Response to Anonymous Referee Comments

We would like to sincerely thank the two anonymous referees for taking the time to review our manuscript and for their helpful comments, which have improved the manuscript. Each referee comment is given below in ***bold italics*** followed by our response to the comment. The line numbers provided in our responses refer to line numbers in the revised manuscript, unless otherwise stated.

Anonymous Referee 1

Comments about the revised manuscript

An Overview of the Vertical Structure of the Atmospheric Boundary Layer in the Central Arctic during MOSAiC

By
Gina C. Jozef, John J. Cassano, Sandro Dahlke, Mckenzie Dice, Christopher J. Cox and Gijs de Boer

First of all, I must say that my comments on the first version of the manuscript were addressed quite thoroughly and the revised version is much clearer and easier to read. A lot of effort has been put into the revision, but it has benefited the quality of the manuscript and is therefore worth it from my point of view. Besides a few minor points - mainly of a technical nature - I still have some concerns about the LLJ analysis which I will describe in detail below and which, in my view, need further discussion.

As already written in my first review: I am not a native speaker but for me some sentences and phrases still sound a bit bumpy but that is certainly a matter of taste should be judged by the native speaker.

We have made an effort to clean up the sentences for increased clarity throughout the paper. If there are certain sentenced you still find to be bumpy, please let us know.

I refer to the lines in the pdf of the revised manuscript version without track changes

General note on wording: in some places the manuscript still contains some very imprecise wording that should be better avoided, just a few examples:

i) A LLJ cannot be fast or slow - only the wind speed (in its core) can be fast (after line 505). In the same paragraph, sometime you just write speed, then LLJ speed - be consistent, in particular when comparing typical wind speeds of the LLJ you could introduce a parameter such as ULLJ or so –

We have removed the section after line 505 based on the comments from the other reviewer. This reviewer argued that the paper was too long, and the information in the LLJs and cloud sections at the end did not add great value for the results. Thus, we removed these sections and rather focus on presenting the important results relating to LLJs and clouds in relation to the SOM. This includes adding a SOM figure which presents moisture and cloud properties. We have still, however, checked throughout the paper to ensure we always refer to the speed of the LLJ (rather than the LLJ) as fast or slow, as well as revised the text to be consistent between saying ‘LLJ speed’ when referring to the speed of the LLJ. Whenever we refer to simply ‘speed’ this is because we are not specifically talking about the LLJ in that case.

or

ii) If you mean the “wind speed” than do not just write “speed” (see line 22)

This has been fixed here and throughout.

Line 46: what do you mean with flux perturbations? Please define!

The original phrasing was taken from Lesins et al. (2012), but we have clarified this, and now say:

"… by dynamically decoupling the surface from the free atmosphere, so that lower atmospheric warming related to increased surface to air (or decreased air to surface) heat fluxes cannot easily spread through the troposphere, and warming is concentrated near the surface (Lesins et al., 2012)." (line 46)

Line 58: the argument about ozone destruction is probably not wrong but a little out of the context – I would skip it

We have removed this sentence.

Line 61: I think decoupling is frequently observed but "typically" is probably going too far as there are have been also coupled ABLs observed

We now say "The ABL is often decoupled from the cloud layer by a shallow stable layer…" (line 71)

Line 86: If I interpret Egerer et al 2023 correctly, they argued that turbulence below the LLJ can be created by shear only during the collapsing phase of the LLJ and decoupling is one of the prerequisites of a LLJ (turbulent exchange coefficient tends to zero).

We do not fully agree with your interpretation of Egerer et al. (2023). Egerer et al. (2023) shows that there is a greater production of turbulence closer to the jet core, but this increased turbulence does not extend to the surface. That does not mean that no turbulence reaches the surface. Thus, wind shear associated with an LLJ can still contribute to turbulence near the surface, just not with as great a magnitude as closer to the jet core. Several other studies (Mahrt, 2002; Mäkirantaetal., 2011) rather argue that the wind shear below the core of an LLJ may be the main source of turbulence for a stable boundary layer (and are one of multiple sources of turbulence for a neutral or unstable boundary layer). Whether the LLJ is coupled or decoupled to the surface does not necessarily dictate whether the wind shear associated with the LLJ can contribute to turbulence in the ABL. To avoid confusion, we have removed the Egerer et al. (2023) citation and added the other references instead:

"Due to enhanced wind shear and subsequent turbulent kinetic energy production, an LLJ can contribute to mechanically generated turbulence below the jet core (Banta, 2008; Mahrt, 2002; Mäkiranta et al., 2011)." (line 79)

Line 109ff: why are the conclusions of the two references so different? (Although I think there should be better references than a textbook contribution). Are they both observations at the same time of year? I think if you are going to make such a strong comment, you should elaborate on this a bit more.

We have added some additional information which may explain the difference between the two results:

"For example, Esau and Sorokina (2010) claims that the central Arctic ABL is stable 70-90% of the time based on lower resolution observational and reanalysis data, while Tjernström and Graversen (2009) found stable and near-neutral conditions to occur with similar frequencies based on higher resolution observations from SHEBA in the Beaufort gyre." (line 96).

Line 132ff: It is hard for me to follow your argumentation: If you want to test if the results of MOSAiC agree with previous observations you should use the same method (SOM) with older data sets – right? There is probably nothing "wrong" with the past observations and only the conclusions drawn from the observations can agree or not. So, I am a little bit confused about your argumentation here – maybe it is only the wording.

We have decided to remove this sentence, as the real argument for the benefit of this study is that it adds to the body of literature on ABL vertical structure in the Arctic and includes measurements to fill gaps in our knowledge, through the use of new methodologies. Thus, we now jump right into these points:

"A self-organizing map (SOM) analysis (which objectively identifies a user-selected number of patterns present in a training data set) was conducted with the radiosonde profiles to reveal the range of vertical structures observed during MOSAiC (differentiated by stability within the ABL and the height and strength of a capping inversion), their relative frequencies, and their correlation to wind and moisture features during the MOSAiC year. The SOM results also were used to develop criteria to define stability regimes characterized by stability both within and above the ABL, such that features related to stability can be analyzed both in the context of the SOM patterns, as well as a more simplified grouping of observations by stability. Through the use of these new methods (i.e., the SOM analysis and detailed stability regime classification), the results provide further constraints on the vertical structure and features of the Arctic lower atmosphere that may be helpful to improve parameterizations of the central Arctic in weather and climate models." (line 116)

Line 156: such a reference should be placed in the bibliography and not inside the main text, maybe a matter of taste.

