Response to Reviewer #1

We thank you for the comprehensive revision of our manuscript and your valuable and constructive comments. These comments have improved the clarity of our manuscript.

Here is a point-by-point response with the original comment in blue and italics. Our responses follow in black with line numbers referring to the revised version.

Evaluation of methods to determine the surface mixing layer height of the atmospheric boundary layer in the central Arctic during polar night in cloudless and cloudy conditions

This paper evaluates methods for determining surface mixing layer (SML) height in the central Arctic using tethered balloon, radiosonde, and tower-based observations from the MOSAiC campaign. The SML height detection methods involve the energy dissipation rate, bulk Richardson number, and Monin-Obukhov similarity theory. The analysis is conducted considering both cloudless (with a surface-based inversion) and cloudy (with an elevated inversion) conditions.

Overall, the paper is well-written and easy to follow, and provides value to the scientific community. However, there are many points throughout the paper where I suggest the authors provide more detailed discussions, so the results can be better interpreted in context of the complex Arctic boundary layer dynamics. While I provide many specific comments below, they are all relatively minor in nature, and after consideration of these points, I would recommend this paper for publication.

General Comments:

Throughout the paper, there is some inconsistency between some methods/descriptions, and actual application, as it related to the time of year for the study period. For example, the title specifies that the study is for polar night, and throughout the intro and methods, ABL processes, and related surface energy budget are discussed in the context of what is relevant/true for polar night. However, Fig. 1 shows that some of the observations occurred during polar day. At other parts of the paper, it is explained that the study is conducted for cases both during polar night as well as during the transition to spring. This all needs to be cleared up. If the study period is really polar night AND spring, processes relevant to spring (e.g., the presence of solar radiation) need to be discussed and considered. Or if you only are using the cases during polar night and the transition to spring before the sun comes up (and not the cases during polar day that are shown in Fig. 1) this needs to be better clarified.

Thank you for your comment. In our study we include both, the polar night and the transition to spring. We agree that the title was misleading and we have changed it accordingly: “Evaluation of methods to determine the surface mixing layer height of the atmospheric boundary layer in the central Arctic during polar night and transition to polar day in cloudless and cloudy conditions”

Further, we have clarified that the two states discussed in this study occur in winter and early spring (see e.g. Line 37). We also added some more description, e.g., “This is most pronounced during the polar night and gradually weakens in the transition to early spring.” (Lines 43f)

Specific Comments:

L16: *This first sentence is a bit weak. The authors mention "intertwined mechanisms and feedbacks" and "Arctic climate parameters" without examples, so it comes across as very vague. This sentence could be strengthened by adding more descriptive terms.*

We agree that this sentence was vague and we have changed it as follows: "Currently, the Arctic climate is changing rapidly driven by intertwined mechanisms and feedbacks, such as the lapse-rate feedback and surface-albedo feedback, leading to an increased near-surface air temperature and corresponding sea ice retreat (Serreze and Barry, 2011; Wendisch et al., 2023a)." (Lines 16f)

L23: *It is unclear what you mean by "The ABL is the atmospheric layer above the Earth's surface whose effects are perceptible on small time scales." Do you mean that the effects of the Earth's surface on the ABL are variable on small time scales? If so, say that more clearly. If you mean something else, please clarify.*

Yes, you are right, this formulation left room for interpretation. We have simplified the sentence to "The ABL is the atmospheric layer above the Earth's surface that is directly influenced by the surface (Stull, 1988; Garratt, 1997)." (Lines 23-24)

L34: *It might be useful to add some more examples of how the Arctic ABL is unique. For example, the lack of convection most of the time, and the absence of a residual layer because there is no diurnal cycle of the sun for most of the year.*

We have added your examples and thank you for your comment.

"The Arctic ABL is formed under unique conditions, such as the strong cooling of the sea ice surface due to the lack of solar radiation during winter, which favors the evolution of stable atmospheric layering. Furthermore, the ABL does not develop a residual layer due to the absence of a diurnal cycle for most of the year and even during the polar day convection typically plays a minor role (Persson et al., 2002; Tjernström and Graversen, 2009; Morrison et al., 2012; Brooks et al., 2017). " (Lines 33-36)

L41: *Perhaps remind the reader here that you are referring to the processes during polar night. Or another option would be somewhere in the Intro to state that the processes described henceforth are characteristic of polar night (while there may be different processes at play during polar day that are not described).*

As we include both, the polar night and also the transition to spring, we added the following:

"This is most pronounced during polar night and gradually weakens in the transition to early spring." (Lines 43-44)

L46: *Clarify that this is the case with low clouds. If there are very high clouds, there can be little to no effect on the ABL or height of the inversion.*

Thanks, we have changed "clouds" to "low-level clouds". (Line 47)

L77: *A more true statement would be that Rib is a measure of the likelihood of turbulence to exist, where Rib below the critical value indicates an atmosphere that is likely to become or remain turbulent and Rib above the critical value indicates that an already laminar layer will not become turbulent.*

Thank you, we fully agree that this was not completely correct. We have followed your advice and changed the sentence accordingly.

