Review of 'The tropospheric response to zonally asymmetric momentum torques: implications for the downward response to wave reflection and SSW events' by Ning et al.

General comments:

I thank the authors for their detailed responses to the comments made by me and the other reviewer. The revised version shows significant improvement and addresses most of my previous concerns. However, a few statements remain imprecise or overly strong, and adjusting them would help ensure that the conclusions accurately reflect what can be supported by the analyses. Please see the detailed comments below.

Major Comments:

1. Some of the statements regarding the causality are too strong. The manuscript frequently uses expressions such as 'causally'. I understand that the main aim is to isolate the influence from the stratospheric forcing, and the authors wish to emphasize this aspect. However, it might be misleading in several contexts. For instance, the last sentence in the abstract, while it is true that the tropospheric response in these experiments originates from the stratospheric forcing, the sentence 'the observed surface response … causally forced by, stratospheric perturbations' (L18-20) is not precise. It is also unclear whether 'observed' refers to the detected model response or observational data. If the latter, the statement is too strong. I suggest rephrasing this as 'can be forced by …' or something similar. In addition, the divergent mass streamfunction can help explain the surface temperature response, but only to a certain extent. The phrase 'causally linked' (L17) might overstate the role of this diagnostic. I recommend softening the tone so that the conclusions more precisely reflect what can be inferred from the present analysis.

2. The statement regarding the consistence with observation/reanalysis is not accurate. The manuscript currently states that the two phases correspond to clusters 4 and 5 of Kretschmer et al. (2018a). However, while the surface temperature response indeed show similarities, the stratospheric circulation does not align in the same way. In particular, wave reflection only appears in cluster 4, where cluster 5 shows a reduced upward wave propagation, but the raw Fz is still positive over Eurasia, unlike the negative raw Fz in the phase-270 experiments here. I understand that we cannot expect the idealized experiments to reproduce every observational feature. However, the comparison should be presented more carefully to avoid implying equivalence where the mechanisms differ. I suggest refining the relevant statements accordingly. In addition, repeatedly referring to cluster numbers may confuse readers unfamiliar with the cited work; it may be clearer to describe their defining characteristics when first introduced and avoid relying solely on "cluster 4/5" labels thereafter.

3. The alignment of timing between branch ensembles and control runs. While the authors noted that 'there is no expectation for the timing of the surface responses to match', the magnitude of response in the branch ensembles after day 13 appears more comparable to those in the control runs after day1. I understand that 'day0' represents different reference points, and strict alignment is not required. But align the timelines based on the peak zonal-wind reversal (e.g., day 0 in the control run and day 12 in the branch runs) may make the comparison more straightforward for readers. Alternatively, omitting direct cross-experiment comparisons at fixed lags, or explicitly noting their limitations, would avoid potential confusion.

Specific Comments:

1. L17. Should 'downward propagation events' refer instead to 'wave reflection events'?

2. L61-63 and L408-410. Previous studies have shown that this type of stratospheric anomaly is linked to preceding tropospheric circulation (e.g., Shen et al. 2023; Tan and Bao 2020) and that similar stratospheric disturbances can lead to distinct surface response depending on the tropospheric processes involved (e.g., Shen et al. 2025). Adding a brief discussion where relevant can be helpful to strengthen the motivation for isolating the role of stratosphere.

3. L152. Change 'present' to 'represent'.

4. L203. Should be 'McIntyre' and 'Edmon et al.'.

5. L231. It is more accurate to state 'averaged over days 6 to 12'. The same applies to other similar descriptions.

6. Figure 3c and d. The tropospheric polar-cap height anomaly peaks almost simultaneously with the stratospheric anomaly. Could the authors clarify why this occurs?

7. L273-275. For the reasons described in major comment #2, I suggest reducing the emphasis on direct comparison with Kretschmer et al. (2018a), particularly for phase-270, which does not closely resemble cluster 5 beyond the surface temperature pattern.

8. L283-284. Here the comparison uses days 1-5, but earlier the authors note that the timings are not expected to match. As mentioned in major comment #3, aligning the timing or avoiding such direct comparisons may reduce confusion.

9. L310-312. Please specify the longitude range of the region discussed for easier interpretation.

10. L321-326. Phase-90 shares characteristics with cluster 4, but phase-270 does not resemble cluster5. Revising this statement for accuracy would be beneficial.

11. L445. Cluster 5 in Kretschmer 2018a does not show a wave reflection. This should be corrected.

Reference:

Shen, X., Wang, L., Scaife, A. A., & Hardiman, S. C. (2025). Intraseasonal linkages of winter surface air temperature between Eurasia and North America. Geophysical Research Letters, 52, e2024GL113301. https://doi.org/10.1029/2024GL113301

Tan, X., and M. Bao, 2020: Linkage between a dominant mode in the lower stratosphere and the Western Hemisphere circulation pattern. Geophys. Res. Lett., 47, e2020GL090105, https://doi.org/10.1029/2020GL090105

Shen, X., Wang, L., Scaife, A. A., Hardiman, S. C., & Xu, P. (2023). The Stratosphere-Troposphere Oscillation as the dominant intraseasonal coupling mode between the stratosphere and troposphere. Journal of Climate, 36(7), 2259–2276. https://doi.org/10.1175/jcli-d-22-0238.1