We would like to thank the referee for reviewing our manuscript and for the many useful comments and suggestions. 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.

A 16-year global climate data record of total column water vapour generated from OMI observations in the visible blue spectral range
Christian Borger, Steffen Beirle, and Thomas Wagner

1. General Comments

This paper assesses a long term record (2005-2020) of monthly mean total column water vapour (TCWV) from the Ozone Monitoring Instrument (OMI) on board NASAs Aura platform. The authors describe adaptations to an existing algorithm used for ESAs TROPOMI instrument, which is seen as a successor to OMI. This includes the rational for the how and why they switch to using earth shine spectra as the reference in the DOAS retrieval setup. This study goes on to present results from an inter-comparison of TCWV against two addition remote sensing products from RSS and ESA and ERA5 reanalysis.

While this study does discuss issues to do with sampling, I feel this could be expanded especially relating to the clear-sky bias. Work on this has been done within the ESA water vapour CCI project, results from which are relevant to this study and would enhance the discussion around the OMI product performance.

Many thanks for this important comment! We agree with the reviewer and have decided to estimate the clear-sky bias within our data set. Moreover, following the suggestions from Reviewer #1, we also estimate sampling errors. As such, we have added the following text to the revised version:

Although satellite observations enable the analysis of trace gas concentrations on global scale, a fundamental problem is that typically a satellite measurement is only taken once a day for one location. Furthermore, satellite measurements are usually only available under cloud-free conditions, especially in the visible or infrared spectral range and thus no continuous time series is guaranteed. Consequently, they cannot provide a complete picture of geophysical variability, which leads to sampling errors in the calculation of averaged values (e.g. monthly means).

Moreover, the question arises to what extent the limitation to cloud-free pixels influences the monthly averages determined from the OMI satellite measurements, i.e. whether in the OMI TCWV data set a so-called "clear-sky bias" exists. Gaffen and Elliott (1993) investigated this bias using radiosonde ascents and found that the TCWV is about 0-15% lower under cloud-free conditions than under cloudy conditions. Similarly, Sohn and Bennartz (2008) found a clear-sky bias between MERIS and AMSR-E of about 10%.

To estimate the sampling errors, we follow the methods of Xue et al. (2019) and Gleisner et al. (2020): we choose hourly-resolved ERA5 data with a spatial resolution of 0.25°x 0.25° as reference data and collocate the ERA5 data with OMI overpass times. These data are then resampled to the 1°x1° resolution of the OMI TCWV data set and the monthly averages are calculated (TCWV_sampled). We then take the complete, original ERA5 data, resample it to the same spatial resolution and calculate monthly means from this data as well (TCWV_true). The difference between the two data sets then represents the sampling error:

$$e_{sampling} = TCWV_{sampled} - TCWV_{true} \quad (4)$$

With this definition, the sampling error summarises the uncertainties due to gaps in the swath, temporal differences or missing data (e.g. due to clouds) (Xue et al., 2019).
Figure 3 shows the annual mean absolute and relative sampling errors for the year 2006. Overall, it can be seen that most deviations are negative, i.e. the actual TCWV is underestimated. Regarding the absolute deviations, the strongest deviations can be seen in the area of storm-tracks in the mid-latitudes (e.g. North Atlantic) and the polar regions with values around -5 kg m-2. The smallest deviations are found in the quasi-permanent cloud-free regions in the subtropics. As expected, the relative differences increase from the equator towards the poles due to the decreasing TCWV values and reach values stronger than -30%.

