Interactive comment on "Mechanisms influencing seasonal-to-interannual prediction skill of sea ice extent in the Arctic Ocean in MIROC" *by* Jun Ono et al.

Anonymous Referee #2

Received and published: 7 September 2017

Summary

The authors present results on Arctic sea-ice extent prediction skill obtained with a MIROC-based forecast system. Further, they explore possible reasons for differences in skill in different times of the year based on lagged correlation and regression patterns, focussing on preceeding states of the (subsurface) ocean heat content and of the sea-ice itself.

In general, The paper is generally well-written and provides interesting results that merit publication. However, there are some points that in my view need further scrutiny. For example, the conclusion that the advection of subsurface water masses from the

Altantic Ocean into the Barents Sea, though plausible, is in my view not sufficiently supported by the results shown. Also, the definition of the subsurface ocean heat content and how it's interpreted deserves additional attention, and the rationale behind performing the lagged correlation/regression analysis primarily based on the hindcasts rather than on the control run, and what might cause differences between them, needs clarification. In addition, there is quite a number of minor issues, listed below.

Therefore I recommend the manuscript should be reconsidered after major revisions.

Specific comments

P1L8: The term "seasonal-to-interannual" should be shifted in front of "predictions".

P1L10: "of up three years" - here is a word ("to") missing.

P1L12: "December SIE_AO can be predicted up to 1 year ahead" - I suggest that this statement should be made more quantitative, e.g., by providing the ACC, and maybe also substantiated with the corresponding p-value.

P1L13-15: The role of advection as indicated here is in my view insufficiently supported by the results shown; see details below.

P1L23: "problem" - just as a side remark, I think this judgmental term adds an unnecessary political dimension to this observation.

P2L1: "or" - I think I know what is meant, but using "or" here seems illogical.

P2L2: "the potential predictability for sea ice extent is continuously one to two years" - I think this statement again needs some numbers; theoretically, marginal (but pratically meaningless) potential predictability should be out there for very long lead times, whereas pratically meaningful potential predictability survives much shorter lead times. At least, something like an ACC threshold which is considered to distinguish "meaningful" from "no" skill should be provided. (Note that "statistical significance" is not necessarily the correct concept needed here.)

P2L5-6: "The observed Arctic sea ice extent based on ensemble hindcasts can be predicted up to 2–7 and 5–11 months ahead for summer and winter" - see my previous remark.

P2L16: Again, I think that the term "seasonal-to-interannual" needs to be relocated, this time in front of "predictability".

P3L7-8: "eight ensemble members produced by perturbing the sea surface temperature based on the observational errors" - I am wondering whether these perturbations are able to generate any meaningful spread, given that the 3D ocean and atmosphere are assimilated towards the same, gap-free, reanalyses. Or, are the differences just very small (and all "assimilations" thereby very similar; note that Fig.2 also shows just one single "assimilation"), but of course sufficient to trigger subsequent divergence during the free forecast/hindcast runs due to atmospheric chaos, so that the same effect could have been obtained with quasi arbitrary small initial perturbations? Maybe the authors can comment.

P3L17-18: "the detrended components were calculated by subtracting monthly linear trends during 1980–2009 from the original monthly data, and anomalies are defined as deviations from the climatology from 1980–2009" - are not the "detrended components" mentioned at the beginning of this sentence already the "anomalies"?

P4L6-7: "September SIE_AO can be dynamically predicted from the previous July" - again, I think this statements needs some quantification; the same holds for the subsequent sentence.

P4L8-9: "The ACC is also significant for the winter SIE_AO, in particular for December, except for the hindcasts started from April 1st, indicating the potential use of dynamical forecasts up to 1 year ahead" - the fact that December SIE_AO is more skillfully predicted by the January hindcasts than by the April hindcasts, also visible in Fig.2, deserves more explanation. While such "reemergence" of skill is often encountered when simple statistical relations - like persistence - are used, in situations with strong

C3

seasonal cycles like given for sea ice, to my understanding this is not to be expected for dynamical forecasts: the closer to the target date they are initialised (taking into account current as well as past observations!), the better should the dynamical forecasts become. To be specific, the OHC content anomalies put into the January hindcasts should also make it into the April hindcasts, although subject to some advection etc. Instead, could this unexpected drop of forecast skill be a mere matter of sampling uncertainty?

