

Cover Letter

Dear Prof. Peter Haynes,

Thank you for your letter and for the reviewers' comments concerning our manuscript entitled "Sorting sudden stratospheric warmings with the downward tropospheric influence using ERA5 and CESM2-WACCM". Those comments are all valuable and very helpful for revising and improving our paper. We have studied comments carefully and have made correction as our reviewers suggest. Revised places are highlighted. The main corrections for the paper and the responses to the reviewer's comments are also uploaded. Please refer to our response files for more details.

Yours,

Jian Rao

raojian@nuist.edu.cn

Response to Reviewer # 1

I appreciate the authors' efforts to respond to my comments and improve their manuscript, making it more accurate and clearer.

I have two comments.

1. I don't think the authors actually answer some of my comments. For example, the comments below:

L212-L217: How do these different behaviors before SSWs determine, influence, or play a role in different types of downward impacts after SSWs?

Figure 7 and the relevant description

My question is about the physical mechanism of the coupling. However, the authors' response is primarily descriptive and repeated as in their own text without any new insight provided. For example, "Consistent with" does not explain the HOW or the mechanism by which the pre-SSW conditions (the "different behaviors") determine, influence, or play a role in the downward impact. "low-level circulation is continuous" or "has not changed much." This is not a satisfying physical explanation.

It is fine if the authors didn't have an answer to these questions. However, I think the authors should be cautious not to overstate their conclusions by implying a causal link where no physical mechanism explanation.

Response: Thank you for your suggestion. We have revised the section to acknowledge our limited understanding of the physical mechanisms and focus on the description and diagnosis of type-dependent dynamical features.

- "These results are intended to document the type-dependent dynamical characteristics of the circulation, rather than to establish a mature causal physical mechanism for the observed differences that are far beyond the scope of the paper." (L448-L450)

For details, please refer to Section 4.

2. Figure 11 and its discussion describe the dynamical features in a chain. I am a bit confused whether the authors implied a causal chain with the black arrows that connect these features. I don't think the causal link can be concluded based on the analyses in the paper. For example, how does the "joint action of wave-1 and wave-2" lead to "three centers in stratospheric height anomaly"? Figures 8-10 show cross-sections, which can only show the different features in the three types of DWs, but there is no analysis linking the wave patterns (spatial map) with the three centers in the height anomaly as shown in Figure 7. Again, I am not asking authors to perform more analyses, but

I suggest the authors should be cautious not to overstate their conclusions by implying a causal link where no physical mechanism explanation.

I would recommend acceptance or a slight modification of Figure 11.

Response: We agree that a clear causal linkage is not easy to establish. We have changed Figure 11 by replacing the black arrows that are usually used to imply causal relationships with dashed lines that emphasize the consistent relation. We have made numerous corresponding revisions to the paper to avoid any causal implications regarding the mechanisms.

Response to Reviewer # 2

I appreciate that the revised manuscript now states more explicitly that this study focuses on classifying SSW types. This seems to address my earlier concern well. I also think the consistency between ERA5 and CESM2-WACCM, including the parts related to SSWs and downward propagation, is a strength of this study. This consistency suggests that those aspects of type-dependent differences are not merely dependent on a few specific cases, but that there are indeed robust distinctions between types. From this perspective, I do see several statements throughout the manuscript noting that results are consistent in CESM2-WACCM. It would be helpful if the manuscript could more clearly remind the reader why this consistency between ERA5 and CESM2 is meaningful.

