Authors' Response to Reviews of

Evaluation of Total Column Water Vapour Products from Satellite Observations and Reanalyses within the GEWEX Water Vapor Assessment

Tim Trent, Marc Schroeder, Shu-Peng Ho, Steffen Beirle, Ralf Bennartz, Eva Borbas, Christian Borger, Helene Brogniez, Xavier Calbet, Elisa Castelli, Gilbert P. Compo, Wesley Ebisuzaki, Ulrike Falk, Frank Fell, John Forsythe, Hans Hersbach, Misako Kachi, Shinya Kobayashi, Robert E. Kursinsk, Diego Loyola, Zhengzao Luo, Johannes K. Nielsen, Enzo Papandrea, Laurence Picon, Rene Preusker, Anthony Reale, Lei Shi, Laura Slivinski, Joao Teixeira, Tom Vonder Haar, and Thomas Wagner

EGUsphere, https://doi.org/10.5194/egusphere-2023-28087, 2023

Please find below our responses to the community comments, We include the original comments in black, our comments in green, and any alterations to text in blue.

This is an important and well constructed study that should in my opinion be published. It's impact could be improved with enhanced motivation to bring out the importance and some more direction to the community in terms of more/less reliable products and strengths/weaknesses or limitations and recommending more or less suitable applications. I just have minor comments listed below.

1) L14 ice free regions change with the season and year - will this alias (slightly) into the variability? TWCV --> TCWV

Our approach is to use the sea-ice mask to remove any grid cells where we have found sea ice throughout the 1988-2014 period, so it is invariant over time in that sense. We also do not use data beyond ±60 degrees. Therefore, the analysis does not consider areas where we would observe this effect. The sea-ice mask then modifies this latitude band range consistently across all records.

2) L15 Why are % changes considered in the fit to temperature but kg/m^2 in the trends? Although ranges are shown, it would be useful to the community to have some expert judgement, such as removing obviously spurious datasets (what is the expected physical range?)

Results from sensitivity to temperature are shown in %/K as the expectation are changes between 6-7.5 %/K as outlined in the new section 3.3.

As in the first phase of G-VAP we only analyse trends in absolute units and per decade. The main motivation is not climate change analysis but trend estimation as a tool to intercompare and characterise data records.

If trends in %/decade are shown a comparison to results from sensitivity to temperature is still not straightforward because results from the sensitivity analysis are computed via regression and not via trend ratios (see also discussion of point 19).

3) L23 seems to be missing text. Also Forster et al. (2021) (IPCC Chapter 7) deals more with radiative effects of water vapour. Some mention of recent updates in the field of water vapour and climate would strengthen the context and motivation of the study e.g. Colman & Soden (2021) RevModPhys doi:10.1103/RevModPhys.93.045002; Allan et al. (2022) JGR doi:10.1029/2022JD036728; Ding et al. (2022) LNEE doi:10.1007/978-981-19-2588-7_27; Douville et al. (2022) Comm. Earth Env. doi:10.1038/s43247-022-00561-z; Wu et al. (2024) GRL doi:10.1029/2023GL107909; Wan et al. (2024) HESS doi:10.5194/hess-2023-301 which build on previous assessments e.g. Trenberth et al. (2005) Clim. Dyn. doi:10.1007/s00382-005-0017-4

We thank you the comment and point us to these references. We have incioprated some of these into updates withihn the introduction.

4) L25 water vapour feedback magnitude should be updated to the latest IPCC report chapter (Forster et al. 2021). It should also note that mid and upper tropospheric water vapour is more important to the feedback strength than lower tropospheric changes that column integrated water vapour is more closely related to. There is however an important link between column integrated water vapour and precipitation as well as downward longwave radiation and atmospheric absorption of sunlight, both of which also impact the energy-water Tim, thank him say we have added some of these to text to enhance the manuscriptcycle coupling (e.g. Douville et al. 2021 IPCC; Fowler et al. 2020 Nature Rev Earth Sci. doi:10.1038/s43017-020-00128-6).

We thank you for highlighting the update in the magnitude of water vapour feedback and have updated the text inline with Forster et al. 2021. We also thank you for the additional comments and references, which we have used to update the introduction

5) L26 water vapour feedback magnitude is not "compared" to greenhouse gas forcing

This statement has been modified as part of updates made to the introduction.

