Response to Reviewer # 1

Review for "Sorting sudden stratospheric warmings with the downward tropospheric influence using ERA5 and CESM2-WACCM" by Lu and Rao.

Summary

This study investigates how different types of downward-propagating sudden stratospheric warming (SSW) events are associated with distinct regional surface cold extremes and classifies them based on their surface impacts, using reanalysis data and CESM-WACCM simulations. According to the abstract, introduction and conclusion, the paper aims not only to classify cases of surface cold extremes associated with SSWs, but also to explain the mechanisms behind how different types of surface cold extremes are caused by difference in SSWs. To this end, the study presents a comprehensive analysis using both observations and CESM-WACCM. One of the main contributions of the paper is the classification of downward-propagating SSW events based on the regional characteristics of the associated surface cold extremes, followed by a global analysis of the related dynamical fields.

Response: Thank you for your positive comments. This study is motivated by the variety of the surface anomalies following the SSWs. As we introduced in the first section of the paper, not all downward propagating SSWs have a cold impact on NA or EA. This classification is different from those in literature, but it really improves our understanding of the SSW.

Firstly, The wave forcings for the SSWs are rich and various: BOTH type is forced by comparable strong wave 1 and wave 2 pulse, EA type is forced by strong wave 2 pulse (EP flux convergence is much larger for wave 2 than wave 1), NA type is forced by strong wave 1 pulse (EP flux convergence is much larger for wave 1 than wave 2), and NDW is associated with the strong total waves (while wave 1 and wave 2 are much weaker for NDW than for DW). This dynamic difference is very clearly shown in Figure 8-10.

Second, the evolution of the circulation during SSWs are diverse: Following the SSW onset, BOTH type is charactered by a complete NAM pattern with the negative height anomaly lobe in midlatitudes extending in a zonal band encircling the earth. The EA type is feature with a NAM pattern but the negative lobe in midlatitude split into two lobes. The NA type is accompanied with the polar anomalous high with the midlatitude negative height anomaly lobe missing. The NDW type has a much weaker NAM like circulation.

Third, as a consequence, the near surface temperature response is diverse. This diversity is not just due to the classification definition, which is consistent with the atmospheric circulation anomalies.

However, the main analyses are mostly descriptive, such as "this type tends to have these features," without sufficient efforts to clarify the causal connections between processes. This creates a mismatch between the intended positioning of the paper and the actual description of the results, which may confuse readers about the central message of the paper. Moreover, even if the main purpose was the classification of SSW types, there are places where the descriptions overinterpret differences between the types that may naturally result from the way they are classified.
Response: Thank you for your criticism comments. During the revision, we considered all of your comments and made corresponding revision.

First, we provide as accurate description as possible and made sufficient effort as possible to clarify the causal connection between the preceding wave forcing, the SSW, the circulation changes, and the near surface response. This chain is finally provided in Figure 11.

Second, all of the mismatch between the results and the explanation has been revised and clarified. Places that might mislead our reviewer and readers have been removed or change with suitable words.

Third, the descriptions overinterpret differences between the types is revised. We also added the difference between CESM2-WACCM and ERA5 in the revision this time.

The reviewer has a viewpoint that the difference is sourced from the way SSWs are classified. We do not deny this. The core of this study is why we classify the SSW according to the surface response, which is what we emphasizes. Thank you!

Nevertheless, I believe the paper has potential, since the analyses are extensive and well-organized. It would be publishable as a descriptive classification study if the authors revise the manuscript to focus on the meaningful dynamical differences revealed by the classification, while excluding the differences that naturally arise from the way the types are defined.
Response: Thank you for your comments. The dynamics differences are consistent with our classification. The near surface difference is the motivation of this study, and we analyzed the difference from several aspects: the wave forcings preceding the SSW, the SSW associated circulation (the NAM pattern structure), the evolution of the downward propagation of the NAM, and the near surface.

