Response to Reviewer Comments

"Origins of Multi-decadal Variability in Sudden Stratospheric Warmings" by Oscar Dimdore-Miles et al.

We thank all the reviewers for providing their comments on our analysis. Their questions and suggestions have helped us to consider the role of ENSO and the PDO in multi-decadal SSW signals more closely as well as make our description of our wavelet methodology and the interpretation of wavelet plots clearer to the intended reader.

Summary of major changes

- Additional analysis using multi-linear techniques (new section 3.2) to explore the comparative roles of ENSO, AL and QBO forcing (Tables 1-3) ; the results support the wavelet results and strengthen our conclusions on the role of the QBO amplitude modulation.
- Section 2 (previously 'Model and Data') renamed 'Methodology' to better reflect its content; now includes an improved description and justification of the wavelet technique;.
- All wavelet analysis figures have been replotted using consistent colour scales for easier comparisons.
- Better justification of which months have been analysed, including an extra supporting figure showing results for NDJFM.
- Improved discussion of the potential role of the PDO and an additional supporting figure.

General Comments: this new study, the authors using a CMIP6 pre-industrial control run from the UKESM global climate model (GCM) to its evaluate internal variability in sudden stratospheric warming (SSW) events. While there is some limited evidence of SSW multi-decadal variability in observations, the atmospheric reanalysis record is too short to completely address this issue (e.g., internally versus externally-driven). Here, the authors use several transformation methods of wavelet time series analysis to investigate SSW variability in the control simulation and the physical sources that may be associated with it. In particular, the wavelet analysis reveals an important connection between the (deep) westerly phase of the Quasi-Biennial Oscillation (QBO) and multi-decadal periods of little to no SSW activity. In agreement with earlier studies, the vertical structure of the QBO is important for addressing the Holton-Tan effect, and thus, also SSW variability.

Overall, this is a very interesting study that addresses an open issue of multi-decadal variability in stratosphere-troposphere interactions, which will be of significant interest to the weather and climate communities. In particular, there have been few attempts to assess SSW variability using a GCM with a well-resolved stratosphere and a realistic internal QBO. However, I think a few more caveats should be more explicitly mentioned given that models still struggle to simulate dynamical coupling between the stratosphere and troposphere. Further, this study is only using one model (UKESM).
Recommendation: The paper will be acceptable for publication in Weather and Climate Dynamics after some major revisions.

Thank you for your supportive comments. We have added some text to the final section outlining the caveats noted (lines 540-547), and also where appropriate throughout the text.

While I am not overly familiar with some of the wavelet transformations employed in this study, more caution (or additional analysis) is needed for interpreting the potential sources of SSW multi-decadal variability (e.g., tropical SSTs – ENSO) that are investigated here.
We have added some substantial new analysis using multi-linear regression (please also see response to a similar issue raised by reviewer 1) including three new tables (tables 1-3) of results as a lead-in to our wavelet analysis, with added text discussing the relative merits of each type of analysis, together with improved description of wavelet analysis and what it can tell us. The new analysis supports our interpretation that while there is some contribution from long-term variability of the ENSO (and PDO, which is also now discussed) there is a stronger link to the deep QBO.

Specific Comments

1.L10-11; Do you think they account for some SSW variability or could it just be coincidental internal noise?
The Aleutian low exhibits coincident power for approximately 100 years which is probably not just internal random noise, but the selected SST regions (particularly tropical west pacific) shows significantly less coincident power with SSWs and is more likely coincidental internal noise. The message we wanted to convey with this statement is that the surface indices we examined cannot sufficiently account for the SSW variability over the whole ~400 years which is why we look elsewhere (namely the QBO region) for additional sources of multi-decadal signals.

