**Authors’ replies to**
**“Comment on essd-2022-111**
**Anonymous Referee #1”**

*Referee comment on "Sea surface height anomaly and geostrophic current velocity from altimetry measurements over the Arctic Ocean (2011–2020)" by Francesca Doglioni et al., Earth Syst. Sci. Data Discuss., https://doi.org/10.5194/essd-2022-111-RC1, 2022*

---

*The authors constructed a monthly mean sea level and dynamic ocean topography dataset using data from CryoSat-2 observations. The main differences from previous studies (but not all) are that the data construction focused on intra-month variability, the handling of data bias in the sea ice region and open water, and the validation using time series data from mooring system observations.*
*Correlation coefficients and RMSDs are calculated using the CPOM DOT as a comparison, but the differences with the CPOM DOT do not lead to the conclusion that the authors' product is truly better. In particular, the sensors in the mooring system that acquire insitu data are discrete in the vertical direction, and the steric height based on the linearly interpolated vertical profile is questionable. Also, the baroclinic Rossby radius is shorter than the L=300 km used in the first step of DIVA, which may be too much smoothing. Other comments are presented below.*

We are grateful to the reviewer for the constructive comments, which helped to improve the manuscript and the validation of our gridded product. Please find here few general remarks on the revision, and further below our elaboration on each comment.

In order to provide more evidence on the quality of the altimeter data used as source data and of our final gridded product, we added two pieces of analysis. First, we gave an overview on the statistics for the AWI and RADS along-track datasets and the merged dataset, furthermore including the latter in the file deposited by PANGAEA. Then, in order to extend the assessment of our gridded product to the western Arctic, we included comparisons with data from the Beaufort Gyre Observing System moorings, part of the Beaufort Gyre Exploration Program (BGEP[1], hereafter used as acronym for the moorings) and from two moorings in the central and eastern Chukchi Sea. We explored the possibility to compare to further data in the Chukchi Sea and Bering Strait but these time series were either too close to the coast to find good comparison points in the altimetry maps, or too short to have significant comparison, or very difficult to get at all.

We also revised our processing of in-situ data to evaluate possible errors deriving from it. We evaluated the reliability of our sea surface height from in-situ data in Fram Strait by comparing these data to continuous hydrographic profiles (for the steric height) and to a nearby bottom pressure recorder (for the bottom pressure equivalent height). We came to the conclusion that most of the variability is resolved, therefore we did not exclude these data from the validation. Detailed documentation is given below. We also found, though, a mistake in the computation of monthly means of in-situ data. When this was corrected, correlations between altimetry and in-situ sea surface height improved, with the greatest improvement in the Fram Strait. Finally, we adjusted the averaging depth of in-situ velocity data in order to exclude the surface Ekman layer, with negligible impact on the results.

We clarify in our replies to your comments (10) and (12) what determines the actual resolution of our dataset. We explain that this is not related to the 300 km radius used in the first gridding step with DIVA, but to the 50 km radius used in the second and final step.

Regarding the comparison with the CPOM DOT, from our point of view, as discussed in our detailed replies, altimetry products for the Arctic Ocean are currently still at the stage of research. For this reason, it is rather useful to have a diversity of products, developed with different methodologies, that can be compared to understand how the quality of these datasets can still improve. We find this approach to be quite common, for instance, in the sea ice community for products of ice velocity, ice thickness, ice concentration.

---

*Major comments:*

*1) Mooring data1*

*The authors state that they are generating sea level data for the entire Arctic Ocean, but the only mooring system data used as validation are those installed in the Fram Strait and Laptev Sea. For example, WHOI has deployed BGOS mooring systems with MMP in the Canadian Basin. If time series data are important, then all available mooring system data should be used.*

Following your advice, we substantially extended our assessment of altimetry-derived sea surface height and geostrophic velocity fields to the western Arctic by comparing to additional mooring data in the Beaufort Sea and Chukchi Sea and hydrographic profiles in the Amerasian and Eurasian basins (see additions in sections 3, 5.2.1 and 6.2).

Regarding the assessment of sea surface height, first, we compared our sea surface height product to time series of the sum of steric height and bottom pressure equivalent height measured at the BGEP moorings A and D in the Canadian basin. Then, we compared our monthly fields to monthly estimates of the sum of steric height from in-situ hydrographic profiles, distributed in the deep basins, and bottom pressure equivalent height from GRACE. Geostrophic velocity was further compared to near-surface velocity from the BGEP moorings A, B and D and moorings S1 and S3 in the Chukchi Sea. In agreement with the comparisons previously done in the Eurasian Arctic, the additional comparisons show that, while there is reasonable agreement between our product and in-situ data at seasonal and longer time scales, significant differences are observed at monthly time scales.

