Dear Reviewer #2,
We are very thankful for the positive feedback and the valuable comments.
We revised our manuscript accordingly and tried to address your comments carefully. Please find our responses to your comments in blue below.

Thank you again and best regards,
Anja Rösel on behalf of the co-authors.

GC1: The study shines through an excellent set of observations. There are parts of the description which ask for improvement, though. One is an improved consideration / discussion of snow depth and sea-ice thickness variations between survey site (local), 5 km circle around Lance (extended in-situ survey) and 10 km circle around Lance (OIB); more details regarding this issue I give in the specific comments. The discussion about scaling and representativity issues is light and could be expanded; elements of this topic I try to express with the following questions. Can one really, as done, combine the OIB data with the in-situ data with such high accuracy? Are EM31 total thickness measurements really that accurate that one can derive highly accurate sea-ice freeboard when combining this data with snow probe snow depth observations? How does the EM31 signal respond to a possibly spatially extended area of flooded sea ice?
We looked again at the data, matched the snow radar measurements with the closest SP data and re-calculated and re-made the Figures 5-8 and improved with it also the discussion about the scaling (local and regional). And yes, we think it's possible to connect OIB data (10s of meter range of the snow radar) with very high-resolution data like SP (cm-range) and a bit coarser EM data (m-range) to get knowledge of the accuracy - therefore we do in addition the inter-comparison with regional ground data (5 km scale) to get the averages of snow depth and not the single point data. Also, the comparison as a pdfs allows to compare different scaled data.

How does the issue that OIB is known to underestimate thin snow / has issues over highly deformed sea ice influence your results? Finally, the one (or maybe two) ice salinity profiles used at the end to conclude that your observations prove that saline snow can be the main cause for the observed biases, are not very well connected to
the rest of the manuscript - even though they are the dominant topic in the discussion and conclusion section. I recommend to give the description of these in situ observations more weight and demonstrate more clearly how well (hopefully) these single point measurements represent the conditions in the detailed survey area.
We rewrote and restructured the discussion and result section to get this more connected. We also moved the Figure of the salinity profile from the Appendix in the main manuscript.

GC2: My impression is that the MAIN issue is that OIB data result in a sea-ice freeboard bias of 0.2 m. This is a SUBSTANTIAL bias but it is not presented and discussed in an overly prominent way in the manuscript. I recommend to be more exhaustive in the discussion of particularly these results in the context of Figures 6 through 8.
We looked again at the data and have to reinterpret statements about a 0.2 m bias. Looking regionally we see a 0.06m difference between OIB radar and MP data (Figure 5b), although it is very difficult to connect these two datasets or identify any features which might explain the pretty good fit between them. The 2D field is the only location where we can reliably pin OIB snow radar to near-coincident in situ observations. Using only radar data inside the 2D field we see a much closer fit between in situ and radar snow depths - more realistically in line with expected brine wicking or flooding depths and also closer to regional values.

GC3: The secondary issue is the snow density of the proposed () two-layer snow setup which is discussed in the context of Figure 8, hi4. There are, to my opinion, at least two areas of improvement. The first one is superposing the drill-hole sea-ice freeboard data onto Figure 7 and discuss the results. The second one is to conduct a sensitivity study which plays around with possible snow layer depths and densities used in

Equation 6 to derive hi4. A pre-requisite for these actions is an appropriate introduction of what you cann the two-layer snow setup, which is not adequately described yet in the manuscript.

In this context, I also kindly ask for clarification of the wording "slushy basal snow layer" versus "snow-ice basal snow layer" because snow-ice is refrozen slush hence hard while slush is soft - which has implications for the snow probe measurements and the interpretation of the measurements.
We revised this in the manuscript to be clear. See also your specific comment for L. 357

Specific Comments:
Line 179++: I am not sure your choice of denoting different variables with 1, 2, 3, 4, using hi (I) and hs (S) and hfb (Fi) and hfbs (F) to name the variables, i.e. without subscripts, and introducing subscripts such as EM, SP or IS is an optimal solution. In any case I recommend to use it in a consistent way. This means that also in the running text hi and hs should be given as used in the formulas (see Lines 174/175 for instance). In addition, I am wondering whether it is necessary to abbreviate "in situ" with IS. To me this is confusing in the zoo of short names. But this is of course your choice. If you keep IS then you need to move its definition from Line 194 to Line 179.
We have removed numbers from all variables and made a naming scheme we think adds some clarity. In this new scheme, the subscript 'IS' is used to mean 'drill hole data', with all other sources using subscripts to denote the instruments used to collect data which values are derived from. We agree, there are many ways to name variables relevant to altimetry on sea ice, and welcome feedback which points out where clarity is needed.

