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:

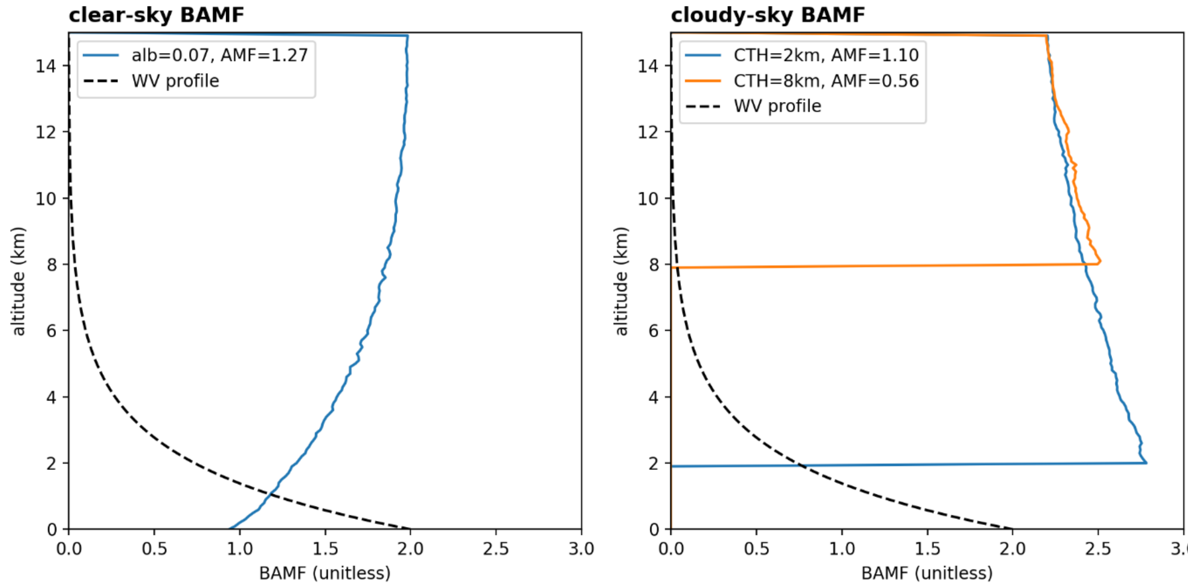
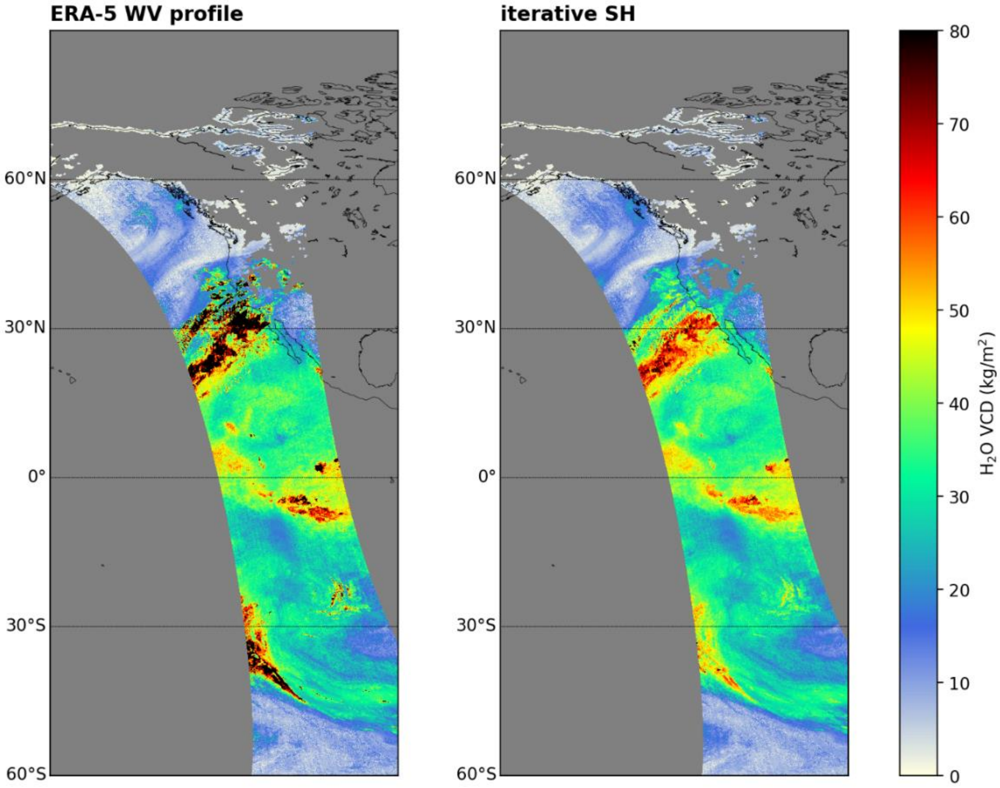