We believe that these comparisons to data in western Arctic, in addition to the ones we conducted with mooring data at three locations in Fram Strait and the Eurasian Arctic (i.e., Arctic Cape and Laptev Sea; see Fig 9), provide an overview of the capability of our satellite product to reproduce in-situ measured sea level variability for a large portion of the Arctic.

Despite differences in the resolution and the nature of in-situ and remote sensing measurements, our gridded sea surface height significantly correlates with mooring data from the ice-covered Arctic and its boundaries, showing correlation coefficients ranging between 0.5 and 0.9. Furthermore, our work entails also the generation and assessment of geostrophic velocity, which now includes comparison with data from nineteen moorings and a discussion of its realism in the spatial and temporal patterns in regions of different dynamical regimes.

*2) Mooring data2*

*The authors calculate steric height using vertically discrete water temperature and salinity from mooring system observations. In the Arctic Ocean, where tilt pressure structures dominate, it is questionable whether the authors' method can correctly determine Steric height. Do the linearly interpolated vertical profiles of temperature and salinity reproduce the results of CTD observations made at the same time?*

We agree that linear interpolation of curved profiles will introduce a bias. However, given that the comparisons between mooring time series and altimetry were aimed at assessing the *temporal variability* of sea level, we tested here how the *variability* of the steric component is affected by the reconstruction of continuous vertical profiles with two different methods (linear and spline, see below). The results of this test showed that most of the variability is recovered using linear interpolation, and that we explained a higher portion of the variability by applying linear interpolation than splines. Following this analysis, we kept our method to compute steric height unchanged. We clarified in the manuscript in ll. 224-229 what is the uncertainty introduced by this approach.

We tested our approach using 432 continuous CTD profiles, deeper than 800 m, collected from the central Fram Strait over 16 years. Following the case of our comparison with mooring data in Fram Strait, we subsampled the profiles at depths of 50 m, 230 m and 720.

Fig R2 shows three series of steric height anomaly computed with different interpolation methods. We see there that the variability of steric height is still well represented by interpolated profiles. Linear interpolation performs better than spline interpolation, with 88% of explained variability. Even though we cannot assume that this result would be the same at all locations where we compared altimetry to in-situ data, the profiles in the Fram Strait were the most critical ones, with only 3 tie points. This suggests that vertical interpolation between discrete measurements can be applied with reasonable confidence as well at the Arctic Cape, in Laptev Sea and in the Beaufort Sea.

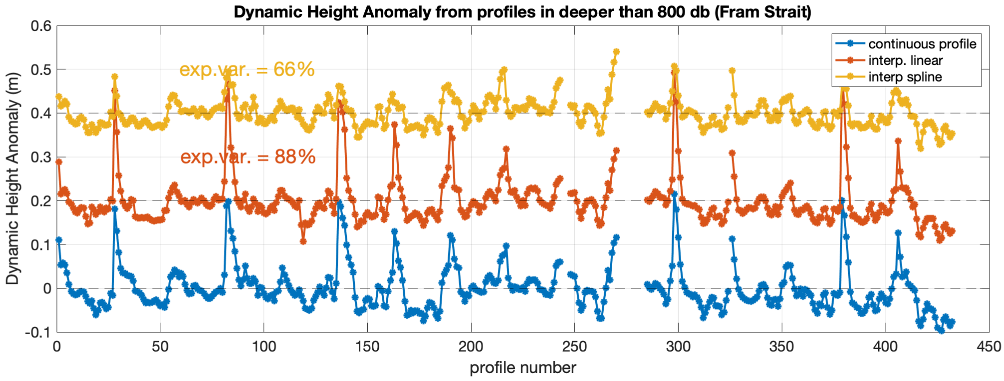

*Figure R 1: series of steric height anomaly computed from the 432 profiles using the continuous profiles (blue), linearly interpolated profiles (red, shifted by 0.2 m) and spline interpolated (yellow, shifted by 0.4 m) profiles.*

