Authors' Response to Reviews of

Arctic sea surface height maps from multi-altimeter combination

P. Prandi et al.
Earth System Science Data, 10.5194/essd-2021-123

RC: ***Reviewers' Comment***, AR: Authors' Response, □ Manuscript Text

We would like to thank both reviewers for the time they allocated to reviewing this paper. Responses to their comments are available below. We hope that the updated version of the paper will be considered suitable for publication in ESSD.

1. Reviewer #1

1.1. General comments

RC: ***The study describes the first multi-mission product from the Arctic Ocean with an impressive 3 days resolution. The whole concept of the paper is very interesting, but the manuscript is at this stage not mature enough for publication. I suggest it should go through a mayor revision. In most circumstances the manuscript can be improved by more information and better description about the procedures, but in some cases the writing is erroneous. The quality of the 3 days gridding is not validated. This is novel about this product, and should be validated. More specific comments: It would be nice with an image of the processing steps. This may be a personal opinion, but I do not like the chosen map projection. The map projection favors the interior of the map. It is difficult to see the data in the Bering Sea. The method for estimating the noise levels are not very strong. I do not follow the method. It may be my lack of knowledge, but the section should be described in more details, and some steps seams rather doubtful. There are no perspective to other studies regarding the comparison results. There are several spelling/grammatical errors though out the text which should be corrected.***

AR: We present a set of data quality assessment metrics to validate our 3-day gridded product. Including comparisons to state-of-the-art global sea level products and in-situ (tide gauges) measurements. I therefore disagree that the 3 day product is not validated. Regarding the projection used for maps, since both reviewers raised this comments, I've switched all plots back to a polar stereo projection. The revised version of the manuscript now features more rationale for noise levels. Although these noise levels are empirically set. Those noise priors shall be revisited with future versions of the product. Regarding the comparisons to other studies, I don't really want to engage into a competition. Currently available datasets serve different purposes: the DTU dataset is unique by its time span and is the only one available to investigate climate time scales. I've added some elements regarding this comparison in the response to reviews, since the open discussion will remain available.

1.2. Abstract

RC: ***- Missing the objectives of why this study is interesting - Include a sentence on how CS and S3A is processed***

AR: The abstract was rephrased to highlight why higher resolution datasets are needed for the Arctic Ocean, and to mention how CS2 and S3A are processed.

RC: ***L7: "A much better performance" is a very weak comment in an abstract. How is it better. Be more clear and precise.***

This was rephrased to mention quantitative comparisons with tide gauges and effective resolution.

1.3. Introduction

RC: ***Processing is a large part of this paper. I am missing a paragraph describing what others have done in this area***

AR: I added two sentences which provide a quick description of the products by Armitage et al., and Rose et al.

RC: ***L26: 1000km spatial resolution. This seams not correct! the whole Arctic is just about 4500 km wide***

AR: 1000km resolution is a bit exaggerated, and this was changed in the revised version to 'several hundred kilometer'. But while the grids the products are interpolated on might be 'high' resolution (down to about 25 km). The effective product resolution is certainly much lower than that. As an example, I've included below two snapshots of the Armitage et al. and Rose et al. datasets. From these snapshots, it is clear that the effective resolution of these datasets is lower than the grid resolution they are interpolated onto.

1.4. Data and methods

RC: ***L64: Please, explain what the CNES S3PP processing is, I am also missing a reference.***

AR: The CNES S3PP is an experimental processing chain. The acronym was explained in the revised text, and a reference to Boy et al. (2017), which introduces the concept of CNES Processing Prototypes, was added.

RC: ***L71: Can you clarify how the different classifiers differs? As I can see S3A is not included in the Longépé et al. (2019) paper.***

AR: I've removed this sentence from the revised version of the manuscript. Each mission has its dedicated classifier (to accommodate different modes and thus waveform shapes), but all classifiers are constructed and trained following the same methodology. The Longépé et al. paper evaluates both S3A dans SARAL classifiers, CS2 is not evaluated though.

RC: ***L81: The Adaptive algorithm...Please, summarize the importance of the algorithm***

AR: I don't understand this comment. For me the importance of using a physical retracker is underlined in the next sentence in the paper. Do you mean I should underline this even more ? or maybe less ?

