Anonymous Referee #1

General information: We found a small bug in our MERRA-2 data reading routine. For every month that was read by the routine the last day of the month was missing. Thus there were three days missing each year. We fixed this. Note that our analysis is **not** affected by this bug. There are only a few more STE events now because there are three more days per year analyzed (39*3=117).

This paper presents a study on the contribution of planetary waves (PWs) to the local stratopause temperature variability based. It is based on the use of global MERRA-2 analysis to estimate the contribution of PWs wave number 1, 2 and 3 at fixed locations where Rayleigh lidar observations are available, Andenes at polar latitude in Northern Norway and Kühlungsborn at middle altitude in Northern Germany. In the first part of the results section the authors compare the stratopause characteristics at these two llocations retrieved from lidar observations and from MERRA-2 analysis. Two cases are considered, the overall wintertime stratopause climatology and the climatology of stratopause temperature enhancements (STEs). The rest of the paper is based only on the use of MERRA-2 data to estimate the contribution of PW wave numbers 1, 2 and 3 to the local stratopause variability. Although I consider that this subject may be interesting, I don't think that this paper brings new interesting information because it does not address the subject the right way as explained below.

The main reason is that the estimation of the contribution of PW 1, 2 and 3 to the total stratopause temperature variability is made using only MERRA-2 data on not taking advantage of having more resolved local lidar profiles. The reanalysis smooth out the small-scale perturbations that can increase the variability, including local temperature perturbations induced by gravity waves breaking and PWs with high wave number. Furthermore there are very few observations assimilated in the model at the stratopause altitude and in the mesosphere. This is not surprising that most of the variability in the reanalysis comes from the PWs with lowest wave numbers but this does not prove that it is the same in the reality. It shows only that MERRA-2 analysis captures mostly the contribution of larger scale PWs. It would have been much more interesting to use the MERRA-2 analysis to compute the PWs contribution to the stratopause temperature at lidar locations and to remove this contribution to stratopause temperature observed by the lidars. However this would imply that MERRA-2 reproduces faithfully the large scale temperature variability.

We only agree partly with the reviewer. To get PW characteristics out of local lidar profiles one would need at least five lidar stations ideally evenly distributed around one latitude. Those five lidar stations have to measure simultaneously just to identify unambiguously wave number 1. The amount of lidars needed to characterize higher wave numbers is even much higher. This is of cause desirable but does not reflect reality. Just because of this restriction not to do the study would definitely not help to better understand the interaction between local and global disturbances. On the contrary, with the help of this study local disturbances can be much better classified as real local disturbances caused by small scale dynamics or as phase shifts of planetary waves. Distinguishing between these two processes is essential. Thus to put local measurements into the global context using global reanalysis data is the best way to do this at the moment.

PW with higher wave numbers are very unlikely in the (upper) stratosphere due to the Charney-Drazin criterion which implies that PWs can only propagate upward in westerly winds that are not too strong. The upper critical limit depends on the zonal wave number, thus the critical zonal wind strength decreases with increasing wave number. For example, it is mandatory that the zonal wind is below ~10m/s for PWs with wave number 4 to propagate into the stratosphere. Thus synoptic-scale waves with wave numbers higher than wave number 3 can only propagate into the stratosphere under very special condition, for example during SSWs. However under normal winter condition the zonal wind at middle and polar latitudes is too strong (20 – 60m/s) and thus synoptic-scale waves cannot propagate to the stratopause.

The reviewer is right, it is possible that local temperature perturbations like STEs can also be induced by small-scale dynamics like gravity waves which might not be captured by MERRA-2. However, as we showed in our study, PWs dominate the STE development. Note that an additional analysis and discussion on the impact of small-scale dynamics on the day-to-day variability of the stratopause region (P12L19 – P13L3) reveal that the impact of small-scale dynamics seems to be larger at Kühlungsborn than at Andenes. We added a much more detailed analysis and discussion on differences between lidar observations and MERRA-2 data in the new subsection 3.3 (P8L7- P9L2) including a new figure (now Fig. 4) directly comparing mean STE profiles derived from lidar and MERRA-2 for both locations.

