Authors' response to the Reviewers' comments on "QBOi El Nino Southern Oscillation experiments: Overview of experiment design and ENSO modulation of the QBO"
by Y. Kawatani et al.

Corresponding author: Yoshio Kawatani (kawatani@ees.hokudai.ac.jp)

We are grateful to the two official referees for their helpful comments/suggestions and to Prof. David Battisti for itemizing the issues addressed in a revised manuscript. Here, we respond to these comments and outline how our proposed revised manuscript will address their concerns. In this reply, we reproduce each Reviewer's comments in blue italics, while our responses are in standard font. <u>We include some figures and tables as part of this response. These figures and tables are labelled "Fig. R1", "Fig R2", "Fig. R3", "Fig. R4" and "Table R1". The other figure and table numbers refer to the original manuscript.</u>

Before presenting our responses to each referee, let us address some overall issues that we would like to share under the general discussion points (I) through (V) below.

(I) Role of this paper within the QBOi program

The QBOi Phase 1 project focused on evaluating various aspects of QBO simulation (quality of the QBO in control integrations, modulation associated with global warming, seasonal projection, wave activity etc; see references below) across a range of global circulation models. This foundation enabled the next phase that involves more targeted studies, such as the present QBOi-ENSO project, to explore specific forcings and their impact on the QBO.

The QBOi-ENSO experiments utilize a simplified framework, adding somewhat typical intense ENSO SST anomalies to the annual cycle climatological SST from QBOi Phase-1 Experiment 2 (QBOi Phase-1 experiment 2 used annually repeated climatological SSTs while Experiment 1 employed observed SSTs from January 1979 to February 2009, see Butchart et al. 2018 in more details).

This deliberate simplification isolates the impact of ENSO on the QBO, allowing for a clearer interpretation of the results. While direct comparison with K2019 and observations is indeed more challenging with this design, the amplified anomalies are expected to increase the signal-to-noise ratio of the response, providing a robust signal to identify key mechanisms and model sensitivities.

This present paper is meant as the initial component of several linked papers, and it specifically aims to provide a comprehensive overview of the QBOi-ENSO experimental design, participating models, and fundamental ENSO influences on the QBO and related meteorological phenomena. This foundational information will be essential for subsequent QBOi-ENSO publications, just as Butchart et al. (2018) served as an overview of the experiment design for QBOi Phase 1. In this paper, Section 2 includes details on the experimental protocol, model descriptions, and data information, similar to Butchart et al. (2018), but also expands upon this with some scientific analysis of ENSO modulation of the QBO, as indicated in the title.

We acknowledge the reviewer's point regarding the complexity of wave-mean flow interactions and the need for more detailed diagnostics. While such in depth analysis was not within the scope of the present paper, it is an important next step. We are currently planning follow-up studies involving detailed wave analyses, similar to Holt et al. (2020) in QBOi Phase-1, using the QBOi-ENSO datasets to address these specific questions. We believe this tiered approach, starting with a broad overview and followed by more specialized investigations, will be the most effective way to disseminate the findings of the QBOi-ENSO project. Therefore, we have reserved more detailed wave analysis for these future publications.

On the other hand, we understand the reviewer’s concern and we plan for significant new analyses that will be presented in our revised manuscript. The additional analysis is described in general discussion points (III) and (IV) below

---QBOi phase-1 project----

<Protocol paper>

Butchart, N. et al., 2018: Overview of experiment design and comparison of models participating in the SPARC Quasi-Biennial Oscillation initiative (QBOi), GMD, https://doi.org/10.5194/gmd-11-1009-2018

<Five core papers>

Bushell, A. C. et al. 2020: Evaluation of the Quasi‐Biennial Oscillation in global climate models for

the SPARC QBO‐initiative, QJRMS, https://doi.org/10.1002/qj.3765.
Richter, J. H. et al., 2020: Response of the quasi-biennial oscillation to a warming climate in global climate models, QJRMS, https://doi.org/10.1002/qj.3749.
Anstey, J. A. et al., 2021: Teleconnections of the quasi‐biennial oscillation in a multi‐model ensemble of QBO ‐resolving models, QJRMS, https://doi.org/10.1002/qj.4048.
Holt, L. et al. 2020: An evaluation of tropical waves and wave forcing of the QBO in the QBOi models, QJRMS, https://doi.org/10.1002/qj.3827.
Stockdale, T. N., et al. 2020: Prediction of the quasi‐biennial oscillation with a multi‐model ensemble of QBO‐resolving models, QJRMS, https://doi.org/10.1002/qj.3919.

---QBOi-ENSO project----
<papers to be submitted within a period of a few months>
Kawatani et al. 2025: QBOi El Nino Southern Oscillation experiments Part I: Overview of experiment design and ENSO modulation of the QBO, EGUsphere, https://doi.org/10.5194/egusphere-2024-3270, 2024. (Current manuscript. Note that "Part I" will be removed in the revision as we asked)
Naoe et al. 2025: QBOi El Niño Southern Oscillation experiments: Teleconnections of the QBO, WCD discussion, submitted
Elsbury et al. 2025: QBOi El Niño Southern Oscillation experiments: Assessing relationships between ENSO, MJO, and QBO, EGUsphere, https://doi.org/10.5194/egusphere-2024-3950, 2025.

(II) Strength of ENSO SSTs

We understand concern regarding the strength of the ENSO SST anomalies used in the QBOi-ENSO experiments as compared to those used in Kawatani et al. 2019 (referred to as K2019, hereafter). We chose to use amplified anomalies to maximize the ENSO signal in the QBO response and clarify the underlying mechanisms. We will explain our rationale and the procedure in detail.

The SST anomalies in K2019 represent a "moderate" ENSO based on observations from all El Niño/La Niña SST anomalies. ENSO SST composites were constructed using data from 1950-2016, based on the Japan Meteorological Agency (JMA) ENSO indices. For each calendar month, El Niño and La Niña events were identified according to the JMA definition. Monthly SST anomalies were then weighted by the corresponding NINO.3 index and averaged to create monthly composite SST anomalies.

This process resulted in "moderate" composite ENSO SST anomalies, as illustrated by the January El Niño example. Seventeen January El Niño events were identified between 1950-2016, with NINO3.4 anomalies ranging from 0.4K to 3.2K. The resulting composite NINO.3 SST anomaly for January was 1.92K, representative of a moderate El Niño event. In the observational record, the highest NINO.3 anomaly value in a January El Niño event is +3.5 K.

To ensure a clear and robust QBO response in our model experiments, we amplified these composite SST anomalies. El Niño anomalies were multiplied by 1.8 and La Niña anomalies by 1.4. These factors were chosen to approximate the peak SST anomalies observed during the strongest El Niño and La Niña events (Table S1, Figure R1). This approach allowed us to better isolate the impact of ENSO on the QBO. We emphasize that SST anomalies employed are not ‘unrealistically’ large in the sense that actual anomalies of this magnitude are observed on occasion.

Figure R1 shows the annual cycle of the amplified composite ENSO SST anomalies, compared to maximum/minimum observed values. While the compositing procedure cannot perfectly capture the evolution of individual ENSO events, the amplified anomalies exhibit realistic seasonal variations, with El Niño peaking during boreal winter. The variability in La Niña development is also reflected in the composite.

Therefore, while stronger than the anomalies used in K2019, our amplified SSTs remain within the realm of observed ENSO magnitudes, representing the peak values seen during strong events. This approach, like the use of amplified forcings in other climate modeling projects, allows us to better discern the QBO response to a substantial ENSO forcing within the constraints of computational resources and project timelines.