Figure 3 depicts typical examples of BAMF profiles for different clear- and cloudy-sky scenarios. The AMFs for the cloudy-sky scenarios were calculated assuming a surface albedo of 7% and an effective cloud fraction of 20%. For the clear-sky scenario (left panel) the sensitivity decreases towards the 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.

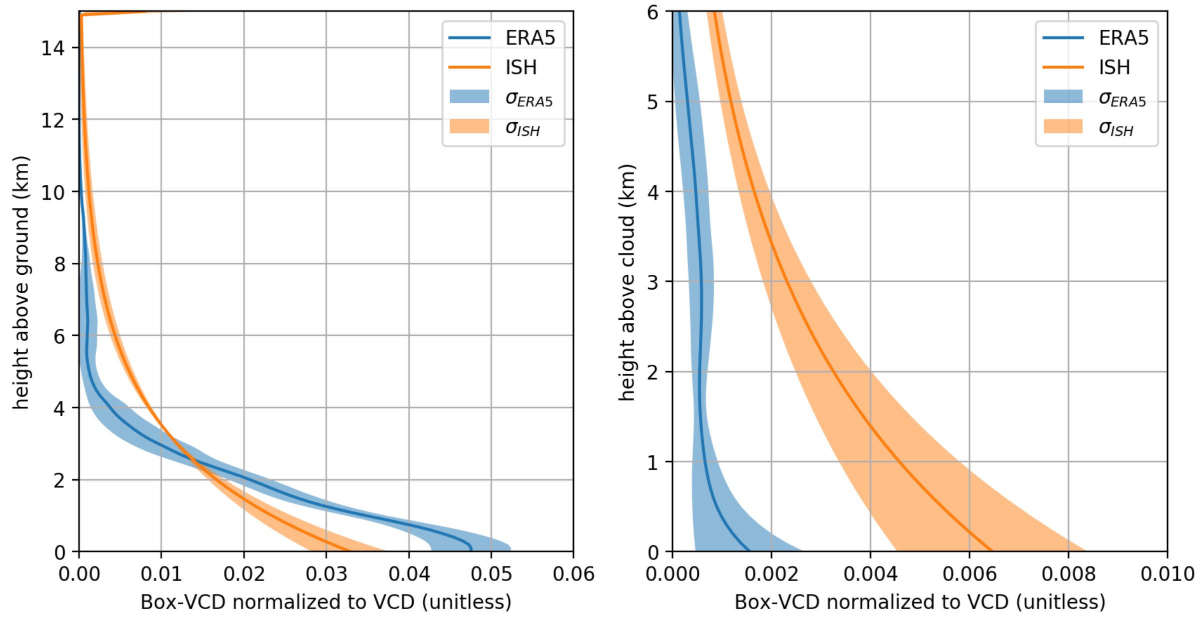
The simplest solution to use a “first-guess” water vapour profile from NWP re-analysis data is simply excluded with a reference to Wang. Then GPS climatologies are used to derive a scale-height parametrization. However, it is not made explicit or clear that the rational to use a scale-height (instead 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.