Responses to Reviewer #2

Reviewer #2

I recommend rejection of this manuscript as it suffers from two fundamental flaws:
1. ERA5 Stratospheric Water Vapor (SWV) is the central dataset in this manuscript. It is studied in sections 3 and 4 and it is used as reference for CMIP6 evaluation (section 5). The manuscript fails to consider the nature of the ERA5 reanalysis of water vapour, i.e. an optimal representation between model state, in-situ humidity observations in the troposphere and satellite radiance observations which are sensitive to humidity only in the troposphere. This appears clearly in Hersbach et al. (2020) where section 5 provides an exhaustive description of the assimilated observations.

Response: Thank you for your criticism. We agree that ERA5 SWV might have some biases. However, we still believe that the bias in ERA5 is smaller than that in models relative to observations.

To well address your concern, we assess the quality of the ERA5 SWV with the SWOOSH as the baseline. It is revealed that the ERA5 quality is not so bad as imagined. The assessment of ERA5 SWV is provided in Figure 1, and our reviewers might find that the SWV in ERA5 and SWOOSH is highly similar especially below 10 hPa.

For your concern, we also add more discussion of the ERA5 assimilation of WV.

- "In contrast to the troposphere, the WV content within the stratosphere is extremely low. Compared to the SWOOSH satellite observation data, the ERA5 reanalysis data provides a longer time span, which provides more samples for revealing the effect of QBO on SWV. There remains uncertainty regarding the performance of ERA5 reanalysis data in depicting SWV. In the ERA5 reanalysis, WV mainly assimilates in-situ humidity observations in the troposphere and satellite radiation observations that are only sensitive to humidity in the troposphere. Therefore, the SWV in ERA5 reanalysis data is a GCCM output with specified dynamics (Hersbach et al., 2020). From 2000 to 2006, ERA5 showed cold deviation in the lower stratosphere. The global average temperature of the stratosphere and tropospheric apex corrected by ERA5.1 was better than that of ERA5 (Simmons et al., 2020). Previous studies have shown a wet bias in the tropical tropopause in the ERA5 reanalysis data (Kruger et al., 2022). Some studies also found that the content of SWV in ERA5 was superior to that of ERA-Interim (Wang et al., 2020). We use SWOOSH satellite monitoring data to validate the applicability and uncertainty of ERA5 reanalysis data in SWV. Figure 1 shows the evolution of the annual mean, summer mean, and winter mean of the WV mixing ratio in the troposphere during SWOOSH satellite data and ERA5 reanalysis from 1992 to 2019. Compared to SWOOSH satellite data, ERA5 can better display the distribution pattern of SWV, and the WV content is basically consistent. The WV content in ERA5 has a 0.3 ppm moisture deviation at the bottom of the tropical stratosphere, which is reflected in the annual mean, summer mean and winter mean, consistent with previous analyses (Kruger et al., 2022). However, at the top of the stratosphere, the WV content in ERA5 is all relatively low." (L104-119)

We hope our reviewer can understand that the ERA5 really provides more QBO samples than other datasets before we have enough observed WV time series. Thank you for your criticism again.

While satellite retrievals from limb-scanning instruments are assimilated after 2002 in the case of ozone (i.e. MIPAS and Aura-MLS), no similar dataset is assimilated for water vapor. SWV has a negligible impact on satellite observations of microwave and infrared radiances because they are observed in a nadir-looking geometry. Hence ERA5 SWV is not influenced by these observations, is fundamentally nothing more than the output of a GCCM with Specified Dynamics, and is thus far from a "true" dataset. This is also indicated by Hersbach et al. in their Fig. 12, where analysis increments of humidity are not shown above 300 hPa while analysis increments of temperature, zonal wind and ozone are also shown at 50 hPa and 3 hPa; and in the discussion of their Fig. 19 (SWV above the South Pole) where it is specifically written that "no humidity observations are assimilated at this level" (850 K isentropic surface i.e. mid-stratosphere).
Response: We agree that ERA5 reanalysis can not represent the real world. However, the ERA5 reanalysis provides more QBO samples. To well address your concern, we give a direct comparison between ERA5 and SWOOSH for the zonal mean WV from 1000 hPa to 1 hPa. It is revealed that the WV above 10 hPa diverges between two datasets. However, the zonal mean WV pattern below 10 hPa shows somewhat similarity between two datasets. Also see the response above.