*3) Reference ellipsoid (line 82 "e.g., WGS84".)*

*From Figure 6d, I see that you used WGS84 instead of TP ellipsoid. Please specify somewhere that WGS84 was used in this study, not "e.g."*

Thank you for pointing that out. The description in section 2 is meant to be theoretical, therefore we changed now the sentence in ll. 82 to include the names of both ellipsoids. Regarding the DTU17MDT field used in this study, we added reference by personal communication with Per Knudsen (ll. 126-127), as we did not find written citable information on the reference ellipsoid used. However, when computing MDT it is usually ensured that the reference system for MSS and geoid is the same, so that the MDT difference will be in-sensitive to the choice. Furthermore, the difference between the T/P and the WGS84 ellipsoid over the Arctic Ocean consists basically of a spatially constant offset (Skourup et al. 2017), which has no impact on ocean dynamics.

*4) Sea ice concentration (line 107)*

*Isn't the 15% sea ice concentration a threshold to avoid so-called pseudo sea ice that misinterprets water vapor as sea ice, for example in the Bering Sea in summer, and doesn't it need to be 15% in the Arctic Ocean? Wouldn't the results be the same if this threshold were set to, say, 5%, 10%, or 20%? Isn't it usually the Waveform, for example Pulse Peakiness, that determines if it is sea ice or sea surface?*

We clarified in the text in ll. 112-114 why there is the need to choose a threshold between ice-covered and ice-free regions.

The 15% threshold is set for retracking of Cryosat-2 data over ice because below this ice concentration the ice is very sparse and the surface type classification (water, ice) is subject of high uncertainties in these areas (see Ricker et al. 2017, and Hendricks et al., 2021). The results would be very similar if a threshold of 20% was used, but in our case the 15% threshold fits well because it coincides with the uppermost limit in ice concentration set for the RADS dataset.

*5) ADCP (line 180)*

*ADCP velocity data is averaged in the upper 50 m. Shouldn't the velocities within the surface Ekman layer be excluded? Also, there must be a momentum flux due to sea ice movement, so shouldn't the surface still be excluded?*

Previous studies indicate that in the Arctic Ocean the Ekman layer extends approximately down to a depth of 20 m (e.g., Hunkins 1966; McPhee 1992; Cole et al. 2014; Peterson et al. 2017). In our previous manuscript version, we were already excluding data from the upper 10 m at the Laptev Sea continental slope. In the Fram Strait comparison, we used data from 75 m depth in order to avoid discontinuities in the timeseries when Current Meters were substituted with ADCPs in the later part of the time series. In order to account for the above comment, we adjusted the averaging depth at the Laptev Sea continental slope to the range 20-50 m, and included a comment about the averaging depth in ll. 200-201. The same averaging range was used for the ADCPs mounted on the BGEP moorings. This modification, however, did not lead any to significant changes in our results.

*6) High-pss filter(line 205)*

*Is there evidence that the high-pass filtered data is valid? If no, should it be excluded from the validation data?*

We did not exclude this timeseries from the validation data. The arguments presented here below to support this choice have been summarized in the manuscript at ll 234-238. While some variability is indeed missing in the in-situ data at this location, we can find strong evidence that the high pass filtered data is reliable (see below). We believe therefore that the in-situ timeseries at the mooring FS_S still includes a large fraction of the total variability. First, as we pointed out in our reply to your comment (2), the steric variability is still well represented at this location. Furthermore, Quinn and Ponte (2012) showed that the coherence between satellite observations of sea level and ocean bottom pressure is highest at timescales shorter than about 2 months, because of the presence of large wavelength – high frequency barotropic waves.

To test whether the high-pass filtered data is reliable, we considered the fact that high frequency variability in ocean bottom pressure decorrelates over large distances (e.g., Peralta-Ferriz et al., 2011). Therefore, we compared the time series at FS_S with the one from a bottom pressure sensor installed on another mooring 150 km away in the same basin (the bottom depth at the two moorings being 3012 m and 2778 m). After high-pass filtering the time series from the two moorings with various cut-off frequencies, we computed correlation and significance. The results are displayed in Table R1.

| **cutoff** | **NO filtering** | **4 months** | **2 months** | **1 month** |
|---|---|---|---|---|
| **Correlation coefficient** | -0.54 | 0.13 | 0.57 | 0.73 |
| **p-value** | 0.88 | 0.16 | **<0.01** | **<0.01** |

*Table R 1: correlation between ocean bottom pressure at mooring FS_S and a nearby mooring (not used in manuscript) high-pass filtered with different cutoff frequencies.*

The correlation coefficient increases by filtering out long time scales and it is significant for thresholds below 4 months, with the best agreement on sub-monthly timescales. This gives us confidence that the high frequencies are not affected by instrumental noise and can be used to validate the month-to-month sea level variability due to mass contributions.