RC: ***L84-85: "...no equivalent of the Adaptive retracking algorithm available..." This is not true! What about ALES+ and SAMOSA+. They are both designed to operate over both surfaces.***

AR: I doubt ALES+ qualifies as a physical retracker for SAR mode echoes, but you are right SAMOSA+ is an alternative. This sentence was removed from the revised version of the paper.

RC: ***L90: Your ocean/lead selection comes after retracking, but you write (L85) that you retrack differently over ice and open ocean?***

AR: I admit this is confusing. I've tried to rephrase the previous retracking paragraph. The retracker choice is

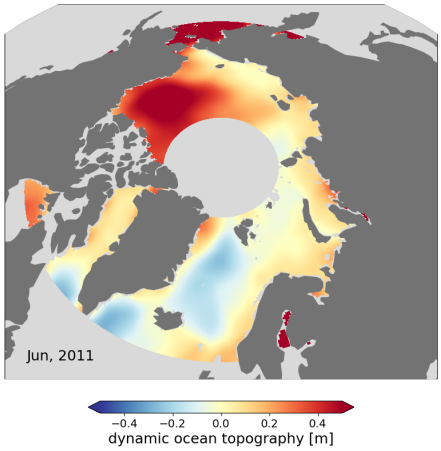

Figure 1: snapshot of the Armitage et al dataset

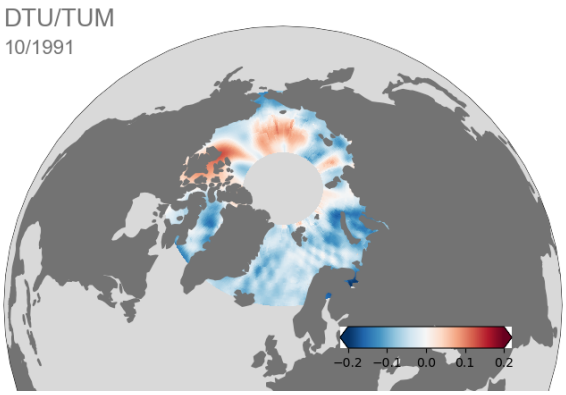

Figure 2: snapshot of the Rose et al dataset

based on waveform classification outcomes, but not all class 2 (peaky) waveforms are leads and additional selection criteria are applied.

RC: ***L95: How has this backscatter threshold been determined?***

AR: The paper was revised to briefly describe the process, and a reference to Longépé et al. (2019) added.

RC: ***L99: Shortly, explain the method used by Poisson et. al. (2018).***

AR: We added a sentence to describe the underlying principle for this hooking detection.

RC: ***L111: Please, explain "(and 2016)". Do you have a reference? I do not believe there exists a MSS DTU16***

AR: This results from a bibliography typo and was corrected in the revised version of the manuscript.

RC: ***L121-122: I do not understand what you do here. What is the iterative process? Could you please explain in more details.***

AR: The sentence was rephrased, the iterative editing simply removes measurements that are too far away from the sample mean value, given sample standard deviation.

RC: ***Fig1: How well are they correlating? It seams like C2 differs more than S3A og SRL, can you explain this?***

AR: I've rephrased this cross-calibration paragraph to better describe the observed signals, and explain what we see in these regional SL comparisons.

RC: ***Fig.2: There are a striping effect in the Norwegian Sea for SRL and S3A. Can you explain this? You have difference maps, but are not referring to them.***

AR: A reference to the difference maps was added, as part of the whole paragraph rephrasing. I don't see the striping pattern you are referring to.

RC: ***L152-153: This may be the striping you are referring to? What is it?***

AR: This cross-calibration paragraph was rephrased to (hopefully) increase clarity. I was referring to large scale patterns north of the Canadian Arctic Archipelago.

RC: ***Fig. 4: Please, comment on the latitude striping south of Greenland.***

AR: A new sentence was added to the text. The variance prior shall be updated in future product versions.