2. L20-21; Reference Domeisen et al. (2020) for importance of SSW to S2S forecasts
Reference added

3. L25-26; Cohen et al. (2009) investigated changes in wave activity/surface forcing on stratospheric variability
Reference added

4. L29-33; Seviour (2017) attributed the recent weakening of the polar vortex to internal Variability
Reference added

5. L91-95; Reword this sentence to improve clarity
We have split up this sentence and reworded this to
"SSTs in other tropical regions also exhibit coherence with the vortex. Rao and Ren (2017) show that Tropical Atlantic SSTs give rise to a vortex response although 100 it is highly variable throughout the season while Fletcher and Kushner (2011), Fletcher and Kushner (2013) and Rao and Ren (2015) propose a tropical Indian Ocean (TIO) connection. Positive TIO SST anomalies lead to a reduced strength of the AL that weakens the Rossby wave forcing of the vortex, an opposite effect to the ENSO-vortex connection where positive SST anomalies leads to vortex weakening."

6. L96-101; A very brief discussion would be helpful here to mention other surface

forcings that may modulate the strength (and perhaps decadal variability) of the stratospheric polar vortex in observation records (Garfinkel et al. 2010). Moreover, recent studies have found that boundary conditions, such as sea ice and snow cover, may modulate the Holton-Tan relationship or even QBO cycle (Hirota et al. 2018, Labe et al. 2019). For example, SSW variability may occur through enhanced vertical wave activity due to Arctic sea ice loss (e.g., Kim et al. 2014; Nakamura et al. 2016) and/or Eurasian snow cover anomalies (e.g., Cohen et al. 2007; Henderson et al. 2018).
An additional paragraph has been added (lines 104-113), outlining the possible role of these different surface forcings.

7. L110; Why is this unexpected? The vertical structure of the QBO has been identified in numerous studies for its importance to variability of stratosphere-troposphere coupling. Perhaps reword to improve reader clarity.
We were referring to the amplitude modulation of a deep QBO, which to our knowledge has not been discussed previously (line 128-129). We have reworded to make this clearer.

8. L130-131; Change to something like: "To compare the climate model with the recent observational record, we use ERA-Interim reanalysis (Dee et al. 2011)."
We have changed this as recommended.

9. L204; How sensitive are the results to your choice of SSW definition?
The number of identified SSWs does not change substantially if we use slightly different variants on the definition we have employed, but we are reluctant to repeat the analysis using substantially different definitions of the SSW such as the moment analysis employed by some recent studies because this would be a large amount of work; Butler et al. 2015 suggest that SSW rates are relatively robust to event definition in reanalyses. Extending the analysis to employ moment analysis and perhaps investigating in terms of split / displaced vortex SSWs would be very interesting, but outside the scope of this current study.

10. L213-214; Restate the definition of the ENSO3.4 index here.
Added definition on line 249.

11. L214-216; Why is this metric chosen as a proxy for the Aleutian Low, instead of something simpler like the central pressure as in Overland et al. (1999)? Reference?
The EOF method is from Chen et al. 2020 (a reference for which we have now added) who inspect the 1st EOF loading pattern of the North Pacific SLP then take a box average over a region where this pattern maximises. We use this measure as opposed to a fixed box method to allow for the possibility that the centre of the Aleutian Low in the model does not line up with observations. Taking the PC as opposed to a box over a maximum region in the EOF field seemed a cleaner method but below we check its robustness. We have amended the method with some better explanation of our metric (lines 250-255).

(Chen et al. 2020 show the SLP variability over the Aleutian islands is indeed the dominant EOF which explains 30% of total SLP variance in the northern pacific region. We have further analysed our metric and find a higher proportion of variance explained by the 1st EOF (38%) and that the 1st PC timeseries is highly correlated (r = 0.95) with a box area average over the maxima of the 1st EOF).

12. L228-229; In my view, there looks to be a statistically significant difference in the number of SSW events distributed per month in Figure 1.

This is true; we have amended the text acknowledging this bias in mid winter SSW rates but also outline that this type of bias in a model is relatively common and may originate in the discrepancy between dataset lengths (following analysis of *Horan and Reichler 2017),* We have also added a discussion of November SSWs, in response to comments from reviewer 3 (lines 270-278).

13. L230-240; Although a comparison between model and reanalysis is great, it should still be noted that the samples are not completely comparable if SSWs are influenced by external forcing (climate change) in the real world.