We no longer include the link in the text, but only in the full reference in the bibliography, while simply providing the in-text citation at this point:

"Instrument specifications and uncertainties for the radiosonde variables are provided in the manufacturer data sheet for the Vaisala Radiosonde RS41-SGP (2017), and are summarized in Table 1." (line 143)

Line 159: additional averaging can reduce noise but not uncertainties – please consider rephrasing and I think you should change "winds" to "wind speeds" or similar. I still have mixed feelings about this issue. All the problems related to wind speed determination, especially in the lowest 100 m or so, cannot be solved simply by vertical averaging.

We have changed 'winds' to 'wind speed and direction', clarified that vertical averaging only reduces noise but not uncertainties, and have added some more details as to why we are confident in the radiosonde wind speeds, even below 100 m:

"It is recognized that the true uncertainties in the wind speed and direction are likely to be greater than those provided in the data sheet, however for the following reasons, we find the original winds provided in Maturilli et al. (2021) to be sufficiently reliable for the current study. First, we determined that our results changed minimally when additional vertical averaging was applied to the winds (beyond the filtering already applied by Vaisala during their data processing), and thus noise in the observations does not bias the results. Second, when comparing radiosonde wind speeds to those measured by the DataHawk2 uncrewed aircraft system which observed the atmosphere during MOSAiC between 5 m and 1 km, Hamilton et al. (2022) found a difference of less than 1 m s^{-1} based on the 95% confidence intervals of observations from both platforms." (line 145).

What exactly convinces you that wind measurements above 35 m are trustworthy? I have my doubts and have heard several comments that radiosondes should not be trusted too much for wind measurements below 100 meters. I would at least make a clear comment that this problem exists and one should be somewhat careful with the interpretation.

In response to your above comment, we share a reason why we are confident in the wind measurements above 35 m:

"Second, when comparing radiosonde wind speeds to those measured by the DataHawk2 uncrewed aircraft system which observed the atmosphere during MOSAiC between 5 m and 1 km, Hamilton et al. (2022) found a difference of less than 1 m s^{-1} based on the 95% confidence intervals of observations from both platforms." (line 150).

Nonetheless, we have added a note that:

"Nonetheless, caution should be taken with interpretation of radiosonde wind speeds in the lowest 100 m." (line 152).

Line 168ff: The explanation of the bulk velocity is a bit misleading and needs a more precise scientific description including the equation to make sure that we all mean the same thing. You have the information from the ultrasonic anemometer to calculate u* based on co-variances how it is developed from theory so what's the problem with this calculation and why do you use a bulk estimate and how does the latent heat flux contribute to the friction velocity? Maybe I have overseen it in Andreas et al., 2010b but he uses in Eq. 2.1a the drag

coefficient and a mean wind velocity to estimate the friction velocity. For me it is not clear what you are exactly doing here and why? The argument that another method is frequently used by modelers is not really convincing.

We have decided to switch from using the bulk parameterization of u_* to the standard eddy-covariance value, calculated using sonic anemometer measurements at 10 m. Therefore, we have removed discussion of bulk u_* from the paper.

"Atmospheric observations of wind speed at 2 m above the surface, as well as friction velocity (a measure of the vertical fluxes of zonal and meridional horizontal momentum, suggesting the magnitude of mechanically generated turbulence; u_*) measured at 10 m, come from a 10 m meteorological tower (hereafter called the met tower) located on the sea ice near the *Polarstern* (Cox et al., 2023a,b). These measurements provide information about near-surface turbulence at the time of each radiosonde launch. Derivation of u_* through standard eddy-covariance methodology, and corresponding uncertainties, follow Persson et al. (2002)." (line 159)

Line 170: maybe you could combine bot Cox citations (Cox, et al., 2023 a,b) or so ?!

As the Cox et al., accepted paper has now been published, we have changed this to Cox et al., 2023b, and have changed the other Cox et al., 2023 citation to Cox et al., 2023a. We have also revised the sentence to read:

"… come from a 10 m meteorological tower (hereafter called the met tower) located on the sea ice near the *Polarstern* (Cox et al., 2023a,b)" (line 162)

Line 176: I think the ceilometer derives cloud base height from backscatter – right? From your wording it sounds like it measures both quantities (backscatter and cloud base height).

You are correct, and thank you for pointing this out. We have clarified the sentence to read:

"… which derives cloud base height (CBH) from measured atmospheric backscatter…" (line 166)

Line 185: please include in the table caption that these numbers are provided by the manufacturer (which might differ from the reality…) to make this clear.

The caption now reads:

"Instrument name and uncertainty for each variable used in this study, as provided by the manufacturer (real uncertainties may differ from those listed)." (line 175)

Line 338: "We make this distinction because we there are different processes that would lead to a shallow versus deep well-mixed layer." - The sentence does not make sense, maybe just delete "we" ?

This was a typo, so thank you for pointing this out. The sentence now reads:

"We make this distinction because there are different processes that would lead to a shallow versus deep mixed layer, which would be better highlighted by differentiating such categories." (line 344)

Figure 2: Still, the labels and numbers are partly quite small; in particular the numbers for the altitude in Fig 2 should be enlarged; same for the labels for the vertical lines in the subplot key of Fig 2. However, it is really appreciated that a subplot key is included to explain the complex figures and the figure quality has improved a lot.

We are glad to hear that adding the subplot key is useful. We have enlarged the vertical and horizontal axes fonts, as well as the font for the labels of the vertical and horizontal lines in the subplot key.

And the same for Fig 3: the upper line in the subplot describing the "Annual" is hardly readable but for the four seasonal cases it is still very tiny.

We have enlarged all of the fonts in Figure 3, and hope the figure is now readable.

Line 536: It is obvious what you want to express here but your wording could be more precise: What cloud property can lead to de-coupling of the sub-cloud layer (better than "below-cloud" layer) from the atmosphere above? This statement is quite vague…

We have changed "below-cloud layer" to "sub-cloud layer" (line 458). We have also added some additional information about cloud properties leading to decoupling:

"Properties of clouds and moisture can impact vertical θ_v structure and stability due to their radiative effect and ability to decouple sub-cloud layers from the atmosphere above (e.g., clouds which form from long-range moisture transport are often separated from the ABL by a stable layer, such that turbulence is not continuous between the ABL and the cloud)." (line 457)

I'm still having some trouble following this line of reasoning: What kind of relationships between mechanical mixing and atmospheric stability have you uncovered? It is probably true that around a LLJ - if the temperature gradient is not too strong - there is shear-induced turbulent mixing, but you have not measured it, so all very speculative. I am not convinced that friction velocity is generally an appropriate measure of potential mixing around the LLJ, or have I misunderstood the reasoning?

