Response to Reviewer 2

We greatly appreciate the reviewer's effort to provide valuable and constructive comments to improve our manuscript. We faithfully addressed the reviewer's comments below. For clarity, our responses are colored in blue.

General Comments:

Reference to Tropopause: Should the definition of tropopause used by the authors be stated? Presumably, a WMO standard lapse-rate definition was used, but many other tropopause definitions do exist. I suggest the authors clarify this from the outset.

The tropopause is defined using the WMO lapse-rate criterion. This is included in the caption of Fig. 1.

References to Hong and Son (2025): Although I appreciate the frequent reference to this study, given it builds upon existing authors' work, I would have expected to see more reference to other literature such as Baldwin et al. (2024), which explore the likely mechanisms for surface amplification of the stratosphere-troposphere coupled response.

For instance, the authors make no reference to the likely role of compression on the polar tropospheric air column, due to polar vortex weakening, aided via lowering of the tropopause height. In relating their results to theory, the combined roles of PV anomaly and stretching should be discussed (Fig. 8b in Baldwin et al., 2024).

Thank you for the comment. The present study reports that tropospheric geopotential height anomalies (negative NAM) are mainly driven by poleward mass transport in the upper troposphere and lower stratosphere. In the SNAPSI experiments, this poleward mass transport is mediated by zonal wind nudging in the initial period and by eddy momentum fluxes in the maintenance period. These processes differ from those proposed by Baldwin et al. (2024), where surface amplification of SSW is associated with compression and spin-down of the tropospheric air column in response to SSW-induced tropopause lowering. In particular, the combined roles of the PV anomaly and vortex stretching in their framework (their Fig. 8b) cannot be directly assessed by the QG mass flux diagnostics employed in the present study. Moreover, a direct comparison is further limited, as the present study focuses on only two SSW cases in nudging experiments instead of multiple SSW cases in long-term reanalysis. In the revised

`

manuscript, Baldwin et al. (2024) is more carefully discussed, as below (L455–466).

The present study complements the recent theoretical framework that explains surface amplification of SSW. Baldwin et al. (2024) proposed that SSW lowers the Arctic tropopause. This compresses the air column below and spins down the tropospheric column via potential vorticity conservation, and is therefore associated with higher surface pressure (see also Ambaum and Hoskins, 2002). They also proposed that higher surface pressure over the Arctic may result from mass moving into the polar cap within the troposphere rather than from the upper stratosphere (see Fig. 7 in Baldwin et al., 2024). In particular, the surface amplification is associated with a poleward mass flux near the surface which is stronger than an equatorward mass flux in the upper troposphere and lower stratosphere. This mass circulation is the opposite of that presented in the present study. The contrasting circulation likely arises from the neglect of vertical mass transport and momentum damping in Baldwin et al. (2024). However, the combined roles of the PV anomaly and vortex stretching proposed in Baldwin et al. (2024; their Fig. 8b) cannot be directly assessed by the QG diagnostics employed here. To better relate the QG framework of this study with the potential vorticity framework of Baldwin et al. (2024), further studies are needed using long-term reanalysis. This is beyond the scope of the present study.

Forecast Initialisation Dates (2018 versus 2019 event): I couldn't help but notice and wonder, if the results shown would in anyway be sensitive to the initialisation dates for each event falling on opposite sides of the SSW onset (4 days before 2018 event and 6 days after 2019 event).

I appreciate the authors make use of the forecast experiments as a community dataset, thus did not perform the experiments themselves and had no control over the initialisation dates. But I would have expected at the very least some discussion over the likely importance (or not) of this difference when comparing between the two events.

Could this also explain (if only in part), why the 2019 NUDGED – CONTROL signal is somewhat weaker than for 2018? Because the initialisation date for 2019 is after the event, the control U10 at 60°N wind ramps up from ~0 to 35 m s-1, so a signal from the SSW state at the beginning of the forecast will also be present in the CONTROL simulations, at least potentially impacting the troposphere during the period of the forecast examined by the authors?