We also compare the timeseries of the tropical WV from 1990-2020 in Figure 2. We noticed that the WV anomalies above 30 hPa are weaker in ERA5 than in SWOOSH. We mainly focus on the WV below 30 hPa (especially in Figure 5).

We also give more discussion on the comparison between both datasets.
- "By comparison, it is found that although there are some differences in the SWV QBO between the SWOOSH satellite data and the ERA5 reanalysis data, the ERA5 reanalysis data reproduce the distribution pattern of WV propagation from the lower stratosphere to the upper stratosphere below 10 hPa. Therefore, the long-term data from ERA5 reanalysis can still be used to diagnose the influence and dynamics of WV in the middle and lower stratosphere below 10 hPa. Our subsequent analysis mainly uses ERA5 reanalysis data." (L175-179)

The accuracy of ERA5 in the stratosphere (and its 2000-2006 correction ERA 5.1) are summarized in section 7 of Hersbach et al. (2020 – see especially their Fig. 15) and extensively discussed by Simmons et al. (2020, not cited in the manuscript). Yet such uncertainties are not considered in this manuscript, which also fails to mention the well-known moist bias in ERA5 at the tropical tropopause (Krüger et al., 2022). An earlier study of stratospheric water vapor in ERA5 did conclude that SWV is better represented in ERA5 than in ERA-Interim (Wang et al., 2020) but this earlier study of the QBO in ERA5 SWV is not cited either.

Response: All those references are mentioned and cited this time. We list all the revisions for your reference.

- "From 2000 to 2006, ERA5 showed cold deviation in the lower stratosphere. The global average temperature of the stratosphere and tropospheric apex corrected by ERA5.1 was better than that of ERA5 (Simmons et al., 2020). Previous studies have shown a wet bias in the tropical tropopause in the ERA5 reanalysis data (Kruger et al., 2022). Some studies also found that the content of SWV in ERA5 was superior to that of ERA-Interim (Wang et al., 2020)." (L109-113)
- "Compared to SWOOSH satellite data, ERA5 can better display the distribution pattern of SWV, and the WV content is basically consistent. The WV content in ERA5 has a 0.3 ppm moisture deviation at the bottom of the tropical stratosphere, which is reflected in the annual mean, summer mean and winter mean, consistent with previous analyses (Krüger et al., 2022)." (L115-118)

The confusion between ERA5 SWV and observed SWV could have been partially overcome by using the SWOOSH dataset, which is entirely based on observations. Yet SWOOSH is only used for Fig.1. After a very superficial comparison between ERA5 and SWOOSH SWV (line 134), the authors seem to believe that ERA5 assimilated HALOE SWV data (lines 135-136) while this is not true.

Response: Thanks very much for your suggestions. Your suggestions have enabled us to have a deeper understanding of the ERA5 reanalysis data of water vapor. We have provided detailed explanations and discussions in the data and methods sections. (L104-124)

The comparison of water vapor in ERA5 reanalysis data and SWOOSH data was also made in Figures 1, 2 and Figure S1. (L164-179)

2. This manuscript attempts to study the seasonality of the QBO signal in water vapor by comparing composite of **anomalies** in "boreal summer" and "boreal winter" (which I understand as JJA and DJF, respectively). While the methodology section severely lacks details, one can still read (line 99) that "The anomaly refers to the deviation of the monthly data from the monthly climatology…". Since these anomalies are the signal **after removal of the seasonal cycle,** it makes no sense to study the differences between their composites for DJF and JJA. Yet this method is at the core of the manuscript, as shown by Figs 4 to 10 and 12 to 14. Furthermore, section 2 states that a "Butterworth first-order bandpass filter was used to extract the water vapor variations at the period of 15-60 months". How can the seasonality of the dataset be preserved after the application of such a filter? The methodology used here would have been greatly clarified (and probably invalidated) by showing time series of SWV prior to calculation of anomalies, and also prior to the application of the bandpass filter.

