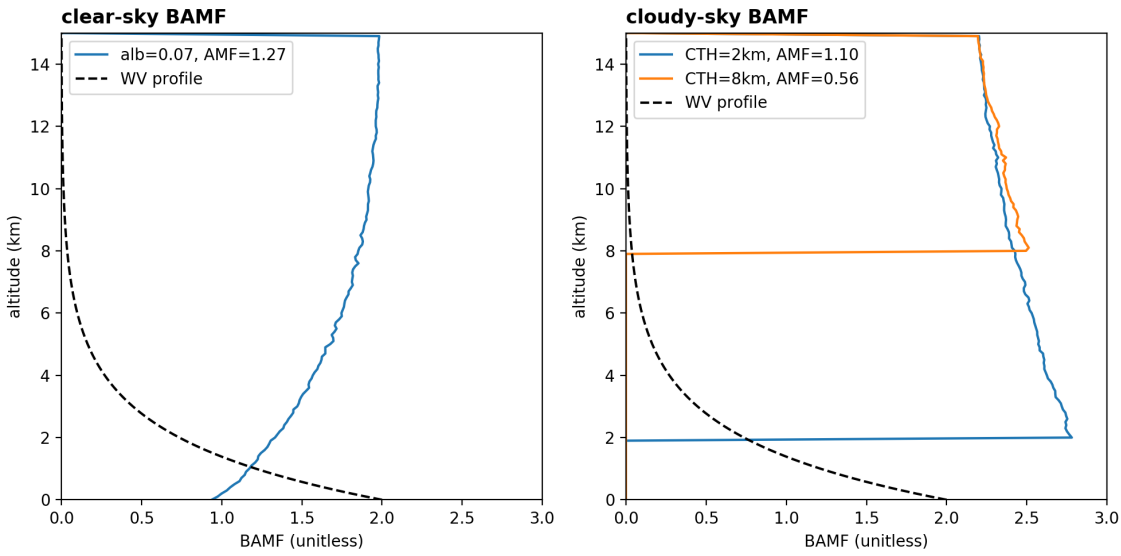


Figure 3. Examples of typical BAMF profiles for different observation scenarios under clear- (left) and cloudy-sky (right panel) conditions. For all profiles we assume a solar zenith angle of 0°, line of sight of −90°, and a solar relative azimuth angle of 0°. For the clear-sky case the BAMF profile is illustrated for a surface albedo 7%. For the cloudy-sky cases the profiles are depicted cloud top heights of 2 and 8 km and the respective AMFs are calculated assuming a surface albedo of 7% and an effective cloud fraction of 20%. The black dashed lines indicate relative water vapour concentrations with a scale height of 2 km.

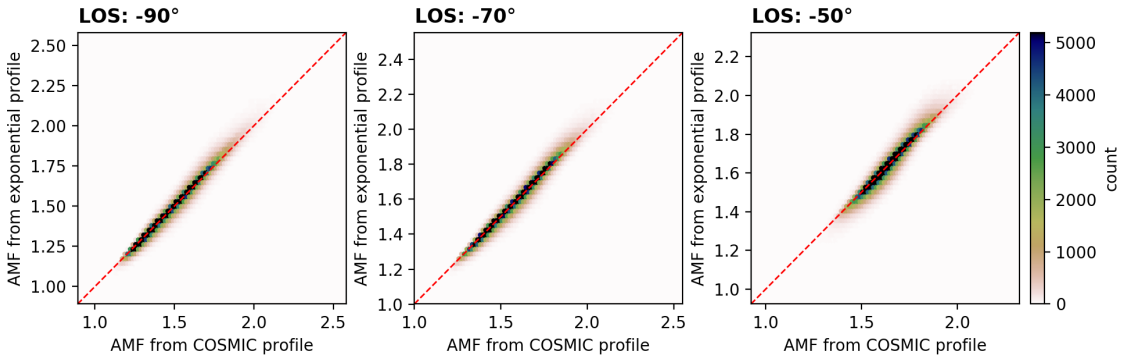


Figure 4. 2D histograms comparing synthetic AMFs (calculated via sum method) for different line of sights angles (−90°, −70° and −50°, from left to right) assuming clear-sky conditions. The color depicts the number of observations within one defined bin of the 2D histogram and the red dashed line represents the 1-to-1 diagonal.

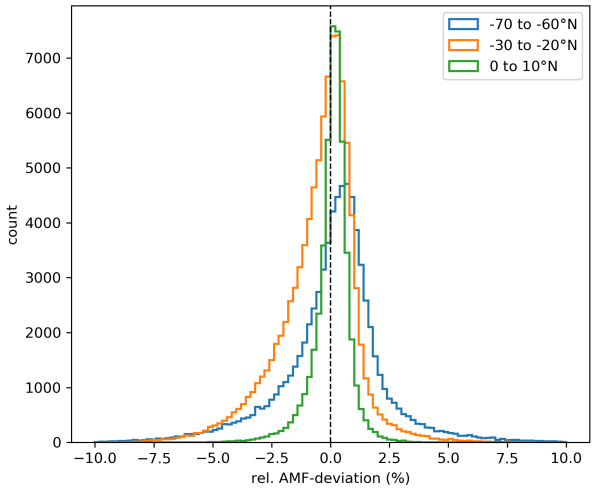


Figure 5. Histogram of the relative deviation of the calculated synthetic AMFs between exponential profile and COSMIC profile for selected latitude bins (0° to 10°N, -30° to -20°N, and -70° to -60°N) assuming clear sky conditions and nadir viewing geometry.

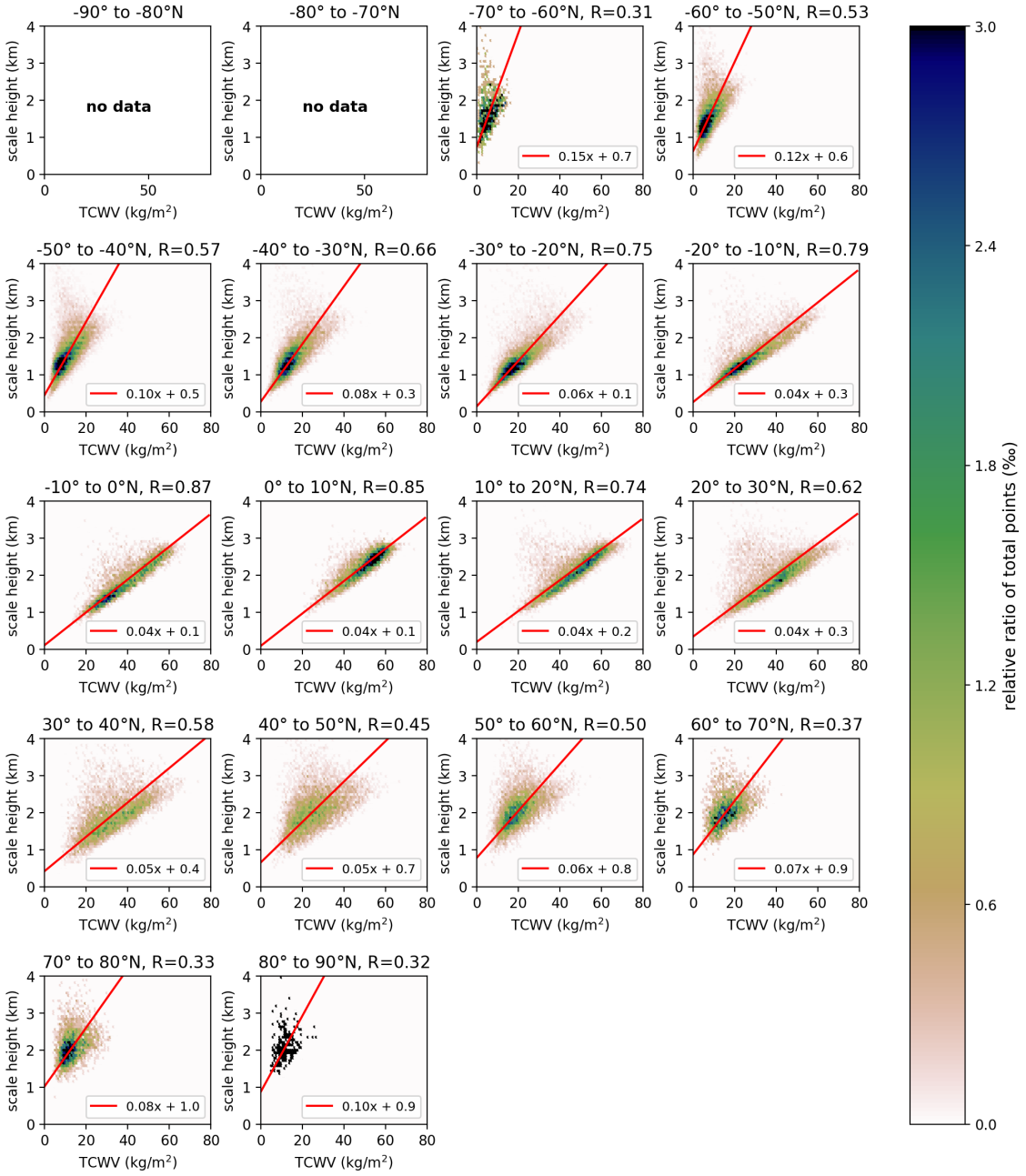


Figure 6. 2D histograms depicting the relation between calculated scale height and TCWV from COSMIC profiles for boreal summer (June, July, and August) only over ocean summarized in 10° latitude bins. Only latitude bins with a sample size of 1500 data points are illustrated. The color indicates the relative share of total points within one bin of the histogram and the red line indicates the fit results of the orthogonal distance regression with detailed results in the legend of each subplot. In addition the Pearson correlation coefficient for each data set is given in the title of each subplot.

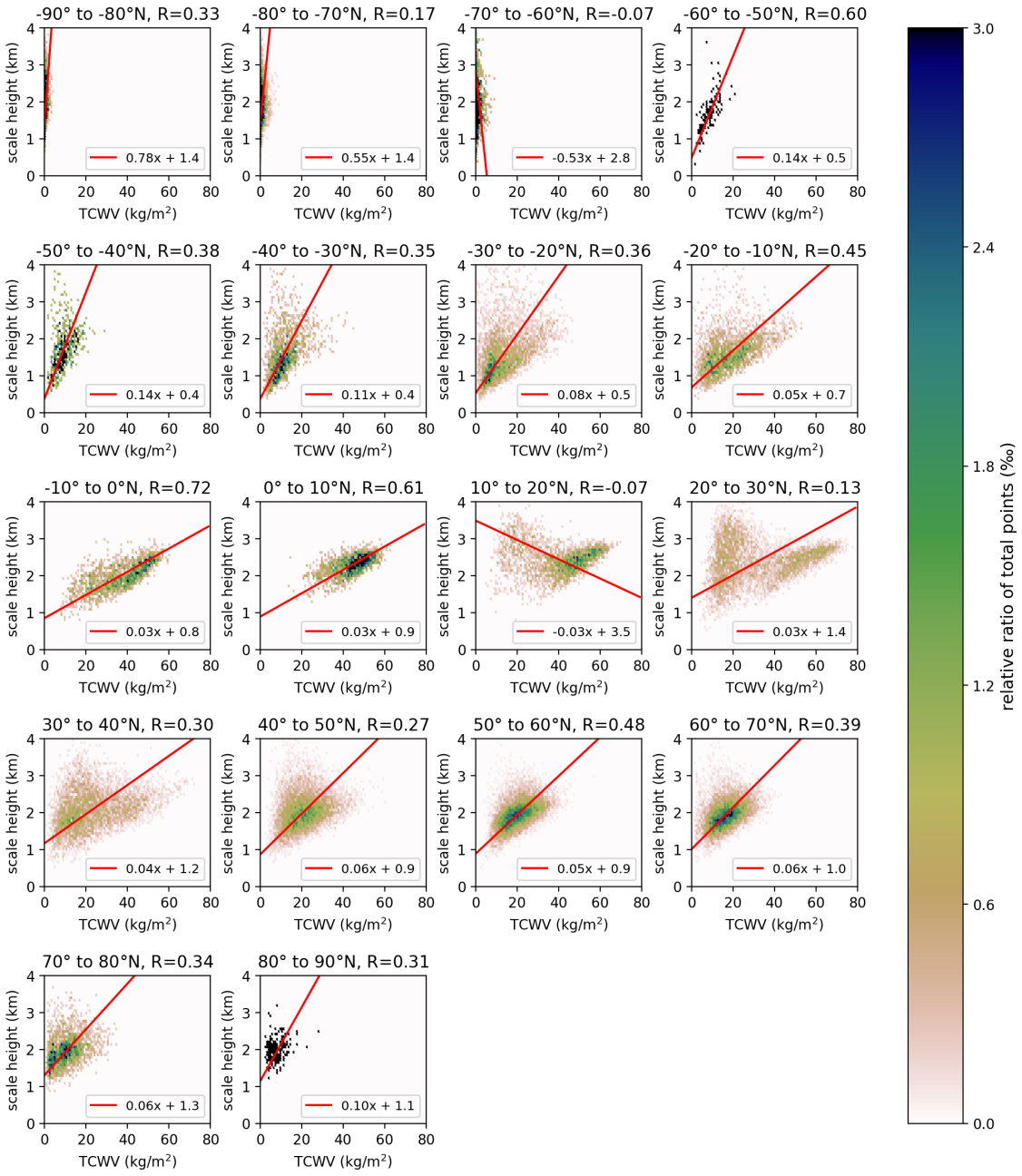


Figure 7. Same as Fig.6, but for data over land.

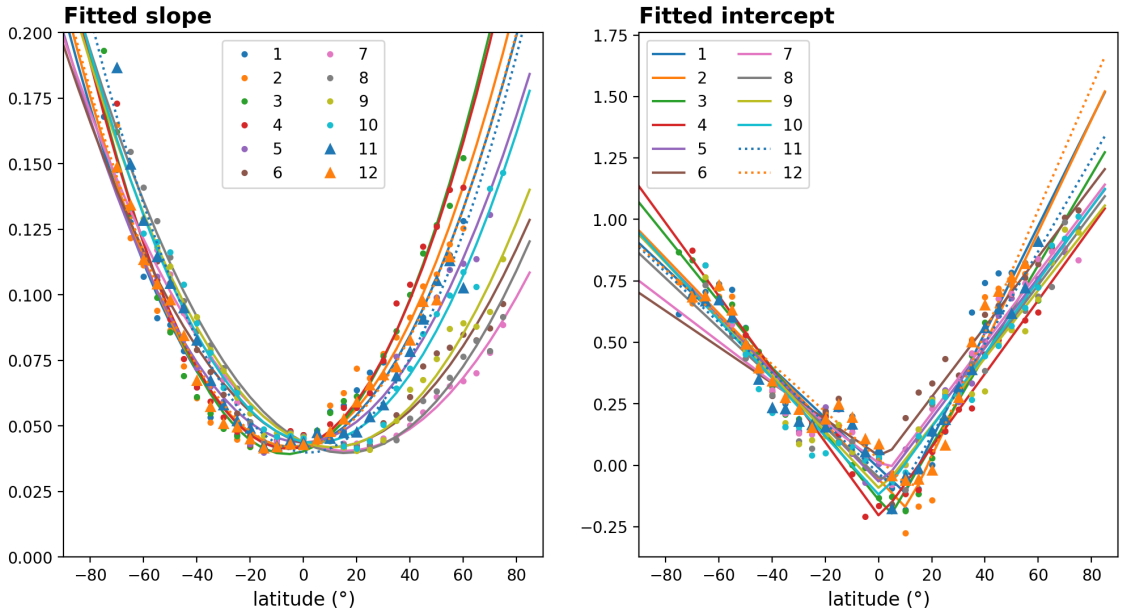


Figure 8. Summary of the results of the ODR fit between COSMIC scale height and COSMIC TCWV as a function of latitude and month for data over ocean. The left panel illustrates the fitted slopes, the right panel the corresponding fitted intercepts whereby the coloured points represent the fit results and the lines represent the approximations for $\alpha(\theta, t)$ and $\beta(\theta, t)$ for each month.