For example, both the QBOi Phase 1 project and various CMIP experiments have employed amplified forcings, such as 2xCO_2/+2K SST and even 4xCO_2/+4K SST, to investigate climate system responses. While a 4xCO_2/+4K SST scenario is unlikely in the near future, the insights gained from such experiments are valuable. Similarly, in the QBOi-ENSO experiments, we prioritize exploring the impacts of strong, yet realistic, ENSO events as a first step. By focusing on the upper end of observed ENSO variability, we can more effectively identify key mechanisms and sensitivities, laying a solid foundation for future research. Therefore, we believe this study offers valuable and meaningful insights into the ENSO-QBO relationship.

It seems that our rationale and conceptual framework were not adequately explained in the initial submission. We intend to address this thoroughly in the revised manuscript. Specifically, we will provide a clearer explanation of the experimental design choices, particularly the decision to use amplified ENSO SST anomalies. We will also articulate more clearly how this study fits within the broader QBOi-ENSO project goals. Furthermore, we will expand on the strategic reasons for focusing on strong ENSO events, given the constraints of computational resources and project timelines. We believe these revisions will significantly improve the clarity and impact of our work.

Also note that we are expanding our study to include analysis of existing AMIP runs with the QBOi models. This analysis will address concern about the imposed SST anomalies in a direct way. See general discussion point (IV) below for more details.

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
El Niño month	17	15	11	13	14	18	18	17	18	18	18	17
Max	3.2	2.6	2.1	1.8	2.1	2.0	2.5	2.9	3.0	3.3	3.6	3.5
La Niña month	16	15	13	14	12	12	15	15	16	16	16	16
Min	-1.8	-1.5	-1.0	-1.3	-1.4	-2.0	-1.6	-1.6	-1.3	-1.6	-1.7	-1.8

Table R1: The number of El Niño and La Niña months during 1950-2016. Max and Min indicate maximum NINO3 anomalies for El Niño and minimum for La Niña (unit: K).

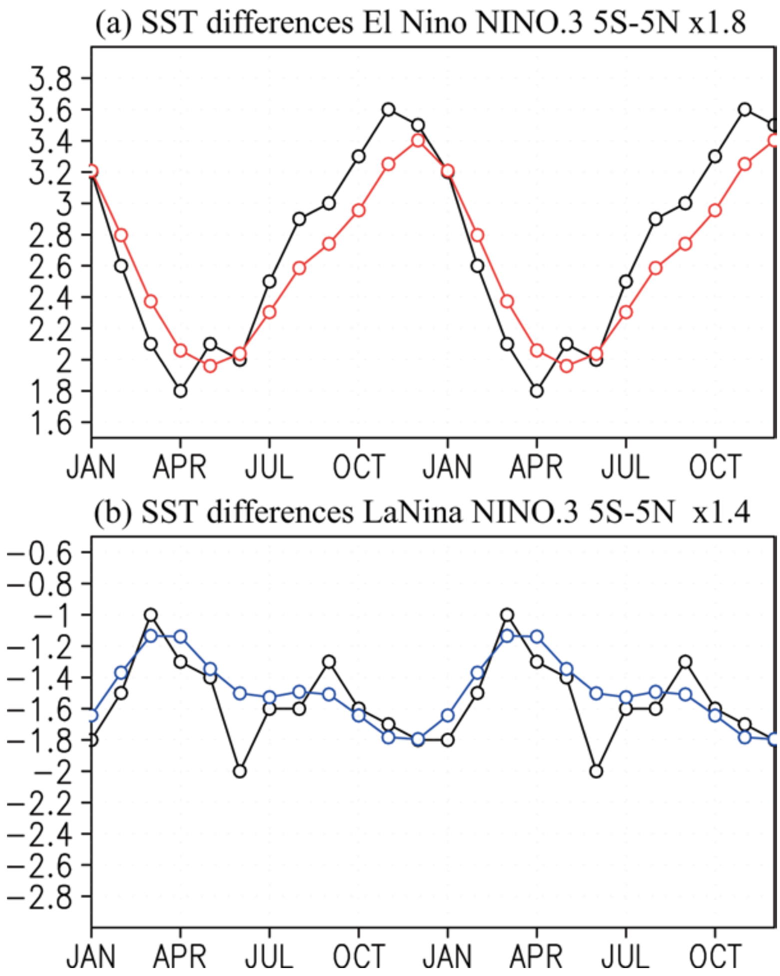


Figure R1: The red and blue lines show the delta SSTs in our (a) El Niño and (b) La Niña experiments and represent typical moderate El Niño and La Niña anomalies multiplied by a factor of 1.8 and 1.4, respectively. These monthly delta SSTs are smoothed in time with a 1-2-1 filter. Black lines represent the maximum/minimum observed monthly values during the entire record as shown in Table R1. For visualization, two (exactly repeated) full cycles are shown.

(III) Adding more detailed analysis of the seasonal and QBO phase dependence of the ENSO influence on the QBO, following Taguchi (2010)

Following the excellent suggestion to provide a more detailed analysis of the ENSO effects on the QBO mean flow evolution, we will present the same figures as in Taguchi (2010) for our QBOi models, as well as for FUB observational data from 1953 to 2022—extending the dataset by an additional 14 years compared to Taguchi (2010). This allows us to examine the modulation of QBO amplitude and period as a function of both QBO phases and seasons during El Niño and La Niña.

We have actually now completed this analysis and Figure R3 presents two-dimensional plots of the mean amplitude and mean rate of phase progression for each category, classified by season and QBO phases at 50 hPa (see Fig. R2 below for the QBO phase definition). In our Fig. R3 red shading indicates stronger QBO amplitudes, or faster downward phase progression, during El Niño compared to La Niña.

FUB observations show a weaker QBO amplitude during El Niño in most seasons and phases. Compared to observations, GISS and LMDz simulate significantly enhanced QBO amplitudes during El Niño, as was also evident in Fig. 4 of the original manuscript.

The QBO phase progression rate in FUB observations indicates that QBO phases generally propagate faster during El Niño than during La Niña across most QBO phases and seasons. Most models capture this characteristic behavior. ECHAM5sh produces results consistent with observations for the W and EW phases, but deviates in the E and WE. EMAC also exhibits phase-dependent variations in QBO behavior.

Much faster phase propagation during El Niño is observed, particularly in westerly phases at 50 hPa. The ENSO-related variability of the QBO is most pronounced during the W phase, when the easterly phase is descending from higher levels. Among the models, EC-EARTH appears to best reproduce these characteristics.

In the revised manuscript we will show all these more detailed results for the seasonal and QBO phase-dependent modulation of the QBO by ENSO. We also will show results applying the analysis to earlier AMIP runs with the same model as described in general discussion point (IV) below.