P4L10: "The RMSE for all 10 hindcasts increases throughout the melting and early freezing seasons (July–October), before decreasing in November–June" - to be precise, it appears that the RMSE does not "increase" and "decrease" during those periods, but that it "is larger" and "is smaller" (with the change happening inbetween).

P4L18-19: See again my comment on P4L8-9!

P4L19: "those started from July 1st, in which only the September SIE_AO is significant" - this statement seems to contradict Fig.2 where there are many "significance stipples" for other target months as well.

P4L22-23: "the SIV_AO is defined as the sum of the grid cell volumes obtained by multiplying the sea ice thickness (SIT) by the SIC for grid cells with SIC greater than 15 %" - if I am not mistaken, the multiplication by the grid-cell areas is missing here, no?

P4L23-25: "the OHC_AO is the vertically integrated temperature multiplied by the density and specific heat capacity of seawater from the mixed layer depth (MLD) to a depth of 200 m, in the area north of 65° N" - (i) The way it's defined here, temperature is vertically integrated instead of averaged, so the distance from the MLD to 200m directly enters the "OHC" and should thereby dominate variations in "OHC" instead of temperature variations, which seems odd. Please clarify. (ii) Why is not the same area used as for SIE_AO, that is, excluding Hudson Bay and Baffin Bay?

P5L7-18 and Fig.4c-f: I am not convinced that the "advection and emergence hypoth-

C4

esis" constructed here is sufficiently supported by the results shown. Firstly, some of the apparent propagation of ("subsurface") OHC anomalies from off the Scnadinavian western coast to the eastern Barents Sea might be simply due to a slight shift of the area with a mixed layer deeper than 200m (areas with quite deep convection): a larger part of the Barents Sea is thereby effectively "masked" in March compared to December in Fig.4. Secondly, the sea-ice edge extends further into the Barents Sea in March compared to December (I assume this is true also in these simulations), and ocean temperatures under ice are subject to weaker variability (with the surface being tied to the freezing point). Thirdly, the rather narrow stripe of anomalies off the Scandinavian coast in March - an important part of the presented explanation - is not present in the control run (Fig. S6). Maybe some clarification could be provided if Fig.S5 was also provided for lags -3, -6, and -9 months? It might also help to clarify things if the integration/averaging was done between fixed depths, so that nothing is masked and the MLD changes do not superimpose temperature anomaly changes. Even more simply, showing just SST anomalies might help.

P5L15-16: "The above features are also found in the control run, suggesting that the advection processes of the OHC in the hindcasts are not due to processes distorted by the influence of initialization or climate drift in MIROC5" - In fact, I do not quite understand the reason why the main figures related to the lagged correlation and regression alaysis are not based on the control run in the first place. Maybe it's just me, but I am somewhat confused why this should be done primarily for the hindcasts, where also the statistical sampling is much worse. If the main analysis was based on the control run, however, it would make sense to show corresponding results for the hindcasts as a supplement, to prove that the shown relations still hold, no?

P5L24: "the persistence of sea ice states initialized in July persists" - the first word maybe should be "anomalies" or similar?

P26-27: "possible mechanisms or sources cannot be detected in the hindcasts started from April 1st (Fig. S8)" - I'd like to repeat my points that this might be partly due

C5

to sampling, and that important regions are "masked" due to the MLD-related OHC definition. I would argue that Fig.S6d, based on the control run (implying better sampling, although showing March instead of April), supports the notion that the April state should be at least as informative as the January state to predict September SIE_AO.

P6L4-6: "Numerical experiments to confirm whether the subsurface OHC anomalies 5 originating from the North Atlantic control the December sea ice extent in the BS and eventually in the Arctic Ocean will be explored in future work." - I am actually quite curious to see results of such interesting experiments!

P6L7-9: The first two sentences of this paragraph seem to contradict each other.

P6L20: "Further improvements in the predictability of sea ice" - here I would recommend to avoid the term "predictability (of)" because in my view "skill to predict" is more accurate.

Fig.2: I would find it helpful if the situations shown in panels c) and d) could be highlighted in panels a) and b), e.g., by black boxes around the corresponding fields of the heat maps. Also, do I understand correctly that panel c) corresponds to a 3 months lead time, whereas panel d) corresponds to a 11 months lead time? That could be stated more clearly in the caption.

C6