Response: We understand your concern. QBO has a spectrum maximum far from the annual cycle. It is more reasonable to remove the annual cycle of focused variables before we study the composite QBO signals. The annual cycle is naturally related to the revolution of the Earth, which is not the focus of our study. In contrast, the QBO has a period of around 28 months. We use a wider period spectrum to keep the QBO signals (15-60 months).

In short, we use the anomaly fields to study with an expectation of removing the interference of the annual cycle.

Compare with the methodology used and explain in much better detail by Wang et al. (2020), where seasonal cycle in ERA5 SWV is clearly shown by **anomalies from the climatological annual mean** (fig. 2) while the QBO cycle in ERA5 SWV is shown by **anomalies relative to the monthly climatology** (fig. 3).
Response: We read the paper by Want et al. (2020). Anomalies should be the deviation from the annual cycle. Please also see the responses below.

The large differences found here between DJF and JJA anomalies are thus quite difficult to interpret. They could be due to inconsistent periods for the removal of the monthly climatologies. In other words: the period used for the climatological seasonal cycle is 1960-2020 (line 98) but are the "composite" SWV anomalies shown from Fig. 4 onwards also computed for that period?
Response: Thanks for your suggestion. As you said, we used monthly climatology to calculate the anomalies so as to show the water vapor QBO characteristics. Wang et al. also showed the water vapor QBO characteristics in the monthly anomaly (Fig.v3). We introduced the calculation method of monthly anomalies in the data and methods.

- "The climatology of a variable is calculated as long-term monthly average over the time period from 1960 to 2020. The anomaly refers to the deviation of the monthly data from the monthly climatology with the trend removed for each calendar month." (L136-137)

Figure 9 in Want et al. shows the water vapor anomalies in northern winter and summer. By extracting several points, such as 1985, 1990, 2005, etc., it can be clearly found that the water vapor anomaly in winter is stronger than that in summer. We copied their figure below for your reference.

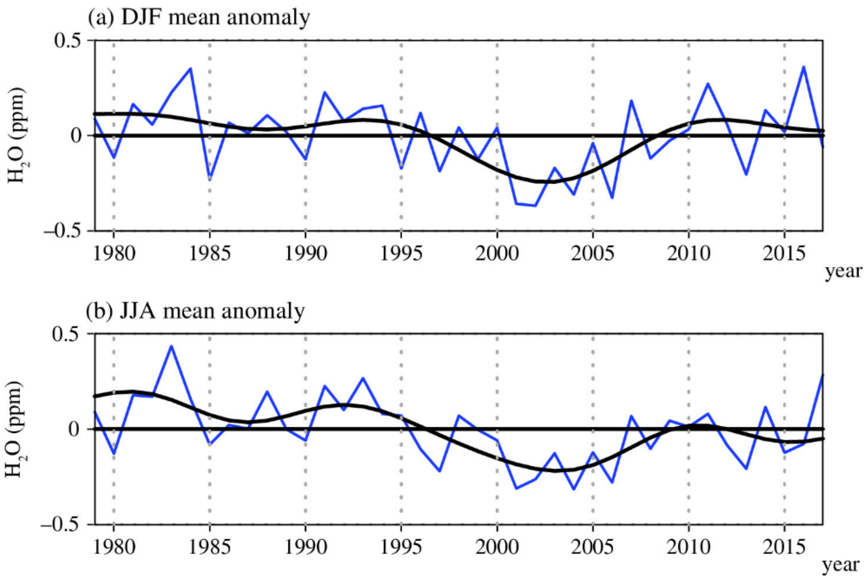

Figure R1. Time series of mean tropical SWV anomaly (15°S–15°N mean at 15–20 km) in (a) DJF and (b) JJA. Data is from ERA5. (Wang et al.'s Figure 9)

The Butterworth filtering is also shown in Wang et al.'s Figure 3 is copied below for your reference. They show that the stratospheric water vapor presents obvious periodic changes of 8-10 years. Therefore, we used the filtering method to extract more significant QBO signals (Fig.2, Fig.S1). You are concerned that the filtering might filter out the seasonal signals, so we showed the original water vapor anomaly in Figure S2, S3.

