Response to Reviewers 1

Title: Synergistic estimation of atmospheric boundary layer height from multi-source observations over the Arctic Ocean during MOSAiC

Author(s): Yuliang Liu et al.

MS No.: egusphere-2026-4501

MS type: Research article

Reviewer Comment:

This paper uses a large subset of the MOSAiC radiosounding data to identify the top of the atmospheric boundary layer, and train a machine learning model using profile data from multiple instruments (microwave radiometers, infra-red spectral radiometer, ceilometer backscatter, and radiosonde wind profiles) to retrieve the ABL depth. The model is tested against a different subset of the MOSAiC data, and an independent data set from the north slope of Alaska, and with both all input data sources, and with individual data types excluded. Testing is evaluated overall and for a variety of conditions (clear skies, clouds, and different stabilities).

Overall the machine learning model does a good job of identifying ABL depth, with accuracy being condition dependent: most effective under clear skies, with larger errors in cloudy conditions.

One conclusion is that the boundary layer depth during MOSAiC (and presumably over Arctic sea ice in general) is controlled mainly by mechanically generated mixing at the surface. This is an important result, and one that I don't think I have seen clearly demonstrated before. Nice work!

The paper is generally well written, but with quite a few (relatively minor) issues relating to missing technical details, or clarity of information, and to the logical flow of the narrative (see detailed comments). These should be easy to fix.

Overall I believe the paper to be acceptable for publication with only minor revision.

Response:

Thank you for your suggestions. We have carefully considered all suggestions and revised the manuscript accordingly. Our point-by-point responses to each comment are marked in blue font, and the corresponding revisions have been made. We believe these revisions have significantly improved the clarity and quality of the paper, and we hope the revised version meets the publication standards of ACP.

Once again, we would like to express our sincere gratitude to the reviewers for their valuable time and insightful comments, which have greatly helped us improve the manuscript.

Response to Comment #1

Reviewer Comment: [One overarching point regarding the presentation of the performance of the ABL depth retrievals is that the various measures (RMSE etc) are given as absolute values in metres. These look good overall (eg, for the MOSAiC 'all' evaluation, RMSE = 42.4m, bias = -10m), but the significance of the uncertainty/bias depends very much on the absolute depth of the boundary layer. A 42m uncertainty is a minor issue for a 1000m deep BL, but much more significant for a BL of 100m or less. It would be useful to also give uncertainties as fractions of BL depth, and to note any functional dependence on depth if

present.]

Response: Thank you for your suggestions. We revised the text to keep the original emphasis on absolute RMSE/bias differences across clear-sky, cloudy, and fog/mist conditions, and added a depth-normalized uncertainty analysis. The revised discussion now makes explicit that the estimation error depends on both weather condition and ABL depth, so absolute metre-level metrics should be interpreted relative to the true depth. The relative RMSE is 20.0 % at 50-100 m, about 10 % at 100-400 m and 16.1 % above 400 m.

Supporting numbers from the current analysis
Depth-resolved RMSE and relative RMSE:
- 50–100 m: RMSE = 12.1 m; relative RMSE = 20.0 %.
- 100–200 m: RMSE = 15.1 m; relative RMSE about 10 %.
- 200–400 m: RMSE = 26.8 m; relative RMSE about 10 %.
> 400 m: RMSE = 87.9 m; relative RMSE = 16.1 %.
In addition, because of the length constraints of the abstract, the detailed depth-normalized uncertainty analysis is reported in Sect. 4.1, where the depth-resolved errors are presented and interpreted in context. This keeps the abstract concise while providing the complete quantitative information in the main text.

Response to Comment #2
Reviewer Comment: [the construction of this sentence is a little odd, "…remain scarce, and ABL structure…": the two statements either side of " , and " are both true, but completely unrelated, while as written the grammar suggests the second statement is in some way dependent on the first. Better to rephrase something like "The Arctic boundary layer (ABL) structure is governed by stability, clouds, and wind shear, but is poorly constrained because of the scarcity of in situ observations"]
Response: Thank you for your suggestions. We revised the original phrase "…remain scarce, and ABL structure…" to a clearer and more reasonable version: "The atmospheric boundary layer (ABL) over the Arctic is governed by stability, clouds, and wind shear, but is poorly constrained because of the scarcity of in situ observations."