Starting in 2004, MERRA-2 assimilates MLS data in the upper stratosphere and lower mesosphere improving MERRA-2 output in that region (Gelaro et al., 2017; their Figure 21). To test the impact of MLS assimilation in MERRA-2 on our results we rerun our analysis restricting to the MLS period but got very similar results. A detailed discussion on that can be found in our answers to Reviewer 2 (her/his main comment 1 and 2). We added a discussion on this topic in section 2.2 and 3.3 including a comparison of climatology and mean STE profiles in the Supplement.

Furthermore there are very few observations assimilated in the model at the stratopause altitude and in the mesosphere. The comparison of the STE characteristics from lidar observations and MERRA-2 analysis made in sections 3.1 and 3.2 is also not convincing. I don't consider that the differences are small as it is claimed at line 8, page 6. For instance, at Andenes, there is a 7-km difference between the climatological stratopause altitude in MERRA-2 analysis (57 km) and in lidar observations (50 km). This is not at all a small difference. A careful comparison of average temperature profiles and stratopause characteristics should have been done. This is also a prerequisite to use MERRA-2 data for embedding the local observations in the global context.

The reviewer is right. The difference in stratopause altitude between the lidar and MERRA-2 is not small especially in Andenes. We corrected the respective text passages and as Reviewer 2 suggested, we performed our MERRA-2 analysis also for the period starting in 2004/05 when MLS is assimilated to MERRA-2 which should improve the quality around the stratopause. Firstly, we compare the vertical profiles of the climatology and mean STE from the complete data set with the period starting in winter 2004/05 (see Figure 1 below) for both locations.
In Andenes the stratopause altitude of the climatological vertical profile is about 5km lower in the MLS period (blue dashed line) than in the complete MERRA-2 data set (blue solid line). The difference to the lidar measurements (green dashed line) is now 2km which is much better compared to the completely available data set. In Kühlungsborn the stratopause altitude is slightly decreased by 2km in the MLS period compared to the completely available period. However the difference to the lidar measurements is now 1km which can be attributed to the vertical resolution of MERRA-2 in that altitude region. Thus the climatological profiles of MERRA-2 in the MLS period and lidar are very similar for Kühlungsborn.
However the difference between both climatological profiles of all available data and of the MLS period is very small at 2hPa the pressure level (see Figure 1 below) where STEs are defined here. We added a much more detailed discussion on the differences between lidar observation and MERRA-2 data set in section 3.3 and in the Discussion.

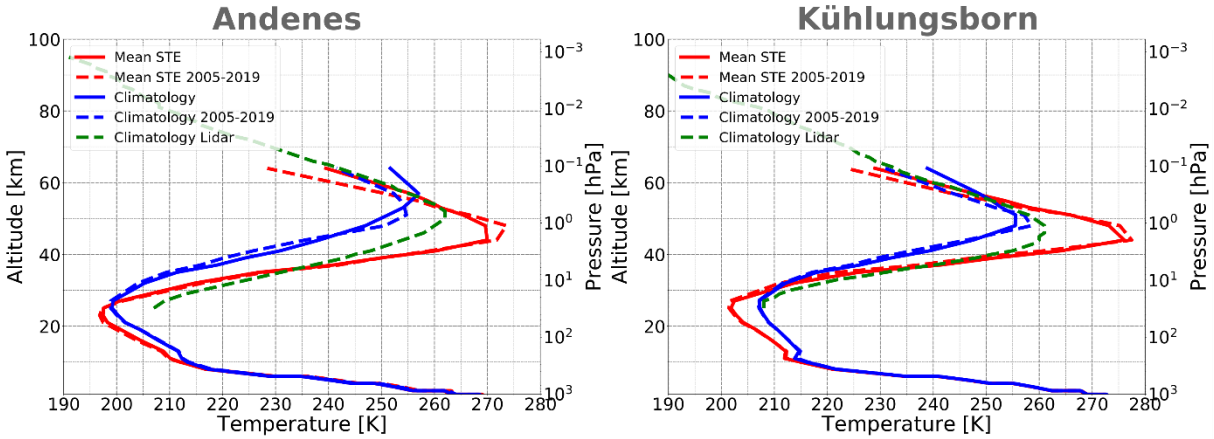

Figure 1 Vertical profiles of the climatologies (blue) based on the whole MERRA-2 data set (solid) and on the period starting with winter 2004/05 (dashed) as well as the mean STE profiles (red) again based on the whole MERRA-2 data set (solid) and on the period starting with winter 2004/05 (dashed). The green dashed line represents the lidar climatology.