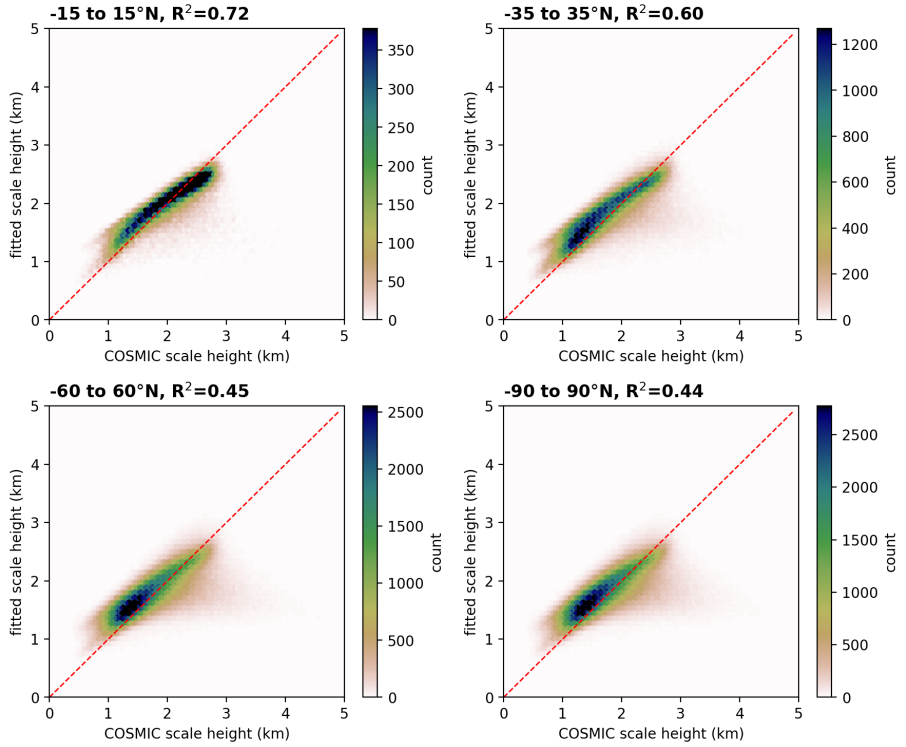


Figure 9. 2D histograms of the distribution between the parameterized scale height and the COSMIC scale height over ocean for selected global latitude zones.

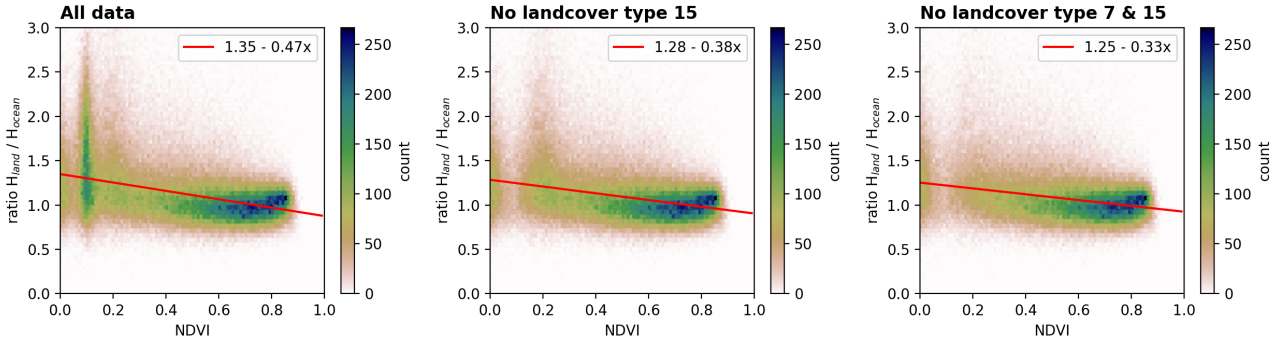


Figure 10. 2D histograms of the distribution between the ratio H_{land}/H_{ocean} against the NDVI for different filtered data sets: The left panel includes all data points, the center panel includes all points except those with MODIS landcover type 15 (corresponding to deserts), and the right panel includes all points except landcover types 7 (corresponding to open shrublands) and 15. The red solid line represents the fit result using the Siegel algorithm with details of the fit results in the legends of each panel.

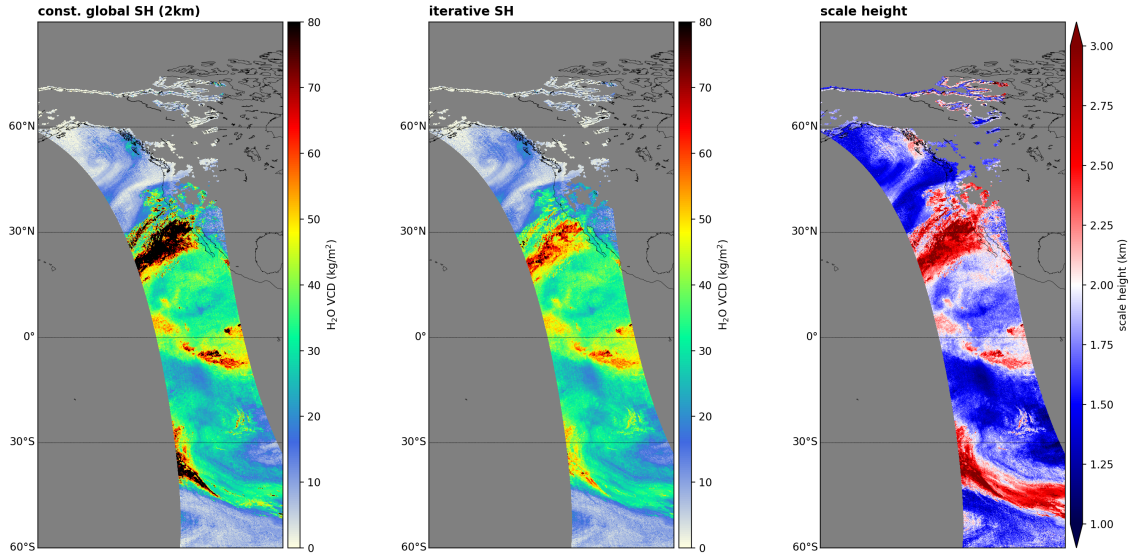


Figure 11. Comparison of the H_2O VCD calculated using a global constant a priori profile shape of 2 km (left) and the iterative scale height method (center) for all-sky conditions. The right panel illustrates the water vapour scale height estimated within the retrieval's VCD conversion. All panels show an atmospheric river hitting the East Pacific/Western US Coast on the 13th February 2019. Invalid pixels are coloured grey.

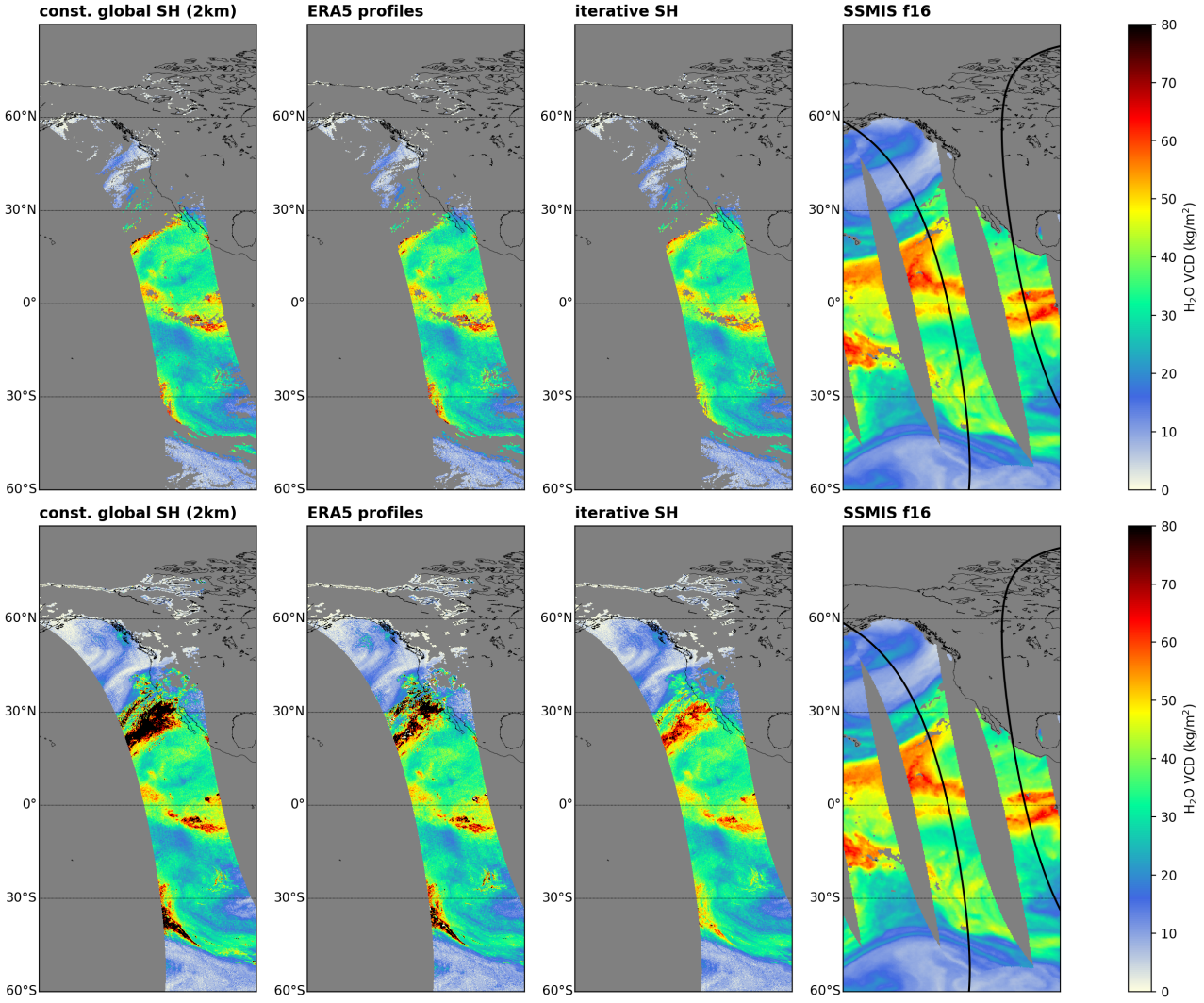


Figure 12. Comparison of the H_2O VCD calculated using a global constant a priori profile shape of 2 km (first from left), ERA-5 profiles (second from left), and the iterative scale height method (~~center~~third from left) for clear-sky (effective cloud fraction < 20%; top row) and cloudy-sky conditions (effective cloud fraction $\leq$ 100%; bottom row) with TCWV from SSMIS f16 (right) for the same scenary as in Figure 11. Invalid pixels are coloured grey. The solid black lines indicate the edges of the TROPOMI swath.

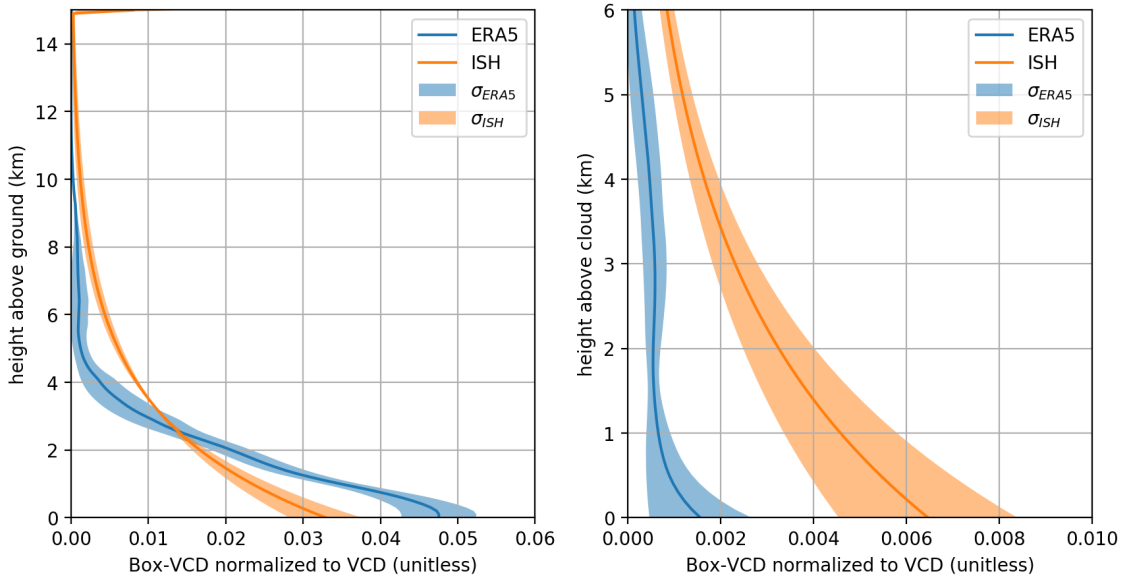


Figure 13. Mean normalized Box-VCD profiles of ERA-5 and the iterative scale height approach for cases of distinctive VCD disagreement within the region of the atmospheric river. The left panel illustrates the mean of the selected profiles from ground up to 15 km and the right panel the mean of the same profiles, but from cloud top up to 6 km above the cloud top. The solid lines indicate the mean profiles and the shaded areas the corresponding 1σ standard deviation.

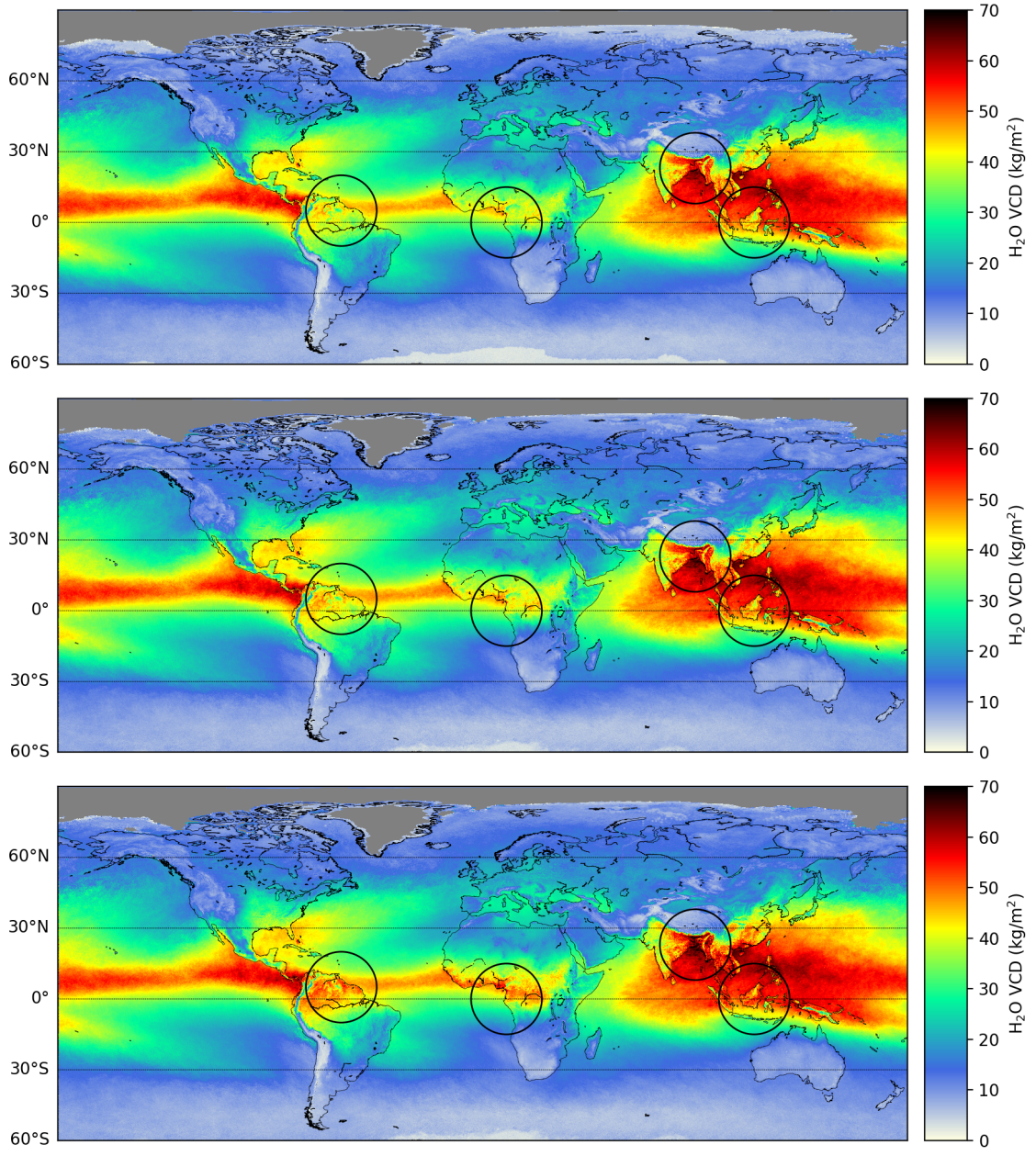


Figure 14. Comparison of the effect of different land albedo input data on the mean H_2O VCD for boreal summer 2018 (top row: OMI monthly mean LER, middle row: OMI monthly minimum LER, bottom row: scaled MODIS Aqua blue surface reflectance). Only pixels with an effective cloud fraction smaller than 20% are included.