Response to Comment #3
Reviewer Comment: [Line 14: "The ABLH annual cycle is found to peak in…". This is ambiguous, what peaks? The depth of the ABL, or the 'cycle' (of some property of ABLH)? -> "The depth of the ABL peaks in May, and has minima in August and December…"]
Response: Thank you for your suggestions. ABLH is the atmospheric boundary layer height, so the ambiguity concerned the height itself, not a separate property of ABLH. The revised sentence now states that the height has a marked seasonal cycle: "The CMNN-estimated ABLH shows marked seasonal variability during MOSAiC, reaching a monthly mean of 356.7 m in May and its lowest monthly means in January (129.9 m) and July (140.6 m) (Figure 9)."

Response to Comment #4

Reviewer Comment: [Line 46–47: the statement "persistent surface-based temperature inversion and longwave radiative cooling are the primary mechanisms maintaining the boundary layer in a near-neutral state" is a little misleading. A surface-based inversion is, by definition, stable, and longwave radiative cooling at the surface will promote stability not near-neutral conditions. Turbulent mixing generated mechanically by wind shear at the surface will tend to oppose this, and maintain a near-neutral condition. Note that a few lines down, at line 53, you cite the same Tjernström paper for the statement "longwave radiative cooling and the absence of solar radiation promote the widespread formation of stable boundary layers": same process (longwave cooling) but different result.]

Response: Thank you for your suggestions. The original statement was physically incorrect and misrepresented Tjernström et al. (2019). The revised text now distinguishes stable conditions maintained by surface-based inversions and longwave cooling from near-neutral conditions maintained by mechanical mixing or reduced near-surface stratification. The revised paragraph reads:

"Previous studies show that stable and near-neutral ABLs dominate in the Arctic. Persistent surface-based inversions, maintained by longwave radiative cooling and the absence of solar heating, favor strongly stable conditions, whereas weakly stable or near-neutral conditions can arise when wind shear or cloud-induced surface warming reduces the near-surface temperature gradient (Tjernström et al., 2019; Jozef et al., 2023a)."

We believe this revision accurately reflects the findings of Tjernström et al. (2019) and corrects the physical inconsistency. We apologize for this oversight.

Response to Comment #5

Reviewer Comment: [Line 51: "…warm and moist air masses transport northward…" -> 'transport' should be 'transported' or better 'advected']

Response: Thank you for your suggestions. We have revised the sentence to replace "transport" with "advected." The revised sentence reads: "In summer, warm and moist air masses advected northward may pass over a colder sea-ice surface, favoring the development of temperature and moisture inversions and shallow stable ABLs (Rantanen et al., 2022)."

Response to Comment #6

Reviewer Comment: [Line 55: "…atmospheric stratification stability…" -> "…atmospheric stability…"]

Response: Thank you for your suggestions. The redundant phrase has been removed; the revised sentence uses "stability" in the list of governing factors: "The atmospheric boundary layer (ABL) over the Arctic is governed by stability, clouds, and wind shear, but is poorly constrained because of the scarcity of in situ observations."

Response to Comment #7

Reviewer Comment: [Line 66–68: this statement start "Several observational techniques…" but then lists instruments. The instruments on their own, or even the measurements from them, don't constitute techniques: the measurements must be processed in some way to derive BL depth; and multiple processing/analysis techniques may exist for any given measurement type. Rephrase something like "ABL depth can be diagnosed from a variety of measurements

from instruments such as…"]

Response: Thank you for your suggestions. The original phrasing conflated instruments with observational techniques. We have revised the sentence to clarify that ABLH is diagnosed from measurements obtained by various instruments. The revised sentence now reads: "ABL height (ABLH) can be diagnosed from measurements obtained by a variety of instruments, including lidar (Jordan et al., 2010), wind-profiling radar (Beyrich and Leps, 2012), infrared hyperspectral radiometers (Ye et al., 2022), microwave radiometers (Cimini et al., 2013), and radiosondes (Peng et al., 2023; Jozef et al., 2023a; Chan and Wood, 2013)."

