Response to Reviewer 2

Overview

We thank the reviewers for their helpful comments on our study. Based on their comments, the manuscript has been substantially improved since its initial submission. We briefly summarise the improvements here in an overview, before responding to the reviewer's specific comments.

The key improvements to the paper include, but are not limited to:

- A new section investigating equatorial waves, in response to comments by reviewers #2 and #3. This includes (a) new analyses of Kelvin wave variance during the QBO disruption (b) new analyses of the symmetric and antisymmetric power spectra, and (c) a consideration of the equivalent depth h_e and vertical wavelength L_z of Kelvin waves. These results help to clarify equatorial wave processes during the disruption, are consistent with our existing material and other studies, and have improved and expanded the scope of the our study.
- Statistical tests which have been applied to our results, including to the comparison of Aeolus with ERA5 and radiosonde measurements, and to our calculation of the equivalent depth.
- A more comprehensive description of important Aeolus and ERA5 biases which might influence the results in this study, with additional information on these for the reader's benefit.

We now respond to the specific comments by reviewer 2 below.

Reviewer 2

In their manuscript "Aeolus wind lidar observations of the 2019/2020 Quasi-Biennial Oscillation disruption with comparison to radiosondes and reanalysis", the authors observe the evolution of the second ever Quasi-Biennial Oscillation (QBO) disruption in novel Aeolus satellite measurements. The observations are validated with radiosonde and reanalysis data. In addition, the reanalysis and satellite data are analyzed with respect to Kelvin wave occurrence before the disruption.

This study is an excellent show case for the use of novel Aeolus wind measurements for the investigation of dynamic phenomena in the upper troposphere / lower stratosphere (UTLS) region. It clearly demonstrates the capabilities and restrictions of this unprecedented dataset for atmospheric research. With this the article is in general of high scientific significance. Nevertheless, I see a major need for improvement with respect to the scope of the journal and the scientific questions answered by the manuscript.

In its current state the paper is more a validation of the new Aeolus dataset and with this more in the scope of Atmospheric Measurement Techniques. However, by applying some changes and extending the discussion the paper could help to address important scientific aspects of QBO research. This would make it very suitable to Atmospheric Chemistry and Physics.

In detail, the paper would strongly benefit from a more detailed discussion on what can be learnt from the differences between ERA5 and Aeolus for the generation of the disruption. Which processes are different (e.g. Kelvin wave activity) and what does this tell us about our current understanding of the physical processes behind?

We thank the reviewer for this overview and for their recommendations.

Specific comments

L27: Any reference to support this sentence? Maybe Smith et al. (2022)?

Yes, we have added Smith et al. (2022) which supports this claim.

L29: Also, here a reference would provide additional information to the reader, e.g. ESA (2020a)

We have added the suggested reference to the paper.

L36ff: This sentence is really long, contains a lot of information and is hard to read. Maybe better split into two or more sentences.

Agreed. We have split this sentence into two sentences.

L39: What is your study adding to the current knowledge / understanding of the QBO? What are the research questions you are trying to answer? For the reader it usually helps to briefly address the "why exactly this research" in the introduction to easier follow the manuscript.

Our study contains the first direct wind measurements of the QBO using a space-borne DWL instrument. The additional scientific analyses suggested by this reviewer have helped to shed light on why forecast models may have struggled to predict the disruption, which we have implemented by finding key differences in the representation of Kelvin waves between Aeolus and ERA5. We have added some text to the introduction to address the reasons for this research, as suggested.

L46f: Global and local are antonyms, but how does reanalysis fit into the picture? Is a reanalysis perspective really a measurement perspective? I would suggest to reformulate this sentence.

We have reformulated the sentence as suggested by the reviewer. The renanalysis of course gives a global perspective, so we've removed the inference that it is neither global nor local.

Section 2.1: Maybe it is worth mentioning that the Singapore Radiosonde Station is commonly used as the gold standard when it comes to QBO analysis as it provides the longest available data record in the tropical stratosphere. Okay, this is somewhere later in the paper. Maybe you could move it forward to the dataset description.

Agreed, we have moved this sentence into the dataset description as suggested by the reviewer.

