Atmospheric Chemistry and Physics Discussions

Atmos. Chem. Phys. Discuss., referee comment RC2
https://doi.org/10.5194/acp-2022-497-RC2, 2022

Review of acp-2022-497

Anonymous Referee #2

Referee comment on "Monitoring sudden stratospheric warmings under climate change since 1980 based on reanalysis data verified by radio occultation" by Ying Li et al., Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2022-497-RC2, 2022

Li et al. use stratospheric temperatures provided from GNSS-RO measurements and ERA5 reanalysis data to propose a new definition for sudden stratospheric warmings (SSWs). They additionally define a series of diagnostics which can be used to further characterize the SSWs by their duration, amplitude, and areal extent. Their results show examples of how events defined using ERA5 compare to those defined using GNSS-RO, the time evolution of their diagnostics for select SSWs, and statistics of events aggregated over the 42 winters from 1980-2021.

The proposed SSW definition and characterization diagnostics are new, but there are a number of issues with the way the authors have described their methods and characterized their results. Their manuscript requires substantial major revisions before being acceptable to publish. Below I provide my questions, comments, and concerns on the manuscript:

General comments:

(GC1) In my opinion, it's not really clear what the manuscript has to do with climate change beyond showing some results in a single figure that show decadal trends. Reading the title of the paper alone would lead me to believe there should be more analysis related to climate change, when in fact the paper is primarily focused on demonstrating a new SSW definition. Personally, I think the paper should use a different title that is more appropriate for the content it provides. For example "Characteristics of sudden stratospheric warmings defined using reanalysis and radio occultation temperature data", or similar.

(GC2) This paper and prior papers have listed valid criticisms of different SSW definitions. However, the new definition proposed here sticks out as being extraordinarily complicated.

For one thing, it's not easily apparent what the authors even propose is their singular definition of an SSW. I would expect this kind of information to be mentioned clearly in the abstract since it's fundamental to the subject of the paper and for others to use the definition. Instead, one has to work backward from all the information in the long and complicated tables 1 and 2. Specifically, in table 2 item 4 we see the SSW detection criterion is based on SSW-MPD >= 7 days. What is MPD? It's main-phase duration (table 2, item 1), which is based on the number of days with SSW-MP-TEA available. What is SSW-MP-TEA? It's the main-phase threshold exceedance area (table 1, item 9), which is the maximum of the PP-TEA and SP-TEA. SP-TEA (table 1 item 8) then depends on LSTA-TEA (table 1, item 3), and PP-TEA depends on MSTA-TEA (table 1, item 2), which each depend on different temperature anomaly thresholds depending on the level of the stratosphere. There's so many steps and acronyms here that makes it difficult to even know what the SSW definition is when you've reached the end. The CP07 SSW definition involves somewhat subtle details (which prevent double-counting and exclude final warmings), but I can at least walk away with the simple "reversal of 10 hPa, 60N zonal mean zonal winds".

It's also not clear whether the proposed definition lends itself to being applied to other datasets. What vertical resolution of data is required for this definition? Climate model and subseasonal-to-seasonal forecast datasets are commonly output on a very limited set of pressure levels -- can this method still be used with, e.g., only 50 and 10 hPa levels? If the definition can only be readily applied to reanalysis data for monitoring purposes, then that really limits its utility.

(GC3) The authors' justifications for the various thresholds they use in their SSW diagnostics are lacking. The authors note: "we here made sure for longterm application that the four SSW TEA key variables are captured and exploited in a way so that they reliably detect and quantify actual SSW warming" and "Extensive robustness and sensitivity testing provided us with due evidence and confidence that these characteristics ... should enable a new level of quality and quantitative insight into SSWs". These statements do not really tell the reader why you specify these many thresholds in Tables 1 and 2:

At least +30 K anomalies for MSTA-TEA
At least +20 K anomalies for LSTA-TEA
At least +40 K anomalies for USTA-TEA
At least 3x10^6 km^2 >= 3 days for PP-TEA
At least 3x10^6 km^2 >= 5 days for SP-TEA
At least 3x10^6 km^2 >= 21 days for TP-TEA
At least 7 days for SSW-MPD for the SSW detection criterion
Below 90x10^6 km^2 days for minor classification
Between 90 and 180 million km^2 days for major classification
Above 180x10^6 km^2 days for extreme classification
Less than/greater than 21 days of TPD to define a non-TC/TC event

This is a large number of "tuning knobs", and it's not evident at all how the various numbers were chosen or how the results would be affected by adjusting them to different values.

