

Answer to Reviewer #3 Comments on “Stratospheric water vapour and ozone response to different QBO disruption events in 2016 and 2020” by Mohamadou Diallo et al.

Dear Editor-in-Chief, S. Fadnavis,

We are submitting our revised article titled “Stratospheric water vapour and ozone response to different QBO disruption events in 2016 and 2020”. We thank the three Reviewers for their detailed and well thought-out comments, which helped to significantly improve the paper. We made substantial changes to the manuscript in order to thoroughly address the Reviewers’ suggestions and comments. The main changes concern:

- Merging of the former Fig S3 with the Figure 1 and Figure 2 in the manuscript
- Moving the Fig S4 into the manuscript as suggested by Reviewer #1 & Reviewer #2 and the related discussion.
- Redone all figures to change wind contours, increase font size and to improve their quality.
- Rephrasing of certain paragraphs in order to clarify the manuscript.

With these changes, we are convinced that the paper is highly relevant for a wide-ranging journal like *Atmospheric Chemistry and Physics*. Please see below our answers point by point to all reviewers’ comments and suggestions.

Reviewers comments are in bold, followed by our respective replies. Changes in the manuscript are in blue, allowing them to be tracked easily.

Kind regards,

Mohamadou Diallo (on behalf of the co-authors)

Reviewer #3 (Comment on acp-2022-382):

Diallo et al. present results from a multiple regression analysis of water vapor and ozone data from MLS and wind and temperature fields from ERA50 spanning 2013–2020 to delineate the impacts of QBO from other natural variations (e.g. El Niño) and from time-varying forcings (i.e. aerosol optical depth). They find distinct planetary wave forcing patterns corresponding to each of the two QBO disruptions, and ascribe the anomalously moist lower stratosphere during the 2020 disruption to Australian wildfires. I have concerns about the robustness of the multiple linear regression and the statistical methods used on such a short time series. I think the paper itself could use editing for flow, clarity, and language. The figures show much more information than is actually discussed, making it a chore for the reader to discern the meaningful results.

Thanks for this critical, but very constructive general comment. We did a substantial rewriting of several parts in order to enhance clarity and language (as also explained in more detail in the responses to reviewers 1 and 2). Regarding the regression there is likely a misunderstanding, as we don’t apply the regression just to the few years shown, but to the entire 2005-2020 period (right??). Hence, we think that the regression results are robust. More specific responses are given to the more specific comments below.

General Comments:

Does the paper address relevant scientific questions within the scope of ACP?

Yes. The objective of this paper is important and germane to ACP.

Does the paper present novel concepts, ideas, tools, or data?

Yes. The multiple regression analysis including both the recent QBO disruptions is timely and useful.

Are substantial conclusions reached?

I think so; the location of the eddy forcing is helpful to know, and the attribution of moist UTLS during the 2020 disruption to the Australian wildfires – an impact that has been discussed in the community – is important to show.

Are the scientific methods and assumptions valid and clearly outlined?

I don't believe they are clearly outlined. I think the multiple regression analysis on 8 years of data – dealing with oscillations that are subseasonal-to-nearly-interdecadal in scope – may be over-determined. Furthermore, while the authors state that a t-test is used to test for statistical significance, they do not specify the parameters used (e.g., the effective degrees of freedom, which I believe to be important in a time series that is temporally over-sampled with respect to the QBO, ENSO, etc.).

Thanks for these remarks. This comment is very likely a misunderstanding of the Reviewer. In Page 4, lines 105-110, we have described that the regression analysis that is applied for the 2005-2020 period e.g. 16 years of data, therefore, there is not issues of temporally over-sampling. In the figures, only the 2013-2020 period is shown to highlight the impact of the two QBO disruption events. The relative changes shown here give us the relance of the QBO effect on these trace gases.

Are the results sufficient to support the interpretations and conclusions?

I don't believe so. Again, I think – considering the shortness of the MLS record – that the multiple regression analysis used here is likely to appear more significant than it is. Also, there is an implicit assumption in the test that the QBO disruptions are well described by a one-dimensional QBO index time series. In other words, that the impacts of the QBO are the same as the impacts of any other transition between QBO phases with a similar index. If this is indeed the case, the QBO disruptions may not be so interesting. If this is not the case, then the interesting impacts of QBO disruptions seem more likely to appear in the residual term (epsilon), but this is not examined in detail.

As explained above, the regression covers the entire 15 years of MLS measurements 2005-2020 (actually 16 years). We agree that it would be better and increase robustness if we'd use an even longer time series, but unfortunately the MLS observations are restricted to that period. We also agree that the linear regression method is not capable of correctly separating non-linear effects. Nevertheless, our MLR results with an automated lag-time show a rather consistent picture between composition and dynamical changes, which in our opinion shows that the MLR method provides meaningful results here.