Response to Comment #8

Reviewer Comment: [Line 74: "…a few tens of metres to less than 100 m". Here 100 m is meant as an upper limit, so "…a few tens of metres to of the order of 100 m". Current phrasing is unclear since a few tens of m is definitely less than 100 m, so both lower and upper limits of range are sort of the same, and upper limit vague.]

Response: Thank you for your suggestions. The original expression "a few tens of metres to less than 100 m" was redundant, as "a few tens of metres" is already less than 100 m, and the upper limit was vague. We have revised the sentence to state the range explicitly: "Because the Arctic stable ABL is often extremely shallow (frequently less than 100 m and sometimes as low as 20 m), these single-instrument approaches may be insufficient to fully characterize near-surface ABLH (Mahrt, 2014; Tjernström et al., 2012; Illingworth et al., 2019)."

Response to Comment #9

Reviewer Comment: [Line 76–77: the sensitivity of infra-red radiometers to the structure above clouds is generally none rather than just 'substantially reduced"; once liquid water path is above the black-body limit for IR (40-50 g m-2) the cloud is opaque.]

Response: Thank you for your suggestions. We have revised the sentence to state that optically thick clouds make the thermodynamic structure above cloud top effectively inaccessible to ground-based infrared radiometers; the revision does not merely reduce the stated sensitivity. The revised text now reads: "Infrared hyperspectral radiometers primarily measure atmospheric thermal-infrared radiance. Clouds can strongly attenuate this radiation, and optically thick clouds with high liquid water content can make the thermodynamic structure above cloud top effectively inaccessible to ground-based infrared radiometers. In addition, large uncertainties remain in temperature and humidity profile retrievals under the strong stable stratification and near-surface inversions common in the Arctic (Devasthale et al., 2011; Turner and Löhnert, 2014; Turner and Blumberg, 2019)."

Response to Comment #10

Reviewer Comment: [Line 114: a reference for the ARM INTERPSONDE product would be useful here.]

Response: Thank you for your suggestions. We have added the reference in Sect. 2, where INTERPSONDE is introduced. The sentence now reads: "Because the routine sampling cannot provide a continuous record, the Atmospheric Radiation Measurement (ARM) Interpolated Sonde (INTERPSONDE) product (https://doi.org/10.5439/1095316; Jensen et al., 2021) is used in addition." The reference list includes: Jensen, M., Giangrande, S., Fairless, T., and Zhou, A.:

Atmospheric Radiation Measurement (ARM) User Facility: Interpolated Sonde (INTERPOLATEDSONDE), ARM Mobile Facility (MOS) MOSAiC, ARM Data Center [data set], https://doi.org/10.5439/1095316, 2021.

Response to Comment #11

Reviewer Comment: [FIGURE 1: the ship track shown on figure 1 is a little misleading since it makes no distinction between the ice drift and the transits (where there are very limited measurements). See Shupe et al. (2022)(fig 2) for a better approach, distinguishing between ice drift and transits. It would also be useful to show the ice extent.]

Response: Thank you for your suggestions. We have revised Figure 1 to distinguish the ice-drift and transit periods: circular symbols denote ice-drift phases, triangular symbols denote transit periods, and the blue and black markers denote the start and end of the expedition. The caption has been updated accordingly. We did not add a sea-ice extent contour, because at the scale of the figure it would overlap the drift track and obscure the drift/transit distinction.

