Earth Syst. Sci. Data Discuss., referee comment RC2
https://doi.org/10.5194/essd-2021-170-RC2, 2021

Comment on essd-2021-170

James Morison (Referee)

Referee comment on "Sea surface height anomaly and geostrophic velocity from altimetry measurements over the Arctic Ocean (2011–2018)" by Francesca Doglioni et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2021-170-RC2, 2021

Review of essd-2021-170, “Sea surface height anomaly and geostrophic velocity from altimetry measurements overthe Arctic Ocean (2011-2018)” by Dogilini et al.

This paper describes the assembly of a high northern latitude, monthly sea siurface height record from CryoSat-2 altimetry. This is a potentially important paper with a worthwhile goal and if sussessful, the data sets could find many users. The effort builds on past work by several groups, but attempts some new approaches. I would probably use the data myself but for several issues.

Probably my biggest concern is the large, 0.2 to 0.5 m, bias between RADS CryoSat-2 heights in the open ocean and the AWI CryoSat-2 heights in the ice-covered Arctic Ocean. In part this may be a natural difference because it is apparently mainly across the East Greenland Current and the two data sources have no overlap. In our work we have made the same comparison, but with slight overlaps between the two sets of data and found the RADS heights were only 10-cm lower than our sea-ice product. These authors do not tell us what is done to reconcile the large RADS v. AWI product bias, and the uncertainty due to it essentially renders their gridded product unusable in the regions that it would otherwise be most useful. The authors could at least determine if the large bias is due to the large separation of the component data sets by making comparisons with dynamic heights and GRASCE bottom pressure over a much larger area as suggested in my detailed comments.

A secondary issue is the treatment of the data to eliminate striping that the authors attribute to sub-monthly variability coupled with the to me novel DVA "poorman's" interpolation scheme. We see the striping in all altimetry products pretty much the same worldwide except for an increased prevalence of striping at lower latitude. As discussed in the detailed comments, the striping or "trackiness" is due to sparseness and inherent uneven distribution of ground tracks particularly for averaging times less than the repeat period of the satellite and would be present with no time variability in the height field at all. The method used by the authors to eliminate the striping is interesting but I think it probably just spreads out an artifract of uneven sampling and it's mix of spatial interpolation and time averaging as a measure of uncertainty is unclear.

This issue is made all the more unclear by the lack of an understandable explanation of the Poor Man's interpolation scheme. I have read that it is a mix of nearest-neighbor and linear interpolation, and I think it could be very useful in altimetry, but the authors don't clearly describe it or the uncertainty that goes with it.

As much as I appreciate the great effort and original thinking the authors have put into this, I can't recommend publication of the paper, nor would I use the database until the issues above are addressed. I just have too many questions. I might even suggest that doing less would be better. Everybody has their own favorite interpolation scheme., so I think the authors database would be useful and more popular if they could simply resolve the RADS v. AWI bias issue with more comparisons with in situ hydrography and GRACE OBP, explain what they do to remove the bias, and then simply bin average the data and corresponding locations in their grid cells. Then as a companion effort, compare different interpolation schemes applied to the bin averaged data and give a do it yourself recipe for each. It would be great to see a comparison of linear, optimal, and poor man's interpolation on data we all care about.

Detailed Comments:

L20 – Minor point, but not all the decline is due to atmospheric warming, A major portion is likely due to enhanced export from the Arctic

L25 – The basic physics of air-ice-ocean interaction are well understood. See [*McPhee*, 2008]. In simple terms the geostrophic velocity of the ice and upper ocean are the same and except at shorter time scales ice velocity is largely geostrophic.

L51-54 - Not true. See Morison et al., (2012, 2019, 2021)[*Morison et al.*, 2021; *Morison et al.*, 2018; *Morison et al.*, 2012]

L105-109 - Morison et al. (2018) find overlap of RADS with the DOT of Kwok and Morison (2011) good down to 15% ice concentration. In the overlap regions RADS is found to be 10 cm lower arguably due to a difference in retrackers. Note, at the time we found it important to use RADS SSH not DOT because the geoid RADS used for their DOT was relative to a different ellipsoid than their SSH.

