

The paper "Estimating snow depth over Arctic sea ice from calibrated dual-frequency radar freeboards" by Lawrence et al. deals with estimating snow depth by combining satellite-based measurements of snow and ice freeboard. The method requires prior calibration with independent freeboard measurements. Here, CryoSat-2 and AltiKa satellite freeboard measurements are calibrated with airborne Operation IceBridge (OIB) measurements.

The latter raises one of my main concerns: The method, as presented here, relies on having reliable independent freeboard data, which at the moment is only provided by OIB data. However, there is disagreement within the science community on how to interpret the OIB radar measurements, i.e. different retrieval algorithms differ in the way air-snow and snow-ice interfaces are detected and localized. A recent paper by Kwok et al. (2017) showed that this caused OIB snow depths as retrieved from different groups to differ on average by up to 7 cm (for first-year ice) and 12 cm (for multi-year ice) for the 2013-2015 data (see Fig. 7 & 8 in Kwok et al., 2017), which is used in the paper presented here. The variability of snow depths is also quite different (so it is not just a constant bias between the different products). Though this problem is briefly mentioned in the paper presented here, this is only done rather late (on p. 14, l. 14), which does not represent how severe this issue is for the proposed method of retrieving snow depth. I think that this should be mentioned and discussed far earlier and with more emphasis because it has major implications on the usability and accuracy of the proposed retrieval method! Ideally, the authors would perform their comparison not only for OIB quicklook data, but also for (at least) one of the other OIB-based freeboard retrieval data sets to estimate how much this can influence the results.

A further concern is that the study of Guerreiro et al. (2016) also uses CryoSat-2 and AltiKa freeboard measurements to retrieve snow depth. Instead of calibrating these Ku and Ka-band measurements with independent data (as done here), they theoretically analyze the penetration depths of both radar altimeters in snow and use snow density estimates to modify the Ku-band radar signal's velocity through the snow. In their study, they compare their retrieved snow depths with OIB snow depths for the same years as in the study presented here (2013-2015). They seem to have somewhat lower RMSDs (4.1...5.4 cm) as compared to the results presented here (4.9...6.7 cm), although their results are independent of OIB measurements, while the results here are not. Why are these results not compared here? Is there any advantage of using the method presented here as compared to the one used in Guerreiro et al. (2016)? This comparison and discussion is missing here!

I found it confusing that the authors first declare that radar altimetry penetrates through to the snow-ice-interface, while laser altimetry does not (p. 2). AltiKa is presented as a radar altimeter (thus suggesting that it penetrates through the snow), but it is later compared with OIB's ATM laser freeboard (section 3.5). From what I understood, Guerreiro et al. (2016) conclude that the radar signal from AltiKa does not penetrate the snow, while Armitage and Ridout (2015) concluded that the AltiKa signal is scattered from roughly the midpoint of the snow layer. This discrepancy is not even mentioned here. What do your results suggest? Please comment/discuss/specify.

Another issue is that I think the structure of the paper could be improved:

a) In an "Introduction" I would mainly expect to read about the importance of the presented study, how it fits into the context of already existing studies and what is the new contribution of the presented study. Instead, we here get a general introduction on the importance of snow (ok) and we are presented the equations used to convert ice/ snow freeboard to snow depth (more appropriate for the "Data and Methods" section?). This is followed by a chapter that lists existing snow depth products, where I would prefer to read more about the differences to the presented study and the implications these have instead of a list of methods.

b) The "Results" section contains a lot of what I would consider discussion (or speculation as some of the statements on p. 11 are not based on citations), while the "Discussion" section on p. 14, l. 20 starts with showing more results...

Otherwise, the manuscript is, in general, well written and I was able to follow the method.

