Review of manuscript

Sea ice leads in the Arctic Ocean: Modelling, inter-annual variability and trends

by

Sanggyun Lee et al.

Summary

The paper describes a new approach for classifying Cryosat-2 waveforms along-track into leads and sea ice classes using a linear unmixing approach based on ideal endmember waveforms representative for each of the two surface types. The resulting classification is validated using MODIS reflectance data and is also compared to previously published Cryosat-2 lead retrievals. The authors find their approach to outperform existing methods and provide monthly lead fraction maps, which they compare to other available lead products. Spatial and temporal lead patterns in their results are discussed in the context of inter-annual sea-ice variability, drift and thickness.

General comment

While the technical approach suggested here is generally valuable and probably expands the toolbox for lead detections from Cryosat-2 data, the presentation, documentation and discussion of the results is not yet convincing and requires improvements. While the authors put forward some very strong hypotheses, some of the discussed findings are not supported by what is being shown in the figures. I think that major revisions are necessary for this study to merit publication in TC, mainly because the comprehensiveness of the paper and the degree of innovation are not yet well balanced. I suggest the paper to be worked over with emphasis being put on the validation and the documentation of the technical approach.

I will try to itemize my main critics and provide suggestions to improve the manuscript in the detailed comments below:

Specific comments

P2, L31: I think that point 3) is not really part of this paper. Some short discussion is provided on this topic, but not really what could be called an "investigation".

P4, L20: Here it is stated that monthly lead fraction maps were calculated. However, no further information is provided on how this was done. How did the authors deal with data gaps in the daily data? How did they convert binary data into lead fractions?

P5, L15: "The large number of ….". This statement needs some explanation.

P5, LL 21-24: What was the criterion for selecting waveforms? More information is required here.

If lead waveforms were selected using the method of Lee et al. (2016) then it should not be surprising that the results are most similar to those of Lee et al. 2016. This requires some discussion.

P5, LL25-27: What is the motivation behind comparing waveforms between months / seasons?

P6, 1st paragraph: The description of how MODIS reference data were computed needs more explanation. Please be more specific about what is compared here. The terms "examined", "assessed" and "evaluation" are used, which leaves the reader somewhat confused.

In general, the processing of the reference data (MODIS) set is not described in a sufficient way. How were MODIS swaths binarized in order to calculate the given performance metrics. Which problems might be involved?

P6, 2nd paragraph: I think this description would merit a figure to document how original waveforms end up in a binary decision on whether leads or sea ice are dominant at a given point along the altimeter track.

P6, 3rd paragraph: More explanation and documentation on the calculation of the metrics is required (equations?). How was the reference data (binary MODIS) computed?

The decision for a 10x10 km grid is not explained. Why did the authors not use a grid similar to Wernecke and Kaleschke (2015). This is especially equivocal because a) the accuracy (not sensitivity) of the lead retrieval seems to be low at lower latitudes with such a grid and b) the authors themselves state "It should be noted that the corresponding lead fraction might not represent actual lead fraction in a 10 x 10 km grid" (P11, L14).

P6, L26: What is meant by "repetitively permuted"?

P6, L27: I think it should read "…, the more sensitive the OBSERVED LEAD FRACTION IS TO THE NUMBER OF AVAILABLE OBSERVATIONS". ?

P7, L8: It is unclear to me, why 2 thresholds are necessary.

P7, LL12-14: I think that everything in parentheses can be skipped here.

P7, L20: "regardless of month": Only single swaths are presented here which does not allow a conclusion about seasonal dependencies.

P7, L21: The fact that results compare best to DT is not surprising as lead waveforms were obviously selected with the same method. This point requires to be mentioned and discussed.

P9, L9: "It is widely … ". I do not understand the subsequent argumentation with "However, the lead …"

P10, L2: It is very uncommon for Arctic surface melt to start already in April.

P10, L3: How is the lead fraction associated with the seasonal cycle of sea-ice thickness?

P10, L4: "April of 2013 and 2016" I cannot see in the figures what is being described here.

P10, Next sentence: This is a strong hypothesis. Are there references which can support it?

P10, L7: "December 2015 to .." Again, the feature described here is not in the Figure.

P10, P11, figures: The subfigures should be much larger. Please consider using one page per figure.

P10, L16: "surface elevation anomaly…" How would this help to detect leads more accurately?

