

Interactive comment on “Year-round Impact of Winter Sea Ice Thickness Observations on Seasonal Forecasts” by Beena Balan-Sarojini et al.

Anonymous Referee #2

Received and published: 8 July 2020

Summary

Balan-Sarojini et al. study the impact of Cryosat2/SMOS winter ice thickness (SIT) observation nudging on a lower-resolution version of the ECMWF ocean/sea-ice reanalysis (ORA) system and on associated coupled seasonal forecasts initialized from that reanalysis system. The SIT constraint suppresses an otherwise too strong annual SIT/SIV cycle in the ORA and provides overall thinner SIT conditions toward the end of the northern winter (except in the perennial ice regions north of Greenland and the CAA), which turn into decreased sea-ice extent in the ORA in summer (despite sea-ice concentration assimilation). The thinner/less extensive initial ice is beneficial for seasonal forecasts initialized before July, but forecasts initialised in late summer tend to be deteriorated. The authors show that this is linked to too-strong spring/summer melt in

C1

the ORAs (when no SIT constraint is available), leading to low-biased ice and warm-biased sea-surface initial conditions in summer, in combination with a too-late/too-weak refreeze in the coupled forecast system. Balan-Sarojini et al. show evidence that the latter points can be explained at least partly with the surface radiation budget in the atmosphere-forced ORAs and in the coupled forecast model.

The study is scientifically sound, well-written, contains appropriate graphics and references, and provides interesting insights into the impact of ice thickness observations on forecasts in the specific system used which might be helpful to understand other systems, too. I do have quite a number of remarks, most of which are however minor. The maybe most demanding recommendation is to compare against simple climatological benchmark forecasts where appropriate. In summary, I recommend publication of this work in The Cryosphere subject to minor(-to-major) revisions as detailed in the following.

Specific comments

L12-13: "we find significant improvement of up to 28% in the September sea ice edge forecast started from April" - From the abstract it does not become clear that the paper is almost completely focussed on biases (and how these affected by constraining SIT) and not on interannual anomalies. In the summary section you state very clearly that this is the case (L441-442), but I think it should be mentioned in the abstract, too. Without that information, the sentence in L12-13 leaves one wondering how such a significant forecast improvement can be reconciled with the "May predictability barrier". In this context, see also my recommendation below to consider comparing with a climatological benchmark forecast where appropriate.

L57: Zampieri et al. 2018 - There's also a follow-on paper demonstrating reasonable skill of ECMWF S2S sea-ice forecasts in the Antarctic: Zampieri et al. 2019 "Predictability of Antarctic Sea Ice Edge on Subseasonal Time Scales".

Eq. 1: It probably doesn't make a big difference, but can you specify whether this

C2

is "floe-thickness" or "effective thickness" (thickness when evenly distributed over grid cell)?

L162-164: "We have also tested the sensitivity to different nudging strengths by running variants of ORA-SIT with a relaxation time scale of 20, 30 and 60 days" - If you mention this, I would expect that you also say something about the impact of the relaxation timescale and why you chose 10 days.

L201-205: "slight underestimation over the central Arctic and overestimation over the Canadian Archipelago still remain in November. This is probably caused by the lack of SIT observations during the months preceding November" - Given the relaxation timescale of 10 days, I assume that this difference goes back almost completely to the first half of November? That would confirm that you could omit the word "probably"; that's a rather obvious link.

L208-209: "The largest impact occurs in March, probably because at this month the SIT observations have been assimilated during the preceding 5 months" - similar to what I say in the previous point, I assume that the SIT state responds according to the relaxation timescale. This implies that, on a monthly scale, the largest impact should occur in the month with the largest bias, no matter for how many months relaxation has been active before that month (as long as it's at least one month).

L210: "with a slight clockwise displacement" - you could mention that this is consistent with the mean climatological Arctic drift pattern (transpolar drift, Beaufort gyre) and thus likely a consequence of the mean advection.

L217-218: "In November [...] the SIT constraint has very little impact on SIC biases" - Could the reason be that (in addition to the fact that no SIT corrections are applied in the previous months) the thickness corrections made in Nov need more time to influence the sea-ice concentration, because that requires a "cross-impact" through other processes (dynamics and thermodynamics)?

C3

L225: "large positive increments from May to October, indicative of strong underestimation of SIC in the ORAs" - To be precise, should "in the ORAs" rather be "in the (hypothetical) forced model without SIC assimilation"? After all, the SIC assimilation makes sure that the SIC underestimation doesn't get too strong.

L232-235: Isn't the even bigger difference in the SIC increments after May (even though these are for the worse) even more strongly showing the long-lasting impact of the SIT corrections on the SIC assimilation?