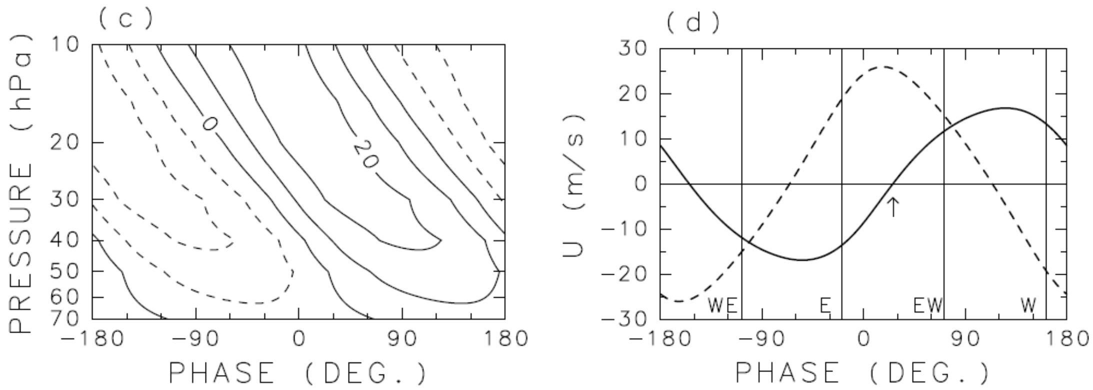
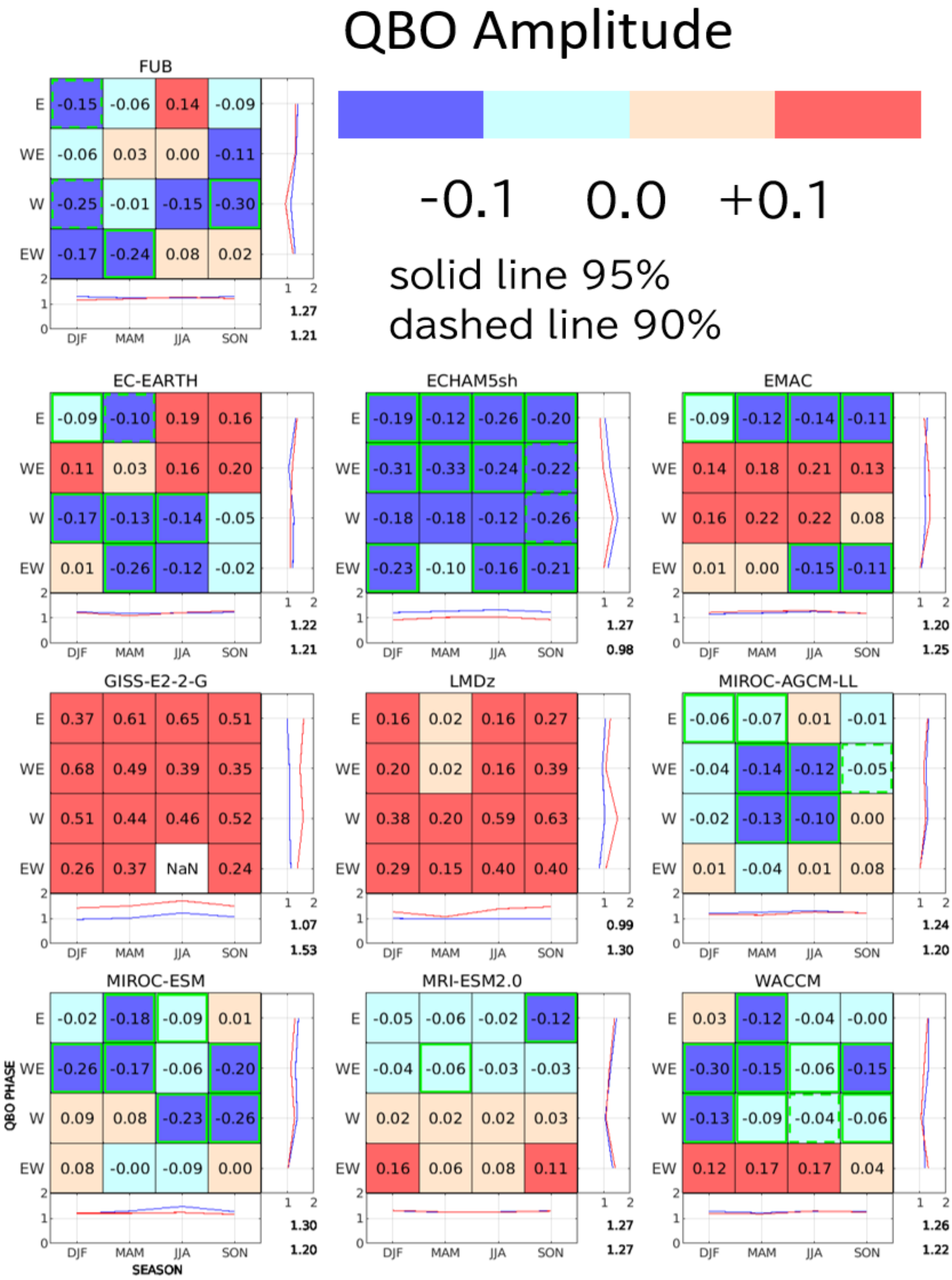


Figure R2: Part of a figure from Taguchi (2010) that we reproduce here to show the definitions of WE, E, EW and W phases. (c) Reconstructed QBO zonal wind. (d) Reconstructed wind at 50 (solid line) and 20 (dashed line) hPa. See more details in Taguchi (2010).


QBO Amplitude
-0.1 0.0 +0.1
solid line 95%
dashed line 90%
FUB
E -0.15 -0.06 0.14 -0.09
WE -0.06 0.03 0.00 -0.11
W -0.25 -0.01 -0.15 -0.30
EW -0.17 -0.24 0.08 0.02
1.27
1.21
EC-EARTH
E -0.09 -0.10 0.19 0.16
WE 0.11 0.03 0.16 0.20
W -0.17 -0.13 -0.14 -0.05
EW 0.01 -0.26 -0.12 -0.02
1.22
1.21
ECHAM5sh
E -0.19 -0.12 -0.26 -0.20
WE -0.31 -0.33 -0.24 -0.22
W -0.18 -0.18 -0.12 -0.26
EW -0.23 -0.10 -0.16 -0.21
1.27
0.98
EMAC
E -0.09 -0.12 -0.14 -0.11
WE 0.14 0.18 0.21 0.13
W 0.16 0.22 0.22 0.08
EW 0.01 0.00 -0.15 -0.11
1.20
1.25
GISS-E2-2-G
E 0.37 0.61 0.65 0.51
WE 0.68 0.49 0.39 0.35
W 0.51 0.44 0.46 0.52
EW 0.26 0.37 NaN 0.24
1.07
1.53
LMDz
E 0.16 0.02 0.16 0.27
WE 0.20 0.02 0.16 0.39
W 0.38 0.20 0.59 0.63
EW 0.29 0.15 0.40 0.40
0.99
1.30
MIROC-AGCM-LL
E -0.06 -0.07 0.01 -0.01
WE -0.04 -0.14 -0.12 -0.05
W -0.02 -0.13 -0.10 0.00
EW 0.01 -0.04 0.01 0.08
1.24
1.20
MIROC-ESM
E -0.02 -0.18 -0.09 0.01
WE -0.26 -0.17 -0.06 -0.20
W 0.09 0.08 -0.23 -0.26
EW 0.08 -0.00 -0.09 0.00
1.30
1.20
MRI-ESM2.0
E -0.05 -0.06 -0.02 -0.12
WE -0.04 -0.06 -0.03 -0.03
W 0.02 0.02 0.02 0.03
EW 0.16 0.06 0.08 0.11
1.27
1.27
WACCM
E 0.03 -0.12 -0.04 -0.00
WE -0.30 -0.15 -0.06 -0.15
W -0.13 -0.09 -0.04 -0.06
EW 0.12 0.17 0.17 0.04
1.26
1.22
DJF MAM JJA SON
QBO PHASE
SEASON

