Response to Reviewer 1

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.

I appreciate the authors' efforts in addressing the reviewers' comments and suggestions, including adding a section on the sensitivity tests of cut-off thresholds for deriving PBLH. However, several key concerns remain inadequately addressed in the responses and/or revised manuscript:

Response: We thank the Reviewer for their first and second rounds of review, and for their useful insights. We have now included a table describing all RO datasets used in our study in the revised manuscript upon the Reviewer's recommendation.

1. Since this study incorporates multiple datasets, and given that the authors acknowledge the impact of processing algorithms on the RO penetration depth, it is essential to provide clear details in Section 2. Please specify which data centers processed the data, the data availability period, the average daily RO counts over the Arctic Ocean, data version, processing modes (e.g., real-time, postprocessed, or reprocessed), and other relevant information. A summary table could be helpful here to enhance clarity.

Response: This is a very helpful suggestion. In the revised manuscript, we now include a table with the requested information (Table 1).

2. Consider modifying the title from "Arctic region" to "Arctic Ocean" to more accurately reflect the study's geographic focus.

Response: Thank you for the suggestion. We have now modified the title to read "Exploring commercial GNSS RO products for Planetary Boundary Layer studies in the Arctic".

3. In Fig. 1, the authors compared the commercial dataset purchased by NASA and NOAA for the same month but not for a common dataset, which could result in sampling differences. It may be premature to attribute the observed differences in the penetration probability solely to the difference of the processing algorithms (Line 170-171). A similar concern applies to Fig. 2b, where the Spire and COSMIC-2 datasets over 30S-30N are compared. The sampling difference between these two missions may be significant due to their distinct RO count distributions with latitudes. The authors may consider using a collocated Spire-COSMIC-2 dataset to replot this figure. It could minimize the impact of sampling difference and provide more robust results. Additionally, if these statistics include the regions beyond the tropical ocean, terrain effects should be accounted for when generating this figure.

Response: We agree that differences between the penetration probabilities of NASA and NOAA GeoOptics data need to be evaluated for a common overlapping subset of RO profiles, as we did for Spire data in Fig. 1(b) of the revised manuscript. However, we think it would be repetitive to have a similar figure, hence we have removed the speculative statement concerning GeoOptics from our revised manuscript.

Figure 2(b) in the original manuscript comparing COSMIC-2 and Spire data over the tropics was only considering profiles over the tropical ocean (not land regions). Even though they were not collocated samples of RO profiles, the similarity in their penetration probability curve suggests that they are likely to have even better agreement when comparing a common collocated subset of ROs. However, this figure no longer appears in our revised manuscript. Instead, the same point is made by comparing Spire NOAA with COSMIC-1 data in Figure 1(a).

4. Fig. 4 shows that the daily RO counts reaching below 500 m for GeoOptics over the whole Arctic Ocean range from a few to 25. Such amount and variability raise concerns about whether GeoOptics data are sufficient to reliably capture the spatial variability of PBLH month by month. Could the authors comment on the reliability of GeoOptics data for deriving monthly PBL structure and variability?

Response: This is a good point. The revised manuscript now includes a table with the average monthly RO count for each satellite product over the Arctic Ocean. NASA GeoOptics has the least number of profiles, averaging roughly 754 per month. Given that nearly 80% of these profiles (~600) reach the altitude of 500 m, we think there are enough observations for deriving monthly mean PBLH maps.

5. The NASA Spire-derived PBLH exhibits lesser spatial and seasonal variability compared to the other two datasets, which the authors attribute to highly smoothed vertical RO retrievals. However, NOAA Spire RO data are not similarly smoothed. Why not present the PBLH derived NOAA Spire data to substantiate this explanation?

Response: The NOAA Spire data are a near-real-time product, and not available for our entire study period. However, we have used an example to show the superior performance of NOAA Spire derived PBLH in comparison to NASA Spire derived PBLH in Figure 10 of the revised manuscript.

6. The last paragraph of the summary lacks scientific accuracy. The discussion is rather weak without reliable justification. For instance, could the author define what constitutes a "smooth" versus a "dramatic" change in the decline rate of RO penetration?

