The Cryosphere Discuss.,
https://doi.org/10.5194/tc-2019-140-RC2, 2019

The Cryosphere
Discussions

Interactive comment on "Estimating Early-Winter Antarctic Sea Ice Thickness From Deformed Ice Morphology" *by* M. Jeffrey Mei et al.

Stefan Kern (Referee)

stefan.kern@uni-hamburg.de

Received and published: 13 August 2019

Review of

Estimating Early-Winter Antarctic sea ice thickness from deformed ice morphology

by

Mei, J. M., et al.

Summary: A suite of very-high resolution contemporary sea-ice draft, snow depth, and surface elevation measurements carried out during the PIPERS expedition into the Ross Sea in May/June 2017 is analysed. First, it is used to obtain a set of very-high resolution sea-ice thickness distribution for the probed sites; this is used as a bench-

mark data set for the rest of the paper. Secondly it is used to investigate ratios and relationships between the three measured and the one computed (sea-ice thickness) parameter paying special attention to a discrimination between level and deformed areas. Various linear models are subsequently applied (and compared to the literature) to figure out which of the parameters measured need to be used and/or combined to obtain an optimal sea-ice thickness product - compared to the above benchmack data set. An attempt is made to relate variations in the coefficients used in the linear models to variations in effective density. Finally, based on these results, a deep learning convolutional neural network is fed with surface elevation information, trained, evaluated and applied to predict sea-ice thickness distributions of independent very-high resolution PIPERS elevation data sets - with quite convincing results with respect to the mean relative error. It is shown that the application of the network trained with PIPERS data is not necessarily succesful in predicting sea-ice thickness from surface elevation measurements of a different expedition.

I found the paper easy to read.

I definitely support publication in "The Cryosphere".

Despite the current quality of the paper I recommend that the authors invest some effort to better explain a few issues, to also improve one or two of the figures, and to better structure it by means of creating a separate methodologies section. I am inclined to rate this a major revision. The main concerns I list in the general comments (following immediately below), whereas the smaller - often less general issues - I noted in the specific comments. You will also find some typos noted.

General Comments: GC1: I note that a dedicated "Methods" section is missing completely. It is Data followed by Results. There are places where this seems ok for the flow of the manuscript but there are other places, e.g. Section 3.2 where this seems not to be optimal. The deeper I stepped into this section the more confused I got. At a certain point I got lost with density values and with regressions with or without intercepts or additional constants. This section would perhaps benefit from a clear up-front explanation of what you did / how you derived coefficients / which density values you choose (and why) / how you derive effective density values (and why)? Such an improvement in structure of the paper would possibly also reduce its length a bit here and there.

GC2: You treat the hydrostatic equation as a form of a linear fit. While one can see this as such a fit it would be very important to mention (even more) that the coefficients as you call them are based on density values and are computed based on physics. This is an important difference to the empirical linear fits used by Xie et al. or Ozsoy-Cicek et al. which are purely mathematical. To my opinion it would add to the understanding of your paper if you would clarify this even better at an appropriate position in your paper. I'd think that interpreting the CNN results into the direction that effective densities can be derived is very hypothetical - especially given the unknown (and non-existing) relationship between sea-ice and snow densities which are both involved. I note in this context that the issue of negative ice freeboards has neither been mentioned nor discussed. I guess it would not hurt to get back to it given the results published in Ozsoy-Cicek et al. (2013) and Yi et al. (2011).

GC3: I am missing the presentation / discussion of more results of the ConvNet approach. What a reader might have loved to see is profiles of sea-ice thickness computed from the draft-snow depth-surface elevation measurements (your benchmark) and of the sea-ice thickness estimated with your approach. Ideally you are able to show at least one representative profile of each PIP used here. That way one will get a better handle on the actually estimated sea-ice thickness distribution compared to the measured one - in addition to the histograms shown.

GC4: ICESat-2 is up since September last year. After having read the paper I am wondering what the ultimate goal of your work is. Is it to create high-resolution validation data sets of the sea-ice thickness which are spatially distributed? Or is it to develop an algorithm which potentially could be applied to ICESat-2 data. For both cases, I believe

the authors could stress the main motivation and future use of their work and product.