All the chains are analyzed and summarized in Figure 11, and all of those results are consistent. We describe our results and give explanation as accurate as possible this time.

Major Comments
Unclear objective: classification or mechanism?

If the goal is to understand the cause of different SSW surface impacts, the paper needs to move beyond descriptive comparisons and provide a clearer explanation of how variables interact across steps. If the goal is to classify the types, the paper should clarify up front that it is intended as a classification study based on observed differences.

Response: The primary objective of this paper is to classify SSWs based on different tropospheric responses and to conduct an investigation into the underlying mechanisms from the dynamic difference.

To well address your concern, we have incorporated this additional text into the manuscript.

- "The primary objective of this paper is to classify SSWs based on different tropospheric responses and to conduct an investigation into the underlying mechanisms from the dynamic differences." (L92-93).

Furthermore, we have expanded upon the discussion of physical mechanisms within the paper.

Insufficient interpretation and unclear physical connections

The authors should avoid wording that may be misinterpreted as implying a causal relationship between features that simply co-occur. It would be helpful to clarify which parts of the discussion are supported by physical reasoning and which are more descriptive.

Response: We have avoided wording that might mislead our reviewer and readers as accurate as possible. Features that cooccur are consistent. The physical reasoning is mainly shown in the dynamics analysis section (Section 4). The description of the difference is mainly shown in the tropospheric response comparison section (Section 3).

1. Some differences reported in the results seem to directly stem from how the cases were defined, but this point is not acknowledged. For example, the EA-type cold extremes are defined to cover a broader longitudinal range. Since the NAM index more reflects zonal-mean changes, it is not surprising that EA cases show a stronger and persistent NAM signal than NA cases. This might be a natural result of the classification.

Response: The EA cold appears in EA and BOTH types, and the NA cold appears in NA and BOTH types. Both continents are compared. Most of the composite is significant. The composite t2m anomalies in Figure 3 are significant, which might imply that our classification is meaningful rather than noisy only.

If the classification lacks this "natural" result, the classification fails. In climate study, those natural results can be broadly seen. We defined the ENSO index, so the composite is a ENSO pattern. We define the QBO index, so the composite is a QBO pattern. We classify the ENSO types, so we get the CP and EP patterns … … All those are predefined, and the dynamics are consistent with this classification. The classification is a useful tool to understand the diversity of our real world.

After careful consideration, we insist on that our classification is meaningful, physically distinguishable, and helpful to understand the SSW diversity.

2. The paper tends to overinterpret co-occurring patterns as physical links. The paper lists many local differences across types but does not sufficiently explain whether these patterns are meaningfully connected.

Response: In our analysis of the geopotential height field, we linked various anomalies through circulation changes. The circulation field provides a more intuitive view of the characteristics of the NAM pattern. The near surface change follows the circulation change on the subseasonal timescale. Furthermore, anomalies in the circulation field, particularly the movement of the polar vortex, can be correlated with changes in surface temperatures before and after various events.

3. the paper lists a number of detailed local differences between SSW types and presents them as if they are physically meaningful, and this seems to depend on the statistical significance of the composite. However, it is not always evident whether these differences are statistically significant. There are figures which show that composite for NDW show widespread dotted areas of statistically significant values (e.g. Fig. 6d), and the statistical testing method is not clearly explained. This raises the concern that some of these apparent differences may not be meaningful physical distinctions, especially in relation to the difference in SSW downward impact.

Response: The significance test used in this paper is the t-test, which test if the resample mean is different from the total mean or if the means of two resamples are different. In some figures, specific regions within the NDW plots do indeed pass the significance test, but these regions are markedly smaller and much scattered than the significant areas observed in the DW types.

To well address your concern, we replotted Figure 6 and give a stricter significant test. The significant area for the NDWs is much limited in ERA5 and CESM2-WACCM after revision. The test method is mentioned in the caption of Figure 6.

4. The anomalies suggested as precursors are not clearly distinguished as being part of the SSW-related signal or independent tropospheric variability.