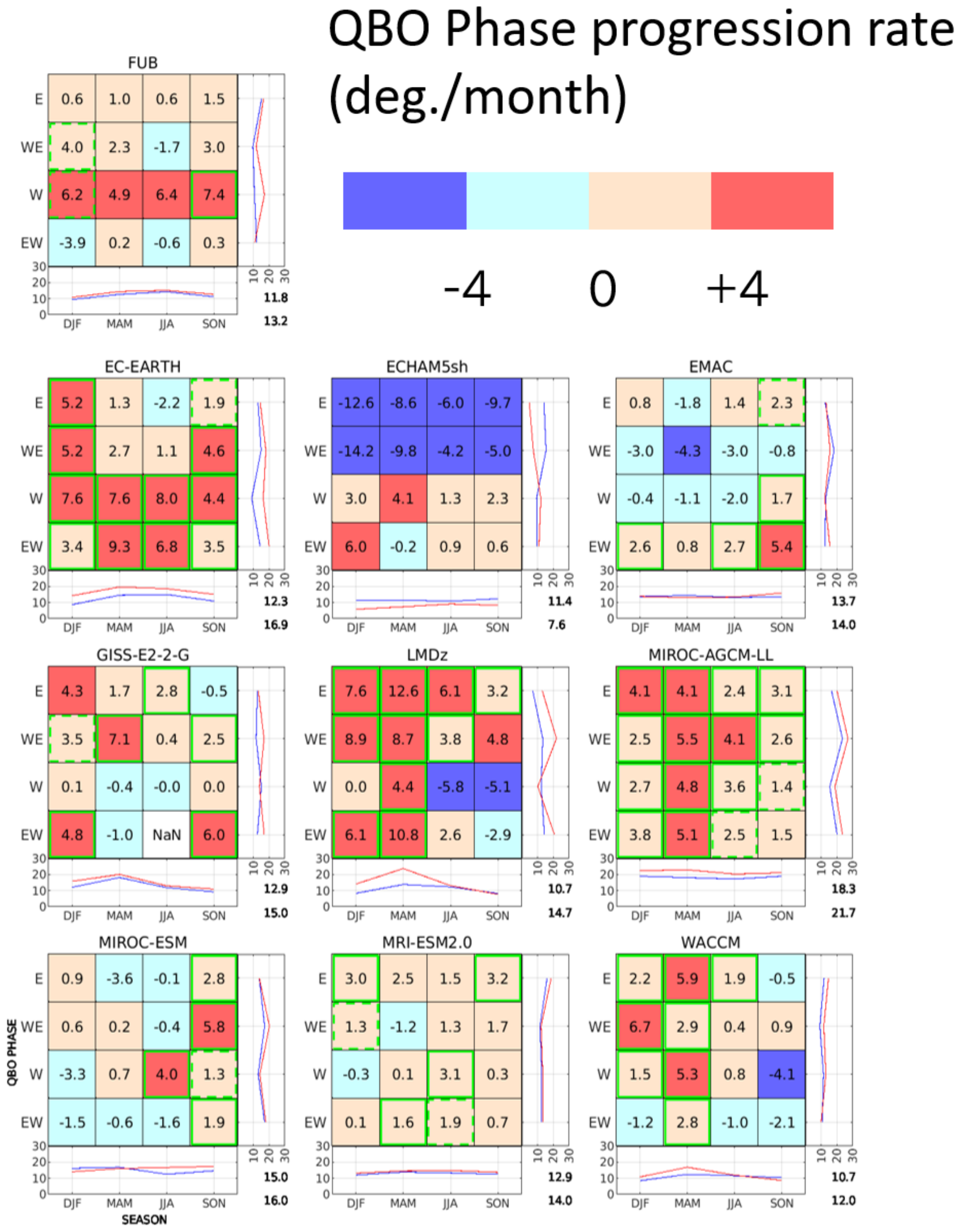


Fig.R3: Two-dimensional plots of mean amplitude and mean rate of phase progression for each category determined by the seasons and QBO phases at 50 hPa. Outlined boxes denote that the numbers are judged to be significantly biased at a 95 % (thick outline) or 90 % (thin outline) confidence level. Line plots for El Nino (red) and La Nina (blue) below the panels show seasonal dependencies when values are averaged in the quadrant (phases). Quadrant (phase) dependencies are similarly shown at the right-hand sides. See more details in Taguchi (2010).

(IV) Relation between the SST amplitude and the response of the QBO

An important consideration is the question of how linear with respect to the strength of the SST anomalies is the ENSO modulation of the QBO. For example, Taguchi (2010, https://doi.org/10.1029/2010JD014325) investigated this issue by repeating his observational analyses but only including strong El Niño and La Niña cases.

Taguchi (2010) didn't show these results but states that "The examination shows that the results are generally insensitive to the definitions of the ENSO cases, since the weaker amplitude and faster phase propagation of the QBO are also obtained for the stronger EL conditions." We feel that the limited observational record presents a difficulty in drawing firm conclusions on this issue, but Taguchi's result at least suggests that the use of strong SST anomalies in our experiments is not unreasonable.

While a comprehensive exploration of this is beyond the scope of the present study, we will add a discussion in the revised manuscript. As the editor suggested, one way to investigate how the SST anomaly amplitude modulate the QBO is to run some of the models for ~20 years with moderate SSTs and compare the amplitude of wave forcing. We have access to relevant results from two of the models, namely MIROC-AGCM and MIROC-ESM, which have been run with the present amplified SST anomalies and with the moderate ones used in K2019. Note that unfortunately, the model versions for MIROC-AGCM used in our present study and in K2019 are not identical. For MIROC-ESM the identical model version was used in K2019.

In QBOi-ENSO experiments, MIROC-AGCM and MIROC-ESM show longer periods of the QBO during La Niña than El Niño by about 3.09 months and 1.55 months, respectively. On the other hand, in K2019 experiments using modest SST anomalies, the differences are 2.2 month and 0.4 month (statistically insignificant) respectively. So in these models the ENSO effect on QBO period seems to depend on the strength of the imposed SST anomalies in an intuitively reasonable sense (stronger forcing associated with bigger ENSO-related period change).

Another approach we believe is to apply the same analyses as Taguchi (2010) to the AMIP runs that were conducted earlier as QBOi "Experiment 1". Experiment 1 employed observed SSTs, providing realistic ENSO conditions, from January 1979 to February 2009 (Butchart et al. 2018). Fig. R4 shows the same results as Fig. R3 but for Experiment

1 using the MIROC-AGCM-LL. Our analysis of the AMIP results has just begun and this result for MIROC-AGCM-LL is based on a single realization of ~30 years. But note that we have available integrations for a total of 3 realizations for MIROC-AGCM-LL and for several other of the models, and in our revision we will include results for all available runs.

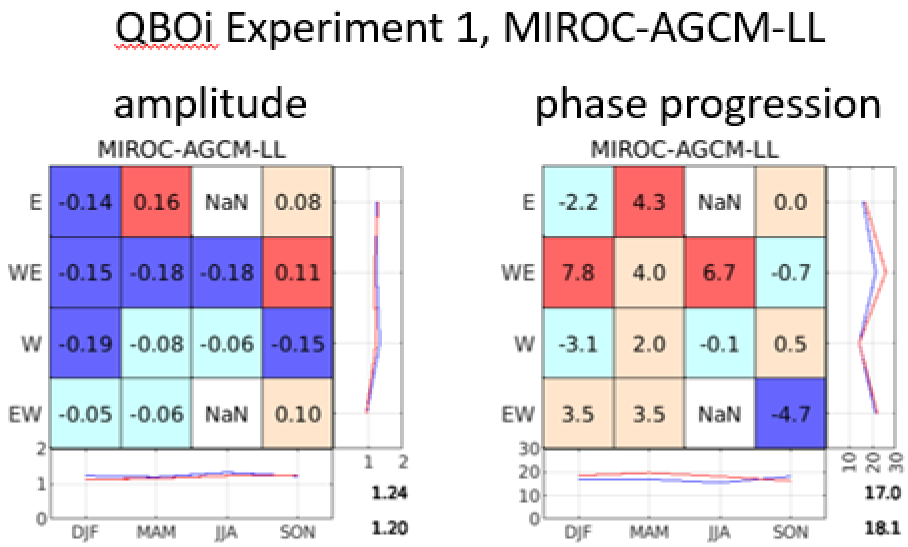


Fig. R4: The same as Fig. R3 but showing results from the QBOi Experiment 1 (an AMIP experiment with observed SST from January 1979 to February 2009) using one realization of the MIROC-AGCM-LL model simulation.