Specific Comments:

Page 2, Line 14: I suggest to cite the paper by Behrendt et al., 2013, Sea ice draft in the Weddell Sea, Earth System Science Data, 5, measured by upward looking sonarsto underline that also this ULS data is a valuable source - even though you write "sporadic"

Page 3, Line 1-9: I guess it would not hurt to perhaps again refer back to Kern and Spreen (2015) who dedicated some work on the uncertainty analysis for ICESat sea-ice thickness retrieval and to Kern et al., 2016, Antarctic sea-ice thickness retrieval from ICESat: Inter-comparison of different approaches, Remote Sensing, 8(7), who inter-compared a number of sea-ice thickness retrieval approaches for the Antarctic and on which results the Li et al. (2018) paper cited is based upon. I am tempted to say that the fits used by Li et al. (2018) are based on the work of Ozsoy-Cicek et al. solely and not on the work of Xie et al. (2011).

Line 14: "do not yet understand the distribution" –> I am inclined to say that we do understand the physical mechanisms forcing the snow depth distribution around ridges very well. What we cannot yet do is, however, to measure this distribution accurately over an large enough area.

Line 10-19: In this paragraph, the work of Weissling and Ackley, 2011, Antarctic sea-ice altimetry: scale and resolution effects on derived ice thickness distribution, Ann. Glaciol. 52(57) might fit as well.

Lines 16/18: This is perhaps a good place to refer to the work of Hutchings et al., 2015, Comparing methods of measuring sea ice density in the East Antarctic, Ann. Glaciol., 56(69)

Lines 20-22: I agree that the unknown snow depth is one factor here. But isn't the fact that we don't know the keel morphology and distribution relative to what we see from above with a LIDAR contributing much more to a potential bias in estimated sea-ice thickness?

Line 23-27: As far as I know, Kern and Spreen (2015) focused quite a bit on ICESat and the uncertainties involved. I doubt, however, that this is the correct citation for the AMSR-E snow depth bias issue. I'd say the first to report this issue were Worby et al., 2008, Evaluation of AMSR-E snow depth product over East Antarctic sea ice using in situ measurements and aerial photography, J. Geophys. Res., 113. Their work was followed later by Ozsoy-Cicek et al., 2011, Intercomparison of Antarctic sea ice types from visual ship. RADARSAT-1 SAR, Envisat ASAR, QuikSCAT, and AMSR-E satellite observations in the Bellingshausen Sea, Ann. Glaciol., 52(57) or Kern et al., 2011, An intercomparison between AMSR-E snow depth and satellite C- and Ku-band radar backscatter data for Antarctic sea ice, Ann. Glaciol. 52(57). In the same Ann. Glaciol. volume you also find the paper by Markus et al., 2011, Freeboard, snow depth and sea-ice roughness in East Antarctica from in situ and multiple satellite data, Ann. Glaciol., 52(57). Another paper about the deficiencies of the AMSR-E snow depth product could be this one: Kern and Ozsoy-Cicek, 2016, Satellite Remote Sensing of snow depth on Antarctic sea ice: An inter-comparison of two empirical approaches, Remote Sensing, 8(6).

Line 33: "fewer such datasets exist" –> This applies to the Antarctic and I would mention this accordingly. In the Arctic there are way more draft measurements available and these have actually been used to develop draft-based sea-ice thickness estimation tools.

Figure 1: I love this figure. It could be even a tiny bit more realistic if the ice floe or sheet would not be continuous in the ridge / keel area.

Page 4: Lines 3-12: I am wondering whether in the context of this discussion the work of Goebell, 2011, Comparison of coincident snow-freeboard and sea ice thickness profiles derived from helicopter-borne laser altimetry and electromagnetic induction sounding, J. Geophys. Res., 116 should be mentioned as well?

Page 6: I suggest to mention / give answers to the following questions here: - Water depth in which the AUV was operating - How many AUV scans per "cake" were stitched together? - If multiples scans: Were all scans carried out into the same direction? Or parallel to each other in opposite directions? X-ing? - Did I understand correctly that per "cake" 4 surface elevation scans were carried out, each from one side of a 100 m x 100 m grid? Or did you actually fly over the area? - I assume the snow depth measurements were the last measurements carried out –> although it is logical it is worth to mention this. - It is not entirely clear how the about 2000 measurements per "cake" are distributed across the "cake" area and how this was technically realized. I assume that the measurements were carried out along parallel transects across the cake with a fixed transect-to-transect distance and that only the sampling along each single transect varies between 5 m and 0.1 m. - How is the reference sea-surface height computed and how accurate is it?

Line 29: "The ice thickness can ..." –> so there were no drillings?

Figure 5: Readability of this figure would improve with an increase of its size.

Page 11:

Line 18: "very similar value of 1.3" –> This very similar value needs two standard deviations (2 time 0.1) to include that 1.5 values from other studies. Perhaps "similar" would do it?