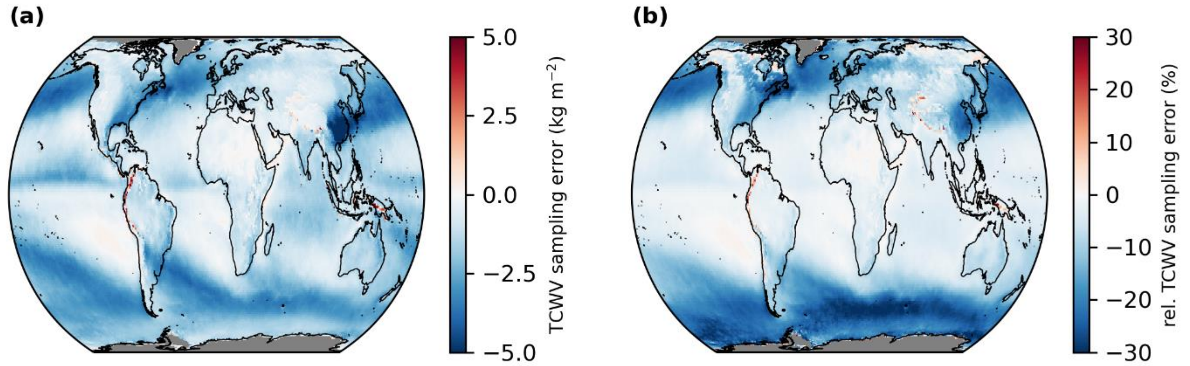

Figure 3. Global distributions of the mean sampling errors derived from monthly mean sampling differences for the time range January 2005 to December 2020. Panel (a) depicts absolute sampling error (i.e. "sampling) and Panel (b) relative sampling error (i.e. "sampling=TCWVtrue). Grid cells for which no data is available are coloured grey.

To investigate to what extent these deviations are related to the clear-sky bias, we proceed similarly to the calculation of the sampling error: we collocate the ERA5 data to the OMI overpass time and once apply a cloud filter (effective cloud fraction < 20%) and once not. Then we resample both data sets to 1°x 1° and calculate monthly means. The difference of both data sets then represents the clear-sky bias:

e_clear = TCWV_clear – TCWV_all (5)

To determine seasonal structures, the global distributions of the absolute and relative clear-sky bias for the different seasons were determined from the monthly differences (see Fig. 4). Overall, the distributions of the clear-sky bias correspond very closely to the distributions of the sampling error, both in strength and in pattern. Moreover, the absolute and relative deviations show only slight changes between the different seasons.