L125-155 - Why no Canada Basin moorings, e.g., BGEP moorings. Using only data from these few moorings severely limits the validation points. Morison et al. (2012) uses repeat hydrographic station steric heights and GRACE ocean bottom pressure change to validate trends in DOT from ICESat. However, repeat hydro stations are not necessary. Morison et al, (2018) compare among open-ocean and ice-covered-ocean ICESat and CryoSat2 against dynamic heights + GRACE OBP at all available hydrographic stations (including Argo) in the Arctic Ocean and Sub-Arctic Seas. They find that when altimetry DOT is plotted versus in situ dynamic height + OBP, the points tend to fall along a common line as long as there are no biases. For example the 10 cm bias between open ocean and sea ice CryoSat2 derived by comparison in overlap regions shows up clearly when the altimeter and hydro station data are plotted as DOT v DH+OBP. For the data of this study where there is no overlapping data to compare, the DOT v DH+OBP comparison would be the only way to properly determine the instrumental bias. Further, any conclusions about circulation would be consistent with the dynamic method of determining circulation from in situ hydrography alone.

L195-207 - This description of lead-open ocean bias is problematic. The bias shown in the figure is large, 0.2 to 0.6 m. Why is it this large? A couple of things pop out at us. There is no overlap and indeed there is a big gap between the AWI and RADS CryoSat-2 data segments. Apparently there are no AWI results for ice concentration less than 70%. Why is this? Especially with this big gap in the ice concentrations covered by the two data sets, the region shown for comparison must be the worst in the whole Arctic Ocean and Sub-Arctic Seas region for making this comparison because it straddles the most energetic current, the East Greenland Current with the biggest natural gradient in DOT. The natural drop in DOT across this feature is probably not 0.4 m but could account for a lot of it. Were other areas included in the along track comparisons? What was done to remove the bias? The text gives no information on this.

We combined RADS CryoSat-2 open water data with our own CryoSat-2 DOT in sea ice (Morison et al., 2018; Morison et al, 2021). We found narrow overlapping regions for these where RADS averaged only 10 cm below the sea ice CryoSat-2 DOT, a result similar to the same comparison in the Antarctic (Armitage et al., 2018). We raised the open ocean values 10 cm. This was consistent with comparing satellite DOT and in situ dynamic heights plus ocean bottom pressure (Morison et al., 2018).

L195-207. This is likely unrelated but the 40 cm average bias is reminiscent of a problem we had comparing the sea ice DOT with open ocean DOT directly from RADS. As I recall RADS used EGM2008 in their DOT computation but by comparing with their raw SSH, we determined that at the time they had used a tide-free EGM2008 relative to the TOPEX/Poseidon ellipsoid rather than a mean-tide EGM2008 relative to WGS84. This resulted in partially offsetting biases (T/P v WGS84 = 71 cm. and tide-free v mean tide = -25 cm for a 45 cm error in DOT. It sounds like this study used RADS SSH and anomalies about a mean SSH, but some results are in terms of DOT, Can it be confirmed that these are in the mean-tide (permanent-tide) system?

L219 – This is a good question that indeed could use more attention. On the basis of studies by Robbins (Robbins et al., 2016), ICESat-2 uses DAC in the open ocean and IB in ice covered seas.

L221-224, Good idea. Robbins et al. (2016) also found DAC and IB to be different in the East Siberian Sea.

L231-232 - To claim consistency of the two data sets we need to resolve the bias issue with respect to Figure 2. In Figure 3, it looks like there is about a 10 cm difference between the ice-free and ice-covered parts. This again raises the question: what was done about the average 40-cm bias?

L252-254 - A latitude-longitude grid? This grid must be coarse at 60N and fine at 88°N. Why use that and not a rectangular grid such as a polar stereographic grid? The standard NSIDC polar stereographic grid extends down to 60°N. If that was used, we could easily compare your DOT to all the NSIDC products.

L258-263 – This is an interesting and perhaps useful idea, but a simple thought experiment shows that the streaks appear even when there is no time variation in the natural field; streaks appear primarily because the data are spatially sparse and for averaging times less than the repeat of the satellite they are unevenly space in streaks. For example, if we gridded up 1 day of a static SSH field, and CryoSat-2 makes 16 orbits per day, we would see 16 streaks where orbits crossed grid cells. Most grid cells shouldn't show any data. With 0.75° longitude grid spacing it would take 2 weeks of orbits to have one pass, ascending or descending, per grid cell. And even then, streaks would appear unless the orbit had a 15-day repeat because the coverage of the ground tracks is uneven and streaky until the orbits start to repeat. Interpolating short time segments of data will smooth out the streaks but will be falsely adding data coverage where there is none in order to smear out the streaks. I think that is probably misleading.