Figure 6: - See comment to Figure 5. - In the caption I would call the black line dashed rather than dotted.

Page 12:

Line 11: "The lidar and AUV data were corrected by ..." –> I don't understand what needs to be corrected here. Is there a way you specify better what you did? Why is this a "correction"?

Page 13: Line 12: "snow = surface elevation assumption" –> I recommend to mention that this is a strong assumption, that it applies to thin, perhaps medium thick first-year ice only (possibly only to the kind grown under quiescent conditions, i.e. not originating from the pancake-ice cycle), that it requires a certain snow load to be present, and that such an assumption can only be made if one is not interested in a really exact sea-ice thickness estimate.

Line 17: "All our coefficients" –> Would you mind to refer to the place where you already mentioned these coefficients?

Line 21: "2.2-3.1 in Ozsoy-Cicek et al (2013)" –> I tried to figure out how you ended up with this range and potentially misunderstood something. If I check that paper, then - in Figure 5, which is possibly the one you got these numbers from, - I find regression lines with a considerable intercept between ∼10 cm and ∼30 cm, depending on the region, paired with this range in factor of F of 2.2 - 3.1. But these values are valid for positive ice freeboards only. When taking all ice freeboards into account, then the black lines (and numbers) in that Figure 5 apply. In addition to that: I could make sense to focus only on the Ross Sea results from that paper?

Lines 30-32: "This means that assuming ..." –> So what you state here basically is, that the linear regression approaches developed by Xie et al. and Ozsoy-Cicek et al. are of limited value? If so you could mention this and also refer to Kern et al. (2016) in Remote Sensing, where it is layed out that the linear regression approaches fail to provide a meaningful circum-Antarctic sea-ice thickness distribution.

Figure 7 - Some data points are annotated "snowy" –> I did not find an explanation of what this is in the text or in the caption. Where is the distinction between "snowy" and "ridged"? - caption: The ice density value given in line 4 of the caption differs from the one given in the text on page 13, line 25.

Page 14: Line 3: "T = 2.45F + 0.21" –> is modified from Ozsoy-Cicek et al., now using unit meters instead of centimetres, correct? "for a winter Ross Sea" –> according to Ozsoy-Cicek et al. (2013) this is data from just one cruise in Sep./Oct. = much later in

the season than PIPERS. In that sense your statement in Line 6 "same region/season" should perhaps be changed? Also the spatial overlap (see Ozsoy-Cicek et al., 2013, Figure 1) is quite small.

Line 4: So your intercept is -0.73 meters or -73 cm? That is quite large.

Line 7: "nonzero freeboard" –> "nonzero ice freeboard"

Lines 15-17: Yes, I agree with your interpretation. However, it might make sense to also mention that the Ross Sea data used in Ozsoy-Cicek et al. (2013) was from a different part of the Ross Sea and from a different season and to my opinion indeed exhibits a totally different characteristics than the PIPERS data set collected 3-4 months earlier.

Page 15: Line 6: "additive constant" –> I don't understand what you mean by this. Did you add an intercept?

Lines 7 & 8: Isn't it surprizing that the coefficient for F fitted over all four PIPs of 10.4 is so close at the upper range of 10.6 for individual PIP fitting? Also: The range for the coefficient for D of the individual PIPs does not include the value found over all four PIPs. Is this logical?

Lines 9-12: - Your measured snow densities are considerably lower than those given by Sturm et al. (1998). Could it be that the latter were obtained in late winter / spring? - While I understand the concept behind the effective sea-ice density (voids filled with water included in the density estimate) I have problems to understand the concept of an effective snow density. What is this? In this context, I find your effective snow density value to be quite high. - I guess it would be good to learn how you ended up with the density values reported in Line 10. When I tried to insert your range for the factor for F (7.9 to 10.6) into an equation where D is zero, then I end up with densities between 897.9 and 931.0 kg/mˆ3. But of course, without further information from your side I cannot reproduce your numbers. - I find it quite surprizing that the standard error for the effective sea-ice density is so low compared to that of the effective snow density.

- You use a water density value which differs from those given at the beginning of this paragraph. Why? Where does this value originate from?

Lines 15-18: Please check these sentences. There is some repetition first and then something is missing.

Lines 20-22: "For example, ..." –> Just to understand this: What you write here in the text is the comparison between using coefficients of ONE of the PIPs to estimate sea-ice thickness in another PIP while in Table 3 you show the comparison between using a joint coeffient of THREE PIPs to estimate sea-ice thickness in the remaining PIP. I just got confused a bit about why you write different things in the text than you actually show in Table 3 (and refer to in the subsequent sentence).