(V) Using ERA5 reanalysis data instead of ERA-I

Both reviewers suggested using ERA5 reanalysis data instead of ERA-I. Accordingly, we will use ERA5 data in the revised manuscript. Additionally, to address Reviewer #2's comment, "*What is new?*", as recommended by the editor, we will incorporate observational results from ERA5 into Figures 11 and 14.

ERA5 provides zonal mean "*Mean eastward wind tendency due to parameterizations*", which includes not only non-orographic gravity waves but also other parameterized forcing. Therefore, we cannot include reanalysis results in Fig. 12 but we can potentially include them in Fig. 14 to illustrate resolved and parameterized forcing. We will include explanations and a note of caution regarding this distinction in the revised manuscript.

Reply to anonymous referee #1

Review of the article QBOi El Nino Southern Oscillation experiments Part I: Overview of experiment design and ENSO modulation of the QBO by Yoshio Kawatani et al.

General comments

This work is a continuation of the publication series produced by the QBOi project, based on an experimental protocol and several models known to simulate the QBO. This work is focused on the El Niño/La Niña effects on the QBO as simulated in 9 models. Most of the article describes the common and different features found in the different model simulations, and its structure follows the work of Kawatani et al. (2019), hereafter K2019, where they investigated the El Niño/La Niña effects on the QBO in MIROC models.

The experimental design chosen here is however deviating from that of K2019. Here they decided to use amplified mean El Niño/La Niña SST anomalies. This makes any direct comparison to K2019 and to observations difficult. Whether or not the QBO response should be linear to the amplitude of the SST anomaly pattern is not discussed. Probably it would have been better to use the same anomalies as in K2019. (Or an entirely different design based on SST fields of selected El Niño and La Niña years.)

Please refer to our general discussion point (II) for this response. We will include additional explanations regarding the comparison between the ENSO SST anomalies in K2019 and the present study in the revised manuscript.

They find that El Niño/La Niña effects on the QBO period are qualitatively similar with respect to the period, with El Niño leading to a shorter period despite of the increased tropical upwelling in the tropical lower stratosphere, from which it is clear that El Niño must also produces a stronger wave mean flow interaction. No common response is found for the QBO amplitude.

An interesting part is the discussion and analysis of the reasons for the described results: The more equatorial precipitation and the weaker Walker circulation found during El Niño conditions. These features are found in all models, and they probably are independent of the skill of a model to simulate a QBO. The discussion of the wave mean flow processes is however more complicated, because of the rather different ways that *resolved and parameterized waves generate the QBOs in the different models. And therefore not so much can be learned from this part, except that there exists still a considerable difference in the way that models generate QBOs. Further, as acknowledged by the authors, more detailed model diagnostics would be needed to learn more about the underlying reasons for the found behaviours. But this additional diagnostics was not part of the protocol, or the modelling groups could not produce these diagnostics.*

We briefly mentioned the datasets used in the present analysis and referenced Butchart et al. (2018). As noted in Butchart et al. (2018), 6-hourly data on temperature, as well as zonal, meridional, and vertical wind, are also available in the QBOi-ENSO experiments. We will provide a more detailed explanation in the revised manuscript. For further on this point refer to our general discussion point (I) above.

Overall I think that the publication is worthwhile, as it creates a baseline for further work on this topic. Some minor corrections are needed before publication.

We sincerely appreciate the Reviewer's positive evaluation and valuable comments. We are grateful that you find our work worthwhile.

Detailed comments and questions

Abstract

L40 Stratosphere-troposphere Processes And their Role in Climate (SPARC) …
As we know QBOi has been started as a SPARC project. But SPARC has changed its name to APARC and QBOi is now listed as an APARC project. Maybe it is worth to add a remark or a footnote on this aspect.

We will fix this.

L45 … models -models … should probably be … models. Models …

We will fix this.

1 Introduction

L64 ... that QBO facilitates ... → ... that the QBO facilitates ...

We will fix this.

L140 … Conducting a common ENSO-QBO experiment across a range of QBO-resolved climate models could help elucidate the role of non-orographic GWP in driving the oscillation. …
The work of Richter et al. (2020) on the climate warming effects on the QBO unfortunately showed that the differences between GWPs are considerable and probably responsible for the rather different QBO responses to the warming. As it seems it was not possible to decide which GWPs were "wrong" or "right". Now a similar exercise is presented aiming at El Niño/La Niña variations in SST as the external forcing instead of a warmer SST and increased atmospheric CO2. Why should we expect a scientifically more robust result if Richter et al. (2020) have shown that differences in parameterizing non-orographic gravity drag can lead to very different results? Simply because El Niño/La Niña cycles exist in the historical period for which the models have been tuned?

We acknowledge your point regarding the challenges in fully elucidating the role of non-orographic GWP. While Richter et al. (2020) examined QBO modulation in a future climate—where direct observational data to validate changes are unavailable—our QBOi-ENSO experiments can be partially validated using existing observations, such as the observed shortening of QBO periods during El Niño compared to La Niña. Although our experimental design is somewhat idealized, it allows us to identify models that produce longer QBO periods during El Niño runs as potentially problematic, prompting a closer evaluation of their GWP parameterizations. We believe this is a key advantage of the QBOi-ENSO experiments over the future climate scenario examined in Richter et al. (2020). Based on your feedback, we will revise the manuscript to clarify this reasoning.

2. Model Description and Experimental Design
L179 … These factors bring the peak composite anomaly SSTs closer to the anomalies observed during the most intense El Niño and La Niña events. …
Using amplification factors is problematic. This makes a comparison to observations or to the work by K2019 difficult. It seems necessary to add some remarks about the linearity between the SST pattern amplitude and the response of the QBO. Can this be assumed? Alternatively you could have chosen specific years with strong El Niño and La Niña SST anomalies. Then there would be no need to amplify the SST anomaly, and there would be less of a risk to construct an SST anomaly pattern that mixes the different types of El Niños, which are discussed in literature.

Please refer to our general discussion point (II) above.

L199 … For clarity and conciseness, we will refer to these models as CESM1, EC-EARTH, ECHAM, EMAC, GISS, LMDz, MIROC-AGCM, MIROC-ESM and MRICESM1, respectively. …
The abbreviated model names are introduced here, but not used consequently. Tables, Figures, and also some sentences use the full model names. Please decide whether short names shall be used or not. But if you decide to use short names, then please use these in all places: tables, figures, and text.

We appreciate your comments and will make sure the revised version uses a consistent system of model names.

L203 … Launch levels for parameterized gravity waves varied across models, ranging from 450 to 700 hPa or 1000 to 100 hPa. …
To which model(s) do the two pressure ranges relate? Please clarify.

Table 1 shows the launch levels. We will add "(see Table 1)" in the revised manuscript.

Table 1. lunched level → launch level

We will fix this.

Table 2. What does the entry for GISS-E2-2G and Residual stream function (5-1115✓) mean?

We will fix this.

L243 … from the ERA-Interim (ERA-I; Dee et al. 2011) reanalyses …
Why is ERA-I used for this comparison, when ERA-5 is now available? Newer reanalyses are generally improved compared to earlier ones.

Please refer to our general discussion point (V) for this response. We will use ERA-5 in the revised manuscript.

