Response to Referee 2

We want to thank Referee 2 for careful reading of the manuscript and thoughtful comments. Below are the point-by-point responses (blue) to each of the comments (black). During the revision process a small timing error was found in the analysis of the spectrometer data. Thus, the temperature and brightness time series plots in Figures 1 and 2 have been re-plotted and have changed a little. This update has resulted in another event to be classified as stable, so there are now 6 cases showing a temperature decrease and 2 showing no change in temperature as an immediate response to the particle precipitation onset. The temperature super-posed epoch plot is re-done in Figure 3, leading to no significant change. Figure 3 now also includes a superposed epoch plot for the airglow brightness as well.

1. General comments: The reviewer evaluates and agrees with the objective and motivation of this study. Quality of data set and analysis direction is also fine. However, the reviewer does not think the current version of the manuscript merits the publication because of lack of substantial verifications to their results and analysis.

   We have extended discussion and analysis according to the suggestions, as detailed below.

2. Discussion and further analysis focusing on variation in OH intensity are required. The authors mainly show variation in T_OH before, during, and after EPP. However, they do not show quantitative verification for intensity of OH airglow (I_OH) before, during, and after EPP as well. Since modulation of height profile of OH* airglow is an essential phenomenon to explain the observed T_OH, the authors have to show more detail and quantitative verifications for observed I_OH. For example, relative amplitude of decrease in I_OH during EPP is necessary to be quantitatively addressed. And then the amplitude has to be verified whether it is enough to change the T_OH with observed level. Empirically Modeled or observed background atmospheric temperature profile and typical profile of OH* intensity would be necessary for this verification. For background temperature profile, satellite data (MLS/AURA or SABER/TIMED) are best to be hired. If coincide temperature profile data are difficult to collect on event days, empirical model (e.g. CIRA) is another choice to know the typical background temperature profile. Anyway, the reviewer strongly recommends authors to check the typical temperature gradient during events weather observed decrease in I_OH can reproduce observed T_OH.

   We agree this being a weakness of the study and have been collecting more evidence. As pointed out by Referee 1 and as also shown by Suzuki et al. (2010), the model temperature profiles tend to be constant over the height range of the mesospheric OH. We searched for SABER/TIMED temperature measurements for the events analysed in this study but found nothing particularly close to the ground-based observation sites. Since the temperature is very variable at the heights of interest in the polar night, a large temporal (> 30min) or spatial (hundreds of km) separation between the ground- and space-based measurements is not giving a reliable comparison. Thus, we looked for random polar night temperature profiles over Svalbard in the SABER measurements. We further prepared a superposed epoch for the OH band brightness. It shows a typical decrease of brightness of about 20% at the EEP event onset. An example of an OH layer profile (red) and temperature (blue) over Svalbard in January 2019 is shown in an attached Figure 1. Here, the OH volume emission rate is an average of the two OH bands SABER measures at 2.0 and 1.6 μm, which corresponds to vibrational transitions above (9–7 and 8–6) and below (5–3 and 4–0) the one measured from the ground (6–2). As OH bands originate from slightly different heights we assume that averaging provides a reasonable estimate. The OH emission rate values are scaled to bring the profile comparable to the temperature values for illustration purposes. The peak of this average layer is at 86 km. A reduction of about 20% in the brightness would deplete about 2 km from the top of the layer. If the weighted average temperature is brought downwards by 2 km, the resulting temperature is about 20 K lower, as the temperature gradient under the peak height is about -10 K/km. This agrees with the typical change suggested by the epoch evolution of the temperature. However, this mechanism alone cannot explain the temperature decrease of 50 K during the first event, which has been pointed out in the discussion. By browsing temperature profiles from SABER measurements during polar night, it seems that an Earthward negative temperature gradient of 5–10 K/km is not uncommon, but a gradient steeper than that is rare. In the bottom panel of Figure 1, as an additional example, the Earthward negative temperature gradient at the heights of the OH peak if about 7 K/km.