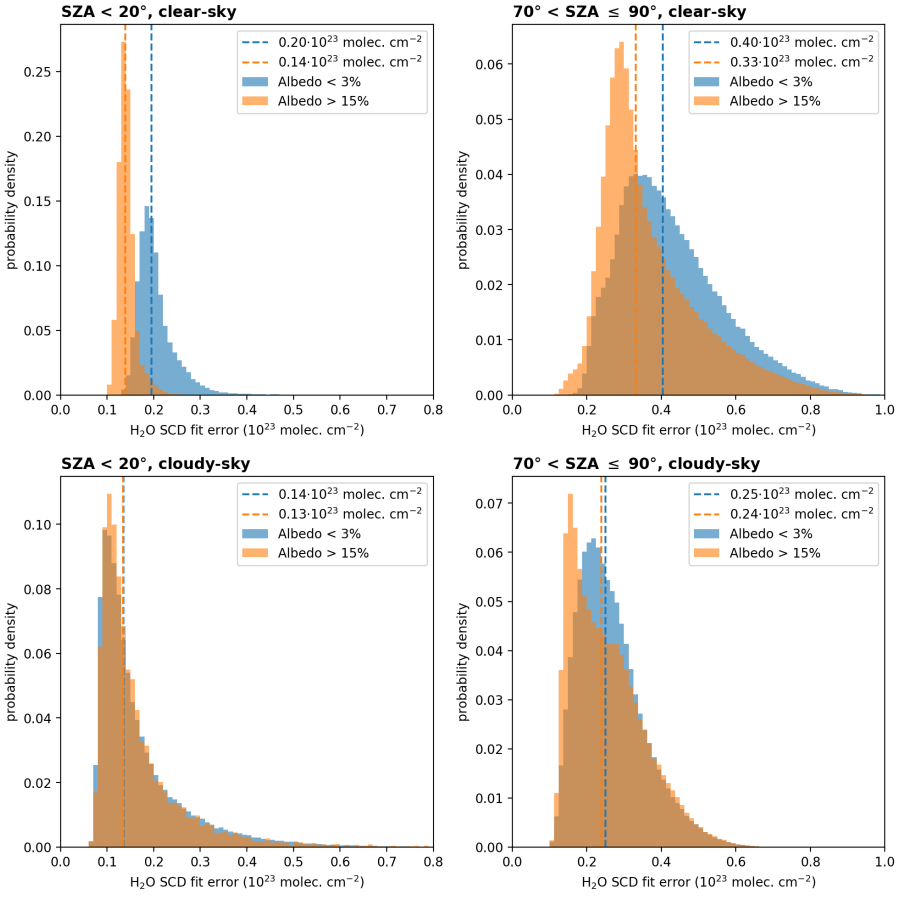


Figure 15. Histograms of the standard H_2O SCD fit error distribution for small (SZA<20°, left panel) and large (70°< SZA≤90°, right panel) solar zenith angles for relatively small (<3%, orange) and high (>15%, blue) surface albedo values for clear-sky (cloud fraction < 5%, top row) and cloudy-sky (cloud fraction > 20%, bottom row) conditions. The coloured dashed lines represent the median of the respective distributions and their values are given in the legend of each panel.

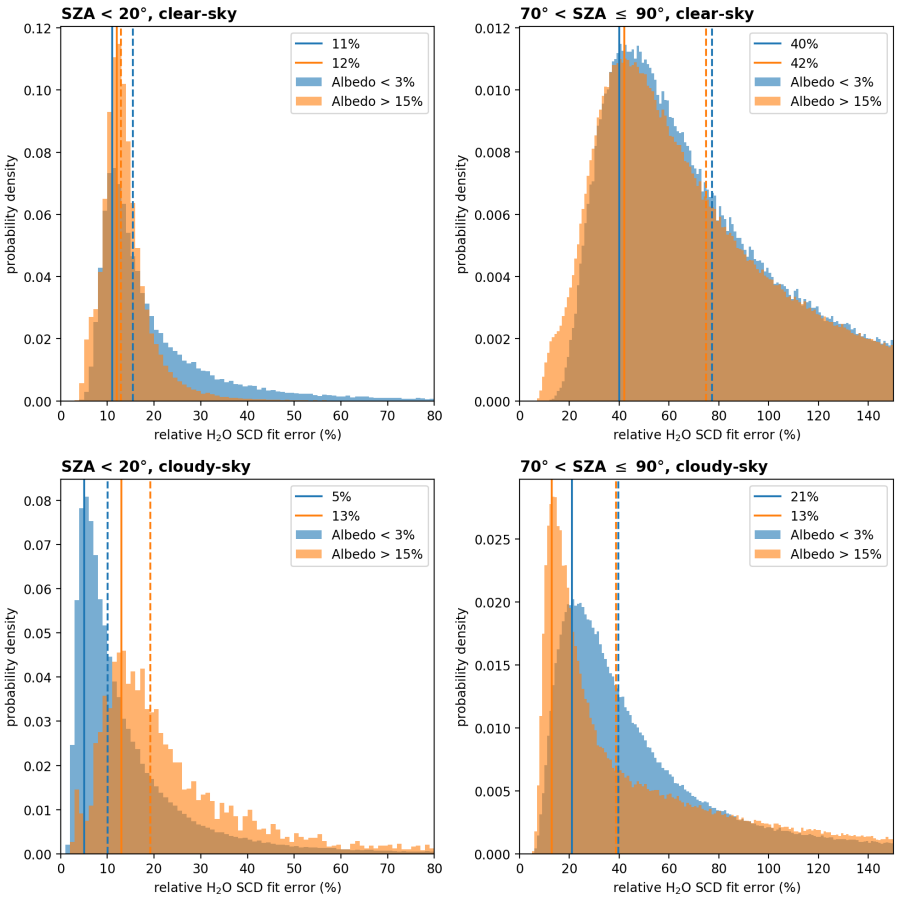


Figure 16. Histograms of the relative H_2O SCD fit error distribution for small (SZA<20°, left panel) and large (70°< SZA≤90°, right panel) solar zenith angles for relatively small (<3%, orange) and high (>15%, blue) surface albedo values for clear-sky (cloud fraction < 5%, top row) and cloudy-sky (cloud fraction > 20%, bottom row) conditions. The coloured dashed lines represent the median of the respective distributions and the solid lines represent the location of maximal probability density (values given in the legend of each panel).

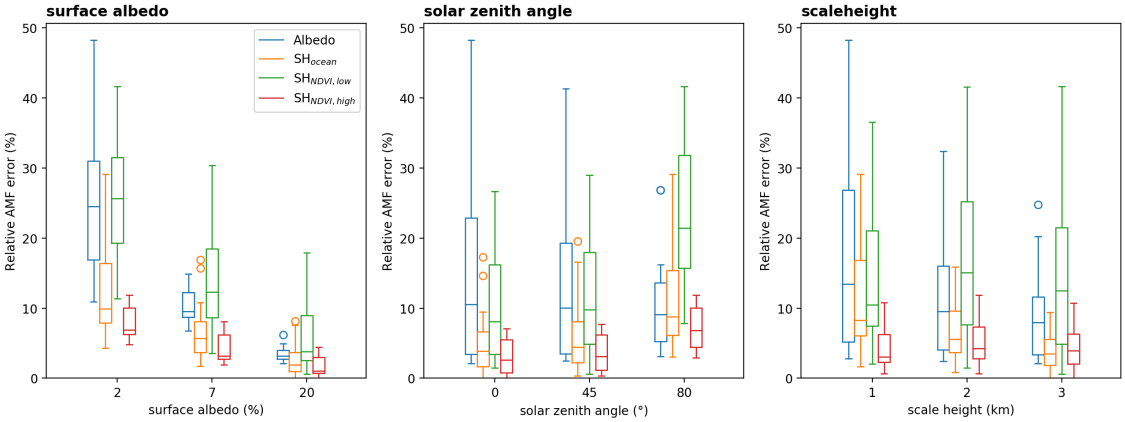


Figure 17. Box-whisker plots of the relative AMF errors for clear-sky conditions due to uncertainties within the retrieval's input parameters (blue: surface albedo, orange: scale height over high vegetation, green: scale height over low vegetation, red: scale height over ocean) according to the uncertainty assumptions in Tab. 6 and simulated for the standard scenarios of the surface albedo, solar zenith angle, and scale height given in Table 5.

.

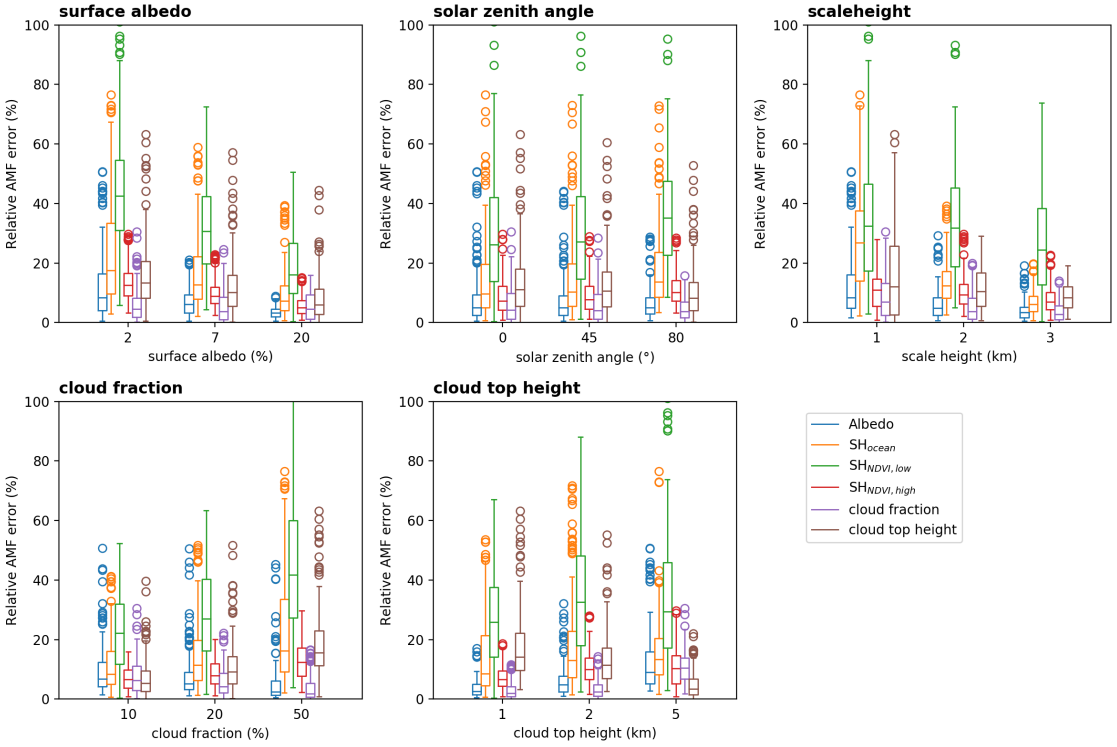


Figure 18. Box-whisker plots of the relative AMF errors for cloudy-sky conditions due to uncertainties within the retrieval's input parameters according to the uncertainty assumptions in Tab. 6 and simulated for the standard scenarios given in Table 5.

.

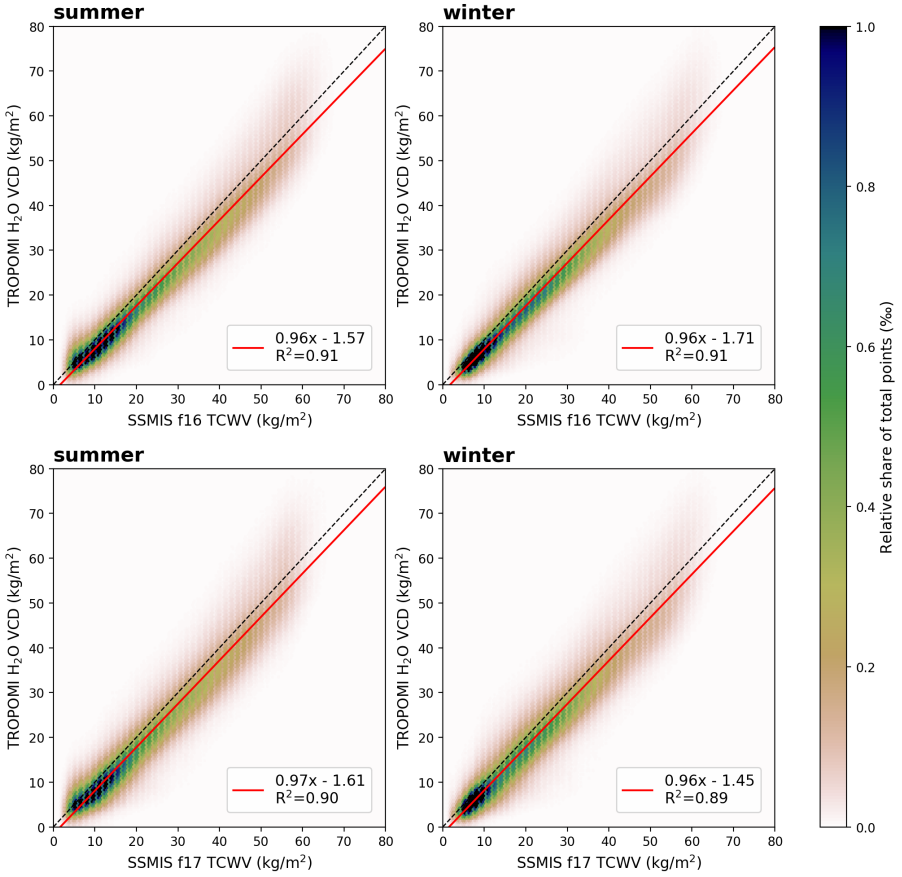


Figure 19. 2D histograms for the comparison between TROPOMI and SSMIS f16 (top row) and f17 (bottom row) for clear-sky conditions (CF<20%) for boreal summer (left column) and boreal winter (right column) where the color indicates the relative share of total data points. The black dotted line indicates the 1-to-1 diagonal and the red solid line represents the results of the linear regression. The parameters of the linear regression and the coefficient of determination are given in the box in each panel.