L247 … Importantly, the composite ERA-I and CMAP data were not scaled …
This is a kind of a flaw in the experimental design. If the response to the SST anomaly patterns is non-linear to the amplitude, then the applied scaling is hindering a direct comparison to observations or analyses. If, however the signal is linear, then the signals derived from ERA-I should be scaled like the SST patterns used for the simulations.

Please refer to general discussion points (II) and (IV) for this response.

3. ENSO Modulation of the QBO and Climatological Mean Field Differences

L266 – L276 These lines discuss deficiencies in the structure of the simulated QBO, as occurring in El Niño or La Niña simulations of ECHAM, GISS, and LMDz. In my opinion it is necessary to point out another deficiency, which is an unrealistic period, although a regular pattern of downward propagating westerly and easterly jets is simulated. Taking the displayed 20 years (Fig. 2) of the El Niño and La Niña simulations together, we have 40 years for which on average (40years / 28 months) we would expect about 17 cycles. A count of the cycles shown in Figure 2 can now serve as an additional measure for the quality of the simulations. If we allowed a range of 15 to 19, then the following models (here excluding ECHAM, GISS and LMDz) would fail: EC-EARTH: 20, MIROC-AGCM-LL: 26, MIROC-ESM: 21. Please extend your discussion of problematic simulations in this direction, so that the reader knows from the beginning which model simulations need to be viewed a bit more critically.

We appreciate your comments regarding the unrealistic QBO periods in QBOi model simulations. A detailed analysis of simulated QBO periods in QBOi “Experiment 2” (with a climatological annual cycle of SSTs) was previously conducted by Bushell et al. (2020). However, we recognize the importance of addressing this issue directly in the context of our current study on the QBOi-ENSO experiment. Therefore, we will incorporate a more extensive discussion of the simulated QBO periods in the revised manuscript. This will provide readers with a clearer understanding of the model limitations and enable them to assess the simulations more critically from the outset.

Bushell, A. C., J. A. Anstey, N. Butchart, Y. Kawatani, S. M. Osprey, J. H. Richter, F. Serva, P. Braesicke, C. Cagnazzo, C.-C. Chen, H.-Y. Chun, R. R. Garcia, L. J. Gray, K. Hamilton, T. Kerzenmacher, Y.-H. Kim, F. Lott, C. McLandress, H. Naoe, J. Scinocca, T. N. Stockdale, S. Watanabe, K. Yoshida, S. Yukimoto: Evaluation of the Quasi - Biennial Oscillation in global climate models for the SPARC QBO‐initiative, Quarterly Journal of the Royal Meteorological Society, https://doi.org/10.1002/qj.3765, 2020

L311 … Next, we consider ENSO modulation of QBO amplitude, which is known less robust … → … Next, we consider the ENSO modulation of the QBO amplitude, which is known to be less robust …

We will fix this.

L323 … GISS, LMDz, and CESM1, all of which have variable GWP sources. …
I think it should be added that MIROC-AGCM-LL has variable gravity waves too, though these are explicitly simulated, within the given resolution, instead of parameterized. Thus variability of gravity waves not necessarily leads to a strong amplitude difference between El Niño and La Niña. And one needs to wonder if a strong change is indicating that the variability of gravity waves is important aspect for a GWP, or whether this effect is rather a result of other aspects of parameterizing gravity wave. Please add some thoughts on this problem.

While MIROC-AGCM-LL can simulate a QBO-like oscillation without parameterized non-orographic GWP, previous studies have indicated that its resolution (T106L72, corresponding to a 1.25-degree horizontal resolution and 550-meter vertical resolution) is insufficient to capture wave forcing as effectively as higher-resolution models, such as T213L256 or even higher-resolution models (e.g., Kawatani et al., 2010). In MIROC-AGCM-LL, most of the unresolved gravity wave forcing that is parameterized in other models is not explicitly simulated. Therefore, we believe that the MIROC-AGCM-LL results do not necessarily support the conclusion that 'variability of gravity waves does not necessarily lead to a strong amplitude difference between El Niño and La Niña.'

Parameterized GWP represents sub-grid-scale processes on much smaller scales than ~100 km. The variable source of parameterized GWP is often linked to cumulus convection, which also generates gravity waves on much smaller scales than those resolved in MIROC-AGCM-LL. We will clarify this distinction in the revised manuscript.

Kawatani, Y., K. Sato, T. J. Dunkerton, S. Watanabe, S. Miyahara and M. Takahashi, 2010: The roles of equatorial trapped waves and internal inertia-gravity waves in driving the quasi-biennial oscillation. Part I: Zonal mean wave forcing, J. Atmos. Sci., 67, 963-980., https://doi.org/10.1175/2009JAS3222.1

L383 … although models tend to simulate the precipitation peak to the east of the observed one over the central Pacific in the El Niño run. …
It should also be mentioned that the precipitation peak in the model simulations is higher than in observations, which indicates that the local forcing by latent heat release in the simulations is higher than that explained by the observed precipitation. Quite likely this is related to the amplified El Niño/La Niña SST strength.

We acknowledge your point that the precipitation peak in the model simulations is higher than in observations. It is important to note that the observed precipitation data we are using represents averaged values weighted by the NINO3.4 index, reflecting precipitation patterns during "moderate" ENSO events. While we could alternatively present precipitation distributions during strong El Niño events, such as the 1997 event, we believe that this level of detail is not essential for the core message of this manuscript. However, we will add further explanation regarding this issue in the revised manuscript to ensure clarity for the reader

L422 ... significantly deep westerly difference ... → ... significantly deeper westerly difference …

We will fix this.

L432 … for (left top) ERA-I …
ERA-I is "center top"

We will fix this.

4 Contrasting wave forcing and residual mean meridional circulations in El Niño and La Niña from QBOi models

L458 … The X term represents any other unresolved forcing. … Do you mean here parameterized momentum diffusion and effects from numerical diffusion and damping operators?

Indeed this represents all the other possible contributions including explicitly parameterized diffusion and any other contributions from the numerical schemes employed.

L470 … in El Niño and La Niña simulations. …
La Niña simulations are nor shown, but differences of the El Niño and La Niña simulations.

We will change the text, thank you.

L477 ... La Niña c annual ... → ... La Niña annual ...

We will change the text, thank you.

L498 … which both use variable sources in their GWP, …
Do you mention this because you think that this is the reason for the differences? Often other differences in the formulation of the non-orographic gravity wave drag parameterizations can cause substantial differences already.

We acknowledge that we mentioned the variable GWP sources in these models based on what was visually apparent in the figure, without directly attributing them as the sole cause of the observed differences. We agree with your assessment that other differences in the formulation of non-orographic gravity wave drag parameterizations could also contribute substantially to these differences. While it is difficult to definitively determine the specific reasons at this stage, we will take this point into account in the revised manuscript.

L509 … averaged over 20°S–20°N …
Maybe it is worth to explain why a band of 20°S – 20°N is chosen, while earlier diagnostics/figures used narrower bands. (I guess this is made in order to remove residuals of the secondary meridional circulation of the QBO.)

We selected the 20°S–20°N latitude band for averaging because this region encompasses the area where the QBO amplitude is most pronounced. While other latitude bands, such as 10°S–10°N or 15°S–15°N, could also be used, they tend to exhibit more noise. However, we have confirmed that our basic conclusions remain largely unaffected by the choice of these latitudinal ranges. In the revised manuscript, we will add a brief explanation regarding this choice, acknowledging the potential rationale behind selecting wider bands to minimize noisier vertical structures.

