Reviewer 3

The authors have done a lot of detailed work to compile these results. The study covers both model and observational parts. However, much of the text describes the figures extensively without giving the reader a clear road map that how these results differ importantly from previous studies. The repetition of text can be seen at some places. My suggestion is for the authors to perhaps shorten the description of results to reflect only their essential messages.

Thank you for your incisive comments, which led to a significantly improved manuscript.

I have the following major concerns before any recommendation on the manuscript:

1. Sometimes, the description of results is too extensive and needs to be shortened. The description related to the mechanics in the results sections should be moved in the discussion section. The repeated text requires curbing. Sometimes, the figures are described randomly within the results section.

   Done as suggested.

2. In the observation part, the authors are only confirming the finding of Taguchi (2010) with the same methodology. I could not observe any advantage of their observation analysis in compare to Taguchi (2010).

   We adopted a different methodology from that of Taguchi (2010).
   Since our preprint was designed to avoid criticizing Taguchi (2010), the advantage was not perceptible to you. Our revision has made clear why our method is advantageous.

3. This study is using 5-month moving averaged deseasonalized data. Is such huge smoothed data (a nearly half-year window) suitable to study the gravity wave generation? I suggest using the monthly data instead of the 5-month moving average. Using the monthly will also be an advantage of your study against Taguchi (2010). A separate section is required for the data and method; currently, there is mixing of data information and result description.

   Done as suggested.

4. All the figures are plotted very causally and not suitable for publication. There is a lot of scope for improvement in almost all the figures. Please see the specific comments for details.

   Please take a look at our uploaded figures rather than the ones inserted in the preprint. We will explain this issue in more detail.

Specified comments & suggestions:
L47-48 "the tropospheric subtropical jet (Garfinkel and Hartmann, 2011a, 2011b)..", can be updated with more recently citation (e.g. DOI: 10.1029/2022JD036691).

Done as suggested.

L48 "the boreal summer monsoon (Giorgetta et al., 1999)", can be updated with more recently citation, Yoden et. al 2023 which shows the QBO modulation on global monsoon system (https://doi.org/10.54302/mausam.v74i2.5948).

Done as suggested.

L103-108: I suggest to add some sentences related to the motivation of this study.

We have added a paragraph before this part.

L121 "We further smooth the deseasonalized zonal winds using a 5-month moving average (for more details, refer to Taguchi, 2010)". It will be better to use the monthly deseasonalized zonal winds instead of the 5-month moving average.

We have adopted your suggestion. Namely, we have employed the monthly anomalies in our observational study.

L 132, I suggest to the authors that the data can be extended for seven more years, i.e., 1953 to 2022.

Anstey et al. (2021) pointed out that the first two EOFs explain no more than 60% of total variance during the 2016 and 2019/20 QBO disruptions. Those two QBOs deserve more studies separately. As to whether outliers should be deleted or not, it is always a controversial issue. We prefer to erring on the cautious side.

L139-142, Why the different base periods? Only one base period can be used, i.e., a de-seasonalized anomaly for the whole time period of the data set.
L149 "identified 21 El Nino and 15 La Nina events between 1953 and 2015". Definitely, using monthly data, the El Nino and La Nina events will increase by two to threefolds. The same can be applied on the model part too.

Those issues are related to climate change and climate variabilities.

1. Why has CPC of NOAA adopted this criterion for many years?
There must be a raison d'être!
Let's conduct a thought experiment by extending the 1pctCO2 CMIP6 experiment (1850-2014) for another one thousand years. What will happen if we only adopt the whole period of model outputs as one base period?
Using current criterion of ENSO events, we will find that in the first 100 years, every month of every year probably falls into La Niña category and that in the last 100 years, every month of every year probably falls into El Niño category because of global warming!
Please refer to Fig. 6 in Latif and Keenlyside (2009) to get a taste of it.

2. Even if we could keep the CO2 concentration at the current level forever, we still need to use the method adopted by CPC of NOAA!
The following is our explanations.
Chapter 8 of Hartmann (2015) details "Natural intraseasonal and interannual variability". Apparently, monthly mean SSTs over the Nino3.4 region does include natural intraseasonal and interannual variability. Adopting your suggestion "Definitely, using monthly data, the El Nino and La Nina events will increase by two to threefolds" will lead to conflating intraseasonal variabilities with the ENSO. It is not acceptable.
Furthermore, adopting the whole period of model outputs as one base period will leads to conflating the PDO with the ENSO. It is not acceptable either. Please also refer to Rao et al. (2019).

