Response to Reviewer #2

We thank the Reviewer for their useful comments and suggestions that have greatly improved our manuscript. The following outlines the changes made to our manuscript in response to the Reviewer's concerns. Reviewer comments are in italics, and responses are in regular font.

Reviewer Comment 1: *There are multiple RO missions data set available from UCAR and ROMSAF websites, including KOMPSAT-5, PAZ, PlanetiQ, TerraSAR-X, TanDem-X, Sentinel-6a and GRACE-FO. These missions are contemporaneous with Spire and GeoOptics covering global area. They also show good penetrating capability and usefulness in PBLH application. Authors need to give explanation of the reason choosing C-1 and MetOp as the counterpart in evaluation for commercial RO.*

Response: The Reviewer makes a good point. The primary reason for not selecting other listed satellites such as, KOMPSAT-5, PAZ, TerraSAR-X, TandDem-X, Sentinel-6a, and GRACE-FO is because RO measurements is a secondary observable. They are primarily radar imaging and radar altimeter type instruments and do not offer consistent sampling of neutral atmosphere RO as their receivers are not always turned on, as opposed to COSMIC-1 and MetOp's GRAS (Global Navigation Satellite System Receiver for Atmospheric Sounding) that are dedicated GNSS RO missions, have similar orbital heights of roughly 800 km, and offer consistent sampling of neutral atmosphere at regular fixed local time. Moreover, we did not select PlanetiQ data for this study as we are focusing on NASA-purchased commercial RO datasets with at least one year of observations.

In the revised study, we have made one change to the source of MetOp dataset. Based on the Reviewer's suggestion, we have used MetOp data from ROMSAF, which is re-processed with an updated, single software version, similar to COSMIC-1 re-processed datasets from UCAR. In the revised paper, we have made this point in section 2.1.2., Line 119 (see below): "To remove ambiguity resulting from software updates - and to ensure consistency - only those RO mission products that have been re-processed with the same software version are compared against Spire and GeoOptics."

Reviewer Comment 2: *The radiosonde profiles from MOSAiC expedition ship campaign provide unique and valuable information on the PBLH detection in Arctic region as independent verification data. I expect to see the radiosonde derived PBLH and the direct comparison of the collocated RO-RAOB PBLH matching pairs. But in this paper the radiosonde data was only treated as source of water vapor for exploring the relationship between moisture and RO penetration. Consider the very limited geographic coverage of the ship campaign, does the conclusion cannot be achieved by using water vapor profile from climate model data, like MERRA-2 or ERA-5?*

Response: As the Reviewer points out, the radiosonde data are limited spatially and not ideal for comparing against GNSS RO-derived measurements which have a coarse horizontal resolution (100-200 km). As a result, the radiosonde profiles are not used for RO-RAOB PBLH comparisons. Instead, we use MERRA-2 reanalyses for comparing RO-derived PBLH. We agree with the Reviewer that the Arctic PBLH comparison should be the main focus of this paper, and we have therefore removed the section comprising the analysis of water vapor and RO penetration probability. MOSAiC radiosonde observations are no longer used as a dataset in the revised paper.

Reviewer Comment 3: *As far as I know, the MERRA-2 reanalysis provides two PBLH values. One is calculated based on the total eddy diffusion coefficient of heat (K_h), and the other one is estimated using the bulk Richardson number method. The PBLH calculated from K_h is usually higher than the one from bulk Richardson number in most regions. It would be interesting to check the difference over the Arctic area and validate it using RO derived PBLH. However, the paper did not provide much description on the MERRA-2 PBLH. Reader don't even know how the PBLH was extracted from MERRA-2 reanalysis.*

Response: The version of the GEOS model used in MERRA-2 includes two PBL parameterization schemes, viz. Lock scheme which is activated for unstable PBLs and the Louis scheme which is activated for stable PBLs. The model PBL depth is defined as the model level where the eddy heat diffusivity coefficient (K_H) value falls below 2 m^2 s^{-1} threshold. At a given time, only one PBL depth value is calculated by the model, either by the Lock scheme or by the Louis scheme. Both schemes use different methods of estimating eddy diffusivity coefficients, and therefore PBL heights. We have included a description of the schemes and a discussion of their relevance for the Arctic Ocean in the revised paper. Please, see Section 2.2.

Reviewer Comment 4: *In the PBLH deriving method section (2.1.3), the method of first minima of the refractivity gradient to exceed -40 N-unit km-1 and the 500 m threshold for RO cut-off height are chosen without justification. A sensitivity study for the threshold, comparison of different methods (minimum gradient, wavelet covariance transformation etc.) and variables (refractivity based, bending angle based) are recommended, according to the discussion for figure 6(b) and 6(c).*

Response: The goal of this study is to compare the PBL heights derived from commercial RO datasets using the previously established and validated methodology (Ganeshan and Wu 2015) for Arctic Ocean. This methodology is found to work well for cold season (Nov-Apr) months over the Arctic. Radiosonde observations from the SHEBA campaign were used to demonstrate that low specific humidity during the cold season months led to a heightened sensitivity of refractivity to temperature gradients (Ganeshan and Wu 2015). The radiosonde data show an empirical relationship between the height of the first local minima in refractivity gradient and the Arctic PBL temperature inversion height. Furthermore, the methodology is found to yield reasonable monthly mean PBLH values when applied to COSMIC-1 observations (Ganeshan and Wu 2015). Thus, in this study, the same methodology is adopted. The sensitivity to cut-off altitude threshold, however, is discussed in section 3.4.

Reviewer Comment 5: *The seasonal variation of RO penetration probability is displayed and discussed, whereas the more important seasonal variation of PBLH was not provided. In my opinion, a big picture of PBLH in north pole region is desirable (seasonal variation, diurnal circle if any, longitudinal variability related to the Atlantic Ocean current and sea ice* *distribution etc.) in the section 1.1, then a statement of how commercial RO can improve the understanding in section 3.3.*

Response: This is a good suggestion. We have included in the revised paper monthly mean RO-derived PBLH for six months of the year, showing seasonality and spatial variability due to the distribution of sea ice and open water (Figures 5-9). Since the RO-derived PBLH retrieval only works well for cold season months as described in Ganeshan and Wu (2015), we only estimate the seasonal cycle for November to April.

Reviewer Comment 6: *In figure 1, the different penetration probability of Spire NOAA and Spire NASA may contributed by the sample noncoincidence, because Spire NOAA is a small subset of Spire NASA (~3500 out of ~12000 in one day). Whereas for GeoOptics, NOAA and NASA are basically covering the same observations. Therefore the explanation of the discrepancy of orange/red lines may be completely different. Since the paper introduced NASA purchased commercial RO, which is processed by vendor, and NOAA purchased commercial RO, which is processed by UCAR, it's ideal to derive PBLH using both NASA and NOAA commercial RO, to help understanding the factors affecting RO penetration.*

Response: This is a great recommendation. In the revised paper, we have included a figure comparing the exact same subset of Spire radio occultations purchased by NASA and by NOAA. Indeed, the difference in the penetration statistics are due to differences in processing software as shown in Figure 2(a). We further drive home this point by including a figure (Fig. 2(b)) showing the similarity in penetration statistics between two different sources of radio occultations, viz. COSMIC-2 and Spire NOAA, over the tropics that are processed by the same software.