Equation 2: I am sure that the freeboard in this equation needs to be the snow freeboard, i.e. hfbs.; this is the classical equation to derive sea-ice thickness I from total (sea ice + snow) freeboard F and snow depth S, isn't it?
The authors agree with the reviewer. We have corrected the equation to using the total (snow+ice) freeboard.

Then equation (3) would be the total freeboard as well and not the sea-ice freeboard. In order to end up with the sea-ice freeboard here I suggest to use the retrieval equation used for radar altimetry (in your notion): hi1IS = (hs1SP * rho_s + hfb1IS * rho_water) / (rho_water - rho_ice)
Solved

In Line 194 it needs to be "1" instead of "2" in the variable name.
Solved

Equation 4: In the context of this equation I'd like to note that you are not consistent with Figure 3. There you denote snow freeboard as hfbs while in Equation 4 and in Line 202 you write hsfb. I like hsfb more.
The authors agree with the reviewer. We have reformulated the equation notations.

Line 206/207: While the accuracy for the drill-hole measurements is clear from the stated measurement accuracy of the measurement device, I am wondering whether you might want to add one sentence about the way you estimated the value of 0.06 m for the combination of snow probe and EM31 measurements.
We added here: "As described in Rösel et al., 2018 the uncertainty of in situ freeboards hfbEM,SP and hsfbEM,SP resulting from the propagation of uncertainties in the snow and ice densities and the sampling uncertainty (represented by the spatial variability) is estimated to be on average ±0.06 m."

Equation (7): I suggest to split this equation into two. Please first introduce hB without the interpretation that it can be decomposed into sea-ice freeboard, basal snow-layer thickness and an error term and subsequently, in a follow-on equation (if necessary) show the decomposition of hB. I note, after having read

the entire manuscript, that there is limited reference to and usage of this equation. Or did I overlook a figure or table where you provide an estimate of the basal snow-layer thickness?
We revised the manuscript and removed Eq (7).

Figure 4: I am a bit confused with what I see here. I appears to me that this plot starts on the left somewhere close to the top right corner of the map shown in Figure 2, takingone of the overpasses shown towards the southwest (crossing the survey area). Fine. Then comes the aircraft turn; the main part of the radar-echogram shown in Figure 4 is related to sea-ice conditions "on ice floe south of the survey site". Now, what I have problems with is the fact that you stated the times for overpass #2 and #3 as 15:37 and 15:43 in the caption of Figure 4 while in the heading line just above the figure it specifies a time range around 15:43, suggesting that this is only overpass #3. Could it hence be that we only look at overpass #3 and that only the small part (left quarter of the echogram) is coincident with the tracks (actually only overpass #3) denoted in Figure 2? If this is the case, I am wondering whether it wouldn't make sense to expand that left quarter of the echogram because this has a direct relation to the survey area.
This is solved; it is only overpass #3 and the survey area is now indicated by a red box

Lines 253-255: "In addition ..." How important is it to explicitly mention the mean of these very few (compared to the other samples) snow depth observations at the drill hole sites in the text? Would it be sufficient to only show this value in the Table? I am suggesting so because you are coming up with quite a number of different snow depth values and it begins to become confusing. Particularly, since you repeat this information in Line 273.
The authors agree with the reviewer. We have removed the sentence and directed the readers to Table 2 in the appendix and Figure 5.

Line 263: What is the motivation to draw a 10 km circle around R/V Lance for the air-borne data when the in-situ observations along the five transects were carried out within a 5 km radius circle?
We chose the 10 km radius for the snow radar data to show the regionality, the in situ transects were only in a radius of 5 km around the ship. Internal data comparison of the snow radar data data do not show significant differences if we choose a 5 km or 10 km radius.

Line 275/276: "Three ... before drilling." –> How do you know?
To avoid confusion, we have rephrased the sentence to 'Three out of the ten drill-holes were found to be flooded.

