

Geosci. Model Dev. Discuss., referee comment RC2
<https://doi.org/10.5194/gmd-2021-394-RC2>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-394

Anonymous Referee #2

Referee comment on "Stratospheric Nudging And Predictable Surface Impacts (SNAPSI): a protocol for investigating the role of stratospheric polar vortex disturbances in subseasonal to seasonal forecasts" by Peter Hitchcock et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-394-RC2>, 2022

Review of "Stratospheric Nudging And Predictable Surface Impacts (SNAPSI): A Protocol for Investigating the Role of the Stratospheric Polar Vortex in Subseasonal to Seasonal Forecasts" by Hitchcock et al.

General comments

This manuscript describes an experimental protocol for multi-model assessment of the contribution of SSW events to surface predictability on sub-seasonal timescales. By adopting the nudging approach, this experimental protocol aims to reveal the influence from the "perfect" stratosphere explicitly. This experimental plan is coordinated by the SNAP working group of WCRP SPARC, and is a plan that the SNAP should have submitted and undertaken earlier. After the Phase-I multi-model experiment of SNAP (Tripathi et al. 2016), this community spared time for the "coordinated" (or dull self-nominated) analyses of S2S prediction data. However, as explained in sections 1 and 2, these Phase-II data analyses were almost impossible to disentangle the stratospheric influence on the tropospheric forecast skill with confidence in the causal relationship (I knew that before they did). Therefore, this kind of experiment is necessary to advance our understanding of the stratospheric influence on the tropospheric circulation and to build a common view of expectable skill contribution from the stratosphere in current prediction systems. I support the importance of this proposal.

However, as a reviewer, I feel a little concerned about achieving the purpose of multi-model inter-comparisons in the current proposed settings. In particular, the author's preference of the nudging only zonally symmetric component and the inclusion of the fourth purpose (about wave evolving process in the stratosphere) may prevent a sound comparison of tropospheric response to the prescribed stratospheric state among prediction systems. I could not convince the propriety of the settings, at least from this manuscript. Moreover, the treatment of tropical coupling seems to be inappropriate. Therefore, I recommend the authors reorganize the priority of scientific purposes and show the validity of experimental settings.

Major Comments

(1) Is it possible to present some evidence for the validity of experimental settings?

I believe that the experimental protocol's main purpose is to share the fixed details of the setting after enough validations (which prevents others from laborious processes checking dependency on settings). In addition, the presentation of a typical (prototype) result would facilitate further participation by others (e.g., Held and Suarez (1994) presented results of two dynamical cores, and it helps the readers and following investigators to deduce the robustness of results). Since this series of experiments depends largely on the nudging parameters, how the authors have fixed the parameters should be explained with enough reasoning. For example, p_b and p_t (and function) are different from those of Hitchcock and Simpson (2014). How did you tune these settings? I guess that the authors have conducted test experiments by using some operational system (the IFS?). How significantly affected the choice of a lower limit of the nudging on the tropospheric ensemble spread and mean difference? Proactive presentations of such information would prevent unnecessary future discussions in the step of inter-comparisons.

(2) Isn't it too greedy to include the fourth purpose?

The zonally-symmetric nudging allows planetary waves to evolve freely (to some extent) even in the stratosphere. This enables the current protocol to address the fourth purpose. However, at the same time, it allows uncertainty of stratospheric state despite that the most important purpose of this experiment is to assess the contributions from the

imposed "perfect" stratosphere and compare them among multi-model results. I feel that the well-tailored nudging of full stratospheric state (e.g., middle-to-upper stratospheric full-nudging with a wider buffer zone below) may be more appropriate to pursue the multi-model inter-comparisons. It would be better if the fourth purpose is placed as an additional scientific goal.

(3) It may be better to change the nudging setting to discuss the tropical coupling.

Although this manuscript roughly touches the coupling between the tropical stratosphere and the troposphere as the secondary science questions (the fifth purpose), this topic has the potential to be a more important target than extratropical coupling. One of the ultimate purposes of the multi-model inter-comparison is to attribute the model's performance to some particular model settings. As many modelers would agree, one of the most uncertain parts of atmospheric models is the representation of clouds. Therefore, it is natural that stratospheric influence on tropical convections should be placed at the highest priority of multi-model inter-comparisons. In such an investigation, the lower limit of nudging in the tropics should be set higher than that in the extratropics (not to interfere in the high tropical cloud directly). However, I am unsure whether the current setting ($p_b = 90$ hPa) is high enough to avoid the direct influence. It may be better to introduce latitudinal dependence in the nudging coefficient if the tropical ensemble spread shows an undesirable distribution. Otherwise, it would be better to plan the nudging experiment focusing on the tropical coupling separately. Sloppy spotlighting may ruin chances of further development.

