Review of 'On the representation of major stratospheric warmings in reanalyses'
by Ayarzaguena
et al. (2019)

This paper has examined how different features of SSWs (e.g., magnitude,
precursors, surface impact)
vary across different reanalysis datasets in both the historical period
(1958-1978) ad post-satellite
era. The authors have also examined the differences in features between
wavenumber 1 and wavenumber 2
SSWs. The paper is overall interesting and is a valuable contribution to the
literature; using historical
data back before 1979 will be useful for the SSW community and this study
suggests that despite the
discrepancies between the pre and post satellite era data, the characteristics
of SSWs in different
reanalyses act fairly similarly. All of my comments below are minor and hence I
suggest only minor
revisions.

One thing to note is that this review was not so convenient to write because of
the line numbering. I
have included a line number and a page number for each comment as it appears
that the numbers ran to
35 before restarting over again continuously!

Specific Comments:

Line 20; Can you just confirm whether by the 'surface fingerprint', you mean
either the
downward impact following the SSW, or the near-surface precursors?
Lines 30-32; It is worth mentioning here that SSWs are not always preceded by
precursory
wave activity in the troposphere (most recently for instance, Birner and Albers
2017,
SOLA; White et al. 2019, J.Clim both found that ~30% of SSWs are preceded by
lower-
tropospheric wave activity in observations and in a GCM, respectively). I don't
mean for
you to go into details regarding this, but it would be good to mention that
sometimes the
source of the anomalous wave activity is in the stratosphere. Another good
citation to add
would be Garfinkel et al. (2010), J. Clim who found that a deepened Aleutian Low
leads to
enhanced upward wave-1 flux. In this part of the text you have only mentioned
about blocking
highs preceding SSWs, when many SSWs are preceded by such an anomalously-deep
Aleutian Low.
Line 9, page 2; this line suggests that all SSWs impact the tropospheric
circulation when in

reality, not all do, and only in the composite mean is there an aggregate impact. It would be better
to make this clearer.
Lines 9-12, page 4; How are SSWs in each reanalysis determined to be 'common'? What is the time
window around the actual SSW date in one reanalysis for which an occurence of a wind reversal in
another reanalysis is deemed to be the same date? You just mention here that four out of seven
reanalyses in the common period must show the same SSW event; but, how is the same event determined?
Line 25, page 4; Can you be clearer here? It is not immediately clear how you chose the common SSWs
to be either D or S here. Did you check each common SSW in each renalysis and then determine if the
majority of reanalyses showed either a D or an S? Or was there some other way?
Line 27, page 4; how sensitive are the results to different levels and latitudes? A sentence or two
would be good to describe the sensitivity. Also, was the 200m difference threshold arbitrarily chosen?
Line 1, page 7; can you better explain how these histograms are calculated? It seems to me that for
each date on the x-axis you take a 21-day window (centred on that date), and count how many SSWs
occurred in that window. You then moved on to the next date and did the same. Is this correct? If so,
it seems to me that by doing this, SSWs are counted multiple times and the histogram may not be a fair
representation. What happens when this window is shortened from 21 days? Shortening the window length
will no doubt be a more accurate way to do this. Just creating bar charts of the #SSWs in each month
would be a fairer and less-ambiguous representation and then just compare the distributions.
In terms of the histograms, it would be useful to test the significance between the individual
histograms using a Kolgomorov-Smirnov test. My guess is that they are significantly different in (a),
but not in (b).
Line 14, page 8; how does the HF look below 100hPa? Say down to 300hPa? Are there any significant anomalies?
Between 300hPa and 100hPa is the comunication region for stratosphere-troposphere coupling that de la Camara
et al. (2017) suggested to be particularly important. 100hPa is already in the stratosphere at high latitudes,
and hence, 300hPa may be a better measure of the upward propagation of wave activity from the troposphere.
Lines 23-24, page 8; This is an interesting result. Is the correct interpretation
that prior to lag -5, the wave activity grows in the stratosphere via constructive interference with the
climatological planetary waves, whereas from lags -5 to 0, anomalous wave growth

