

Response to Reviewer 3 Comments

General comments: *This paper proposes a new application of GNSS-RO for detecting stratospheric sudden warming (SSW) events. I appreciate the devoted efforts of the authors for developing an interesting analysis technique with the COSMIC GNSS-RO data and applying it to the major SSW event that occurred in 2009. The results are impressive in visualizing the horizontal distribution as well as the time evolution of the 2009 SSW event. However, I have several concerns about the data analysis procedure and the usefulness of the proposed techniques.*

Overall response: We thank Reviewer 3 for his/her comments. We have carefully addressed all the comments as stated below and also in the Revised Manuscript (RM).

Major comments:

Point 1: *This method is successfully applied to the SSW event of January 2009, which is a well-studied case. I am afraid that Section 3 is too descriptive. For comparing the analyzed results with earlier studies, any new scientific findings on SSW behavior should be reported. The technique should be evaluated, showing any new features of the SSW that can be uniquely resolved by GNSS-RO data.*

Response 1: Thanks for the comments. As we have discussed in the introduction section of this paper, one key motivation of this initial paper is to pave the way towards an improved standard definition of SSW, and such a definition does not exist yet due to lack of sufficiently reliable observational data and also the diversity of application purposes. Therefore, we intend to propose a robust method to detect SSW and subsequently to propose a standard definition of SSWs.

For our initial study here, we used this typical 2009 SSW event. The new method shows its basic potential for reliable SSW detection. The results obtained using this new method are consistent with results shown by previous researches. We will apply this new method in future to longer-term data records, and have just very recently started this work, to refine our specific settings and also thresholds to finally make a broadly useful SSW detection method. Based on this we also intend to introduce a standard definition of SSWs.

The prime new features of this study are hence in our method design. It can satisfy the requirements that were outlined by Butler et al. (2015), reading as follows (cited from line 33 to 34 in page 3 and from lines 1 to 2 in page 4 in the RM):

“The new definition should be proposed primarily for the purpose of describing polar winter variability. Secondly, it should be easily calculated and applicable to reanalysis and model outputs, both in post-processing and in real time. Finally, the new definition should not be highly sensitive to details, such as an exact latitude, background climatology, threshold wind speed, spatial extent, or pressure level.”

Our proposed method can describe the variability over all polar region before, during

and after the occurrence of SSWs and also over different stratospheric layers. The new features of our new method include that:

- (1) It can be applied to RO and similar profile data, and also gridded reanalysis data such as from ECMWF.
- (2) It uses anomalies over selected height layers and also introduces the concept of TEAs, robustly quantifying warming or cooling areas in polar region. This design makes our method properly insensitive to any perturbing fine details.
- (3) RO bending angles, and density profiles, are found useful auxiliary variables for SSW detection, but it similarly works for temperature-only data as well.

Based on our new method, Section 3 also shows all new findings obtained using our new method, including our metrics, TEAs and also RO bending angles anomalies during SSW, which have not been analyzed by any other previous researchers.

To make our new features clearer for readers to comprehend, as also suggested by Reviewer 1, we have added some discussions at the end of Section 3 (from lines 8 to 15 in page 13 in the RM), 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.”

Point 2: *Before proposing this method to monitor major/minor SSW events from long-term records, more cases should be tested to confirm that it is fully robust.*

Response 2: Thanks for this comment. Yes, we will make sure this method is fully robust by applying this method to nearly 1.5 decades (14 years) of RO data and also several decades of ERA-5 data, as we stated in the original manuscript (and see now lines 16 to 18 in page 14 in the RM, where we updated that the winter half year 2019/20 is now as well already available):

“Based on the encouraging demonstration in this study, follow-on work will apply the method to long-term RO and reanalysis datasets (RO overlapping 2006–2020 with reanalyses over 1979–2020) and assess its utility for long-term SSW monitoring.”

Point 3: *SSW is defined by the temperature anomaly in this study, but it is also characterized by zonal wind reversal. As GNSS-RO can provide only the former*

information, it may not clarify the entire behavior of SSW; therefore, this method may not be considered as primary. (See comment 1.)

Response 3: Thanks for the comment. As what has been discussed by Butler et al. (2015, 2018) and also in this paper, there is currently no standard definition of SSWs. There are currently nine often used definitions, in terms of zonal-mean winds at 10 hPa and 60° latitude, geopotential height anomalies, and also temperature anomalies, etc. Among these definitions, SSWs characterized by zonal wind reversal is one of the common used definitions. However, this does not mean that SSW definitions must be characterized by wind reversals.

