Response to Reviewer 3

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 3 below.

Reviewer 3

The manuscript by Banyard et al. presents wind measurements of the 2019-2020 QBO disruption from the first spaceborne Doppler wind lidar ADM-Aeolus. The general topic is relevant for publication in ACP and the paper is interesting, demonstrating that the QBO is generally well captured in Aeolus data. However, I find the manuscript very focused. It would benefit from including a broader analysis and discussion of the new information brought by Aeolus on tropical lower stratospheric dynamics compared to a state-of-the-art reanalysis. Some differences between the ERA5 and Aeolus are already mentioned but not much commented.

For these reasons, major revisions (additions) are required before the paper can be considered for publication. My recommendation to the authors would be to make more quantitative statements and expand the comparison between Aeolus and ERA5 to other modes of variability than the QBO (e.g., equatorial waves) which are only alluded to in the present manuscript. The authors should also consider the recent preprint by Ern et al. (ACPD) on a related topic.

We thank the reviewer for their comments, and in response we have made significant changes to make the paper more quantitative and have expanded the comparisons, as described in the Overview at the top. We have considered and referenced the recent preprint by Ern et al. (ACPD) where appropriate.

Main comments

1) Lack of significance test : in a few instances the author claim that a bias/difference is significant, but do not include a statistical significance test. For instance the near-tropopause wind bias at Singapore (l238-240) is not clear to me, since it is also present in the ERA5-Singapore radiosonde comparison.

The authors agree that a statistical test is useful to quantify, for example, the near-tropopause wind bias at Singapore (L238-240), and we have included a Student's t-test of p<0.001 to highlight altitudes where the bias is statistically significant. A t-test is suitable here since the distributions are near-normal (as shown in the violin plots). As mentioned in the response to reviewer

#1's comment at L208, we suggest that much of the tropopause bias seen in Fig. 4b is a result of biases in ERA5, rather than solely being a consequence of biases in Aeolus. This is particularly likely since a bias is also seen between ERA5 and the radiosonde, as reviewer #3 mentions here.

2) The authors touch the topic of equatorial wave representation in ERA5, but they could make more quantitative statements. They might also wish to change the colormap used for Figure 6. Given that one of the advantages of Aeolus is global sampling, the authors could without much effort quantify the differences as a function of longitude, beyond the location of Singapore. This is of interest, in particular since earlier studies (e.g., Baker et al., 2013 ; Podglajen et al., 2014) found that reanalysis uncertainties and wind errors in the tropical UTLS were larger in regions without assimilated radiosonde observations, resulting in a pronounced zonal structure. Note that the need for wind observations to better constrain the flow in tropical regions is what motivated Aeolus in the first place.

The authors agree that the topic of equatorial wave representation should be explored further, especially in the context of the disruption and with more quantitative statements. To this end, a new section has been added examining the differences in equatorial wave spectra between Aeolus and ERA5 and differences in Kelvin wave equivalent depths with height. We have considered the reviewer's suggestion of further analysis of the differences between Aeolus and ERA5 as a function of longitude; however the recent preprint by Ern et al. (2023, ACPD) covers this longitudinal variation extensively and we have thus taken a different path, whereby we focus on the QBO disruption and Aeolus' capability to capture it and the QBO. The colormap for the middle panel of figure 6 has been changed to make the structure clearer.

3) Regarding the last sentence of the abstract : "This analysis highlights how Aeolus and future Doppler wind lidar satellites can deepen our understanding of the QBO, its disruptions, and the tropical upper-troposphere lower-stratosphere region more generally." , it is not shown in the paper that Aeolus can help clarify the mechanisms of the disruption, at least directly. I imagine spaceborne lidars would help deepen the understanding of the QBO by putting observational constraints on equatorial waves and tropical gravity waves, but this is not really the focus of the paper.

The authors have made changes to the paper to make it more suitable for publication in ACP by exploring the role of equatorial waves in the development of the disruption to a greater depth. This better justifies this sentence in the abstract compared with previously, so we have kept it as in the original manuscript.

Line 1 : The abstract could mention some key number (bias, etc.)

We have added this information to the abstract, along with the key results from the new material.

Line 102 : please describe the method briefly here

We have moved these details a little earlier in the section and modified the text so that it describes the correct method.