RC: ***L165-174: You are referring to equation terms in other papers. It is difficult to follow. Please, write the equation.***

AR: I feel that deriving the full OI process is out of the scope of the paper. Mathematical elements are well described in Bretherton (1976) while some implementation details are described in Le Traon and Dibarboure (1999) and Ducet et al. (2000).

RC: ***L178: Missing reference on high latitude errors in MSS***

AR: I don't know any published work that specifically addresses htis question, but I've added a reference to Pujol et al. (2018) that show (their figure 5) that large discrepancies between models remain at high latitudes.

RC: ***L188: Which SLR file?***

AR: This refered to the prior for SARAL used in the DUACS processing, but this paragraph is certainly confusing. The whole paragraph was rephrased and simplified.

RC: ***L189: Why*** $5cm^2$? ***Please quantify.***

AR: I've added a short comment about this choice, but this remains empirical and should be revisited in future product versions.

RC: ***Fig. 5: Why are there data everywhere, when they are separated in leads/ocean?***

AR: Priors are defined over the whole domain. During the optimal interpolation the right prior will be selected and used at each measurement depending on its type (leads or open ocean).

RC: ***L194: Why choosing*** $10cm^2$***?***

AR: This comes from the analysis of empirical orbit error variance levels in the region (which are available from DUACS global products) and provide a proxy for correlated error levels. This was clarified in the text.

RC: ***Fig. 6: Please explain the latitudinal striping.***

AR: I've added a sentence to comment on this latitudinal pattern. Again this shall be revisited in future product versions, but does not appear to have large impacts on the dataset quality.

RC: ***L198: It is very interesting with a 3 day product, but what is the data sampling in grid cells, ex. in the sea ice cover? I think this is the first time a 3 day product in the Arctic is available, but how is the 3 days quality?***

AR: For each grid point, the zone of influence (the time/space box in which observations will be considered) is pretty large, with individual measurements weighted according to their distance with respect to the interpolation position. We loose the measurement count information in the process, but assessing the quality of the 3-day product is the goal of the the data quality assessment section.

1.5. Product Description

RC: ***L200: Missing a reference, and again what does "(and, 2016 means?)"***

AR: The missing reference to Andersen et al. (2016) issue was corrected

1.6. Data quality assessment

RC: ***L235-247: You should use the same MSS in the comparison, or else it does not make sense. All the differences you see could be from the MSS.***

AR: The MSS is likely the main source of differences between the polar product and the global product, and this is acknowledged in the text. Using the same MSS in both products is not possible. It would require either to produce a DUACS-like global dataset with MSS DTU15 or an Arctic product with a CNES/CLS MSS, known for its limitations at high latitudes. Both solutions require more work than we can afford in this context. The goal here is simply to check that the Arctic dataset is not 'too far away' from a state-of-the-art open ocean dataset.

RC: ***L54-: There is no information about removing dac to altimetry or adding it to the tide gauges? This should be done before comparisons. Are you correcting for GIA? How are your results compared to other studies?***

AR: I agree that the comparisons to tide gauges paragraph doesn't provide enough details on the method. This has been corrected in the revised version and all relevant information is now provided.

RC: ***L254: What does co-located altimeter data means? How are you comparing altimetry to the tide gauges? Closest grid cell or...?***

AR: This information has been added in the revised manuscript version, we average altimeter grid points within 50 km (150 km) of the tide gauge position when comparing to the multi-altimeter (single-altimeter) dataset.

RC: ***L:257-261: How are your results compared to other studies? Both (Armitage et al., 2016; Rose et al., 2019) have better single mission correlation at Prudhoe Bay.***

AR: I don't really want to get into comparing correlations with tide gauges across papers. There are many details that can impact the correlations. Among them:

- both Armitage et al. and Rose et al. used monthly PSMSL data, while we use hourly GLOSS/CLIVAR data,
- Armitage et al. apply a 300km filter on altimetry before selecting the closest grid cell, 350 km for Rose et al., while we are at 50 and 150 km,
- temporal periods do not overlap,

We did however perform a quick comparison between our hourly tide gauges time series and CPOM and DTU datasets. The resulting time series at Prudhoe Bay are included in this response as figure 5

1.7. Conclusion

RC: ***L277: "...performs better than any single-altimeter analysis" I do not completely agree (See above), maybe it should be: "...performs better than our single-altimeter analysis"***

AR: This was corrected in the revised paper, and I've added a reference to Athanase et al. (2021) who used the dataset to characterize new water pathways north of Svalbard.