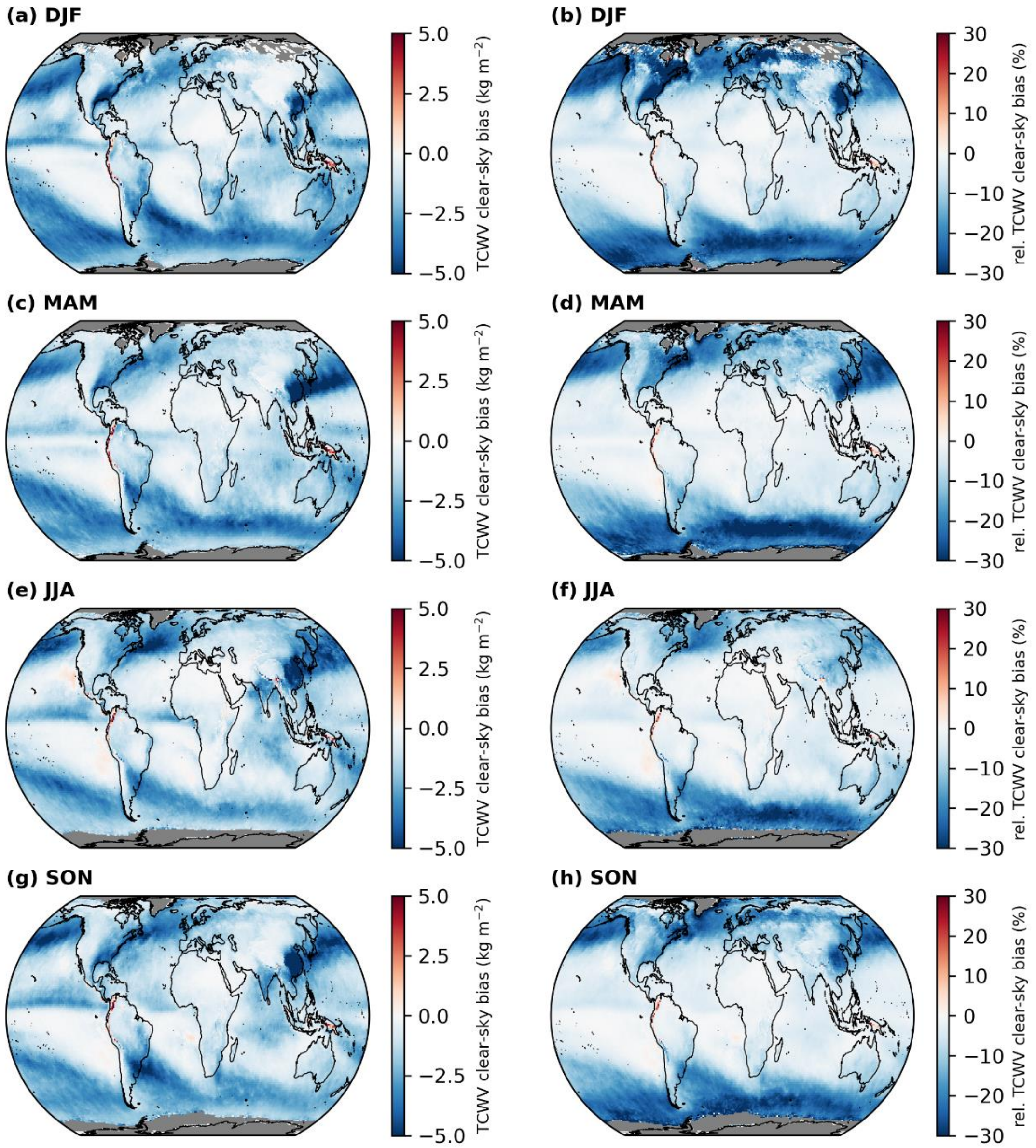

Figure 4. Global distributions of the absolute differences ("clear; left column) and relative differences ("clear=TCWVall; right column) of the monthly mean differences between clear-sky and all-sky ERA5 based on the OMI cloud information for winter (DJF; a & b), spring (MAM, c & d), summer (JJA, e & f), and autumn (SON, g & h) for the time range January 2005 to December 2020. Grid cells for which no data is available are coloured grey.

Figures 5 and 6 summarize the sampling error and clear-sky bias distributions, respectively. For the sampling error we obtain a mean absolute deviation of -1.6 kg m-2 (median -1.4 kg m-2) and a mean relative deviation of -9.5% (-6.2%) and for the clear-sky bias we get a mean absolute deviation of -1.7 kg m-2 (median -1.3 kg m-2) and a mean relative deviation of -10.0% (-6.0%). However, the distributions of the absolute and relative deviations for the sampling error and the clear-sky bias are highly left-skewed and thus the mean value in particular is influenced by the long tails of the distributions. Nevertheless, for the clear-sky bias the obtained values agree well with the findings of Gaffen and Elliott (1993) and Sohn and Bennartz (2008). Since the effect of the clear-sky bias is already included in the sampling error and the results for both errors are very similar, it can be assumed that the spatial and *temporal sampling errors play only a minor or negligible role compared to the clear-sky bias.*

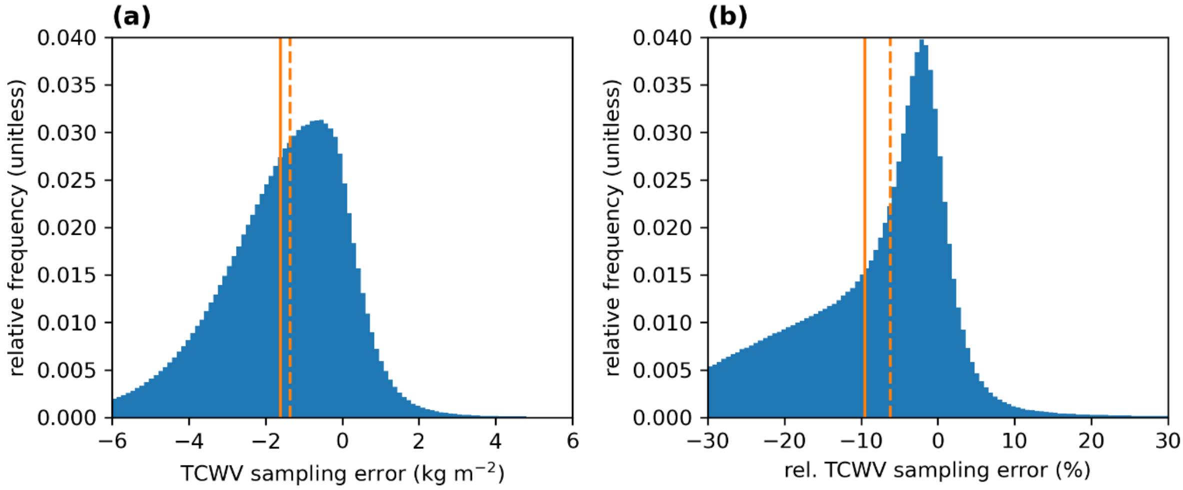

Figure 5. Distributions of the absolute differences ("sampling; Panel a) and relative differences ("sampling=TCWVtrue; Panel b) of the monthly mean differences between clear-sky and all-sky ERA5 data based on the OMI cloud information. The solid and dashed orange line indicate the mean and the median of the distributions, respectively.

