

## To the Editor

Dear Prof. Haynes,

We appreciate the feedback from referees 2 and 3, and have made further changes to strengthen our manuscript in light of their comments.

Referee 3 expressed concerns about statistical significance, which we have addressed with new figures. Since referee 2 already had difficulty following the statistical reasoning, before we introduced these figures, we believe that placing them in the supplement makes the manuscript more readable and more sensible physically.

We have also addressed referee 2's request to streamline the discussion. We believe that the manuscript has improved during the review process, and hope that the current form is suitable for publication in ACP.

Sincerely,

Tiehan Zhou on behalf of all co-authors

# Response to Reviewers

## Reviewer 2

### General comments

The authors reorganized the manuscript to address the reviewers' concerns and clarified aspects of the methodology. The manuscript has improved, even if I would suggest to be more concise wherever possible as repetitions are still present (see comments below).

The plot titles are sometimes inconsistent. For example, in Fig. 6 you use 'response' while in previous panels 'composite', I suggest using similar titles instead for clarity.

Done as Suggested.

I find the model description section a bit too long and difficult to follow, such as the discussion on the various periods/subperiods analyzed.

This is a dilemma. A longer description would be easier to follow. However, it is already a bit too long.

In the discussion I would link momentum fluxes and OLR differences more clearly, based on the plots you reported.

Neither momentum flux differences at the source levels nor OLR differences can satisfactorily explain the ENSO modulation of the QBO periods on their own. It is the combination of momentum fluxes at the source levels, OLR differences, the walker circulation that determines the contrast in momentum fluxes that enter into stratosphere. Thus, we cannot link momentum fluxes and OLR differences in a clearer/simpler way.

Please correct 'Walker circulations' -> 'Walker circulation'.

Done as Suggested.

### Specific comments

37 'apparent cancelation' of what?

The phrase "because of the apparent cancelation" has been removed.

121 'deformation' of what?

'deformation' has been changed into 'wind deformation'.

198 perhaps 'composition interactivity' -> 'atmospheric chemistry'?

Done as Suggested.

239 a blank space seems missing

The blank spacing between adjacent words is automatically controlled the Microsoft software. We don't find any missing space there.

398 you used 'am', now 'A', I guess they both mean amplitude. Please choose one

By 'am' we meant the QBO amplitude calculated from monthly mean FUB data. In other words, 'am' denotes the monthly QBO amplitude .

By averaging monthly 'am' over the duration of an ENSO event, we obtain the "evently" QBO amplitude.

In other words, we can regard 'am' as an "instantaneous" amplitude and 'A' as a mean amplitude over an ENSO episode.

However, we don't want to use an overbar at this moment. Instead, we reserve it for a further averaged value, i.e., the "evently" amplitude averaged over the number of El Niño (or La Niña) events, which is called 'sample mean' in the manuscript.

435 to me this point was already clear from the methods, 435-9 can be removed

This several lines serve as a smooth transition from section 2 to section 3 so that the readers can easily follow the storyline.

491 why the varying significance level compared to 484?

The revised manuscript uses the same 5% significance level.

497 I think I already asked that. For clarity I would use say A and lowercase psi for amplitude and phase speed, but use an overbar or brackets to indicate averages (right now A and uppercase psi are not defined?)

You pointed out the confusing part near line 398.

Accordingly, we have revised that paragraph so that A and  $\Psi'$  could be easily understood..

743 this seems a repetition from the Methods section (same for 825)

This repetition has been removed.

1043 'Niño'

Corrected.

1069 Is it shown? Perhaps if not the paragraph can be shortened or deleted.

As suggested, the paragraph has been shortened.

### Reviewer 3

This manuscript has been revised nicely, but I still have concerns related to statistically significant test for the differences between El Nino and La Nina (Fig.2c, Fig3c., Fig.5c and d, Fig6c, and d... and the rest of the figures contain the El Nino- La Nina panels). I have also advised the same during the previous revision but the authors are showing their inability by responding that sample spaces consist of monthly data.

In the text, the authors have mentioned the number of months in El Nino and La Nina sampling. The statistical significance of the composite difference can be evaluated between these sampling using a two-sided Student's t-test, assuming two independent samples in each period. Without significant test, the panels for the composite difference between El Nino- La Nina are irrelevant.

Supplement has been created to show that the various composite differences between El Nino and La Nina are statistically robust.

To avoid seasonal variability within the cluster, the de-seasonalized anomalies must be use for the composite difference analysis.

In lines 276-278, the newest tracked version of the manuscript contains "For the sake of consistency, we also apply this same filtering procedure to all other fields simulated by GISS E2.2 models such as OLR, zonal winds, resolved wave forcing, parameterized GW forcing, absolute convective momentum flux, etc." In other words, the de-seasonalized anomalies were used for the composite difference analysis in this study. Note that only observed zonal winds (i.e., the FUB data) were not de-seasonalized. However, the FUB data were not used for any composite analysis in this paper.