L267-269 - How can we say, " variational inverse method that we use here is an alternative which does not assume an a priori behavior of the analyzed field." and then turn right around and say, " minimizing a cost function dependent on the data and few input parameters". In principal, when using optimal interpolation, the data tell the level of noise and correlation length. Here L and lambda are essentially tuning parameters chosen to optimize a cost function that we make up and applied to comparisons to a select bit of in situ data. Then when we find the cost function is mostly independent of L and lambda, we pick the values that we like best based on other criteria. The variational inverse method seems a lot more subjective and arbitrary than optimal interpolation.

L276-278 - Sorry, the interpolation technique in simple terms. Interpolation involves weighting observations usually near the location of the point we are interpolating to, but I can't find how the weightings are determined by L and lambda.

L314 – For those of us unfamiliar with it, can the "poorman's" estimate and the rationale for its use be explained in simple terms? I have read of the poor man's interpolation using a combination of linear interpolation and nearest neighbor interpolation of segmented data where there are problems with linear interpolation, but sea surface height is not segmented. It would be helpful to describe the method and give a better rationale for its use even in favor of linear interpolation. Without these, a user of the data is not going to have confidence that there aren't strange artifacts of the interpolation method in the results.

L315-335 - This is an unusual but interesting approach. Standard error ~ standard deviation of samples divided by the square root of the number of samples assumes the samples are independent, so this sounds good for observational uncertainty, but figure 4 suggests this isn't true for the "sub-monthly component". More importantly my comment regarding L258-263, would suggest the sub-monthly component is an artifact of the treatment of spatially sparse data and has no clear relation to uncertainty.

L327-332 - Another reason to use a rectangular grid.

L338 – Ahh, now using good old linear interpolation. I know how that works. Rhetorical question: what happened to “poorman's” interpolation?

Figure 6a - it is almost impossible for me to gain much of a sense for the data looking at height anomaly for one month. For me it would be better to look at the dynamic ocean topography because we have some idea of what that should look like.

L379-384 - The relative error exhibits a strikingly symmetric latitudinal dependence due to the convergence of ground tracks farther north, but one wonders what part the latitude-longitude grid plays. Would the pattern look less like a bull's eye if a rectangular grid were used?

Table 5 and Figure 8 - What are the RMS differences between the altimetry and in situ height anomalies?

501-508 - This is a serious flaw associated with the limited extent of the sea-ice region CryoCsat2 data, but it could be helped by better characterizing the bias between the sea-ice and RADS heights using large numbers of hydro station dynamic heights plus GRACE OBP (Morison et al., 2018)

Discussion - One could compare with winter DOT from ICESat, CryoSat-2 in Morison et al., 2012, 2018, and 2021.

References

Armitage, T. R. Kwok, A. F. Thompson, and G. Cunningham,

“Dynamic topography and sea level anomalies of the southern ocean:

Variability and teleconnections,” J. Geophys. Res., Oceans, vol. 123,

- 1, pp. 613–630, 2018.

McPhee, M. G. (2008), *Air-Ice-Ocean Interaction: Turbulent Ocean Boundary Layer Exchange Processes*, Springer.

Morison, J., R. Kwok, S. Dickinson, R. Andersen, C. Peralta-Ferriz, D. Morison, I. Rigor, S. Dewey, and J. Guthrie (2021), The Cyclonic Mode of Arctic Ocean Circulation, *Journal of Physical Oceanography*, *51*(4), 1053-1075, doi:10.1175/JPO-D-20-0190.1.

Morison, J., R. Kwok, S. Dickinson, D. Morison, C. Peralta-Ferriz, and R. Andersen (2018), Sea State Bias of ICESat in the Subarctic Seas, *IEEE Geoscience and Remote Sensing Letters*, *15*(8), 1144-1148, doi:10.1109/LGRS.2018.2834362.

Morison, J. H., R. Kwok, C. Peralta-Ferriz, M. Alkire, I. Rigor, R. Andersen, and M. Steele (2012), Changing Arctic Ocean freshwater pathways, *Nature*, *481*(7379), 66-70, doi:http://www.nature.com/nature/journal/v481/n7379/abs/nature10705.html - supplementary-information.

Robbins, J. , T Neumann, R. Kwok, J. Morison, 2016. ICESat-2 Oceanic and Sea Ice Response to Atmospheric Forcing , poster presentation at the 2016 Fall AGU Meeting