Perhaps a line number was missing from your comment, and we are not entirely sure which section you are referring to with your comment about the relationship between mechanical mixing and atmospheric stability. Were you referring to lines 599-601 in the previously-submitted manuscript? We will address this comment given this assumption. First, we have changed the word "reveal" to "suggest" as much of the results are inference, based on the fact that wind shear contributes to mechanical mixing. Additionally, we were not intending to say that friction velocity tells us about potential mixing around the LLJ. Friction velocity and the LLJ are two separate characteristics of the atmosphere which both separately suggest the potential for mechanical generation of turbulence, where friction velocity tells us about mechanical turbulence near the surface, and the presence of an LLJ tells us about mechanical turbulence at higher altitudes. The point of this sentence was to list all of the variables we looked at for inferring the mechanical generation of turbulence. We have attempted to clarify this:

"In the following discussion, we first summarize the relationships between wind speed features and atmospheric stability. Average wind speed and LLJ characteristics for each SOM pattern (Fig. 4), and wind shear and u_* within the ABL (Fig. 6c-d) suggest important relationships between mechanical mixing and atmospheric stability and vertical structure. Wind shear and u_* within the ABL quantify mechanical turbulence near the surface, while the presence of an LLJ can enhance mechanical generation of turbulence aloft." (line 593)

About the LLJ discussion:

In the revised version, you explain the two different ways an LLJ can form, but their characteristics and causes are quite different, and you have no way to analyze exactly what kind of LLJ you observed. Also, for the SOM analysis it is not possible to distinguish between the two LLJ variants although they have very different characteristics or are actually two completely different phenomena. On the other hand, this could be important for the interpretation of a possible correlation between the friction velocity and the properties of the LLJ.

However, from my point of view the friction velocity and the LLJ have not necessarily to do with each other - especially not for the LLJ which can be described by an inertial oscillation. The friction velocity only describes the momentum transfer from the surface to the lower atmosphere, and probably a low friction velocity is a prerequisite for the occurrence of one LLJ variant to minimize the turbulent vertical energy exchange (for the "inertial oscillation LLJ"). Therefore, I also have some concerns about the described situation and interpretation with stronger LLJs (faster wind speeds within the LLJ combined with higher friction speed) in a near neutral ABL, as you mention in line 603.

If a LLJ is caused by baroclinicity and is directly above the ABL according to your reasoning, how can this LLJ be coupled to the surface since most ABLs are covered by at least a weak inversion? By the way, to which section in Brümmer and Thiemann are you referring exactly?

The detailed description of the two LLJ types helps in the text, but what consequences it has for the analysis is just not clear to me yet. I think that this point of the manuscript needs some deeper discussion and especially the differences of the two LLJ types and in particular their consequence for the analysis needs to be better discussed.

You make a good point that with the data we use, we really have no way of definitively knowing whether an LLJ was formed by baroclinicity or inertial oscillations, so our discussion of such things is only speculative. Therefore, based on your comments here, as well as comments from the other reviewer, we have chosen to eliminate discussion of these two LLJ formation mechanisms from the introduction and results. Ultimately, the main conclusions we want a reader to draw from this paper are independent of the LLJ formation mechanism. With such edits, we have removed the Brümmer and Thiemann citation. Next, it seems one of your greatest concerns has to do with our discussion of the relationship between LLJ features and friction velocity. However, we did not intend to argue that the LLJ and friction velocity have anything to do with each other. We were simply trying to share that LLJ speed and stability have a similar relationship to each other as do friction velocity and stability. We hope by adding the following sentences, this is now clear:

"In the following discussion, we first summarize the relationships between wind speed features and atmospheric stability. Average wind speed and LLJ characteristics for each SOM pattern (Fig. 4), and wind shear and u_* within the ABL (Fig. 6c-d) suggest important relationships between mechanical mixing and atmospheric stability and vertical structure. Wind shear and u_* within the ABL quantify mechanical turbulence near the surface, while the presence of an LLJ can enhance mechanical generation of turbulence aloft." (line 593)"

Line 612: "These results suggest…" is this a surprising result? By the way, the argumentation is somewhat misleading: in a stable environment you need more wind shear to develop turbulence and finally mix the ABL – right? The amount of turbulence to mix the ABL is the same – independent of stability. It is a classical Richardson problem.

We have revised the sentence to better clarify our point:

"One possible explanation is that when the atmosphere is initially strongly stable (e.g., in the absence of clouds during winter), more wind shear is required to produce enough mechanically generated turbulence to fully mix out the near-surface layer than if the atmosphere is initially weakly stable (e.g., in the presence of clouds). Then the stable layer becomes elevated, separated from the surface by the mixed layer." (line 609)

Line 652: It sounds to me like you consider a "storm" and an LLJ to be equivalent? Maybe it is more about the wording but please clarify!

Based on comments from the other reviewer, we have removed the figure which shows LLJ frequency depending on the season, and thus have removed this paragraph from the Discussion and Conclusions section.

Anonymous Referee 2

This is a revised version of a paper dealing with an extensive analysis of the vertical structure of the Arctic lower troposphere, and especially the boundary layer (BL), from the MOSAiC year-long field campaign. Let me start by saying that the revised version is a clear improvement to the original manuscript and there is a lot of interesting new information.

That being said it is still not a great paper and this is sad because it could be a great paper. So, at this juncture the choice is whether to accept an extensive paper that has several flaws, in which case some revisions are still necessary, or if yet another major revision is required to make this the really great paper it could become. I will recommend major revision, because I want to maintain a high standard, especially when it is possible, but ultimately this is an editorial decision.

Major concerns

The first thing that becomes obvious is that the extensive scope of the paper is a problem. For one it is quite long;

close to 40 pages including figures and references (although the review manuscript is longer than necessary depending on how the figures are set). There are three reasons for this: 1) A quite long introduction and methods section; 2) that the manuscript is actually two studies merged together, and; 3) the inclusion of some other related data to the analysis of vertical structure (low-level jets (LLJs) & clouds) at the end.

We have made an effort to significantly reduce the length of the paper based on your comments. 1) In the originally submitted version of this manuscript, the introduction and methods sections were a bit shorter, but previous reviewer comments were asking for more details from prior literature on Arctic ABL stability and structure, as well as about the SOM methodology. Thus, in response to those comments, we added additional content to both the introduction and methods sections. Therefore, we do not want to remove too many details at this point, to remain consistent with the previous reviewer comments. However, we have shortened the introduction by about half a page and have worked to reduce the length of the methods section throughout, where possible. 2-3) We have removed the LLJ and clouds sections at the end, and instead have added some additional content/discussion regarding the SOM, so that the paper comes across as more of just one study, and is not including unnecessary content at the end. The amount of content added is less than that removed, so ultimately this reduces the length of the paper.