L243: "low bias" could be mistaken for "negative bias", maybe better say "weak bias" or "small bias" or similar

L250-262: To compute the IIEE, do you use the ensemble-median ice edge (50%-contour of sea-ice probability where SIC=15% is used to determine "presence" or "absence" of sea-ice in each ensemble member) or do you compute it for each member individually and average the IIEEs afterwards? That would make a difference, so this should be specified. Related, note that there's a probabilistic version of the IIEE ("Spatial Probability Score", Goessling and Jung 2018 "A probabilistic verification score for contours: Methodology and application to Arctic ice edge forecasts") that you could apply to your ensemble forecasts directly, which would have the advantage that changes in uncertainty/reliability would be captured, too.

Fig. 6 caption and throughout the paper: DelSole and Tippett (2016) just apply the sign test (a special case of the binomial test with $p=0.5$), only that they visualize how the outcome develops from forecast case to case like a random walk. I would recommend to refer to the test simply as the sign test (which in fact dates back to 1710!).

Sect. 3.2 and Fig. 7: 1) Can you please explain how the bias correction is performed? Is this simply done for each gridcell individually? Do you just subtract the mean concentration bias (difference as a function of time of the year and lead time, averaged over 2011-2016/17), possibly with a correction that makes sure concentration values remain bound between 0 and 1? Or is quantile normalization involved? 2) Related to

C4

the bias correction, I would find it very useful if the forecast errors could be compared against a climatological benchmark forecast. The latter could be based simply on the same period (2011-2016/17), or on the preceding decade (to make it more "operational"). I would expect that the uncalibrated forecasts are worse than climatology for most lead times (except the first one or two months?), but the calibrated might beat the climatology for a few months? In the summary section you say very clearly that you are "yet to demonstrate the benefit of interannual varying data on bias-corrected forecast scores", but I think it would be rather easy and revealing to add a climatological reference (even if it reveals clear limitations of current sea-ice forecast skills).

Fig. 8, top: Can you provide an explanation why the SIV in the ORAs converge from May to September, so that the large SIT difference in spring is completely "forgotten", whereas the coupled forecasts maintain much of the initial offset? Is there some fundamental reason why the forced (vs. coupled) atmosphere would cause such a difference, or can it be linked to the continued assimilation of SIC (or ocean variables)?

Eq. 2: The way the melt energy tendency is defined, it seems to be really just the derivative of (area-averaged) SIT (times a constant factor), right? Also, maybe it's better to use partial d's to make clear that these are not material (Lagrangian) derivatives. Related, you could also mention that changes in SIT through divergence as well as advection are included, implying that the "melt energy tendency" can in principle also change through dynamics. I understand that, by averaging over a large area (almost the whole Arctic), most of any dynamical effects would be compensating each other, but being clear about this would be good.

L314-316 & Fig. 9: The plot caption reveals that for the forecasts you look only at the first-month MET, but you do not mention/explain/motivate this in the text. Further, do I understand correctly that, by considering just the first month of the respective forecasts instead of a "closed" seasonal cycle, the annual integral of MET is not expected to be zero (while it should be zero for the ORAs)? In fact it looks a bit like it's rather negative (average build-up of sea-ice volume), can you confirm this?

C5

Fig. 10 and corresponding text: I am wondering to what extent turbulent fluxes (in particular sensible) could also play a role, for example, with stronger downward spring/summer sensible heat fluxes in the forced ORAs compared to the coupled forecasts (acknowledging that there might not be a corresponding observational dataset to compare against). Too high near-surface temperatures that could generate too strong downward sensible heat flux would be consistent with a positive downwelling longwave bias in ERA-I, even if clouds also seem to play a role there. If differences in turbulent fluxes are too small to be important, please mention that.

L351-352: "Significant cold biases are present in forecasts for most of the start months and lead months" - Is this also true over Arctic sea ice in winter? If so, how can it be reconciled with Batrak and Müller (2019) "On the warm bias in atmospheric reanalyses induced by the missing snow over Arctic sea-ice"? I thought that the surface coupling is similar in the system studied here?

Fig. 12: I was a few times slightly confused when looking at this figure, intuitively thinking that the lower panels show differences between FC-SIT and FC-REF that could be directly combined with the biases shown in the upper panels. But the lower panels show the differences in mean absolute error, which is alright but easily misleading. I suggest to use a different colourbar for the lower panels so that the different flavours of "temperature" (signed vs. unsigned) is more intuitively reflected.

Technical corrections

L25: last -> lasts

L80: "as cross-check variables evaluation" - I recommend to reformulate.

L91: These -> This

L168: "differ on" -> "differ in" / "differ regarding"

L208: "gradients on" -> "gradient in the" or "gradients of the"

C6

L212: "end of melt season" -> add "the"

L217: "reduced up to" -> "reduced by up to"

L227: "indicates" -> "indicate"

L228: "at marginal seas" -> "in the marginal seas"

L232: "on an average" -> "on average"

L232-233: "in ORA-SIT analysis" -> add "the"

L265: "that to be" -> add "are"

L288: "is smaller" -> "are smaller"

There are a few more such tiny things, please check carefully!

Interactive comment on The Cryosphere Discuss., <https://doi.org/10.5194/tc-2020-73>, 2020.