According to literature, the SSW phenomena was first noted by its sudden and quick temperature increase measured by radiosonde since the 1950s. Wind reversal is one of its important characteristics occurring during this sudden warming. It is a useful metric for SSW. However, we also noted that selection of certain altitude and also latitude for measuring of wind reversals can result in different detection results. The SSW climatology obtained from previous studies, by using wind reversals at 65°N and 60°N, could result in very different results in some cases, with the center day varying more than 2 weeks. From this, we can see that if characterizing SSWs by wind reversals, results may suffer from detailed selection of latitudes.

Our method, which is relying on anomalies over several selected altitude layers, and also large threshold exceedance areas (TEAs), is not sensitive to such detail selection. The method is also not only suitable for RO or similar profile data, and also allows for real-time applications, and can likewise be applied to reanalysis data. Since reanalysis data also provide wind information, users who prefer to use reanalysis data could also extract wind information together with our method for SSW detection.

For our treatment in the manuscript, we tried to indeed provide an informative introduction and overview of this (without excessive length); and please see also our response to comment 1.

Point 4: *The analysis procedure is a bit complicated. It employs three anomaly parameters: temperature, density, and bending angle, in four altitude ranges. Out of a total of 12 values, only five parameters listed in Table 1 are used to monitor the SSW characteristics. It is not clear whether the selection of these five parameters will generally be adopted for any SSW event, or this set is used specifically for the 2009 SSW event.*

Response 4: Thanks for the comment. For formulating our metrics, we select five TEAs based on five thresholds. These thresholds will be further refined after we apply our method to longer data records, a work we just recently started. The method will then can be adopted for any SSW event.

Point 5: *Assuming that the COSMIC GNSS-RO data is assimilated into ECMWF, the 2009 SSW naturally appears similar in both GNSS-RO and ECMWF. Therefore, agreement of the SSW characteristics, as shown in Fig. 4, does not necessarily*

confirm the validity of the proposed GNSS-RO method.

Response 5: Thanks for the comments. The proposed method does in fact not just rely on RO data. As noted in a response above already, it is a method that can use both RO data and/or analysis data from model, such as ECMWF data. If other observation data can offer global dense distributed and high quality atmospheric profiling data over the stratosphere, the data can also be applied to this method.

The reason we incorporate RO data, but not just the model data is because RO data can provide accurate profiling for climate-type monitoring, but also available in form of real-time observation of the atmosphere. Furthermore, RO data have several distinctive advantages, such as the high vertical resolution in stratosphere (~1 km), global coverage with very repeatable quality, and the availability of useful auxiliary variables such as RO bending angles. It is such advantages that make RO data quite suitable for SSW detection on their own.

The impacts of assimilation of RO data into ECMWF is that ECMWF analysis data have higher accuracy in regions, where RO data of high density and accuracy than other observation data. It does not affect the validity of our method to ECMWF data at all.

Point 6: *The accuracy of the GNSS-RO data in the upper stratosphere and mesosphere (above about 40–50 km) should be tested carefully, because the error in the bending angle due to ionospheric effects could dominate, depending on the ionospheric conditions. Moreover, it is noteworthy that the bending angle profile at high altitudes is heavily optimized by referring to a model atmosphere profile, reducing the deviations from a climatological profile.*

Response 6: In our original manuscript, though we used so-called optimized bending angle, we have internally tested that using the ionosphere-corrected atmospheric bending angles directly (i.e., non-optimized bending angles) are very similar to the results from using optimized bending angle. In order to avoid this confusion in our revised manuscript, we only used the atmospheric bending angles for the calculation of bending angle anomalies now, so that “optimization” cannot play any role. In this form, atmospheric bending angles are indeed independent observations not affected by background model atmospheric profiles. In order to avoid somewhat higher short-scale variability of these non-optimized bending angles, we use a 2 km vertical averaging for smoothing over the short-scale variations. We tested again, and found confirmed, that the results from using the atmospheric bending angles directly are similar to the ones using optimized bending angles. The RM includes these updates accordingly. The figures related to bending angle are all updated to use non-optimized bending angle. We have also used describe this in the text from lines 22 to 23 in page 6 as below:

“In order to avoid the impacts of background model on bending angles, we used non-optimized bending angle in our study.”