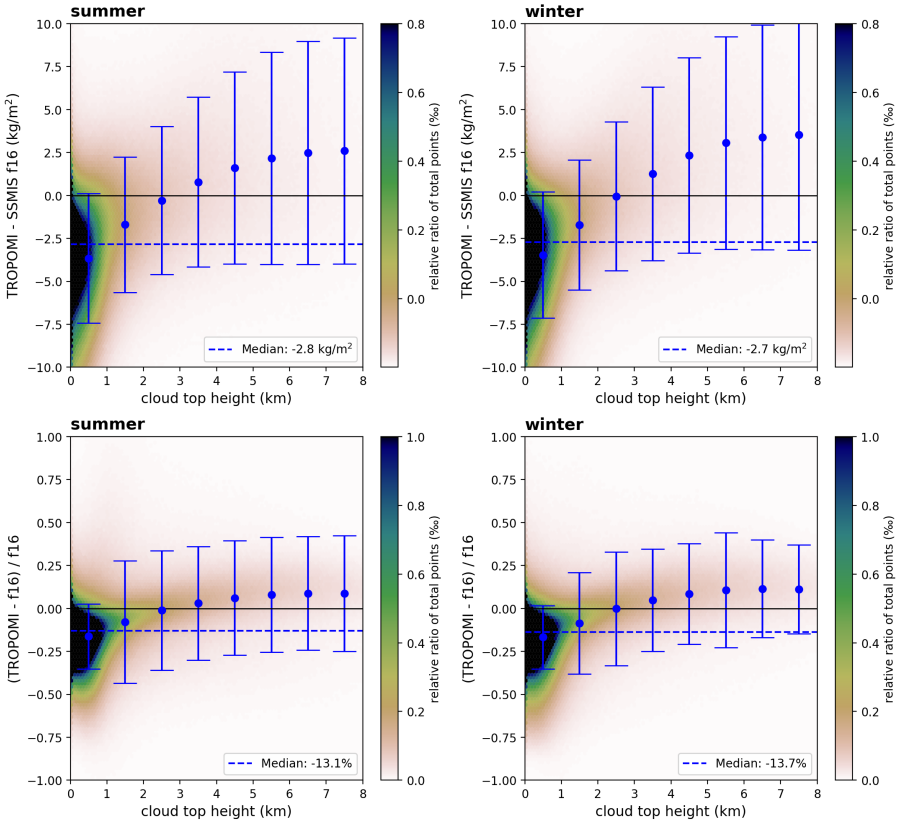


Figure 20. 2D histograms of the difference (TROPOMI-SSMIS f16, top row) and relative difference (TROPOMI-f16)/f16 (bottom row) as a function of the input cloud top height (CTH) for clear-sky conditions (CF<20%) for summer (left column) and winter (right column). The blue dashed line represents the median over the whole CTH range. The blue dots represent the median within a 1 km CTH and the error bars represent their respective 1σ standard deviation.

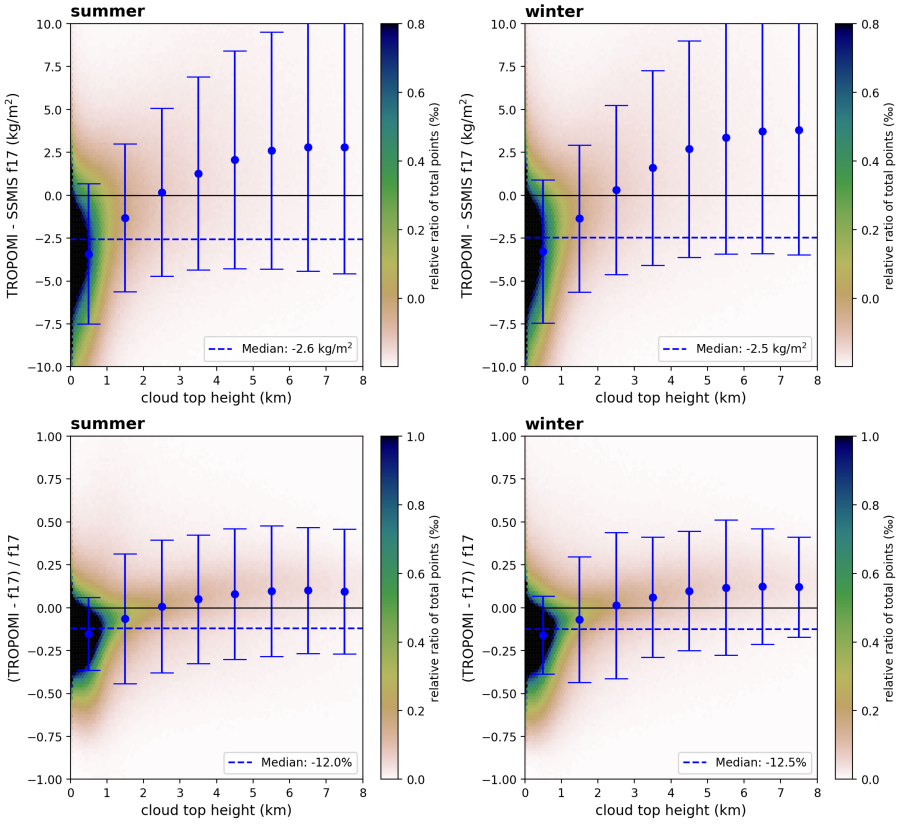


Figure 21. Same as Fig. 20, but for SSMIS f17.

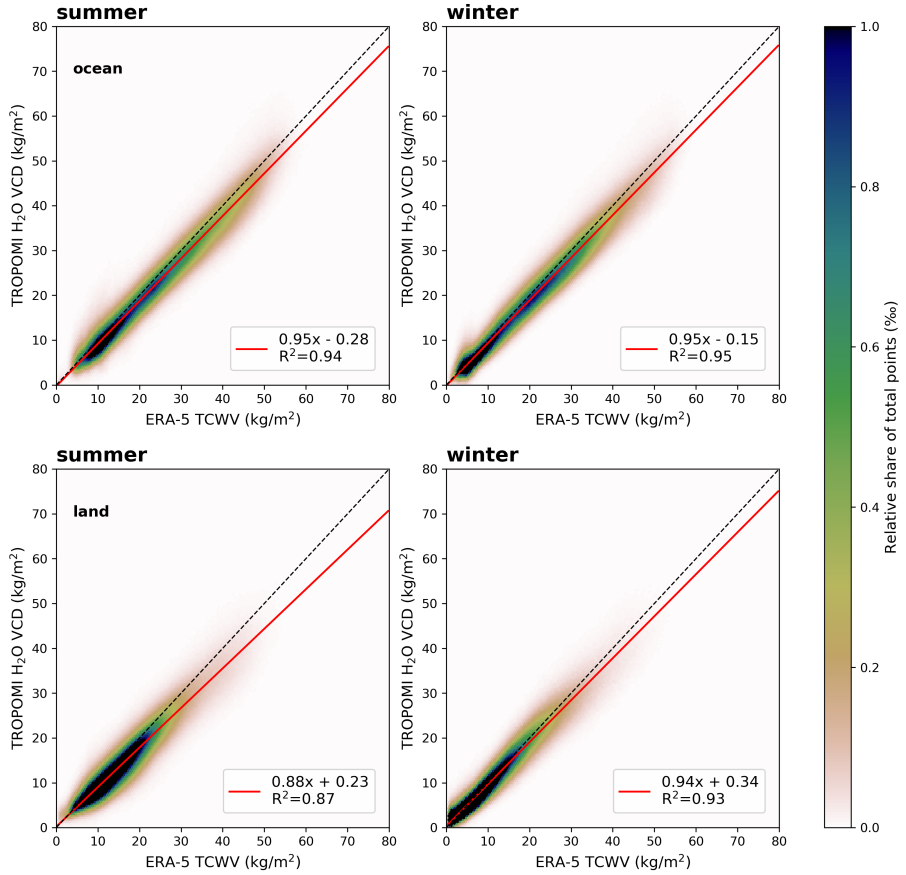


Figure 22. 2D histograms for the comparison between TROPOMI and ERA-5 for data over ocean (top row) and over land (bottom row) for clear-sky conditions (CF<20%) for boreal summer (left column) and boreal winter (right column) where the color indicates the relative share of total data points. The black dotted line indicates the 1-to-1 diagonal and the red solid line represents the results of the linear regression. The parameters of the linear regression and the coefficient of determination are given in the box in each panel.

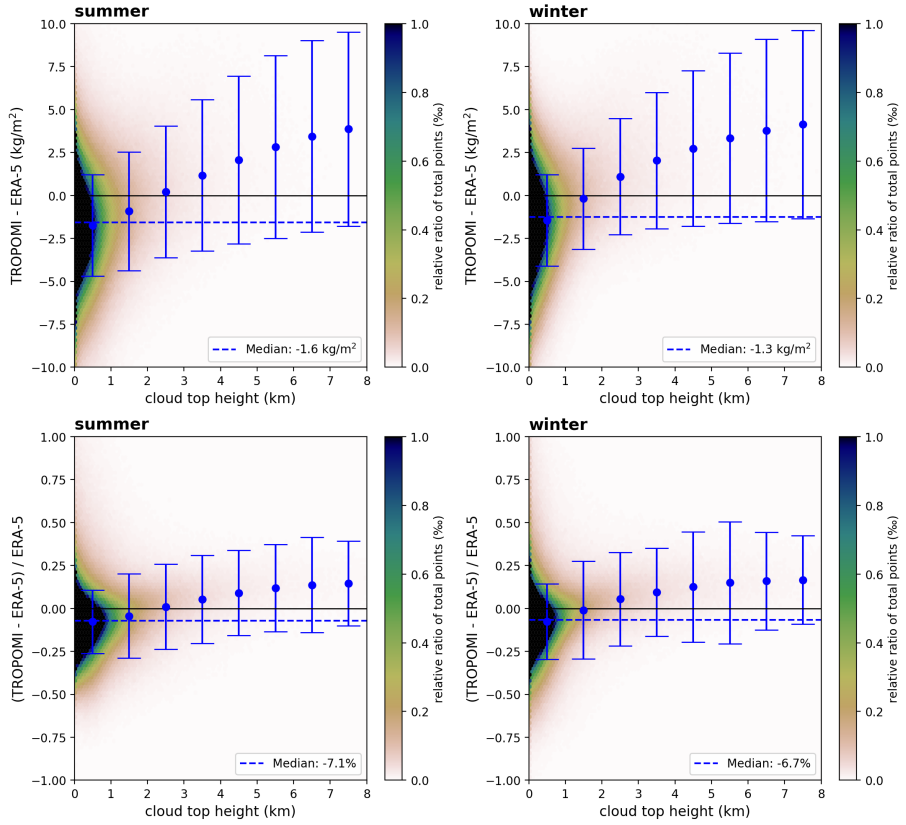


Figure 23. 2D histograms of the difference (TROPOMI-ERA-5, top row) and relative difference (TROPOMI-ERA-5)/ERA-5 (bottom row) as a function of the input cloud top height (CTH) for clear-sky conditions (CF<20%) for summer (left column) and winter (right column) for data over ocean. The blue dashed line represents the median over the whole CTH range. The blue dots represent the median within a 1 km CTH and the error bars represent their respective 1σ standard deviation.

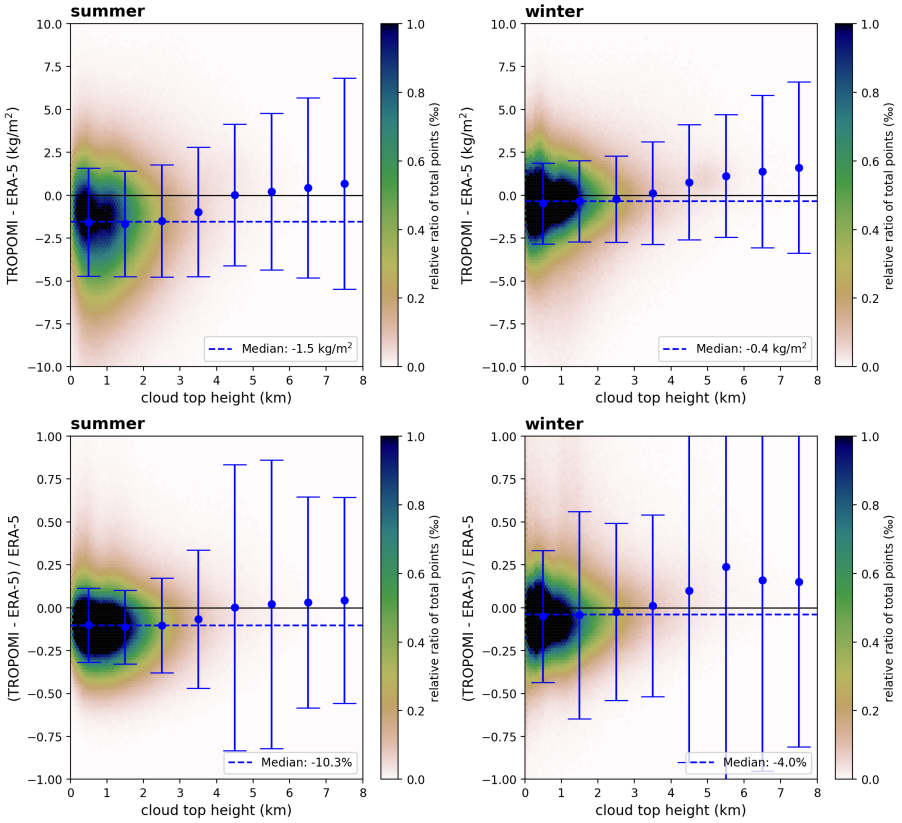


Figure 24. Same as Fig. 23, but for data over land.

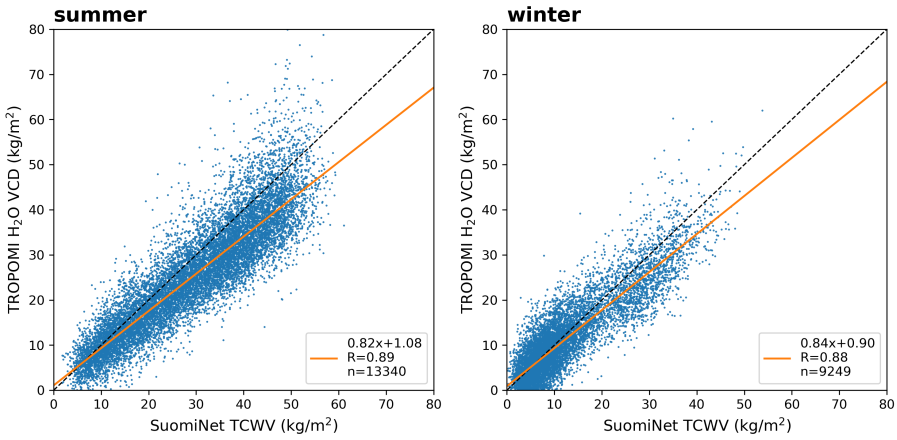


Figure 25. Scatter plots of the intercomparisons between TROPOMI and SuomiNet for clear-sky conditions (CF<20%) for boreal summer (left panel) and boreal winter (right panel). The black dashed line indicates the 1-to-1 diagonal and the orange solid line represents the results of the robust regression based on Siegel (1982). The parameters of the regression and the coefficient of correlation are given in the box in each panel.

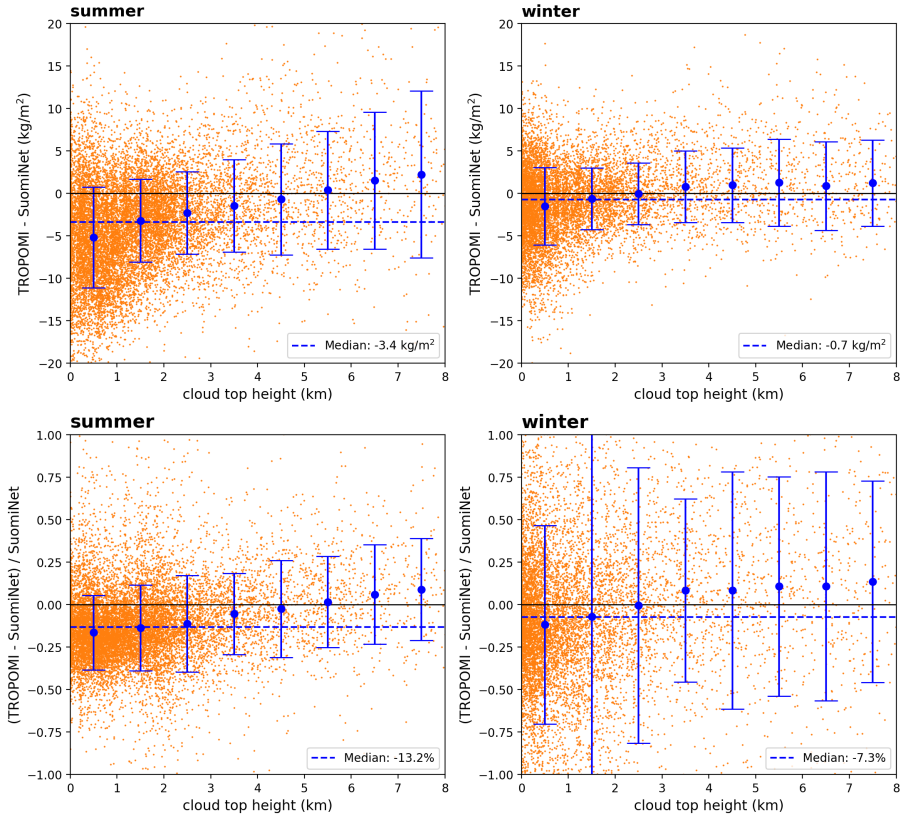


Figure 26. Scatter plot of the difference (TROPOMI-SuomiNet,top row) and relative difference (TROPOMI-SuomiNet)/SuomiNet (bottom row) as a function of the input cloud top height (CTH) for clear-sky conditions (CF<20%) for summer (left column) and winter (right column). The blue dashed line represents the median over the whole CTH range. The blue dots represent the median within a 1 km CTH and the error bars represent their respective 1σ standard deviation.