*7) Specific volume anomay*

*The reference depth the authors used is 400 dbar. Why? It is the depth of upper Atlantic Water. R. Kwok and other scientists used about 750 dbar (deepest depth of ITP).*

We thank the reviewer for pointing to this important issue. The main component of steric height changes on timescales up to a few years are in the so-called "Polar Mixed Layer" (PML, e.g. Korhonen et al., 2013), that part of the Arctic water column that may be seasonally mixed during ice formation and concurrent brine rejection. This layer is bounded underneath by the so-called "lower halocline", which is denoted approximately by the 34-isohaline (e.g. Korhonen et al., 2013). The lower halocline acts as a barrier between the PML and the deeper-lying warm Atlantic Water, which is largely isolated from surface influence in much of the Arctic basins. The lower halocline, and thus the Polar Mixed Layer, resides within the top 200 m across the Arctic. Hence, using the depth-range of 0-400 m should capture most of the steric height variability up to decadal timescales. This approach has been used in several studies, e.g. on Arctic freshwater (Rabe et al., 2011; 2014). Even though this range probably does not capture all multidecadal changes associated with changes in the Atlantic Water inflow, the dataset we're using to assess altimetry does not cover those timescales. Thus, a comparison between the steric height derived with this procedure and the altimetry data is appropriate.

Another reason to discard deeper data when studying variability is that much of the data over the time period under study here has been measured by WHOI-ITP (autonomous CTD profilers). These systems have to be quality-controlled by applying a conductivity correction. The only way to achieve that, for instruments that are usually not recovered, is to compare to historical data. The range used for that is 400-800 dbar (their profiles usually reach to a maximum of 760 dbar). Even though the historical reference data field is updated in time, it is not likely that variability deeper than about 400 m is captured by ITPs (see also Sumata et al., 2018). Furthermore, other studies assessing altimetry data used in the past steric height from hydrographic profiles integrated to a reference pressure shallower than the full ITP depth (760 db), e.g. 500 db, for instance Kwok et al. (2011), Morison et al. (2011), Armitage et al. (2016).

Following our reasoning above, we have modified the manuscript in ll. 265-267 to explain what component of the variability we expect to capture integrating over this depth range.

*8) Figure 2b*

*The INSET PANEL in Figure 2b is difficult to understand. How about color-coding the grid points by AWI and RADS?*

We changed Figure 2 by including two panels with color-coded grid points, green for AWI data and blue for RADS data.

*9) FIgure 4 inset panel*

*The color of the inset panel in Figure 4 indicates the number of crossovers, which basically increases as you go to higher latitudes since these are polar-orbiting satellites.*
*If this inset panel is independent and the color of each season indicates the difference of η at the crossover point, it is easy to understand where and in which season there is a difference.*

We are grateful for the suggestion but decided not to change the content of Figure 4. Even though we understand that there is an interest to know more about the error budget in the different seasons, our main aim in the inset of this figure is to show what the statistics for the error calculation is. To make this more explicit, we added one sentence in ll. 334-335 specifying the total number of crossovers used in this analysis. We will take however your comment in consideration for future investigation.

*10) Spatial resolution*

*Is the resolution setting just following the CPOM DOT, or if you want to differentiate yourself from CPOM, is there some strategy to change the resolution? At the moment, it is no different from CPOM except in the Siberian Sea.*

We consider that the dataset resolution is set by the data coverage and the decorrelation length used in the interpolation. In our case, the formulation of the DIVA gridding method is derived from a continuous equation, so the solution depends only on the decorrelation radius and the data density, regardless of the chosen grid, as long as this has a fine enough resolution. Even though the output grid for our dataset and the CPOM DOT has the same resolution, we used a radius of 50 km in the final gridding with DIVA while Armitage et al. (2016) smoothed the CPOM DOT with a larger radius of 100 km. This introduces a difference in the resolution between the two datasets. We additionally provide an analysis of the spatio-temporal resolution based on comparison with in situ data (see sections 5.2.2b and 6.3). From this analysis emerges, in summary, that our monthly geostrophic velocities can resolve seasonal to interannual variability of boundary currents wider than about 50 km.