You appear to dislike the practice of CPC of NOAA with regard to the definition of the ENSO, i.e., you seem to strongly object the widely adopted method: filtering out the intraseasonal and interdecadal variabilities in order to define the ENSO.

This is surprising given the well-known fact that if Lewis Fry Richardson had applied filtering to his data, he would have fulfilled his dream of numerical weather prediction 100 years ago (Lynch, 2006).

Even now, various national weather centers still conduct various filtering on daily basis (Houtekamer and Mitchell, 1998). In his chapter 7 "Filtering and Data Assimilation", Sullivan (2015) pointed out "it is not bigotry to be certain we are right; but it is bigotry to be unable to imagine how we might possibly have gone wrong".

L143 "CDC" to "CPC"

Corrected.

L144 As suggested in the comments line 121, if authors consider the use of monthly data, then monthly ONIs can be used to define periods of El Niño and La Niña whenever it exceeds the threshold values ± 0.5 K (+ El Niño, − La Niña). Sometimes the SST lies in ENSO phase for 2 to 3 months, and the generation of gravity waves for such a short period will be washed out in the 5-month moving average and cannot be ignored.

We have adopted your suggestion. Namely, the monthly FUB zonal wind anomalies are used. We even skipped deseasonalizing the FUB data.

L154-156 Not a justified reason. If we go beyond the 2015 period, the QBO disruptions (2016 and 2019/20) will not have a significant impact on the total variance of the leading two EOFs. In our own analysis for the period 1979–2022, the two leading EOFs account for 94.73% of the total variance (58.07% by EOF1 and 36.66% by EOF2). If authors are worried about these QBO disruptions, then the time period of disturbance can be excluded if lies in El Nino and La Nina sampling.

This has already been answered.

L157 Instead of two data sets (FUB and ERA5), the authors may also think of using only the ERA5 data for all observational analysis.

Pawson and Fiorino (1998) pointed out that the QBO westerlies from both NCEP–NCAR and ERA-15 reanalyses were generally weaker than the radiosonde winds at Singapore (1.8°N, 104.8°E) at 30 hPa and below.
This seems to be also applicable to the QBO winds from ERA5 presented by Pahlavan et al. (2021). Thus, the FUB data are superior.

L168: This is programming language. Proper mathematical expression should use here.

According to Wikipedia, atan2 is a well-grounded mathematical function. Specifically speaking, the webpage points out that atan2 ranges from $-\pi$ to π while arctan ranges from $-\pi/2$ to $\pi/2$, which is the very reason why we have adopted atan2.

L207: If possible, unit "radians/month" to Km/month. Same in sequent text.

We haven't found any relevant reference to address this issue.

L241: This is a clear mixing of analysis description and data information. The data information should be in a separate section.

Corrected.

L336 -345: As mentioned in the above comment (L139–142), why the different base periods to calculate the SST anomalies?

Already answered.

L356: For the reader's convenience, it will be nice to include the Fig.1 EOFs vectors in Fig.4 also (same line format as in the Fig.1).

We have merged Fig.1 and Fig. 4.

L530 -532: "Comparing Figs. 8a and 8b with Figs. 3a and 3b". I suggest to add one more row at the bottom of Fig. 8 for the difference between the model and observed amplitude of QBO during El Nino and La Nina.

Done as suggested.

L 588-590: The description of the ERA5 reanalysis should be moved into the new suggested data and method section.

Done as suggested.

L593: Is the composite difference in Fig.11c passes the statistically significant test (>95% confidence)?

Since the sample spaces consist of monthly data, we cannot know the effective sample sizes of the El Niño and La Niña sample spaces. Thus, it is a bit hard to rigorously conduct a significance test.

L604-628: This paragraph can be rewritten more precisely by focusing on the comparison between ERA5 and models.

Done as suggested.

L687-693: This paragraph seems unfit here and can be shifted to an appropriate place. The discussion part can start with Paragraph 2.

This paragraph is the punchline of our paper, which is also emphasized in the abstract.