1.8. Minor corrections

RC: ***L56: PDGS: What does this stand for?***

AR: The acronym was expanded in the revised manuscript

RC: ***L86: TFMRA: Abbreviation - spell it out***

AR: The acronym was expanded in the revised manuscript

RC: ***L87: MLE3: what is this, missing reference***

AR: MLE3 is the 3 parameter maximum likelihood estimator retracker. This is the standard ocean retracker.

RC: ***L167: figure 4 comes before figure 3***

AR: Figure 3 is now reference before figure 4 in the cross-calibration section

RC: ***L209: Error in ref (??)***

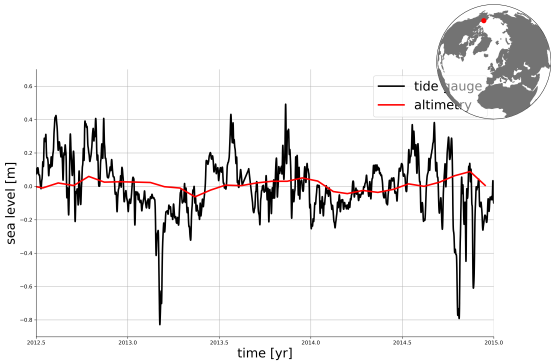

Figure 3: CPOM dataset

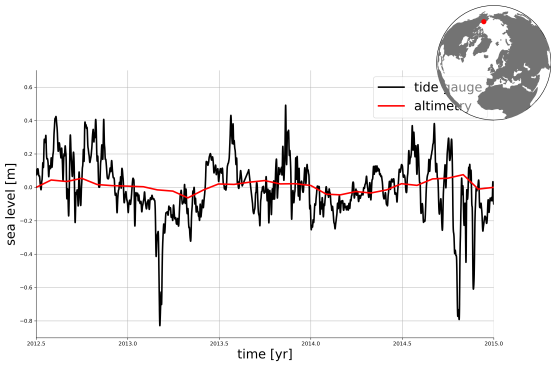

Figure 4: DTU dataset

Figure 5: comparisons to the Prudhoe Bay record used in this paper

AR: This reference error was corrected in the revised version of the manuscript

RC: ***L258: Fig ref error***

AR: This reference error was corrected in the revised version of the manuscript

2. Reviewer #2

RC: ***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.***

2.1. General

RC: ***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?***

AR: The starting point of this study is really to see if we are able to leverage the current altimeter constellation to improve our sea level products in high latitude areas. While existing datasets are enough to study the Beaufort Gyre extension, we know that beaufort gyre staibility mechanisms involve much smaller structures. I'm not claiming we are able to resolve such structures, but this is a step in the right direction. It's true that at very high latitudes this falls back to a mono-mission dataset here, but this could be used to include ICESat-2 data (if funding allows). Future versions of the product are planned and we hope to be able to inlcude new missions in the constellation. Following a similar comment by the other reviewer the abstract was rephrased, and some quantitative data quality assessment metrics were included. The time span of the dataset is limited by upstream data availability. Extensions are planned, when new upstream data are available.

2.2. Specific comments

2.2.1 SARAL/Altika

RC: ***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)?***

AR: We don't rely on cross-overs to perform the cross-calibration, as explained in the cross-calibration section. So the drifting phase of SARAL has little impact on the dataset.

2.2.2 Cryosat-2

RC: ***L56 Please change 0 padding to zero-padding.***

AR: this was corrected in the revised version of the manuscript.

2.2.3 Sentinel 3A

RC: ***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.***

AR: zero-padding was corrected. Following a similar comment by reviewer 1, we've added a reference to Boy et al. (2017) which introduces the processing prototype concept. The big advantage of the S3PP here is that it implements zero-padding while providing state-of-the art ocean data in one dataset.

