

Interactive comment on “Defining metrics of the Quasi-Biennial Oscillation in global climate models” by Verena Schenzinger et al.

Anonymous Referee #2

Received and published: 19 January 2017

This study evaluates the QBO as represented in recent climate models, using the small number of CMIP5 and CCMVal models that represent the QBO. The main point of the study is establish the set of metrics that are used here to characterize the QBO, and the authors advocate that these metrics be used in future multi-model comparisons such as those expected from the SPARC QBOi activity. An interesting finding is that the models, on average, have QBOs that are shifted upward and are meridionally too narrow, in comparison to reanalyses.

The proposed metrics are potentially a timely and useful contribution. However I have some issues with the way in which they are presented:

1. The method for calculating the metrics should be presented in a crystal-clear, algorithmic fashion. Since the point is for future studies to repeat these calculations on

C1

different models and/or reanalyses, it needs to be very clear how to do this. I don't think the description of the calculations is sufficiently clear in the present draft. Please see detailed comments in the line-by-line remarks, below.

2. The metrics are presented as-is, with virtually nothing being said on why these choices were made and not others. For example, other ways of defining the QBO amplitude have appeared in the literature, such as Baldwin and Gray 2005. It would be useful for the authors to make the case as to why they settled on these particular choices. Otherwise I speculate that later authors might choose different metrics to characterize the QBO, if this paper hasn't convinced them that the choices made here are well founded. For example, why not just simply use the RMS monthly-mean zonal-mean wind amplitude at a set of standard pressure levels as the measure of QBO amplitude?

3. The metrics in Tables 3 and 4 have no uncertainty estimates associated with them and I see no reason for that omission. The results are mostly given to three significant figures but there is no sense of how meaningful this precision is. Table 6 does give estimates, associated with the multi-model ensemble spread. But for single models (and reanalyses), shouldn't it be possible to give uncertainties based on the internal variability? That is, the variation between QBO cycles.

Based on these issues, and other detailed comments below, I recommend major revisions.

Some other suggestions:

1. Plots for individual models would be useful as supplemental material. For example you could make Fig. 4 for each model individually.
2. Amplitude of the QBO in temperature at the tropical tropopause would be a useful metric. You would need to define the tropopause, but perhaps even just providing the amplitude at 100 hPa would be a simple and useful way to do it.

C2

3. It might be useful to state, in your discussion section, what interesting properties of the QBO are not captured by these metrics. For example some characterization of the zonal momentum budget would be interesting. I'm not suggesting the paper needs to include that, but it would good to state why it doesn't. Data not available in the CMIP5 archive? A desire for simplicity?

4. Histograms showing the distribution of QBO period in each model could be useful (a multi-panel plot, one panel per model). Fig 6 is useful, but the models might show interesting variations amongst themselves. It would show whether some models tend to be more synchronized with the annual cycle than others.

Comments/suggestions by page and line number:

Page 1

6: "QBO like" -> "QBO-like"

12: "ERA-Interim" -> "ERA-Interim reanalysis"

16: insert "known" after "repeatable"

16: ", beyond" -> " outside of"

17: From Osprey et al and Newman et al I think we've learned that whether the QBO remains as regular in present-day climate is also an outstanding question!

Page 2

2-3: Join these two paragraphs together, since they both describe the basic QBO theory.

7: insert "reasonably" before "realistic"

8: "Follow on" -> "Follow-on"

8: delete "signal"

C3

9-10: "depends, amongst others," -> "is affected by"

14: Four out of thirty sounds pretty bad, but on the other hand many of these models might have poor stratospheres in general, with model lids below the stratopause. Do you have an estimate of how many of the CMIP5 models can be regarded as "stratosphere-resolving" but still don't produce a QBO?

17: "aims" -> "aim", "are" -> "is"

20: "An additional purpose is to provide" -> "The purpose is to provide"

21: "the future QBO simulations" -> "new QBO-resolving" (so as not to suggest that only future projections are of interest)

29: "Merra" -> "MERRA"

Page 3

1-2: Suggest deleting this first sentence, it doesn't really add anything. You might instead start this section by introducing Figure 1, since otherwise the figure is first introduced in parentheses near the end of the paragraph, which is easily missed.

