Second review for The Cryosphere https://doi.org/10.5194/egusphere-2022-214

"Arctic sea ice radar freeboard retrieval from ERS-2 using altimetry: Toward sea ice thickness observation from 1995 to 2021." Bocquet et al., 2022

The authors have taken my comments into consideration and made changes to their manuscript which improve the clarity and presentation.

Another reviewer, Robbie Mallett, spent some time investigating the dataset provided by the authors (which I did not) and identified concerns with the presentation of the results as a "reconciled time series of sea ice radar freeboard". His analysis appeared to show that the corrected ERS2 radar freeboards reflected little of the variability of the underlying raw retracked ERS2 freeboards and the correction could be drawing more information from other inputs to the neural network, such as the ice age or ERS waveform parameters. These concerns were echoed by the third reviewer, who submitted their review after seeing Dr. Mallett's analysis.

In response to Dr. Mallett's concerns, the authors argue that the lack of covariance between sea ice radar freeboards obtained from the same processing chain (TFMRA50) applied to different radar altimeter missions was a known motivation of their study. The different sensing geometries of the altimeters encourage varied impacts of surface roughness, snow properties, etc, on the backscattering of the surface and consequently the retracked heights of floes and leads. Therefore, the correction obtained from the NN is highly nonlinear, so there should be no simple linear correlation expected between raw and corrected freeboards. The partial dependency plot (Fig 2 of the review) they provide illustrates that the corrected freeboards – at least their mean patterns – are sensitive to all the input parameters of the NN, including the raw radar freeboards. However, there are several other parameters, such as the leading-edge slope of the waveform, that show stronger dependency than the raw freeboards.

In general, my view is that the corrected freeboard time series is only as good as its evaluation against independent sea ice freeboard/draft/thickness observations. This is the case for all satellite sea ice freeboard and thickness products, including single sensor products, calibrated multi-sensor records, and proxies for thickness derived from other datasets related indirectly to the ice thickness. In my original reading of the manuscript, I paid careful attention to the Validation section 4.2 and felt that the assessment of Envisat and ERS2 freeboards against various independent datasets was convincing enough to back up the main conclusion of the paper. I.e., that the time series of pan-Arctic radar freeboard volume in Fig 13 has had a significant negative trend since the 1990s. The biases wrt reference observations in Fig 12 are affected by snow loading assumptions but, if anything, imply the negative rfbv trend could be even steeper. However, results from the analysis by Dr. Mallett question whether the full spatial and temporal patterns of the radar freeboard anomalies in corrected grids from ERS2 and Envisat should be trusted.

<u>I would recommend the authors carefully reframe their manuscript to highlight the issues with their time series raised by Dr. Mallett's analysis while emphasizing what they think can be gained from it and why.</u> What should the gridded record of radar freeboard be used for and what do they think it shouldn't be used for by the community, in future studies? Along these lines, I would suggest the authors consider renaming their product a radar freeboard "proxy" from ERS-2 and Envisat (ideally a CryoSat-2 radar freeboard proxy – which relates to another comment by Dr. Mallett on the definition of radar freeboard – but this is another matter..). This change would alleviate the problem with a correction/calibration that is larger in magnitude than the raw freeboard and provide a strong caveat to prevent 'misuse' by the research community when the product is made available. There is absolutely nothing wrong with a freeboard proxy and, in my view, will not diminish the impact of the

obtained climate-relevant trend. (n.b. I also like the addition of the green 'climatology freeboard' line on Fig 13, which reinforces the importance of freeboard/thickness compared to ice area)

I also have a few more comments on points that came up during the review and response:

- Radar freeboard as a geophysical quantity with an uncertainty or not.
  I think this depends entirely on the definition of the radar freeboard, and it has not been consistently defined across different studies, even if some have tried. As soon as the radar freeboard is assumed to represent the mean height of the snow-ice interface within the radar footprint, with some displacement caused by unknown wave propagation delay in snow, then it is a measurable geophysical quantity with a characterizable uncertainty. However, if you do not make geophysical assumptions on the backscattering height represented by the radar freeboard, then it does not have a geophysical uncertainty. It is not necessarily a precise (i.e., repeatable) value though; this depends on the precision of the retracking algorithm which, for waveform fitting methods, may not be insignificant.

- Ideally, we will develop more sophisticated retrackers that can accurately account for sensor-related and mode differences between missions, without needing to statistically calibrate satellite records to one another. This *may* be sufficient to raise the signal above the noise and produce a robust 25-year time series of gridded radar freeboard observations with trustworthy spatial and temporal variations. (It may not, of course, if the 'target'-related noise is too high between sensors, not just the sensor-related noise). However, I consider this calibrated proxy record of CS2 radar freeboard to still be valid within the degree of uncertainty constrained by the validation exercise.

- On the analysis of 20-30% uncertainties in freeboard caused by roughness and whether they are separable from partial penetration.
  These values come from the application of different retrackers to the same CS2 SAR/SARIn mode data, so in reality it is quite an arbitrary representation of the range/freeboard uncertainty, affected by the assumptions and subjective decisions of those applying each retracking algorithm. This uncertainty spread is attributed to roughness because the roughness has a first-order impact on the variability of the retracked range: a larger roughness produces a larger difference between retrackers. The authors are correct that part of this variability in range between retrackers comes from the fact we are uncertain about the radar penetration in snow – is our basic assumption of full snowpack penetration correct? However, this is challenging to evaluate on a pan-Arctic seasonal basis without independent reference observations (although note: Nab et al, GRL, 2023). The most conservative approach would be to independently consider roughness, partial penetration, and SLA-related errors on the freeboard (assuming the final radar freeboard = the ice freeboard), but this will likely overestimate the total uncertainty if numbers from Landy et al 2020 Fig 8 are used.

Unfortunately, I do not have time to provide another thorough review of the revised manuscript with minor/grammatical suggestions. So, I ask the authors to consider my general comments here when reframing the paper. Congratulations again on a really valuable study – it is critical we work towards a long-term record of Arctic sea ice thickness from multi-sensor altimetry that will constrain the thinning(?) trend on a climate-relevant timescale.

Feel free to get in touch if you have any questions, Jack Landy