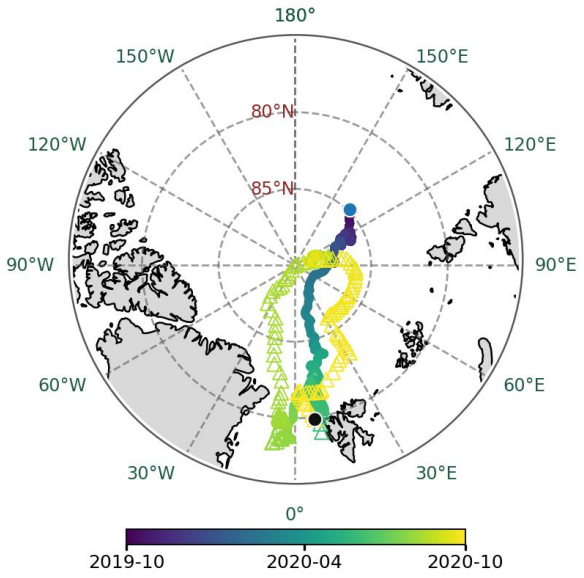


Figure 1. Map of the central Arctic showing the ship track during MOSAiC, color-coded by date. Circular symbols indicate ice-drift phases, and triangular symbols indicate transit periods. Blue and black dots mark the start and end points, respectively.

Response to Comment #12

Reviewer Comment: [Line 127: '…unified time and height grid'. The time resolution (1 minute) is given, but there is no information on the vertical resolution given, please provide this.]

Response: Thank you for your suggestions. We have added the vertical resolution in Sect. 2. The revised sentence reads: "It interpolates the MOSAiC soundings to a uniform grid with 1 min temporal and 20 m vertical resolution; the analysis uses the 10–1990 m range."

Response to Comment #13

Reviewer Comment: [Line 130: '…the Richardson number method.' On its own this doesn't provide the reader with enough information to know exactly what was done, there are

multiple possible approaches. Is this a gradient Richardson number? A bulk Richardson number? What is vertical resolution for former and lower/upper limits for latter. If following a previous study then give a reference.]

Response: Thank you for your suggestions. We have clarified that the reference ABLH is diagnosed with the bulk Richardson number method following Peng et al. (2023). The method is described in Sect. 3.2, including the critical value Ri_c = 0.35, the 30 m reference level and the persistence requirement of two consecutive levels on the 20 m grid; the INTERPSONDE profiles used for the diagnosis span 10–1990 m. The sentence in Sect. 2 now reads: "The temperature, humidity, pressure and wind profiles of INTERPSONDE are used to diagnose the reference ABLH with the bulk Richardson number method introduced in Sect. 3.2."

Response to Comment #14

Reviewer Comment: [Table 1. This might be better moved to an appendix of supplementary information. For most readers the precise details of the channels used will be of little relevance.

Also, in all the column headings 'wavenumber' has a typo and misspelled 'wavenymber']

Response: Thank you for your suggestions. The table has been removed from the main text and now appears as Table S1 in the Supplement, with the column headings corrected. The main text points to Table S1 where the channel set is introduced.

Response to Comment #15

Reviewer Comment: [Line 171: again for the INTERPOSONDE product we're told that the winds are provided at '100 vertical levels' but not what the vertical resolution is, or what vertical range those 100 levels span: is that all the levels, or just a low-atmosphere subset.]

Response: Thank you for your suggestions. The revised text now specifies the vertical spacing and the range of the 100 levels used. It states that the input includes "the INTERPSONDE wind profile on 100 levels with 20 m vertical spacing", and that "the analysis uses the 10–1990 m range." These are the lowest 100 levels of the product, starting at 10 m and extending to 1990 m, which is the part relevant to ABLH diagnosis.

Response to Comment #16

Reviewer Comment: ["Line 181–184: the list of features used to enhance the model's ability is rather general, e.g. 'characteristic height information…from backscattered profiles', but insufficient for a reader to repeat the analysis. While a detailed explanation of all of these features might be too long to include in the main text, it is relevant. I would encourage its provision as supplementary information."]

Response: Thank you for your suggestions. The description in the main text was too general for the analysis to be reproduced. We have added a detailed description of all input features, including the characteristic heights extracted from the backscatter profiles, together with the training settings, in Sect. S1 of the Supplement. The main text now states: "The complete set of derived features and the training settings is provided in Sect. S1 of the Supplement."

Response to Comment #17

Reviewer Comment: [Line 197: 'zs': the s should be subscript.]

Response: Thank you for your suggestions. The notation has been corrected to z_s, with s as a subscript.

Response to Comment #18

Reviewer Comment: ["Figure 2.: the figure shows several example profiles of temperatures, wind and humidity from radiosonde profiles with the height of the ABL top, derived from the Richardson number, indicated. It would be really useful to the reader to also show the profile of Richardson number from which the ABL top is diagnosed."]