Thank you for the insightful comment. The lack of sensitivity test to initialization dates is one of the

`

limitations of the present study. The nudging experiment for SSW2018 fully captures the rapid breakdown of the polar vortex, whereas the experiment for SSW2019 only encompasses the recovery phase of the polar vortex. This may lead to different surface impacts. Moreover, because the nudging is gradually ramped up over the first five days, the CONTROL experiments for SSW2019, initialized after the SSW onset, retain initial SSW signal. This may partly explain smaller differences between the NUDGED and CONTROL experiments for SSW2019 compared to those for SSW2018. However, without model experiments, it is difficult to identify the impact of initialization dates or SSW phase. In the revised manuscript, we have added a description of the nudging ramp-up in the Methods (L87–88) and a short discussion on differing initialization timings between SSW2018 and SSW2019 experiments and their potential impacts, including the retained SSW signal in the CONTROL experiments for SSW2019, as below (L467–476).

Although the mechanisms are similar, subtle differences in tropospheric anomalies remain between the SSW2018 and SSW2019 experiments. For example, the tropospheric response to SSW2019 is weaker than that to SSW2018 when comparing the NUDGED and CONTROL experiments (Figs. 3e,f). This could partly result from weaker stratospheric anomalies during the recovery phase of SSW2019 (Fig. 3f). Note that while the SSW2019 experiments are initialized six days after SSW onset and thus cover only the recovery of the polar vortex, the SSW2018 experiments are initialized four days before SSW onset and cover both the rapid deceleration of the polar vortex and its slow recovery (Fig. 1). In addition, as the nudging is ramped up over the first five days (Section 2), the CONTROL experiments for SSW2019, initialized after the SSW onset, retain the SSW signal during the initial period. This may reduce the differences between the NUDGED and CONTROL experiments compared to those in SSW2018. The possible impact of SSW phase at initialization could be tested by changing initialization dates, especially for SSW2018 experiments, in future study.

Specific Comments:

L34: “This leads to surface extremes…” --> Perhaps not always, certainly on a regional level, so maybe: “This leads to enhanced risk of surface extremes…”

Modified.

`

L74-75: The authors should add/cite here that zonally symmetric nudging has been found in free-running models to reproduce closely the modelled tropospheric response to SSW events (Hitchcock and Simpson, 2014). And importantly, zonally symmetric nudging protocols in the stratosphere have been verified in earlier studies to not induce any significant artificial properties in tropospheric circulation (Hitchcock and Haynes, 2014).

Thank you for pointing out this. In the revised manuscript, we have incorporated Hitchcock and Simpson (2014) and Hitchcock and Haynes (2014) to support the validity of the zonally-symmetric nudging method (L91–93).

Fig. 1 & 2: I wonder if a vertical line as opposed to a small gray triangle on x-axis might help the timing of SSW onset to stand out better, certainly for Fig. 1. The transparency/linewidth can be customised to ensure this does not obscure the filled contour field?

Thank you for the suggestion. The gray triangles are now replaced with vertical orange lines in Figs. 1 and 2.

L219-220: As mentioned under 'General Comments', the mechanisms for surface pressure increase over the Arctic, in response to downward coupling of SSW-induced anomalies, was the subject of much attention in Baldwin et al. (2024). This study and ideally other relevant literature should be cited and discussed.

As in response to major comment 2, we have more carefully compared the QG framework of this study with the PV framework of Baldwin et al. (2024) (L455–466).

L363-371: I found myself a little confused when reading this, the pressure ranges the features of note correspond to. I suggest the authors could improve by specifying pressure ranges in this paragraph where possible or alternatively 'upper/mid/lower stratosphere or troposphere' to help the reader out.

We have added specific pressure levels.

Technical Corrections:

L94: "SSW" --> "SSWs"

L139: "…one-moth-averaged…" --> "…one-month-averaged…"

`

L318: “(Figs. S4)” --> Singular as its just one Figure?: “(Fig. S4)”

Corrected.

`