Figure 5: In the left panel you denote hs1 and hs3 with the respective methods while in the right panel you write "regional". My suggestion would be to be consistent. At this point the reader possibly knows that hs1 is based on snow probe data and that hs3 is based on the snow radar. Hence you could use "survey area" in the left panel and
keep "regional" in the right panel.
Thanks for the suggestion, we revised the figure and changed the labeling accordingly.

Lines 294-298: These sentences about brine wicking etc. fall a bit from heaven. I suggest to start a new paragraph here and motivate these further considerations by again stressing that near-zero ice-freeboard supports flooding of the ice-snow interface and subsequent upward wicking of seawater and/or brine into the basal snow cover. The paragraph about the c-profile of the salinity given at the beginning of Section 3 could be placed here in a much more logical way - as this piece of information seems a bit lost where it is located currently.
Thanks for the comment. We have reorganized the introduction of snow salinity caused from flooding and brine wicking in a separate paragraph. We have also moved the salinity c-profile to the discussion section.

Figure 7: I suggest to overplot the freeboard values shown in Table 2 of the appendix onto the maps in each panel by using, e.g., a color-filled circle, and discuss what you see.
Figure 7 was re-done and might be clearer now.

Lines 303-313 / Figure 8:
- Line 305: I don't find hi_REGIONAL in Figure 8.
This now changed to hiATM,SR(all) (also in the updated Figure)

- Line 307: "slightly thinner" –> It might be a matter of taste but a difference of 0.3 to 0.4 m in a thickness range between 1.1 and 1.5 m I would not call "slightly".
We have removed 'slightly' in the revised manuscript.

In addition, those 5 regional survey lines, were these laid out without checking the ice conditions beforehand?
No, this was done afterwards. They were a combination of first year ice and second year ice. We tried to stick to a triangular shape, to cover randomly everything without pre-choosing the ice type.
What I mean by this is, that based on the large sea-ice thickness standard deviation it seems likely that these 5 lines had a substantial fraction of thin ice from refrozen leads while the bulk of the sea ice inbetween might have had a similar thickness than the ice of the survey site. Please check and, if need be, re-phrase statements.

- I am missing commenting on hi3.
Thanks for the comment. We considered this.

- Line 309-313: I suggest to re-write this part. hi4, if computed using Eq. 6, uses the sea-ice freeboard (hfb) which is derived from airborne ATM total freeboard (accurate) and the in-situ snow depth (accurate as well); in addition, for the snow part it uses (again) the in-situ snow depth. The densities you used are those which you measured in the field = accurate. Hence, at first glance it is not clear why even with the densities you measured the hi4 is biased compared to hi1 and hi2. The statement that the bias in sea-ice thickness between hi1 and hi4 of about 0.4 m is consistent with the bias in hfb of 0.03 m and that this is in part due to the not sufficiently well considered "two-layer snow set up" is not backed-up yet by your writing or the figures. If, as you suggest, snow densities across this two-layer snow set up are the main cause, then I strongly suggest to play around with potential snow densities (you measured some in the field, didn't you?) and layer thicknesses to figure out whether your hypothesis is true. In other words: Which snow density values of the proposed two-layer snow setup explain
the observed difference between hi1 and hi4? Are these realistic?
Good point. We recomputed ice thicknesses for figure 8 and found a much closer match between datasets.

Discussion section:
General comments: I don't see a reason why to write "total snow depth" or "total snow thickness". I guess "total" can be omitted.
Thanks. "Total" omitted before snow. Also, throughout the manuscript, we have replaced snow thickness by snow depth.

Line 345: This is a reminder that it might make sense to provide a scientific motivation for using a 10 km radius for the airborne data versus a 5 km radius for the in situ data. If by chance the airborne data observed a comparably large fraction of thin ice, then the airborne snow depth would naturally be smaller and hence its bias to the in situ snow depth.

As mentioned in our reply to a comment above we did some analysis of a 10km vs a 5 km radius of the airborne data which do not show significant differences.

Line 357: "slushy, snow-ice formation in the basal layers of the snow pack" vs. Line 367: "formation of highly saline and saturated slush in the basal snow layers" –> It is a difference whether one speaks of wet / saturated slushy snow ... which is a rather soft material, or whether one speaks of snow-ice which - at least to my understanding - is refrozen ... and hence hard. It is, to my opinion, important to distinguish between those because I'd think that the SP measurements would penetrate slush and hence INCLUDE the thickness of the slush layer into the snow depth reading while these would not penetrate snow ice and exclude that part from the snow depth reading. Please be clear what you mean and observe to avoid misunderstandings.
As already mentioned above, we clarified and revised this already.