"Ri_b is derived from the ratio between shear and buoyancy and is a measure of the likelihood of turbulence to exist. A Ri_b below the critical value indicates an atmospheric layer that is likely to remain or become turbulent. Turbulence cannot be sustained and laminar layers will not become turbulent if the Ri_b is above a critical value." (Lines 78ff)

Throughout: *Sometimes you use the whole phrase 'bulk Richardson number' and other times you use the abbreviation 'Rib'. Should be consistent.*

We agree that there was still some inconsistency and have added: "In the following, Ri_b always refers to the surface bulk Richardson number." in Line 203f for clarification.

L97: *The statement "in different fields" is vague. Do you mean to say that the measurement included those of the atmosphere, ocean, sea ice, biogeochemistry, and ecosystem? If so, say that.*

Yes, this is exactly what we meant and now we have added the respective details following your suggestion.

"The MOSAiC expedition facilitated measurements onboard RV *Polarstern* and on the ice floe covering the atmosphere, the Arctic ocean, sea ice, ecosystem, and biogeochemistry throughout an entire seasonal cycle." (Lines 98-99)

Figure 1: *I assume the white background in panel c indicates the brief period when there was a full diurnal cycle including day and night? You should clarify this by adding that to the legend, or stating that in the figure caption.*

We have added the information in the caption of Figure 1: "The background shading indicates the respective daylight conditions during the observation period with the white background indicating the period with a diurnal cycle."

L124: *Specify the distance between Met City and the balloon operations. This is important, considering that the greater the distance, the more likely that the two instrument sites are sampling a different or evolved airmass.*

This is a very good point and important for interpretation. The distance between Met City and the operation site of the balloon was always about 300 m but changed over time (according to Shupe et al., 2022, Figure 4 and Table 1 therein).

"There were about 300 meters distance between *Met City* and the balloon operation site increasing over time due to ice flow dynamics." (Lines 130-131)

L128: *It would be good to note that the net irradiance you are referring to is the net longwave (or as you say, terrestrial) irradiance, which is a proxy for radiative energy budget at the surface only during polar night when there is no shortwave (solar) radiation. However, regardless of the presence or lack thereof of solar radiation, the net longwave irradiance can be used to differentiate between cloudless and cloudy conditions. Clarify all of this in the text.*

We have added "terrestrial" here (Line 134). The explanation for using the net terrestrial irradiances as cloud indicator is given in Lines 148-149. We have followed your advice and added following sentence: "F_{net} can be used to distinguish between cloudless and cloudy conditions during both polar day and night."

L139: *Jozef et al. (2022) showed that over 3 hours, when comparing UAS to radiosonde observations, the ABL height did not change significantly at the 5% significance level. You could add this reference to support your choice to compare coinciding balloon and radiosonde profiles, despite that atmospheric structure can change over this 3 hour time span.*

We agree and added the reference, thank you.

"The time difference between both launches is at most around 3 hours, in which the SML height typically did not change significantly as shown by Jozef et al. (2022)." (Lines 141-142)

L140: *A cloudless ABL can be even shallower than a cloudy ABL. This should also be mentioned when discussing observational challenges.*

While we fully agree that the even shallower cloudless ABL is challenging, we would like to point out here the cloud remote sensing challenge that very low Arctic clouds pose. The low-level clouds are often below the lowest detection limit of the cloud remote sensing instruments. To ensure that the reader relates shallower ABLs with cloudless conditions, we have edited the sentence: "The Arctic ABL can be very shallow, especially in winter and during cloudless conditions. However, also the cloudy Arctic ABL is often shallow, and hence poses challenges on remote sensing approaches to detect the very low cloud layers (Griesche et al., 2020)." (Lines 144f)

L177: *Some more description should be provided about how you examined the thermodynamic profiles and settled on the threshold value.*

We agree and have added the following more precise description:

"The highest threshold almost always yields the lowest h_ε values, and the two lowest thresholds are very close and result in nearly identical values of h_ε. To identify the appropriate threshold value, we have compared the SML heights derived by different thresholds with the potential temperature profiles. Additionally, we have examined whether the SML height coincides with a significant change in ε." (Lines 180 ff)

L185: *Please clarify that values below the critical value imply turbulence.*

We agree and modified accordingly: "This stability measure describes whether there is a tendency for turbulence to weaken or strengthen. Ri_g smaller than the theoretical value of 0.25 refers to an turbulent atmospheric layer." (Line 191)

L200: *It should be noted in the text that there could be a temperature offset between the met tower and the balloon due to a variety of reasons, e.g., the airmass evolved as it was advected between the two sites, or the two sites are sampling airmasses that were differently impacted by upwind features such as leads. This is a source of uncertainty in the Rib method that is explained.*

This is probably true but we consider this as a minor effect because the distance between *Met City* and the balloon operation site is only about 300 m.

L246: *How do you know that the elevated inversion base indicates cloud top in this example? You should provide some evidence to confirm this speculation, for example, from the MOSAiC ceilometer measurements. While what you say is likely true, this should be confirmed if you are going to make the definitive statement. I make this point because many studies have concluded that there is often a shallow stable layer between the ABL and the cloud, decoupling the two (e.g., Brooks et al., 2017).*

You are fully right that we have not confirmed our statement here and we decided to delete it because the exact location of the cloud is not important in this context.