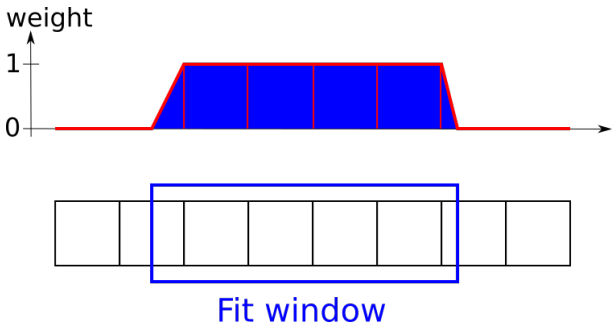


Figure A1. Schematic illustration of the weights used during the retrieval's spectral analysis.

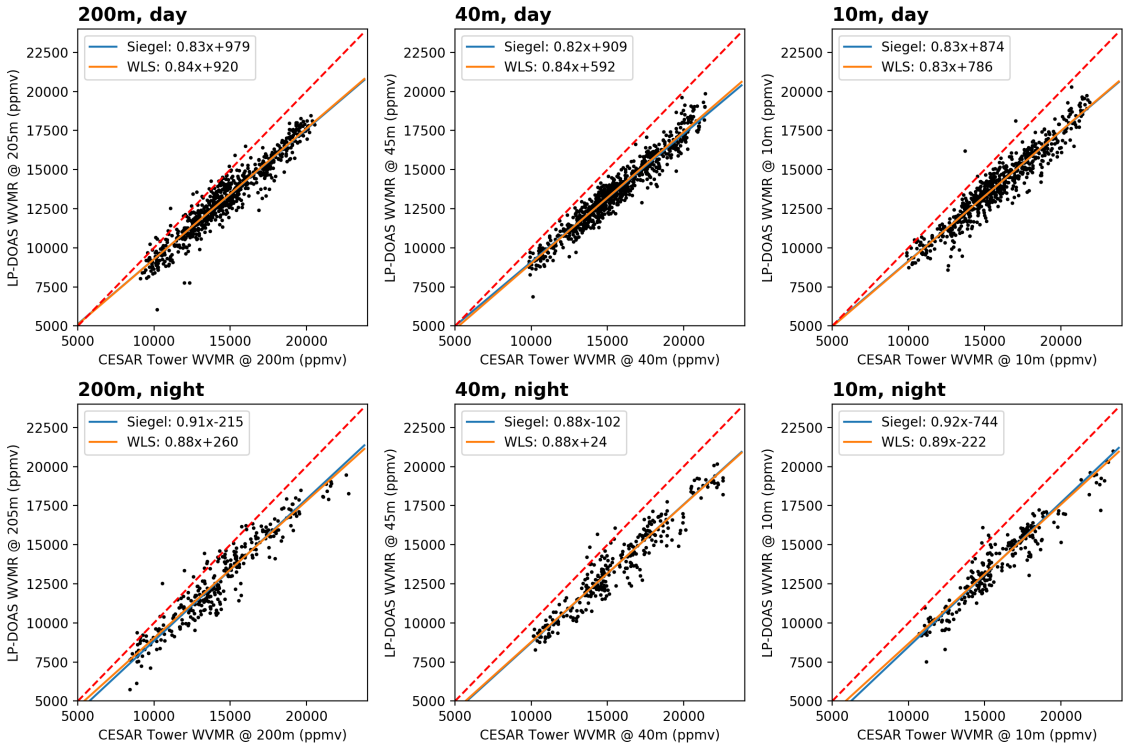


Figure A2. Scatter plots of water vapour volume mixing ratios (WVMR) derived from LP-DOAS measurements and meteorological measurements at different altitudes (10 m, 40 m and 200 m) at the CESAR Tower for day and night during the CINDI-2 campaign. Water vapour absorption cross sections have been calculated from HITRAN 2012 line list. The dashed red line represents the 1-to-1 diagonal, the solid blue line the results from the robust regression (Siegel, 1982), and the solid orange line the results from the weighted linear regression.

~~Mean water vapour profile shapes for data in 2013 for latitude bins of 10°. The solid lines represent the results from the exponential scale height approaches (blue: non-linear fit; orange: sum method) and the black dots represent the COSMIC measurements.~~

~~Bias of the profile shapes with respect to COSMIC profile shapes for the same data as in Fig. S4 (blue: non-linear fit; orange: sum method). The dashed black line represents the zero bias line.~~

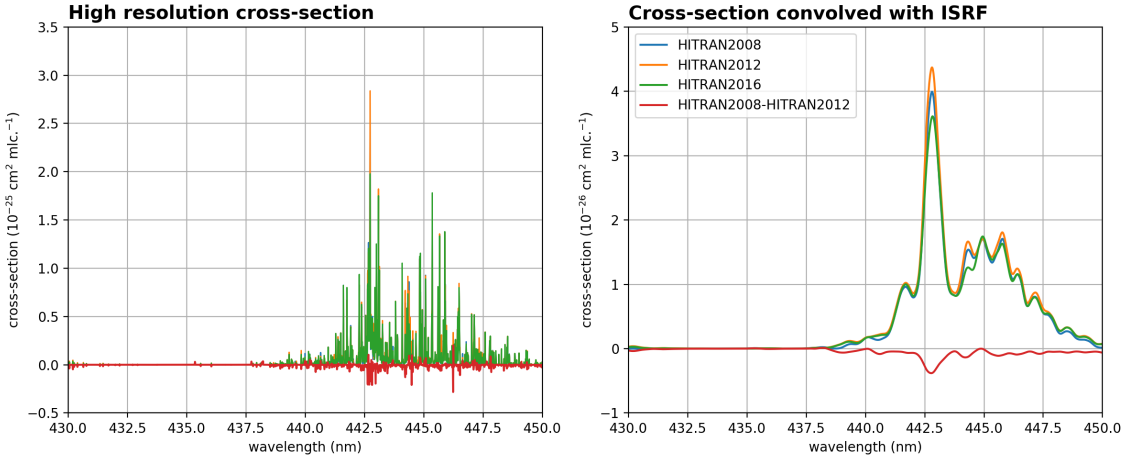


Figure A3. ~~Mean absolute error~~ Comparison of ~~exponential profile shapes with respect to COSMIC profiles~~ water vapour absorption cross section derived from different HITRAN versions (2008, 2012, and 2016) for a temperature of 296 K. The left panel depicts the ~~same data as in Fig~~high resolution cross-sections and the difference between HITRAN2008 and HITRAN2012. ~~S4~~ The right panel depicts the same cross-sections, but convolved with a typical TROPOMI Super-Gaussian ISRF (~~blue: non-linear fit; orange: sum method~~values from Beirle et al. (2017)).

~~Standard deviations of the exponential and measured COSMIC profile shapes for the same data as in Figure S4. The blue solid line represents the results for the non-linear fit and the orange line for the sum method. The dashed black line represents the results for the COSMIC measurements.~~

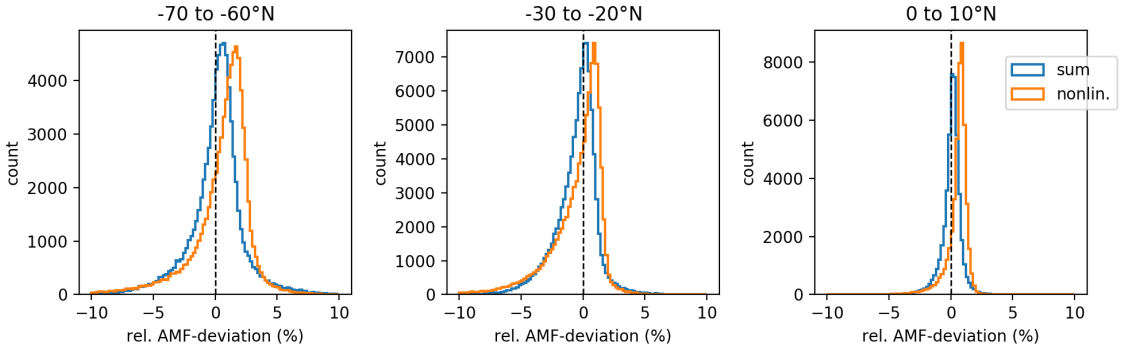


Figure A4. Histogram of the relative deviation of the calculated synthetic AMFs between sum method (blue)/non-linear fit (orange) and COSMIC profile for selected latitude bins (0° to 10°N, -30° to -20°N, and -70° to -60°N) assuming clear sky conditions and nadir viewing geometry.

~~Histogram of the relative deviation of the calculated synthetic AMFs between sum method (blue)/non-linear fit (orange) and COSMIC profile for latitude bin from -30°N to -20°N assuming different cloudy-sky conditions (cloud fraction 10, 20 and 50 % (left to right); cloud top height 1 km, 2 km and 5 km (top to bottom)) and nadir viewing geometry.~~

~~Monthly/seasonal dependence of the fit parameters (a_0, a_1, a_2) and $(b_0, b_1, b_2, \theta_0)$ for the functions $\alpha(\theta, t)$ and $\beta(\theta, t)$ in Eq.(11).~~

~~2D histograms for the comparison between TROPOMI and SSMIS f16 for July 2018 for different cloud fraction bins (left to right column) and cloud top height bins (top to bottom row). The color indicates the amount of points within one bin of each panel. The black dotted line indicates the 1-to-1 diagonal and the red solid line represents the results of the linear regression. The parameters of the linear regression and the coefficient of determination are given in the box in each panel.~~

~~Same as Fig. S10, but for SSMIS f17.~~

~~Same as Fig. S10, but for ERA-5 TCWV data over ocean.~~

~~Same as Fig. S10, but for ERA-5 TCWV data over land.~~

~~Scatterplots for the comparison between TROPOMI and SuomiNet for boreal summer 2018 for different cloud fraction bins (left to right column) and cloud top height bins (top to bottom row). The black dashed line indicates the 1-to-1 diagonal and the orange solid line represents the results of the robust regression. The parameters of the regression and the correlation coefficient are given in the box in each panel.~~

Table 1. DOAS fit settings for the H_2O slant column retrievals

Parameter	Description
Fit window	430 - 450 nm
Absorption cross sections	Water vapour, 296 K (Rothman et al., 2009)
	NO_2, 220 K (Vandaele et al., 1998)
	O_3, 243 K (Serdyuchenko et al., 2014)
	O_4, 293 K (Thalman and Volkamer, 2013)
Ring effect	2 Ring spectra calculated from daily irradiance
Polynomial	5th order
Pseudo-absorbers	intensity offset (inverse spectrum)
	shift & stretch (Beirle et al., 2013)
	ISRF parameter changes (Beirle et al., 2017)

Table 2. Parameter list and nodes for the BAMF profile simulations

Parameter	Nodes
Wavelength (nm)	442
Sensor altitude (km)	720
Surface altitude (km)	0.0, 0.5, 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 6.0, 8.0, 12.0
Surface albedo	0.0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.12, 0.14, 0.16, 0.18, 0.20, 0.25, 0.30, 0.40, 0.50, 0.80, 1.0
Solar zenith angle (°)	0, 10, 20, 30, 40, 50, 60, 65, 70, 80, 85, 87, 88
Line of sight angle (°)	-90, -86, -82, -78, -74, -70, -66, -62, -58, -54, -50, -46, -42, -38
Solar relative azimuth angle (°)	0, 20, 40, 60, 80, 100, 120, 140, 160, 180

Table 3. Fit results of the robust regression between ratio of scale heights H_{land}/H_{sum} and the NDVI for different filtered data sets.

Data set	slope	intercept
All data	−0.47	1.35
No landcover type 15	−0.38	1.28
No landcover type 7 & 15	−0.33	1.25

Table 4. Summary of the different error sources considered in the H_2O SCD uncertainty.

Source	Type	Parameter uncertainty	Estimated uncertainty in SCD
Absorption cross section	Systematic	10%	10%
DOAS fit error	Random	-	SZA<20°: 0.15×10^{23} molec cm^{-2} (~10%)
	-	-	SZA>70°: 0.30×10^{23} molec cm^{-2} (~30%)

Table 5. Standard retrieval scenarios for the estimation of AMF error.

Parameter	Values
Surface albedo	2%, 7%, 20%
Solar zenith angle	0°, 45°, 90°
Water vapour scale height	1 km, 2 km, 3 km
Cloud fraction	10%, 20%, 50%
Cloud top height	1 km, 2 km, 5 km
Line of sight angle	-90°

Table 6. Summary of different error sources considered in the AMF uncertainty.

Parameter	Type	Parameter uncertainty	Source	Estimated uncertainty in AMF	
				clear-sky	cloudy-sky
Surface albedo	Random+Systematic	0.02	Kleipool et al. (2008)	5-25%	5-10%
Scale height (ocean)	Random	0.45 km	-	2-10%	5-20%
Scale height (NDVI, low)	Random	0.73 km	-	5-25%	20-50%
Scale height (NDVI, high)	Random	0.34 km	-	2-7%	5-15%
Cloud fraction	Random+Systematic	0.05	Veefkind et al. (2016)	-	2-10%
Cloud top height	Random+Systematic	0.5 km	-	-	5-15%

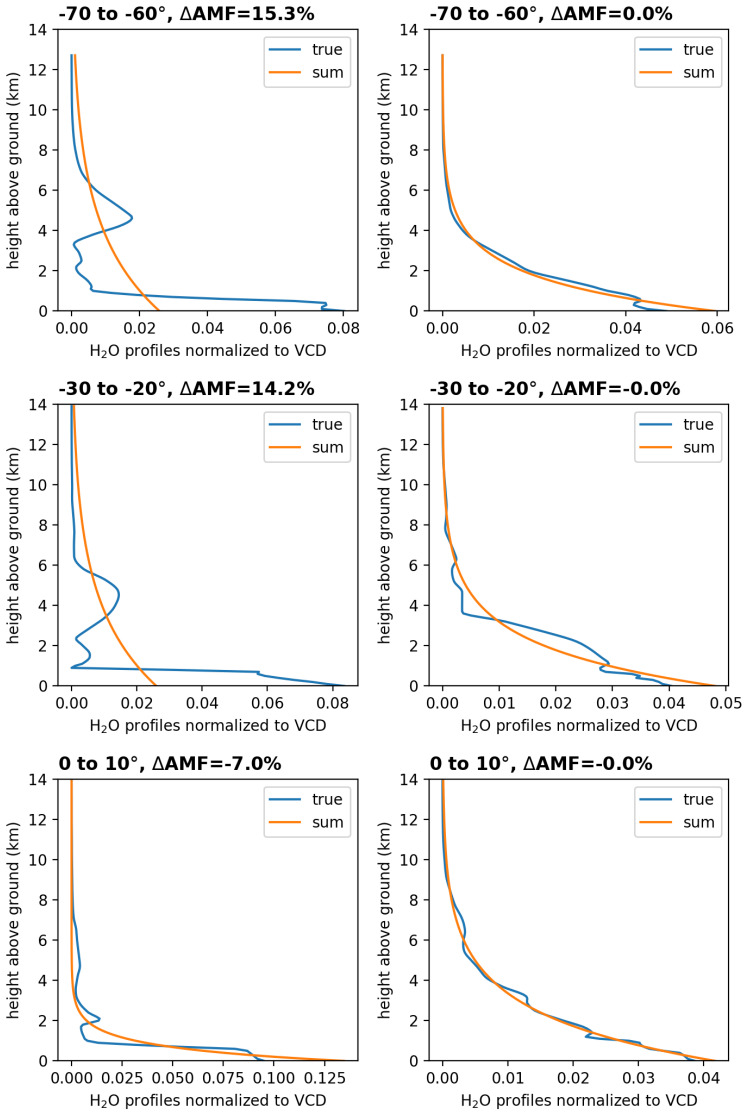


Figure S1. Comparison of profile shapes for selected latitude bins illustrating the maximal (left column) and minimal (right column) absolute relative AMF deviations. The blue line represents the "true" water vapour profile shape as measured from COSMIC and the orange line represents the exponential profile with a scale height H calculated from the sum method.