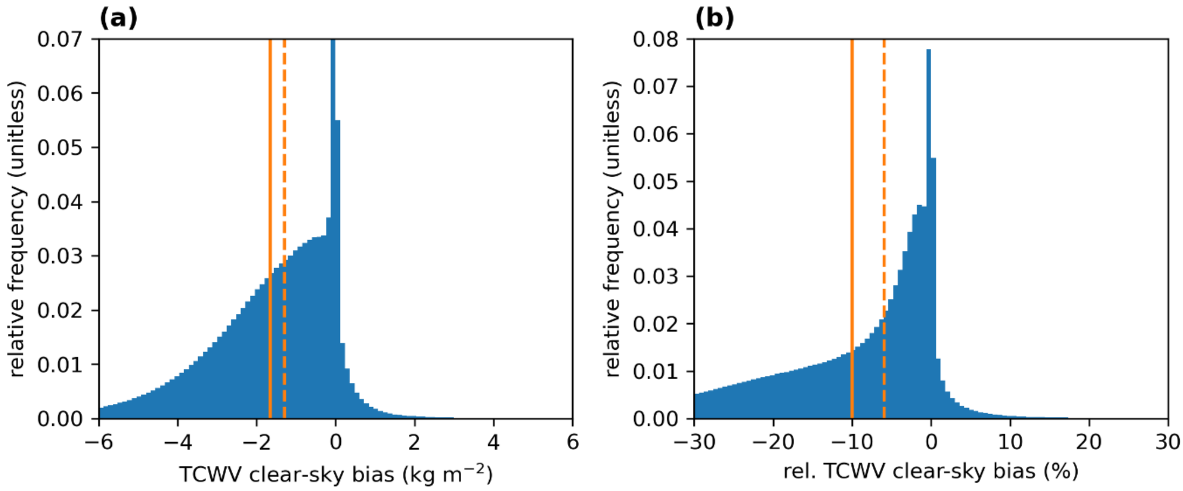

Figure 6. Distributions of the absolute differences ("clear; Panel a) and relative differences ("clear=TCWVall; Panel b) of the monthly mean differences between clear-sky and all-sky ERA5 data based on the OMI cloud information. The solid and dashed orange line indicate the mean and the median of the distributions, respectively.

In addition to the sampling error and the clear-sky bias, we also examined in Appendix C to what extent the monthly means would change if no RA-filter is applied, i.e. if all data of the complete OMI swath were available. It turns out that although deviations arise due to the RA-filter, these deviations are almost an order of magnitude smaller than those of the clear-sky bias and the global distribution of the deviations is mostly noisy. Due to this small influence of the RA-filter, we conclude that the filtered OMI TCWV data are a good representation of the actual TCWV values.

Furthermore, there is no mention regarding the quality of the datasets chosen for the inter-comparison exercise. Addition of this information at the beginning of section 3 would help inform a reader unfamiliar with these data sets to why they were used by this study.

We thank the reviewer for this suggestion and have restructured Section 3 accordingly. Moreover, we added the following text to the beginning of Sect. 3:

To evaluate the overall quality of the OMI TCWV data set, we conducted an intercomparison study for which we use the merged, 1-degree total precipitable water (TPW) data set version 7 from Remote Sensing Systems (RSS) (Mears et al., 2015; Wentz, 2015), TCWV data from the reanalysis model ERA5 (Hersbach et al., 2019, 2020), and the ESA Water_Vapour_CCI (WV_cci) climate data record CDR-2 as reference.

The RSS data set consists of merged geophysical ocean products whereby the values are retrieved from various passive satellite microwave radiometers. These microwave radiometers have been intercalibrated at the brightness temperature level and the ocean products have been produced using a consistent processing methodology for all sensors (more details in Wentz, 2015; Mears et al., 2015). The major advantages of microwave TCVW retrievals are their high precision and accuracy and that they are insensitive to clouds, so that TCWV values can also be retrieved even under cloudy-sky conditions. A disadvantage, however, is that these retrievals are (mostly) only available over the ocean surface.