Point 7: *As the analyzed values are a weighted mean over three days, the time resolution is longer than one day. Therefore, the time evolution of the SSW event,*

such as its duration and onset date, cannot be precisely determined at a daily resolution.

Response 7: What we do (as we describe) is that we use a weighted mean over three days, just to allow more RO events for a reliable statistical calculation. However, in order to get our daily sampling to achieve a near-daily resolution, we give the data of the middle day a twice-high weight of 0.5, while the weights of the predecessor and successor days are only 0.25. In this way the daily sampling is indeed still meaningful, while we statistically improve the robustness of the average profiles from enabling a bigger ensemble of profiles for the weighted averaging, by taking 3-day windows.

We note also that we have made detail sensitivity tests to make sure that our selection of temporal resolution is robust. We have calculated anomalies by just using one day's data and we have also calculated anomalies by using same weight of all three days (each day 1/3), etc. We found that if just using one day's data, the magnitudes of the resulting anomalies are similar to our weighted mean over three days, but somewhat noisier. If using three days' mean with even weights for all days, the dynamic daily evaluation of SSWs would be somewhat more blurred. So at the end we chose, as a most suitable trade-off, the weighted averaging that favors the center day's weight as noted above.

In order to make this clearer also for readers, we have added the following statements from lines 16 to 19 in page 7 as below:

“In this way the daily sampling is indeed still meaningful, while we statistically improve the robustness of the average profiles from enabling a bigger ensemble of profiles for the weighted averaging, by taking 3-day windows.”

Point 8: *I would encourage the authors to extend the latitude range below 60°N, as the effects of SSW on the middle latitudes and equatorial regions have been the subject of recent research. I would also suggest the use of ionospheric electron density data with GNSS-RO to identify SSW effects on the upper atmosphere.*

Response 8: Thanks for the comment. However, in this initial paper we prefer not to study at the same time the effects of SSWs on the middle latitudes and equatorial regions, but rather focus on introducing a robust new SSW detection method. Therefore, we decided on this focus on polar region here, though in future we are indeed interested to analyze also “teleconnection effects” towards the lower latitudes. Similarly, using of ionospheric electron density data to identify SSW effects are also beyond the focus in this study.

Specific comments:

Point 9: P2, L27–30: Show some references on the limitations of other satellite missions.

Response 9: OK. We have added two new references, i.e., McInturff, et al. (1978) and

Manney et al. (2008), which have discussed the limitations of conventional satellites data. Please refer to line 29 and 30 in page 2 in the RM.

Point 10: P3, L1–3: *Similarly, explain the limitations of the reanalysis data, referring to the relevant papers.*

Response 10: OK. We have added the Butler et al. (2015) also as a reference here, which have discussed the limitations of the reanalysis data. Please refer to line 3 in page 3 in the RM.

Point 11: P5, L19–25: *GNSS-RO data are neither distributed evenly nor regularly, but randomly with a relatively high horizontal density.*

Response 11: Thanks, yes. In order to make this description more accurate, we have updated the corresponding statements from lines 20 to 24 in page 5, now stated as follows:

“The upper panel shows the distribution of RO observations in the study domain from 50° N to North Pole, within which strong warmings were found by previous studies of the SSW event (Labitzke and Kunze, 2009; Harada et al., 2010; Kodera et al., 2011; Taguchi et al., 2011). It can be seen that most of the polar region is covered by RO observations. Similarly, high observation density also applies to the other days of the study period.”

Point 12: P6, L23–25: Is the GNSS-RO data assimilated into the ECMWF? If so, the agreement of the climatology is reasonable. (See comment 5.)

Response 12: Thanks, yes. And as responded to Point 5, assimilation of GNSS-RO data does not affect the application of our method, which can be applied to both RO and ECMWF data.

Point 13: P6, L32: Remove one of “the”.

Response 13: Thanks. We have deleted this “the” in the RM.

Point 14: P7, L9–11: Temporal resolution is lower than one day, which affects the description of the time evolution of SSW, such as its duration and onset date. (See comment 7.)

Response 14: Thanks. Following up to our response to Point 7 above, our strategy of three days’ averaging by given the center day twice the weight does not appreciably affect the quantification of the time evolution of SSWs, based on our sensitivity tests. We have also applied our method to the regularly sampled ECMWF data (2.5*2.5 grid points and four time layers a day), without such three days’ averaging for these data, but the conclusions are still the same as the other results discussed.