Response: The revised manuscript does not show the rate of decline of RO penetration (Fig. 10 in old manuscript) as it was no longer relevant to our conclusions. We have removed the discussion related to the drastic decline in RO penetration observed for NASA Spire data.

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.

With the PBL finally receiving its deserved attention, this presents an interesting assessment of different instruments and data streams for PBLH detection in Arctic regions. I do however have several major comments. This is primarily related to my in-depth knowledge of Metop data (though that information is also publicly availablefrom the ROM SAF website), but also due to the presentation of results, use of figures, omission of some ROM SAF data.

Response: We have carefully considered all comments by the Reviewer. Particularly, we have taken into account the different nuances of available MetOp data streams, and have chosen the most relevant product for our study. We thank the Reviewer for bringing this to our attention.

Major Comments:

This reprocessed ICDR ROM SAF data represents a rather old processing setup, developed sometime in 2017 and frozen in time (it is an ICDR data set, as also mentioned on the ROM SAF website). The latest, Metop NRT data, is using a much improved processing setup. E.g. ICDR's lowest altitude reached for refractivity processing, using the first week of 2024, is on average 1.5km, while NRT's is 0.87km. The distributions also look very different. And, if you were to split this up further,you'll also find a different setting vs. rising distribution (due to the use of raw sampling tracking on Metop, not a "full open loop"). Thus, what you primarily see in Metopdata is similar to what you see in the 2 COSMIC data streams (Figure 1). Improved processing is available with more recent data sets. Maybe, to show that Metop datahas similar penetration improvements, it might be worth to include the ROM SAF NRT data stream too? Or, at least make it much clearer that the Metop data is notrepresentative of the current processing.

Response: We thank the Reviewer for sharing insightful information on the available MetOp data products.
Figure R1 (left panel) compares MetOp NRT versus MetOP ICDR for reference, and there is indeed improvement in RO penetration with the former.
Our study is focused on evaluating commercial RO datasets against established climate data records using stable, long-term observations. This goal prevents us from using NRT products as they lack consistency due to frequent software updates. This is explained in Line 114 in the revised manuscript (see below):
"Although MetOP near-realtime (NRT) product from ROM SAF has more advanced processing setup with improved lower tropospheric penetration, the goal is to compare with a consistent climate record to avoid ambiguities resulting from frequent software updates. Therefore, the ICDR data are used in this study."

Moreover, based on the Reviewer's recommendation, we additionally inspected the differences between rising and setting occultations for the MetOp ICDR dataset (right panel of Figure R1), and found that setting occultations have improved RO penetration, which is consistent with the findings of Innerkofler et al. (2023). As a result, we have only used setting occultations in the revised study, and the resulting penetrations statistics are much improved and comparable to other RO datasets.