Line 363: "the 1-m thick FYI floe" –> In lines 241-244 you already provided some information about these observations. I note that these are not coherent and not specific enough. What did you mean by "in the vicinity of the 2D site" in those lines? There you gave one date (March 5), here you give two dates. The thickness of that floe (is it representative?) is just 1 m, i.e. substantially less than the surveyed floe with 1.4-1.5 m thickness. I suggest to comment in your manuscript about these differences and slight misfit in information. I also interpret from your writing that snow depth (on that single floe) increased by 0.08 m between March 5 and March 23 and that that snow depth is considerably smaller than the average or modal snow depth of the 2D survey site. Comments?
Thanks for the comment. The ice salinity measurements are from March 5. We have removed March 23 in the sentence, in the revised manuscript.

Lines 379-382: I have a stupid question in this context: How do you compute the seaice thickness from estimates of sea-ice freeboard and snow depth using the classical equation (e.g. equation 6) when the sea-ice freeboard is negative? Then the first term in equation 6 is negative. Is the retrieval using that equation defined at all?
No, the second term would always be positive (and with a higher number since it uses snow depth). For the ice thickness calculations, we used eqn 2. Eqn 6 was not used and removed from the manuscript to avoid confusion.

Line 383++: "Our study shows that saline snow conditions can ..." –> I suggest to formulate more clearly how the conditions met in this study differ from those encountered on Canadian Arctic fast ice. There the snow cover was dry, the sea-ice freeboard positive and the brine concentration in the basal snow layers solely caused by the high
sea-ice salinity near the ice-snow interface. Here, during N-ICE2015, the situation appears to have been completely different, with a substantial amount of negative sea-ice freeboard, hence flooding of the ice-snow interface and an (unknown?) amount of slush at the ice-snow interface from which large amounts of brine can be wicked up (how high?) into the overlying snow.
Thanks for the comment. We have rephrased the previous section (before Line 383) differentiating between landfast FYI (with positive freeboard) and ice with negative freeboards. There were no direct measurements of how much slush there was nor the slush salinity. Nandan et al. (2020) (Figure 4b) reported high salinities towards the basal snow layers (up to 25 ppt), sampled during the N-ICE 2015 campaign. However, we cannot confirm if these high basal salinities are caused by brine wicking from the underlying slush layers.

Line 407: Please explain why there are two different snow thickness values (here in the text and also in Fig. 9 a)
The two different values result from the Warren climatology used for the Kurtz algorithm. We rephrased as follows: 'Modelled snow depths of 0.15 m and 0.37 m (derived from Warren et al., 1999; Kurtz et al., 2014, Figure 9a) are underestimated when compared to the observed in situ snow depth, which averaged 0.55 m.

Line 408: I suggest to add a statement about the fact that the survey cite is located in an area where the shown CS-2 products (and also the snow depth) show large spatial variability. In light of the fact that sea ice is not static but drifts, your "verdict" about the quality of the CS-2 product could perhaps formulated in a less harsh way. I note in this context, that you completely ignored the CPOM results published by Tilling and co-workers, and issue which I kindly ask you to amend in your manuscript.
Thank you for the comment. We have rephrased the paragraph with more constructive criticism for the CS-2 product quality and have also referred to Tilling et al. (2018) where their work showcases the impact of sea ice drift affecting the quality of CS-2 products.
"However, presently operational CryoSat-2 retracker algorithms or empirical models (e.g. Hendricks et al., 2010; Ricker et al., 2014; Kurtz et al., 2014) do not account for snow pack flooding as a source of error, affecting the accuracy of sea ice freeboard and thickness estimates. Moreover, since our survey site was also drifting, we acknowledge the impact of sea ice dynamics also affecting the correlations between in situ measurements and satellite-derived estimates, both acquired at different times (Tilling et al., 2018)."