Point 15: P7, L14–17: For the four altitude regions, the exact height ranges should be provided here, even though they are shown in Table 1.

Response 15: Thanks. We have added the corresponding altitude ranges we used for these five TEAs (in lines 31 to 32 in page 7 in the RM).

Point 16. P7, L30–34: Is selection of the thresholds intended to be applicable to any SSW events, or specific to the 2009 case? (See comment 4.)

Response 16: These thresholds are currently applied to this 2009 case. We will refine these thresholds after we have applied our method to longer-term data records, which is a work just recently started. However, the method will be the same as now, but only with some refined threshold definitions.

We have discussed this in the original manuscript and now further updated the corresponding discussion (in lines 3 to 6 in page 9, in the RM and in lines 16 to 23 in page 14 in the conclusion part of the RM), as follows:

“In a follow-on work using long-term RO and reanalysis datasets, these indicators will be used to detect SSW events, for example by requiring a minimum main-phase duration of 7 days or so to qualify as an SSW, and to record the strength of the events. 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.”

“Based on the encouraging demonstration in this study, follow-on work will apply the method to long-term RO and reanalysis datasets (RO overlapping 2006–2020 with reanalyses over 1979–2020) and assess its utility for long-term SSW monitoring. In this way, the most suitable settings to use for the duration, area, and overall strengths indicators for robust SSW detection, monitoring, and classification can be determined.

In addition, we will be able to learn how the possible advantages in long-term stability and accuracy of the RO data play out or not in SSW monitoring in comparison to reanalysis data, including for different variants of RO processing and reanalysis. Overall, we expect the approach to be valuable for monitoring how SSW characteristics unfold event by event but also, and in particular, how they possibly vary under transient climate change and how they tele-connect to lower latitude regions.”

Point 17. P8, L2: Isn't 50% of the density deviation reasonable? It seems too large. (See comment 6.)

Response 17: Thanks for the comments. We have double checked all data and computations and find the density anomaly magnitudes up to roughly these levels, which are also consistent with bending angle anomaly sizes. Since many existing studies show polar mean anomalies or anomalies smoothed over multi-day periods, the anomaly magnitudes are somewhat smaller with those stronger smoothings. Here we exploit regional mean anomalies which are based on relatively small bin areas ($5^{\circ} \times 20^{\circ}$) on a daily sampling basis, therefore, the magnitudes are somewhat larger. See also the comment appended at the end of this response document; a technical correction (using the correct long-term background climatology) also made the anomalies a bit smaller, though in general for strong SSW events like this one in 2009 they reach these magnitudes.

Point 18. P8, L13: The word “then” can be read as “than”, right?

Response 18: Yes, thanks for pointing this out. We have corrected “then” to “than” in the RM (line 23 in page 8 in the RM).

Point 19. P8, L19–20: “number of days”. (See comment 14.)

Response 19: The “number of days” here suggests the number of days that our main-phase records last. Regarding the concern about the daily temporal resolution, please see our response to Points 7 and 14 above.

Point 20. P8, L22–23: How is the technique adjusted for long-term data? It sounds like this method is not fully robust, and a specific tuning is required for each SSW event. (See comment 4.)

Response 20: We have not applied this method to longer-term data in this initial study, we just very recently started such work. From our researches on this 2009 event, and also some complementary cross-check analyses (on the polar-cap mean, on daily mean anomalies of other years, etc.) we found that our selection of thresholds is basically reasonable. Yet, as mentioned already above in our response to Point 2, in follow-on work we will apply our method to almost 1.5 decades (14 years) of RO data and several decades of ERA-5 data, to refine our specific thresholds and make our method fully robust for the detection of multiple SSW events of various strengths.

Point 21. P9, L26–28: “differences above 50 km”. (See comment 6.)

Response 21: Please refer to our Response to Point 6.

Point 22. Section 3, P11, L24–26: I encourage the authors to show any new scientific findings obtained with the GNSS-RO data. (See comments 1 and 3.)

Response 22: Please see our extended response to Points 1 and 3. The method itself is all new so we prefer to keep the length and scope of this manuscript in its current form. We will be interested in such extended scopes in follow-on work. As to the new content of this manuscript, we note that Section 2 and also the results in Section 3 are all essentially new developments and findings of this paper.

We thank Reviewer 3 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.