Review of 'The role of the stratospheric state in upward wave flux prior to Sudden Stratospheric Warmings: a SNAPSI analysis' by Ayarzagüena et al.

General comments:

This study analyzed how the stratospheric condition influences the upward wave flux, and thus the occurrence of SSWs, based on a set of S2S model experiments in SNAPSI. The S2S forecasts are still struggling to predict SSW with long lead times. Thus, the current results provide helpful insights into understanding the role of stratosphere, which could potentially help improve the forecast skill in the future. The manuscript is very clean and well-written. However, since there are various diagnostic metrics and a huge amount of information, a major revision is needed to help with the presentation and delivery of the main conclusions.

Major Comments:

1. Balance in presenting the model mean and model spread. Currently, most analyses are based on the model mean, while the model spread is not fully presented. For example, in Section 3, while Section 3.1 shows the mean skill of models in predicting zonal wind, Section 3.3 focuses on the difference in Z500 between 'weaker u' and 'strong u' groups. This 'inconsistency' poses a mismatch in the obtained information. It would be helpful to link the spread in Z500 with the spread in zonal wind. Thus, the model spread in predicting the zonal wind can be added in Fig. 1, and the model mean of Z500 should be included in Figs. 5-7. Although the fraction of members forecasting an SSW is shown in Table 3, this cannot accurately reflect the spread. For the latter, including the model mean will help in interpreting Figs. 5-7 as discussed in L360-365.

2. Discussion on the stratospheric state. While I understand that the main focus of this study is to provide a general understanding of the role of the stratospheric state, it would be worthwhile to add more discussion on what might constitute the stratospheric state. For instance, for the 2018SSW, the poleward shift of the PNJ is missing in the models, which is one possible candidate. In addition, the manuscript did not discuss about the QBO (although very briefly in L686), which is also included in the stratospheric condition and reflected by the experiment design. Although at such a short timescale the bias in the QBO may not be evident, the three SSWs occurred under different QBO states (e.g., Butler et al. 2020; Shen et al. 2020), and it would be worthwhile to reflect this and discuss the potential implications. In addition, as briefly discussed in Section 3.1, the intensity of the polar vortex relative to climatology also differs, which could also serve as preconditioning. The related discussion should be added to provide general implications.

Reference:

Butler AH, Lawrence ZD, Lee SH, Lillo SP, Long CS. Differences between the 2018 and 2019 stratospheric polar vortex split events. Q J R Meteorol Soc. 2020; 146: 3503–3521. https://doi.org/10.1002/qj.3858

Shen, X., Wang, L., & Osprey, S. (2020). Tropospheric forcing of the 2019 Antarctic sudden stratospheric warming. Geophysical Research Letters, 47, e2020GL089343. https://doi.org/10.1029/2020GL089343

3. Visualization of the plots. There are quite a few plots which are all informative. However, for some plots it is difficult to identify the regions/features being discussed in the main text. I suggest the authors adjust the plots to make the information more straightforward, which will help readers grasp the key information more quickly. Please see the detailed comments in the Specific Comments.

Specific Comments:

1. L210: Suggest explaining that the negative eddy heat flux indicates upward propagation in the SH. Despite the figure caption of Fig. 4, there is no explanation in the main text. Please add this.

2. L227: The classification of 'weakest u' and 'strongest u' is a bit counterintuitive, as the 'weaker' group corresponds to a better forecast, whereas the 'stronger' group corresponds to a worse forecast. Would it be better to define them as something like 'SSW-like' and 'no SSW', which is more straightforward?

3. L260, Figure 1: Suggest also showing the spread in predicted [u] as in my major comment 1.

4. L278: Suggest briefly stating that WN2 is mainly responsible for the 2018 SSW, otherwise it is a bit abrupt to focus on WN2 directly.

5. Figures 2-4: Since there are lots of lines, I suggest bolding the model being discussed in the main text for visualisation.

6. Figures 5-7:

- Suggest adding the variance of forecasted [u] among the members in the subtitles after the model's name. This could help provide a straightforward comparison and understanding of the linkage between the spread in tropospheric circulation and stratospheric response.
- Suggest also adding the box to indicate the region in ERA5, as the map is not very visual and thus takes time to identify the region of focus.
- The climatological PWs are from ERA5. While I understand this is what one can do, it is also possible that the model bias in the climatological PWs can influence the interpretation of linear interference. Perhaps it would be better to include a brief discussion.

7. L348: Please clarify what 'this center of action' refers to.

8. L380-410: Suggest checking the HF300 as well. As stratospheric wave forcing does not entirely originate from tropospheric forcing as discussed later on and shown in Yessimbet et al. (2022). It would be helpful to establish a linkage between the tropospheric circulation, the tropospheric wave forcing (HF300), and the stratospheric wave forcing (HF100).

Reference:

Yessimbet, K., Shepherd, T. G., Ossó, A. C., & Steiner, A. K. (2022). Pathways of influence between Northern Hemisphere blocking and stratospheric polar vortex variability. Geophysical Research Letters, 49, e2022GL100895. https://doi.org/10.1029/2022GL100895

9. Figures 8 and S1. I like Fig. S1, which is very informative. It not only confirms the linear relationships among the three variables but also indicates the spread among the model members. I suggest the authors move it to the main manuscript, perhaps merging it with Fig. 8 as they are related, and add more discussion on it. For instance, we can see that for SSW2018, the scatters are densely located in the upper left, indicating that the forecast [u] is overall quite strong and related to the weak HF (Fig. S1a). Moreover, for the HF and Z500, the scatters are located in the lower panel (Fig. S1d). Despite the weaker Z500-HF linkage, the Z500 is also quite diverse. For comparison, it would be better to use the same range for x and y axes across different cases, also add the corresponding ERA5 variables. This will also help to interpret the conclusion from Fig. 8.

10. L432-437. Please briefly explain how the quantitative changes are computed.

11. L440: It is interesting that for 2019SSW SH, although the models have relatively good performance in capturing the tropospheric wave forcing (i.e., for the multiEnsM, the observation is within the 1.5IQR), they fail to capture the observed stratospheric wave forcing. This seems to imply that the stratospheric wave forcing does not completely come from the tropospheric wave forcing. Whereas in the other two events, the skill for HF300 and HF100 is more similar.

12. L477: Suggest adding 'for the majority of models' for clarity.

13. L480: Suggest changing 'model simulates very similar anomalies to ERA5' to 'the ERA5 value lies within the IQR ..'

14. Figure 11: Suggest marking the period of interest.

15. L537-538: Suggest changing 'relatively weak wave activity that is simulated by reanalysis and models' to '…wave activity seen in the reanalysis and simulated …'

16. Figures 12-13: Suggest adding the box to indicate the regions for EP flux budget and changing the y-axis labels to hPa for consistency with the main text. In addition, suggest extending the latitudes to 10S to show the QBO structure, related to my major comment 2.

17. Figure 14: Please make sure the vertical magenta line mentioned in the figure caption is visible.

18. Figure 15: Suggest changing the color of the median value line for visualization. In addition, while currently we can see the difference in Fz and Fy, we cannot find their relative contribution to the net EP flux convergence. According to the previous analysis (Figs. 12, 13), the difference in the horizontal wave forcing is evident, however, it does not appear to play an important role according to this plot. This poses a gap between these two sections. I suggest the authors add a scatter plot of Fz100 vs Fy for each member, similar to Fig. S1. This might help in understanding their relative roles.