2.2.4 Waveform classification

RC: ***L69: "In Arctic SL studies, classification generally relies on the pulse peakiness…" This sentence ie use of high resolution measurements. Open and ice-covered ocean noise levels are shown on figure 5s 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.***

AR: The sentence was rephrased to acknowledge these more sophisticated methods. A reference so the unsupervised clustering proposed by Muller et al. (2017) was added.

RC: ***It's not clear to me why Brownian echoes are discarded? Why it's not possible to use them?***

AR: Due to the C2 mode mask, not many open ocean SAR echoes are available. This is a simplification on our side, but we have plans to include C2 everywhere in future versions.

2.2.5 Sea level anomaly estimation

RC: ***L112: Why do you decide for DTU15MSS? What does "(and, 2016)" mean? Why don't you use a newer DTUMSS?***

AR: The reference to Andersen (2016) was corrected. We stick with DTU15MSS after investigating the benefits of switching to the DTU18MSS. It turns out (I'm not MSS specialist) that the DTU18MSS has large biases in the ocean, due to use of RADS data. DTU15MSS provides better consistency with our global ocean product.

RC: ***Can you be more specific which ECMWF model or mapping function you use? I'm missing a reference.***

AR: I don't know of a specific reference for the ECMWF model. This is the same model and methods than the ones reported in Taburet et al. (2019) https://os.copernicus.org/articles/15/1207/2019/#section2

2.2.6 Data editing

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

AR: This is not a quantity we've estimated with great accuracy. Typically between 40 and 90% of the data considered valid (with variations across missions and over time). This is considering a baseline of class 1 and 2 waveforms only. Including all measurements would lead to a much lower (unknown) number.

2.2.7 Ocean/lead bias correction

RC: ***L 135. Why are only lead returns processed?***

AR: This is explained in the paper under the waveform classification section and in a response to a previous comment. Due to the mode mask, there not many Ocean echoes, and they are discarded to simplify the processing.

2.2.8 Cross-calibration

RC: ***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°N)? Could you please add some sentences for that case?***

AR: We don't apply any empirical radial orbit error correction in this product, as this would require the estimation of a global cross-overs dataset. Instead we allow for more room during the optimal interpolation for geographically correlated errors. The first paragraph of the section was rephrased to make this appear more clearly (hopefully).

2.2.9 Optimal interpolation

RC: ***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).***

AR: Describing the OI is, in my opinion, out of the scope of this paper. The mathematical framework is well described in Bretherton (1976) and implementation details (regarding geographically correlated error mitigation for example) are available in other papers that are cited in the manuscript.

2.2.10 Product description

RC: ***For a better overview, please add a table with the NetCDF attributes.***

AR: Do you mean NetCDF variables ? I'm not sure adding a table with all attributes is going to provide clear overview of the dataset... From a quick look through ESSD papers, this does not appear to be a typical feature.

RC: ***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?***

AR: The rationale for using 10 days/75 km grid cells is a compromise between grid resolution and data coverage. Smaller grid size would lead to more empty grid cells. This is set empirically.

2.2.11 Mono-mission products comparisons

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

AR: The sentence was rephrased.

2.2.12 Regional statistics

RC: ***L222: "The variance distribution…" I can't find a plot or a deeper comparison between the existing datasets and the proposed one.***

AR: We didn't include plots from other datasets, and I'm not sure this necessary for the paper. Since the open discussion will remain visible, I'll include them here (figure 8). Please be aware that these are estimated on different time spans.

RC: ***L225: "This is expected…" How do you deal with this issue in your processing from section 2? What resolution do you reach here?***

AR: The effective spatial resolution is not estimated here. You are absolutely right though that we are in a mono-mission configuration in these areas and the benefit of the OI is low. Estimating the effective resolution would require to leave one mission out of the analysis, we will try to characterize the resolution when adequate Sentinel-3B data becomes available.

RC: ***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.***

AR: SARAL/AltiKa provides the baseline for the product, and no open ocean/lead bias is removed on this mission. Therefore I think this is a result of accurate classification and retracking.

