REVIEWER 1

Ayarzagüena et al. investigate how the stratospheric background state influences upward wave activity preceding Sudden Stratospheric Warmings (SSWs). Using ensembles of different models from the Stratospheric Nudging And Predictable Surface Impacts (SNAPSI) project, the study compares free-running, stratosphere-nudged and control simulations to ERA5 for three events: the February 2018 boreal SSW, the January 2019 boreal SSW, and the September 2019 austral minor SSW. The study aims to isolate the effect of the stratospheric state on the triggers of SSWs. Overall the paper provides a detailed study of the processes and wave fluxes that go into contributing to the different SSWs and attempts to separate tropospheric and stratospheric influences. The paper itself is rather long but the analysis is thorough and the authors walk the reader through the plots and their interpretation. I would recommend publication after addressing the points below.

Thanks a lot for your comments. Here is our reply to the major and minor comments in blue:

1) Figure 1 and Table 3. I have difficulty reconciling the fact that the ensemble mean line for CNRM in Fig 1(a) doesn't show an SSW but 62% of the ensemble members do. Do you have a suitable plot to illustrate the spread please?

First, it is important to note that we calculate the percentage of ensemble members predicting an SSW event during the 45 days of the FREE experiment. Therefore, it is not restricted to a specific day or a short period around the observed SSW date. In the CNRM-CM6-1 ensemble, SSW events occur between 8th February and 11th March 2018 (the final day of the simulation), with a higher probability between 15th and 28th February (Figure R1.1). This temporal spread may then partially explain why the ensemble mean of u60_10 does not cross the zero-line on any of the 45 days of the FREE experiment, but it falls below the ERA5 climatological values from 14th February onward as shown in Figure 1a of the manuscript.

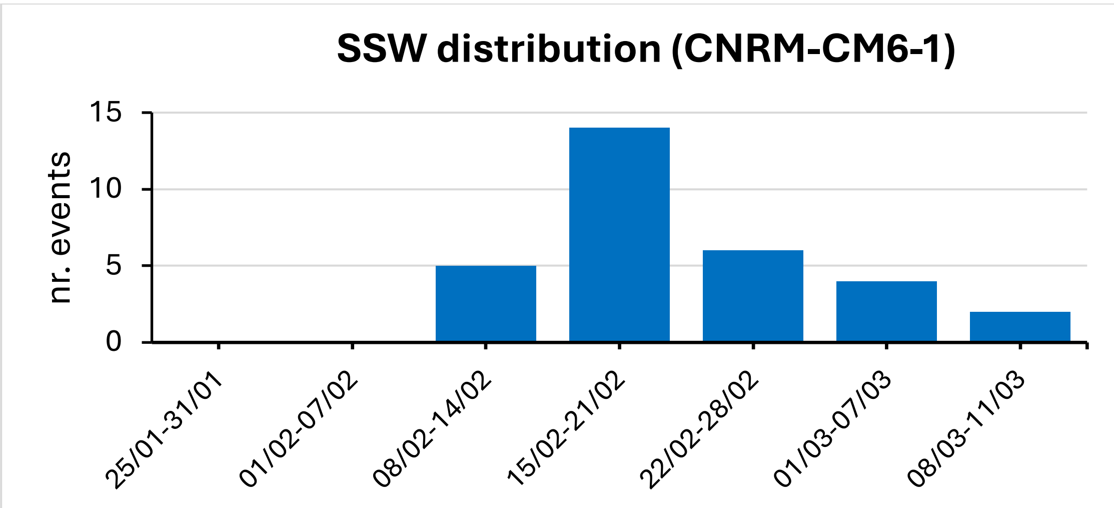

Figure R1.1. *Distribution of SSW events in the CNRM-CM6-1 ensemble members during the 45-days period simulated in the FREE experiment for the SSW2018.*