Line 238-240 : Is this a significant bias ? It does seem smaller than the difference between Aeolus and the radiosonde (panel a), which by the way, also maximizes near the tropopause.

As we have previously mentioned, we have performed a Student's t-test on these results to demonstrate that the bias is statistically significant. This is the case both for the difference between Aeolus and the radiosonde (panel a) and the differences between the observing instruments and reanalysis (panel b and c). Though significant, this does not give any direct indication as to the cause of the biases in relation to each dataset, i.e. it is not possible to say that one bias is the addition of the other two since the distributions are not additive.

Line 264-265 it is suggested that the sampling by aeolus does not induce much bias. You could easily prove this with a figure comparing ERA5 with aeolus sampling and the actual zonal mean

The biases in ERA5 that are caused by regional differences in observation sample sizes are discussed by Kawatani et al. [2016] and are mentioned earlier in our study. Given that Aeolus' data sampling is mostly homogeneous (aside from data gaps due to instrument downtime or quality controls), we believe it is fair to make this assumption. We have not added the suggested figure since we do not think it adds much strength to the paper. Mapping the locations and density of input observations for ERA5 would take quite a bit of work to complete, and is outside the scope of this study.

Fig 4 : the pressure scale is wrong in this figure. What do the dots correspond to? Some of the numbers in the legend would find their place in the main text.

The pressure scale has been corrected and the dots have now been referenced in the figure caption. We are not sure which legend the reviewer is referring to, but we have placed all relevant numbers off the figure in the caption. Median differences and standard deviations have been updated to correspond to the new 150 km constraint.

Fig 5 : Could you rather plot the mean wind in contour and the difference in colors ? This would be more quantitative. There seems to be a delay, does it hold for the later period or is this just a feature of the disruption?

The authors agree with the suggestion by the reviewer and have changed the plot so that it shows the mean wind for Aeolus and ERA5 in contour and the difference between them in colour. In general, Aeolus has stronger winds than ERA5, but the zero wind line does not always follow the line of zero difference.

Fig. 6 : Is there a zonal-mean structure in the error (see main comment 2) ? Also, I do not understand the need for time filtering. Eastward propagation appears clearly in your unfiltered plot.

We disagree with the reviewer that time filtering is not necessary to more clearly show the eastward propagation of waves, since in the unfiltered plot the standing wave of the Walker Circulation dominates and masks the Kelvin waves. (There is also a contradictory comment from reviewer #2)

lines 271-277 : The phrasing is a bit odd, suggesting that this long period was not an optimal choice for Kelvin waves. You cite a few papers which show that the typical stratospheric Kelvin waves commonly seen in Hovmoller diagrams indeed have planetary wavenumber 1-2, periods of 10-20 days or more, as first described by Wallace and Kousky and reported in many papers and textbooks. Convectively-coupled Kelvin waves in the OLR are higher frequency but this is not what dominates at tropopause altitude where free-travelling waves are prominentl. I would shorten this discussion.

The reviewer makes an important point, and we have rephrased this discussion accordingly. The distinction between convectively-coupled Kelvin waves in the OLR (and therefore covering the large altitude range of 10-18 km as mentioned by Alexander et al. [2008]), travelling with periods of 7-10 days, and free-travelling waves in the TTL and stratosphere, with a larger range of periods, is an important one to make. We have tried to keep the discussion concise whilst still mentioning the details we think are relevant. Our choice of filter is optimal for the dominant wave periods seen in the Aeolus data, and is broad enough to capture the vast majority of the Kelvin wave spectrum at these altitudes.

References

S. P. Alexander, T. Tsuda, Y. Kawatani, and M. Takahashi. Global distribution of atmospheric waves in the equatorial upper troposphere and lower stratosphere: Cosmic observations of wave mean flow interactions. *Journal of Geophysical Research: Atmospheres*, 113(D24), 2008. doi: 10.1029/2008JD010039.

Y. Kawatani, K. Hamilton, K. Miyazaki, M. Fujiwara, and J.A. Anstey. Representation of the tropical stratospheric zonal wind in global atmospheric reanalyses. *Atmospheric Chemistry and Physics*, 16(11):6681–6699, 2016. doi: 10.5194/acp-16-6681-2016.