L66ff: Is Banyard et al. (2021) really an appropriate reference here? Maybe remove this part of the sentence. The references before already support this sentence sufficiently.

This sentence is a close-quotation from Banyard et al. (2021) since this describes what we intend to say well and we do not wish to rephrase this, so a citation is necessary here to avoid plagiarism.

L71: I would suggest to better reference ESA (2020b) here.

We have added the reference ESA (2020b), although we have left the original reference of Legras (2022) in place since we did use this resource to help give us information about the Hunga-Tonga RBS.

L77f: This is in general correct, but the sentence is confusing here. 35°off-nadir (as mentioned in the sentence before) per se gives mainly the vertical wind component. Only because the vertical wind component is much smaller than the horizontal component, the zonal wind can be derived, which you actually mention, but only one sentence afterwards. So, the order of the sentences here confused me. Maybe just remove this sentence on the advantage for QBO observation?

Yes. We have rephrased this paragraph to avoid confusion.

L108: ERA5 is available at a temporal resolution of 1 hour:
https://www.ecmwf.int/en/forecasts/dataset/ecmwf-reanalysis-v5

We used 3 hour resolution data in order to reduce unnecessary computational load. In practice such a change makes only an extremely small difference to our dataset and does not affect our results significantly.

Figure 1: How did you fill the data gaps in this plot? E.g. between the weekly QBO settings or during instrument down times (e.g. spring 2021).

A 7-day boxcar filter is used to fill the gaps above 20 km following the QBO 2020 RBS change, and where there are data gaps elsewhere, these are filled using a broader 20-day boxcar filter. We have added this text to the figure caption. As this zonal-mean timeseries requires less stringent constraints for accuracy, we have been able to apply the intended restrictions for NWP centers with some flexibility.

L116f: This would probably be better visible if the Aeolus dataset would be extended until launch in 2018.

We have only used data from after the switch from laser FM-A to FM-B occurred in June 2019, since at the time this analysis was conducted, data from the first FM-A data period was not trustworthy. Since then however, data from this period has been reprocessed to remove the existing biases, so if this analysis was done in future we would indeed extend back until launch in 2018 as the reviewer suggests.

L161: Why only for reanalysis?

We have added 'and NWP models' to the text.

*L185: ..., as *mentioned / described* on Kawatani et al. (2016), ...*

We have changed the text as suggested.

L185ff: I had to read these sentences a couple of times, before understanding their meaning. Maybe rephrase?

Agreed, we have rephrased this paragraph to make it clearer, and moved the sentence about the importance of using the Singapore radiosonde into the Data and Methods section.

L193: Why 3D?

Removed to avoid confusion.

L204f: This is a courageous statement. Weiler et al. (2021) and Abdalla et al. (2020) only look into this on global average. According to ESA (2021), there could be higher regional biases. In addition, there are small bias differences between ascending and descending orbits in certain months (also ESA, 2021).

We agree with the reviewer and have added a distinction between global and regional biases in the text. We have also modified the text to change "kept to a minimum" to "kept low".

L206: I fully agree to use negative HLOS in figure 3, but please also state this in the figure or caption.

We have clarified this in the figure caption.

L211ff: You nicely describe the evolution of easterly winds in the profiles. The weakening of westerly winds is also nicely visible in profiles b-e. Maybe you could also guide the reader through this point first, before coming to the evolution of the easterlies. Otherwise, why do you show the profiles b-e?

Yes, we agree that this is important. We have added a sentence to guide the reader through the weakening of westerly winds which occurs before the disruption easterlies develop, so that the full context of the disruption itself is provided.

L219: Aeolus measurements are performed separately for each range bin and, unlike the derivation of temperature and aerosol backscatter and extinction, the derivation of wind from lidar measurements does not rely on an iterative profile reconstruction. Thus, cloud contamination in higher atmospheric bins cannot originate from clouds in lower atmospheric bins. To me it seems more that your limit of 250km is quite a large range (in validation studies often 100km is used as colocation criterion) and the atmosphere varies within this distance. There are many points in your profile comparisons where two or more Aeolus winds in the same range gate at the same time are quite far from each other (not only in figure 3g & 3h). This is probably due to the variability of the wind in latitudinal direction. Maybe the altitude offset in figure 3f and 3i are also due to a not perfect match in location.