The introduction (~3 pages) reviews almost anything that one could think of being done previously on the topic and is written quite long. It also cover topics that are not at the center of this paper, such as decoupling, and I'm, sure it could be shortened by 30%. There is a 12-line paragraph on page 2 that deals with decoupling, which is important – indeed suggested even to be "typical" (I would have settled for common)"- however, the methods used in this paper makes it impossible to separate out decoupled BL clouds in this study. So even if it is indicated to be important in the introduction, there is no feed-back to this from the results in the study, nor are there any comments on this. By using a bulk-Richardson approach to determine the BL depth only the lower coupled layer will be detected. When this is later used as a criterium in the profile analysis, that seals the deal, although several of the upper-left SOM nodes clearly indicate what this reviewer interprets as a clear case of decoupling; a layer of high static stability – presumably a capping inversion – much higher than the indicated BL depth. At the very least this should be discussed. Also, while focusing on thermal stability, very little is discussed about the moisture structure. For example, what about moisture inversions; another feature where the Arctic seems to be special.

It was in response to a previous reviewer comment on the originally submitted manuscript that so much information had been included in the introduction. Therefore, we have worked to reduce it some, but did not want to remove too much, as we want to honor the original reviewer comment which called for a more detailed introduction. We have addressed your specific comments about the introduction and discussion. First, we now say that "the ABL is often decoupled from the cloud layer" (line 71) instead of "typically decoupled" as originally stated. Second, we have reduced the amount of text spent explaining coupled vs. decoupled clouds in the introduction, and have added some results and discussion which delve deeper into whether cloud may be coupled or decoupled for the various SOM patterns, which now better supports the inclusion of this information in the introduction . This is largely done by adding a SOM figure which looks at mixing ratio profiles, cloud, and liquid water path characteristics for each SOM pattern.

"This points to the varying cloud coupling or decoupling states, with respect to the surface. For example, cases in the lower right of the SOM (VSM stability) in which CBH is well above the ABL, are more likely to reflect the cloud-surface decoupling state, whereas cases in the upper left of the SOM (NN-stability) in which CBH is just above the ABL, are more likely to reflect the cloud-surface coupling state." (line 489)

"The varying depth of a well-mixed layer is likely a function of whether the ABL is coupled to a stratocumulus cloud layer: a coupled cloud supports a deeper ABL that is well-mixed up to cloud base (with the mixed layer extending to cloud top) whereas a decoupled cloud is separated from a shallower ABL by a θ_v inversion below cloud base (Brooks et al., 2017). Therefore, comparing ABL height to CBH suggests whether the surface and the cloud may be coupled or decoupled. The patterns in the lower right of the SOM with VSM stability have a shallower ABL capped by a θ_v inversion, with CBH several hundred meters above the ABL, suggesting surface-cloud decoupling. The patterns in the upper left of the SOM with NN stability have a deeper ABL with CBH often below the altitude of the θ_v inversion, suggesting surface-cloud coupling. For these patterns, which also have stronger winds, we theorize that the combination of relative warming of the near-surface atmosphere from the clouds, as well as the mechanical turbulence generated from wind shear allows for vertical mixing of the near-surface layer, which is strong enough to reach the level at which downward-propagating buoyant turbulence from cloud top cooling is present, creating a well-mixed layer between the surface through the cloud. Conversely, for stronger stability cases with a high CBH, the cloud and surface are completely decoupled such that the cloud is unlikely to impact the surface, and the strong stability persists. This discussion agrees with Sotiropoulou et al. (2014) which found that decoupled clouds typically occur at higher altitudes. However, the aforementioned discussion is at this point only educated speculation, and additional analysis based on the equivalent potential temperature profiles is required to confirm the cloud coupling or decoupling state." (line 640)

Lastly, to address your comment about lacking discussion on moisture structure and moisture inversions, the newly added figure of moisture profiles for each SOM pattern (Figure 5 in the revised manuscript), and associated discussion now cover this topic. For example, we now state:

"For most SOM patterns, the θ_v inversion corresponds with a moisture inversion (increase in mixing ratio with altitude) at about the same height, with the strength of the moisture inversion proportional to the strength of the θ_v inversion in most cases. There are however some exceptions where the moisture inversion is weak despite a moderate to strong surface-based θ_v inversion (e.g., patterns 5, 9, and 10). For the well-mixed layers (i.e., VSM, WS, and NN), below the elevated θ_v inversion, mixing ratio is relatively constant with altitude, or slightly decreasing. For pattern 15 (NN-WSA stability), there is no moisture inversion." (line 465)

"There is little relationship between ABL height, CBH, and the height of the moisture inversion. For patterns with surface-based θ_v and moisture inversions (SS and MS), CBH is typically over 2 km above the top of the ABL, but a correlation between the variables is not found. For patterns with elevated θ_v and moisture inversions (VSM, WS, and NN), in some cases the CBH is just above the ABL, at a similar level with the elevated inversions. In other cases, the CBH is well above the ABL, which could be at, above, or below the level of the elevated inversions." (line 485)

"The SOM showed that there is typically a moisture inversion at the same altitude as the θ_v inversion, with moisture inversion strength proportional to θ_v inversion strength, which agrees with previous studies (Naaka et al., 2018; Devasthale et al., 2011; Nygård et al., 2014)." (line 631)

This brings me to the duality of the study; the SOM part and the criterium-based analysis. It seems to this reviewer that the former is not at all necessary for latter and that instead this mix causes problems with the narrative of the paper. That doesn't at all mean the SOM analysis is pointless; in fact, I think that the SOM analysis, with a more in-depth discussion of the results, would make a very nice paper all by itself. The text argues that the SOM analysis is a prerequisite for the formulation of the criteria later used, but I see very little of that.

To address this comment, we have removed most of the criterium-based analysis, and have added an additional figure related to the SOM analysis (Figure 5 in the revised manuscript). We have also added additional discussion to explain why the SOM was a prerequisite for the formulation of the stability regime criteria.

"The SOM made the development of these stability regime criteria possible, as it revealed a manageable number of physically meaningful patterns representative of the entire training dataset, such that important variations between profile θ_v structures could be discerned. Based on these SOM results, stability regime criteria could be developed that were applicable to any θ_v profile, and thus were applied to each SOM pattern (using the average of all radiosonde profiles mapped to a given SOM pattern) as well as to individual radiosonde profiles." (line 313)

The SOM was crucial for revealing the range of stability types ultimately defined in Table 2, as if we were to look individually at nearly 1400 radiosonde observations to try to understand the range in stability, it would have been very difficult to make sense of all of it. The SOM, however, makes this possible by distilling the information in all of the observations into 30 patterns which represent the range in features present in the entire dataset.

