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.

The present manuscript analyzes the sources of multi-decadal variability in major Sudden Stratospheric Warmings (SSWs). To do so, the authors apply a wavelet spectral decomposition method to the output of a 1000-yr pre-industrial control simulation of the UKESM model. The results reveal a periodicity of SSWs variability of approximately 60- 90 years during 450 years of the simulation. Among the studied phenomena, long-term variability of the (deep) QBO amplitude seems to be the most important. In contrast, variability in tropical sea surface temperatures and Aleutian low do not seem relevant to explain long-term SSW variability. The analysis of multidecadal variability is a very interesting topic that is recently receiving increasing attention. However, only a few papers have addressed it due to the unavailability of long observational data record and the relatively low number very long model simulations with daily output. Thus, a 1000-yr pre-industrial control simulation as that used in this paper provides a good opportunity to perform such a type of analysis. The findings identified in this study, in particular, the connection between long-term variability of QBO and SSWs, are certainly promising. However, they are mainly focused on the results provided by the wavelet analysis, and I miss an attempt to explain the detected relationships. Based on this and my following specific comments, I think there are major corrections that the authors should address before recommending its publication.

Thank you for your supportive comments. We agree that the majority of our results are derived from the wavelet spectra analysis. We have added additional text explaining why we believe wavelet analysis is the best technique to use, and we have also added a new section (3.2) where we show results from a multi-linear regression analysis (Table 1) and discuss its limitations (as a lead-in to our wavelet analysis). We also present some further regression analysis (Tables 2-3) later in the results section that provide supporting evidence for the wavelet analysis results. We hope the reviewer agrees that this provides more balance to our study, as well as strengthening the conclusions.

We do not believe it is possible to explain the detected relationships without substantially more research (and we would prefer not to speculate at this stage). At this stage of the research, we report on a comprehensive investigation of the relationships that we expected to see (primarily from the surface in terms of SST and associated variability) and found them to be inadequate. We then explored alternative relationships that involved the QBO, in particular an amplitude modulation of the QBO that had not previously been discussed to our knowledge. We are planning a follow-on paper where we will explore the origins of this QBO amplitude modulation (which may still be related to SSTs via equatorial wave forcing – we are not sure) but the current paper is already overly long. Similarly, explaining the relationship between the amplitude modulation of the deep QBO index and the strength of the polar vortex is also a challenging task which will require targeted model experiments to understand the wave mean flow interactions. We have some experiments underway that will be reported separately.

L21: There are earlier papers that relate the occurrence of SSWs to cold air outbreaks in mid-latitudes such as Tomassini et al. (2012) or Lehtonen and Karpechko (2016).
References added.

L130-131: The historical simulation might be better for model validation than the preindustrial one. Changes in GHGs concentrations may affect the distribution of SSWs.
We have added a caveat stating the possibility of an anthropogenic warming signal in the ERA-Interim warming rates (lines 270-277). We also refer to Andrews et al. 2020 for validation of the historical simulation of the model and agree that consideration of such work is good practice.

We caveat this however, with the fact that the nature of climate change signal in SSWs is not well understood (and varies significantly across different models) as outlined in Ayarzagüena et al. 2020.

L214-216 and L339: In the first sentences the authors indicate that the Aleutian low index is computed for September-November, but along Section 3.4 it is written that it is computed for December-March. From my point of view the later would be more accurate, as it will be simultaneous to the months considered for the occurrence of SSWs.
The statement in the first sentences was a typo and should read Dec-Mar as this is the month range used in the figures. We tested a range of Autumn and Winter month ranges and the spectra are relatively robust across the season on the timescales we consider (60-90 years). As the reviewer points out, Dec-Mar covers the same months used to define SSWs and was a natural choice to use encompassing mid-late winter. We have added a statement of this testing of different month ranges on lines 254-255.

L221-229: I am surprised that the frequency of SSWs in November does not appear in the figure. Moreover, the sum of the monthly frequency of SSWs for December-March in the model does not give 0.54 events/winter but approximately 0.1 events/winter less.