L510 … ranging from approximately 0.2 mm s^1 in MIROC-AGCM to approximately 0.4 mm s^1 ⁻⁻in LMDz. …
This strong difference in the tropical upwelling implies also a strong difference in the strength of wave mean flow interaction that is necessary to simulate a QBO with a realistic period. This aspect is not discussed here, and maybe this El Niño/La Niña related article is the wrong place. Still it directly shows that the wave mean flow interaction must work at different strengths.

Weak tropical upwelling in the MIROC model is discussed in Kawatani et al. (2010). In addition, the tropical upwelling is also different among reanalysis, as discussed in the S-RIP final report SPARC (2022). We will add a brief discussion about tropical upwelling related with a QBO in the revised manuscript.

SPARC, 2022: SPARC Reanalysis Intercomparison Project (S-RIP) Final Report. Masatomo Fujiwara, Gloria L. Manney, Lesley J. Gray, and Jonathon S. Wright (Eds.), SPARC Report No. 10, WCRP-6/2021, doi: 10.17874/800dee57d13, available at www.sparc-climate.org/publications/sparc-reports.

L522 … However, the specific altitudes at which $w*$ changes would most strongly influence the overall QBO period remain unclear. …
Sentence unclear.

We will modify this sentence to
"However, it remains unclear which specific altitudes where changes in w* occur have the strongest influence on the overall QBO period."

L579 … While output data of parameterized gravity wave fluxes in LMDz were not available at the time of this analysis, this model, which also uses variable parameterized wave sources related to precipitation activity, showed similar structures affected by precipitation distributions (Dr. Lott, personal communication). …
Francois Lott is a co-author of this study. Please include the LMDz results in Figure 12.

Francois Lott informed us that the datasets were quantitatively incorrect due to inadequate processing. However, we have confirmed that the qualitative distribution is related to precipitation, similar to what is observed in CESM1.

L620 … parameterized wave forcing below is stronger …
What does "below" refer to? Maybe the sentence needs to be rephrased.

We delete "below" in this sentence.

L639 ... As discussed for Fig. 14, ... → ... As discussed for Fig. 13, ...

We will fix this.

5. Summary and concluding remarks

L685 … remained consistent …
"consistent" seems to be the wrong term, because this could have different meanings. (If El Niño/La Niña influences the ozone distribution, then the same ozone field cannot be consistent with El Niño and La Niña conditions at the same time.) "unchanged" would express more clearly that these fields simply have not been changed.

We will change from consistent to unchanged following your suggestion, thank you.

Reply to anonymous Reviewer #2

Review of the manuscript "QBOi El Nino Southern Oscillation experiments Part I: Overview of experiment design and ENSO modulation of the QBO" by Kawatani et al.

Summary

The manuscript presents an overview of the experimental design from the new SPARC QBOi project and examines the modulation of the QBO by ENSO using nine climate models. The findings indicate that the QBO period is longer during La Nina compared to El Nino across all models, consistent with observations. However, changes in the QBO amplitude remain inconclusive. Overall, I find the experiment intriguing, the manuscript well-written, and the results clearly explained. Most of my comments are minor and focus on improving consistency between different parts of the manuscript and aligning the figures with the text.

We sincerely appreciate the Reviewer's positive evaluation and valuable comments. We are grateful that you find our work worthwhile.

One major comment, however, concerns the lack of deeper insights into ENSO modulation of the QBO. The authors attribute this to the simplicity of the analyses and the limited availability of model output data, which they suggest prevents a full explanation of the quantitative differences in QBO between El Nino and La Nina. While future studies are mentioned as a potential avenue to address this, I argue that if a more detailed analysis is feasible, it should be included in this paper, as it was the primary motivation for the experiment and study.

Please refer to our general discussion point (I) above for a consideration of the role of this paper within the broader QBOi program which helps motivate our work. Our plans for revision include significantly more analysis. In particular, following the Editor's suggestion we will repeat the more detailed analysis of Taguchi (2010) to elucidate the seasonal effects in the QBO-ENSO connections. We also are in the process of expanding our study to include analysis of the earlier AMIP runs (QBOi "Experiment 1") to supplement the new "annually repeating" runs that was the main focus of our paper.

Major Comments

1) Lack of additional insight into the mechanisms of ENSO modulation of the QBO

My primary concern is whether this paper and the associated experiments provide any additional insight into the mechanisms by which ENSO modulates the QBO. At the start of the paper, I had hoped—likely in line with the motivation behind designing and implementing these experiments—that this study would offer a deeper understanding of these mechanisms. However, the study appears to be an extension of Kawatani et al. (2019), with potentially more models included beyond MIROC, yet missing important analyses due to data limitations.

There are repeated statements such as: "Further investigation of these models is hampered by the incomplete model variables in the available data sets", "This simple analysis with limited model output data cannot fully explain quantitative differences in QBO periods between El Nino and La Nina", and "Detailed zonal-time spectral analyses of model fields, like those performed in Kawatani et al. (2019), remain a subject for future study."

If such analyses are indeed possible, this paper is the appropriate venue to present them, rather than postponing them to future studies. For example, as the authors mentioned, detailed spectral analyses of the EP flux, gravity wave parameterization fluxes, precipitation, or momentum budgets based on the TEM framework could offer crucial insights into the intermodel spread of QBO period and amplitude.

To provide further context on my disappointment, Kawatani et al. (2019) noted: "It would be interesting to analyze the ENSO modulation of the three-dimensional wave forcing as well as tropical upwelling, which must show large differences between El Nino and La Nina. This may be investigated in a future study". Now, five years later, this study states: "A detailed investigation of the three-dimensional distributions of parameterized wave fluxes modulated by ENSO, including model dependence, would be of interest and remains a topic for future research."

It feels like an opportunity has been missed to address these outstanding questions. If there is a way to conduct these analyses, I strongly encourage the authors to include them in this paper.

We do believe that our manuscript already contains a substantial amount of useful material and it provides the background and introduction for related QBOi studies please refer to our general discussion point (I) above. However, we understand the reviewer's concern and will include further analysis of the simulations following the editor's suggestion, as explained in our general discussion point (III) above as well as a substantial expansion of our study to include analysis of earlier AMIP runs as explained in general discussion point (IV) below

2) Lack of use of recent data and citations of recent studies
Some aspects of the study, including citations and the data used, feel somewhat outdated. For instance, the use of ERA-I instead of ERA5. Additionally, the study only uses observed data up to 2012. If this limitation is due to avoiding the QBO disruptions, there are still several years of data available between 2012 and the end of 2015, as well as between 2020 and 2024. While including these additional years may not change the main conclusions of the paper, it would enhance the robustness of the analysis, particularly for slowly evolving phenomena like ENSO and QBO, where even a few more samples could provide valuable insights. Moreover, the citations miss some relevant and recent studies, such as Zhou et al. (2024), and a few others noted in my review.

We appreciate this concern and we will repeat all our analyses with ERA5 data as explained in general discussion point (V). We will also include the recent references suggested.

Minor comments
L47-49: It can also be mentioned that "all models simulate stronger equatorial tropical upwelling in El Nino compared to La Nina up to ~10 hPa".

We will mention this.

L85: Small-scale gravity waves also contribute significantly to the forcing of the QBO westerly (e.g., Pahlavan et al. (2021))

We will refer to this paper.

L95: As a good reference on this you can cite Coy et al. (2020).

We will refer to this paper.

In general, the figures can be significantly improved by reducing redundancy, which would allow for larger, clearer panels. For instance, in Figure 2, use "El Nino" as the title for the left column and "La Nina" for the right, rather than repeating them for each panel. Similarly, list model names only on the left side of the figure and show the y-axis (0–20) only on the bottom panels, instead of repeating it in every panel. These changes can enhance readability and apply to other figures as well.

We will fix these issues in the revised manuscript, thank you.