occurs? I am wondering if this
is indicative of the Plumb (1981), JAS idea of self-tuning resonance? i.e., a standing climatological wave and a
transient anomalous wave interact constructively to give a growing-in-amplitude wave in the stratosphere? This
wave then grows to very large amplitude and eventually splits the vortex. This is more of a probing statement,
as I do not know for sure. But some interpretation as to why the earlier lags are dominated by the interference
term and the lags closer to zero are dominated by the anomalous term, would be appreciated here.
Lines 12-13, page 9; How sensitive are the results in this figure to this lag window? I ask because the lag
window you have chosen is based on figure 5 which only extends down to 100hPa. In figure 7 you present 500hPa.
Do the significant HF anomalies below 100hPa extend further back in time to before lag -10? If so, then this
would suggest increasing the length of the lag window.

Technical Comments:

Line 26, page 2; what is the 'second one' here?
Line 32, page 2; Here seems a good place to start a new paragraph when you start talking about
the aims/methods of this paper.
Lines 3-4, page 3 (top of page); I think you also examined the downward impact of S and D events,
right? Unless you are classifying S and D, and WN1 and WN2 events as the same (although I don't
think you are)
Line 9, page 3; typo. I think you mean: 'The former analyses the momentum budget during SSWs...'
or something to this effect!
Line 23, page 3; did you perform the interpolation yourselves? A sentence or two describing the
method used would be useful - was it a simple linear interpolation? Or something more complex?
Line 32, page 3; Just to clarify, the anomalies are calculated as the departure of the field from the
daily climatology for EACH reanalysis? Or do you mean the anomaly from the daily climatology over
ALL reanalysis products (i.e., away from the MRM)?
Line 33, page 4; why is the 1981-2010 baseline used instead of the full 1979-2012 period?
Line 8, page 4; imposing --> requiring. Also, I think the Charlton and Polvani (2007) paper must
be cited here! As this is, as I recall, the definition from their paper exactly.
Line 24, page 4; Perhaps better would be: '...with respect to the occurrence of an SSW, according
to the definition in section 2.2'

Line 4, page 6; 'two-folded' --> 'two fold'
Lines 22-24, page 6; somewhere it should be mentioned that only the historical period is considered
in figure 1.
Line 6, page 7; What is meant here by 'traced back to the PNJ'? You haven't previously explicitly
calculated the PNJ (which from section 2.4 I understand to be the difference in wind strength prior to
and following the SSW central date). Are you here referring to the PNJ as just the strength of the U at
60N and 10hPa as shown in figure 1c? If so, then the PNJ as defined in section 2.4 needs to be better
articulated.
Line 23, page 7; I think you mean to compare Fig 3,b,c with Fig 3,f,g?
Line 13, page 8; is this area-averaged? i.e., weighted by the cos(lat)?
Lines 23-24, page 8; 'precedent' --> 'preceding'.
Line 28, page 8; Change to 'historical period'
Figures 4-6; Negative contours would be easier to identify if they were dashed rather than solid. This is
particularly true if there is no significance (and hence no shading)!
Figure 7, caption; Only gridpoints with stat sig values are shaded right? The contours are the full anomalies?
If so, line 4 on page 25 needs to be updated (i.e., change 'plotted' for 'shaded') as it is not clear.
Further, the density of anomaly contours is very high, especially considering that much of the plots are
insignificant. Seeing as the WN1 and WN2 climatological centres of action are important in your description,
it would be useful to put one or two contours (say, in green) for each centre on the plot. Hence, I suggest
to reduce the density of anomaly contours and to just put a couple of contours representing the climatology,
which should not clog up the plot.
Line 21, page 9; 'MMR' is meant to be MRM?
Line 8, page 10; so the bottom row should equal the sum of the top two rows? Further, are the units of the colorbar
percentages?
Line 9, page 10; 'al' --> 'all'
Line 19, page 10; 'non-significant' --> 'insignificant'
Line 7, page 12; 'but at much less extent' --> 'to less of an extent,'
Line 32, page 12, change to 'pre- and post satellite eras.'