Firstly, we concur with the reviewer's assessment of the Aeolus measurement profile retrieval. Regarding the comment on cloud contamination issues, we speculate that the higher random error in the higher atmospheric bins might result from cloud contamination at those altitudes, rather than from the clouds at lower altitudes, which were removed through quality control. Aeolus' along-track accumulation spans approximately 87 km, and data may still be flagged as clear-sky even if some clouds are present. Considering the presence of convection at lower levels, isolated deep convection might also occur along this 87 km track, albeit insufficient to flag high-altitude data as cloud-contaminated.

Secondly, we agree that 250 km is quite a large range so we have redone the analysis for 150 km colocation. We found that this is a good trade off between sufficient data for our analysis and close colocation (The average minimum proximity of the Aeolus overpass to the radiosonde station is $\sim$80 km). A similar radiosonde comparison by Ern et al. (2023), albeit implemented for a slightly different purpose, imposes colocation criteria of $\pm 2^\circ$ latitude and $\pm 10^\circ$ longitude which, although a larger area, should still provide a good comparison of the larger-scale dynamics. Interestingly we find that the altitude offset in figure 3f and 3i is still present in the 150 km colocation results, which suggests it may not be due to latitudinal variability. In addition, the 150 km colocation shows, if anything, a slightly weaker similarity to the radiosonde, especially in figure 3e, likely due to the reduced number of profile points used. Nevertheless, Fig. 3 and 4 have now been updated to constrain data to a maximum of 150 km from the Singapore radiosonde launch site.

Figure 4: Which colocation method is used for the comparison of these data points? Aeolus vs radiosonde, probably the same 200km, but what did you do for the comparison to ERA5? Interpolate the model onto the Aeolus and radiosonde location? This might explain the better agreement between ERA5 and Aeolus vs radiosonde and Aeolus.

We have used the same colocation method as for figure 3, now updated to contain all colocations less than 150 km. The profiles in figure 3 are a selected subset of all the profiles used for figure 4. The ERA5 data has been projected on the HLOS measurements as described in the Data and Methods section.

L251ff: Maybe a good idea for a next study. ECMWF forecast / analysis data with and without Aeolus assimilation is available at ECMWF.

We thank the reviewer for their good suggestion, this is indeed an idea for further study which the authors would be interested in carrying out.

L255: What exactly do you mean with like-for-like comparison? Are you interpolating ERA5 data to Aeolus measurement locations?

Yes, that's precisely what we do. Although other studies may compare data from Aeolus against ERA5 on different grids, since we're focusing on the perspective of Aeolus we have interpolated ERA5 onto each Aeolus measurement location. We have altered the text in the Data and Methods section to clarify this, and added the formula used to convert ERA5 u and v into a synthetic ERA5 HLOS wind. In the text here, we have removed the phrase 'like-for-like' to avoid confusion, and have referred the reader to the ERA5 subsection for clarity.

Figure 5: Maybe you could show the ERA5 contour lines in the whole plot.

We can only show ERA5 contours over the same area as the Aeolus measurements since we are effectively making a comparison between (a) Aeolus measuring the real atmosphere and (b) Aeolus measuring the 'ERA5 atmosphere'. This means that the 'ERA5' dataset has the same spatial and temporal extent as the Aeolus dataset.

L257: You not only see this difference in onset time, you also clearly see stronger winds and wind gradients in the Aeolus data in the troposphere before the disruption (e.g. -5 m/s line is at a higher altitude in Aeolus data). These stronger wind gradients might have an impact on the Kelvin wave propagation (you discuss afterwards), so I think they should be mentioned here.

We thank the reviewer for this useful comment. It has partially motivated some of the further analysis added later in the paper. We agree, and the updated version of figure 5 shows this more clearly. We have added more information about the relationship between these stronger wind gradients and Kelvin wave propagation to the text.

L270: ... highlight only symmetric wind structures ... (antisymmetric waves with respect to the equator are removed due to your averaging from -5° to +5° latitude; for a detailed analysis of equatorial waves in the Aeolus dataset, you could have a look at Ern et al. 2023)

That is correct. We have added this clarification.