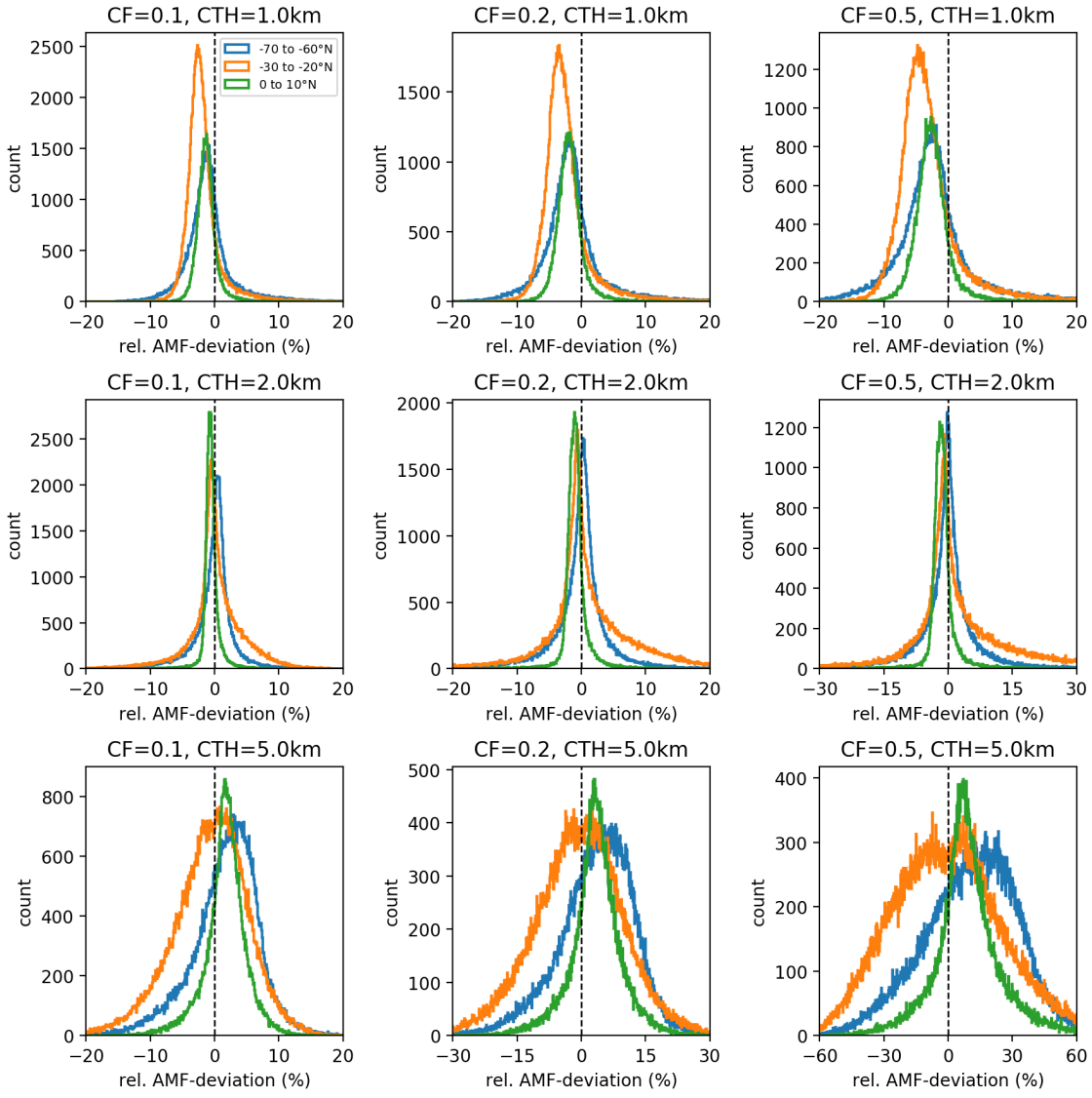


Figure S2. Histograms of the relative AMF deviations between exponential profile and COSMIC profile for the same latitude bins as in Fig. 5 for different cloud scenarios (cloud fraction 10, 20 and 50 % (left to right); cloud top height 1 km, 2 km and 5 km (top to bottom)) and a nadir viewing geometry.

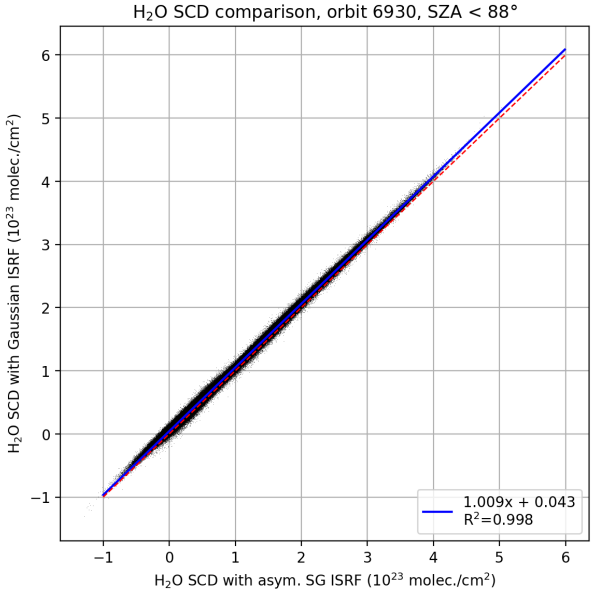


Figure S3. Comparison of H_2O SCDs fitted assuming a Gaussian and an asymmetric Super-Gaussian ISRF. Values are taken from orbit 6930 for a solar zenith angle < $88°$. The blue solid line indicates the results of the linear regression and the red dashed line the 1-to-1 diagonal.

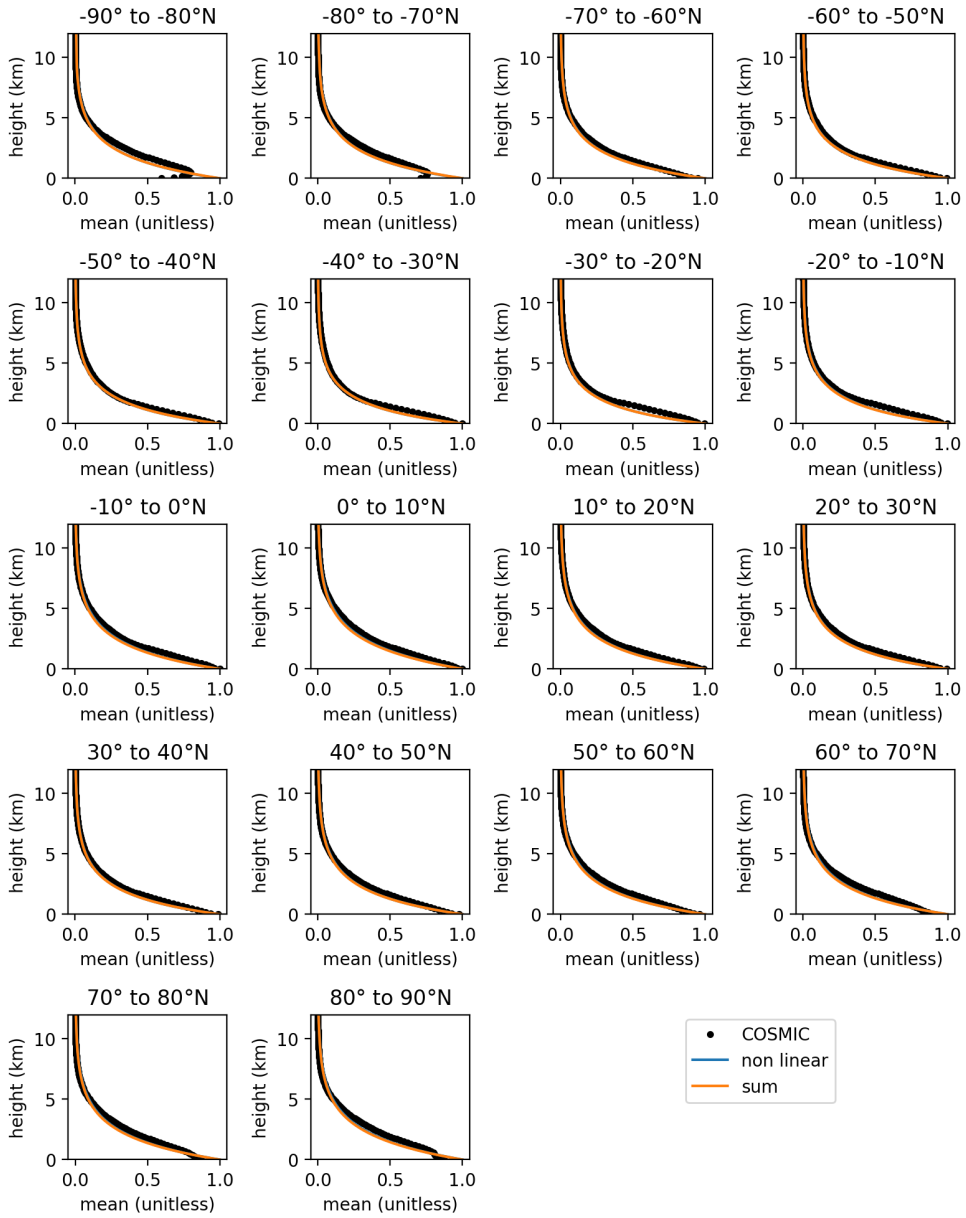


Figure S4. Mean water vapour profile shapes for data in 2013 for latitude bins of 10°. The solid lines represent the results from the exponential scale height approaches (blue: non-linear fit; orange: sum method) and the black dots represent the COSMIC measurements.

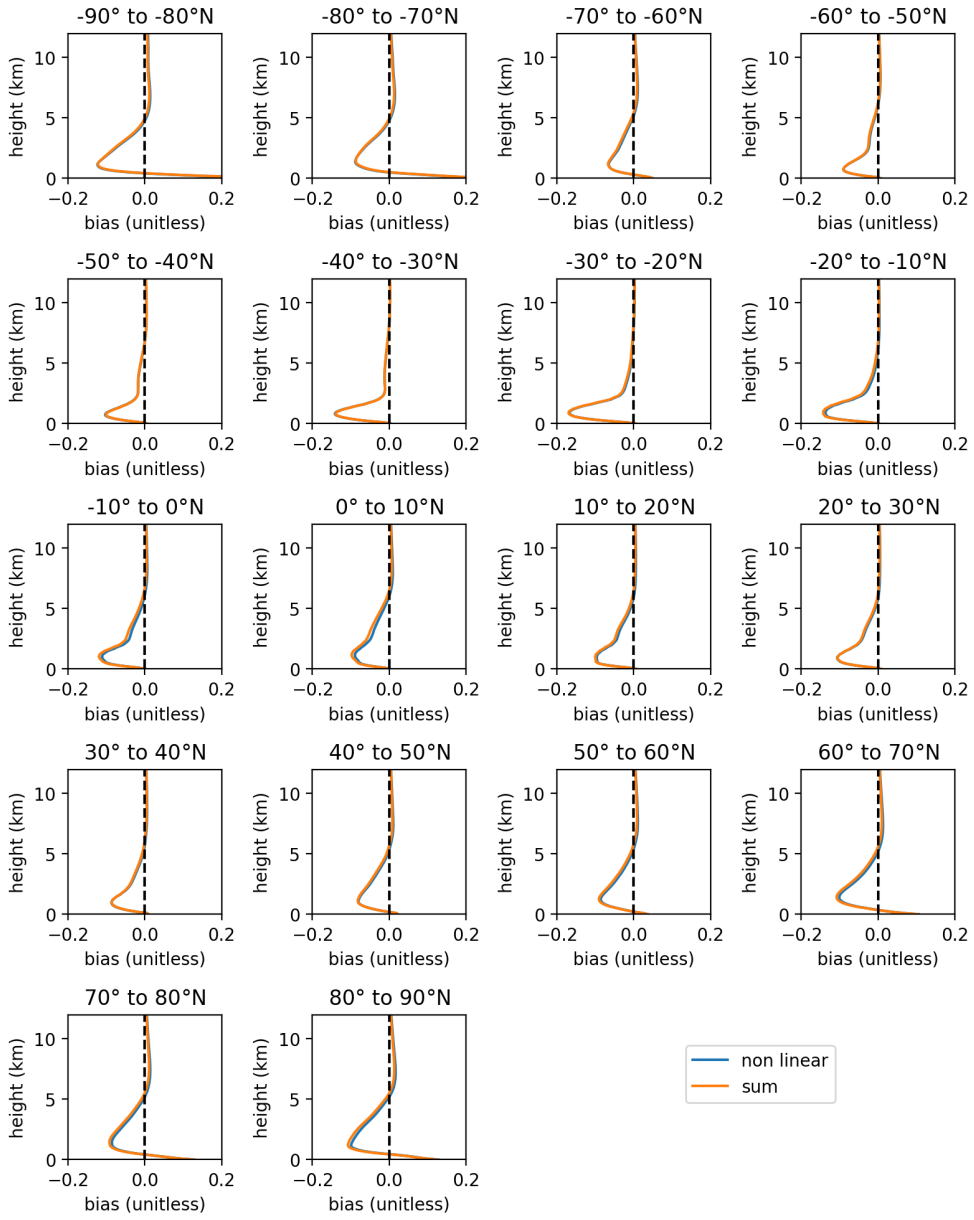


Figure S5. Bias of the profile shapes with respect to COSMIC profile shapes for the same data as in Fig. S4 (blue: non-linear fit; orange: sum method). The dashed black line represents the zero bias line.

