

The Cryosphere Discuss., referee comment RC2  
<https://doi.org/10.5194/tc-2022-92-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Comment on tc-2022-92

Anonymous Referee #2

---

Referee comment on "Reconstruction of Arctic sea ice thickness and its impact on sea ice forecasting in the melting season" by Lu Yang et al., The Cryosphere Discuss.,  
<https://doi.org/10.5194/tc-2022-92-RC2>, 2022

---

In this manuscript the authors present a statistically reconstructed dataset of Arctic summer sea ice thickness (SIT), which they create using SIT & SIC from the CMEMS Arctic reanalysis system, TOPAZ. The resulting SIT reconstruction, BRMT, is then compared with several model products, and thickness estimated by BGEP Eulerian moorings and IMB Lagrangian buoys. Finally they BRMT dataset is assimilated into a short forecasting experiment for the period September 2011.

There are some interesting ideas and concepts here that are worthy of publication in The Cryosphere. However, there are some major issues that will need to be addressed before the manuscript can be published.

### Major comments/concerns

- The quality of English language used throughout the manuscript leaves a lot to be desired. As well as reducing the overall readability of the paper, there are also several cases in which it is hard to understand what is being said (or why).
- One of the main motivations listed for this study, and the BRMT dataset, is that summer satellite SIT data is not available. This issue is mentioned several times and the authors go so far as to state that it is "impossible". However, the authors do not take into account the fact that summer satellite SIT has been being developed now for many years. The first dataset of summer Arctic radar freeboard (10-years) was published at the beginning of this year (Dawson et al., 2022). (NB. conversion from radar freeboard to SIT has also been done but that paper is still in press.)  
I'm not necessarily saying that the existence of these new datasets invalidates the

motivation for this study, but it should definitely be referenced and included in the discussion. How does/would the BRMT compare with the Dawson et al. (2022) freeboard?

- The paper manages to be both too long and detailed, and too short and vague at the same time. By this I mean that, despite the paper being rather long, sufficient details are not provided of the methods used to create BRMT. Given the length of the paper, I recommend that the authors consider dropping Section 5 and focussing properly on the BRMT reconstructed dataset – both creation and evaluation.
- The “retrospective forecast experiments” in Section 5 are really only cursory, with less than one month of forecasts performed for only one particular year. To properly assess the impact of assimilating the BRMT SIT a much more comprehensive assessment would be needed – including the whole summer period on at least 2 different years. Many would also question the fact that BRMT uses information from the future (i.e., 2012-2018) in the reconstruction. This makes it unusable for real-world forecasting situations. How much skill would be lost if you were to only use past data?
- I struggle with the concept and motivation for the reconstructed SIT dataset. BRMT essentially uses the relationship between SIC and SIT in the TOPAZ reanalysis. Much of the motivation for using TOPAZ in this way is not included and so I am left with so many questions in my head: So why not just use TOPAZ? What extra is BRMT bringing to the table? Why do you trust the relationships in TOPAZ so much? How much difference would using a different reanalysis make to the SIT reconstruction? I’m also concerned that there is some horrible kind of circularity in the analysis here, whereby desired traits – such as the relationship between SIC & SIT – are included in the design of the system and then used as part of the evaluation.
- The comparison of the model-based and reconstructed SIT observations with the in-situ observations is either not performed carefully enough or not described adequately. I am not convinced that these comparisons are being performed or interpreted correctly for the following reasons:
  - The authors do not specify anywhere how they define Sea Ice Thickness (SIT). Is it the “floe thickness” (i.e., the average thickness of all the sea ice floes present in the grid-cell) or the “grid-box-mean thickness” (i.e., sea ice volume per unit grid-cell area)? The former is certainly what your point/in-situ observations (BGEP/IMB) are measuring. However, the latter is the prognostic used in the sea ice continuum models formulation and so very likely what you are using from the model-based products (TOPAZ/PIOMAS/GIOMAS/etc.). Obviously for SIC=1 these definitions are the same but not for SIC<1. Another consideration is that if using the “grid-box-mean thickness” definition, SIT will likely be much more correlated with SIC than using the “floe thickness” definition.
  - There is no discussion of how much one would expect agreement between the model and the in-situ observations. In particular the BGEP data are point observations (of ice draft converted to thickness) and are being compared with the modelled thickness in a large grid-cell. Even for the case SIC=1, when the SIT definition issue above is not present, a direct comparison is not obvious. It might be that spatial and temporal averaging of the BGEP data makes it comparable to the mean of the model ITD but that is not discussed. The same is true for the IMBs which will only ever model a single floe. Furthermore there is no discussion of the sampling issues one would expect in the IMB dataset. The IMBs are of course Lagrangian in nature and are permanently attached to the same ice floe. So one would expect changes in thickness to be relatively slow, given that they are purely driven by thermodynamics. Meanwhile the dynamical nature of the Eulerian model could have huge changes in thickness from one time-step to the next. Finally, there are known sampling biases in the IMB dataset that should be discussed. IMBs are normally deployed just before the freeze-up in ice that has survived the summer. Mid-thickness floes are normally chosen – avoiding thick floes for practical reasons and thin floes to limit the chance of losing expensive equipment too quickly.

## Minor comments/concerns

- In many cases results seem somewhat overstated. In particular the performance of BRMT in the East Greenland region, described as 'outstanding', is based upon comparison with only 2 IMB buoys – although the Lagrangian trajectory will include several individual measurements, they will only be of 2 individual ice floes! (NB. the same is true for the forecast results, which are based on only a few forecasts performed in a single year, but this is already mentioned above.)
- Too little information is provided about the model/reanalysis products being used. In particular, what observations are assimilated and what surface forcing is being for the reanalyses. This applies for both the reanalyses datasets in Section 2 and the MITgcm model in Section 3.
- RMSE of SIC is not a very good metric for sea ice forecasts because of the errors in the passive microwave satellite observations. This is particularly true in the summer when the SSMIS cannot distinguish surface melting/ponds from open water. The SSMIS accuracy is also lower in areas of low concentration or thin ice. These points are the motivation for people using sea ice "extent" to compare with satellites.
  - In particular, I would drop the MIZ analysis in 5.2.1/Figure 15 because the SSMIS satellite is likely not able to resolve that. If you redid this analysis using AMSRE2 observations (which is higher resolution and more able to resolve thin/low concentration ice) then the results could be quite different.
- Some of the figures do not bring any useful information and so could either be removed or reformulated. For example, the data in Fig 13 can be understood easily from Fig 12b. Similarly, Fig 14, for which all the panels in look the same, could be improved by changing the model fields on the lower rows to be model-obs differences.

## Typos and technical comments

I attach and annotated version of the original pdf with technical comments.

I do not highlight all instances where the English language needs to be improved only the cases where the language is unclear to the point that the scientific understanding is inhibited.

Please also note the supplement to this comment:

<https://tc.copernicus.org/preprints/tc-2022-92/tc-2022-92-RC2-supplement.pdf>