Following Reviewers 1 and 2 recommendations, we have looked for a systematic way of displaying the ensemble spread of u60_10 for all models. Doing this for each day and model in Figure 1 may not be straightforward, given that we include eight models plus the two ERA5 lines. However, we agree on the importance of illustrating the model spread for at least a relevant period, such as the dates surrounding each SSW in ERA5 (12th-16th February 2018, 2nd-6th January 2019 and 18th-22nd September, Table 2 of the manuscript). To address this, we have added three new panels to Figure 1 (Fig.1d-f) that show the ensemble distribution of u60_10 for these periods. A brief discussion of the ensemble spread in each case has also been included in the revised manuscript (L251-260 of the revised manuscript). For instance, these panels reveal that easterly wind values fall within the 1.5 interquartile range of the ensemble distribution for the CNRM-CM6-1 model during the SSW2018, whereas this is not the case for the other models. This helps then reconcile the wind values and the % of ensemble members simulating an SSW.

2) Figure 5 to 7: I think all the subplots should use the same colorbar within each figure. For ERA5, am I right in thinking that this is a difference from climatology whilst for the models it is a difference between the strongest and weakest ensemble members? As such I would expect that the model composite differences shown are larger than if you were able to do a comparison to the free running climatology of each model (like for ERA5). I would like to see a more careful discussion of what is being shown in these figures around line 320. Whilst there are similarities between the patterns in some models and ERA5, the strength is much weaker in all cases.

Figures 5 - 7 display the difference of the anomalous Z500 for the “weakest u” members (the 15 members with the weakest polar night jet (PNJ) at the time the SSW was observed in ERA5) minus the anomalous Z500 for the “strongest u” members (the 15 members with the strongest PNJ at the time the SSW was observed in ERA5). These composite maps of Z500 were computed for the period corresponding to the strongest burst of upward-propagating wave activity at 300 hPa in each case (3-13 Feb. 2018; 18-30 Dec. 2018; 1-10 Sep 2019). This approach allows us to assess whether the tropospheric state was different between the two groups of members and, consequently, whether it may have influenced the occurrence of SSWs. Thus, the plots for each model do not show the full Z500 pattern, as the ERA5 maps do, but rather the differences in tropospheric anomalies between the two groups of ensemble members. This explains, at least in part, why the amplitude of the model composite differences is weaker than that in the ERA5 plots and so, the colorbars should not be the same for ERA5 and the models.

Additionally, the comparison of the tropospheric circulation between the two groups of models helps identify the tropospheric structures that may modulate the likelihood of SSW occurrence in the models. However, in some cases, the stratospheric state between the two groups of ensemble members may not differ substantially as it happens, for instance, in GLOBO for SSW2018 or CNRM-CM6-1 for SSW2019 (Fig. 1e). As a result, even if the tropospheric pattern plays a key role in triggering the SSW, the corresponding anomaly maps may display lower magnitudes than those in ERA5.

In the revised version of the manuscript, the description of what Figures 5 to 7 show has been extended and clarified in L315-322 of the revised manuscript, explicitly highlighting the differences in the computation of ERA5 and model plots. The ensemble means of anomalous Z500 of the FREE simulation of each model have also been added to Fig. 5-7 in contours to help the interpretation of the mentioned difference maps. In those plots one can see that the magnitude of the anomalies is comparable to the ERA5 ones.

3) Figure 9: would it be helpful to add the multimodal mean?

The multimodel mean of HF at 100 and 300hPa for the three different experiments was already included in Figure 9 in gray. However, it was not clearly stated. In the revised version, we have included direct references to this multimodel mean in the figure caption and the associated description in the manuscript (L449-450).

4) Refractive index. The authors acknowledge around line 715 that the resonant growth is likely non-linear. It is my view that since the refractive index is derived from linear theory, it has serious limitations in how it can be applied and I would prefer if Fig 12 and 13 and associated discussion were omitted.

We thank the Reviewer for the suggestion. We have decided to omit the refractive index in Figures 12 and 13 and the associated description. Indeed, Section 5 has been simplified to bring the main points more clearly as also recommended by Reviewer 3.