Typos / editoral remarks:

Line 24: "wicking and saturation into" –> "wicking into and saturation of". Later in the sentence: Would it be sufficient to write "causing the airborne radar signal ..."? I would read more fluently. If "more diffuse scattering and influenced" shall be kept then I suggest to split the sentence into two.
Thanks for the comment. The sentence is split into two as follows: These conditions caused brine wicking into and saturation of the basal snow layers. This causes the airborne radar signal to undergo more diffuse scattering, which results in the location of the radar main scattering horizon to be detected well above the snow/sea ice interface. This leads to a subsequent underestimation of total snow depth, if only radar-based information is used.

Line 33: "it may result ..." –> I am not sure I understand what this refers to.
We have revised the sentence as follows: "Our results suggest that sea water flooding of the snow/ice interface leads to underestimations in snow depth or overestimations of sea ice freeboard, measured from radar altimetry, in turn impacting the accuracy of sea ice thickness."

Line 43: You could add a sentence stating the importance of the thickness of the snow layer on sea ice on the in-ice and under-ice biological processes.
Thanks for the comment. We have added a sentence in the revised manuscript as follows: "In addition, snow cover controls the amount of transmittance of photosynthetically active radiation affecting the productivity of primary algae and phytoplanktons (Mundy et al., 2007)."

Line 73: I suggest to cite the work of Willat et al. (2010) here, also from the Antarctic but a different region: Digital Object Identifier 10.1109/TGRS.2009.2028237
Willatt et al., 2009 added to the citation.
Willatt, R. C., Giles, K. A., Laxon, S. W., Stone-Drake, L., & Worby, A. P. (2009). Field investigations of Ku-band radar penetration into snow cover on Antarctic sea ice. IEEE Transactions on Geoscience and remote sensing, 48(1), 365-372.

Line 85: Perhaps switch "Atlantic Sector of the Arctic Ocean" and "Southern Ocean" since your primary focus is in the Arctic Ocean?
Changed.

Lines 112/113: "... accuracy ... higher ... " –> I know what you mean but a reader might stumble at first glance being surprized that the EM31 accuracy is "better" for rough and deformed ice - which, I guess, quite some readers automatically imply into "higher". Perhaps it might make sense to write "worse"?
We have deleted the accuracy phrase for rough and deformed ice to avoid confusion.

Lines 122/123: "We use the results of the independent snow transects from Floe 2 to provide the regional context ..." –> I am a bit confused. You have that 400 m by 60 m survey area on the floe. And then you have those additional (>5000) measurements within 5 km of the ship. These are - at least this is my assumption - not necessarily on floe 2. But these are the measurements which provide the regional context, am I right?
Thanks for the comment. We have rephrased the sentence in the revised manuscript, as follows: "We use the 2-D grid snow depth measurements and those sampled via transects within a 5 km radius, to provide a spatial representativeness and context from local- to regional-scales".

Figure 2: What is "WAV snow depth"?
WAV refers to the snow radar-derived snow depths using the NOAA Wavelet technique (Newman et al., 2014). We have added a sentence in the figure caption on this.

Line 153: "During the survey ..." –> Does this refer to the OIB survey or to the groundbased survey?
We changed it to 'During the OIB survey'

Line 161: "..., and the ..." –> delete "and"
done

Line 171: Any measurements of the density of second-year ice?
We do not have any confirmed and accurate sea ice density measurements for second year ice from N-ICE.

Line 178: "for flooding" –> "to flooding"
done

Line 190: "can calculated" –> "can be calculated"
done

Line 250: "second mode" –> "mode"
done

Line 254: "0.08 m than" –> "0.08 m larger than"
Done

Lines 263-265: "with the ... of the ship" is kind of a repetition of the end of the previous paragraph. I suggest to simply refer to the above-mentioned transects.
We have deleted the repetition of the shup usage and referred them to the transect measurements.

Line 273: "(FB2)" can be deleted, I guess.
Deleted.

Line 287: "of+-0.06 m" –> two spaces are missing
Space added.

Line 290: "Figure 5" –> "Figure 6" "lie in the negative range, to -0.1 m." –> "are negative with magnitudes up to 0.1 m."
Rephrased.

Line 291: "Results in the same range, with ... 0.09 m, are obtained ..." –> "Results in the same range are obtained ... elevation, resulting in an average value of hfb4 ... 0.09 m (FB4, see Figure 6)."
We have rephrased the sentence in the revised manuscript as suggested.