Table 2: - "no int." means what? - Are you sure the AIC is a monotonic function even on the negative value range? I am just wondering whether the "smallest" AIC criterion does not need to be applied to absolute values? Could it be that these negative values have no proper meaning in case that the correlations are so low? - The subscript "adj" stands for "adjunct"? "adjusted"? If the latter adjusted to what? - What is the unit of the constant? It seems to be in meters?

Page 17: Table 3, caption, last line: "zero freeboard = zero thickness condition"? Do you refer to ice freeboard here? Do you perhaps mean "zero snow depth"?

Line 1: This equation is something you could use in a "Methods" section (should you include one) to tackle my general comment GC1. I note however, that seemingly with this equation one can explain only parts of the entries in Table 2; the c3 times sigma part is not represented in Table 2.

Lines 5-9: It is still not clear to me how you discriminate between "snowy" and "level". Please add.

Page 18: Lines 3-6: I can in principle follow your argumentation that zero surface elevation (= zero snow depth) means zero sea-ice thickness. I would sign this if we

consider larger scales. But on the scales investigated with the PIPs this is not necessarily true because under cold conditions and hence impermeable sea ice there will be many places with a negative ice freeboard. Even if we assume for simplicity that most of these will have a snow cover and hence potentially have a non-zero surface elevation, it is still likely that especially in the vicinity of ridges and/or where the ice is under lateral stress - at the scales of your measurements - you will have surface elevations close to zero or even negative ones paired with a non-zero sea-ice thickness. - This paragraph is again a good place to comment and/or underline the difference between the physically based coefficients used by Zwally et al. and similar papers and the empirically based coefficients used by Xie et al and similar papers. One could argue that the physically-based coefficients are more dependent on the validity of the hydrostatic assumption while the empirically based ones are not ... but I am not sure this holds.

Page 19: Line 13/14: "20% of the data ..." –> this refers to the randomly selected data? If so, please stress so in the text.

Line 13 vs. Line 16 and remainder of the text: Please check your usage of "floe". From the text until here I got the impression that the PIPs are subsets of one floe. Here I get the impression that PIPs comprise several floes out of which a few are selected. Please clarify your terminology here.

Line 20: You state PIP8 here but in Figure 9 it seems you refer to PIP9. Please check.

Line 24: "epoch 881" –> does it make sense to refer to the Appendix here? Otherwise this information is perhaps a bit out of context.

Line 31: "are all negative" –> except for level ice.

Line 33-35 and beyond: I doubt that this comparison should be presented as is. Aren't these data sets quite different? I wrote about the sub-set of data for the Ross Sea used in Ozsoy-Cicek et al. (2013) already. In Li et al. (2018), the data basis is ICESat footprint-scale estimates of the freeboard - hence we talk about one value for one footprint of which we do not know how well it covers how many different surfaces. The data set used in Ozsoy-Cicek et al. (2013) is at least based on multiple measurements conducted on one or more transects across a single floe.

Figure 9: The description / caption of the figure needs to be improved. - Please annotate the images with a), b), c) ... - What is the value behind showing a continuous fit in addition to the bars? It extrapolates the bars towards non-existing data values. - What are the bin-sizes used? Are these the same in all three histograms shown? Do they always have the same borders (i.e. minimum and maximum value included into the count of a respectiv bar)? - The peak counts are obscured by the legend. This needs to be changed. - I suggest to add in each row which PIPs are used for what. - I suggest to stress in the caption that the last row shows a different range of thickness values. - The caption in line 3 says PIP7-9 but in line 2 it is PIP4, 7 and 8. What is correct? - The caption in line 4 says PIP9. True? - What is the unit at the y-axis in the histograms? - The mixed colors in the histograms originating from overlapping bars of different data sets are not easy to interpret. Perhaps you could either add these in the annotation (which is possible if you increase the size of Figure 9) or find a different way to show the counts of the different data sets. One way to do this would be that you use substantially narrower bars which you do not let overlap each other and center these at a specific thickness; then in the caption you might need to state that you display three bars centered at a specific thickness, separated horizontally for better visibility. That way the real differences in the distributions would become more clear.

Page 20: Lines 2-4: Perhaps you could put these numbers in context with the number of data points used to get these uncertainty estimates? I guess, in case of Ozsoy-Cicek et al. (2013) we are talking about 23 floes with an actually unknown number of measurements per transect. About how many measurements are we talking in your case?

Page 21: Line 14: I suggest to remove the "see" –> at least I cannot see these results.