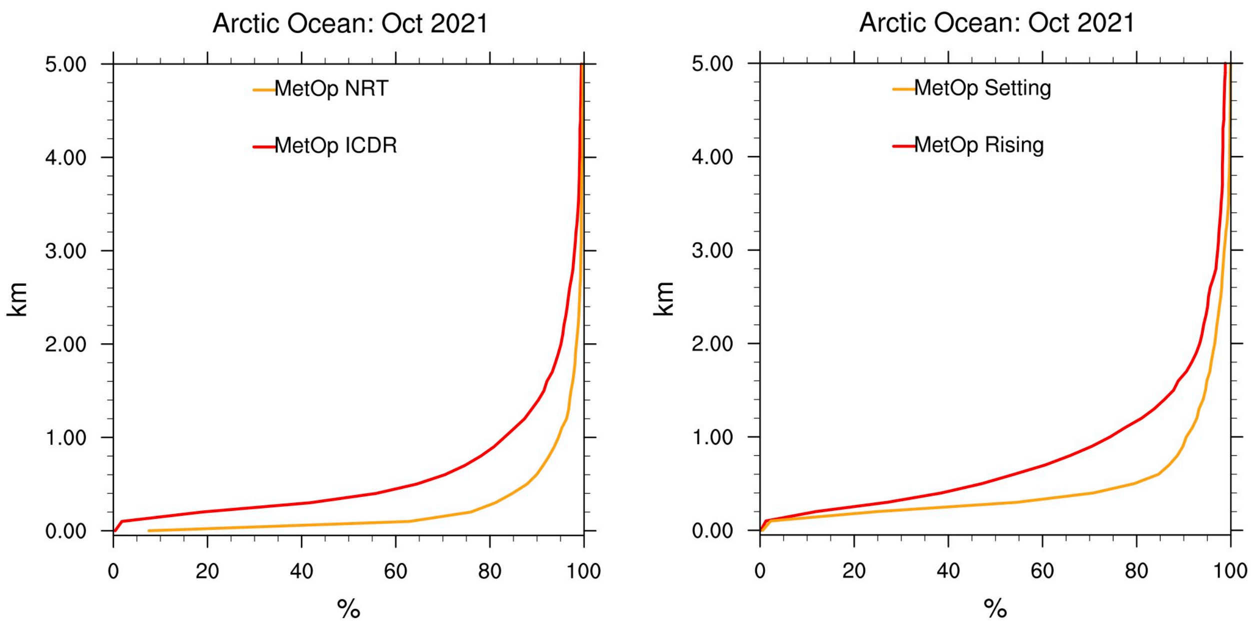

Fig. R1 RO penetration loss as a function of altitude over the Arctic Ocean (north of 60°N) for the month of October 2021 comparing **(left)** MetOp NRT and MetOP ICDR products and **(right)** rising and setting occultations from MetOp ICDR product.

Is the nine-point local smoothing applied in longitude, or latitude, or both? And if used in longitude, given that the difference between 2 longitude points is getting very small the closer one gets to the pole, what is the impact of this? And wouldn't it in general be better to apply this smoothing to all data sets, so that one compares likes with likes? And use a smoothing over a fixed area, not fixed latitude/longitude?

Response: We thank the Reviewer for pointing out the disadvantages of using grid-point based smoothing for polar regions. In the revised manuscript, we have produced new figures without using any smoothing. The results are not changed much.

P5L160: also due to the viewing geometry, as rising is more difficult to track, and sometimes even has a different antenna.

Response: This line (Line 178 in revised manuscript) has been modified to include the differences due to viewing geometry as suggested by the Reviewer.
"Penetration loss can also be different for measurements from the same instrument due to the viewing geometry, as rising occultations can be more difficult to track (Innerkofler et al. 2023), as well as due to inherent disparity in excess phase computations and bending angle retrieval algorithms".

P5L163: as stated above, this applies in the same manner to Metop data.

Response: Agreed. This has been noted in Line 114 of the Revised manuscript.

P8L216: It might be instructive to mention that reductions in Metop/GeoOptics are visible, as they show all data from those satellites. Spire on the other hand likely has a contract to fulfil, and can select from more satellites.

Response: This statement corresponds to Figure 4 in the old manuscript which is no longer shown in the revised manuscript. Instead, Table 1 shows the average monthly available observations for each RO product, and Spire has the maximum count. Figure 4 in the revised manuscript shows the annual time-series of the **percentage** of available RO observations at 500 meter altitude. A seasonality is clearly seen for GeoOptics and MetOp which is related to the sensitivity of RO penetration to atmospheric moisture. This is observed for all RO products (COSMIC-1, NOAA Spire, NOAA GeoOptics), however it is not seen in NASA Spire data, which is duly reported in Line 244 of the revised manuscript.

P9L229: There are 5 figures covered in this one paragraph. Each has 6 sub-figures. I assume this can be shortened to 2 figures, and only pointing out differences withrespect to a few representative ones?

Response: We thank the Reviewer for this suggestion, however, we find it useful to look at the actual PBLH spatial distribution and seasonal evolution as a measure of the qualitative performance of each product.

P15L281: Is there actually anywhere in the manuscript some information on what data amounts / number of occultations you compare? And is the Spire NOAA data set the same size / number of occultations as the Spire NASA one? Are they covering the same occultations, sometimes, or always? Somewhere above you mentionoverlapping Spire data, but is that only in time, or are these also the same occultations?

Response: In the revised manuscript, we now include a table (Table 1) which lists the monthly average count of all available RO profiles over the Arctic Ocean for each product. The Spire NOAA dataset has fewer samples compared to Spire NASA, and the daily profile count is dependent on their respective purchase agreements with the vendor. There are some overlapping RO profiles between the two datasets on a given day. In Figure 1(b) of the revised manuscript, a common sub-sample of radio occultations from both Spire NOAA and Spire NASA are compared, which are essentially the same occultations.