Thus, we also compare the OMI TCWV data to the ESAWV_cci CDR-2. At the moment of preparation of this manuscript, the CDR-2 is a beta-version of the combined microwave and near-infrared imager based TCWV data record (COMBI). The CDR combines microwave and near-infrared imager based TCWV over the ice-free ocean as well as over land, coastal ocean and sea-ice, respectively. The data record relies on microwave observations from SSM/I, SSMIS, AMSR-E and TMI, partly based on a fundamental climate data record (Fennig et al., 2020) and on near-infrared observations from MERIS, MODIS-Terra and OLCI (Danne et al., 2022).

Within comparisons between different satellite data sets a major drawback is the influence of sampling errors due to different observation times, pixel footprint sizes or orbit patterns. To minimise this source of error, data from reanalysis models are useful. ERA5 is the fifth generation ECMWF reanalysis (Hersbach et al., 2020) and combines model data with in situ and remote sensing observations from various different measurement platforms. For our purpose, we use the "monthly averaged reanalysis by hour of day" from the Copernicus Climate Data Store on a 1°x 1° grid. To account for OMI's observation time (around 13:30 LT), we first calculate the local time for each longitude in the ERA5 data set, then select the TCWV data for the time period between 13:00-14:00 LT and finally merge the selected data.

[…]

Finally, what is not clear is to whether this new data record from OMI is meant to complimentary to the existing TROPOMI data set? By this, I mean could the records be used sequentially to bring the time series out the end of the TROPOMI mission? I this is the case, how does the performance of these two record compare?

Indeed, in the future we plan to merge the TCWV datasets of OMI and TROPOMI or to continue the OMI dataset using TROPOMI data. For the time being, however, we are refraining from doing so:

- the TROPOMI cloud algorithms are not yet fully developed and there are currently repeated jumps in the TROPOMI TCWV dataset (see for example Küchler et al., 2021).

- the OMI radiances have been processed with the TROPOMI processor since 2022 and will also be reprocessed with it after the end of the mission. This should lead to an improvement of the irradiance and radiance spectra, so that it may be possible to switch from an Earthshine fit to a solar irradiance fit.
- OMI will soon run out of fuel and thus the mission will end soon (2023 or 2024). By then, there should also be enough overlap between OMI and TROPOMI.

Overall, I find that this study is of scientific value and recommend it for publication, after all the issues that I have highlighted are addressed.

2. Specific Comments

- Section 3: I think the term validation here is incorrect as you are performing inter-comparison of the OMI performance against other gridded products at monthly time scales. For this to be a validation study you would need to perform this on the level 2 swath data against ground truth sites. Alternatively, accurate (fiducial) characterisation of these reference products on monthly time scales would need to be done, and this would be a major undertaking in itself.

We agree that the term "validation" is not adequate and will instead refer to an "intercomparison study".

- Lines 157-158: What is the assumption you base the relative error estimates on? From the literature, or results not included in this paper? Are you actually describing uncertainties or do you mean errors? Further elaboration here would make this clearer to the reader.

For SSMI, we followed the results of Mears et al. (2015), who found that the uncertainty for TCWV = 10mm was around 1mm and for TCWV = 60mm around 2-4mm. Thus, we have revised the uncertainties again and specify that the uncertainty is 5% or at least 1mm.
For ERA5 and ESA CDR we can assume similar uncertainties over ocean, since the TCWV values there are also mainly based on microwave observations. For the ESA CDR, "average retrieval uncertainties" are given, but these are unrealistically low over land (<0.1kg/m^2). Therefore, we decided to make a compromise and, for simplicity's sake, set the uncertainty about twice as high as the uncertainty over ocean.
We added the following text to the revised manuscript:
Mears et al. (2015) found that the uncertainty daily microwave TCWV observations for TCWV=10 kg m-2 was around 1 kg m-2 and for TCWV = 60 kg m-2 around 2-4 kg m-2. Hence, we assume that the uncertainty of the RSS data set is 5% or at least 1 kgm−2. For ERA5 and ESA CDR-2 we can assume similar uncertainties over ocean, since the TCWV values there are also mainly based on microwave observations. Unfortunately, no uncertainties are provided for TCWV over land. Thus, for the sake of simplicity, we assume that the relative errors of the reference data sets over land are twice as high as over ocean, i.e. 10% or at least 2 kg m-2. For the OMI TCWV data set we assume an uncertainty of 20% (Borger et al., 2020). We also tested other variants of error assumptions and it turned out that the exact choice of errors is negligible for the regression results as long as the ratio of uncertainties remains similar.