This also leads to another peculiarity. First in the methods section (section 2.4), where the SOM analysis is referred to in a "hand-waiving" manner long before any of the SOM results are even presented, let alone discussed. Second, in Section 3.1 where the SOM results are now described already having defined the SOM nodes in terms if the stability classes, that here have not yet been presented and discussed. Either the stability classes a re dependent on the SOM analysis, and then you need to present those, followed by how they inform the

stability classes and then discuss those. Or you define the stability classes first, then do the SOM analysis and then identify where they fit. Now you're trying to do both and the result is confusing.

In order to help with the flow of the paper, we now include Fig. 2 in Section 2.3 when it is first introduced, and we explain that:

"The full range of vertical structures revealed by the SOM was used to develop a set of criteria for classifying stability of any given observation that distills the details of the SOM to the most critical factors of stability within and above the ABL, which will be discussed in Sect. 2.4." (line 296).

There are, however, many details in Fig. 2 which are not discussed until later in the manuscript (e.g., how we determined the stability regime of each SOM pattern). Therefore, we have also added that:

"Additional details included in Fig. 2 will be discussed later in the manuscript." (line 299).

Ultimately, it is up to the typesetter where this figure is inserted in the manuscript, but we hope by moving it to the methods section right after the SOM methods are discussed, as well as adding some additional discussion about how we got from the SOM to the stability regime classifications (see our response to your previous comment, and our response to your following comment) help to make this all more clear.

In the end, very little quantitative information from the SOM makes it into the criterium selection. If the authors actually did use quantitative results from the SOM directly informing the criterium selection in the vertical-structure statistics analysis, that needs to be much discussed in detail and how this is handled in the narrative needs to be clearer. That this doesn't seem to be the case is borne out by the results. There are very few WS cases; for some seasons there are only a hand-full or less and even annually there's one class with only 9 cases. Is there a point in having a specific criteria that hardly ever happens? To me this seems to be proof that the criteria was not determined objectively.

We have added some more details about how the SOM was used for the criteria selection:

"The stability regime definitions were developed alongside a similar SOM-based analysis of ABL profiles in Antarctica (Dice et al., 2023). An iterative process was conducted by visually inspecting the MOSAiC and Antarctic SOMs to identify groupings within each SOM which appeared to be substantially different from other groupings in that SOM, based on the near-surface and aloft stability. Then, thresholds (based on prior literature where possible) were determined to differentiate each grouping that made sense for the MOSAiC SOM and all the Antarctic SOMs. This process was completed considering both Arctic and Antarctic SOMs to support the robustness of these methods for classifying stability in either polar region, and to reduce subjectiveness. Further details about the determination of thresholds are provided below." (line 317)

Though the process cannot be completely objective, we attempted to make it as objective as possible by considering both Arctic and Antarctic SOMs, as well as by applying the same thresholds for categorizing near-surface stability as for categorizing stability aloft. This is why we end up with a class such as WS which has very few cases. In our opinion, it would be more subjective to decide that this WS category should be grouped together with another category, than to let it stand alone based on the thresholds for near-surface and aloft stability. Therefore, we stand by our decision for the stability regime criteria even though there are some classes that, especially when separated out seasonally, are rare. To find that these regimes are rare is in itself an important result. We have added some additional text in Sect. 2.4 to explain how we end up with the WS regime, for clarity:

"For the regimes listed as WS and NN, this means that the stability aloft does not fall into a category with greater stability than near the surface." (line 355)

At the end, the inclusion of the LLJ & clouds into the analysis is also not uninteresting, but somewhat superficial and doesn't add much to the results in general. Almost a page of the introduction (page 3) discusses LLJs and how they can form and LLJs are also taking up almost half a page in the methods section. Also, the cloud information if very superficial; how much clouds (a few octas, scattered or overcast) and how are multi-layer

clouds treated? Together the LLJ and cloud section adds 3 pages and two complex figures to an already long paper without adding too much new information on the physics, aside from the frequencies of occurrence.

We have removed the LLJ and cloud sections included at the end of the paper in the previously submitted version of the manuscript. We have also removed details about LLJs and clouds in the introduction and methods which are no longer relevant based on the removal of these final sections. Instead, we use the SOM to show the important results about how LLJs and clouds correspond to vertical structure and stability, and therefore some of the discussion on these topics is kept in the Discussion and Conclusions section.

In summary, the content of this paper holds great potential. The novelty of the SOM analysis is however underutilized and the SOM results could have been discussed in much more detail. The criterium-based vertical structure analysis, which I think is the core of this paper, does not really build on the SOM analysis, although some of the results refer back to it. Criterium-based studies are not really very novel but is here more detailed and extensive. But having both in the same paper in this way is at best confusing and is causing the paper to be very long. The addition of the LLJ and cloud analysis at the end makes it even longer without adding very much; both these aspects deserve better. So even if I'm typically not a fan of 2-part-papers, this is a case where I would advocate that method: Part I with the SOM analysis and a Part II with the vertical structure analysis, extended with the moisture profiles and a deeper analysis of the LLJs and cloud information.

We would like to first note that this work has already been separated into 2 papers: the one which you have reviewed, and another paper recently published in ACP which focuses on the thermodynamic and kinematic turbulent forcings which influence stability regime (i.e., delving deeper in the stability regime-based analysis introduced in the current paper). This paper (Jozef et al., 2023b) is referenced in multiple places (lines 115 and 660 of the previously submitted manuscript). Therefore, we do not wish to further separate the current paper into two papers. However, a lot of the major takeaways from the final LLJ and clouds sections in the current paper which you suggest that we remove are reiterated and explored deeper in Jozef et al., 2023b. Therefore, we are ok with removing these sections from the current manuscript, and have thus done so.

To address your other concern that the SOM is underutilized and that moisture profiles should be considered as well, we have added an additional figure and additional discussion to the SOM analysis which looks at moisture profiles and cloud characteristics in the context of the SOM. This is all discussed in more detail in response to one of your previous comments. We do choose to keep the figure and associated discussion which shows stability regime frequency distribution and ABL characteristics when stability regime identification is applied to individual radiosonde observations (Fig. 6 of revised manuscript), as we think these results are important to highlight some key takeaways from this study in a more straightforward, concise, and accurate way than can be understood from the SOM.

And possibly as a side issue, and not being a native English speaker, the wind speed can be "large" or "small". The wind may be "strong" or "weak". Bu neither are "fast" or "slow".

This has been fixed throughout.