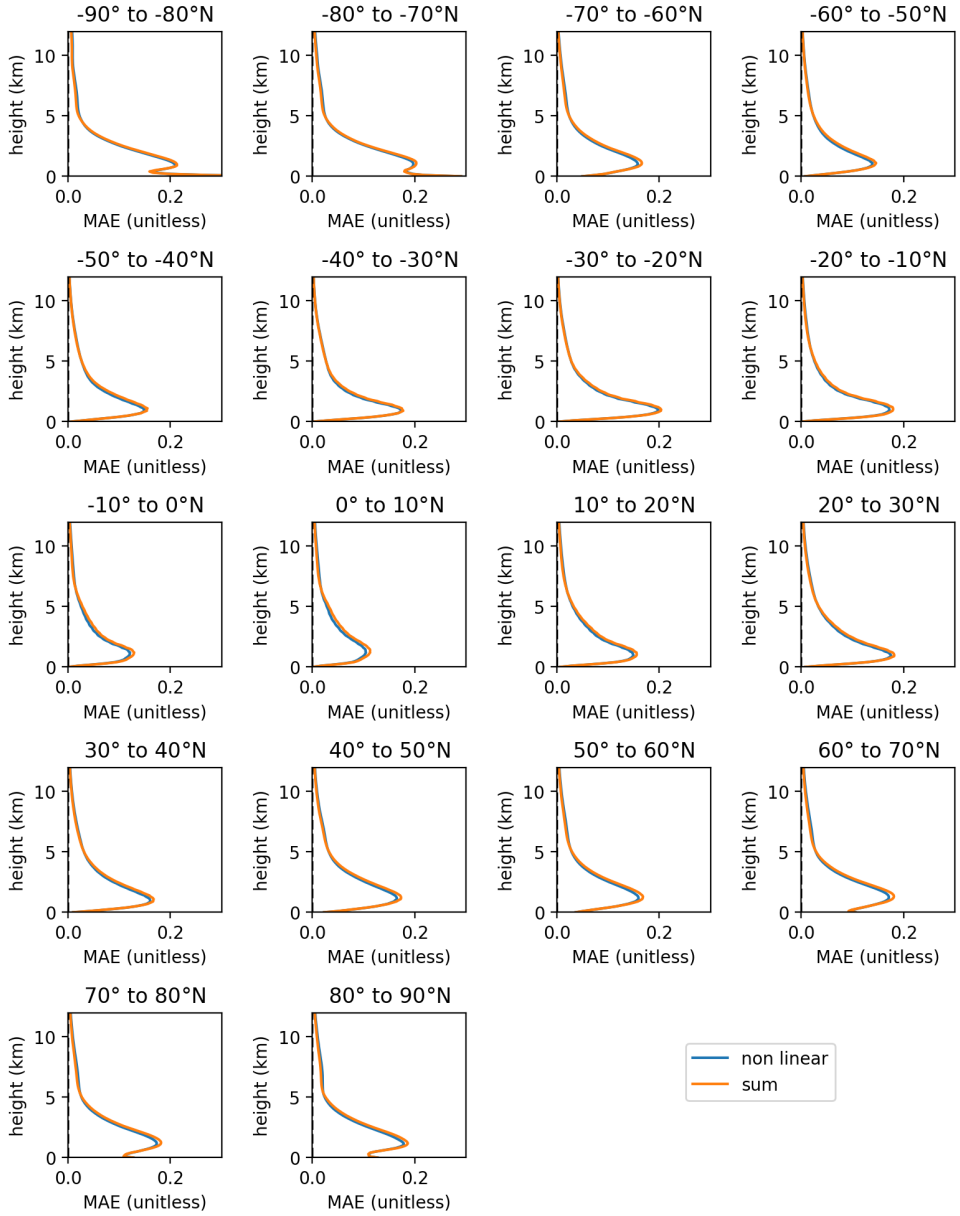


Figure S6. Mean absolute error of exponential profile shapes with respect to COSMIC profiles for the same data as in Fig. S4 (blue: non-linear fit; orange: sum method).

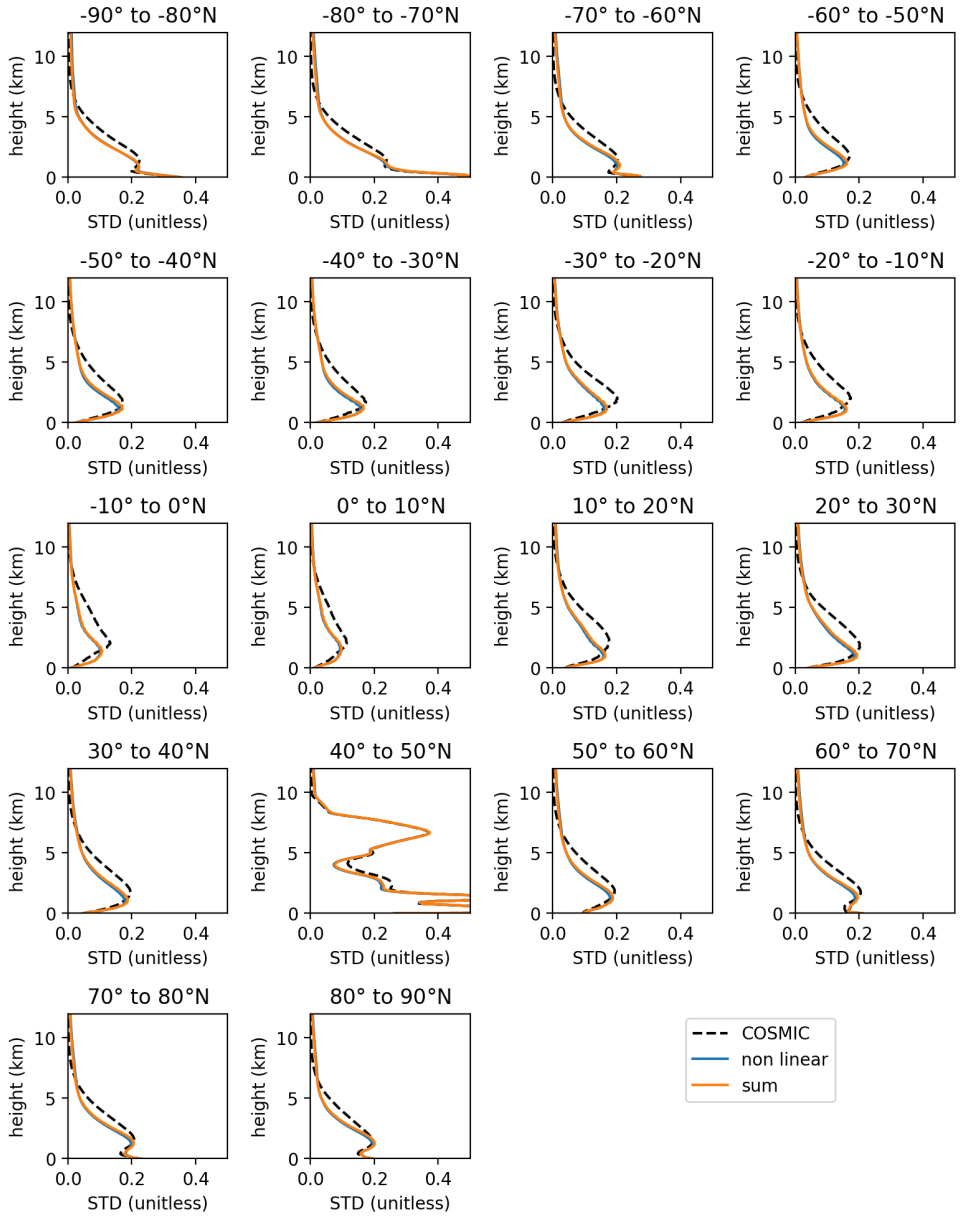


Figure S7. Standard deviations of the exponential and measured COSMIC profile shapes for the same data as in Figure S4. The blue solid line represents the results for the non-linear fit and the orange line for the sum method. The dashed black line represents the results for the COSMIC measurements.

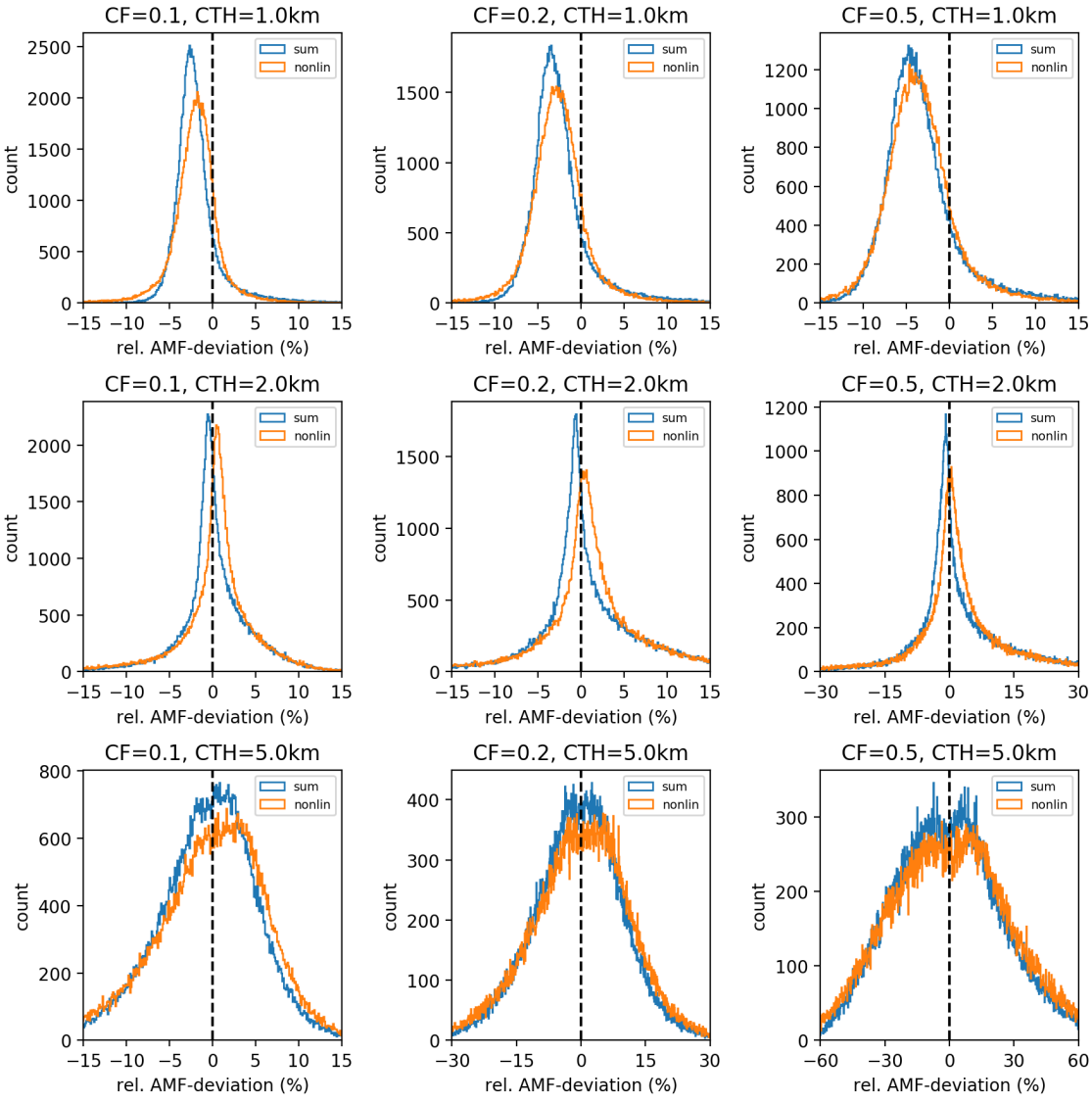


Figure S8. Histogram of the relative deviation of the calculated synthetic AMFs between sum method (blue)/non-linear fit (orange) and COSMIC profile for latitude bin from -30°N to -20°N assuming different cloudy-sky conditions (cloud fraction 10, 20 and 50% (left to right); cloud top height 1 km, 2 km and 5 km (top to bottom)) and nadir viewing geometry.

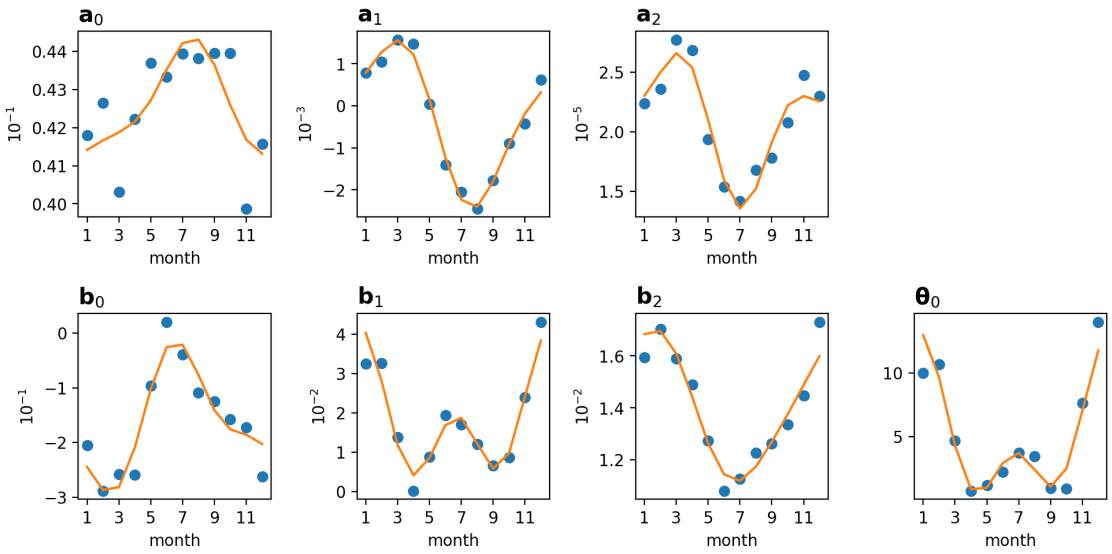


Figure S9. Monthly/seasonal dependence of the fit parameters (a_0, a_1, a_2) and $(b_0, b_1, b_2, \theta_0)$ for the functions $\alpha(\theta, t)$ and $\beta(\theta, t)$ in Eq. (11).

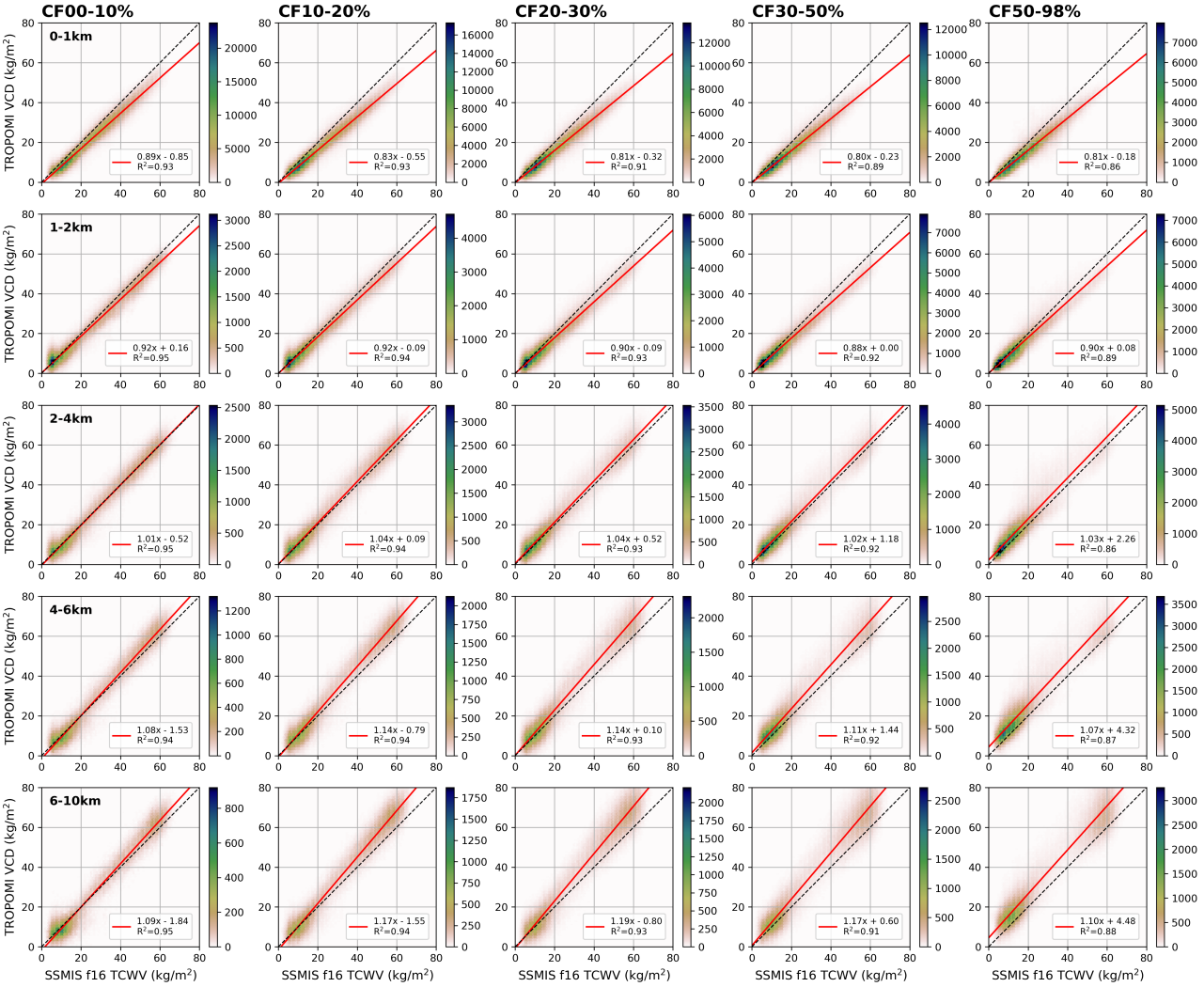


Figure S10. 2D histograms for the comparison between TROPOMI and SSMIS f16 for July 2018 for different cloud fraction bins (left to right column) and cloud top height bins (top to bottom row). The color indicates the amount of points within one bin of each panel. The black dotted line indicates the 1-to-1 diagonal and the red solid line represents the results of the linear regression. The parameters of the linear regression and the coefficient of determination are given in the box in each panel.