- Section 3.2: For the ERA5 did you take the hourly data and interpolate to the local over pass time or the monthly mean data on hourly time steps? Slight rewording to clarify is needed. Additionally, did you consider using the ensemble output which would have given the spread in the reanalysis rather than assigning a relative estimate of the uncertainty?

For the ERA5 data we used the "Monthly averaged reanalysis by hour of day" from the Copernicus climate data store (CDS). Starting from the OMI overpass time (13:30LT), the local time was determined for each longitude and then the TCWV data for the period 13:00 to 14:00LT were selected and merged.
We added the following text to the revised manuscript:
For our purpose, we use the "monthly averaged reanalysis by hour of day" from the Copernicus Climate Data Store on a 1° x 1° grid. To account for OMI's observation time (around 13:30 LT), we first calculate the local time for each longitude in the ERA5 data set, then select the TCWV data for the time period between 13:00-14:00 LT and finally merge the selected data.

Regarding the use of ERA5 ensemble data, we have to admit that we are not experts in this field and therefore cannot completely understand how the different ensemble members come about. Nevertheless, we have taken a look at the ensemble data and calculated the ensemble spread and the relative spread as follows:
Spread = max(ensemble) - min(ensemble)
Relative spread = Spread / mean(ensemble)

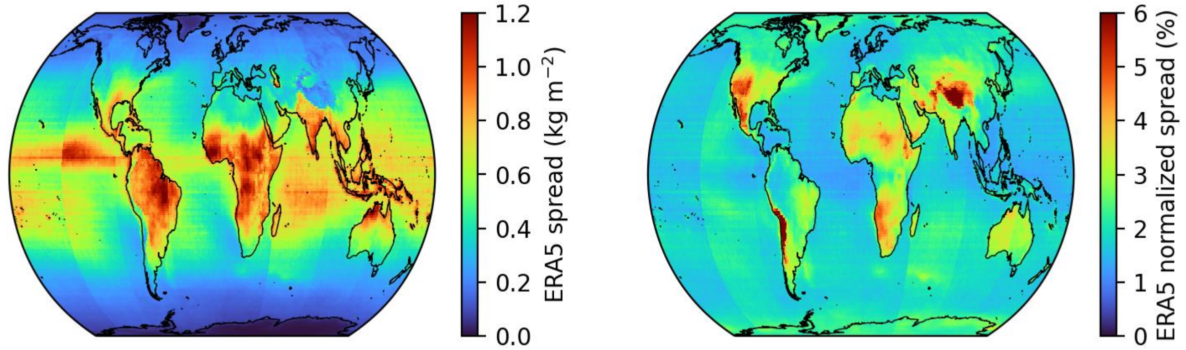

Fig.RC1: Spread (left) and relative spread (right panel) of ERA5 TCWV ensemble data at around 13:30LT derived from 3-hourly resolved monthly mean by hour of day TCWV data for the time range January 2005 to December 2020. The local time was determined by longitude. Then the data was selected for a time between 12:00 and 15:00LT and finally merged.

In our opinion, however, these spreads underestimate the actual uncertainty of ERA5: Over the ocean, the uncertainties are smaller than in the SSM/I (see comment above), although similar input data should be used. And over land they are in some places only slightly larger than over the ocean, which is not entirely understandable, since much less observational data is available than over the ocean, or over land the (satellite) measurements typically have a larger uncertainty.
Therefore, we believe that our assumption for the uncertainty, is a good estimate of the uncertainties in ERA5.

- Figure 7/B3: The comparison to ESA CCI over land – did you also apply stricter cloud filtering to the OMI data as well as the common mask? The improvement in representativeness can be seen in figure B3 but there could still be additional cloud in the OMI data which is biasing the data. The common mask from the ESA data will be for 10:00 hrs LST, while with OMI overpasses at 13:30 hrs LST which will have an impact in convective areas. Finally, is there an improvement in the Hovmöller time series when the common mas is applied?

Following the suggestion of the reviewer, we have also applied a more stringent cloud filter (CF<5% instead of CF<20%) and have carried out the comparisons with it. For the sake of simplicity, we only show the comparisons with ERA5.