Line 29: "isostatic assumption may no longer be valid" –> may be so. What do we know about spatial scales over which the isostatic assumption is valid? No too much I'd say - particularly for ridged ice. Perhaps this sentence could be deleted.

Figure 10: - this figure belongs to section 4 and should be located within section 4 not before it. - Why do we have 3x3 imagettes for layer 1 but 16 for layer 3? - The size of 4 m and 8.8 m given in the caption, do these refer to the pixel size in these imagettes or to the imagette size itself? It seems as if the pixels in layer 1 are indeed smaller than in layer 3. - Instead of "as the lidar" you might want to write "as the surface elevation" - Is it in this context correct to assume that layer 3 has the unit meters while layer 1 is unitless? - What do the bright and dark pixels in layer 1 mean?

Page 22: Line 17 through Page 23: Line 7 and Figure 11: - Please provide a), b), ... in Figure 11; it aids referring to the images. - I suggest to mention Figure 11 before Figure 12. - You refer to Figure 11 in Line 5 but should perhaps also do it in Line 3 (strong correlation for feature #0) and again in Line 6. - I have difficulties to understand the continued mentioning of "effective densities". I doubt that with the CNN you can (and should) derive any conclusions about the effective density - especially because the densities for sea ice and snow do not necessarily co-vary. This brings be back to GC2.

Page 23: Line 10: "snowy surfaces" –> which still need to be defined in comparison to "level surfaces".

Page 24: Line 4: "ridged and level surfaces are clearly distinguishable" –> I don't agree when I look at Figure 13 - unless I have perhaps misunderstood what the used tool is able to show. But my interpretation of this figure is that level, ridged and snowy symbols overlap well.

Page 25: Lines 9/10: That prediction of snow depth from lidar input is possible as also been shown by Ozsoy-Cicek et al. (2013) and Kern and Ozsoy-Cicek (2016).

Lines 11-13: I guess these two sentences could be deleted.

Page 26:

Lines 1-8: I am not a fan of these attempts to try to relate CNN features to (effective) snow density variations which may or may not be realistic and physically meaningfully linked to input parameters. To my opinion, this really requires a careful analysis and description of how the CNN "learns" from the input data and whether there is (within the CNN) a link to physics - which I doubt is the case.

Lines 18: "thickness of a new dataset" –> you seem to have applied your approach to a different PIPERS data set. It might be really beneficial to show this example in the paper and not to just mention it. Particularly because you come back to this in your conclusions (Line 33, "unseen floe").

Page 27: Line 5: "it can account for a varying ice/snow density" –> I'd say that this is a hypothesis. It may be that the ConvNet is able to account for the different densities and perhaps even provide additional information about these - but the evaluation of whether this is the case and/or whether this is at all meaningful physically based on deep learning is not known and might not be over-stressed here.

Lines 12-15: "Our error ..." I suggest to not overstress these inter-comparisons because these are based on completely different data sets and scales. After all, a real quality measure of your method will be its application to ICESat-2 data which should be the overall goal here - as is finally mentioned in the last paragraph.

References: You need to go through the references list and complete it with respect to page numbers and journal volume and issue numbers. Also doi's are generally missing. Some journal abbreviations are not in place.

Page 28: Line 31: I am not sure but I guess Figures in the Appendix need to be named differently to the main text. See the instructions for authors.

Page 29: Line 2: Would you mind commenting on the layer sizes being first 4 m, then

8.4 m and subsequently 8.8 m? Do these "strange" values have to do with the pixel size of 0.2 m?

Figure 15: - What is a "training loss"? - It appears that after epoch $\sim$550 there is a small jump in validation and training error from a certain level before that epoch to a certain, lower level afterwards. Any explanation to this? - What explains the sudden increases in the training error from a low background of $\sim$15-16% MRE to the level of the validation error? It seems as if the result of the ConvNet even after that many epochs is still not stable?

Typos:

Please replace "e.g" by "e.g." (a few incidences) Please check usage of "climatology" and replace all incidences in the paper by a more appropriate term.

Page 2: Line 9: witeh –> with, interrannual –> interannual

Page 9: Line 26/27: "beyond beyond" –> "beyond"

Page 15: Line 11: "which" –> "who"

Page 21: Line 28: "slighly" –> "slightly"

Page 28: Line 23: "assigining" –> "assigning"

Figure 14, caption, line 3: "optimzer" –> "optimizer"

Interactive comment on The Cryosphere Discuss., https://doi.org/10.5194/tc-2019-140, 2019.