5) The authors mention gravity wave drag as a source of uncertainty in models and as being important for triggering wave resonance. Have you looked at this in the SNAPSI models?

We agree with the Reviewer that examining the gravity wave drag would be highly valuable for the analysis. Unfortunately, only a limited number of the analyzed models provide gravity waves-related output variables. Specifically, three models include the eastward and northward acceleration by parameterized orographic gravity waves and only one model provides corresponding data for non-orographic gravity waves. Furthermore, two of these three models (GRIMs and CESM2-CAM6) are low-top models and are also those that present the largest challenges for representing wave propagation in the stratosphere during SSW2018. This is the reason why we propose this analysis for the future in the framework of a similar initiative as SNAPSI.

6) Line 745: I do not think interactive chemistry plays a role in the onset of SSWs but does later in the year. All the time periods analysed are in polar night.

We acknowledge that stratospheric ozone plays the most relevant role in shaping the polar stratospheric state after the polar night. However, several studies, including Haase and Matthes (2019) and Oehrlein et al. (2020), have shown that interactive chemistry also modulates the mean polar vortex state during the second half of winter (January-April), leading to a stronger vortex. Consistent with this strengthening, Haase and Matthes (2019) reported that interactive chemistry influences vortex variability, with a tendency toward fewer SSWs when chemistry-climate interactions are included. More importantly, both Haase and Matthes (2019) and Oehrlein et al. (2020) found that the time evolution of the polar temperature around the DJF SSW central dates differs depending on whether the interactive chemistry is considered or not. In particular, the inclusion of interactive chemistry tends to amplify the stratospheric warming associated with the onset of the SSW, likely due to increased dynamical heating linked to the greater wave activity necessary for an SSW to occur with a stronger mean vortex state. This is extremely important for the SSW forecast as it can determine whether a minor or a major SSW occurs.

These previous results suggest that interactive chemistry might play an active role in the onset of SSWs and therefore, represent an important aspect to investigate in future work.

<u>Minor comments:</u>

A general comment here is that the typesetting of the maths could be improved. For example, subscripts are used for both the z and \phi components of F and for partial derivatives.

The typesetting of the maths has been improved. In the new version the subscripts are restricted for the z and ϕ components of F.

Equation (1): Bold F for vector here and across manuscript. Added

Equations (1.1) and (1.2): \overline{\theta}_z rather than \overline{\theta_z} Modified

Line 190: Define z. Also de-italicize 'and \theta'. Done

Equation (2) actually comes from Kushner and Polvani (2004) Eq (7). The way they present it is much easier to read. Modified

Equation (4): Definition of q. Which PV? Quasi-geostrophic? Yes, we referred to the quasi-geostrophic vorticity. However, the equation has been removed in the new version of the manuscript following the recommendations of this Reviewer.

Table 3

- First column. Consider using the event names introduced in Table 2. Modified

- Second row: 2018 should be 2019.

  It referred to the first initialization of the FREE experiment for studying the SSW2019 that was on 2018-12-13. In any case, this has been replaced by the event names as suggested.

- Third row. Maybe round 47.5% to 48% for consistency. Rounded

Supplementary figures S2 to S4 are far too small.

Following Reviewer 2's suggestion we have included the ensemble means of anomalous Z500 of the FREE simulations in Figures 5-7 of each model. Thus, the new figures S2 to S4 only show the model means of this variable of the NUDGED and CONTROL simulations. Consequently, the figures look bigger than previously.

References:

Haase, S. and Matthes, K.: The importance of interactive chemistry for stratosphere–troposphere coupling, Atmos. Chem. Phys., 19, 3417–3432, https://doi.org/10.5194/acp-19-3417-2019, 2019.

Oehrlein, J., Chiodo, G., and Polvani, L. M.: The effect of interactive ozone chemistry on weak and strong stratospheric polar vortex events, Atmos. Chem. Phys., 20, 10531–10544, https://doi.org/10.5194/acp-20-10531-2020, 2020.