2.2.13 Comparisons to tide gauges

RC: ***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 valida-***

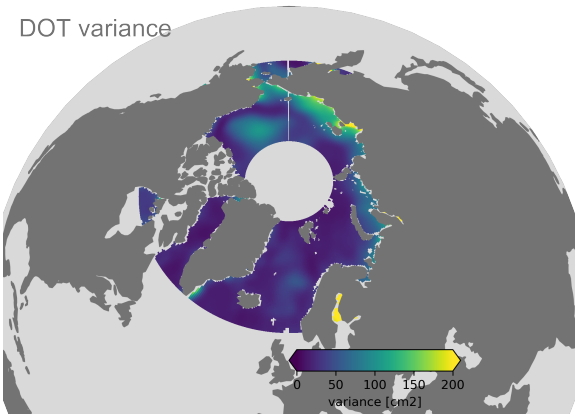

Figure 6: CPOM dataset

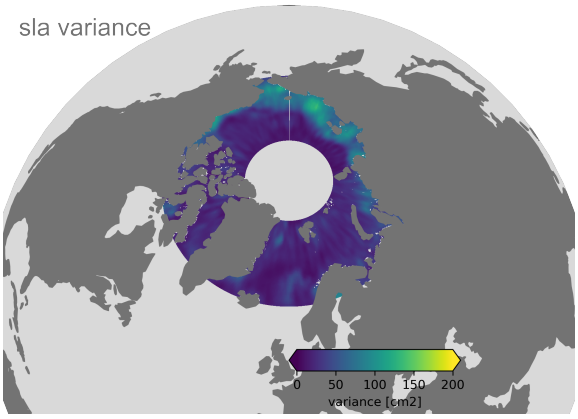

Figure 7: DTU dataset

Figure 8: SL variance levels

tion? If yes, how does your dataset compare to the external datasets?

AR: Following a similar comment by the other reviewer, this section was rephrased to provide more details.

2.2.14 Conclusion

RC: ***L274: "(from mid 2016 to mid 2019)" I guess the dataset ends in April 2019, which is not the middle of the year.***

AR: The sentence was changed to mention actual product dates.

RC: ***Is there a plan to extend the combined dataset? If yes, maybe you can add some words to the outlook of the conclusion?***

AR: Yes, extensions are planned and this was mentioned in the revised version of the manuscript.

2.2.15 Additional comments

RC: ***Table 3: References are missing in case of the tropospheric corrections and mean sea surface.***

AR: The MSS reference was updated, I don't know any reference of the ECMWF model, and no reference is given in Taburet et al. (2019).

RC: ***Table 3: Is there an official name of the SRAL product?***

AR: I don't know of an official name for the product, but the orbit is a CNES POE. This was updated in the table.

RC: ***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.***

AR: I really liked this orthographic projection... but since both reviewers raised this comments, I've switched all plots back to a polar stereo projection.

RC: ***Figure 2: The caption is kept very short. What are the investigation periods of the three plots?***

AR: When comparing missions together, the common period (Jun 2016-Apr 2019) is used. This was made clear in the caption in the revised version of the manuscript.

RC: ***Figure 3: What are the investigation periods of the overlap?***

AR: When comparing missions together, the common period (Jun 2016-Apr 2019) is used. This was made clear in the caption in the revised version of the manuscript.

RC: ***There are some Latex errors (e.g. ??).***

AR: missing LaTeX references were corrected in the revised version of the manuscript.

RC: ***Could you please add all used datasets as a reference (e.g. tide gauge data) to the data availability section?***

AR: I've added a reference to the UHSLC data repository in the data availability section.

2.2.16 Comments to the NetCDF dataset

RC: ***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.***

AR: Thank you for you're feedback on dataset format and content,

RC: ***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.***

AR: The format for these global attributes was enforced by the data repository, and I thought this was some kind of convention for the expression of time periods. I will check with them.

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

AR: I don't really get how this can be OK in the multi-satellite file and not in the single-satellite file, but I'll correct this and upload a new version of dataset as soon as possible.

RC: ***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.***

AR: Thank you. I hope that reviewers comments are correctly addressed and that the revised manuscript will be considered acceptable for publication.