3. Insufficient discussions to explain the observed variations in T_OH. As the authors mentioned in the manuscript, atmospheric parameters are highly variable in polar mesopause mainly due to existence of many kinds of atmospheric waves. In particular, small scale (∼10–100 km) atmospheric gravity wave is known to be major source causing large fluctuations with a period of hours to minutes. Authors excluded this possibility since the correlation between observed I_OH and T_OH is poor and amplitude of T_OH is greater than 10 K for all cases. Nevertheless, the authors also say that decrease in I_OH is shown in most cases (L260). In addition, authors also say that 'While a positive correlation can be seen between the two parameters in case of the fourth and fifth event, no significant correlation across the entire event set was found (data not shown).' (L243). Thus, it seems little bit inconsistent in their context explaining a relationship between T_OH and I_OH. Thus, the reviewer recommends the authors to re-organize their discussion about relationship between T_OH and I_OH. As the reviewer already pointed in former comments, authors must show more details for observed I_OH during EPP events. The amplitude over 10 K is possible and often seen in T_OH due to atmospheric gravity waves in a polar mesopause region. The phase between I_OH and T_OH is roughly positive but can shift each other depending on a vertical wavelength, damping factor, and a sign of vertical wavenumber of atmospheric gravity waves [Liu and Swenson, 2003]. Thus, authors should discuss more carefully to evaluate and exclude the effect of atmospheric gravity waves. For example, in a first event (29 Dec, 2007), there seems large fluctuation with period of 2-hours over the night in both T_OH and I_OH. In this case, phase of T_OH seems to lead the I_OH. This kind of feature is very common and typically observed in variation of T_OH even on no EPP days [e.g. Suzuki et al. EPS., 2010].

This part was indeed confusing and has been re-phrased to:
"*As the scatter plot includes data points from one hour before to two hours after the onset time, the lack of scatter correlation suggests that there is no longer-term or periodic coherent behaviour between temperature and brightness within the examined time period. The synchronous decrease in temperature and brightness seen in the epoch curves is a short-term feature, which does not dominate the scatter. A periodic out-of-phase relationship between temperature and brightness, which has been observed for non-EPP conditions (Suzuki et al. EPS (2010)) would result in low correlation but would not explain the synchronous decrease at onset.*"
The length of the time series of T_OH and I_OH data of the first event in Figure 2 is not long enough for reliably detecting periods of 2 hours. Even if it may seem that the temperature is leading the brightness variation, they still both show a minimum value at the same time after the particle precipitation onset. Time shifted scatter plots were explored and do not show an improved correlation.

4. Lack of verification on auroral contamination to OH spectrum data. During the night with active EPP, bright aurora feature would covers entire sky in typical. Since Meinel OH(6–2) band sits on wavelength between 825 nm and 860 nm, strong contamination from aurora light (including strong OI line at 844.6 nm) can disturb OH spectrum. Since T_OH is very sensitive to relative intensity of P lines, authors should address the how they judge the spectrum data is free from auroral contamination. The authors mentioned about accuracy of T_OH observation in section 2.2 as ±2 K. However, data shown in Fig 1 and Fig 3 have much larger error than this. The authors also should clarify about this point.

This is a good point. The auroral contamination part has been clarified in the new version: "*An oxygen auroral emission line at 844.6 nm lies in the middle of the OH(6–2) spectrum. The times when that emission intensity overtakes the OH emission intensity (fit covariance greater than 0.5) are excluded in the temperature analysis due an inaccurate fit. Other things causing poor fits and missing temperature values are cloudiness, high background illumination (e.g. scattered moonlight) or technical issues with the instrument. The threshold values for the fit variances have been determined empirically by viewing and fitting large datasets over decades. For consistency we have employed the same threshold values as in the earlier work (Sigernes et al. 2003, Holmen et al. 2013).*"
During the events studied here the auroral activity reached near the zenith where the radar beam was looking along the magnetic field line (about 7 degrees south of zenith), but did not fully cover the zenith-pointing spectrometer field-of-view. The auroral contamination is responsible for some of the missing data points (such as the one close the onset time of event 8), but mostly the optical measurements are collected right next to the energetic precipitation.

The error values have been further clarified to distinguish between the accuracy of the method itself and the range given in STD: "*The accuracy of the method in estimating rotational temperatures is ±2 K. The error bars and uncertainties given for the data in this study represent the standard deviations (STD) over the averaged time, which are typically somewhat larger (see values in Table 1).*"

5. Minor comments:
   Fig 1. Include a plot of I_OH as well as Fig 3. Fig 1. Include a vertical line to show the onset time in the plot.
   OH band brightness evolution has been included in Figure 1, and a vertical line showing the onset time as been added.

   Table 1. Add uncertainties in each value.
   Temperature values in Table 1 have been supplemented with ±standard deviation.