Detailed issues:
Lines 45-47: If this is an explanation of the so-called lase-rate feedback, it needs another attempt. The surface heat fluxes have very little to do with Arctic warming; that is determined at the top of the atmosphere. The key issue here is the almost-total absence of deep convection.

The original phrasing was taken from Lesins et al. (2012) which discusses that Arctic lower atmospheric warming is disproportionately felt right near the surface. We have clarified this, and now say:

"… by dynamically decoupling the surface from the free atmosphere, so that lower atmospheric warming related to increased surface to air (or decreased air to surface) heat fluxes cannot easily spread through the troposphere, and warming is concentrated near the surface (Lesins et al., 2012)." (line 46)

Lines 56-60: This was a really looong sentence; I'm sure it can be broken up in two or three.

This sentence has been broken up, with some additional information added, and now reads:

"Near-neutral or weakly stable conditions can occur in the presence of stratiform clouds (Intrieri et al., 2002a; Tjernstrom, 2007; Curry and Ebert, 1992; Liu and Key, 2016; Shupe et al., 2011; Tjernström, 2005, Tjernström et al., 2012; Wang and Key, 2004; Zygmuntowska et al., 2012). Increased near-surface temperatures associated with enhanced downwelling longwave radiation caused by cloud cover can erode the surface inversion (Tjernström et al., 2019), which is sometimes supplemented by downward mixing from the cloud itself (forced by cloud-top radiative cooling) (Morrison et al., 2012). This is common in Arctic summer (Walden et al., 2017) when ample moisture is advected north either into the Arctic or from the broader ice-free areas across the pan-Arctic region. (Sotiropoulou et al., 2016; Tjernström et al., 2019)." (line 60)

Line 61: To me, "typical" means "very often", almost always. Studies indicate maybe 30% of the time, so I would say "often" instead.

This now reads:

"The ABL is often decoupled from the cloud layer by a shallow stable layer…" (line 71)

Lines 66-67: And how would this work? To me its not evident so a little more here would be nice

In order to shorten the introduction, this sentence has been removed.

Lines 68-71: Maye also a bit long; shorten or divide if you can.

In order to shorten the introduction, this sentence has been removed.

Lines 92-93: The very few direct studies of turbulence on top of a low-level jet that this reviewer has seen doesn't seem to indicate very much mixing.

Information given about LLJs in the introduction has been reduced per your other comments, so this sentence has been removed.

Line 97: Does "current paper" mean this manuscript? That begs the question, similar how?

To clarify, the sentence now reads:

"A study conducted using some of the same measurements as this manuscript found LLJs to be present more than 40% of the time in the central Arctic (Lopez-Garcia et al., 2022)." (line 84)

Line 94: Is the fact that models have a lower frequency surprise? Is there an explanation for this; comment please!

The explanation for this is provided in the Discussion and Conclusions when comparing our results to those from previous literature:

"Lastly, the difference in frequency from Tuononen et al. (2015) is likely because the much lower vertical resolution of the Arctic System Reanalysis (ASR-Interim) data used in Tuononen et al. (2015) would miss shallow LLJ cases." (line 626)

Lines 133-134: Strange mix of "firstly" and "additionally". If you use "first" there has to be a "second". If you use "additional", "first" is not necessary

Based on the comments from another reviewer, we have removed this sentence from the manuscript.

Lines 144-145: Change "with the result being being" to "resulting in".

This has been changed:

"During the MOSAiC year, many measurements were taken to observe the atmosphere (Shupe et al. 2022), sea ice (Nicolaus et al. 2022), and ocean (Rabe et al. 2022), resulting in the most comprehensive set of observations of the central Arctic climate system to date." (line 131)

Line 151: Discussion about Level 2 and 3 sounding data is meaningless without a description.

We have added a further descriptor for each product:

"We use the level 2 radiosonde product (Maturilli et al., 2021) for this analysis, as the level 2 Vaisala-processed product is found to be more reliable in the lower troposphere than the level 3 GRUAN-processed product (Maturilli et al., 2022)." (line 138)

Lines 157-161: This discussion is meaningless without more information. What data sheet is that?

This discussion was included in response to previous concerns about the validity of the radiosonde winds. We have now added some more information:

"Instrument specifications and uncertainties for the radiosonde variables are provided in the manufacturer data sheet for the Vaisala Radiosonde RS41-SGP (2017), and are summarized in Table 1. It is recognized that the true uncertainties in the wind speed and direction are likely to be greater than those provided in the data sheet, however for the following reasons, we find the original winds provided in Maturilli et al. (2021) to be sufficiently reliable for the current study. First, we determined that our results changed minimally when additional vertical averaging was applied to the winds (beyond the filtering already applied by Vaisala during their data processing), and thus noise in the observations does not bias the results. Second, when comparing radiosonde wind speeds to those measured by the DataHawk2 uncrewed aircraft system which observed the atmosphere during MOSAiC between 5 m and 1 km, Hamilton et al. (2022) found a difference of less than 1 m s^{-1} based on the 95% confidence intervals of observations from both platforms. Nonetheless, caution should be taken with interpretation of radiosonde wind speeds in the lowest 100 m." (line 143)

Figure 1: The makers are not seen with this resolution, so maybe use a line instead. Also, some of the yellow parts of the legs fade away into the white paper.

We have adjusted the symbol size and color for better readability.

Lines 168-170: This was a weird explanation of the "friction velocity". It is not a "theoretical wind speed"; it is a velocity scale derived directly from the momentum flux. It does not express the "magnitude of turbulence"; it is perfectly possible to have high turbulence and friction velocity, for example of turbulence is dominated by buoyancy.

We have revised the description of friction velocity to now read:

"… a measure of the vertical fluxes of zonal and meridional horizontal momentum, suggesting the magnitude of mechanically generated turbulence… " (line 160)

Lines 172-175: Using the eddy-covariance momentum flux gives exactly the momentum flux and nothing else; this has nothing to do with latent heat flux! There may be other reasons for using the bulk flux formulas, but this is not one of them!

We have decided to switch from using the bulk parameterization of u_* to the standard eddy-covariance value, calculated using sonic anemometer measurements at 10 m. Therefore, we have removed discussion of bulk u_* from the paper.

"Atmospheric observations of wind speed at 2 m above the surface, as well as friction velocity (a measure of the vertical fluxes of zonal and meridional horizontal momentum, suggesting the magnitude of mechanically generated turbulence; u_*) measured at 10 m, come from a 10 m meteorological tower (hereafter called the met tower) located on the sea ice near the *Polarstern* (Cox et al., 2023a,b). These measurements provide information about near-surface turbulence at the time of each radiosonde launch. Derivation of u_* through standard eddy-covariance methodology, and corresponding uncertainties, follow Persson et al. (2002)." (line 159)

Lines 211-212: Would be useful to know how much data was lost when retaining 1377 soundings. Does this number include any inetsive period with more than 6-hourly? 1377 divided by 4 gives a little over 344; very close to the 352 of a full year.