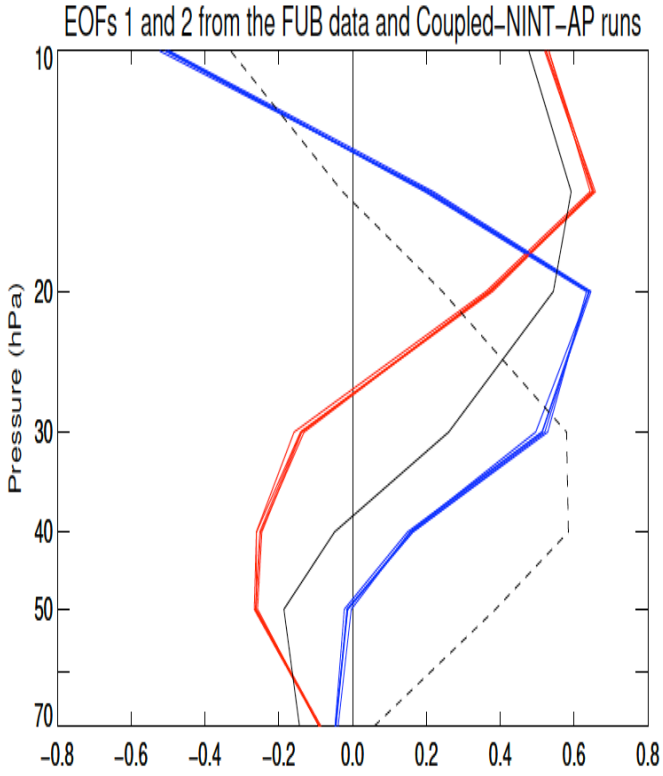
Kawatani et al. (2019) and Serva et al. (2020) pointed out that parameterized gravity waves are either unable to simulate the ENSO modulation of the QBO or harmful to simulating this modulation in high-resolution models. This paper shows that the properly parameterized gravity waves can simulate the ENSO modulation of the QBO even though the horizontal resolution is not high. Since horizontal resolutions in many climate models are not high and those models still employ various parameterization schemes of gravity waves, our results might be useful for the further improvements of these models.

Comments of Figs.

All the figures are plotted causally and not suitable to be considered for publication.

Fig.1 It is too elongated along the x-axis, for the best view, the aspect ratio X:Y should be ~=1:1.3

Note that we have also separately uploaded all individual figures which are plotted as vector images. Vector images can be arbitrarily resized, rescaled, and reshaped infinitely without losing any image quality. Those uploaded figures are ideal in that the professionals of ACP can design the layout of each page in any way without worrying about the issue of losing resolution. The following rescaled figure has the aspect ratio ~1:1.3 (X:Y) as you suggested.



Note that we arbitrarily inserted those figures into the manuscript in such a way that the reviewers can read them clearly. We ignored such things as golden ratio, believing that they are the job of the ACP staff.

Fig.2 This figure is also too elongated along the x-axis, the aspect ratio X:Y should be =4:1.

Same as above.

I suggest to interchange the panels (a) and (b) for the systematic representation, i.e., top (a) should be for El Nino and (b) La Nina and (c) same (El Nino- La Nina).

You look at this issue from an aesthetical point of view.
However, we regard it as a cognitive issue because the pattern of El Nino is almost identical to that of (El Nino - La Nina). The repeated pattern will facilitate our visualization and enhance our memory.

It is an inconvenience to compare all the panels in the current color scale as it is different for different panels. The color bar should have the same scaling on both the positive and negative sides for all panels (here for this figure -45 to 45 Wm-2).

We have adopted the “substance over form” principle.

Had your suggestion been adopted, we would have missed the very important information in each panel: maximum and minimum values of OLR anomalies.

Currently, panel (a) and Panel (b) share some common color bars to the fullest extent. It is very easy for us to compare and contrast them:
(a) From the first order of approximation, the negative OLR anomalous pattern during El Niño is roughly the mirror image of the positive OLR anomalous pattern during La Niña.
(b) Further looking at the maximum and minimum values in panel (a) and panel (b), we can spot the implied asymmetry between El Niño and La Niña. Namely, the amplitude of El Niño is stronger (up to 4.5°C, as measured by the spatially averaged SST anomalies over the eastern equatorial Pacific) than that of La Niña (up to −3°C). This issue has been studied extensively such as Rao and Ren (2014) and Zhao and Sun (2022) concerning the CMIP5 and CMIP6 models, respectively. Please also refer to Timmermann et al. (2018).
Due to the first-order symmetry, it is natural that the contour interval is adopted as two times that used in Panels (a) and (b).

The fine and coarse contour intervals can be used for visualization of smaller and larger signals (e.g. please see Fig. 21 of Hitchman et. al 2020, https://doi.org/10.2151/jmsj.2021-012). The same can apply for other color figures too.
There is no need for more fine contour intervals. Adding several dozens of more contours will make Fig. 2 flashier and less elegant. Einstein pointed out: “Everything should be made as simple as possible, but not simpler”.

Fig. 21 of Hitchman et. al 2020, (https://doi.org/10.2151/jmsj.2021-012) is a great figure!
Note that it is a raster image. In other words, if you rescale Fig. 21 many times larger it will become blurred with many pixels showing up. This is the situation where the right aspect ratio X:Y should be taken into account.
When you rescale our uploaded figures, you will never encounter such a problem. We would like the ACP professionals to decide how to rescale our figures.