4: "was quickly established": not sure what you're referring to here. A previous project comparing QBOs in different models?

5-6: "a typical oscillation with one constant restoring force": I'm not sure what this means. Perhaps you mean "a single restoring force"? For a simple pendulum, $F = -kx$ (Hooke's Law), so F is not constant (its magnitude and direction change). And "typical" is an odd choice in this context: do you mean in comparison to other atmospheric oscillations? It might be simpler to just say that the QBO period is variable, and then go one to explain (as you do from line 6) what might be the causes of the variable period.

8: "these different aspects" -> "the different aspects of the QBO"

C4

8: "Figure" → "e.g., Figure"

9: suggest delete ", for example,"

10: add comma after "extent".

11-17: On p. 2 you say, "The aims of this paper are to establish a set of standard metrics that comprehensively characterise the QBO." To be used by subsequent studies, the procedure for calculating these metrics needs to be unambiguous. I suggest you provide here a very clear algorithm (set of steps) that you used to calculate the metrics. Something like Charlton and Polvani 2007 ("A new look at SSWs, Part I"), Sec. 2b, is ideal: a numbered list of clearly described steps. Otherwise the reader has to fish through the text for the details, and it is easy for you to inadvertently omit some details. For example, in the caption of Fig 2 you say, "The Fourier harmonics around 2 years are averaged". You need to define the exact range of periods used. They are indicated by vertical lines in the left middle panel of Fig 2, but numbers need to be given so that the diagnostic is reproducible by others. It would also be worth mentioning that this introduces a dependence on the QBO period into all subsequent metrics that are based on the averaged Fourier amplitude, depending on the degree to which a given model's QBO period (which is variable) falls within the chosen range.

13: "height" → "altitude". Similarly in Fig. 2 title of bottom panel.

14: "QBO period" → "distribution of QBO periods"

16-17: What is the min/max amplitude "from each QBO cycle"? Is it just the min/max wind, or wind shear? If so then remove "amplitude", or otherwise define how amplitude is calculated for a single QBO phase. Also state explicitly whether it's a wind amplitude, or vertical wind shear amplitude, or both, that you're calculating. You say "shear zone", but you're discussing a time series of the wind at a single altitude.

18: I think you mean the sum of squared amplitudes of Fourier harmonics that fall between the min and max QBO periods? State how the min/max QBO periods are

C5

determined: are these assumed values? (see comment for lines 11-17, above). This is potentially misleading because in the previous paragraph you said that the min/max QBO period is determined from the timeseries of u at h_{\max} . But I assume you can't be referring to these periods here because h_{\max} hasn't yet been defined, since you're describing here how you determine the latitude-height structure. So the order of presentation between the previous paragraph and this one is confusing. A clear, algorithmic description of how the metrics are calculated could fix this.

22: "maximum amplitude" → "maximum"

23: Why is a fitted Gaussian used? Why not just use the latitude-altitude structure itself, as was done for the vertical depth? If a Gaussian is required for some reason (the reason should be stated), is it always a good fit? Does the fit quality vary amongst models? I'm worried that in comparing the values of this metric for different models, if a Gaussian is a good fit for one model but not another then the comparison may be less meaningful.

23-24: "The QBO Fourier amplitude...": this sentence seems out of place here, since you have already referred to the maximum. Also, still unclear what is "maximum amplitude": is it just the maximum? The term "amplitude", here and leading up to this point, seems to be used carelessly. Amplitude is itself a metric, which can be defined in various ways, e.g. RMS amplitude of a time series.

27: "subsequent u values of opposite sign" → "values of u having opposite sign at adjacent gridpoints" (or similar. "subsequent" seems the wrong word here)

Page 4

6: "The progress... is noticeable": Do you mean from older to newer models in your set of models? If so, you could refer to Table 1 as indicating the vintages of the different models (by the year of the references given). Or, if you mean with respect to earlier results in the literature, please provide some specific comparisons.