Response: Thank you for your suggestion. We have incorporated the significance of this consistency between ERA5 and CESM2 into Sections 2.2, 3.1, and the Conclusions respectively. We listed those places of revisions as follows:

- “By examining the same classification framework for both ERA5 and CESM2-WACCM, this study aims to assess whether the identified SSW type distinctions are sensitive to different data sources, rather than being dependent on a specific dataset product.” (L125-L127)
- “The similar evolution patterns in ERA5 and CESM2-WACCM indicate that the characteristics differences are not confined to the fewer events in the observational data, lending confidence to the robustness of the conclusions.” (L268-L270)
- “The agreement between ERA5 and CESM2-WACCM does not imply that the underlying physical mechanisms are fully captured, but it does indicate that the identified type-dependent characteristics are robust across distinct datasets.” (L614-L616)

My main remaining concern is that several places still read as if coincidences or composite-mean features are being mechanistically confirmed. Of course, given that the surface temperature responses differ across types, it is appropriate to describe how the temperature anomalies for each type arise through differences in the NAM evolution or circulation structures. This aligns with examining mechanisms of downward influence. However, as your first sentence of the Abstract suggests, a “downward propagation mechanism” of SSWs generally refers to the full chain of processes—from precursors, to SSW onset, to downward propagation, and to the subsequent surface influences. In the current manuscript, the mechanism remains somewhat unclear from the pre-SSW stage through the tropospheric circulation differences. At present, it seems clearer how the unique circulation patterns are linked to different surface temperature responses, while the mechanism leading to those pre-SSW differences is less well established.

Response: We agree with the reviewer that, while the manuscript clearly documents circulation characteristics and surface response features, the physical mechanisms are not fully established. We have therefore carefully revised the manuscript to avoid mechanistic interpretations of these pre-SSW signals and to focus the paper on the description and diagnosis of type-dependent dynamical characteristics rather than causal explanation. Revisions are as follows:

- “These results are intended to document the type-dependent dynamical characteristics of the circulation, rather than to establish a mature causal physical mechanism for the observed

differences that are far beyond the scope of the paper.” (L448-450)

- We have changed Figure 11 by replacing the black arrows that are usually used to imply causal relationships with dashed lines that emphasize the consistent relation.

Results expected from the type definition

Also, I still notice a tendency in the main text to overstate results that are essentially expected given how the types are defined. My point is not that composites should not be shown or discussed, but rather that outcomes that follow naturally from the definition should not be framed as discoveries. For example, if a CP type were defined as cases where Niño4 is stronger than Niño3, one would not frame it as a “finding” that Niño4 SST is higher for the CP type.

For example, in explanation of Figure 3, the text states: “To compare the possible different impact of DWs on the near surface, the composite t2m anomalies in the 40-day intervals before and after the SSW onset are shown in Fig. 3. In both ERA5 and CESM2-WACCM simulations (Fig. 3a-d), the cold anomalies appear over different regions during the SSW for DWs and the NDW.”

As I have emphasized, these impact differences are inevitable because the types are defined based on impacts. This phrasing could lead readers to interpret the result as unexpectedly large or newly revealed. A more straightforward framing such as “the temperature impacts are naturally well separated by the type classification” would be more appropriate. In particular, the phrase “possible different impact” is difficult to justify here, because different impacts are expected by definition.

Response: Thank you for your suggestions. We revised the description of Fig. 3 to explicitly state that the near-surface temperature differences are expected by construction of the DW types, and we avoided framing these features as new findings. Revisions are listed as follows:

- “Due to the definitions of the SSW types, the near-surface temperature impacts exhibit marked differences naturally. Figure 3 shows the composite t2m anomalies and their differences in the 40-day intervals before and after the SSW onset.” (L311-313)
- “In the post-SSW period, the t2m anomaly patterns differ among the DW types and the NDWs, consistent with the t2m-based definition of SSW events. DW events exhibit cold anomalies over their respective defining regions, while NDW events lack coherent surface temperature signals. These characteristics are similarly represented in ERA5 and CESM2-WACCM.” (L327-330)

Emphasis on detailed differences between the types

1. The statement around Line 25 “NDWs show relatively weaker impact on the troposphere, which is primarily related to the weaker stratospheric disturbance amplitude” reads as if the amplitude difference is being treated as a deterministic driver of type differences. In other words, does this imply that when an SSW is strong it will always propagate downward, or vice versa? Please consider revising this sentence to avoid that implication.