(4) Changing the priorities of the experiment will allow more models to participate.

Among the experiments listed in I.75-89, the "free" and "nudged-full" are free from the artificial relaxation procedure with shocks and relatively easy to conduct even by models with grids that are not necessarily harmonic with the zonally-symmetric nudging. I think it is better to set these two experiments as the first step request. Then, other experiments ("nudge," "control," and "control-full") should be requested as the second step. Such a division would increase possible participants, at least for the first step. Since the "free" experiments would approach the model's climatological state if the initialization date is set far enough from the SSW onset date (although there are exceptions, of course), the purpose of deducing stratospheric contribution to the troposphere can be roughly achieved by just comparing the "nudged-full" and "free" experiments. I agree that there are large merits of conducting the zonally-asymmetric nudging and comparing it with the "control" experiment. However, I wonder which should we place the priority in the multi-model inter-comparison.

Minor Comments

Title: the Stratospheric Polar Vortex → e.g., "Recent Weakening Events" of the Stratospheric Polar Vortex

Since this protocol covers just only 3 SSWs which are mainly touched by recent publications of quick S2S data analyses, it is inappropriate to use the term representative of various behavior of stratospheric polar vortices.

Is the "control-full" setting appropriate?

Unlike the "control" experiment, the "control-full" experiment would strongly damp the stratospheric wave components due to the sample-averaged smooth structure of the climatological state. Is this as you intended? I think the true "control-full" experiment should construct its ensemble by changing $T_C(t)$ to $T_{\text{year}}(t)$ (each year's state of ERA5). In this case, at least a 40-member ensemble can be obtained using the reanalysis data from 1979 to 2018.

Figure 1:

Is it possible to arrange this figure as a more straightforward form for this manuscript? I think the histogram of the split SSW is unnecessary (e.g., Figure 11 of Maycock et al. 2020: Removing CTL_ADJ is more desirable...).

Figure 2:

This figure needs to be brushed up. The observation should be changed to the same format as the forecasts. It seems that the temperature anomalies of the forecasts are limited over the land. How many ensemble members are used to plot in each panel?

The caption of Figure 3:

In my understanding, Butler et al. (2020) does not describe the calculation method of NAM indices in detail. They have just cited Gerber and Martineau (2018). I do not really like such an inappropriate citation. It is better to write such as "ERA5 version of Figure 5(a) in Butler et al. (2020)."

Table 6 and Authorship:

I doubt the necessity of Table 6 and authors from operational centers since the numerical integrations are not performed, and any early results are not provided in this manuscript. They have just only expressed the intention to participate. The authorship of these people should generate when the data are submitted and the model settings are described in some data journals (e.g., ESSD?). Therefore, the contribution of these types should be noted in the acknowledgement.

Typos, etc.

l.320: 60 N --> 60° N

Make consistency in the use of abbreviation terms (NAO, SAM, MJO, QBO). For example, l.396 and l. 403 uses "Southern Annular Mode" although the SAM is already defined in l.307. Also, "NAO" is used in l.129- before the "North Atlantic Oscillation" in l.308.

References

Gerber, E. P. and Martineau, P. (2018) Quantifying the variability of the annular modes: reanalysis uncertainty vs. sampling uncertainty, *Atmos. Chem. Phys.*, **18**, 17099–17117, <https://doi.org/10.5194/acp-18-17099-2018>

Held, I. M., and Suarez, M. J. (1994) A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models, *Bull. Amer. Meteorol. Soc.*, **75(10)**, 1825-1830. [https://doi.org/10.1175/1520-0477\(1994\)075<1825:APFTIO>2.0.CO;2](https://doi.org/10.1175/1520-0477(1994)075<1825:APFTIO>2.0.CO;2)

Maycock, A. C., Masukwedza, G. I., Hitchcock, P., and Simpson, I. R. (2020) A regime perspective on the North Atlantic eddy-driven jet response to sudden stratospheric warmings, *J. Climate*, **33(9)**, 3901-3917. <https://doi.org/10.1175/JCLI-D-19-0702.1>

Tripathi, O. P., Baldwin, M., Charlton-Perez, and co-authors (2016) Examining the predictability of the stratospheric sudden warming of January 2013 using multiple NWP systems, *Mon. Wea. Rev.*, **144(5)**, 1935-1960. <https://doi.org/10.1175/MWR-D-15-0010.1>