Is the description of experiments and calculations sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)?

I don't believe so, for reasons stated above. Also, Figure 3E is said to capture the aerosol optical depth (AOD) impacts of the Australian wildfires, but since there are only two disruptions, of differing character, it seems possible that some of the impact of the QBO disruption itself may be wrapped up in the AOD term, in addition to the residual term.

Again, I'd suggest to argue with the meaningfulness of our results, the consistency between composition and dynamics variability. As far as we know there was not a wildfire during the 2015-2016 disruption. Figure 3e shows only the AOD impact. In addition, the regression is suitable and well established and has been applied in a previous aerosol impact study (Diallo et al. 2017, GRL).

Do the authors give proper credit to related work and clearly indicate their own new/original contribution?

I believe so. The authors cite a good variety of work, including very recent work, and the authors' own related contributions. It seems strange to me that Taguchi (2010) is being cited to back up the statement that there is not yet a clear understanding of how QBO disruptions are linked to anomalous sea surface temperatures, since QBO disruptions had never occurred when that paper was written.

We have rephrased the sentence.

Does the title clearly reflect the contents of the paper?

Yes.

Does the abstract provide a concise and complete summary?

Yes

Is the overall presentation well structured and clear?

The paper could use some work in this regard. Paragraphs are rambling, having at times several different subjects. There are no subsections to aid quick navigation of the paper.

We have rephrased many paragraph in order to improve clarity.

Is the language fluent and precise?

The paper needs editing to be publication ready. Like the paragraphs, the sentences are long and often have weak structure. Definite/indefinite articles are often missing, singular/plural disagreement is prominent. Many other phrases (e.g. "the quasi-periodic QBO cycle of about 28-month period" on line 37) are just a bit awkward and could use careful reading by a native English speaker.

As said above, we did a substantial effort of rewriting to improve language and clarity. We think that the text is substantially improved now.

Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?

Yes.

Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?

I think many of the figures show 2D anomalies, but only certain levels or time periods are described. Perhaps reducing the dimensionality of the figures would make the authors' message clearer. Figure production is rough, with uniformly small labeling for ticks labels, and figure panel titles. The figure lettering is added after the fact, is large, and is removed from the figure panel title. Some figure labels are missing. All color scales are the same.

We have increased the ticks labels and figure panel titles also suggested by the reviewers 1 and 2. We think that the visibility and clarity of figures is much improved now.

Is the amount and quality of supplementary material appropriate?

I think the supplementary figures are of similarly rough quality to the figures in the main text. The figures I would be most interested in (some of the other regression terms) are not shown.

The main focus of the paper is on the differences in BDC and the impact of the QBO disruption on H2O and O3.

A few small edits/suggestions follow:

1. **Line 7: "circulation anomalous responses" needs re-wording.**
We have rephrased the sentence.
2. **Lines 9–11: this sentence needs to be split or otherwise clarified.**
We have rephrased the sentence.
3. **Line 31: "Considered as a..." could just be "Considered a..."**
I am not a native speaker, but starting the sentence with just "Considered a ..." does not sound odd to me, and could be used.
4. **Line 37: "The quasi-periodic QBO cycle of about 28-month period" needs rewording.**
We have rephrased the sentence following the suggestion of Reviewer 2.
5. **Line 44: "study" should be "studies".**
We have rephrased the sentence by removing the citation of Osprey et al (2016) as the study did not use CMIP6.
6. **Line 50: comma needed after amplitude.**
We leave it as it is.
7. **Line 74: "planetary (PWD) and gravity (GWD) wave drag" would be better as "planetary wave drag (PWD) and gravity wave drag (GWD)." It's just two words more, but much easier to read.**
We have rephrased the sentence.
8. **Line 99: "In in"**
We have rephrased the sentence.
9. **Lines 104–106: Sentence beginning with "Both" is then used to contrast the two QBO events. Re-word.**
We have rephrased the paragraph.
10. **Lines 161–163: This sentence is confusing. I keep reading "fits" as a verb, and it breaks everything. Make it clear.**
We have rephrased the sentence.
11. **Lines 282–285: The last sentence of the previous paragraph and the first sentence of the subsequent paragraph are a bit redundant—they both serve to introduce the topic of gravity wave drag.**
We think that the first sentence is for transition and the second for actually underlying what we do. We keep the sentences for clarity.

12. **Line 327: The word “finally” is repeated. Check how many times it’s used. Use it once.**

We have checked the use of the word “finally”. We used it only 5 times in the whole manuscript, and in different context. Therefore, we keep the wording as is.