Response to Reviewer 1 Comments

Overall comments: *This is a well-done manuscript outlining the details of and application of a novel method for detecting and evaluating sudden stratospheric warmings (SSWs). The authors aptly couch their work in the context of the ongoing discussion within the SSW community about SSW definitions. They demonstrate that their method and definitions, at leastfl for the 2009 SSW, agree with established metrics and provide additional objective information. While the context of their work is centered around the use of radio occultation data, the authors show that using a selected model's data results in complimentary analysis, showing that this work may readily be applied to long-term reanalyses. I find this work to be properly placed in the literature and a novel contribution to the community. I do have a few comments I would like the authors to address prior to publication.*

Overall Response: We thank Reviewer 1 for his/her overall positive comments. We have carefully addressed all the comments as stated below and also in the Revised Manuscript (RM).

Minor comments

Point 1: *My main concern about the manuscript is on how clear the authors are in letting the reader know that the particular threshold choices are determined based on this one anomalous event. I appreciate that they do make this clear in the conclusions section, but that clarity was missing in Section 2.3 where the threshold values are introduced. In particular, I think the paragraph beginning on page 7, line 27 could use an additional statement(s) on this topic.*

Along these lines, I think some additional clarity in the statement on pg. 12, line 1 is warranted. Certainly, this SSW is known for being strong, but as-written, the authors seem to suggest that their method is sufficient to determine that this event is strong.

Given that the work in this manuscript is based off a single SSW, it's not obvious how that can be determined independently of other SSWs.

I think the authors should critically consider other areas of the text that would benefit from further discussion about this topic.

Response 1: Thanks for these comments. Yes, we agree that it is an improvement to discuss some more details how we select our thresholds for calculating the five TEAs and for formulating the metrics. We have hence inserted more statements from lines 3 to 7 in page 8, and from lines 18 to 20 in page 8 in the RM, with text as follows:

"*As the thresholds for calculating these five TEAs, we use those defined in Table 1, (4)–(8); for example, the thresholds for MSTA are 30, 35, and 40 K as seen therein. The selection of these thresholds was mainly guided by results on the polar-mean and regional mean anomalies shown in Sections 3.1 and 3.2. We examined the temporal variations of the magnitudes of warming and cooling of the five TEAs by sensitivity checks and finally chose suitable thresholds as summarized in Table 1 for illustration*

of this 2009 event."

"*The thresholds for formulating the metrics are selected based on the condition that the TEAs calculated for the chosen thresholds can suitably capture the main features of warming or cooling of the SSW event.*"

Yes, with the statement in the first line of page 12, we should not indicate that this event is a very strong or non-so-strong one, just based on this single event. Hence we modified this from lines 14 to 16 in page 12 in the RM, as follows:

"*In lines with previous studies on this particular event, and also on recent preliminary studies on several other events, the values of main-phase duration and also of the strength indicate that this is a very strong SSW event.*"

Point 2: *The authors bring up the Butler et al. (2015) requirements for a standard definition of SSWs. Missing from the manuscript is the authors' discussion on how their definition fits these three proposed criteria. These are criteria the SSW community has agreed upon, so providing additional contextualization of their method in light of these should be done.*

Response 2: Thanks for the comments. We have added discussions from lines 8 to 15 in page 13 in the RM to state how our method meets the requirements of standard definitions of SSWs. The updated statements are as follows:

"To summarize, the metrics proposed in this study for monitoring the SSW events can well satisfy the conditions that Butler et al. (2015) suggest for proposing a standard definition (cf. Section 1). Firstly, our approach well captures the sudden warming of the main phase and also its downward propagation into the lower stratosphere as well as the cooling occurring after the warming phase in the upper stratosphere. Secondly, the approach can be used for both RO and other suitable profile data and likewise for reanalysis data, and can be applied for both post-processing and in real time. Finally, the new approach is using anomalies over several height layers, and TEAs over larger area, and hence the detection and monitoring results are not sensitive to details such as exact latitude or pressure level. Potential further refinements of the thresholds for our metrics will be determined from recently started work on multiple SSW events, using longer-term data over the recent decades."

Specific comments

Point 1: *Do the authors report somewhere that bending angle and density are given in normalized units? This is apparent, but the reader would benefit from a definitive statement in the manuscript. As well, please state how the normalization is performed (normalized with respect to what?).*

Response 1: We have in fact shown the equations of how to calculate anomalies from (1) to (3) in our Table 1, which is used to concisely summarize our methodology of the full approach. To make this clearer, we have added the following statements from lines 20 to 22 in page 6 in the RM:

"*Temperature anomalies are calculated as absolute values, while density and bending angle anomalies are calculated as relative (percentage) values, by dividing the absolute-value anomaly profiles by the collocated climatological profiles.*"

Point 2: *Abstract, line 20: recommend "has strong potential."*
Response 2: Thanks, we have corrected to "has strong potential" in the RM.

Point 3: *Abstract, lines 17 and 22: is it necessary to introduce these metrics – the 3 Mio. km^2 threshold and the MSTA-TEA40 metric – here? I'm not sure that the abstract benefits from either the specificity of the former or the raising of the as-yet undefined metric and abbreviation of the latter. I would recommend removal unless the authors have strong objections.*
Response 3: Thanks for this suggestion. In the RM, we have avoided to introduce the metrics in the abstract. We improved the sentence from lines 22 to 23 in the first page in the RM as follows:

"temperature anomalies over the middle stratosphere exceeding 35 K cover an area more than 10 Mio. km^2."

Point 4: *Pg. 8, lines 25-26: I'm not quite sure I follow what's being said here. Is it that the specific definitions the authors have proposed may change as more systematic study is performed?*
Response 4: Yes, we try to say that the we can tell whether the SSW event is a strong one or a more minor one from our recorded duration and also strength. However, we agree that we can only know the specific values or thresholds of such determination of SSW events by applying our method to longer-term data records. By then, we can more robustly propose a definition for SSW based on the determination thresholds. To make our statements clearer for now in this initial introduction of the method, we have corrected the corresponding sentences from lines 5 to 6 in page 9 in the RM as follows:

"*However, the specific thresholds for our metrics and indicators for SSW detection, monitoring, and classification can only be determined after the new approach is applied to longer-term data containing multiple events.*"

We thank Reviewer 1 for his/her comments again. In addition to the comments and suggestions from the Reviewers, we inform that based on our careful rechecks of the full computation we found that a technical bug remained in our original calculation of the background climatology. It implied that we inadvertently had used the mean climatology of January and February of the year 2017 only, instead of the long-term mean climatology from the full range 2007 to 2017. We have corrected this in the updated RM.
We found that all our basic conclusions are still robust and valid; just the detailed quantifications changed, since the long-term climatology yields a somewhat different (and even more smooth) reference background. Figures 1 to 5 are very similar to the

original ones. The main impact is that the TEAs from the original thresholds are now smaller, while the TEAs temporal characteristics are strongly similar. Therefore, in the RM, we generally selected somewhat lower thresholds for calculating TEAs and for formulating our metrics, to keep their sizes and magnitudes rather similar. The Figures 6 to 7 reflect this and confirm that there was no qualitative change, just limited quantitative change that left the conclusions robust. Please find the RM for further details where we made small adjustments to the text accordingly.