Fig.3 Same comments as for Fig.2

Same as the reply to the comments on Fig.1.

Fig.4 Same comments as for Fig.1

Same as above.

Fig.5 Same comments on color scale as for Fig.2. The representation of panel numbers should have the same order in the caption of all figures. In Fig. 3, it is before starting the description [see L 1111 (a) La Nina and (b) El Nino] but here it is after starting the description [see L1152 La Nina (a) and El Nino (b)]. The same corrections must be applied for other figures too.
Fig.6 Same comments on color scale as for Fig.2.
Fig. 7. The aspect ratio X:Y should be similar to Fig.5 and 6. Same comments on color scale as for Fig.2.
Fig.8 too compressed along x-axis. the aspect ratio X:Y should be =4:1.
Fig.9 and 11: Same comments as for Fig.2.

Fig. 12, and **Fig. 13**: Same comment on the color scale as for Fig.2. Fig13. the aspect ratio X:Y should be =4:1.

Those similar issues have already been addressed in the above replies.

References

Anstey, J. A., Banyard, T. P., Butchart, N., Coy, L., Newman, P. A., Osprey, S., and Wright, C. J.: Prospect of Increased Disruption to the QBO in a Changing Climate, Geophys. Res. Lett., 48, e2021GL093058, https://doi.org/10.1029/2021GL093058, 2021.

Hartmann, D. L.: Global physical climatology, International Geophysics, Elsevier Science, pp498, 2015.

Houtekamer, P. L. and Mitchell H. L.: Data assimilation using an ensemble Kalman filter technique, Mon. Weather Rev., 126, 796–811, https://doi.org/10.1175/1520-0493(1998)126<0796:DAUAEK>2.0.CO;2, 1998.

Latif, M. and Keenlyside, N. S.: El Niño/Southern Oscillation response to global warming, P. Natl. Acad. Sci. USA, 106, 20578–20583, https://doi.org/10.1073/pnas.0710860105, 2009.

Lynch, P.: The Emergence of Numerical Weather Prediction: Richardson's Dream, Cambridge University Press, pp292, 2006.

Pahlavan, H. A., Fu, Q., Wallace, J. M., and Kiladis, G. N.: Revisiting the quasi-biennial oscillation as seen in ERA5. Part I: description and momentum budget, J. Atmos. Sci., 78, 673–691, https://doi.org/10.1175/JAS-D-20-0248.1, 2021

Pawson, S. and Fiorino, M.: A comparison of reanalyses in the tropical stratosphere: Part 2: The quasi-biennial oscillation, Clim. Dynam., 14, 645–658, https://doi.org/10.1007/s003820050247, 1998.

Rao, J. and Ren, R.: Statistical characteristics of ENSO events in CMIP5 models. Atmos. Oceanic Sci. Lett., 7, 546–552, https://doi.org/10.3878/AOSL20140055, 2014.

Rao, J., Garfinkel, C. I., and Ren, R.: Modulation of the Northern Winter Stratospheric El Niño–Southern Oscillation Teleconnection by the PDO, J. Climate, 32, 5761–5783, https://doi.org/10.1175/JCLI-D-19-0087.1, 2019.

Sullivan, T. J.: Introduction to Uncertainty Quantification, Springer, pp354, 2015.

Timmermann, A., An, S.-I., Kug, J.-S., Jin, F.-F., Cai, W., Capotondi, A., Cobb, K. M., Lengaigne, M., McPhaden, M. J., Stuecker, M. F., Stein, K., Wittenberg, A. T., Yun, K.-S., Bayr, T., Chen, H.-C., Chikamoto, Y., Dewitte, B., Dommenget, D., Grothe, P., Guilyardi, E., Ham, Y.-G., Hayashi, M., Ineson, S., Kang, D., Kim, S., Kim, W., Lee, J.-Y., Li, T., Luo, J.-J., McGregor, S., Planton, Y., Power, S., Rashid, H., Ren, H.-L., Santoso, A., Takahashi, K., Todd, A., Wang, G., Wang, G., Xie, R., Yang, W.-H., Yeh, S.-W., Yoon, J., Zeller, E., and Zhang, X.: El Niño–Southern Oscillation complexity, Nature, 559, 535–545, https://doi.org/10.1038/s41586-018-0252-6, 2018.

Zhao, Y. and Sun, D.-Z.: ENSO asymmetry in CMIP6 models, J. Climate, 5555–5572, https://doi.org/10.1175/JCLI-D-21-0835.1, 2022.