Response: Thank you for your suggestions. We have revised Figure 2 by adding the Richardson number profiles as solid orange lines overlaid on the specific-humidity panels (c), (f) and (i) for the three example cases, and the diagnosed ABLH is marked by a dashed horizontal gray line in each panel. The caption has been updated accordingly to describe the overlaid Richardson number profiles and the ABLH markers.

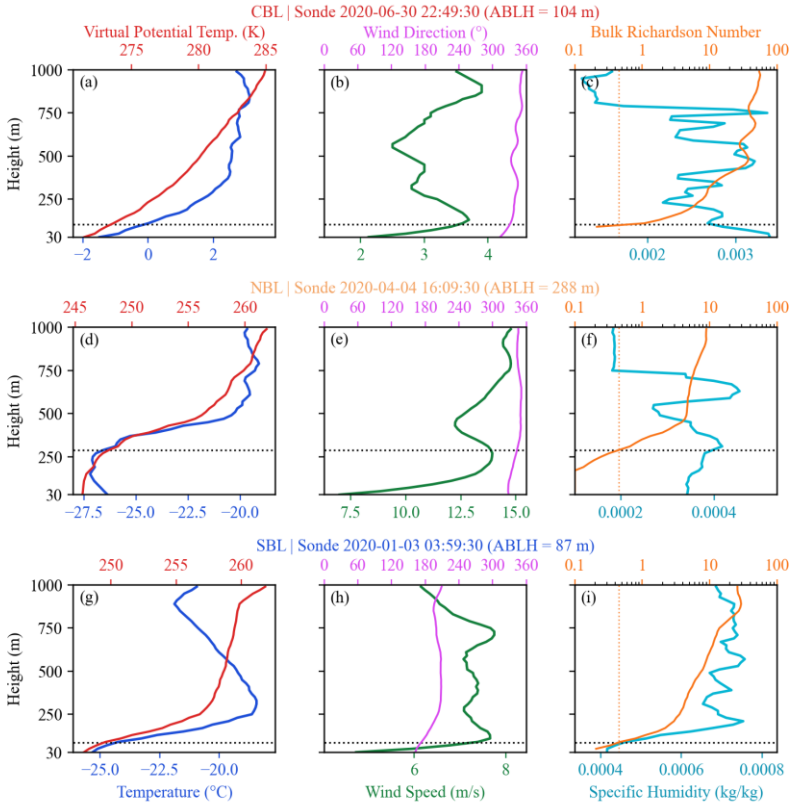


Figure 2. Panels (a), (d), and (g) show air temperature (blue) and potential temperature profiles (red); panels (b), (e), and (h) show wind speed (green) and wind direction (pink); and panels (c), (f), and (i) show specific humidity (solid lines) with the corresponding Richardson number profiles overlaid (solid orange lines). Panels (a–c), (d–f), and (g–i) correspond to 22:49 UTC on 30 June 2020, 16:09 UTC on 4 April 2020, and 03:59 UTC on 3 January 2020, respectively. The boundary layers at these three times represent a convective boundary layer (CBL), a near-neutral boundary layer (NBL), and a stable boundary layer (SBL), respectively. The dashed horizontal gray lines indicate the atmospheric boundary layer height estimated from the Richardson number method shown in each panel.

Response to Comment #19

Reviewer Comment: "Line 268: it is not clear from the text whether the RMSE's given here refer to the test data set of the validation data set. It can be inferred they apply to the test data only after reading further, but I found it easy to confuse the two when looking back for values referred to in later discussion. Make this explicit."

Response: Thank you for your suggestions. We have made the distinction explicit by naming the validation and test sets in the same sentence. The revised sentence reads: "In the validation set, the RMSE is 12.0 m under clear skies, 17.8 m under cloudy conditions and 15.0 m in fog/haze; in the test set (834 launch times), the corresponding values are 15.3, 44.7 and 25.1 m, respectively, and R^2 exceeds 0.9 in all three categories (Figure 5). Differences in sampling, temporal dependence and reference-label uncertainty may contribute to the validation–test gap."