P16L309: The "contemporaneous" is not the correct term here, as the data you looked into are frozen in 2016. The contemporaneous Metop data (NRT) does show verydifferent characteristics.

Response: The word "contemporaneous" is no longer used in the revised manuscript.

P19L404: This ROM SAF data set contains PBL height estimates from bending angles, refractivity, etc. Have those at all been looked at within this investigation?

Response: No, we did not compare against the PBLH variable within the MetOp data files. Our investigation is focused on the Arctic Ocean, for which a region-specific refractivity-based PBLH retrieval algorithm has been previously developed and tested (Ganeshan and Wu 2015). Hence, we chose to use the established region-specific algorithm to derive PBLH rather than use generic global PBLH data.

Minor Comments:

P4L127: The cited work looking at the 500m cut off height is based on COSMIC-1 only. Here you look at different processing algorithms, different instruments, can youfurther justify that the 500m is applicable here too? ... Actually, reading further, I realized there is justification below. I would recommend to mention this justification herealready, otherwise one is in the blue for a while on why 500m is chosen and mentioned several times.

Response: Thank you to the Reviewer for this comment. We have now included a line in the Data and Methodology section (Line 167 of the revised manuscript), explaining that the sensitivity to cut-off altitude threshold will be explored.
"In this study, sensitivity of commercial RO products to the choice of cut-off altitude threshold will be additionally explored."

P9L239: Is the Spire data N. Bowler looked at the same as the one you are using, and this statement applies? Spire also updates their processing algorithms.

Response: The Level 2 Spire data in the Bowler (2020) study is also provided by the vendor (Spire), similar to the NASA Spire Level 2 profiles. While we cannot verify the version of the Level 2 product, we expect similarities between the two.

P15L292: Again, I think statements that are general about Spire in comparison to other missions need to be treated with care. If I got enough money, I could ask Spire to give me 10k occs a day that reach 300m, and they will provide this every day. So no seasonal variability is visible in that data set, as Spire has selected the occultations from a larger set, that might show the seasonal variability. On the other hand, one cannot easily task a Metop satellite to provide more occultations, so that any variability is reduced.

Response: We agree that seasonal variability in the number of RO profiles should only be considered along with the caveat that Spire is contractually obligated to provide a certain number of L2 profiles throughout the year. However, the seasonal variability in the penetration probability (which is a percentage of total profiles reaching a certain altitude) is more likely due to atmospheric conditions. All RO products, including COSMIC-1 and NOAA Spire, show a seasonal reduction in penetration probability during summer months at 500 m altitude and lower. Hoewever, there is no such seasonality in NASA Spire observations which is duly reported in the revised manuscript (Fig. 4).

Fig 11: Isn't the left plot already shown above? Might be thus better to just show the difference between 300 and 500m cut off.

Response: We no longer show this figure. Instead, the revised manuscript has a similar comparison (Fig. 10) for a different month (Feb 2024) along with a panel showing NOAA Spire derived PBLH for the same month.

Editorial Comments:

P4L114: I assume the reference to this data set should be stated here (it is in the references section, but never used).

Response: This reference has been added. We thank the Reviewer for their attention to detail.

P17L352: I assume this should be Bowler, N.E.?
Response: This has been corrected.

P18L367: Is Jarraud cited anywhere?
Response: This reference has been deleted.

P18L383: Is Maturilli cited somewhere?

Response: This reference has been deleted.

P19L386: Is Maennel cited somewhere?
Response: This reference has been deleted.

P19L415: Isn't this the "THE NASA PBL INCUBATION STUDY TEAM REPORT" document, e.g. available here: https://smd-cms.nasa.gov/wp-content/uploads/2023/05/NASA_PBL_Incubation_Final_Report_2.p

Response: Yes. It has been duly noted.

References:

Innerkofler, J., G. Kirchengast, M. Schwärz, C. Marquardt, and Y. Andres (2023). GNSS radio occultation excess-phase processing for climate applications including uncertainty estimation, *Atmos. Meas. Tech., 16,* 5217–5247, https://doi.org/10.5194/amt-16-5217-2023, 2023.