Response: Prior to the occurrence of the SSW, significant changes have already taken place in the atmospheric circulations, which drive the onset of SSWs. It is too controversial to say this change is part of the SSW or the tropospheric variability. Since they are mixed together and indistinguishable, to the best of our knowledge. Despite this, we still believe that these precursor anomalies are intrinsically linked to the SSW.

Methodology

1. Even if the main objective of the study was the type classification, more careful methodological design would be necessary. For example, EA DW cases already show strong cold anomalies before the SSW occurred, which raises questions about whether the cold extremes are truly caused by SSW. If the anomalies are computed based on a fixed climatology, long-term signals may be included. However, there is also no clear method for determining whether post-SSW cold events are actually caused by the SSW. This makes it difficult to rule out the possibility that some tropospheric anomalies (or anomalies of longer timescale) developed independently of the SSW.

Response: Thank you for your suggestion. We agree with your viewpoint. Regarding the significant cold anomaly response observed at the surface following SSW events, many studies have been conducted within the academic community, although the controversy still exists whether the stratospheric and tropospheric variability can be split. Since the climate is a nonlinear system, we do not think it is easy to distinguish the stratospheric and tropospheric variability, which is far beyond the scope of our study.

In our study, through analysis of the circulation (Figure 7), we show that the polar vortex has already begun to weaken and deviated from the pole before the SSW occurrence. The region towards which the polar vortex migrated coincides with the area where surface cold anomalies emerged before the event. Therefore, we conclude that although the timing of the surface cold anomaly was earlier, the stratospheric influence can not be ignored.

2. How CESM-WACCM contributes to the overall interpretation is not clear. Is it simply to provide more cases, or to test mechanisms? For example, when CESM-WACCM and observational results differ, the physical meaning of the differences is not discussed. Although the study analyzes both ERA5 and CESM-WACCM, the interpretation is mostly based on ERA5.

Response: At the very beginning of our paper, we only show the ERA5 reanalysis, which has really limited sample sizes. As our reviewers suggest in the last round review, we provide the CESM2-WACCM results. We hope to find a balance between the reviewers. We wish our reviewer can understand us.

The primary reason for using the CESM-WACCM model was to validate the characteristics identified in ERA5. The two datasets exhibited consistent behaviours in key findings and conclusions, although minor differences exist. In the revised version, we have expanded our discussion of the differences between the two datasets. After all, no models are perfect.

Lack of structural and editorial refinement

1. Although the analysis is limited to the Northern Hemisphere, this is not explicitly stated in the methodology.

Response: Thank you for your suggestion. We have provided additional explanations in the methodology section. (L122)

2. The structure of the methodology section also lacks consistency. For example, the CESM-WACCM model description is under the subsection "2.1 Reanalysis Data."

Response: Thank you for your suggestion. Adjustments have been made. (L109)

3. While the study defines four geographical domains (Europe, Asia, the United States, and Canada) the classification later becomes just two types (NA and EA), without explanation of the criteria used to group these regions. Furthermore, there is no explanation for why those particular domains were chosen for each region.

Response: One of the main purposes of the article is to distinguish SSW events with different impacts on two continents (North America and Eurasia). However, a single SSW will only have an impact on a portion of the mainland, making it difficult to cover the entire continent most of the circumstances.

For this reason and as the other reviewer suggests, we further divide the mainland. We choose these specific areas as representatives because cold tends to appear in those subregions following the SSWs. We made more explanation this time:

- "Considering that cold surges are more active over the Eurasia (Europe + Asia) and North America (US + Canada) following the DWs, a comparison between the cold anomalies over the two regions can further classify the DWs into three types …" (L159-161)
- "For the mainland, as long as one district meets the criteria, it is considered to be associated with the downward impact of the preexisting DW." (L168-169)

4. A few grammatical errors are present in the manuscript.

Response: Inspections and modifications have been made.