Response to Comment #20

Reviewer Comment: ["Line 287–290: the logic of the statement ... seems flawed: it implies the importance of the information from the wind profiles is demonstrated by a reduction in the accuracy of the retrieval. I agree the information from the wind profiles is important, but it should be unaffected by low visibility. I think what is meant here is that the importance of this information is demonstrated by the fact that while the retrieval performance is degraded, it is still reasonable, even where important information on thermodynamic structure (e.g. IR spectrometer profiles in low visibility conditions) is lost, and the algorithm relies much more on the winds. Consider rephrasing."]

Response: Thank you for your suggestions. The original logic was flawed: degraded performance in low visibility does not by itself demonstrate the importance of wind information. The revised text now separates the two points. It attributes the wind contribution to mechanical mixing and vertical shear, while noting that part of the contribution reflects the shared wind input used in the Ri labels. It also reports AERI attribution and ablation separately. The revised text states: "Wind-related features have the largest SHAP contributions, and removing the wind input produces the greatest increase in estimation error. This dependence reflects both the information carried by the wind profiles and their shared use in constructing the Ri-based reference labels; the two contributions cannot be separated by these analyses."

Response to Comment #21

Reviewer Comment: ["Figure 7: it would be interesting here to also indicate the height of the ABL top derived directly from the Richardson number profiles, maybe just at the times of radiosondes. Y-axis label for panel (d) 'PS' should be 'Ps'."]

Response: Thank you for your suggestions. We have corrected the Y-axis label in Figure 8(d) from "PS" to "P_s". Black dots in Figure 8(a) mark the ABLH derived from the Richardson number method at radiosonde launch times, and the figure caption has been updated to describe these markers.

Response to Comment #22

Reviewer Comment: ["Line 367: 'In spring and summer, higher near-surface temperature generally corresponds to higher ABLH…'. This should make a distinction here between a higher (seasonal) mean temperature, and a higher temperature within the diurnal cycle

(which is what is really referred to here). As stated it is ambiguous as to which is meant."]

Response: Thank you for your suggestions. The original sentence did not make clear whether the comparison concerned seasonal mean temperature or temperature within the diurnal cycle. The revised text now refers explicitly to the diurnal cycles: "The diurnal variations in CMNN-estimated ABLH and near-surface air temperature correspond most clearly in spring, with ABLH reaching its maximum later than near-surface air temperature." For summer, the text states that melting sea ice limits surface warming and that the relationship between near-surface air temperature and ABLH is less pronounced.

Response to Comment #23

Reviewer Comment: [Line 369–371: "In autumn, the diurnal pattern becomes less regular, suggesting that the boundary layer gradually shifts from being thermally dominated to being jointly controlled by thermodynamic, dynamical, and radiative processes." This is not obvious to me. Why is The ABLH diurnal cycle here considered 'less regular' than that in spring? Both show significant variation over the day. Or do you mean a less obvious relationship to temperature?

I find the diurnal temperature record here a little odd: for all other seasons the cycle looks more or less continuous across midnight (the jump from 2300 to 0000), but here there is a large temperature jump across that 1-hour boundary…is that real?]