Response: We have revised this sentence as follows:

- “On average, the relatively weaker tropospheric impact of NDWs is consistent with their weaker stratospheric disturbance amplitude.” (L25-L27)

2. Some of the detailed descriptions in the Figure 1 section read as though the type separation is mechanistically rooted in those pre-SSW characteristics. If the statistical test was conducted against

the full wintertime distribution, I suggest modifying the approach to directly test differences among the SSW types themselves. That is, without considering non SSW cases, please test whether the differences among the SSW groups shown here are statistically significant. For example, are the pre-SSW stratospheric signals for BOTH and NDW truly significantly different? Since the pre-SSW signals also appear less evident in CESM2-WACCM, it would be helpful to test the difference explicitly, or to show additional evidence (e.g., a scatter plot of cases using a NAM averaged between lag# and lag# at level#) to confirm how well these groups are separated. I also suggest stating in the text that these pre-SSW signals are not clearly seen in CESM2-WACCM.

Response: We have revised both the figure and the paragraph. We have plotted the differences between SSW types in the figure and showed the significance level. Revisions are listed as follows:

- “The dashed contours represent the difference between (a, d) BOTH and NDW (BOTH - NDW) and between (b, c) EA and NA (EA - NA). The solid contours mark the difference at the 95% confidence level.” (L276-279)
- “In contrast, type NA exhibits characteristics markedly distinct from other types: the positive NAM dramatically develops from day -50 in the upper stratosphere to day -15 in the lower troposphere, forming a pronounced vertical structure extending from the upper stratosphere to the lower troposphere (Fig. 1c1).” (L235-238)
- “The process of NAM reversal from positive to negative exhibits relatively minor differences for the DWs and NDW...” (L246-248)
- “Further, it evolves from intensely positive to intensely but shortly negative for type NA.” (L251)

3. For Figure 7, the precursor patterns appear not always consistent between observations and the model. Yet authors still explain the details. I think the discussion should focus more on the common features. Given that the analysis pathway is presented as a sequence addressing the overall mechanism (precursors → pre-SSW NAM → NAM propagation → tropospheric circulation/PV → surface impacts), it seems important to demonstrate whether the differences between types are statistically significant, for the precursors as well (Figures 7, 8, 9, and 10). The goal here is not to contrast SSW vs. non-SSW conditions, but to emphasize differences among types given that SSWs have already occurred. In that sense, estimating within-type variability (e.g., standard deviations around the composites) and testing differences among type composites would help describe these results as type specific characteristics. This would help the manuscript more convincingly claim not only a useful classification but also type-dependent dynamical behaviors. My intention is not to reject the overall framework, but to encourage the text to focus on the most meaningful and statistically supported signals.

Response: Thank you for your nice suggestions, which we have considered in our revisions this time. We have revised the discussion of the precursor patterns in Fig. 7 to focus on the common, robust features shared between ERA5 and CESM2-WACCM, rather than on detailed differences that are not consistently reproduced.

Further, we also revised Figs. 7, 8, 9, and 10 to demonstrate the significant differences between the various types, emphasizing this point in the main text. Places of revisions are listed as follows:

- “...as well as the composite difference between (e, k) BOTH and NDW, and between (f, l) EA and NA...” (L514-515; L534-535)
- “These characteristics are markedly different from EA (Fig. 7f). For NDW, compared to DW, the polar vortex at 100 hPa is not significantly disturbed, and no significant low-pressure systems developed at sea level (Fig. 7d, e).” (L462-464)
- “The model results also suggest a weakening of the polar vortex at 100 hPa prior to DW events. At the surface, the spatial structure of the NAM, including the sign and location of the anomaly centers for different event types, broadly agree with those in the observational analysis.” (L470-473)

Refer to the revisions for more details. Thank you!