When looking at the difference maps for each season, a clear reduction of the systematic overestimation in the tropics or in the Amazon region can be seen on average, but we now also see clear underestimation in areas with frequent cloud cover (e.g. India or Southeast Asia). This indicates that such a stringent filter leads to a strong clear-sky bias. Furthermore, the filter leads to many gaps in the TCWV data set, so that it would no longer be optimal for time series studies.

Therefore, we think that the current configuration with CF<20% is a good compromise between spatiotemporal coverage and quality.

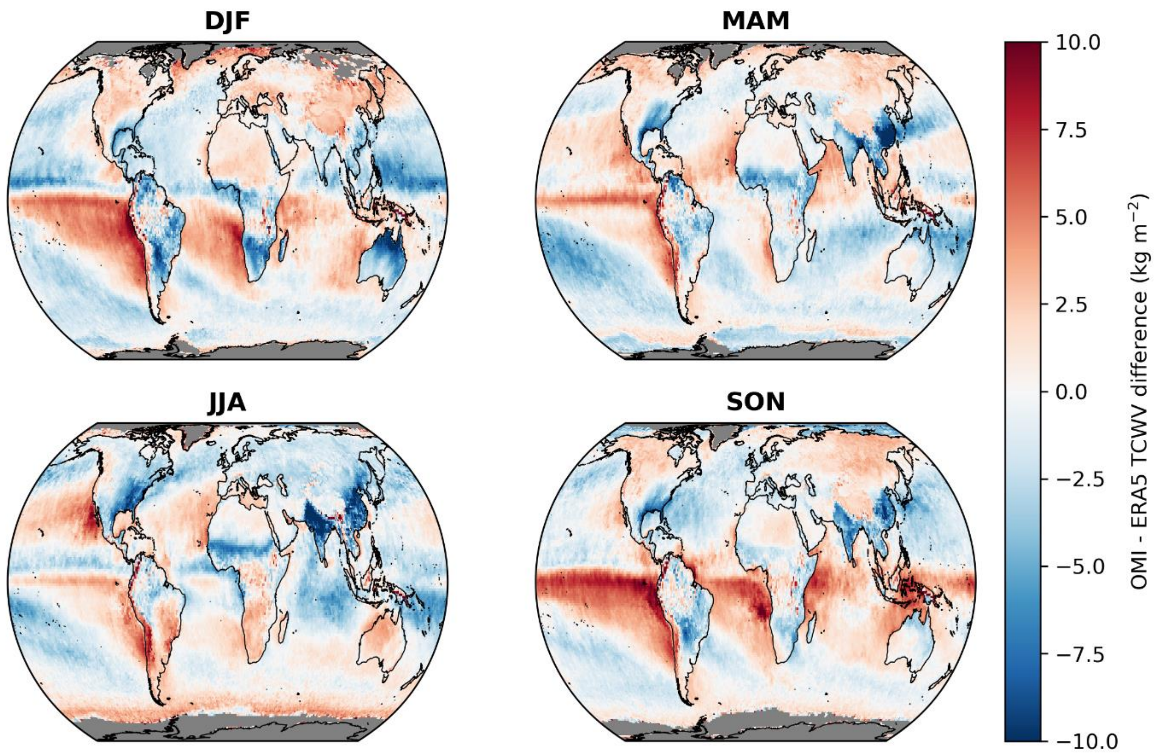

Fig. RC2: Global distributions of the absolute differences between the OMI TCWV data with a stricter cloud filter (eCF < 5%) and ERA5 for winter (DJF), spring (MAM), summer (JJA), and autumn (SON). Grid cells for which no data is available are coloured grey.