We have added some description about when the most data were lost:

"The 132 profiles which were removed from analysis are dispersed throughout the year, but many of them were observations from early October, mid-April (an intensive period when radiosondes were being launched every 3 hours), or mid-September." (line 202)

Lines 213-215: First, a bit more detail would be appreciated; the term "bulk" here can refer either to a finite difference across two layers (in contrast o a real derivative) or between a layer and the surface. Second, if the latter (which I tend to believe) this means, as far as I can see, that the BL top detected will be that of the lower surface-based layer and will not capture a decoupled but turbulent layer aloft, unless the decoupled layer is less than 20 meters on top. You need to be clear about this. The text in lines 218-220 is not enough, since this is such a central issue.

The bulk Richardson number profile for ABL height detection is calculated over a rolling altitude range of 30 m, calculated every 5 m. This way, a decoupled, but turbulent layer aloft should still be captured and considered as within the ABL. It is only when the layer truly is consistently having a Ri_b value above 0.5 that the ABL height is found. This is explained in detail in Jozef et al., (2022), but additional information has been added to the text:

"Ri_b profiles were created by calculating Ri_b across 30 m intervals in steps of 5 m, rather than using the ground as the reference level, in order to isolate local likelihood of turbulence rather than that over the full depth from the surface (Jozef et al., 2022)." (line 207)

Line 222: What do you mean by "below"? The wind at the surface below is zero, right?

Most previous literature phrases the description of an LLJ as a peak in the wind speed that is 2 m/s faster than the minima above and below, which is why we too phrased it like this. However, you are right that the wind speed minimum below is always zero at the surface, so how we phrased it may be confusing. We have changed it to say:

"LLJs were identified from each radiosonde, where there was a maximum in the wind speed that was at least 2 m s^{-1} greater than the wind speed minimum above (Stull, 1988)." (line 215)

Lines 231-234: This doesn't make sense to me. If you include cases where the LLJ is less than 25% larger than then the next minimum, how does that make you include high-wind environments? Or should it be "low" on line 234.

We did mean to say "high wind speed environments" on this line. For example, if the wind speeds are high throughout the entire profile, we may have a wind speed maximum of 20 m/s, and the wind speed minimum above is 17 m/s. This meets the criteria of the wind speed maximum being 2 m/s faster than the minimum above, though 20 m/s is not at least 25% greater than 17 m/s. However, if wind speeds are lower throughout the entire profile, we may have a wind speed maximum of 10 m/s, and the wind speed minimum above is 7 m/s. This meets the criteria of the wind speed maximum being 2 m/s faster than the minimum above, AND being least 25% greater than 7 m/s. We have added this example to the text:

"Our analysis differs from that by Lopez-Garcia et al. (2022) as they only considered LLJs in which the jet core speed was at least 25% greater than the wind speed minimum above the jet core, whereas we do not include this criterion, and thus our analysis also includes LLJs which occur in ubiquitously high wind speed environments (e.g., a wind speed maximum of 20 m s^{-1} would be 2 m s^{-1} faster, but not 25% faster, than a wind speed minimum of 17 m s^{-1} above)." (line 225)

Line 254: Is this correct? The figure seems to indicate the SOM is applied to the gradient of θ_v (see Lines 284-285).

You are correct, thank you for pointing out this typo. This has been fixed:

"Here, the SOM analysis is applied to radiosonde profiles of θ_v gradient to identify vertical structure and stability in the lowest 1 km of the atmosphere over the Arctic ice pack during MOSAiC." (line 248)

Line 277: "in the 1377"

This has been fixed:

"In this study, a 30 pattern SOM was used to describe the range of lower atmospheric stability profiles, defined by θ_v gradient ($d\theta_v/dz$), present in the 1377 MOSAiC radiosonde profiles." (line 271)

Line 285: Subtracting the value at 1 km does not make the result an "anomaly".

We have changed this to say:

"… in the form of the θ_v difference with respect to that at 1 km, to remove seasonal temperature dependence…" (line 280)

Line 287: What do you mean by "distinguished"? Maybe the wrong word?

We have changed this to say:

"… but found that the range in height and strength of the θ_v inversion, as well as the differentiation between a weakly stable or near-neutral layer below a θ_v inversion, were not as evident." (line 281)

Line 289-230: Wouldn't it be easier to calculate the specific humidity, then θ_v and if necessary linearly interpolate that directly, rather than calculating the pressure separately with the hypsometric equation?

In order for the value of θ_v to be most accurate, the individual variables should first be interpolated properly before calculating θ_v. Though perhaps easier to calculate θ_v first and then interpolate that linearly, this is not the most accurate method.

Line 300: Again, subtracting the value at 1 km from a profile doesn't make the result an "anomaly". Just a difference.

We chose to use the word 'anomaly' to represent that the θ_v profiles are not the raw θ_v, but rather the value of θ_v with respect to that at a standard altitude of 1 km. We understand that 'anomaly' is commonly used to represent a difference from a climatological mean, but there are other uses of this word in atmospheric sciences. For example, in Cassano et al. (2015), they use 'anomaly' to refer to the value of SLP at a given grid cell minus that averaged across the domain on that day. Previous studies similar to the current study (Cassano et al., 2016; Dice and Cassano, 2022) use 'anomaly' to represent the value at a given altitude minus the value at a standard altitude, similar to what we do. To avoid confusion, we have included additional description of our definition of 'anomaly' in this work:

"… with the mean profiles of $d\theta_v/dz$ and θ_v anomaly (where 'anomaly' refers to the value at a given altitude minus the value at 1 km) for all radiosondes mapped to a given pattern."(line 295)

Section 2.3 starts out by discussing relationships with a SOM analysis from which no results have yet been presented and discussed.

We now include Fig. 2 in Section 2.3, so that the SOM results have been shown by the time we begin the discussion of stability regimes, and how they relate to the SOM patterns, in Section 2.4.

Lines 326-335: Maybe just language, but in the definition of the criteria, how come "mixed" is used for something more stable than "weakly stable"? In my book "mixed" is a synonym to "near neutral".

We did not intend to communicate that something more stable than "weakly stable" is "mixed." We have revised the text for improved clarity:

"The second step for stability regime identification is only applied to cases with a near-surface regime of WS or NN and is carried out to differentiate such mixed ABLs (where NN is well-mixed, and WS is almost well-mixed) that are very shallow, from those that are deeper." (line 342)

Line 338: "… we there …" – drop "we".