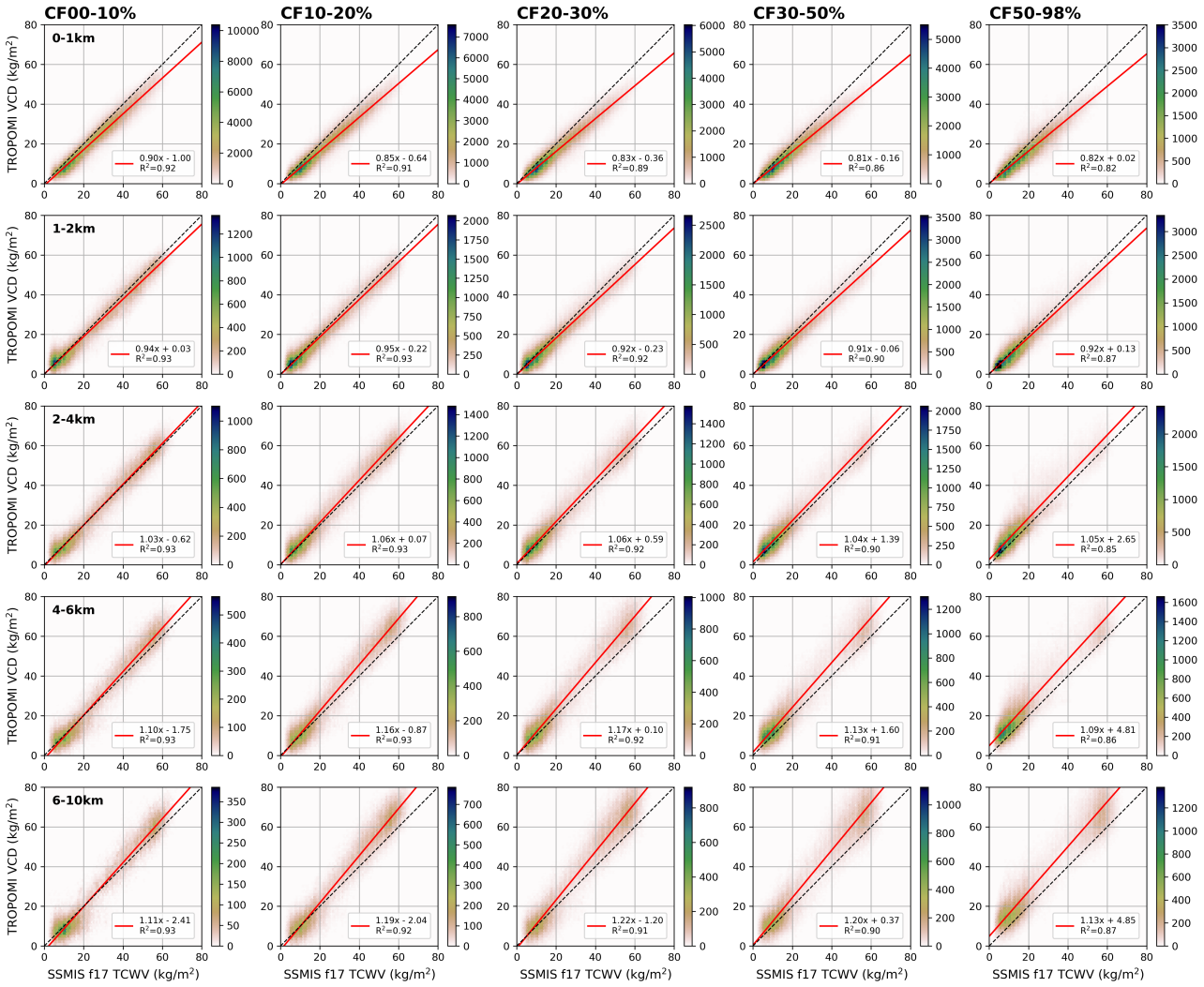


Figure S11. Same as Fig. S10, but for SSMIS f17.

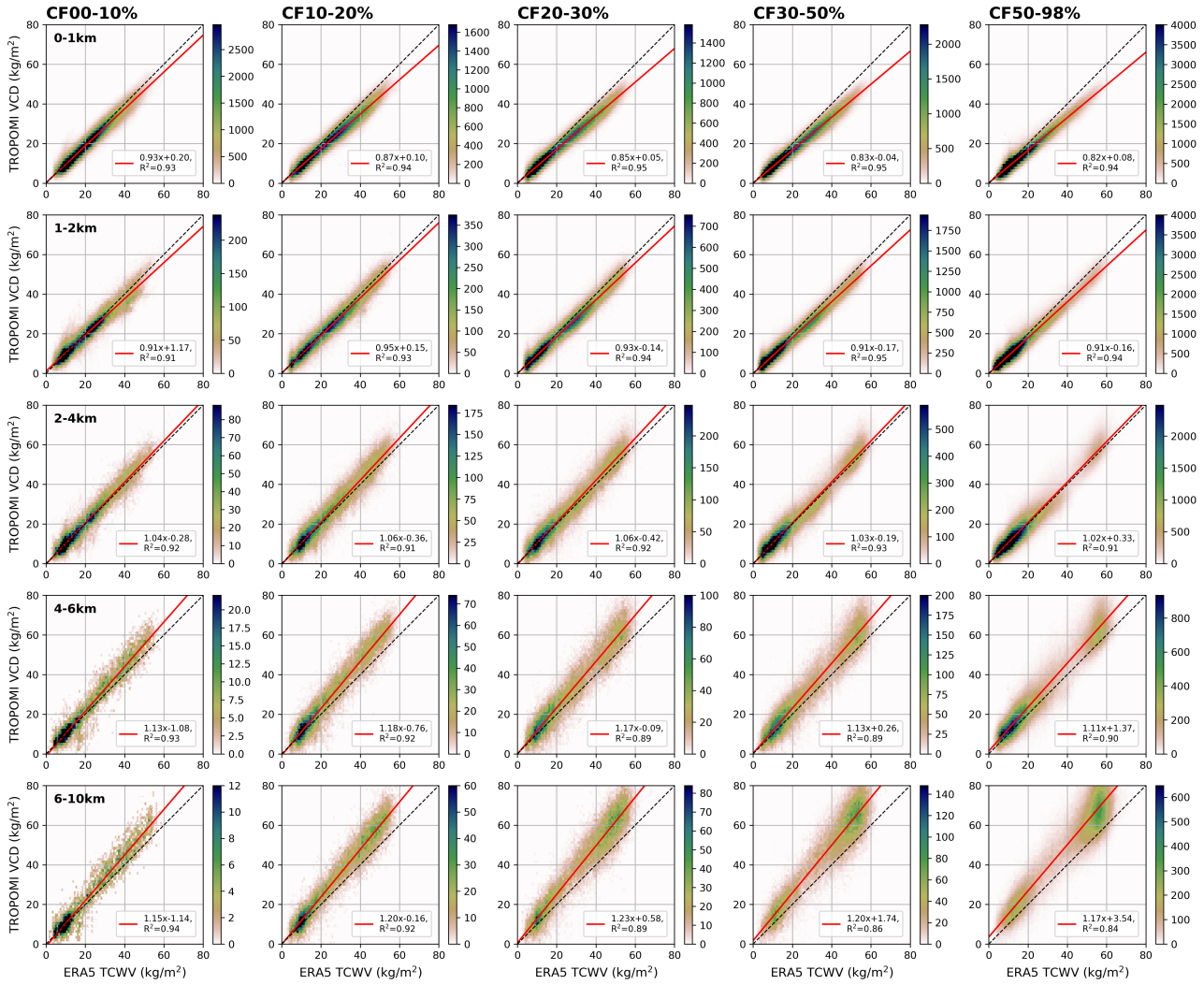


Figure S12. Same as Fig. S10, but for ERA-5 TCWV data over ocean.

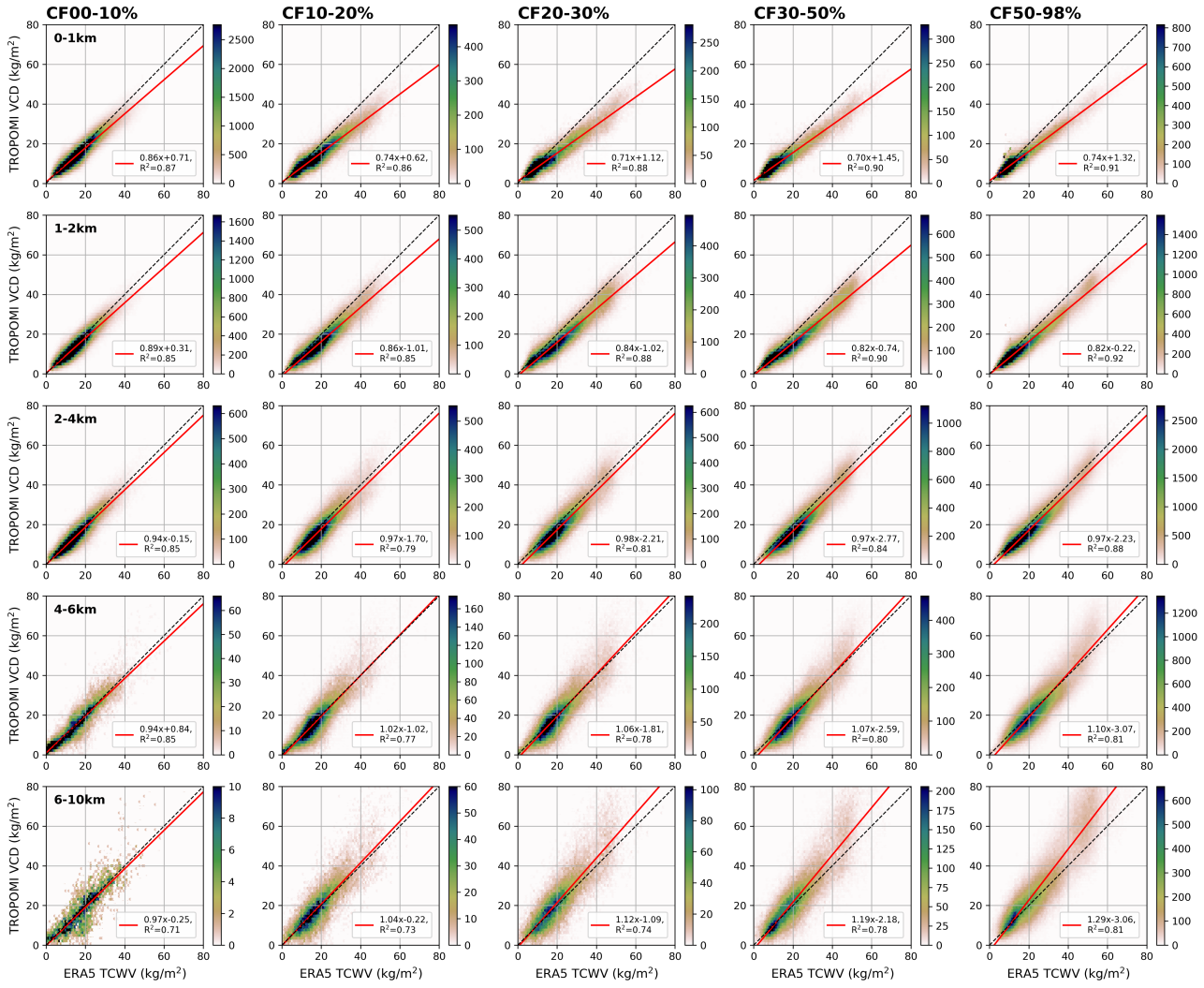


Figure S13. Same as Fig. S10, but for ERA-5 TCWV data over land.

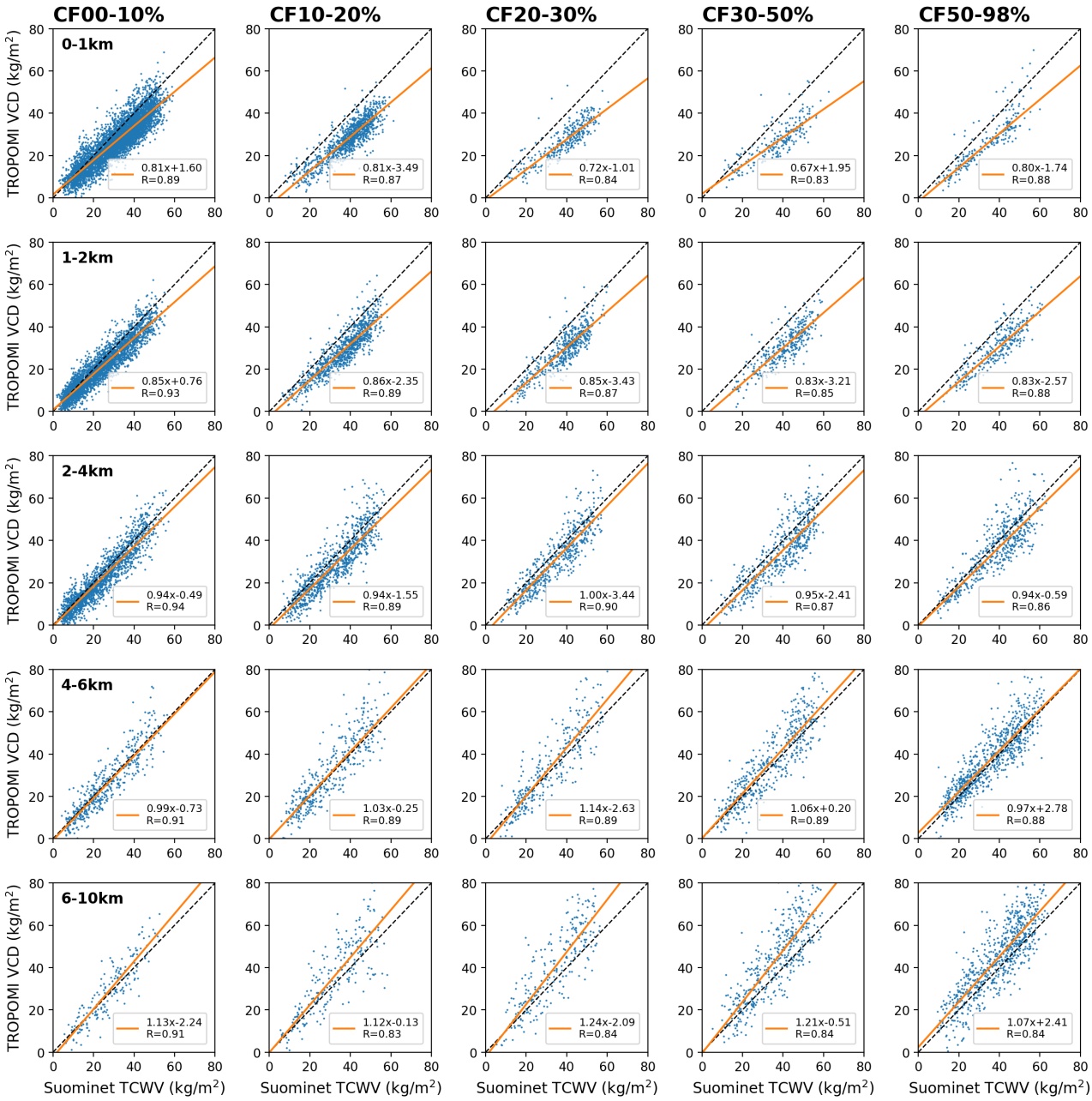


Figure S14. Scatterplots for the comparison between TROPOMI and SuomiNet for boreal summer 2018 for different cloud fraction bins (left to right column) and cloud top height bins (top to bottom row). The black dashed line indicates the 1-to-1 diagonal and the orange solid line represents the results of the robust regression. The parameters of the regression and the correlation coefficient are given in the box in each panel.