Furthermore, our work differs from the what done in Armitage et al. (2016) in several aspects that go beyond the product resolution. We clarified these aspects in section 6.1 of the manuscript and used them to guide our discussion. A summary is provided in the list below:

- the source data (ellipsoidal heights from CryoSat-2) used in ice covered areas, which have been derived using different algorithms;

- the approach used to correct and/or minimise the impact of unresolved high frequency variability due to wind and tidal forcing and the estimate of its impact on the error;
- the interpolation method.

In addition, we extended the number of regions and dynamical regimes covered by the validation, which includes now both the eastern and western Arctic circulation regimes, the central Arctic Ocean, Arctic shelf seas and the main exchange gateways of the Arctic. This provides an improved analysis and validation of the realism in the spatial and temporal patterns of our dataset with respect to the work done by Armitage et al. 2016 and 2017.

Given that altimetric data products for the Arctic Ocean are currently still at the stage of open research, rather than development and improvement of a mature product, we think it is not appropriate to benchmark these products against each other. We rather think it is useful to compare them and find how they differ to understand potential uncertainties in the data. It is quite common in other areas of research, for instance in the sea ice community, to have a set of products (e.g., sea ice concentration or sea ice thickness), developed with different methodologies by different teams. Products can have different strengths also depending on the application. In future developments, methods and approaches from different products might find their way into a new, more mature DOT product.

*11) lines 358 & 360*

*I don't understand it because there is no detailed explanation of where the 4.2 cm and 8.2 cm came from.*

Thank you for pointing to this ambiguity. Regarding the 4.2 cm, this was a previous estimate of the error on the along-track data, which should have been updated in the current version to 3 cm (see ll. 339 of the revised manuscript); we corrected in this paragraph the number to 3 cm (ll. 403). The quantity 8.2 cm is instead our estimate of $\sigma$, which is then used to compute the signal to noise ratio used in DIVA (given by $\sigma^2/\varepsilon^2$). This estimate was derived by taking the data signal ($\sigma^2$) equal to the average spatial variance of weekly subsets of along-track data in the period 2011-2020. We agree that the formulation in the previous manuscript was not clear and rephrased this sentence to explain where this number comes from (ll. 404-405).

*12) Local gradients between 7W and 4W (line 462)*

*The authors used L=300 km, so they just applied too much smoothing. Why not take into account the bathymetry and reduce the value of L if there are steep velocity changes in the horizontal direction, for example?*

Regarding your first comment, we believe there is a misunderstanding about the meaning of the decorrelation scales used for the first and second interpolation steps, and would like to clarify that here below.

We used the decorrelation scale of L1=300 km to compute the background field (first step of interpolation). This step was necessary so that we could remove the mean field from the data and obtain anomalies relative to the period covered by the data. In the second step, where we apply the final gridding of the anomalies and obtain the final monthly fields, we used a decorrelation scale of L2=50 km. The scale used to compute the background field (L1) does not influence the smoothness of the final monthly fields, which is instead related to L2 and to the signal-to-noise ratio. The scales resolved in the dataset are further discussed in sections 5.2 and 6.3, where it results that the dataset does not resolve variability at scales shorter than L2=50 km (e.g. ll. 751-754). We added few sentences in the method section to clarify what above and furthermore highlight that a decorrelation scale of 50 km was chosen as the shortest radius possible based on the number of data points available to constrain the interpolation (ll. 389-394).

Regarding your second question, we agree that probably a spatially homogeneous decorrelation radius for the whole Arctic region is not the most accurate approximation. However, there is to date not enough knowledge to determine an appropriate alternative. We tested a built-in option in DIVA to apply anisotropic correlation length, where the anisotropy was based on a long-term average geostrophic velocity field from model output. This test however showed no significant change in the monthly fields. For this reason, and given that we do not know the accuracy of the model field used for anisotropy, we decided to exclude this feature in the final approach.

*13) Figure 9a*

*I guess relatively low correlation is due to incorrect steric height based on linear interpolation.*

In the previous version of the manuscript, we unfortunately made a mistake in computing the monthly averages of in situ SSH. This was actually in part the reason for the low correlations. As you can see in the updated manuscript, in Fig. 9a the correlation coefficient now reaches 0.7, which we do not consider as low. We agree on the fact that part of the variability is not captured by the interpolated in-situ data at this location, for the reasons discussed following your comments (2) an (3). We think, however, that data from this mooring still resolve most of the variability, one clear example being the steric height seasonal cycle. Therefore, as discussed under comment 2, we consider the interpolated mooring data a valid dataset in the context of our evaluation, given also the scarcity of data available in the Arctic Ocean.

