Review of “*Sorting sudden stratospheric warmings with the downward tropospheric influence using ERA5 and CESM2-WACCM*” by Lu and Rao for Atmospheric Chemistry and Physics

Summary:

This study analyzed different types of sudden stratospheric warming (SSW) events based on their downward impacts. Besides only classifying SSW events into those with (DW) and without (NDW) downward impacts, the authors further classified the downward events into three different subgroups based on the region where the cold air outbreaks occurred and looked into the differences in various dynamical quantities for different types of DW events. The study provides an interesting aspect to further study the downward impacts of SSWs and documents different behaviors for four types of SSWs. The paper is well-written in terms of language and structure. However, there are two main concerns listed in the major comments below regarding the analysis and description of the results.

Overall, this study is interesting and meaningful for its detailed examination of DW events and their impacts on various regions. However, the analysis is somewhat shallow, and some of the descriptions and interpretations are not accurate or clear. I suggest a major revision before publication.

Major comments:

1. The authors want to use the CESM-WACCM data to support the findings obtained from the ERA5 data, given the small sample size of SSWs in ERA5. However, the authors described it too generally for the analysis using the CESM-WACCM model data in many places in the paper, by only saying that the model results are consistent with or resemble the ERA5, which, in many places, is not true or accurate (see specific comments below). There are quite some differences in the analyses between ERA5 and model data, which could raise questions and concerns about the robustness of the results found in ERA5 (although the model is not perfect and has its own biases). I don’t think the authors discuss the differences and their implications carefully enough.

2. There is no in-depth analysis beyond purely describing the figures, which can, in the authors’ words, “enhance our understanding of the mechanisms driving surface impacts”. For the discussion in almost all figures, the authors described the differences in the dynamical variables between the three types of DWs in detail, but without explaining how these patterns or unique processes play a role in the development of the three types of DWs. In fact, it is not even clear whether these patterns are critical for the cold air outbreaks in different regions. Besides, some descriptions of the figures are not accurate or consistent with what the figure shows (see specific comments below).

Specific comments:

L110-111 and L112-113: The introduction of the CESM2-WACCM model simulation for the SSWs and downward impacts is too general. It would help to add a bit more detail. For example, does it simulate SSWs (and the downward impacts) well in terms of frequency or magnitude or duration/persistence? Is CESM2-WACCM a high-top model? What is its vertical resolution?

L153: What is the definition of the onset of DWs?

L155-160: The regions for NA and EA are too broad. Within Eurasia and North America, it is not necessary that the entire region experiences the cold air outbreak after SSWs. For example, very often it could be some specific part of the US that experiences extreme cold weather. Taking the average for such a large region could weaken the cold responses. It is also not clear about the classification of the event types. When an SSW event is classified into the EA type, does it mean that both Asia and Europe meet the criteria? Or is only one region (either Asia or Europe) meeting the criteria good enough and counted?

L161: How did the authors decide the -0.3℃ here? Does it correspond to a certain standard deviation or percentile threshold? Besides, -0.3℃ seems relatively small, could averaging over such a large region as mentioned above make the overall anomalies small?

L211 and the entire discussion related to Figure 1: Are these differences statistically significant across these three types of DWs? The same questions apply to all other analyses shown below.

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

L225-230: Are the NAM anomaly contrasts before and after events for each type of SSW event robust? For example, the contrast in the NA events is not shown in the model. I understand that the model is not perfect and has its own biases. However, since the sample size in ERA5 is also small while the model has more samples, it could also be that the features in ERA5 are not robust and due to a specific individual strong event.

Besides, are the differences in these contrasts for different types of SSWs statistically significant? If the contrast is important, how does it play a role in influencing the development of different types of SSW events?

L264-265: I agree that the authors can interpret that the main features in the PDF of the model are similar to those of ERA5. However, there are also some differences that may require a second thought. For example, we can see the shift of the peak in EA and a higher probability with smaller negative NAM values of EA in the model compared to ERA5, and a clearer peak and the overall shape of the PDF of NA in the model differ from ERA5. This would make me wonder about the robustness of the features found in the ERA5. For instance, is the difference in the mean or median between BOTH and EA actually significant or distinguishable, because it seems not to be the case in the CESM2-WACCM model (the PDFs of BOTH and EA have larger overlap in the model than in ERA5)?

L300-301: It seems like a seesaw pattern in t2m between Eurasia and North America, which is interesting. Is this a firm and robust relationship in which, when one region is anomalously cold, the other region is anomalously warm? If so, why and what causes this seesaw pattern?