The other general issue with the figures is the presence of too many contours, which reduces readability. In particular, in Figures 4, 8, 9, 10, and 13, the contours over the shadings can be removed, similar to Figure 12, to improve clarity.

In our revised version we will adopt these suggestions to improve legibility of these figures.

Figure 3: Have you analyzed each phase of the QBO separately? For example, do both phases of the QBO become shorter during El Nino?

This is an excellent suggestion and was also made by the Editor. We have now repeated the analysis segregated by QBO phase and season and will include these results in the revised version. Please refer to general discussion point (III) for details.

Figure 3: Have you considered using a Fourier Transform to determine the period instead of relying on zero wind line crossing (e.g., as done in Lee et al. (2024))? While it likely won't change the conclusions, it might be a better option, particularly when the QBO becomes more irregular/unrealistic, as seen during El Nino in ECHAM.

We appreciate the suggestion to use a Fourier Transform to determine the QBO period, as demonstrated in Lee et al. (2024). We agree that FFT methods are particularly useful when defining the phase transition of the QBO is challenging (we employed FFT methods in our QBOi phase-1 paper, Richter et al. 2020). However, in the current QBOi-ENSO experiments, all models have a clearly defined QBO phase transition at 20 hPa. Therefore, we prefer to determine the QBO period by phase transition, following the approach of K2019. As you noted, this choice is unlikely to alter the overall conclusions of our study.

L412: Will the cooler anomaly around 60°N–90°N in ERA-I, which is not observed in the models, change if more data is included, such as using ERA-5 from 1940 to 2024?

We did not find this even in ERA-5, as we will show in the revised manuscript.

For Figures 10, 11, 13, and 14, you could consider including results from reanalysis (e.g., ERA5) as a reference, similar to what is done in Figures 8 and 9.

We will do this. Once again refer to our general discussion point (V).

Figure 12: Could you add the total flux for El Nino and La Nina (i.e., averaged over 10°S–10°N and all longitudes) to the bottom panels? If so, is it consistently larger during El Nino?

Thank you for the suggestion. We appreciate the value of including the total flux for El Niño and La Niña (averaged over 10°S–10°N and all longitudes) in Figure 12. While we are open to adding this information, we believe it would be best presented in a separate supplementary figure to avoid overcrowding the existing figure. We will carefully consider this option and its potential impact on the clarity of the manuscript.

Editorial comments

L84: "respectively" seems redundant.

We will fix this.

L107: You can cite (Richter et al., 2020) again to avoid ambiguity.

We will fix this.

L108: SST is not yet defined.

We will fix this.

L140: "QBO-resolved" -> "QBO-resolving"

We will fix this.

L164: GWP is already defined.

We will fix this.

L165: What is experiment 2?

This means what is defined as "experiment 2" in the QBOi phase 1 project. We will explain this in the revised manuscript.

L165: SST needs to be defined at L108.

We will fix this.

L198: The model name "CESM15-110L" is mentioned here, while "WACCM5-110L" is used in the results (figures and tables). I suggest selecting one naming convention for consistency.

We will fix this, thank you.

L200: Using the concise version of model names is a great choice, but it would be helpful to maintain this approach consistently in the results (figures and tables) as well. Currently, there is a discrepancy where the text uses concise names while the results use the full model names, making it harder to follow.

As we mentioned in our response to the other reviewer, we will have a consistent nomenclature for the models in the revised version.

L198: "CESM15-110L" is mentioned here but in the results (figures and tables) "WACCM5-110L" is used. I suggest choosing one for consistency.

We will fix this.

L200: It is great to use the more concise version of the model names but it would be great to use the concise version in the results as well (figures and tables) to make it easier to follow. Now, there is this discrepancy between the text and results, the former using the concise version, and the latter the full name of the models.

As mentioned above, we will fix this and have a consistent nomenclature in our revised manuscript.

L220: Palmerio et al. (2022) is not in the bibliography.

We will fix this, thank you.

L238: TEM is already defined.

We will fix this.

Table 2: what is "5-1115" in front of GISS.

This is our mistake. We will fix this.

L284: "larger" -> "longer"(?)

We will fix this.

L295: "with" -> "for"

We will fix this.

L375: ITCZ not defined yet.

We will fix this.

L378: "(left-top)" -> "(center-top)"

We will fix this.

L392: Any reference for this statement?

We will refer to the relevant paper in the revised manuscript.

Figure 7: “PRCP” not defined.

We will fix this.

L427: BDC is already defined.

We will fix this.

L431: “(left-top)” -> “(center-top)”

We will fix this.

L442: A point after 4 is missing. "In El Nino and La Nina from QBOi models" is redundant. Also, capitalize the first letter to maintain consistency with the other titles.

We will fix this.

L443: “eddy forcing” -> “wave forcing”, to be consistent with the other sections.

We will fix this.

L443: “mean zonal” -> “zonal mean”

We will fix this.

L444: TEM is already defined.

We will fix this.

L445: “eddies” -> “waves”

We will fix this.

L566-567: “WACCM” -> CESM1

We will fix this.

L577: ITCZ should be defined earlier at L375.

We will fix this.

L620: “below” should be removed.

We will fix this.

L621: “below” -> “above” (?)

We will fix this.

L639: “Fig. 14” -> “Fig. 13”

We will fix this.

Caption of Fig. 14: “eddy” -> “wave”. “resolved motions” -> “resolved forcing”.

We will fix this.

References

Coy, L., Newman, P. A., Strahan, S., & Pawson, S. (2020). Seasonal Variation of the Quasi-Biennial Oscillation Descent. Journal of Geophysical Research: Atmospheres, 125(18), e2020JD033077.https://doi.org/10.1029/2020JD033077

Geller, M. A., Zhou, T., & Yuan, W. (2016). The QBO, gravity waves forced by tropical convection, and ENSO. Journal of Geophysical Research: Atmospheres, 121(15), 8886–8895. https://doi.org/10.1002/2015JD024125

Kawatani, Y., Hamilton, K., Sato, K., Dunkerton, T. J., Watanabe, S., & Kikuchi, K. (2019). ENSO Modulation of the QBO: Results from MIROC Models with and without Nonorographic Gravity Wave Parameterization. https://doi.org/10.1175/JAS-D-19-0163.1

Lee, H.-K., Chun, H.-Y., Richter, J. H., Simpson, I. R., & Garcia, R. R. (2024). Contributions of Parameterized Gravity Waves and Resolved Equatorial Waves to the QBO Period in a Future Climate of CESM2. Journal of Geophysical Research: Atmospheres, 129(8), e2024JD040744. https://doi.org/10.1029/2024JD040744

Pahlavan, H. A., Wallace, J. M., Fu, Q., & Kiladis, G. N. (2021). Revisiting the Quasi-Biennial Oscillation as Seen in ERA5. Part II: Evaluation of Waves and Wave Forcing. Journal of the Atmospheric Sciences, 78(3), 693–707. https://doi.org/10.1175/JAS-D-20-0249.1

Richter, J. H., Butchart, N., Kawatani, Y., Bushell, A. C., Holt, L., Serva, F., et al. (2020). Response of the Quasi-Biennial Oscillation to a warming climate in global climate models. Quarterly Journal of the Royal Meteorological Society, n/a(n/a). https://doi.org/10.1002/qj.3749

Zhou, T., DallaSanta, K. J., Orbe, C., Rind, D. H., Jonas, J. A., Nazarenko, L., et al. (2024). Exploring the ENSO modulation of the QBO periods with GISS E2.2 models. Atmospheric Chemistry and Physics, 24(1), 509–532. https://doi.org/10.5194/acp-24-509-2024

We appreciate these relevant recent references. We will include them in the revised manuscript.