C6

11: insert "on average" after "QBO structure"

11-12: Table 5 shows that the models and reanalysis disagree on h_{\max} , i.e. the model error bars do not overlap the reanalysis value. So it seems incorrect to say that h_{\max} in the models is realistic. This is also clear from Table 3, first column (h_{\max} is 10 hPa for all but three models). The disagreement is consistent with your general result that the QBO in the models is shifted upward with respect to reanalyses.

16: insert " (Figure 4)" after "temperature amplitude"

Page 5

2-3: Does the timing of phase transitions agree better between obs and reanalyses if you exclude some of the older reanalyses, such as NCEP1/2 and perhaps also JRA-25?

5: In Table 5 I count ten models and eight reanalyses. Also, you assessed the observations (FUB winds).

8: "was established" and "was assessed" (previous paragraph used past tense - be consistent)

11: I'm not sure where you commented on the variability of the QBO period in the models. Table 3 shows the min/max period, but plots of the distribution of periods would be more informative.

12: "narrows" → "is narrower", and "stronger than" → "than in"

14: I'm not sure this the correct way to state Haynes (1998)'s result. That paper shows that the QBO width not set by the width of the forcing when the imposed wave forcing is prescribed to have a very wide latitudinal distribution, designed not to impose a latitudinal scale on the QBO. I don't see that it rules out the actual forcing having a latitudinal distribution that might affect the QBO width. You note that the width of the ITCZ and/or imposed gravity wave sources may play a role, and I agree.

22: "coupled" → "coupled to"

C7

26-27: In Table 5, the standard deviation of descent rates for the models is the same for westerlies and easterlies. Either this statement is wrong or Table 5 is wrong.

31: If you mean that increased resolution leads to better representation of the wave forcing, perhaps change "(subsequently)" to "concomitantly"

Table 1

- According to the text (p. 2), there are four CMIP5 models, not three. I believe CMCC-CMS is also a CMIP5 model, and shares many similarities with MPI-ESM-MR. Please correct the caption.

Table 3

- are confidence intervals for some of these columns appropriate? e.g.mean period.
- why are the descent rates reported with fewer significant figures than the other metrics?

Table 4

- for temperature, lowest level (as in Table 3 for wind) would be a useful metric.

Table 5

- For the reanalysis column, a number of the error values are zero.
- "Values are means and standard deviations of the metrics in Tables 3 and 4" → "The mean and +/- one standard deviation of the metrics in Tables 3 and 4 are shown."
- "excluding CMCC-CMS and both" → "excluding both CMCC-CMS and"
- change "Depth" in the table to "Lowest level", to be consistent with Table 3
- why are the min/max periods not included? (all other metrics from Tables 3,4 are included)

Fig 1

C8

- It would be helpful to expand this figure in the vertical (pressure) direction. Right now all the panels look kind of squished.
- Label the middle panel to indicate that h_{max} is the blue horizontal line.
- The blue and red lines in the middle panel are helpful. It's good how they correspond to the colours of the lines in the top, right, and bottom panels. But the dashed line style makes it easy to miss the colours. Perhaps make these solid lines.
- It would help to add arrows between the panels indicating the algorithm for calculating the metrics. That is, an arrow from the left (Fourier spectrum) pointing at the middle panel (latitude-altitude QBO amplitude), and then arrows from the middle panel point outward at the other three panels.

Fig 4

- Since the filled contours show the model bias (with respect to reanalyses), it would be more conventional to show the model-minus-reanalysis difference.

Fig 6

- This is subjective, but I find it very hard to compare the shapes of the three datasets in this format of plot. You might consider using a six-panel plot to show these results. You could have the phase transition direction as the row and the datasets as the columns (the plots could be narrower with only one dataset shown on each one).

Interactive comment on Geosci. Model Dev. Discuss., doi:10.5194/gmd-2016-284, 2016.