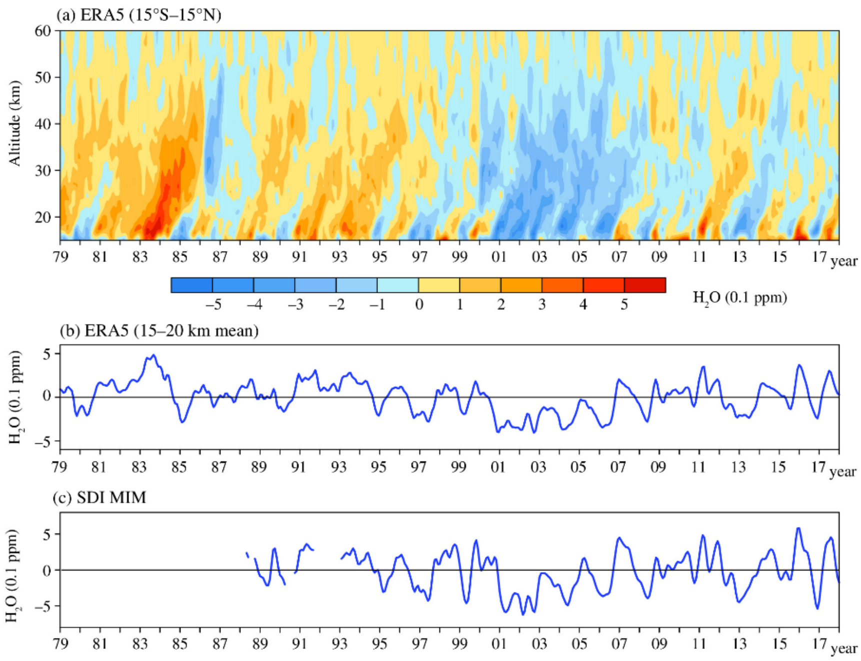

Figure R2. Tropical SWV anomalies (15°S-15°N mean) relative to the monthly climatology using (a, b) ERA5 and (c) SDI MIM. The panels (b, c) are based on the 15-20 km mean. Unit: 0.1 ppm (Wang et al.'s Figure 3)

Incorporating feedback from other reviewers and considering the lagged influence of the QBO on lower stratospheric water vapor, we found that the impact of the 30 hPa QBO index on 100 hPa water vapor peaks at a six-month lag. Although the lag effect remains stronger in winter than in summer, the response sign is identical between two seasons now. (Figs. 4, 5). (L220-230)

Other comments

In case the manuscript undergoes major revisions, the following issues should be addressed as well:

1. The abstract should better introduce the key concepts, i.e. the QBO and "the seasonal difference in the water vapor QBO".

Response: The abstract has been rewritten to clearly distinguish between the QBO and the water vapor QBO. (L10-23)

2. The introduction itself is not written in a rigorous manner:

The very strong statement of the first sentence ("Water vapor is the dominant greenhouse gas in the atmosphere") is not clearly supported by the two provided references (Dessler et al., 2013; Solomon et al., 2010) which are about the feedback between SWV and tropospheric climate.

Response: Revised.

- "The feedback of water vapor (WV) has an impact on global temperature changes (Held and Soden, 2000; Dessler et al., 2013; Solomon et al., 2010)……" (L25-26)

The modulation of chemical processes by SWV (line 31) is actually a weak feedback, as explained by Wohltmann et al. (2024). Tian et al. (2023) does not address specifically this statement; I was not able to check Tian et al. (2009) as this reference is incomplete.

Response: Thanks for your suggestion. This sentecen was revised.

- "……and the increase in SWV, for example, has a slight impact on ozone depletion (Tian et al., 2009, 2023; Wohltmann et al., 2024)." (L29-30)

Admittedly, "the seasonality of the water vapor QBO signal has been seldom studied" (line 65). But why is it interesting to study this seasonality? The introduction fails in its primary aim.

Response: Thanks very much for your positive comments. We added the motivation of this study.

- "Serva et al. (2022) found that there are seasonal differences in temperature and SWV in tropical regions. In the northern summer, the temperature at 100 hPa and the WV at 85 hPa reach their peaks, while in winter, they reach their lowest levels (Serva et al., 2022)……" (L70-)

3. Many more details are necessary about the datasets (section 2a). What is their time resolution at download time (i.e. prior to derivation of the anomalies): hourly, daily, monthly? Has ERA-5 been downloaded on model levels or (much less accurate) on pressure levels? What are the uncertainties of SWV in SWOOSH (see comment 1 for

ERA5)? Two generic papers are cited for CMIP6, but what are the specific references about the CMIP6 historical simulations? Could a reference be found about the overall quality of SWV in these simulations?