Figure 4: *Is the red line the average? Please specify this in the caption for with a legend.*

Yes, this is the average and we have changed the figure and added a legend.

L294: *So have you shown here that the critical Richardson number does not necessarily vary based on atmospheric conditions (cloudless vs. cloudy)? Maybe state this as a conclusion, and note whether this agrees with any previous work, or is a new finding.*

We noted this in our conclusion, but it is difficult to say whether this is a new finding because other studies have used different definitions for the bulk Richardson numbers for cloudless and cloudy conditions (see Peng et al., 2023, for example).

L307-308: *What are the implications of this statement? If you are going to mention this, you should explain the potential impact on the results.*

We agree, this sentence was incomplete and we have added our intention: "It should be noted that the reference height h_ϵ was not determined at the same time as the height based on the radiosonde ascents, which may explain some of the observed differences." (Line 313)

L315: *Change to "… around half of the profiles contained an SML with height less than 150 m"*

We edited the sentence accordingly, thank you.

L317: *What conditions do the tethered balloons miss? Based on previous discussion, it seems the balloon would miss the stormy conditions, where wind speeds are too high for the balloon to fly. In these cases, the SML is much deeper, likely related to the high winds and also the presence of clouds. Explain this, and the implications for the results. For example, your cloudy conditions in this paper are all those with relatively low winds. What might you expect to see when you have clouds AND high winds? Do you think that would change the critical Rib number?*

We agree and modified as follows: "The tethered balloon operations primarily cover periods of lower wind velocities and shallower SML heights, and during storm events the SML can be much deeper. Whether Ri_{bc} may change during these events remains open and can not be answered in this study." (Line 321-323)

L325: *While different cloud characteristics can influence the SML differently, that is not the only factor at play here. Another important factor for SML depth is wind speed, which can also be highly variable. You could suggests here that the variation in SML during cloudy conditions may be affected by variations in wind speeds.*

This is a very good point and we directly added this thought to the manuscript:

"Furthermore, the wind velocity can play a role in the variation of the SML heights, as the wind velocity can vary significantly during cloudy conditions." (Lines 332-333)

Figure 11 caption: *"Further, shading refers to similar to daylight condition." Something sounds weird there.*

Yes, that was error-prone. We have changed the text in the caption of Figure 11 to: "[…] Further, shading refers to daylight conditions similar to Fig. 1. […]"

L355: *Do you have an explanation for this atmospheric vertical structure? Perhaps advection of a warmer airmass at higher altitudes contributed to the elevated inversion?*

We don't have an explanation but this stratification is frequently observed in the high-latitudes during winter. We added a reference and modified the sentence: "With the beginning of the twilight (15 February 2020, profiles 34 to 36, period III), the cloudless ABL was characterized by a surface inversion and an elevated inversion above, separated by a weakly stable to neutral stratified layer in between, as frequently observed in high-latitudes during winter Mauritsen (2007)." (Lines 361ff)

L377: *This sentence doesn't make sense. Is there a typo?*

Yes, we corrected it.

L382: *It should also be noted that for cloudy conditions, warming of the near-surface atmosphere, relative to in the absence of clouds, lessens the suppression of shear-driven turbulence by the static stability (because the longwave cooling is reduced).*

According to your comment, we have added following sentence: "In addition, surface cooling, and hence stability, is reduced in the presence of clouds, leading to less suppressed wind-shear-driven turbulence." (Lines 391-392)

L388: *Rather that saying the cause is irrelevant, perhaps say something like "the cause for turbulence generation is not of concern for the current study."*

Thank you, we adjusted the formulation. (Line 397)

L390: *This method also requires the collection of skin temperature or 2 m temperature, in addition to the radiosondes, correct? This could be seen as a drawback as well, and should at least be noted, perhaps with a suggestion of what to do if no such measurements are collected.*

This a good suggestion and we added it as follows: "However, the Ri_b method requires skin temperature measurements or reliable temperature measurements at a height of 2 m."

L395: *I find it hard to believe that not a single decoupled cloud situation was observed, as I would expect such conditions to occur occasionally throughout winter and spring. But you also did not include actual cloud height observations in this study to be able to make the claim that you make. So a more correct statement might be "For example, we were not able to quantify coupled versus decoupled clouds." I wonder if cloud coupling/decoupling could also explain some of the variability in SML height under cloudy conditions. In addition to repeating this study during the summer, you might also suggest for future work to add cloud observations so that the coupling/decoupling state can better be quantified.*

You raised a very good point and we fully agree that, first, our statement was too strong, and second, that cloud observations at the balloon would definitely add value to the discussion. We have changed the formulation (Line 406) and have extended the text with the following suggestion: "A valuable supplement to understanding the influence of coupling/decoupling is in situ cloud observations on the balloon." (Line 410f)

L425: *The sentence would read better as "… but profile measurements with radiosondes can also be useful to either…"*

Thanks, we changed it accordingly. (Line 436)