L291 – 296: This paragraph mixes different things and draws conclusion which I either do not understand correctly or a very far-fetched. I would suggest to remove the whole paragraph.

We have removed most of this paragraph, and only retained the sentence relating to the comparison between the different QBO disruptions, since we believe that this is particularly relevant to this study.

animation S2: Title of plot and caption are not in line with each other. It should be +-5° latitude as everywhere in the manuscript.

We thank the reviewer for pointing this out. The Hovmoller and cross-section plots are plotted from the equatorial bin of a 5° latitude resolution gridded dataset, so these should both be +-2.5 degrees. A gridded dataset is used to make the analysis easier to run. The zonal-mean time series of the QBO for figure 1 and 5 is for +-5° since this Aeolus data has not been binned in the same way; in order to maximise the SNR and accuracy as much as possible. Re-running code to ensure these match would take a long time and is unlikely to change the timeseries figures significantly.

Animation S2 and Figure 7: Why don't you apply the same temporal filtering as in the Hovmöller plots? You want to show the Kelvin waves, but these are barely visible due to the dominating feature of the Walker circulation. By applying a similar filter as in Figure 6, this strong dipole should vanish and the Kelvin waves should become clearly visible.

We agree with the reviewer here and have updated animation S2 and figure 7 to apply the same temporal filtering as in the Hovmöller plots as suggested. The result of this is that the Walker circulation no longer dominates and the Kelvin waves have become more clearly visible.

L328ff: You could perhaps reformulate the sentence to stress the importance of a future wind lidar measuring up to at least 30km or higher for QBO research.

We have added this to the text and given two references in support of this.

L336: ... this change in random error ... (the high random error itself is a problem especially for short analysis periods and perturbation analysis, for these a bias would be less of a problem)

We have edited the text to add this, although this paragraph has now been moved to the Data and Methods section.

L348: Why only reanalysis model?

Yes, this is the case for NWP models as well. We have added this to the text.

L349: Why is your analysis not spanning the whole Aeolus measurement period, so why is the data before summer 2019 missing?

As mentioned earlier, we have only used data from after the switch from laser FM-A to FM-B occurred in June 2019, since at the time this analysis was conducted, data from the first FM-A data period was not trustworthy. Additionally, our analysis does not span the whole Aeolus measurement period for a very important reason, which is that we are focused primarily on the QBO disruption of 2019/2020. It is important that the focus of the paper does not shift away from the disruption (and therefore relevant for publication in ACP) towards a more general techniques paper on the characteristics of Aeolus measurements in the tropical UTLS (which would be more relevant in AMT). Please see comments by reviewer #3 for more details.

Discussion: This is not a general discussion of the results of the study. What you describe here are drawbacks of the Aeolus mission. Thus, I would suggest to either rename the section or revise its content.

We agree with the reviewer that there is an overemphasis on the drawbacks of the Aeolus mission here, and so we have revised the contents of the discussion section accordingly. We think it is important to still mention key limitations of the study which may affect the results, but some of the content has been moved to the Data and Methods section, partially in response to a comment by reviewer #1 as well.

L379f: Why is this in agreement with the data validation?

We agree that this was confusing, and since we discuss the data validation in the previous paragraph, we have removed this clause for clarity.

L385: Maybe better: ... have been discussed.

We have changed this as suggested.

L389: Maybe good to stress here the importance of measurements up to at least 30km.

We have added this to the Discussion in response to an earlier comment by this reviewer.

Conclusions: You nicely describe the data, but what have we learned from a scientific point of view (except that Aeolus is a well-suited dataset for observing wind related phenomena in the UTLS)? Why did it come to this QBO disruption? Where is the difference between ERA5 and Aeolus and what can we learn from this difference for our understanding of the underlying physical processes? These are questions I would expect to be answered in an ACP manuscript.

We thank the reviewer for this comment, which has motivated much of the additional scientific analysis we have conducted. We have learned that there are key differences in the vertical wind shear and Kelvin wave variances between ERA5 and Aeolus which help to explain the observed lag in disruption onset in the reanalysis. We have therefore made the manuscript more suitable for publication in ACP due to the greater focus on the dynamics of the QBO disruption.