Review of manuscript essd-2021-319 entitled “A 16-year global climate data record of total column water vapour generated from OMI observations in the visible blue spectral range” by Christian Borger, Steffen Beirle, and Thomas Wagner.

General comments

This manuscript presents a total column water vapour (TCWV) data set derived from 16 years of OMI observations. The retrieval method was developed in a previous publication (Borger et al., AMT, 2020). It has been slightly improved in order to meet the long term stability requirements for a climate data record. The manuscript describes briefly the modified aspects of the retrieval algorithm and gives additional details on the data quality control applied for this specific purpose. The latter seems to reject a significant fraction of the raw data, although this number is not indicated. Most of the manuscript is devoted to a comparison/validation analysis of the OMI TCWV results with respect to satellite microwave radiometer data (SSM/I), reanalysis data (ECMWF’s ERA5), and the ESA/CCI/CDR-2 water vapour product. The comparison results are fairly documented, including scatter plots, Hovmoeller diagrams and maps, although some synthetic statistics are missing (see the specific comments below). However, the conclusions sound far too optimistic to me, given the poor agreement found between the OMI data and the validation data. Especially, the large positive biases over land and near the coastlines in the tropics are striking and not sufficiently commented or explained. Two main reasons are hypothesized: too low land surface albedo and incorrect cloud information, both leading to an underestimation of the AMF. These paths should be further explored in order to achieved a more reliable product meeting the climate data quality requirements. Although it is shown that the results are improved when a special could mask is used, this is only an artificial way to improve the quality of the product.

Regarding the temporal stability, it is not clear how the significance of the global mean bias, RMSE, and trend differences are established. It seems to me that the numbers are beyond the limits usually required for water vapour climate data (e.g. an error of 0.1%/decade in the global mean TCWV trend represents nearly 20% of the signal). Moreover, the uncertainty due to different time and space sampling with the different reference products should also be quantified.

In conclusion, it is my feeling that the proposed data set has significant defects that are not well understood. I recommend first a more insightful analysis of the error sources, especially over land and, if possible, the elaboration of an improved version of the data set, and second, a more comprehensive discussion of the validation results in a revised version of the manuscript.

Specific comments

More should be said about the “row anomaly” which affects the OMI observations throughout almost the whole period analysed in this paper. Figure A2 shows that a large fraction of “rows” are discarded. Is it sufficient to discard these rows or could adjacent rows also be affected in some way? What is the impact of this screening on the representativeness of the final observations?

Why are two regression methods (OLS and ODR) used? In principle, a single statistic is sufficient, unless the difference of results from the two are discussed, but this is not done in this manuscript. I suggest either to choose one or to better justify the choice of two and analyse the obtained differences.

L77: replace “I and I0” by “I0 and I”

L129: define also Delta_SCD

L144: indicate which fraction of raw data is remaining

L154: “ESA Water Vapour CCI climate data record CDR-2” needs a reference

L154: “For the correlation analysis” is misleading or incorrect if referring to regression analysis. Please reword (e.g. For the intercomparison…)

L155: add a reference for the ODR method

L155-156: “In the case of the ODR it is necessary to use reasonable ratios of the relative errors of the compared data sets instead of using absolute errors in order to obtain meaningful results”. This statement needs to be justified by an adequate explanation or reference.

L156-159: these sentences sound in contradiction with the previous statement. Moreover, the sensitivity of the regression results to the relative errors should be discussed in more detail (e.g. in an Appendix) and the choice of 5%, 10%, and 20% for the three dataset (which appear quite arbitrary) should be clearly motivated/discussed.

L169-171: “In general the deviations are quite low with values between +/- 2.5 kg/m2” be more specific in quantifying the differences here, e.g. indicate which fraction of data lie in the range of +/- 2.5 kg/m2, or use quantiles or other statistics (mean, standard deviation, etc.). Note also that the correlation coefficient is not much relevant when the seasonal variations are included.

L173-178: Be more quantitative again, here in the comments on Fig. 4. I would also suggest to include the coastlines of Africa and Indonesia in the list of regions with significant positive deviations.

L176: Be more specific on the impact of the “cold tongue” and “too low albedo” on the observed deviations.

L182: “the slight overestimation of 3-5%”: it is not clear what these numbers represent exactly. Is it a mean difference (bias)? Is it computed over all data or only a fraction? (Note that a slope of 1.03 does not mean that all the values are 3% higher, this depends also on the intercept value).

L194: how is the change-point at 26 kg/m2 selected in the piecewise linear regression?

L210: satellite measurements in the thermal infrared are NOT available/reliable in cloudy conditions.

L209-215: I’m not convinced that the ERA5 uncertainty over tropical land areas contributes much to the huge bias observed in the differences (above 10 kg/m2). This idea should be further documented or discarded (also in the Conclusion).

L227: is there any update on the publication of the ESA CDR-2 data set?

L225-254: Similar comments as for ERA5 apply here to the CDR-2 comparison (lack of statistics, etc.).

L261: More details are needed on the linear regression method and significance tests.

L274: The stability requirement for water vapour climate data is rather at the level of 0.3 %/decade (GCOS – 112, April 2007).

Figure 3: the fit results would be more understandable if given as an equation: $y = 1.03\ x + 0.18$ rather than just two numbers.

Figure 3: indicate that the OMI results here are over ocean (it is only obvious if one knows that SSM/I data over only over the oceans).

Figure 4: add the piecewise linear regression lines (mentioned L194) on the plot.

Figure 9: the red dashed lines are not visible in the plots.