Specific comments:

p. 1, l. 3: "...*can be applied to any coincident freeboard measurements*" -> to any coincident snow and ice freeboard measurements? (would be clearer)

p. 1, l. 19: "...*snow depth estimates could be usefully assimilated...*" -> "usefully" is a vague (and strange) expression here...

p. 1, l. 23-24: "*The implications ... is*" -> The implications ... are

p. 2, l. 4: Eq. (1) -> Is this formula from Beaven et al., 1995?

p. 2, l. 27-28: "*The granular nature of snow acts to scatter and dissipate microwave energy radiating from the Earth's surface, reducing the surface brightness temperature.*" -> This statement is only true for part of the frequency spectrum of microwaves! Not true for low microwave frequencies.

p. 2, l. 30: "*for a given frequency*" -> Too vague, I'd prefer to see the frequency (range) that you mean here.

p. 3, l. 30-31: "*AltiKa was designated with a maximum penetration depth of 0, i.e. no penetration, and CS-2 a maximum penetration of 1, i.e. full snow penetration...*" -> What does this mean? Is it possible to retrieve snow depth using this method? Could you compare these with your method?

p. 4, l. 8-14: You write about the issues raised by different satellite footprint sizes, please also give the CS-2 footprint size here to make the comparison easier.

p. 4, l. 28: "*retrieves surface elevations up to 81.5°*" -> a) Please add "latitude" (to avoid confusion with "geometrical elevations", which can also be given in degrees). b) I think this should be mentioned earlier in the manuscript because it constitutes a major limitation for polar applications of AltiKa.

p. 6: References for statements in l. 10-15 ?

p. 6, l. 20: "*It*" -> it + "*this criteria*" -> this criterion

p. 6, l. 21: Is "*snagging*" a word generally used for this? (just asking)

p. 6, l. 22: "*To overcome these problems,...*" -> refers to which problems? the off-nadir ranging of leads or also roughness?

p. 6, l.26: "... *we instead adopt an approach...*" -> Did you come up with this approach? Or did you take up an existing approach? (If yes, which one?/Reference?)

p. 6, l. 30-31: "*the appeal of this methodology is its applicability to any freeboard data sets*" ->

Why would this (i.e. applying to any freeboard data sets) not be possible for the method described in Guerreiro et al. (2016), for example? Wouldn't both have to be re-evaluated for their performance with different freeboard measurements anyways?

p. 6, l. 31: "*By calibrating satellite freeboards with an independent data set, biases are systematically corrected for*" -> I think this statement is too "optimistic"/general. Whether or not biases are systematically corrected for depends to a large extent on the quality, accuracy, and temporal + spatial resolution of the independent data. Not to mention that the bias is not the only thing to worry about...

p. 7, l. 7: "*snow depth, retrieved with the Kansas Snow Radar to within 5 cm accuracy*" -> Here (and also already in the introduction) it should be mentioned that different snow depth retrieval algorithms give very different snow depths! (Kwok et al., 2017)

p. 12, l. 5: Asterisk too high?

p. 14, l. 22: "*Spring*" -> spring

p. 16, l. 3-5: Did you use 2016 OIB data for calibration when comparing with the 2013, 2014 and 2015 OIB data? If not, why not?

p. 16, l. 14-15: remove parentheses around "*Kwok et al., 2017*"

p. 17, l. 1: "*onto a onto a*"

p. 17, 9-10: "*Snow depth agrees with expected spatial distribution and magnitude*" -> Compared to what? How do you know? Or do you mean just with regard to the statement that follows (on thicker snow over multi-year and thinner snow over first-year ice)?

Fig. 18, l. 8: "*...this evaluation does demonstrate the ability to up-scale OIB snow depths to the wider-Arctic*" -> Do you mean ability as in "we do not get nonphysical snow depth values" or how is the ability for this demonstrated here without comparing the snow depths to independent data?

Fig. 1 & 2: For the sake of completeness, it would be good to mention what the dashed grey line is (the zero line I guess).

Fig. 3: Why are the snow depths smoothed? Is there a physical reason for this? The spatial variability contains information too (about real variability or about the "consistency" of the method, for example), why not show this?

Fig. 6: It is very hard to see the OIB measurements on top of the snow depth map. Maybe zooming into the campaign area would be useful? L. 3 of caption: "*each grid cells*" -> each grid cell"

Fig 6 & 7: In none of the scatter plots there is snow depth values <0cm or >60cm, why would you show the data for a range of -20 to 80cm? This raises the question whether this was made to make the regression look "better"... and also creates unused white space that could be used for information instead.