(GC4) While I can generally understand what the authors intend to say throughout the paper, the text will overall require a significant amount of editing for grammar. In my specific comments below, I have primarily focused on asking substantive questions/comments rather than pointing out grammatical issues.

Specific Comments (L# refers to the line numbers of the manuscript):

L34: It's not clear in this sentence, but the westerly winds are the polar vortex. The westerly zonal mean zonal winds of the polar vortex can reverse during a strong (or major) SSW, but the three-dimensional polar vortex can undergo a displacement or split.

L35-36: This sentence is very unclear. The planetary waves from the troposphere can be modulated by the QBO, ENSO, etc.

L77-80: I don't agree with these assessments. Plenty of studies have been done that are able to draw robust conclusions about weather and climate phenomena based on simple SSW definitions.

L84: Earlier on you make the argument that reanalyses have inhomogeneieties and irregularities due to observing system changes; apparently this is not a big enough deal to prevent the usage of ERA5 for purposes of defining/characterizing SSWs?

L181: How do previous climatologies lack quality during the 1990s?

Figure 1: How are you determining the number of profiles for ERA5? With a resolution of 2.5x2.5 degrees, you have 144 lons and 73 latitudes, which corresponds to (144*71) + 2 = 10226 profiles over the entire globe (the poles only count as single points).

L248-255: This description of how you made decisions is not really sufficient. How were the thresholds in table 1 and table 2 chosen? These details should be available to the reader within this paper, since you are proposing your definitions be standard.

L265-269: Doesn't ERA5 assimilate the RO measurements? If so, then it is not really surprising that they agree well.

L283-287: Why do your diagnostics not characterize the depth of significant warming? SSWs that lead to persistent temperature anomalies in the lower stratosphere are thought to be the ones most likely to lead to coupling with the troposphere.

L291-292: This is by definition of how you characterize minor, major, and extreme events.

L349-351: Your statistics count all of your SSWs, including minor events. The numbers you are comparing are not directly comparable since, e.g., the 0.6 event/year figure corresponds to a particular definition for *major* SSWs. It would be more useful and more interesting if you listed the individual frequencies for your classification of minor, major, and extreme events (or major+extreme).

L365: I do not think this is correct. The 90s were a notably cold/quiet period for winters in the NH stratosphere. While minor-warming-like events did occur, there were not many that were associated with zonal wind reversals. This shouldn't have much, if anything, to do with the number of radiosonde stations.

Table 3 and Figure 9: Is it really appropriate to "double-count" with events that occur within 1-2 weeks of others? Most definitions of SSWs take into account that radiative timescales are pretty long in the stratosphere such that potential events that occur within ~3 weeks of another are considered to be part of the same event. By my reckoning of Table 3, there are 4 events which occur very closely in time to others (e.g., 1989-02-12 and 1989-02-20, 2000-12-07 and 2000-12-18, 2003-12-24 and 2004-01-04, 2006-01-11 and 2006-01-21). How are the results of Figure 9 impacted if these are not "double counted"?

Figure 9e: I am confused at what is plotted in these panels. Each of the points are decadal averages? How do you account for gap years (i.e., years in which you do not detect anything) and years with more than 1 event? You always divide by 10 years, rather than the number of events in the given decade? How sensitive are these results to the fact that you require a MPD of 7 days for an event to even be considered? I.e., how much do your results change if you instead considered events with at least a MPD of 5 and 9 days?