Response: All those details are added.

- "A horizontal resolution of 1° (latitude) × 1° (longitude) at 37 pressure levels from 1000 to 1 hPa in the vertical direction was collected for this study. The variables used include zonal and meridional wind, specific humidity, and air temperature on pressure levels." (L92-94)
- "In contrast to the troposphere, the WV content within the stratosphere is extremely low. Compared to the SWOOSH satellite observation data, the ERA5 reanalysis data provides a longer time span, which provides more samples for revealing the effect of QBO on SWV." (L104-)
- "Previously, some studies found that the CMIP6 models underestimate the WV content at the bottom of the tropical stratosphere (Keeble et al., 2021; Ziskin et al., 2022)." (L133-134)

4. The base climatology is calculated for the period 1960-2020. Does it make sense two very different periods of ERA5 (compare fig. 1a with fig. 1b) to compute its base climatology? See also comment 2. The period 1960-2020 presumably applies only to ERA5, as the CMIP6 simulations end in 2014. Does it make sense to compare anomalies for different periods?

Response: Thank you for your suggestion. Below 10 hPa, the water vapor QBO signal of ERA5 is relatively consistent in the two periods. We have modified Figure 2 to visually display the water vapor below 10 hPa.

Actually, the filtering does not depend on the climatology.

5. Line 104-105 define criteria of +/- 5 m/s for QBO "events". Are these the criteria used to build the W QBO and E QBO composites? Please clarify.

Response: Clarified.

- "QBO events are selected when the QBO index is greater than 5 (less than -5) m/s to build the westerly QBO (easterly QBO) composites, following previous studies (e.g., Rao et al. 2020a, 2020b)." (L142-144)

6. All water vapor anomalies are shown in ppm units. Does this refer to water vapor volume mixing ratio (a.k.a. mole fraction) or mass mixing ratio?

Response: All figure captions modified to "(mass mixing ratio, units: ppm)".

7. Figure 2 shows only two contour lines for the zonal wind, with no labels. Yet lines 142-147 discuss the numerical values, which are impossible to read from the figure.

Response: Figure 3 caption added. "The contours are shown at ±15 m/s and ±30 m/s." (L193)

8. Similarly for fig.4: lines 179-180 and 187-188 discuss a difference in tropopause pressure between the W QBO and the E QBO, but this can not be seen on the figure. BTW, how is defined the tropopause here? Are you using a thermal, dynamical or SWV-based definition?

Response: Thanks for your suggestion. We have removed the discussion regarding the tropopause height differences. The original Fig. 4 has been moved to the supplementary materials.

9. What is the meaning of the dotted regions on Fig. 4?

Response: Moved to Fig. S2.

- "Dots denote statistical significance at the 95% confidence level based on Student t-test."

10. Figures 12 is not explained at all, and discussed very succinctly with figure 13 and Table 1 (lines 344-359). This should be expanded.

Response: Thanks for your suggestion. In response to your feedback and in consideration of the other reviewers' comments, the original Figures 12, 13 have been removed to reduce the figures.

Additional references

Krüger, K., Schäfler, A., Wirth, M., Weissmann, M., & Craig, G. C. (2022). Vertical structure of the lower-stratospheric moist bias in the ERA5 reanalysis and its connection to mixing processes. *Atmospheric Chemistry and Physics*, *22*(23), 15559-15577.

Simmons, A., Soci, C., Nicolas, J., Bell, B., Berrisford, P., Dragani, R., ... & Schepers, D. (2020). Global stratospheric temperature bias and other stratospheric aspects of ERA5 and ERA5. 1.

Wang T, Zhang Q, Hannachi A, Hirooka T, Hegglin MI. Tropical water vapour in the lower stratosphere and its relationship to tropical/extratropical dynamical processes in ERA5. *Q J R Meteorol Soc*. 2020; 146: 2432–2449. https://doi.org/10.1002/qj.3801

Response: Thanks for your suggestion. We learned and cited the references.