Author response to Referee 3 comments

I thank the referee for reviewing the manuscript. I have tried to address all of the comments and revise the manuscript, and hope the revision is satisfactory. Before responding to each comment, I note several changes:
-Figures 10 and 11 now include results from CONV.
-Figure 12 in the previous version has been removed. The figure and discussion were a bit lengthy compared to the messages obtained. I mention the results from the figure in the text only, so that this does not affect the overall argument in this study.
-I have corrected some existing figures (Figs. 2c, 5b, 9a,b), but this does not affect the argument. Figure 5b in the previous version used the (CL, AR) data for STDD on the MSSW onset dates identified in CONV. It now uses the (CL, AR) data for STDD on the MSSW onset dates identified in STDD. In Figs. 2c and 9a,b, a few color shades did not appear as intended.
Please also note that the pages and numbers in my response below refer to those in the (unformatted) manuscript.

Major comments
Style: The paper takes a largely descriptive approach to comparing the datasets until section 6. I really was hoping that the author might incorporate a little more discussion throughout the manuscript, particularly when comparing the JRA-55 and JRA-55C datasets, since there is a great deal more that might be said about the differences here, particularly since they seem to be confined to the upper stratosphere and sometimes significantly influence the timing of SSW events. As noted above, the paper describes the comparisons in quite a lot of detail leading to unnecessary figures and text and making the paper overly long.
The descriptive style of this study is by intention, since we compare dynamical variability in the extratropical stratosphere especially from a morphological point of view. I think the present analysis and manuscript are reasonably organized: survey climatology and variability (MSSWs) among the three products in Section 3, describe RMSD distributions and compare them to the MSSWs in Section 4 (STDD vs CONV), and examine wave forcing and vortex response in Section 5 (STDD vs AMIP).
It will be possible to compare vertical structures of the polar vortex between STDD and CONV for various dynamical conditions. Then, one may find that the vortex has largely different structures especially in the upper stratosphere between STDD and CONV even when the two products have a MSSW on a same onset date. However, it will be difficult to combine such an analysis to the current manuscript, and I therefore retain the overall structure (and most figures, except for Fig. 12) of the manuscript. I have added a discussion paragraph to mention this extension as possible future work (p. 12, l. 20-24).

Methodology: In a number of places the author compares the SSWs detected in the standard JRA-55

dataset with those detected in the other two datasets. I can see the merit in this, but wouldn't it make more sense to compare the same dates when SSWs are detected in the standard re-analysis with the other datasets. Then one can really assess the extent to which tropospheric information contribute to producing the SSW seen in the standard assimilation. Similarly, I think it could be very informative to compare the JRA55-C and JRA55-AMIP simulations in this way.

While Table 1 compares the MSSW onset dates identified in STDD to those identified in CONV and I mention their time differences in places in the text, I also compared/compare features between STDD and CONV on the MSSW onset dates identified in STDD, such as in Figs. 8 and 9f-j. Figures 10 and 11 also use the maximum of the 21 day mean heat flux for DJF identified in each product.

Figure 11, which now includes results from CONV, will contribute to the latter point. It shows that AMIP is an outlier in terms of the zonal wind evolution, while all three products are similar in the other quantifies: AMIP underestimates a vortex response, or a zonal wind deceleration at 10 hPa to a wave forcing of similar strength from the troposphere (which can be regarded as a model bias). STDD and CONV are similar in all quantities of interest, but the zonal wind in CONV is slightly stronger than that in STDD. This seems consistent with the result that CONV misses or delays several MSSWs compared to STDD.

I didn't find the following figures (and associated analysis) added much and suggest they could be removed or added to a supplement. Figures 6, 7, 9, 11, 13.

Please refer to my response to the first comment.

Minor comments

Throughout I find the name 'conventional' a little confusing for JRA55-C since this is really just 'non-satellite'

I follow the paper (Kobayashi et al. 2014) of the developers of the JRA-55C product for the use of the word "conventional". I also explain the meaning of "conventional" at the beginning (p. 1, l. 10).

p2 l15 Could you say more about the AMIP-style integration? Is this a continuous run with just SST and other boundary forcing?

I have added an explanation about AMIP at p. 3, l. 19-20.

p3 l15 Why did the author introduce new nomenclature for the datasets? I think using the standard names would be more useful for comparison with other studies.

Because this study deals with the JRA-55 family data only, the "JRA-55" part in JRA-55 (standard), JRA-55C, and JRA-55AMIP are obvious and redundant, and hence could be removed. One could then refer to the products as standard, conventional, and AMIP: I simplify these by using the acronyms (STDD, CONV, and AMIP), and I think these are clear and easy to remember when reading the manuscript.

p8 l3 What do you mean by ’envelope’

I have removed the phrase using the word “envelope”, as the sentence is clear without it.

Fig. 2 etc. Please remove the wedge where the shading and contours don’t wrap around the zero line

I have improved the figures as suggested.

Fig. 5 I found the lines between the two dates (observed and observed in JRA55-C) tended to obscure the points. I think these could be removed without much loss of information.

I think that the lines are useful since it enables one to compare/connect corresponding cases between STDD and CONV. I have changed the lines and STDD data points into grey, so that they do not disturb the main distribution for CONV in (b).

Fig. 8 The arrows and labels in the panels made the quite hard to read and to distinguish different points. I like this plot but could it be simplified?

I have somewhat simplified the figure.

Fig. 10 Could you show the heat flux distribution at the bottom of each panel for the different cases to make the point made in the text about the shift between the two datasets more clearly?

I have changed the figure accordingly.