6) L60 The 'long period' was also presumably chosen to commence at the start of the SSM/I record and for consistency with previous analyses e.g. Allan et al. (2020) NYAS doi:10.1111/nyas.14337

The start was chosen as a compromise between number of data records and maximised length. Thus, we decided already in the first phase of G-VAP to start the “long-term period” in 1988.

7) L75 - are the AIRS + AMSU v6 Obs4MIP data set (Tian & Hearty, 2020 Earth & Space Science, doi:10.1029/2020EA001438) version also evaluated? These were developed to remove systematic biases and allow better comparison with CMIP simulations.

No we only used the science products for consistency with other data records used in the study.

8) L102 do all products vertically integrate to the top of atmosphere or are some cut off at a certain level?

All are vertically integrated to TOA values

9) L105 I didn't understand "lower, respectively higher spatial resolution"

In order to avoid confusion we included the actual resolutions and removed “lower” and “higher”.

10) L122 it would be useful to mention limitations of the datasets. For example ERA5 and other reanalyses are subject to a changing observing system that can introduce spurious changes, though water vapour now seems quite robust in ERA5 after the mid-1990s (e.g. Allan et al. 2022). The 20CR only assimilates SST and surface pressure so water vapour is determined by the model based on these contraints and so is for all intents and purposes an atmospheric model "amip" type simulation nudged towards realistic atmospheric circulation. For satellite datasets, degradation in sensors, orbital drifts and intercallibration present a challenge

We do not detail performance limitations for every record within the assessment, as these are usually documented elsewhere. Instead, we test the performance of data records and where issues are identified, investigate them further. We have added additional text to further explain the process we follow on the assessment:

“It should be noted that all water vapour records will have limitations based on their underlying assumptions or operational frameworks. For example, satellite sensors can experience degradation (often corrected through recalibration efforts e.g. Tabata et al. (2019)), reducing the sensitivity of an instrument, while reanalysis records can experience introduce shifts in the time series due to changes in observing systems assimilated (Schröder et al., 2017; Allan et al., 2022). Individual data record performance assessments are usually detailed in publications or via technical documents such as the Product User Guide (PUG) or Validation Report (VR) and are not provided here. Through the assessment, we can highlight performance issues (e.g. breakpoints) and attempt to map them to know issues. Where we cannot identify the cause, our results can be used by the data record teams in future product updates.”

11) L204 - is a consistent land/sea/ice mask (e.g. Fig. 2) applied to all datasets (if not, this could introduce differences in variability). Is the mask a climatology as in Figure 2 (though ice varies seasonally and interannually) or does it vary from month to month?

See comments above

12) Figure 3 - is this a median across datasets? A fuller caption may help

This is the median across all datasets that span the common long-period, the caption has been modified to make this clearer to readers

13) Figure 4 - which of these correlations are significant or not?

Figure 4 has been updated to highlight which correlations are significant. This is presented a s heatmap to compliment the exisiting cortrelation plot.

14) L293 - seasonal range usually means range over the year but I think intra-seasonal range is meant?

Thank you for spotting this; the text has been updated accordingly.

15) L305 - large (e.g. 2-sigma deviations) could usefully be reported to suggest outliers

An addition of +/- 2 sigma lines have been added to figure 6

16) L312 ERA5 does not seem significantly wetter (e.g. probably depends on years chosen)

Yes, this could be the reason. However, we only analyse set time periods within the assessment and do not have scope to extend the analysis at this time. This definitely something we should consider in the future.

17) L317 IR estimates presumably sample clear-sky regions which are systematically drier than cloudy regions e.g. John et al. (2011) JGR doi:10.1029/2010JD015355 (presumably visible records are also susceptible). I think this is discussed later but could be flagged earlier.

This forms part of the clear sky bias discussion in section 5.

18) L339 missing reference

This has been rectified.

19) L343 Are these annual trends? Were trends in %/decade also computed? This could remove mean bias effects (e.g. wetter datasets may vary more in absolute terms but not percent) and it would be useful to quote % changes for consistency with other analysis (e.g. sensitivity to temperature) and the literature

As in the first phase of G-VAP we only analyse trends in absolute units and per decade. The main motivation is not climate change analysis but trend estimation as a tool to intercompare and characterise data records.

If trends in %/decade are shown a comparison to results from sensitivity to temperature is still not straightforward because results from the sensitivity analysis are computed via regression and not via trend ratios.