*Minor comments:*

*Line 82 : e.g, --> e.g.,*

Changed

*Line 129: CPM --> CPOM*

Changed.

*Table 3: The table shows FES2014, but the caption describes it as FES2004.*

Changed.

---

References:

[1]Beaufort Gyre Exploration Program: https://www2.whoi.edu/site/beaufortgyre/

Armitage, T. W. K., S. Bacon, A. L. Ridout, S. F. Thomas, Y. Aksenov, and D. J. Wingham, 2016: Arctic sea surface height variability and change from satellite radar altimetry and GRACE, 2003-2014. *Journal of Geophysical Research: Oceans*, 121, 4303–4322, https://doi.org/10.1002/2015jc011579

Cole, S. T., M.-L. Timmermans, J. M. Toole, R. A. Krishfield, and F. T. Thwaites, 2014: Ekman Veering, Internal Waves, and Turbulence Observed under Arctic Sea Ice. *J Phys Oceanogr*, 44, 1306–1328, https://doi.org/10.1175/jpo-d-12-0191.1

Hunkins, K., 1966: Ekman drift currents in the Arctic Ocean. *Deep Sea Res Oceanogr Abstr*, 13, 607–620, https://doi.org/10.1016/0011-7471(66)90592-4

Korhonen, M., B. Rudels, M. Marnela, A. Wisotzki, and J. Zhao, 2013: Time and space variability of freshwater content, heat content and seasonal ice melt in the Arctic Ocean from 1991 to 2011. *Ocean Sci*, 9, 1015–1055, https://doi.org/10.5194/os-9-1015-2013

Kwok, R., and J. Morison, 2011: Dynamic topography of the ice-covered Arctic Ocean from ICESat. *Geophysical Research Letters*, 38, 1–6, https://doi.org/10.1029/2010gl046063

McPhee, M. G., 1992: Turbulent heat flux in the upper ocean under sea ice. *J Geophys Res Oceans*, 97, 5365–5379, https://doi.org/10.1029/92jc00239

Morison, J., R. Kwok, C. Peralta-Ferriz, M. Alkire, I. Rigor, R. Andersen, and M. Steele, 2011: Changing Arctic Ocean freshwater pathways. *Nature*, 481, 66–70, https://doi.org/10.1038/nature10705

Peterson, A. K., I. Fer, M. G. McPhee, and A. Randelhoff, 2017: Turbulent heat and momentum fluxes in the upper ocean under Arctic sea ice. *J Geophys Res Oceans*, 122, 1439–1456, https://doi.org/10.1002/2016jc012283

Rabe, B., M. Karcher, U. Schauer, J. M. Toole, R. Krishfield, S. Pisarev, F. Kauker, and T. Kikuchi, 2011: An assessment of Arctic Ocean freshwater content changes from the 1990s to the 2006–2008 period. *Deep-Sea Research Part I*, 58, 173–185, https://doi.org/doi:10.1016/j.dsr.2010.12.002

Rabe, B., and Coauthors, 2014: Arctic Ocean basin liquid freshwater storage trend 1992–2012. *Geophys Res Lett*, 41, 961–968, https://doi.org/10.1002/2013gl058121

Ricker, R., S. Hendricks, L. Kaleschke, X. Tian-Kunze, J. King, and C. Haas, 2017: A weekly Arctic sea-ice thickness data record from merged CryoSat-2 and SMOS satellite data. *The Cryosphere Discussions*, 11, 1607–1623, https://doi.org/10.5194/tc-11-1607-2017

Skourup, H., and Coauthors, 2017: An Assessment of State-of-the-Art Mean Sea Surface and Geoid Models of the Arctic Ocean: Implications for Sea Ice Freeboard Retrieval. *Journal of Geophysical Research: Oceans*, 122, 8593–8613, https://doi.org/10.1002/2017jc013176.

Sumata, H., and Coauthors, 2018: Decorrelation scales for Arctic Ocean hydrography – Part I: Amerasian Basin. *Ocean Science*, 14, 161–185, https://doi.org/10.5194/os-14-161-2018