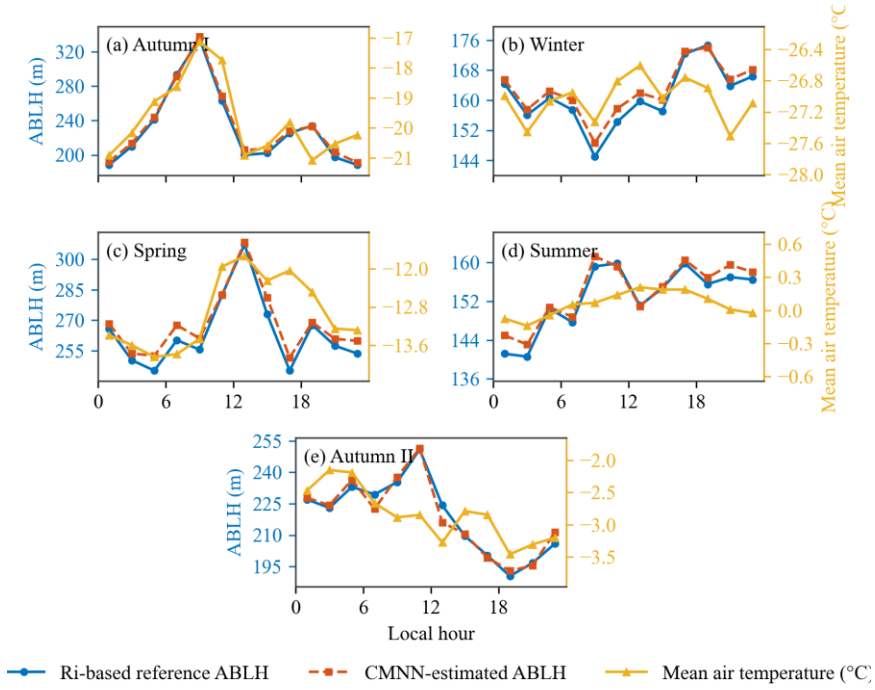
Response: Thank you for your suggestions. The revised text no longer describes the autumn cycle as "less regular". It states that near-surface air temperature alone provides only a partial explanation for the diurnal variation in ABLH. The temperature jump at the 23:00–00:00 boundary arose from combining the two autumn periods; Figure 10 now separates autumn I (2019) and autumn II (2020), so the discontinuity is removed. The revised sentence reads: "Near-surface air temperature alone provides only a partial explanation for the diurnal variation in ABLH."

Response to Comment #24

Reviewer Comment: [Figure 4: How are the seasons defined here? The simple, and widely used 3-month groupings June/July/August = Summer, Sept/October/Nov = Autumn, don't map neatly to the changes in surface energy budget in the Arctic. Shupe et al. (2022) define seasonal periods (and transitions) via a consideration of the energy budget and meteorological conditions; this has seasons of unequal lengths which might then impact statistical robustness (OK, but needs acknowledging if similar here).

And a sanity check since it is not stated. These diurnal cycles are based on local times not the UTC time given in all the archived data files…right?]

Response: Thank you for your suggestions. The simple three-month grouping does not map neatly onto the Arctic surface energy budget. Figure 10 now uses five seasonal periods following the energy-budget-based definition of Shupe et al. (2022): autumn I (2019), winter, spring, summer and autumn II (2020), and the two autumn periods are separated. The diurnal cycles are computed in local time. Because the periods differ in length and sampling, they are compared qualitatively and are not treated as formal seasonal contrasts. The revised text and caption state this.



Response to Comment #25

Reviewer Comment: ["Line 381: you divide the results into 3 subsets, 'coupled, CBH-above-ABLH, and 'ABLH-above-CBH'. It is not clear what the definition of 'coupled' is here."]

Response: Thank you for your suggestions. We have defined "coupled" explicitly in Sect. 4.3: "coupled (cloud base within 100 m above the ABL top)". The other two categories are "CBH-above-ABLH (cloud base more than 100 m above the ABL top)" and "ABLH-above-CBH (cloud base below the ABL top)". The 100 m threshold is adapted from the 200 m criterion of Zhang et al. (2024) to account for the shallower Arctic ABL.

Response to Comment #26

Reviewer Comment: [Figure 11: this is an interesting and informative analysis, but I wonder what the rationale is for choosing latent heat flux as one of the independent variables that may influence ABLH rather than (or as well as) sensible heat flux. Latent heat flux is almost always small in the Arctic, and has a small influence on convective forcing than the sensible heat flux. Where convection is a potential influence, perhaps the buoyancy flux (from the virtual potential temperature flux) is more relevant, including the effects of both sensible and latent heat.]

Response: Thank you for your suggestions. Sensible and buoyancy fluxes are also relevant to convective forcing. We therefore examined latent heat flux, sensible heat flux and buoyancy flux together, and all coefficients are listed in Table S3. Among the three, latent heat flux has the highest correlation with ABLH (r = 0.398), compared with 0.301 for sensible heat flux and 0.296 for buoyancy flux, which is why it is shown in Figure 12(f). Its association with ABLH may partly reflect covariation with wind speed and the surface–air humidity difference, and it is not interpreted as an independent effect of moisture exchange.