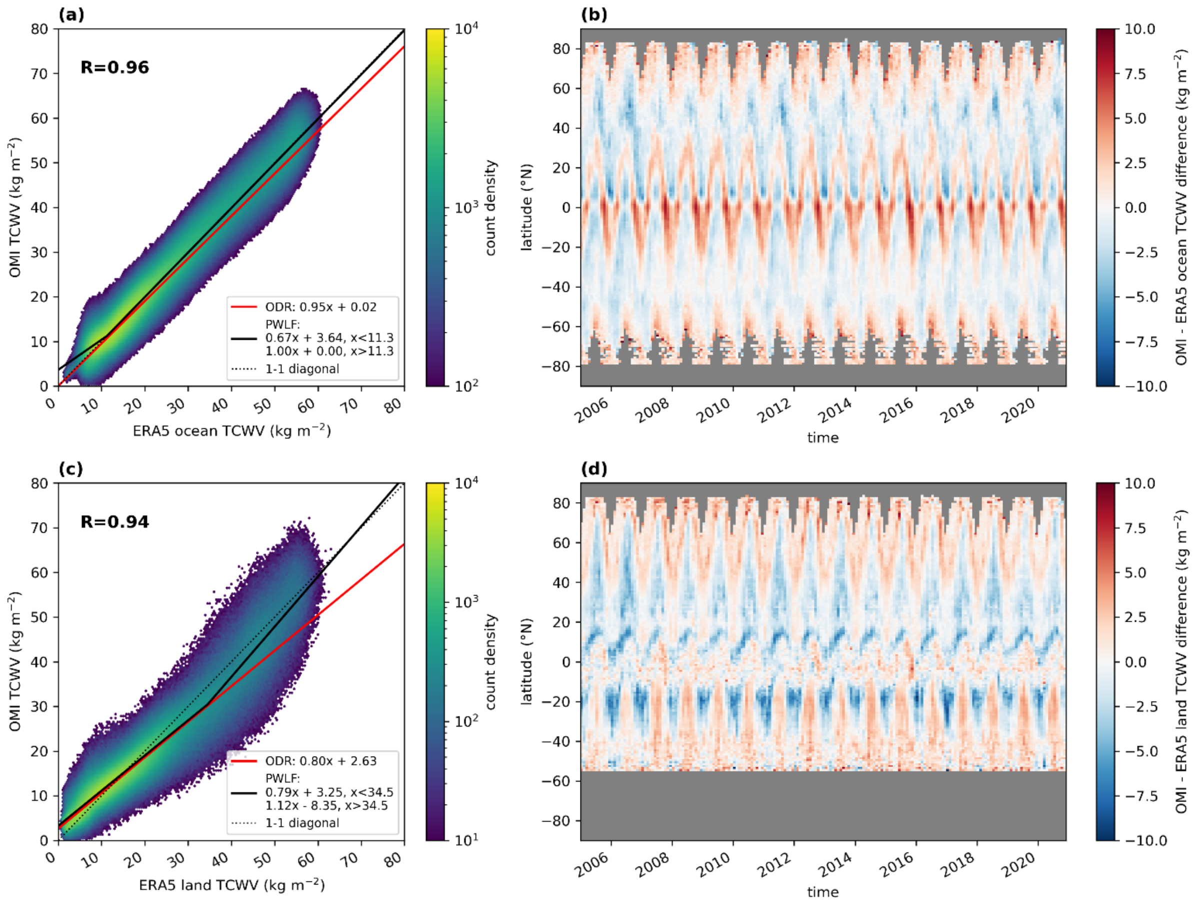

Fig. RC3: Intercomparison between monthly mean TCWV from OMI (with the stricter cloud filter) and ERA5 for data over ocean (top row) and land (bottom row). Panel (a) and (c) illustrate a 2D histogram in which the colour indicates the count density; the red solid line represents the results of the orthogonal distance regression (ODR) and the solid black line the results of the piecewise linear regression (PWLF). The results of the respective fits are given in the bottom right box and the correlation coefficient in the top left corner. The dashed black line indicates the 1-to-1 diagonal. Panels (b) and (d) depict the TCWV difference of OMI minus ERA5 within the latitude-time space; reddish colours indicate an overestimation, blueish colours an underestimation of the OMI TCWV data set.

3. Technical Comments

- Line 38: the reference Susskind et al. 2003 is for joint microwave and infrared retrievals from AIRS. Therefore, is not an explicit reference for IR water vapour retrievals. There is also an extra ')' on line 39 after the reference, did you mean to have the 2003 in-cased in parenthesises?

Thank you for the clarification regarding the Susskind et al. reference! We have added the references of Schlüssel et al. (2005) and Schneider and Hase (2011) as examples of TIR retrievals.

- Line 39: both your references here are for near infrared retrievals from MERIS, missing a shortwave infrared reference e.g. SCIAMACHY (2.3 µm), GOSAT (1.6 +2.1 µm), or TROPOMI (2.3 µm).

We have added the references of Schrijver et al. (2009), Dupuy et al. (2014) and Schneider et al. (2020) as examples of SWIR retrievals.