20) L347 is this the interannual regression or does it include the seasonal cycle (which is determined by very different processes)? Or is it the trend in TCWV divided by the trend in temperature? For example in Allan et al. (2022) the ERA5 global TCWV sensitivity to T2m is 5.76 +- 0.35%/K for 1988-2014 while the trend is 0.78+-0.08 %/decade which combined with a warming of 0.17 K/decade gives a lower sensitivity of 4.6 %/K. It was also noted that ERA5 decreases in TCWV over the ocean before the mid-1990s are at odds with the SSM/I record. Ocean and land estimates are also available in the paper.

For trend estimates we fit the annual cycle and ENSO strength simultaneously, while for the regression we remove the seasonal cycle from both the surface temperature and water vapour data. We understand that this may give different answers compared to other studies; however, for the assessment we do not claim that these are climate estimates, rather estimates of performance. This also keeps the analysis between different phases of the assessment consistent.

21) Figure 9 - if microwave values are masked does this mean there are variable numbers of datasets in each grid point? MIssing reference in caption.

This is indeed true. We did not include a corresponding statement in the caption as it should be obvious and because it does not impact the interpretation of the figure. The reference now points to Figure 2.

22) Figure 10/11 could be combined (and enlarged). It may also be useful to have a zoom in on the more homogeneous datasets since the outliers dominate somewhat

We may either include a single column plot as it is now or a double column plot. We may change to double column and ask the editor if this is acceptable. Else, we prefer to have separate figures to emphasise the focus of land versus ocean results.

23) L395 do any of the ERA5 breakpoints coincide with the early 1990s low latitude ocean trends identified in Allan et al. (2020, 2022) and Hersbach et al. (2020) that were also in previous versions of this dataset and linked with changes in the observing system? These seem linked with decreases in surface relative humidity and 850 hPa specific humidity over the ocean in the late 1980s-early 1990s.

Within the scope of our study, we are not able to determine regional trends. Therefore, we only focus on analysing global trends. Previously, these were done for global ice-free ocean only, but this time, we have also included land estimates. We think this would be an interesting avenue for future research and will consider

24) L402 - can changes in the mix of SSM/I satellites introduce spurious variability since they are observing at different overpass times. Some studies use particular SSM/I satellites with more stable or consistent overpass times to avoid this (e.g. Allan et al. 2022).

In general, a mix of satellites may introduce spurious variability even when carefully intercalibrated and optimally merged. One of the main aspects of the results presented in this paper is actually the impact of a changing observing on homogeneity. In line 402 we conclude that the observed break point does not coincide with a change of the observing system. However, we did not analyse if changes of observing system impact variability. We propose to add in line 402: "It can be a topic of future G-VAP efforts to analyse this feature further, e.g., by comparing the full SSM/I and SSMIS climatology to a climatology of near constant equator crossing times (similar as in Allan et al., 2022)".

25) Fig. 12 - the stippling seems to show discontinuity south of Alaska and south of India?

On closer examination we do not see this effect, we did notice issues on some draft printouts but not in the electronic pdf. Thus, we cannot clarify this issue.

26) L435 but only some of the breakpoints are matched to phyiscal causes?

Within the scope of this study, we are only able to identify most of the causes of the observed breakpoints. Where we were unable to identify a cause we mark these as 'Unclear'. We still include the information to 1) inform potential users, and 2) so data product teams can investigate these in greater depth during product updates. We update the sentence to be clearer:

“Most data records are affected by breakpoints, where some of the physical causes can be identified.”

27) L460 although clear-sky sampling introduces dry biases, it only affects moisture variability if the clear-sky regions vary in a different way to the cloudy regions (e.g. Allan et al. 2003 QJRMetS doi:10.1256/qj.02.217)

Thanks for this comment. We did not explicitly analyse this aspect, except that we looked at spatial variability of the bias and at its annual dependencies. In both cases, we observed changes from month-to-month changes and regions to region. This variability is usually driven by changes in cloudy patterns.

28) Conclusion - the impact of this considerable work could be enhanced with some recommendations to the community with regard to better and worse datasets for particular applications (e.g. climatological, regional variability, interannual variability and long-term trends).

This part is outside the scope of the exercise. We indirectly do this by showing the results, this is intentionally left to the reader and depends on the type of analysis people do. We provide the information for users to make their decision

Richard Allan

We thank you the time taken to review the manuscript and many helpful comments and suggestions.