   L251 Reference Maeda [1967] is old. The reviewer suggests to add a recent paper modeling O3 destruction during EPP events. (e.g. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016JD0250154010.1002/28ISSN292169-9402.EEL15)
   The old reference has been replaced by the recommended one.

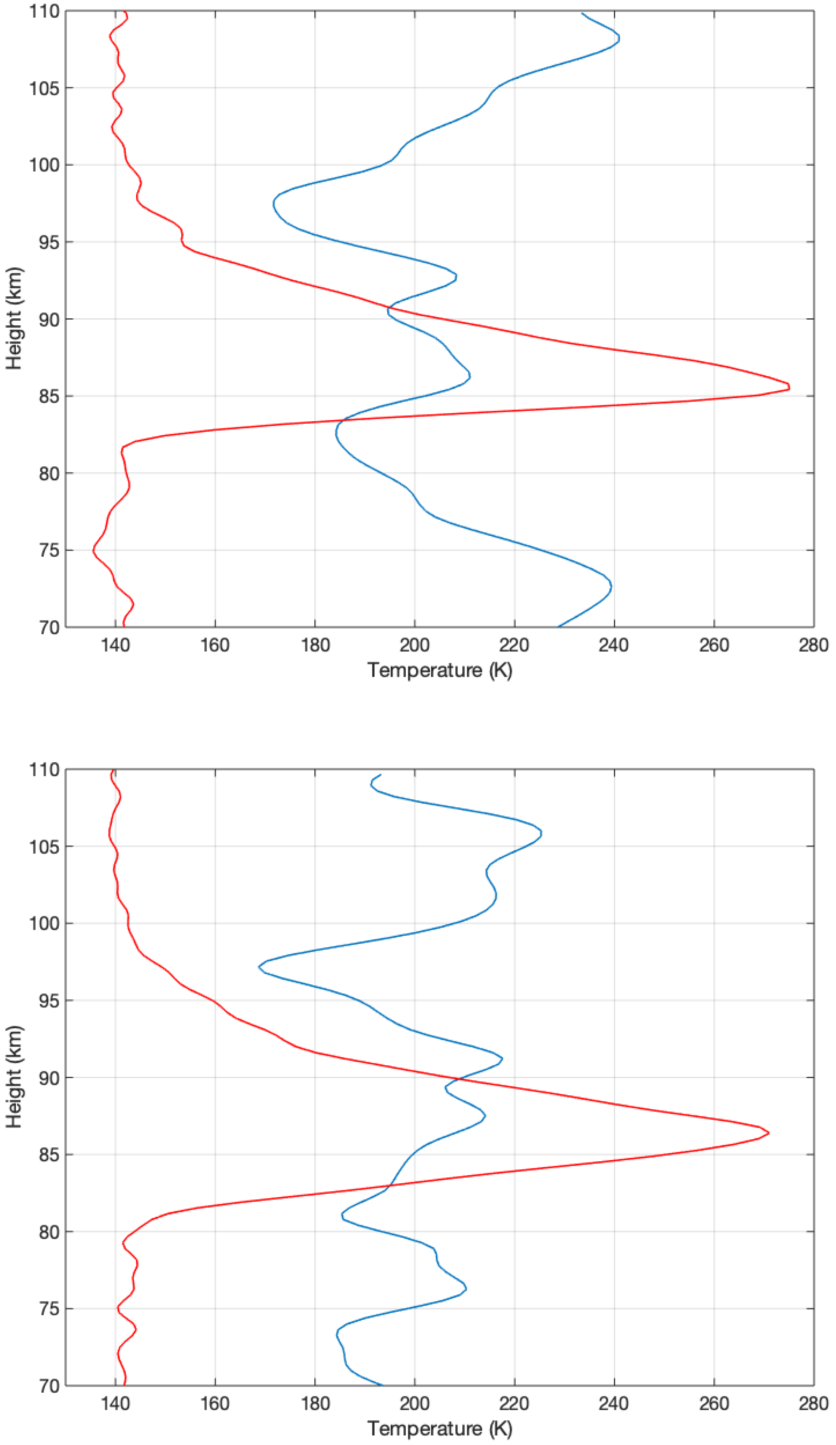

Figure 1. Two examples SABER/TIMED measurements of OH emission rate (red, scaled to temperature axis) and temperature (blue) over Svalbard in January 2019. Top panel: Orbit 92668, bottom panel: orbit 92624.