This was a typo, so thank you for catching this. We have revied the sentence to read:

"We make this distinction because there are different processes that would lead to a shallow versus deep mixed layer, which would be better highlighted by differentiating such categories." (line 344)

Lines 364-365: I don't understand "… was never observed in an individual MOASiC profile …".. There seems to be several NN-like profiles in the upper left of the SOM.

We were intending to say that NN with no enhanced stability aloft (i.e., talking about the last row in Table 2 labeled as NN, not including NN-SSA, NN-MSA, or NN-WSA) was not observed in any individual profile, nor in any SOM patterns. We have revised the text to make this more clear:

"While NN with no enhanced stability aloft (last row of Table 2) was never observed in an individual MOSAiC profile (in the case of near-surface stability of NN, stability aloft was always weakly to strongly stable), we include its definition in Table 2 to support the use of these criteria for observations from other campaigns." (line 370)

Line 394-396: Isn't the number of SOM-patterns representing a certain stability a bit beside the point, as they are not equally populated?

We have added a sentence to explain why this is an interesting/valuable result:

"We note this, as a greater number of patterns of a given stability regime highlights greater variation in vertical structure within that stability regime category." (line 393)

Line 458: "Drop "in descending order" – pretty obvious if you read.

This has been removed:

"The most frequent near-surface regime observed was NN (37% of profiles), followed by VSM (27% of profiles), MS (14% of profiles), and SS (13% of profiles)." (line 507)

Line 475: What is df?

We explain in that same sentence that df is degrees of freedom. We have attempted to clarify this:

"The determination uses a two-tailed t-test when degrees of freedom (abbreviated as 'df') $\leq$ 100 and a two-tailed z-test when df > 100." (line 524)

Figure 5 and corresponding discussion: Well, now when you look at the results, compared to the other classes, there are very few WS cases, correct? For some seasons there are only a hand-full or less and even annually

there's one class with 9 cases. Does that tell you anything about the validity of the stability criteria? Is there a point in having a criterion that is so specific that it hardly ever happens?

As explained in a response to one of your previous comments, we included the WS regime for the sake of consistency in separating out weak to strong stability both near the surface, and aloft, following the same thresholds. To then discover that one of the regimes (WS) is very uncommon is an interesting result because it tells us that first of all, weak stability near the surface is the most rare of all near-surface regimes, and when we do have weak stability near the surface, the stability is usually enhanced aloft. We argue that to learn that a certain regime is very uncommon is just as important as learning which regimes are most common, as it adds to our overall understanding of the Arctic ABL. We have added some text to highlight the importance of this finding.

"Weak stability either near the surface or aloft is the rarest condition (demonstrated by few WS or -WSA SOM pattern and low frequencies of the WS and -WSA regimes). Thus, a near-surface regime of WS may represent a transition state between the stronger (SS, MS, and VSM) and weaker (NN) stability regimes, and there are rarely conditions to support weak stability aloft (-WSA). Discovering both the most common and the least common stability regimes are equally as important for our overall understanding of the Arctic lower atmosphere." (line 575)

Line 577: in a coupled system, the cloud is mixed all the way to the cloud top; not the cloud base.

We now clarify this:

"… a coupled cloud supports a deeper ABL that is well-mixed up to cloud base (with the mixed layer extending to cloud top)…" (line 641)

Line 582-584: This statement is so weak that it is entirely useless. Essentially all boundary layers are capped by an inversion, except for the stable boundary layer which is inside an inversion. Hence, there is always a stable layer in the lower troposphere regardless of what the boundary layer looks like. The distinction here is not the stable layer; it is the height. In the tropics you may have to go to several km to find the capping inversion, over the extratropical land maybe 1-3 km on a sunny summer day. The stability of the boundary layer is determined by the stability in the boundary layer. And that is not dominated by stable conditions. How hard is that to say?

Our main point here is that the stable layer, either within the ABL or capping the ABL, is usually strongly or moderately stable, but not often weakly stable. And this stable layer (capping inversion in the case of a mixed ABL) is located within the lowest 1 km of the atmosphere in the Arctic, which is in contrast to much of the planet where the capping inversion is located at much higher altitudes (as you point out). We have revised the sentence to make our points more clear:

"The most frequent stability regimes were those with strong or moderate stability either near the surface (SS and MS) or aloft (VSM-SSA, VSM-MSA, NN-SSA, and NN-MSA). Thus, we conclude that the central Arctic atmosphere over sea ice is inclined to include a strongly or moderately stable layer somewhere below 1 km AGL, and usually below 400 m (this contrasts with the mid-latitudes and tropical regions where the capping inversion is often as high as 1 to 2 km). Sometimes this strongly to moderately stable layer is within the ABL and sometimes it caps a well-mixed ABL, with the latter scenario occurring with higher frequency than the former, consistent with Tjernström and Graversen (2009). In the latter scenario, the depth of the well-mixed layer is highly variable, ranging from 38 m (minimum ABL height of a VSM case) to 914 m (maximum ABL height of an NN case). Weak stability either near the surface or aloft is the rarest condition (demonstrated by few WS or -WSA SOM pattern and low frequencies of the WS and -WSA regimes). Thus, a near-surface regime of WS may represent a transition state between the stronger (SS, MS, and VSM) and weaker (NN) stability regimes, and there are rarely conditions to support weak stability aloft (-WSA). Discovering both the most common and the least common stability regimes are equally as important for our overall understanding of the Arctic lower atmosphere." (line 568)

Lines 616-628: This paragraph is very speculative, to the point that I think it should be dropped. First, the discussion about how a LLJ is formed in relation to the LLJ core height and BL depth is very hand-waiving to say the least. Second, it is also dependent on the definition of the BL-height which here is not the same as the capping inversion depth. In all cases where there is a decoupling, one may find the LLJ at the top of the turbulent layer associated with the decoupled cloud and that would not necessarily be a sign of baroclinicity.

We have removed any discussion here and throughout the paper on the LLJ formation mechanisms, as you make a good point that this is all speculative, and we lack the analysis needed to really answer this question in the current study.

Line 638: If you mention the ASR, you will have to tell the reader what that is and why its resolution is lower.

We now include the full name of ASR, which reveals this is reanalysis, making it obvious as to why the resolution would be lower (reanalysis data usually have lower resolution than current observations):

"Lastly, the difference in frequency from Tuononen et al. (2015) is likely because the much lower vertical resolution of the Arctic System Reanalysis (ASR-Interim) data used in Tuononen et al. (2015) would miss shallow LLJ cases." (line 626)