REVIEW of "On the representation of major stratospheric warmings in reanalyses" by Ayarzagüena et al.

SUMMARY: This paper discusses the representation of SSW events in different reanalysis products. This is an important contribution given the increased use of SSWs for long-range prediction of surface quantities, which are often initialized from and compared against different reanalysis products. This is a timely contribution for the S-RIP project of comparing reanalysis products for the stratosphere.

OVERALL ASSESSMENT: The paper is well written and addresses an interesting and worth-while problem. I have some comments that I hope will improve the manuscript, see below.

We thank the reviewer for the useful comments that have contributed to improve the manuscript. Please see below our replies in blue color

SPECIFIC COMMENTS:
Page 1:
Line 22: "surface fingerprint": does this refer to the signature after the SSW event? Please specify.
Yes, it does. We have replaced "fingerprint" for "response" to clarify it.

Line 26: "lead to": this is not a causal effect, but effects that are linked through thermal wind balance
We understand what the reviewer means and it is true that both the vertical shear of zonal wind and the meridional temperature gradient are connected through thermal wind balance, so it is not easy to determine what is causing what. However, in the specific case of SSWs, changes in the meridional heat flux are the forcing that leads to changes in the wind. Moreover, in many SSWs it is the change in the polar temperature that precedes the maximum deceleration of the wind. Thus, we think the "lead to" is justified in this case.

Line 31 – 34: The literature is rather split about this issue, see e.g. Birner & Albers 2017, Sjoberg & Birner, 2014.
We agree with the reviewer that recent studies have already shown that the enhancement of wave activity prior to SSWs tends to happen within the stratosphere and/or is related to a preconditioning of the mean stratospheric flow. We have included a comment about this in the Introduction of the revised manuscript (new lines 32-34).

Page 3:
Line 12: "analyzes the SSWs the momentum budget": unclear
We have slightly modified the sentences to make it clear. The new sentence reads like this: "The former analyzes the momentum budget during SSWs restricted to the post-satellite period"

Page 4:
Lines 24 – 28: since K. Shibata is a co-author, it would help to clarify the algorithm used in the manuscript in case it's not (yet) published.
According to the reviewer's suggestion, the description of the algorithm has been extended in the new version (see new lines 128-134).

Page 5:
Line 25: anomalies from climatology?
Yes. We added "daily" to make it clearer.

Page 6:
Line 17: The deviation in the results of NCEP from other reanalysis products is not surprising. There's an artificial trend in the stratosphere – we found it in Badin & Domeisen, 2014 (pages 1498/1499). I could imagine there's also an S-RIP publication that documents this problem?
Thanks for the reference. Unfortunately, most of the S-RIP publications (or even earlier papers) that document the worse performance of NCEP/NCAR in comparison to the other reanalyses are focused on the post-satellite era (e.g.: Manney et al., 2003, Long et al., 2017). In contrast, in this part of the manuscript we are addressing the inter-reanalysis differences in the historical period. We are not aware of other S-RIP publications reporting this issue, and hence we have mentioned the artificial trend in the stratosphere found by Badin & Domeisen

(2014) in the first 50 years of the data record and related that finding to our results by the end of Section 3.1 (lines 214-217) and in the Conclusions (lines 389-390).

Page 6/7: I'm wondering if it would be helpful to list the classification for all events, not just the ones that are common
Thanks for the suggestion. However, we think it is not necessary for different reasons. First, as indicated in the text, most of the differences are more likely due to specific thresholds or methodological issues rather than relevant biases in the reanalyses. In addition, we are using this information in Table 2, as a brief overview of the reanalyses' performance when different events are considered based on fixed criteria. The remaining analyses in the manuscript are based on the events shown in Table 1. As the classification requested by the reviewer is a result of the SRIP initiative (to be included in Chapter 6 of the SRIP report), we have just added an additional reference to that chapter in Section 2 when talking about the classification of SSWs.

Page 7:
Line 8: "can be traced back to the PNJ": this does not sound like an explanation, rather a symptom
We have modified the sentence to avoid confusion. In particular, we have only related the NCEP/NCAR peak of SSWs in early winter to a weaker climatological PNJ in this reanalysis than in the other two.

Lines 15/16: given the large uncertainties in the pre-satellite period this is difficult to state. However, there are indeed changes in decadal variability of SSW frequency in Domeisen, 2019, JGR, maybe this is helpful?
Yes, it certainly helps. We have included the reviewer's comment and some references to previous studies that reported a multi-decadal variability of SSW frequency (including Domeisen 2019). Multi-decadal changes in SSW frequency could also translate to the intra-seasonal distribution of SSWs. Indeed, in the new version of the manuscript, we have confirmed that the SSW distributions of the historical and satellite periods are statistically significant, according to a Kolmogorov-Smirnov test.