P10, last paragraph: It is stated here that different spatial resolutions of the sensors are the main cause for spatial differences in monthly lead fractions. I doubt that this is the main reason. The passive microwave and thermal infrared sensors are just sensitive to different surface features as compared to the altimeter.

It is stated that there is predominantly Pancake and Nilas in the Chukchi Sea are and that this is the main cause for differences between methods. However, this is pure speculation. No support is provided for this statement. In the next paragraph, the authors even state that "Altimeter-based monthly fraction maps might be insufficient to represent monthly lead fractions in the coast line of the Arctic Ocean…", which again makes the previous discussion about the Chukchi Sea obsolete.

P12, Figure 6: It is not really clear to me, how the standard deviation is to be interpreted in terms of uncertainty. Would a higher variability of lead occurrences not also cause a higher standard deviation?

P14, L4: "well" in comparison to what?

P14, L4: What is "the shear zone"?

P14, LL10-14: Doesn't the better resolution come at a price which is too high, when the data in lower latitudes cannot be trusted?

P14, last paragraph: Again, it is stated that altimeter data are well suitable for a detection of pancake and nilas, which is by no means supported by any of the presented analysis.

P15, Figure 7: It is read to recognize details in the subfigures. Please enlarge this figure.

P15, L10: It is a bit misleading to compare Fram Strait and the Beaufort Sea in this context. The first is already part of the marginal ice zone in my mind, where lead dynamics are not primarily a result of the Transpolar Drift.

P15, L13: Six years of observations do not allow to draw conclusions about a trend.

P15, L17: "suddenly" … should be removed

P15, L18: "March and April …" The figures do not really reproduce the feature that is described here.

P15, L18 … "This unexpected decrease …. sea ice thickness " This statement needs some clarification. Why is this the case? What is actually meant here?

P16, Figure 8: If averaged data is shown and no data are available for summer, then this figure should not show continuous lines, but rather points or bars.

P16, L11: I would rather call this chapter "Discussion". Moreover, I think it could be merged with chapter 5.

P17, L4-5: "… which implies that sea ice becomes more vulnerable …". This is a strong statement which would hold only if the "recent years" could be compared to a longer time series or at least a reference.

P17, L9: How was the improvement in spatial resolution achieved? Wasn't that an arbitrary choice, which is hard to justify given the high uncertainty in the lower latitudes?

P17, L18: Please consider talking of "fractions" instead of "abundances"

P17, L19: Not clear what is meant by "regardless of month"? This is not what can be inferred from Figure 2.

Technical corrections

P1, L12: "…detect leads."   Add "from Cryosat-2 data"

P1, L20: "…from the literature". It is rather "…from previously published data sets."

P1, L22: replace "known as" with "referred to as"

P1, L22: replace "between sea ices" with "in sea ice"

P1, L26 and throughout the manuscript: the "u" in "Lupkes" must be "ü". This holds also for "Röhrs" and "Bröhan"

P1, L27: rephrase "…could make near surface temperature up to a 3.5 K…"

P1, L29: "atmospheric boundaries". Better: "atmospheric boundary layer" ?

P2, L4,5: remove commas around "…with 1 km spatial resolution…"

P2, L10: replace "produced" with "to produce"

P2, L11: add "channels" after "(AMSR-E)"

P2, L13: Replace "could" with "can"

P2, P2, LL13-18: This paragraph needs rephrasing.

P2, L19: "…pixel ARE a linear combination…"

P2, L22: remove "the" after "However"

P2, L23: replace "enough" with "sufficient"

P2, L24: SIRAL instead of SAR ?

P2, L25: replace "…in January…" with "from January…"

P2, L26: Please add here from which data set N-FINDR selects endmembers. What is meant here by "mitigate"?

P2, L27: "evaluated" or rather "validated"?

P2, L30: replace "2) identify…" with "2) compute…"

P3, L7: "SAR and SIN". These abbreviations are not introduced. Explanation is required on what is the difference.

P3, L9: Merge parentheses after PRF

P3, L12: Remove "…and this is called multi-looking"

P3, L13: Replace "The results in the …" with "Exemplary results of possible…"

P3, L14: "…with a condition". What is meant? Please rephrase.

P4, L21: The data source should be: "These maps are available at the PANGAEA Data Publisher for Earth and Environmental Science data (http://dx.doi.org/10.1594/PANGAEA.854411)."

P4, L28: remove "an"

P5, L7: No URL is required here.

P5, LL24-25: Too many references.

P6, L3: remove "having"