Is it possible that November shows a frequency of around 0.1 SSWs/winter? If so, that means that UKESM is one of those models that presents a too weak vortex in November and an artificially high frequency of SSWs in that month. I was wondering if this is the reason why the authors restrict the analysis to the period between December-March. I would agree that SSWs in November are unrealistic but I have some concerns about not considering them. The occurrence of SSWs in November will precondition the state of the vortex in December-January as it will be recovering and probably anomalously strong. In that case, it will prevent the occurrence of an SSW. This might distort the results for the whole winter. I think it might be worthy to repeat the analysis considering the November SSWs to check if conclusions remain the same.

This is a fair point and the reviewer is correct in suggesting that UKESM has an unusually high SSW rate in November and that this was the reason for omitting them from the analysis – an artificially high number of November warmings may have obscured the QBO, ENSO and AL signal in SSW_5yr if the origin of this bias was elsewhere (such as due to early formation of the vortex as suggested in Mennary et al. 2018 which analyses HadGEM3GC3.1, a sister model of UKESM1). we have discussed the November bias in the text on lines 275-277 and have carried out checks to ensure that excluding November does not significantly influence our main conclusions; in view of this we decided that it was not necessary to amend the figure

We have calculated the spectra for SSW_5yr which includes November warmings and included it as supporting figure A1. It shares the key features of the spectrum presented in the paper – persistent power at 60-90 years for ~450 years of the simulation.

L 315-318: The mature phase of El Niño events is reached from November to January (Wang 2002). I agree with the authors that it will likely remain in the same state between early and mid-winter but I think it will be important to check it.

We agree with the reviewer that this is an important check. When we evaluate ENSO3.4 index in Nov-Jan (as Wang (2002) recommend), we obtain an index that correlates highly with the corresponding ENSO3.4 evaluated in Sep-Nov (r = 0.94). Furthermore, this correlation applies on multi-decadal timescales shows by the correlation between the 2 ENSO indices after they have been fourier filtered to retain only power corresponding to periods greater than 60 years (r = 0.92). This suggests our results regarding ENSO are robust to the month range chosen. We have added text in the methodology section to say that sensitivity checks were performed to ensure that results were not sensitive to the choice of months (line 255).

L318-321: Some authors have related PDO and ENSO (Verdon and Franks, 2006). Is the low frequency of ENSO related to the PDO in this case?

Thank you for raising this possibility - the PDO was not discussed sufficiently in our analysis. We have added text (within section 3.4, lines 398-400) to note the possibility that the ENSO signal at 90 years may be a manifestation of the PDO. We have added supporting figure A3 to discuss its importance. However, while we agree that it merits further investigation, we believe this to be outside the remit of the current study, given that it does not account for a large portion of SSW_5yr power (and not as much as the QBO metrics can).

L337-353: I was wondering if the lack of correlation between the variability of Aleutian

Low and SSW can be explained by the fact that the region selected for the Aleutian Low does not coincide with that associated with precursors of SSWs. In this sense, Garfinkel et al (2012) investigate the reason why the SSW frequency in El Niño and La Niña winters is similar in observations. The find that both La Niña and El Niño lead to circulation anomalies of the same sign in the area associated with SSW precursors.
I think it would be important to identify the areas of precursors of SSWs following Garfinkel et al. (2012) or Garfinkel et al (2010) and compare with the spatial pattern of the Aleutian Low that the authors compute.
This is another fair point as we have used a PC based method to pick up the variability of the AL as opposed to a box in Garfinkel et al. (2012) which is associated with SSWs. The box suggested by Garfinkel et al. (2012) (52.5°N–72.5°N, 165°E–195°E) overlaps slightly but lies slightly further north of the centre of the 1st EOF of SLP so merits some analysis to check if this may be a better metric to use.

An index defined as the area weighted average de-seasonalised SLP anomaly within the box exhibits a significant but somewhat small correlation with our existing metric ($r = 0.4$, $p < 0.01$) suggesting there may be some variability missed by our metric. However, the average over this box exhibits lower correlation with our SSW time series than the existing AL index: $r = -0.21$ for our AL and $r = -0.13$ for the box-based AL. This suggests we are not missing a region of high influence over the vortex with our existing measure.

Technical comments:
L44: north –> North
Changed to North.

In some figures such as Figure 11 or 5, the authors use lower case a), b) and so on to refer to the different panels of a figure and in other figures such as 8 or 9 the authors use upper case A) and B) for the same purpose.
Changed all subfigure labels to lower case.

L404: figure 12 b
Remains unchanged (see previous comment).

L409: please include (not shown) at the end of the sentence.
(not shown) added.