Page 8:
Lines 1-6: maybe it would be helpful to indicate the changes in stratospheric representation between the different NCEP reanalysis tools, or maybe refer to the Hitchcock, 2019 paper?
NCEP/NCAR and NCEP-DOE reanalyses are using basically the same model although with different versions, 1995 and 1998, respectively. Most of the improvements made in NCEP-DOE from NCEP/NCAR are related to changes in the lower levels (troposphere), except for the prescription of a new climatology of ozone (Kanamitsu et al., 2002; Long et al., 2017). Other differences in the concentrations of CO_2 or radiation schemes might also explain the small differences in results between both NCEP reanalyses.
In the revised manuscript, we have briefly extended the description of differences in the setup and models of both NCEP/NCAR and NCEP-DOE based on Kanamitsu et al. (2002), Fujiwara et al. (2017) and Long et al. (2017) (new lines 245-254).

Page 9:
Lines 24 – 26: yes, indeed, this is why it is so difficult to trace waves from the troposphere to the stratosphere. This is not so counterintuitive given the literature on the stratospheric contribution to SSWs.
Following the recommendations of Reviewer#1, in the revised text we have extended the discussion and inserted references to the recent literature on this topic (lines 312-318). In particular, we have stressed the special importance of the initial state of the polar vortex for the occurrence of WN2 SSWs (e.g. Albers and Birner 2014), the type of events discussed in this part of the study. In those cases, an initial vortex structure close to its resonant point is prone to lead to the split of SSWs with a small increase of tropospheric wave forcing.

Line 29: at which level?
We first checked at 10 and 20hPa, where we found the largest values of anomalous heat flux. However, they are probably not the best levels if we are trying to connect those changes with tropospheric structures. We have removed this sentence from the discussion.

Page 11 / Figure 7 / Page 23, line 31: are these differences significantly different from each other? i.e. not just significantly different from climatology?
Yes, they are. Panel c shows WN1-minus-WN2 differences and the shading indicates that these differences are statistically significantly different from each other. We have corrected the figure caption.

MINOR COMMENTS:
Page 1:
Line 30: I would suggest using Charlton et al (2007) as the authoritative reference here.
Modified

Page 2:
Line 8: Martius et al (2009) seems like the perfect reference here, it's already included in a different place in the manuscript
Included

Lines 10 – 16: would it make sense to include the classification into reflective and absorptive events here (Kodera et al, 2016)?
We prefer to keep it as it is, because we are not referring to these events later on.

Line 18: given the very limited number of studies of stratospheric effects on the ocean I would not call the assessment of oceanic phenomena based on the stratosphere a "common metric"
We just meant just the other way, i.e. oceanic effects on the stratospheric variability. Actually, we were mainly referring to the ENSO effects on the polar stratosphere or other phenomena that have also been recently explored such as PDO or MJO. We have slightly modified the sentence to clarify it.

Line 21: leave out "interestingly", and "largely"
Done

Line 22: "assimilation data sources": do you mean the data used for the assimilation of observational data into the reanalysis products?
Yes. This has been modified

Line 27: "than in the second one". Do you mean "than during the satellite era"?
Yes. We have changed it

Page 3:
Line 6: is made on > is given to
Changed.

Line 26: do you mean "across different reanalysis products"?
We meant across different reanalyses, not products. It has been corrected and clarified.

Page 4:
Line 29: "similarly": do you mean the identification was similar or it was also included in the table?
We meant that the identification was carried out in a similar way as for the common dates. We have clarified it.

Page 5:
Line 28: I'm not sure what is meant by "discrepancies" (also: page 6, line 14)
In the first case we have clarified that it means to the lack of consensus on the precursor role of blockings in SSWs. As for page 6 (now line 184), we have just replaced discrepancies for reanalyses results.

Page 8:
Line 9: ones -> SSWs
Corrected

line 19: "reanalysis deviation": not clear what this means
Differences across reanalyses.

Lines 23 – 26: be more clear which terms this corresponds to in the equation
Done

References

Albers, J. R. and Birner, T.: Vortex preconditioning due to planetary and gravity waves prior to sudden stratospheric warmings, J. Atmos. Sci., 71, 4028-4054, 2014.

Badin, G. and Domeisen, D. I. V.: A search for chaotic behavior in Northern Hemisphere stratospheric variability, J. Atmos. Sci., 71, 1494-1507, 2014.

Fujiwara, M. and coauthors: Introduction to the SPARC Reanalysis Intercomparison Project (S-RIP) and overview of the reanalysis systems, Atmos. Chem. Phys., 17, 1417–1452, 2017.

Kanamitsu, M. and coauthors: NCEP-DOE AMIP II reanalysis (R-2), Bull. Amer. Meteor. Soc., 83, 1631-1643, 2002.

Long, C. S., Fujiwara, M., Davis, S., Mitchell D. M., and Wright, C. J.: Climatology and interannual variability of dynamic variables in multiple reanalyses evaluated by the SPARC Reanalysis Intercomparison Project (S-RIP), Atmos. Chem. Phys., 17, 14593–14629, 2017.

Manney, G. L. et al.: Lower stratospheric temperature differences between meteorological analyses in two cold Arctic winters and their impact on polar processing studies, J. Geophys. Res., 108, 8328, doi: 10.1029/2001JD001149, 2003.