L314-317: First of all, the differences between ERA5 and the model are not minor, especially from day -40 to day -10. The t2m anomalies are almost zero in this time period in the model. This again goes back to my previous concern about the robustness of the results in ERA5 and brings to the next question: do the t2m anomalies before SSWs precondition the cold anomalies after SSWs? This analysis cannot really show or support the role or importance of the t2m anomalies before SSWs in developing BOTH DWs. The same questions for other types of DWs.

L322-323: Does the evolution of t2m here really relate to or result from SSWs for regions of Europe and Asia, given that the anomalies are not statistically significant and the t2m anomalies are almost unchanged during the period in Europe?

L355-356: It seems like the high IPV center is consistently over Europe throughout the whole period. There is no clear movement of high IPV from the figure, and there is a sudden strengthening of the anomalies, which indicates the IPV does not seem to be conserved but is under external forcing. Thus, it cannot tell the source of the IPV based on the movement of high IPV.

Figure 7 and the relevant description: After reading the whole description of Figure 7, it is good to see different patterns in different types of DWs and NDW, but it is still not clear to me what dynamical processes or what and how these features and patterns before the SSWs influence the development of different types of DWs. The analysis is just descriptive, not really in-depth enough to reveal the differences in the underlying dynamics.

L480-481: From the vectors in figure 8e1 and g1, they seem to be pointing downward over the high latitudes, which indicates the opposite of what the authors describe here. Besides, the convergence is very weak in the stratosphere for BOTH and NA. Are the anomalies here statistically significant?

L486-487: How does the short lifetime of the wave forcing explain the lack of impacts on the surface?

L503-505: The description is not accurate. Only Figure 9g1 shows this. The differences between the model and ERA5 are quite dominant.

L512-514: There is a divergence in the lower stratosphere for the NA DW. Could the author provide any explanation for this? How does this feature lead to or play a role in the NA surface impact? Please explain this.

L514-515: I disagree. The differences between the model and ERA5 are quite dominant to me.

L517: “Both ERA5 and model simulations” Again, I think the differences between the model and ERA5 are large. For example, the model does not show the EP flux convergence anomalies for EA.

L517-518: The arrows of the vectors are hard to see. This is also the question for Figures 8 and 9. In fact, the authors did not mention whether there is any scaling when plotting the EP flux, which is usually the convention. The small magnitude of the vectors indicates the anomalies are small. Are the EP flux anomalies statistically significant?

Discussion for Figures 8-10: I can see the differences in the EP flux and EP flux divergence across the three types of DWs. However, I think there is a lack of in-depth analysis of how these differences influence/determine the types of DWs. For example, after reading the whole section, I still do not know how the anomalous upward propagation of wave-2 wave activity after SSWs leads to the coldness over EA. Are the displacement or split of the vortex more important, or the wave activity itself more important?

Figure 11 and its summary: It is clear to summarize the main differences using this schematic. However, there is a lack of connection between all these dynamical processes and wave activities flux, and coupling, which lead to or contribute to the final development of different types of DWs. There is no clear interpretation to put all these components into a thread, which makes the study purely descriptive and lacks in-depth insights.

L599-600: I don’t see any improved understanding of the mechanisms. Please see my comments above.

Minor comments:

L155 and Table 2: It will help to summarize the percentage of each type of event for both ERA5 and the model.

L164: It would help to briefly describe the method by Maycock and Hitchcock (2015) so that it would be clearer when you mention the parameters you modified below.

L238: What is Fig. 1e1 and f1?

L240: I agree there is nice consistency between the model and ERA5 in most cases here. However, the magnitude of the negative NAM is apparently weaker from the lower stratosphere to the mid-troposphere in the model than in ERA5. Something like this can be mentioned in the section where you introduce the overall model performance in downward impacts.

L249: I don’t think the authors introduced how to compute or obtain the NAO index before.

L254-255: How can DWs modify the PDF of the NAO?

L291-292: Keep in mind that the t2m anomaly pattern contrastingly different for different types of events is not something new because this is how you define the three types of DWs and NDW.

L329-330: Figure 4b2 does not fully support this statement for Europe.

L331-332: The Europe region in Figure 4c2 does not support the description.

L333-334: This sentence is confusing. I don't understand which features or lines in the figure the authors talked about.

L442: Isn't the 100 hPa height anomaly over the North Atlantic also negative?

L525-529: Isn't the finding opposite to the previous studies?