



Comment on **essd-2021-123**

Anonymous Referee #2

Referee comment on "Arctic sea surface height maps from multi-altimeter combination" by Pierre Prandi et al., Earth Syst. Sci. Data Discuss.,
<https://doi.org/10.5194/essd-2021-123-RC2>, 2021

General comment:

The proposed study describes the generation of a multi-mission altimetry 3-day resolution dataset for the Arctic Ocean. The work combines different techniques concerning waveform classification, retracking and grid interpolation. The title of the study sounds very interesting and motivates to learn more about the combination of different altimeter missions in a very challenging area. However, following the remarks of referee 1, the paper lacks much detailed information on the individual processing steps. I would like to see this study published, but some points are not clear to me or need more explanations. Therefore, I decide for a major revision. Based on the comments of referee 1, I have additional comments, structured in general and section specific remarks.

General:

- What is the deeper motivation behind that study? What are the major research questions? It would be useful to indicate a bit more why you have generated this dataset.
- I recognize the motivation for high temporal resolution by using multiple altimetry missions at the same time. However, this only holds for areas between 50°N and ca. 81°N. At higher latitudes there is only Cryosat-2. How can you be sure that your multi-mission solution performs better in areas, which are only covered by one single mission?
- Referring to the comments above, I think the last sentence of the Abstract (L6-8) must be rephrased.
- I'm generally missing some key numbers about the dataset performance in the abstract. It would be nice for the reader to have some key numbers directly visible in the abstract.
- Another general comment is related to time span of the dataset. Why does the dataset end in April 2019? It would be nice to have 3 complete years covering 3 full sea-ice melting and freezing cycles. Is it possible to extend the dataset to June, 2019?

Specific comments:

2.1.1 SARAL/Altika

- L46-48 Nothing is said about the drifting phase of SARAL. Maybe you can add some more information.
- Does the drifting phase have an influence on the general processing, for example on the multi-mission cross calibration (e.g. number of cross-over points)?

2.1.2 Cryosat-2

- L56 Please change 0 padding to zero-padding.

2.1.3 Sentinel 3A

- L63 Please change 0 padding to zero-padding.
- Are there any processing schemes or any other information about the CNES S3PP processing prototype? What are the significant differences compared to the standard dataset? It would be nice to have some references here.

2.2.1 Waveform classification

- L69: "In Arctic SL studies, classification generally relies on the pulse peakiness..." This sentence is not entirely true. In your introduction you cite Quartly et al., 2019, who introduces several methods for open water detection. In addition to threshold-based methods that use only a simple pulse peakiness value, there are also methods that use more sophisticated thresholds or rely on unsupervised classification approaches (e.g. Müller et. al, 2017 or Lee, 2016). Please rephrase the sentence.
- It's not clear to me why Brownian echoes are discarded? Why it's not possible to use them?

2.2.4 Sea level anomaly estimation

- L112: Why do you decide for DTU15MSS? What does "(and, 2016)" mean? Why don't you use a newer DTUMSS?
- Can you be more specific which ECMWF model or mapping function you use? I'm missing a reference.

2.2.5 Data editing

- Can you specify how many outliers (percentage) are rejected or removed from the dataset?

2.2.6 Ocean/lead bias correction

- L 135. Why are only lead returns processed?

2.3.1 Cross-calibration

- I'm not sure, if I have fully understood the calibration process. But, how do you apply radial orbit errors in areas, where do you only have CS2 observations ($>81^{\circ}\text{N}$)? Could you please add some sentences for that case?

2.3.3. Optimal interpolation

- Since the OI is the core of the grid processing, it would be nice to have some more information about the major processing steps of the OI. I understand the interpolation is too complex to explain it in every detail in the scope of this article, but I would like to read some information about the basic processing in order to introduce the reader to the interpolation method (e.g. maybe adding a flow chart or some easy equations).

- Product description

- For a better overview, please add a table with the NetCDF attributes.
- Maybe I missed it, but what are the reasons or decision for the 10 days (time step) in case of the mono-mission in table 4? Where does this number come from?

4.1 Mono-mission products comparisons

- L215-218: Could you please rephrase the sentence "Differences...". I think something is wrong with that sentence (missing punctuation).

4.2 Regional statistics

- L222: "The variance distribution..." I can't find a plot or a deeper comparison between the existing datasets and the proposed one.
- L225: "This is expected..." How do you deal with this issue in your processing from section 2? What resolution do you reach here?
- L228 ", suggesting that our lead detection and retracking algorithms perform well". I can imagine that is more related to a good performing transition between ice and open ocean areas. I think this does not necessarily result from a well-working lead detection.

4.3. Comparisons to tide gauges

- How do you perform the tide gauge comparisons? Could you please give more information?
- Have you used tide gauge stations that are also used in other Arctic SL datasets (e.g. Rose et al., 2019) for validation? If yes, how does your dataset compare to the external datasets?

Conclusion:

- L274: "(from mid 2016 to mid 2019)" I guess the dataset ends in April 2019, which is not the middle of the year.
- Is there a plan to extend the combined dataset? If yes, maybe you can add some words to the outlook of the conclusion?

Additional comments:

- Table 3: References are missing in case of the tropospheric corrections and mean sea surface.
- Table 3: Is there an official name of the SRAL product?
- Geographical maps and plots: I recommend a stereographical projection for a better representation of all Arctic regions. At the moment areas like the Bering Strait are hard to identify.
- Figure 2: The caption is kept very short. What are the investigation periods of the three plots?
- Figure 3: What are the investigation periods of the overlap?
- There are some Latex errors (e.g. ??).
- Could you please add all used datasets as a reference (e.g. tide gauge data) to the data availability section?

Comments to the NetCDF dataset:

- Flagged values are indicated by a very high value. It would be good to tell the user

what is the flagging value or to include this information to the NetCDF file itself.

- Referring to the global attributes:
 - `time_coverage_duration = 'P1032.0D'`
 - `time_coverage_resolution = 'P3.0D'`

What does this mean? I think the information must be changed or additional explanations need to be made.

- In the case of the single satellite datasets, there is a problem in the maximum value for the coordinates.
 - `geospatial_lat_max` = 9.969209968386869e+36
 - `geospatial_lon_max` = 9.969209968386869e+36

As a final remark, I support the publication of the manuscript if the above comments and those of reviewer 1 are answered and text passages are modified.