We have added a caveat stating this (lines 274-278). We also mention that the nature of climate change signal in SSWs is not well understood as outlined in Ayarzagüena et al. 2020.

14. L239-246; Again, additional caveats about the use of one model for this analysis are needed... i.e., difference in QBO period (common in high-top models), which could affect the overall conclusions.

We have added a caveat and acknowledgement of possible influence of these biases (lines 540-545)

15. L321-322; This 90-year periodicity looks somewhat large though in Figure 8? Is there something physically-related to this or is it internal noise? Have you done any lead-lag or regression composites to further investigate any ENSO-SSW relationship at this time-scale?

Apologies, not all the spectra used the same color scale and our discussion focussed more on interpretation of the significance contours and not the power values. We have re-plotted all spectra figures to show the same shading levels. When this is done one can see that ENSO exhibits significantly weaker power at relevant periods than the QBO metric. It is also worth pointing out that each spectra is normalised by the variance of the time series (this was probably not clear from our existing text and we have also rectified this in section 2.3).

We have also included new analysis using multilinear regression analysis, comparing the contributions of ENSO, the AL and the deep QBO amplitude (tables 1-3) to SSW_5yr however we stress that, while results from this approach are easy to interpret, they do not directly tackle the problem posed here for two main reasons:

- The signals observed on the wavelet spectra are non-stationary (only persistent for ~450 years of the simulation) and therefore a regression analysis of the entire series may not fully reveal relevant signals.

- Regression analysis considers variability in time series on all timescales whereas we focus on 60-90 year variability when examining our cross spectra. This could be overcome by filtering the timeseries and this is included in analysis and mentioned briefly.

The regression analysis has quantitatively verified the results indicated by wavelet spectra and has strengthened our interpretation – we thank the reviewer for this suggestion.

16. L360; It could also be just noise in the short reanalysis record.
This may be the case, however both Lu et al. papers present a relatively compelling statistical case for fluctuations in HT strength as well as evidence for a physical explanation for the variability. We have added some text to note the issues of extracting signal from noise in such a short data sample on line 274.

17. L366-367; This is difficult to see in Figure 10. Could the left six panels be modified Slightly?
The plots have been modified to make the lines thinner so that the amplitude modulation is more obvious (and also figure 11 for consistency).

18. L375-378; This conclusion seems particularly sensitive to the QBO definition. Any Thoughts?
Yes this is a fair assessment. We want to be transparent in our choice of QBO measure so have demonstrated the differences in results when different QBO metrics are used. All definitions show power at the 2-4 year periods (figure 10) but the longer-term variability is much noisier and more sensitive to the QBO level employed. Some of this may simply be noise (especially the shorter-lived responses) but some are likely to be real sporadic connections, further emphasising the presence of non-stationarity in the signal. It may be that the HT relationship is sensitive to different QBO levels at different times, perhaps depending on the strength of the planetary wave forcing from the troposphere, or on the predominance of wave 1 or wave 2 forcing, which could vary over time in some way. However, the presence of an extended response at ~60-90 years only in the 20 hPa and the deep QBO index plots supports our suggestion that the amplitude modulation of the westerly phase of the deep QBO (which can be seen by eye in fig 10a-f) is important at these timescales.

19. L409-412; Where is this shown?
We have not shown this explicitly in a figure - it was calculated from the ERA-Interim data shown in figures 2 and 3.

20. L416-418; Reword sentence to improve clarity.
We have reworded this to "In particular, the deep QBO index exhibited significant signals coincident with those in SSW_5yr corresponding to periodicities of around 90 years." (now found on line 533)

Technical Comments:
1. L6; ". . . coupled Atmosphere-Ocean-Land-Sea ice model." to "coupled global climate model."
Changed

2. L46; ". . .and the [stratospheric polar] vortex."
Added in stratospheric polar

3. L49; "link" to "effect"
Changed

4. L74 and throughout; Unless you are talking about the vertical structure of the Aleutian Low, change “depth” to something like “strength” or “intensity”
Changed to intensity throughout.

5. L91; Lowercase “tropical”
Changed to tropical