Response to Comment #27

Reviewer Comment: [Line 417: 'indicating that the ABL deepens when the surface is warmer than the air above it rather than when the surface is warm in absolute terms': this statement relates to the relationship with Ts-T2, so is misleading as given, implying a relationship with absolute surface temperature rather than near-surface temperature gradient.]

Response: Thank you for your suggestions. The original statement was misleading. We have removed the implication that absolute surface temperature controls ABLH and now state that the positive correlation applies to the surface–air temperature difference, Ts − T2m. The revised sentence reads: "Downward longwave radiation and the surface–air temperature difference, Ts − T2m, show positive correlations with the CMNN-estimated ABLH (r = 0.409 and 0.372, respectively)."

Response to Comment #28

Reviewer Comment: [Line 418: "Latent heat flux shows a moderate correlation with ABLH (r = 0.38), which may reflect the promoting effect of enhanced surface-atmosphere coupling under moist or cloudy conditions on boundary-layer deepening": this seems unlikely. Under moist conditions (and in the presence of low cloud the near-surface will be relatively moist) we would expect a smaller latent heat flux. A higher flux would be expected in low humidity conditions.]

Response: Thank you for your suggestions. The original interpretation was physically incorrect. The revised text no longer interprets latent heat flux as promoting ABLH. It states that latent heat flux has a moderate positive correlation with ABLH (r = 0.398), compared with 0.301 for sensible heat flux and 0.296 for buoyancy flux, and that its association may partly reflect covariation with wind speed and the surface–air humidity difference.

Response to Comment #29

Reviewer Comment: [Line 423-428: the discussion of the relationship between low level jets and ABLH claims the LLJ's promote deeper boundary layers. The arguments are physically reasonable, and I think this probably is the case…but. The discussion lacks any recognition of potential confounding, or at least complicating, factors. Figure 12a, and the associated discussion, link jet core speed with ABLH. There is certainly a strong correlation, but no recognition that jet core speed is dependent on the background wind speed, and a strong background wind drives a higher wind stress, stronger surface mixing, and a deeper ABL. Figure 12a is more convincing, but there remains the possibility that jets might be more common for deeper boundary layers, biasing the statistics.

Overall, I think LLJs probably do contribute to maintaining deeper boundary layers, but the discussion needs more nuance.]

Response: Thank you for your suggestions. The original discussion was overly absolute and did not account for potential confounding. We therefore use the minimum wind speed below each core as the lower-layer reference. Both core speed and ABLH are regressed on this below-core minimum, and the residuals correlate at 0.505 (Table S4), so core speed remains conditionally associated with ABLH. This minimum correlates with core speed (r = 0.788) and with ABLH (r = 0.534), and contributes to the jet amplitude.

Response to Comment #30

Reviewer Comment: "Line 437: the discussion here refers to the no-microwave-band model, but it's not stated why this is used…is it simply that no microwave radiometer is available for this data set?"

Response: Thank you for your suggestions. Yes. NSA does not operate a 183 GHz radiometer, so the microwave channels available during MOSAiC have no counterpart there, and the no-microwave configuration is used for the transfer experiment. The revised text states: "Because NSA and MOSAiC differ in underlying surface, climate background and ABL structure, and because NSA does not operate a 183 GHz radiometer, the microwave channels available during MOSAiC have no counterpart there; the no-microwave configuration is therefore used."

Response to Comment #31

Reviewer Comment: "Line 444–445: 'As shown in figure 13…'. You mean figure 4, the values quoted are for the MOSAiC validation set in figure 4."

Response: Thank you for your suggestions. The incorrect "Figure 13" citation has been removed. The values n = 834, RMSE = 35.67 m and $R^2 = 0.95$ refer to the no-microwave configuration on the MOSAiC launch-time evaluation and are reported directly in the text. Figure 5 shows the all-input reference model, not the no-microwave configuration, and Figure 14 shows the NSA transfer. No figure is cited for the MOSAiC no-microwave values.